# الإحصاء الوصفي وتطبيقاته

## باستخدام الحزمة الإحصائية

## SPSS


**دكتور: فايز عبد الهادي البتانوني**

**دكتور: عصام الدين محمد**

منشورات جامعة عمر المختار 2022

# الإحصــاء الوصفــي وتطبيقــاته

# **باستخدام الحزمة الإحصائية spss**

دكتور: فايز عبد الهادي البتانوني

دكتور: عصــام الــدين محمـــــد

منشورات جامعة عمر المختار

ليبيا – 2022

اسم الكتاب: الإحصاء الوصفي وتطبيقاته باستخدام الحزمة الإحصائية spss

اسم المؤلف: دكتور: فايز عبد الهادي البتانوني، دكتور: عصام الدين محمد

رقم الإيداع: 2022/106م.

**دار الكتب الوطنية بنغازي – ليبيا**



منشورات جامعة عمر المختار البيضاء

**الترقيم الدولي**

**ردمك ISBN 978-9959-79-118-4**

بِسْمِ اللّهِ الرَّحْمَنِ الرَّحِيمِ

(قُلْ إِنَّ صَلاَتِي وَنُسُكِي وَمَحْيَايَ وَمَمَاتِي لِلّهِ رَبِّ الْعَالَمِينَ) ﴿162﴾

من سورة الأنعام

## مقدمة

الحمد لله رب العالمين والصلاة والسلام على رسوله الكريم سيدنا محمد صلى الله عليه وسلم، وعلى آله وصحبه ومن اتبع هداه إلى يوم الدين وبعد...

فقد أصبح الإحصاء ضرورة لا غنى عنه في البحوث العلمية والتربوية والنفسية وغالباً لا تجد هذه البحوث سبيلها إلى الكمال إلا بالجانب الإحصائي.

ويهدف هذا الكتاب للمساهمة في التعرف على الإحصاء الوصفي لفهم بعض الأسس والمفاهيم التي يكثر ذكرها في البحوث والتي ستكون المدخل لدراسة الإحصاء الاستدلالي "التحليلي".

ويشتمل الكتاب على عشرة أبواب يناقش أولها المفاهيم الإحصائية، ويتناول الثاني مصادر البيانات، ويتناول الثالث الاحتمالات، ويوضح الرابع التوزيعات التكرارية، ويوضح الخامس كيفية العرض البياني للبيانات الإحصائية، ويشرح الباب السادس مقاييس النزعة المركزية، والباب السابع مقاييس التشتت، والباب الثامن مقاييس الارتباط الخطي بين متغيرين، والباب التاسع الانحدار الخطي البسيط، بينما الباب العاشر الإحصائيات الحيوية.

هذا ونسأل العلي القدير أن يكون هذا الجهد المتواضع معيناً للطلاب والباحثين.

**والله ولى التوفيق**

# الفهرس


| الصفحة | العنوان |
|---|---|
| | **الباب الأول: مقدمة للمفاهيم الإحصائية** |
| 3 | مقدمة........ |
| 3 | (1-1) أهداف علم الإحصاء........ |
| 4 | (1-2) علم الإحصاء........ |
| 7 | (1-3) المجتمع والعينة........ |
| 10 | (1– 4) بعض المفاهيم الإحصائية........ |
| 28 | (1-5) أنواع المتغيرات........ |
| 34 | (1- 6) شكل الانتشار........ |
| | **الباب الثاني: مصادر البيانات** |
| 41 | مقدمة........ |
| 41 | (2 – 1) جمع البيانات........ |
| 42 | (2– 2) مصادر البيانات........ |
| 45 | (2 – 3) طرق جمع البيانات ميدانياً........ |
| 58 | (2–4) الأخطاء الشائعة في جمع البيانات........ |
| 60 | (2 –5) أساليب جمع البيانات الإحصائية........ |
| 64 | (2 – 6) العينة الإحصائية........ |
| | **الباب الثالث: الاحتمالات** |
| 91 | مقدمة........ |
| 91 | (3–1) بعض التعريفات الهامة........ |
| 91 | (3–1.1) التجربة العشوائية........ |

| العنوان | الصفحة |
|---|---|
| (3-2.1) فضاء (فراغ) العينة........................................ | 92 |
| (3-3.1) الحدث........................................ | 93 |
| (3-4.1) تصنيف الأحداث........................................ | 97 |
| (3-1.4.1) الأحداث المستقلة........................................ | 97 |
| (3-2.4.1) الأحداث المتنافية (المانعة)........................................ | 98 |
| (3-3.4.1) الحدث المؤكد........................................ | 98 |
| (3-4.4.1) الحدث المستحيل........................................ | 99 |
| (3-2) العمليات على الأحداث........................................ | 99 |
| (3-1.2) تقاطع الأحداث (أ ∩ ب) ........................................ | 99 |
| (3-2.2) اتحاد الأحداث (أ ∪ ب) ........................................ | 100 |
| (3-3.2) الفرق بين حدثين........................................ | 101 |
| (3-4.2) الحدث المكمل........................................ | 102 |
| (3-3) الاحتمال والحدث........................................ | 103 |
| (3-2.3) الأحداث غير المنافية........................................ | 109 |
| (3-2.3) الحوادث المتنافية........................................ | 110 |
| (3-4) التعبير الرياضي عن قواعد حساب الاحتمالات........................................ | 114 |
| **الباب الرابع: التوزيعات التكرارية** | |
| مقدمة........................................ | 125 |
| (4-1) التصنيف والعرض للبيانات الإحصائية........................................ | 125 |
| (4– 2) طرق كتابة الفئات........................................ | 127 |
| (4-3) الجدول التكراري........................................ | 129 |
| (4-4) التوزيعات التكرارية المتجمعة........................................ | 131 |

| العنوان | الصفحة |
| --- | --- |
| (4-5) التوزيعات التكرارية المزدوجة........................................ | 136 |
| **الباب الخامس: العرض البياني للبيانات الإحصائية** | |
| مقدمة........................................................................ | 143 |
| (5-1) العرض البياني لبيانات غير مبوبة........................................ | 144 |
| (5– 2) العرض البياني للبيانات المبوبة........................................ | 169 |
| (5– 3) التوزيعات التكرارية........................................ | 172 |
| (5-4) التمثيل البياني للبيانات المتقطعة........................................ | 197 |
| **الباب السادس: مقاييس النزعة المركزية** | |
| مقدمة........................................................................ | 207 |
| (6 – 1) الوسط (المتوسط) الحسابي........................................ | 207 |
| (6-2) الوسيط........................................................ | 220 |
| (6-3) الرُبيعات........................................................ | 225 |
| (6-4) المنوال........................................................ | 231 |
| (6 – 5) العلاقة الاعتبارية بين الوسط والوسيط والمنوال.................... | 237 |
| (6 – 6) أنواع المنحنيات........................................ | 238 |
| **الباب السابع: مقاييس التشتت** | |
| مقدمة........................................................................ | 243 |
| (7-1) المدى المطلق........................................................ | 248 |
| (7 – 1.1) المدى للبيانات غير المبوبة........................................ | 248 |
| (7-2.1) المدى للبيانات المبوبة........................................ | 248 |
| (7-2) المدى الربيعي........................................................ | 250 |
| (7-3) نصف المدى الرُبيعي (الانحراف الربيعي) ........................ | 250 |

| العنوان | الصفحة |
|---|---|
| (7–4) التباين والانحراف المعياري | 258 |
| (7–5) معامل الاختلاف | 269 |
| (7–6) الدرجة المعيارية | 272 |
| **الباب الثامن: الارتباط الخطي بين متغيرين** | |
| مقدمة | 276 |
| (8–1) تحديد العلاقة بين المتغير X، Y | 279 |
| (8–1.1) العلاقة الخطية | 279 |
| (8–2.1) العلاقة الخطية (الموجبة) | 280 |
| (8–3.1) العلاقة الطردية غير التامة | 280 |
| (8–4.1) العلاقة الخطية الطردية التامة | 283 |
| (8–5.1) العلاقة الطردية (الموجبة) بين الظاهرتين ولكنها ليست قوية | 284 |
| (8–6.1) العلاقة الخطية العكسية التامة | 285 |
| (8–7.1) العلاقة العكسية غير التامة | 286 |
| (8–8.1) صور متعددة للانتشار الخطي | 287 |
| (8–2) معامل الارتباط الخطي البسيط | 288 |
| (8–1.2) معامل ارتباط بيرسون | 288 |
| (8–2.2) حساب معامل ارتباط بيرسون للبيانات غير المبوبة | 290 |
| (8–3) الارتباط المتعدد والجزئي | 305 |
| (8–1.3) معامل الارتباط المتعدد | 305 |
| (8–2.3) الارتباط الجزئي | 314 |
| (8–4) معامل ارتباط الرتب لسبيرمان | 318 |
| (8–5) ارتباط الصفات غير القابلة للترتيب | 322 |

| العنوان | الصفحة |
|---|---|


(8-1.5) مقياس معامل الاقتران.................................................... 322
(8-2.5) معامل التوافق.......................................................... 327

## الباب التاسع: الانحدار الخطي

مقدمة.................................................................... 335
(9-1) تقدير خط الانحدار للبيانات غير المبوبة.................................. 336
(9-1.1) خط انحدار y على x.................................................... 336

## الباب العاشر: الإحصائيات الحيوية

مقدمة.................................................................... 345
(10-1) أهمية تعداد السكان.................................................... 346
(10-2) أهمية دراسة السكان.................................................... 346
(10-3) مفاهيم ومصطلحات...................................................... 348
(10-4) الهرم السكاني.......................................................... 351
(10-1.4) خصائص الهرم السكاني.................................................. 351
(10-2.4) أهمية الهرم السكاني.................................................. 352
(10-3.4) العلاقة بين جغرافيا السكان والديموغرافيا.............................. 352
(10-5) العوامل المؤثرة في توزيع السكان........................................ 353
(10-6) البيانات السكانية...................................................... 353
(10-7) مصادر البيانات الديموغرافية............................................ 355
(10-1.7) التعداد السكاني...................................................... 355
(10-2.7) أساليب التعداد السكاني................................................ 356
(10-3.7) خصائص التعداد........................................................ 357
(10-4.7) أهمية التعداد السكاني................................................ 359

| العنوان | الصفحة |
| --- | --- |
| (10-5.7) التسجيل الحيوي | 361 |
| (10-6.7) المسح الميداني | 361 |
| (10-8) تقدير السكان | 362 |
| (10-1.8) التغير المطلق في عدد السكان بين تعدادين | 362 |
| (10-2.8) عدد السكان في نهاية فترة زمنية | 362 |
| (10-3.8) الطريقة الحسابية | 362 |
| (10-9) تقدير العوامل الديموغرافية | 367 |
| (10-1.9) التغير المئوي في عدد السكان بين تعدادين | 367 |
| (10-2.9) معدل الزيادة السنوية في عدد السكان | 368 |
| (10-3.9) الكثافة السكانية | 368 |
| (10-4.9) العوامل المؤثرة على كثافة السكان | 369 |
| (10-5.9) تقدير كثافة السكان وتوزيعهم | 370 |
| (10-6.9) كثافة السكن | 371 |
| (10-7.9) الخصوبة | 372 |
| (10-8.9) إحصائيات الخصوبة | 373 |
| (10-9.9) إحصائيات المواليد | 374 |
| (10-10.9) إحصائيات الوفاة | 377 |
| (9-11.9) إحصائيات الزواج | 381 |
| (10-12.9) إحصائيات الطلاق | 381 |
| (10-13.9) إحصائيات الهجرة | 381 |
| المراجع | 385 |

# الباب الأول

# مقدمة للمفاهيم الإحصائية

مقدمة

(1-1) أهداف علم الإحصاء

(1-2) علم الإحصاء

(1-3) المجتمع والعينة

(1- 4) بعض المفاهيم الإحصائية

(1-5) أنواع المتغيرات

(1- 6) شكل الانتشار

## مقدمة: Introduction

الإحصاء كغيره من العلوم المختلفة له مصطلحاته وتعاريفه الخاصة به، ولذلك كان لزاماً علينا أن نضع تعريفاً واضحاً لكلمات معينة سوف تستخدم في هذه الدراسة، ونتناول في هذا الفصل تعريف علم الإحصاء وبعض المصطلحات الإحصائية المهمة مثل المجتمع، العينة، ... الخ.

فالإحصاء Statistical يقصد بالإحصاء العد أو التعداد أو عدد الأشياء أو جمع بيانات عنها، وهو يشير إلى إحصاء السكان بمعنى عدد السكان في وقت معين، وكلمة أحصى تعني: عد وعلم عدد الأشياء وربما خصائصها، وبذلك تعني هذه الكلمة جمع البيانات بالإضافة إلى تلخيص وتنظيم وتحليل البيانات وعرضها في جداول والتوصل إلى استنتاجات عن معنى البيانات، وعادة ما تكون هذه الاستنتاجات في شكل تنبؤات. والإحصاء فرع من فروع العلوم التي تتعامل مع البيانات وتحليلها وتنظيمها للإجابة عن التساؤلات والاستدلال منها، وبذلك يستخدم الإحصاء في فهم الكثير من المشكلات، وأحياناً يساء استخدام الإحصاء في عرض البيانات بشكل خاطئ أو خادع للاستدلال.

فالإحصاء هو تطوير وتطبيق الأساليب لجمع وتحليل وتفسير المعلومات المرصودة (البيانات) من التحقيقات المخطط لها.

## (1-1) أهداف علم الإحصاء: The objectives (Aims) of the statistics

**1**. وصف البيانات Data Description بجمع البيانات لعرضها في جداول أو رسومات بيانية، وذلك لتسهيل فهمها وتحليلها.

**2.** التعبير عن الحقائق بصورة عددية واضحة ودقيقة بدلاً من عرضها، والتعبير عنها بطريقة إنشائية (أرقام بدل الجمل).

**3.** الاستدلال الإحصائي Statistical Inference بمقارنة المجموعات المختلفة وإيجاد العلاقة بينها، وتفسيرها Inferential واستخلاص النتائج واتخاذ القرارات المناسبة بعد قيام الباحث بتحليل البيانات المتوفرة لديه.

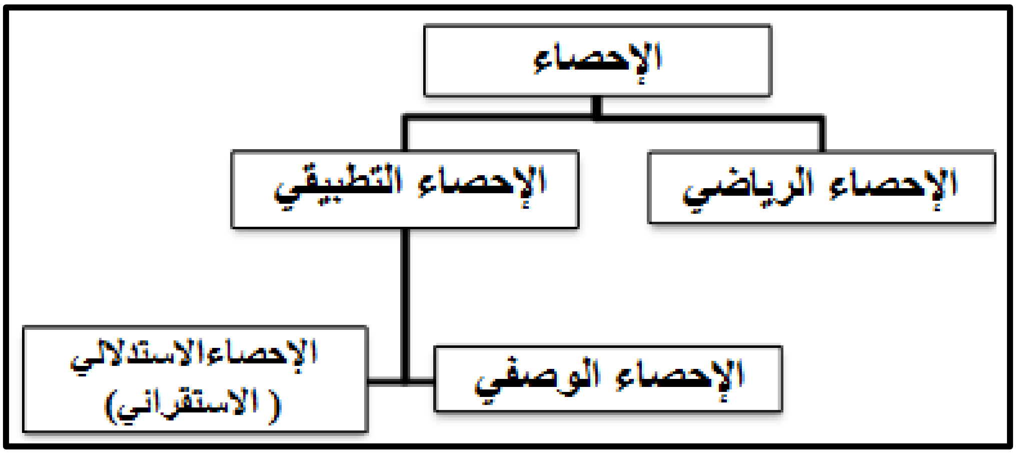
4. التنبؤ Forecasting

## (2-1) علم الإحصاء: Statistics

هو العلم الذي يختص بالطرق العلمية لجمع البيانات وتنظيمها وعرضها وتحليلها للظواهر العلمية المختلفة وتلخيصها بطريقة يسهل معرفة اتجاهات هذه الظواهر وعلاقة بعضها ببعض بهدف الوصول إلى نتائج وقوانين تحكمها واتخاذ القرارات المناسبة.

أي أنه علم استنباط الحقائق من الأرقام بأسلوب وطريقة علمية.

أي العلم الذي يبحث في جمع البيانات، وتنظيمها، وعرضها، وتحليلها، واستقراء النتائج، واتخاذ القرارات بناء عليها **ينقسم علم الإحصاء إلى**:




شكل (1) يوضح أقسام علم الإحصاء

- **الإحصاء الرياضي: Mathematical statistics**

هو دراسة الإحصاء من وجهة نظر رياضية باستخدام نظرية الاحتمالات وفروع رياضية أخرى مثل الجبر الخطي والتحليل.

- **والإحصاء التطبيقي ينقسم إلى قسمين:**

**أولا: الإحصاء الوصفي: Statistics Descriptive**

يهتم بجمع البيانات وتبويبها وعرضها باستخدام جداول تكرارية أو رسوم بيانية وتحليلها ببعض المقاييس العددية أو الوصفية التي تصف توزيع البيانات بغرض إظهار خصائصها المميزة.

ويحقق الإحصاء الوصفي أهدافه من خلال ما يلي:

**أ**. الطرق الإحصائية لجمع البيانات وتجهيزها وتبويبها.

**ب**. العرض البياني أو الهندسي وإمكانية عرضها في جداول ورسومات بيانية.

**ج**. الإحصاءات الوصفية تشير لخصائص التوزيعات مثل:

**1**. مقاييس النزعة المركزية وتتضمن: الوسط الحسابي، الوسيط، المنوال.

2. مقاييس الانتشار أو التشتت وتتضمن: المدى، التباين، المدى الربيعي، الانحراف المتوسط والانحراف المعياري.

3. مقاييس الوضع النسبي وتتضمن: الدرجة المعيارية، الربيعيات والمئينيات.

4. مقاييس الارتباط وتتضمن: ارتباط بيرسون، ارتباط كندال، ارتباط سبيرمان.

**ثانياً: الإحصاء التحليلي (الاستدلالي، الاستنتاجي): Analytical of inferential Statistics**

هو استخلاص النتائج العامة من النتائج الجزئية.

أو هو العلم الذي يختص في تحليل بيانات المجموعة والملخصة بهدف الوصول إلى نتائج تفيد في اتخاذ القرارات عند ظهور حالة عدم التأكد.

أي يهتم بطرق اتخاذ القرار التي تُكْتَشف، ويستدل منها على المجتمع اعتماداً على ما يتوافر من بيانات خاصة بالعينة المأخوذة منه، وتتناول ما يعرف بنظرية التقدير واختبارات الفروض ومستويات الدلالة بهدف إيجاد تقديرات لمعالم مجهولة أو الإجابة عن بعض الأسئلة البحثية أو التحقق من بعض الفروض المسبقة حول هذه المعالم المجهولة.

فالإحصاء الاستدلالي Inferential يعد مكملاً للإحصاء الوصفي إذ يقدم الإحصاء الوصفي خصائص الظاهرة ومواصفاتها، في حين يؤكد الإحصاء الاستدلالي على صحة حدوث الظاهرة.

ويعتمد على افتراضين أساسيين هما:

1. التوزيع الاعتدالي للمتوسطات.

2. العشوائية في اختيار العينة المستخدمة في الدراسة.

**ومن مقاييسه:** اختبار "ت" – تحليل التباين – اختبار مان ويتني – النسبة الحرجة – فريدمان – كروسكال واليز – ولكوكسون – كا2.

ويعتمد على فكرة اختيار جزء من المجتمع تسمى عينة بطريقة علمية مناسبة، بغرض استخدام بيانات هذه العينة في التوصل إلى نتائج يمكن تعميمها على مجتمع الدراسة، ومن ثم يهتم الاستدلال الإحصائي بموضوعين هما:

### 1. التقدير: Estimation

فيه يتم حساب مؤشرات من بيانات العينة تسمى إحصاء Statistics تستخدم كتقدير لمؤشرات المجتمع وتسمى معالم Parameters، ويطلق على المقاييس الإحصائية المحسوبة من بيانات العينة في هذه الحالة بالتقدير بنقطة Point Estimate، كما يمكن أيضا استخدام المقاييس الإحصائية المحسوبة من بيانات العينة في تقدير المدى الذي يمكن أن يقع داخله معلمة المجتمع باحتمال معين، ويسمى ذلك التقدير بفترة Interval Estimate.

### 2. اختبارات الفروض: Tests of Hypotheses

وفيه يتم استخدام بيانات العينة للوصول إلى قرار علمي سليم بخصوص الفروض المحددة حول معالم المجتمع.

**(1–3) المجتمع والعينة:**

● **المجتمع: Population**

هو مجموعة معرفة تعريفاً دقيقاً تحتوي على عدد محدود من العناصر تشترك في صفة أو صفات معينة تهم الباحث.

أو هو كل القياسات أو المشاهدات التي يقوم الباحث بجمعها لدراسة ظاهرة أو متغير يهتم بدراسته.

أو هو المجموعة التي تتكون من كل المفردات محل الدراسة.

● **المجتمع الإحصائي: Statistical Population**

هو الذي يتكون من تجمع عدد من المفردات تشترك في صفات وخصائص محددة يمكن ملاحظاتها وقياسها، كما أن هذه المفردات محددة ومعرفة بحدود زمنية ومكانية واضحة.

أو تلك المجموعة الأصلية التي تؤخذ منها العينة وقد تكون هذه المجموعة مدارس، كتب، سكان أو أية معدات أخرى.

● **المجتمع المحدد: specified community**

هو المجتمع المكون من نتائج محددة، مثل المجتمع المكون من إنتاج مصنع ما في يوم معين.

●**المجتمع غير المحدد: Non-specific community**

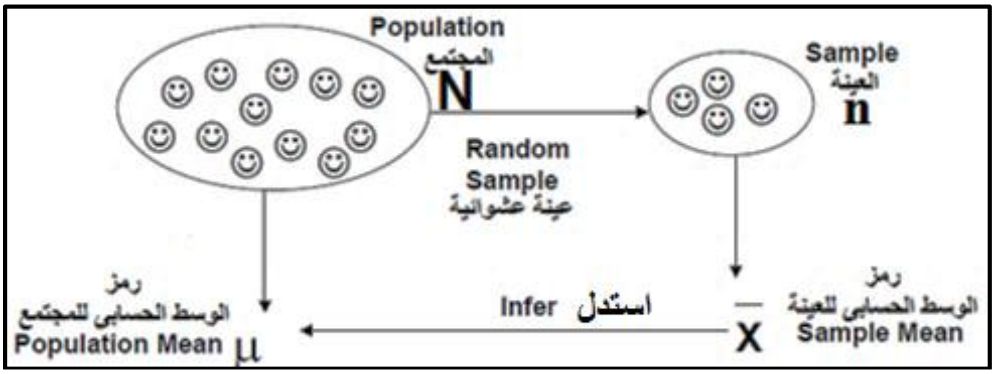
هو المجتمع المكون من جميع النتائج الممكنة، مثل نتائج رمية العملة (صورة، كتابة).

●**العينة: Sample**

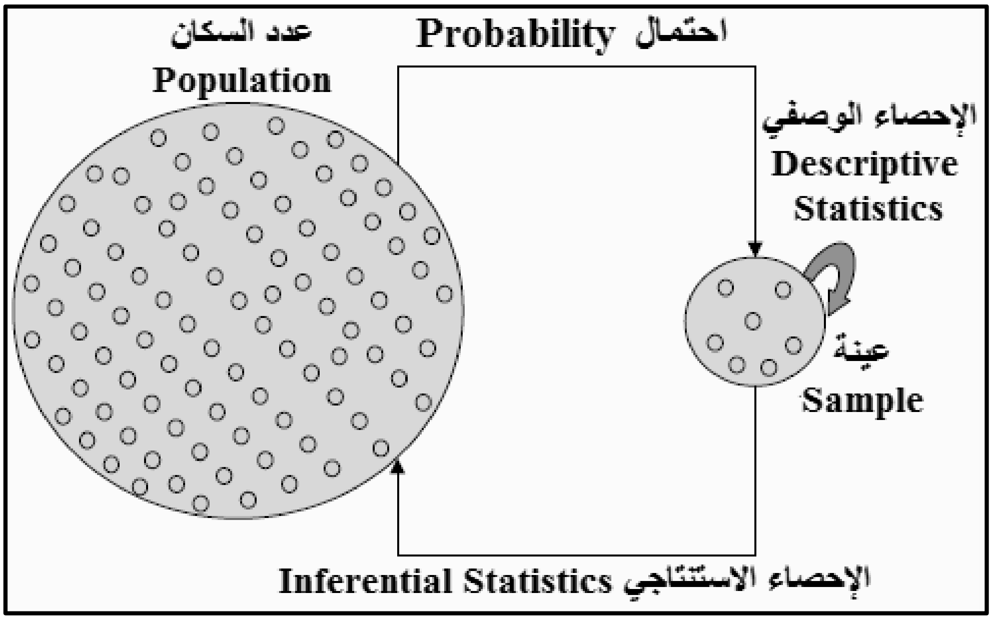
هي مجموعة جزيئية من مجتمع البحث.

أو ذلك الجزء الذي يتم اختياره من المجتمع بهدف تعميم نتائجه على المجتمع كله.

أو هي جزء من المجتمع تختار بطريقة أو بأخرى بغرض دراسة خصائص المجتمع الأصلي من خلال خصائص تلك العينة.




شكل (2) يوضح العلاقة بين المجتمع والعينة




شكل (3) يوضح العلاقة بين المجتمع والعينة والإحصاء الوصفي والاستنتاجي

- **حجم العينة: Sample size**

هي عدد الحالات الدراسية الإحصائية (المشاهدات) في العينة، وتعني عدد مفردات البحث المطلوب قياس متغيراتها، ونرمز لحجم المجتمع بالرمز N، ولحجم العينة بالرمز n.

**(1- 4) بعض المفاهيم الإحصائية:**

● **المفردة: Single**

تعني كائناً متحركاً (إنسان، حيوان) أو جامداً يمكن أن يقاس أو يعد عداً، والمفردة هي وحدة العينة Sample Unit.

● **المتغير: Variable**

مصطلح متغير يتضمن شيئاً يتغير، ويأخذ قيماً مختلفة أو صفات متعددة، فهو مفهوم يعبر عن الاختلافات بين عناصر فئة معينة مثل: النوع "الجنس"، والتحصيل، والدافعية، والانتباه، والمستوى الاقتصادي والاجتماعي، والجنسيات "مصري، سعودي، كويتي"، وطرق التدريس.

فالمتغير مصطلح يدل على صفة محددة، تأخذ عدداً من الحالات أو القيم أو الخصائص وتشير البيانات الإحصائية التي يقوم الباحث بجمعها إلى مقدار الشيء أو الصفة أو الخاصية في العنصر أو المفردة أو الفرد إلى متغيرات، وقد يشير المتغير إلى مفهوم معين يجري تعريفه إجرائياً في ضوء إجراءات البحث، ويتم قياسه كمياً أو وصفه كيفياً، فالذكاء مثلاً صفة عقلية لدى الأفراد بدرجات متفاوتة وهو لذلك متغير، لأنه ليس بنفس القيمة أو الدرجة أو المستوى عند جميع الأفراد.

ونلاحظ ضرورة اختلاف عناصر الفئة لكي نطلق عليها اسم متغير، أما إذا كانت العناصر من نفس النوع فإن هذه الخاصية تعد مقدارًا ثابتاً وليست متغيرًا، ومثال ذلك إجراء دراسة على الذكور فقط، ويعنى هذا أننا نثبت متغير الجنس (أي يصبح مقدارًا ثابتاً)، وبذلك يمكن تعريف المتغير بأنه: اختلاف الأفراد في قيم أو درجات خاصية معينة، ويهتم الباحثون بدراسة المتغيرات المختلفة وكذلك دراسة الثوابت.

فالمتغير هو نتائج تجربة لدراسة ظاهرة معينة يمكن قياسها وتأخذ قيماً مختلفة سواء رقمية أو وصفية مثل الدخل أو النوع.

- **المتغير الثابت: constant variable**

هو الذي يأخذ قيمة واحدة ثابتة لا تتغير طوال فترة الدراسة أو المتابعة.

- **المتغير المستقل: Independent Variable**

هو المتغير المعالج بغرض اختبار تأثيره على المتغير التابع وغالباً ما يسبق المتغير التابع ظهوره.

أي المتغير الذي يمكن للباحث التحكم فيه وهو صاحب التأثير الرئيسي والأساسي والمسبب للناتج.

- **المتغير التابع: Dependent Variable**

هو المتغير الذى يظهر أثر المتغير المستقل فيه، ومطلوب اختباره وقياسه ونحاول تفسيره ومعرفة أسباب حدوثه وتحديد مدى إمكان التنبؤ بها.

- **متغير دخيل: Variable Intruder**

هو متغير لا أثر له في التجربة أو البحث، وليس من ضمن المتغيرات المستقلة أو التابعة لأنه ليس محور الدراسة ولكنه قد يؤثر فيها لذا يجب تثبيته.

- **المعلمة: Parameter**

هي شيء يميز المجتمع ككل، أي مقدار أو خاصية (يتميز بها المجتمع) مثل متوسط أعمار طلاب كلية ما، أو نسبة غير المتعلمين في مجتمع معين، ويطلق عليه أيضاً بأنه قيمة أو خاصية للمجتمع، ويرمز له عادة برموز يونانية وهي قيم ثابتة للمجتمع مثل:

جدول (1) يوضح بعض الرموز الإحصائية

| الخاصية | Parameter المعلمة | Statistic الإحصاء |
|---|---|---|
| الوسط الحسابي Mean | للمجتمع (ميو) $\mu$ | للعينة (x-bar) $\bar{\chi}$ |
| التباين Variance | للمجتمع $\sigma^2$ | للعينة $s^2$ |
| الانحراف المعياري Standard Deviation | للمجتمع (سيجما) $\sigma$ | للعينة S |

**التناسب: Proportion**

كسر يُعبّر عنه عادة في صورة عشرية وفيه يتم المقارنة ما بين عدد صغير من الحالات البحثية والعدد الكلي للحالات = 30 ÷ 10 = 0.33.

- **النسبة المئوية: Percentage**

مقدار (قيمة) مضروباً في الرقم 100.

$$\text{Percentage} = \frac{X}{Y} \times 100$$

- **النسبة: Ratio**

نسبة X إلى Y تساوي عدد الحالات البحثية لـ X مقارنة مع عدد الحالات البحثية لـ Y. $\frac{X}{Y}$

- **الانحراف المطلق للمتوسط الحسابي: Mean Absolute Deviation**

يعني قياس التغاير للوسط الحسابي (أي انحرافه عن المعدل الطبيعي) باستخدام متغير واحد، بغرض بيان تجانس أو عدم تجانس مجموعة الحالات البحثية (العينة) الخاصة بالبحث، ولكنه ليس في مستوى فائدة الانحراف المعياري، كمقياس للتغاير.

● **الإحصاء البارامتري (المعلمي): parametric Statistical**

هو الذي يهتم بالبيانات الرقمية الحقيقية للمتغيرات لدى عينات كبيرة الحجم ممثلة للمجتمع الأصل الذي سحب منه.

أو أنه قيم ثابتة ممثلة للسمات المرغوبة والمهمة في المجتمع وهو في اللغة العربية معلمة، ويتطلب شروطا خاصة وهي: اعتدالية التوزيع، والتجانس، والعشوائية. ولهذا يفترض أن تكون عينة الدراسة مسحوبة طبقا للمنحنى الاعتدالي، ويتطلب افتراضات أو معلومات حول خصائص التوزيع الأساسي للمجتمع، وأن يكون حجم العينة كبير، وتم اختيارها عشوائيا. وهو أكثر ملاءمة لمعالجة البيانات من المستوى الفتري والنسبي (المتغيرات الكمية) لاعتماده الدرجات الخام، والتي يتم تحليلها كما هي، ويستخدم مع العينات كبيرة الحجم (اختبارت، الارتباط الخطي، تحليل التباين).

● **الإحصاء اللا بارامتري (الاختبار اللا معلمي): parametric Statistical Non**

هو أسلوب إحصائي يستخدم في التحقق من صحة الفرضيات المتعلقة بقيم مجتمعات بارامتراتها غير محددة، أي لا يعتمد على معالم المجتمع فهو لا يتطلب أية افتراضات، أو معلومات حول خصائص التوزيع الأساسي للمجتمع، أي لا يعتمد على توزيعات معينة، ويمكن استخدامه مع العينات الصغيرة الحجم (أقل من ثلاثين)، وهو أكثر ملاءمة لمعالجة وتحليل البيانات من المستوى الاسمي والرتبي (المتغيرات النوعية)، وتسمى الأساليب اللا بارامترية

أحيانا باختبارات الرتبة order test، raking test ؛ لأنها ترتكز على رتبة أو ترتيب الدرجات، وليس على القيم العددية ، ومثال ذلك اختبار كا$^2$($\chi^2$ )، و مان ويتني.

● **التقدير: Estimation**

هو إعطاء تقييم على درجة عالية من الكفاءة لقيمة الظاهرة أو متغير في المجتمع من خلال بيانات العينة التي تمت دراستها، وهذه التقديرات إما أن تأخذ شكل قيم محددة (الوسط الحسابي – الوسيط) أو شكل تقديرات داخل فترة محددة (بالحد الأعلى والحد الأدنى) وذلك بدرجة ثقة أو احتمال معين.

● **الفروض: Hypotheses**

هي علاقات متوقعة بين متغيرين أو أكثر، أو هي توقعات الباحث لنتائج دراسته، وتعد الفروض حلولاً محتملة للمشكلة موضع الدراسة.

● **الفرض الإحصائي: Statistical Hypothesis**

عندما نعبر عن الفروض البحثية والصفرية والبديلة بصيغة رمزية وعددية، فإنها تسمى عادة الفروض الإحصائية.

وقد يكون الفرض الإحصائي "فرض موجه Directed" وهو صياغة للفرض مع تحديد اتجاه العلاقة "موجبة أو سالبة"، أو تحديد اتجاه للفروق بين المجموعات في المتغير التابع، وتوجد ثلاثة أنواع من الفروض وهي:

● **الفرض البحثي: Research Hypothesis**

يشتق الفرض البحثي عادة اشتقاقاً مباشراً من إطار نظري معين، وهو يربط بين الظاهرة المراد تفسيرها وبين المتغير أو المتغيرات التي استخدمناها في هذا التفسير ومن

أمثلة الفروض البحثية (توجد علاقة بين الرضا عن العمل والإنتاجية لدى العاملين بالمؤسسات الصناعية).

- **الفرض الصفري (العدمي): Null Hypothesis**

يظن البعض أن الفرض الصفري عكس الفرض البحثي، لكن هذا غير صحيح فالفرض الصفري يعبر عن قضية إذا أمكن رفض صحتها فإن ذلك يؤدي إلى الإبقاء على فرض بحثي معين.

فهو الفرض حول معلمة المجتمع التي نجري اختبارًا عليها باستخدام بيانات من العينة، والذي يشير أن الفرق بين معلمة المجتمع والإحصائي من العينة ناتج عن الصدفة ولا فرق حقيقي بينهما، ونرفضه عندما تتوفر دلائل على عدم صحته، وخلاف ذلك نقبله.

**صياغة الفرض الصفري (العدمي):**

**الصياغة في حالة الاختبار من جانب واحد: One Tailed Test**

$$H_0: \mu > \mu_0 \quad \text{Or} \quad H_0: \mu < \mu_0$$

**الصياغة في حالة الاختبار من جانبين: Tow Tailed Test**

$$H_0: \mu - \mu_0 = 0 \quad \text{Or} \quad H_0: \mu = \mu_0$$

- **الفرض البديل: Alternative hypothesis**

يرمز له بالرمز ($H_1$) ويأخذ صيغة الإثبات، أي وجود فرق معنوي أو وجود علاقة أو وجود تغير أو وجود دلالة إحصائية.

وتكتب صيغة الفرضية البديلة عكس صيغة فرضية العدم وحسب نوع الاختبار كالاتي:

**صياغة الفرض البديل:**

**الصياغة في حالة الاختبار من جانب واحد: One Tailed Test**

هو الاختبار الذي تبين فيه الفروض البديلة أن المعلمة للمجتمع أكبر أو أصغر من إحصائية العينة، فهناك تحديد للاتجاه.

$$H_1: \mu < \mu_0 \quad \text{Or} \quad H_1: \mu > \mu_0$$

**الصياغة في حالة الاختبار من جانبين: Tow Tailed Test**

هو الاختبار الذي لا تبين فيه الفرضية البديلة أن معلمة المجتمع تساوي أو لا تساوي من إحصائية العينة.

$$H_1: \mu - \mu_0 \neq 0 \quad \text{Or} \quad H_1: \mu \neq \mu_0$$

- **اختبارات الفروض: Tests Of Hypotheses**

تهدف إلى اتخاذ قرار معين بشأن الفروض الأولية التي تمت صياغتها لتفسير ظاهرة معينة تجري دراستها، وذلك من خلال الاستعانة بالبيانات الإحصائية التي تم تجميعها؛ لتحديد إجابات واضحة بقبول أو رفض هذا الفرض.

- **الخطوات الأساسية في اختبارات الفروض:**

1. صياغة الفرضيات (الفرض الصفري $H_0$ Null hypothesis والفرض البديل $H_1$ Alternative hypothesis)، وتحديد الافتراضات Assumptions المتعلقة بتوزيع المجتمع (T Distribution، Z Distribution).
2. تحديد مستوى معنوية الاختبار $\alpha$ significance Level.

3. تحديد إحصاء الاختبار The test statistic (توزيع T، توزيع Z) المستخدم وتحديد توزيع المعاينة لهذا الإحصاء في ضوء الافتراضات، أي تحديد نوع توزيع المجتمع ما إذا كان المتغير العشوائي الذي يتم دراسته يتبع التوزيع الطبيعي، أم توزيع بواسون، أم توزيع ذو الحدين، أم غيره من التوزيعات الاحتمالية المتصلة أو المنفصلة، معظم التوزيعات الاحتمالية يكون توزيعها مشابهاً للتوزيع الطبيعي خاصة إذا كان حجم العينة كبيراً، وهناك نوعان من الطرق الإحصائية التي تستخدم في اختبار الفرضيات:

( أ ) **الاختبارات المعلمية:** وتستخدم في حالة البيانات الرقمية التي توزيعها يتبع التوزيع الطبيعي.

(ب) **الاختبارات غير المعلمية:** وتستخدم في حالة البيانات الرقمية التي توزيعها لا يتبع التوزيع الطبيعي طبيعي، وكذلك في حالتي البيانات الترتيبية والوصفية.

4. صياغة قاعدة القرار بتحديد منطقة الرفض بناءً على مستوى المعنوية واتجاه الفرض البديل.
5. جمع البيانات من العينة، وحساب قيمة دالة الاختبار الإحصائية.
6. إيجاد القيمة الجدولية لإحصاء الاختبار في ضوء الافتراضات المشار إليها في الخطوة (1).
7. اتخاذ قرار برفض أو قبول فرضية العدم $\mathbf{H_0}$.

- **الدلالة الإحصائية: Statistical Significant**

تعد العينة الإحصائية ذات دلالة إحصائية، وذلك عندما تقع داخل منطقة الرفض في حالة اختيار الفرض الصفري، ويرفض الفرض الصفري طالما أن العينة الإحصائية تقع بعيداً عن معلمتنا المفترضة.

- **المستوى الدلالي الإحصائي: Level of Significance**

هي وصف توصف بها نتيجةً لتجربة أجريت عندما تكون القيمة الاحتمالية (p-value) أقل من مستوى الدلالة.

وتمثل نسبة المساحة الواقعة تحت توزيع المعاينة الإحصائية، والذي يمثل ما يمكن اعتباره منطقة الرفض، أو القيم غير المحتمل ظهورها بناء على افتراضنا الصفري، ويعد المستوى الدلالي هو الاحتمال بأن نحصل على قيمة غير محتملة الظهور أو أن احتمال ظهورها أقل من احتمال ظهور القيم الواقعة في منطقة الرفض.

فهي درجة الاحتمال الذي نرفض به الفرض الصفري (العدم) $H_0$ عندما تكون صحيحة أو هو احتمال الوقوع في الخطأ من النوع الأول ويرمز له بالرمز $\alpha$، وهي يحددها الباحث لنفسه منذ البداية وفي معظم العلوم التطبيقية نختار $\alpha$ مساوية 1% أو 5 % على الأكثر.

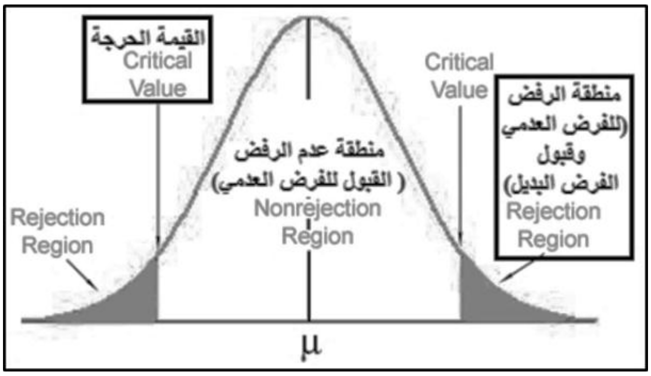
فمستوى الدلالة الإحصائية 0.05 يعني أننا لو أعدنا هذا البحث أو الاختبار 100 مرة فإننا سوف نحصل على نفس النتيجة 95 مرة، وسوف تكون نتيجة مختلفة هو 5 مرات من أصل المائة (0.05)، وكذلك بالنسبة لمستوى الدلالة 0.001 تعني أعدنا هذا البحث أو الاختبار 1000 مرة فإننا سوف نحصل على نفس النتيجة 999 مرة في مقابل مرة واحدة خطأ.

- **القيمة الاحتمالية: (Sig. or P-value)**

احتمال الحصول على قيمة أكبر من أو تساوي (أقل من أو تساوي) إحصائية الاختبار المحسوبة من بيانات العينة أخذاً في الاعتبار توزيع إحصائية الاختبار بافتراض صحة فرض العدم $H_0$ وطبيعة الفرض البديل $H_1$ ويتم استخدام القيمة الاحتمالية لاتخاذ قرار حيال فرض العدم.

● منطقة الرفض: **Rejection Region**

في توزيع المعاينة الإحصائية، تتضمن هذه المنطقة العينات الإحصائية غير المحتملة الظهور بناء على الفرض الصفري، وعموماً تعني المناطق الطرفية أو قيم العينة التي تظهر نادراً إذا كان الفرض الإحصائي (الصفري) صحيحاً.



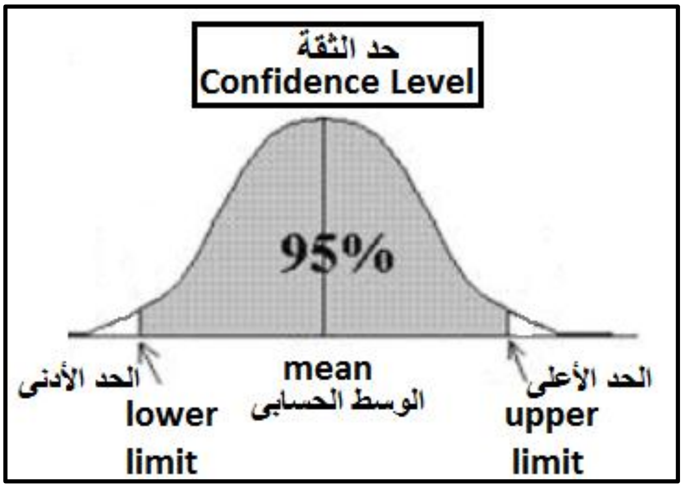

شكل(4) يوضح مناطق القبول والرفض للفروض الإحصائية

● معامل الثقة: **Confidence Coefficient**

يرتبط مستوى الاحتمال بوجه عام باحتمال ($1-\alpha$)، وتشير القيمة 0.99 (0.95 أو 0.9) إلى أننا لو كررنا إجراءات المعاينة الإحصائية، لعدد غير محدود من المرات، فإننا نتوقع أن قيمة المتوسط الحسابي لمجتمع الدراسة سوف يقع في كل مرة، داخل حد الـ 99 % (95% أو 90%) من المرات أن يقع المتوسط الحسابي للعينة داخل الحد (حد الثقة) الذي قمنا بتحديده.

● حد الثقة: **Confidence Interval**

هو الخط المستقيم، أو الحد الفاصل الذي تسجل عليه معلمات مجتمع الدراسة، والإحصاءات الأخرى للعينة.




شكل(5) يوضح حدود الثقة

- **حدود الثقة**: **Confidence Limits**

القيم الخارجية لحد الثقة لو كان حد الثقة يعني بالمتوسطات الحسابية، فالقيم الخارجية تساوي قيم $(\bar{X})$ المتوسط الحسابي.

- **التنبؤ**:

يتم استخدام نتائج الاستدلال الإحصائي، والتي تدلنا على سلوك الظاهرة في الماضي في معرفة ما يمكن أن يحدث لها في الحاضر والمستقبل وهناك العديد من الأساليب الإحصائية المعروفة التي تستخدم في التنبؤ، ومن أبسطها أسلوب الاتجاه العام، وهي معادلة رياضية يتم تقدير معاملات باستخدام بيانات العينة، ثم بعد ذلك استخدام المعادلة المقدرة في التنبؤ بما يمكن أن يحدث للظاهرة في المستقبل.

- **ملاحظة**:

مصطلح فئة وفترة الفئة يستخدمان في أغلب الأحيان للدلالة على نفس المعنى على الرغم من أن فترة الفئة هي في الحقيقة رمز للفئة.

- **تكرار الفئة: frequency class**

هو عدد المفردات أو القيم التي تقع في مدى تلك الفئة ويرمز لها بـ (ك)، ويجب أن يكون مجموع التكرارات دائما مساويا للعدد الكلي لقيم الظاهرة (ن)، وتوجد علاقة عكسية بين طول الفئة وعدد الفئات، أي كلما اتسعت الفئة قل عدد الفئات والعكس صحيح.

- **الفئات**: هي المجاميع التي قسمت إليها قيم المتغير، وتأخذ كل فئة مدى معيناً من قيم المتغير.

- **مركز الفئة (منتصف الفئة): (Class midpoint (mark**.

هي القيمة التي تقع في منتصف الفئة ونحصل عليها بجمع الحد الأدنى والحد الأعلى للفئة ونقسم المجموع على اثنين.

مركز الفئة (منتصف الفئة) = (الحد الأدنى للفئة + الحد الأعلى للفئة) ÷2

**Class midpoint = [upper + lower] ÷ 2**

أو تكتب على هذه الصورة

$$\text{مركز الفئة (منتصف الفئة)} = \frac{\text{الحد الأدنى للفئة + الحد الأعلى للفئة}}{2}$$

$$\textbf{Class midpoint} = \frac{\textbf{[upper + lower]}}{\textbf{2}}$$

- **الحدود الحقيقية للفئات: Exact Interval Limit**

**الحد الحقيقي (الفعلي) للحد الأدنى للفئة = الحد الأدنى للفئة – 0.5**

Lower class boundary:

Lower class boundary = True Lower limits - 1/2 unit of measurement.

Or

Lower class boundary = Real Lower limits – 1/2 unit of measurement.

**الحد الحقيقي (الفعلي) للحد الأعلى للفئة = الحد الأعلى للفئة + 0.5**

Upper class boundary:

Upper class boundary = True upper limits + 1/2 unit of measurement

Upper class boundary = Real upper limits + 1/2 unit of measurement

والجدول التالي يوضح الحدود الحقيقية للفئات:

جدول (2) يوضح الحدود الحقيقية للفئات

**Examples of true (real) limites**

| Numbers | True(real) Lower limits | True(real) upper limits |
|---|---|---|
| 0 | 0 – 0.5 =-0.5 | 0 + 0.5 = 0.5 |
| 5 | 5 - 0.5= 4.5 | 5 + 0.5= 5.5 |
| 30 | 30 – 0. 5= 29.5 | 30 + 0. 5= 30.5 |
| 2.3 | 2.3 – 2.0.05= 2.25 | 2.3 +0.05= 2.35 |
| 152.13 | 152.13 – 0.005= 152.125 | 152.13 + 0.005= 152.135 |

جدول (3) يوضح تكوين جدول تكراري و الحدود الحقيقية للفئات

| Class limits حدود الفئة | Exact Interval Limit الحدود الحقيقية للفئات Or Class boundary حدود الفئة | Tally عصا الحساب( علامة ) | Frequency التكرار |
|---|---|---|---|
| 318 - 335 | 317.5 – 335.5 | IIII | 4 |
| 336 - 353 | 335.5 – 353.5 | 卌 | 5 |
| 354 - 371 | 353.5 – 371.5 | II | 2 |
| 372 - 389 | 371.5 – 389.5 | III | 3 |
| 390 - 407 | 389.5 – 407.5 | III | 3 |
| 408 – 425 | 407.5 – 425.5 | II | 2 |
| 426 - 443 | 425.5 – 443.5 | I | 1 |

مثال: إذا كانت درجات الذكاء سجلت فإن الفئة (140–145) تتضمن من الناحية النظرية كل القياسات من (139.5–145.5).

فهذه الأرقام تسمى الحدود الحقيقية للفئة، حيث أن:

الحد الأدنى الحقيقي للفئة = 140 - 0.5 = 139.5

Class lower boundary= 140 – 0.5 = 139.5

الحد الأعلى الحقيقي للفئة = 145 + 0.5 = 145.5

Class upper boundary = 140 + 0.5 = 145.5

- **عرض الفترة (فترة الفئة): Interval Width(Class Interval)**

عرض الفترة = الحد الأعلى للفئة — الحد الأدنى للفئة.

Class Interval= Upper Class limit – Lower class limit

عرض الفئة = الحد الأعلى للفئة - الحد الأدنى للفئة

Class width = Upper class limit – Lower class limit

ملاحظة: تسمى طول الفئة، ولكن الأصح تسمى Class width (عرض الفئة)، أو فترة الفئة (Class Interval) حالياً يتم حسابها طبقاً للعلاقة التالية.

**طول الفئة (مدى الفئة) = الحد الأعلى الحقيقي للفئة – الحد الأدنى الحقيقي للفئة.**

**Class Interval= Upper Class boundary – Lower class boundary**

**واقترح (فايز البتانونى) تطبيق العلاقة التالية:**

**طول الفئة (مدى الفئة) الحقيقي = [الحد الأعلى للفئة – الحد الأدنى للفئة]+ 1**

Class Interval= (Upper Class limit – Lower class limit) +1

**كما يتضح من الجدول التالي:**

جدول (4) يوضح حساب طول الفئة بناء على الحدود الحقيقية للفئات

| Age (years) | Frequency | Class Interval | Class boundary |
|---|---|---|---|
| 0 - 9 | 15 | 10 | -0.5 – 9.5 |
| 10 - 19 | 18 | 10 | 9.5 – 19.5 |
| 20 - 29 | 17 | 10 | 19.5 – 29.5 |
| 30 - 49 | 35 | 20 | 29.5 – 49.5 |
| 50 - 79 | 20 | 30 | 49.5 – 79.5 |

**Class Interval= Upper Class boundary – Lower class boundary**
Class Interval= [9.5 – ( - 0.5)]= 10
Class Interval= [19.5 – 9.5] =10
Class Interval= [29.5 – 19.5] =10
Class Interval= [49.5 – 29.5] =20
Class Interval= [79.5 –49.5] =30

**Class Interval= [Upper Class limit – Lower class limit] +1**
Class Interval= [9 – 0]+1= 9+1=10
Class Interval=[ 19 – 10] +1= 9+1=10
Class Interval= [29 – 20] +1= 9+1=10
Class Interval= [49 – 30] +1= 19+1=20
Class Interval=[ 79 –50] +1= 29+1=30

● **التكرار النسبي: Relative Frequency**

التكرار النسبي = التكرار الأصلي للفئة ÷ مجموع التكرارات

**Relative Frequency** $= f \div n$

التكرار المئوي = التكرار النسبي× 100%

و الجدولين التاليين يوضحان مفهوم التكرار النسبي والمئوي.

جدول (5) يوضح مفهوم التكرار النسبي والمئوي

| مستوى الهيموجلوبين في الدم فترة الفئة | التكرار f | التكرار النسبي f/n | التكرار المئوي $(f/n)^{*}100\%$ |
|---|---|---|---|
| 12.95 – 13.95 | 3 | 3/50 = 0.06 | 100× 0.06=6% |
| 13.95 – 14.95 | 5 | 5/50 = 0.10 | 10% |
| 14.95 – 15.95 | 15 | 15/50= 0.30 | 30% |
| 15.95 – 16.95 | 16 | 16/50 = 0.32 | 32% |
| 16.95 – 17.95 | 10 | 10/50 = 0.20 | 20% |
| 17.95 – 18.95 | 1 | 1/50 = 0.02 | 2% |
| المجموع | n=50 | 1.0 | 100% |

جدول (6) يوضح مفهوم التكرار النسبي والمئوي لتكرار متجمع صاعد

| Class الفئة | Frequency ( f ) التكرار | Cumulative Frequency التكرار المتجمع الصاعد | Relative Frequency (f / n) التكرار النسبي = تكرار الفئة ÷ مجموع التكرارات |
|---|---|---|---|
| 40-49 | 1 | 1 | 0.05 |
| 50-59 | 2 | 3 | 0.10 |
| 60-69 | 3 | 6 | 0.15 |
| 70-79 | 4 | 10 | 0.20 |
| 80-89 | 6 | 16 | 0.30 |
| 90-99 | 4 | 20 | 0.20 |
| Σ | 20 | | 1 |

● **الجدول التكراري المعدل:**

هو الجدول الذي يتم تعديل تكراره حسب طول الفئة وذلك لعدم انتظام طول الفئة في الجدول، وفيها يتم إضافة عمود في الجدول باسم طول الفئة، ويتم حسابه عن طريق الفرق بين الحد الأعلى والأدنى للفئة، وعمود آخر باسم التكرار المعدل، ويتم حسابه عن طريق ناتج

قسمة التكرار الأصلي على طول الفئة الذي تم حسابه مؤخرا، ويكون الجدول المعدل عبارة عن الفئات الموجودة من الأصل والتكرار المعدل الذي تم حسابه.

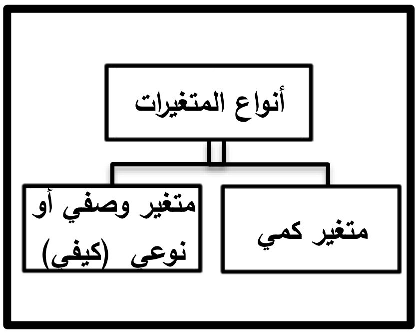
جدول (7) التكراري المعدل

| الفئات | التكرار | طول الفئة | التكرار المعدل = التكرار ÷ طول الفئة |
|---|---|---|---|
| 5-0 | 10 | 5 | 2=5÷10 |
| 8-5 | 20 | 3 | 6.67=3÷20 |
| 15-8 | 10 | 7 | 1.4 =7÷10 |
| 23-15 | 20 | 8 | 2.5 =8÷20 |
| 25-23 | 10 | 2 | 5 = 2÷10 |
| 30-25 | 20 | 5 | 4= 5÷20 |

## (5-1) أنواع المتغيرات: Types of Variables

المتغيرات تشير إلى الخصائص التي تشترك فيها أفراد المجتمع الإحصائي ولكنها تختلف من فرد إلى فرد آخر فالعمر، درجة الذكاء، وطول القامة، واللياقة البدنية والقدرة على القراءة، والدخول التي يحصل عليها الأفراد أمثلة للمتغيرات وتتميز هذه المتغيرات بأنها قابلة للقياس الكمي وبإمكانية تحديد قيمة معينة لها.

كما يمكن التعبير عن النوع، سنوات التعليم والعمر، والدخل السنوي بقيم مختلفة. فالمتغيرات عبارة عن ظواهر أو صفات تختلف قيمها باختلاف الحالات، والشكل التالي يوضح نوعي المتغيرات:




شكل (6) يوضح أنواع المتغيرات الإحصائية

**أولاً: متغير كمي: Quantitative Variable**

عبارة عن أرقام أو قيم عددية تدل على كميات، أو أعداد أشياء معينة لأفراد المجتمع.

ومن أمثلة المتغيرات الكمية (طول الشخص، عدد أولاد الشخص) وتنقسم المتغيرات الكمية إلي قسمين هما:

- **التوزيع المتصل: Continuous Distributions**

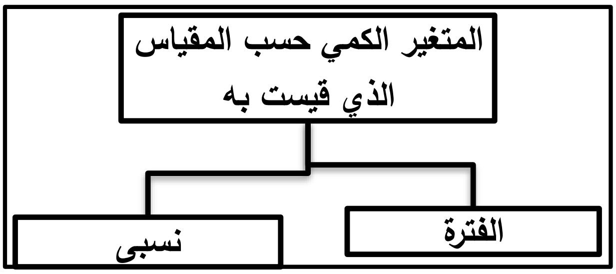
يتضمن مختلف القياسات التي بالإمكان إجراؤها على الأفراد أو الأشياء، ويمكن أن يأخذ قيمًا صغيرة وأرقاماً لا نهائية، أو ينحصر مداه داخل فترة مغلقة أو مفتوحة من مجموعة الأرقام الحقيقية، وفيه يتم التعبير عن المتغير بصورة عشرية أو كسرية متصلة مثل التعبير عن الطول ب 5. 10 م أو الوزن أو العمر ... وغيرها.

فهو القياس الذي يمكن أن يأخذ أي قيمة في فترة من الخط الحقيقي (مثل وزن).

- **التوزيع غير المتصل(المتقطعة): Discrete Distributions.**

وفيها يتم التعبير عن المتغير بأعداد صحيحة منفصلة لا يوجد بينها تداخل مثل فئة الأعداد 0،1،2،ن.

ويأخذ قيم ثابتة محددة أو أرقاماً منفصلة دون كسور ومعروفة، ويمكن تعيينها (وهي عادة قيم صحيحة فقط)، غير أنها تحمل قيماً بينها ترتيباً بين هذه القيم أو الفئات وتشبه المتغيرات المنفصلة المتغيرات النوعية والمتغيرات الرتبية من حيث أنها تسمح بتقسيم الأفراد في فئات منفصلة، ويكون المتغير منفصلا عندما يأخذ قيماً قابلة للعد، أي أنها تكون محدودة أو معدودة مثل: عدد أفراد الأسرة، وينقسم المتغير الكمي حسب المقياس الذي قيست به إلى قسمين:




شكل (7) يوضح تصنيف المتغير الكمي حسب المقياس الذي نستخدمه في القياس

**1. الفترة (الفاصلة، المسافة): Interval**

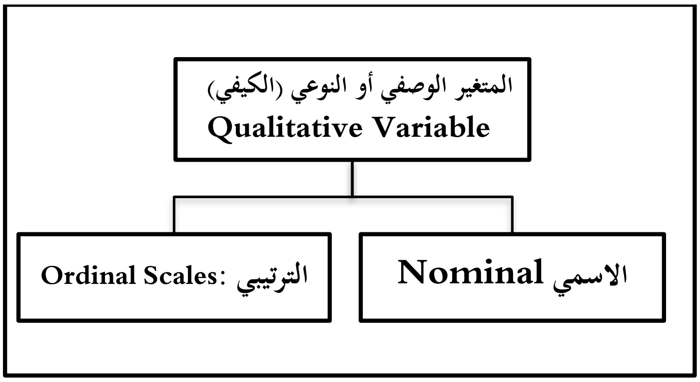
هي بيانات رقمية، تقاس بمقدار بعدها عن الصفر، أي أن للصفر دلالة على وجود الظاهرة، وتستخدم الأعداد في مقارنة قياس أو درجات الأفراد، وتسمح بالمقارنة بمدى الفروق بين قياسين درجة أ > ب درجة بمقدار 20 درجة في الاختبار س مع ملاحظة أن المقياس الفتري يؤخذ عليه عدم وجود نقطة الصفر المطلق بمعنى أن الصفر هنا لا يقيس حالة انعدام الخاصية. ومن أمثلة ذلك درجة الحرارة (0) ليس معناه انعدام الظاهرة، ولكنه يدل على وجود الظاهرة، ودرجة الطالب في الاختبار(0) لا يعني انعدم مستوى الطالب.

**2. نسبى (قياس نسبي): Ratio scale**

تستخدم الأعداد في تحديد العلاقات الوثيقة بين الأشياء أو الأحداث أو الأشخاص، أي يسمح بإجراء النسبة بين قيم المتغيرات، ومن الأمثلة على ذلك الأوزان والأطوال.

**ثانياً: متغير وصفي أو نوعي (كيفي): Qualitative Variable**

وتسمى أيضاً المتغيرات المصنفة أو النوعية Qualitative (non-numeric, categorical) data وهي تلك المتغيرات التي تصنف مفردات (أو عناصر) المجتمع الإحصائي أو العينة في عدة مجموعات، تشترك كل مجموعة في صفة معينة مثل لون العين، أو البشرة، وينقسم المتغير الوصفي أو النوعي (الكيفي) إلى قسمين:




شكل (8) يوضح تصنيف المتغير الوصفي

**1. الاسمي: Nominal**

هي بيانات وصفية مقاسة بمعيار اسمي Nominal Scale وغير رقمية تتكون من مجموعات متنافية، كل مجموعة لها خصائص تميزها عن المجموعة الأخرى، كما أن هذه المجموعات لا يمكن المفاضلة بينها أو ترتيبها، ويمكن وصفها بكلمة أو عبارة (مثل فصيلة الدم، ولون، ولون العين)، وتستخدم الأعداد في تقسيم أو تصنيف بالاسم فقط مثل الجنس (ذكر، أنثى) فيمكن إعطاء الذكر كود (1)، وإعطاء الأنثى كود (2)، والحالة الاجتماعية يمكن إعطاء المتزوج كود (1)، وإعطاء الأعزب كود (2)، والمطلق كود

.(3)

فالعدد لا يدل على كم أو مقدار ولكن يحل محل الاسم، ولا يمكن إجراء العمليات الحسابية على الأرقام.

## 2. الترتيبي: Ordinal Scales

هو استخدام الأعداد في ترتيب الأشياء أو الأشخاص ترتيباً تنازلياً أو تصاعدياً مثل: أ > ب، أي تتكون من مستويات، أو فئات يمكن ترتيبها تصاعديا أو تنازليا فهو كم لا يشار إليه بعدد ويهتم بترتيب الأفراد في الخاصية، ومن الأمثلة على ذلك تقديرا لطالب A, B, C,D.

ويمكن تحليل البيانات الترتيبية مثل البيانات النوعية، ولكن يتطلب حقا تقنيات خاصة تسمى أساليب غير بارامترية.

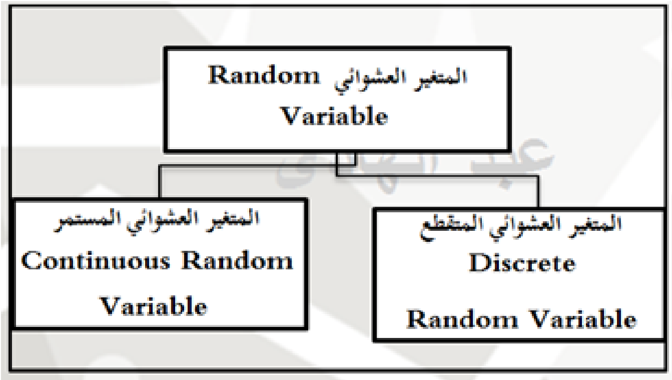
والجدول التالي يوضح بعض المعالجات الإحصائية للمتغيرات:

| مستوى القياس Level of Measurement | الوسط Mean | الوسيط Median | المنوال Mode | التباين/ الانحراف المعياري Variance / Standard Deviation |
|---|---|---|---|---|
| الاسمي: Nominal | x | x | √ | x |
| الترتيبي: Ordinal | x | √ | √ | x |
| الفترة: Interval أو نسبي Ratio | √ | √ | √ | √ |

## ●المتغير العشوائي: Random Variable

هو المتغير الذي يتم الحصول على قيمته نتيجة لتجربه عشوائية، فمثلاً إذا ألقينا عُملة معدنية ذات وجهين عدة مرات متتالية فإن نتيجة كل مرة سوف تكون متغيرة كما أنه لا يمكننا

قبل إجراء التجربة (المحاولة) معرفة ما سيظهر أولا، هل ظهر العملة أو وجهها، فإذا كانت س تمثل ما يظهر عند إلقاء عملة فإن س متغيراً عشوائياً، وينقسم المتغير العشوائي إلى قسمين:




شكل (9) يوضح تصنيف المتغير العشوائي

### 1. المتغير العشوائي المتقطع: Discrete Random Variable

إذا كانت كل القيم التي يمكن أن يأخذها المتغير يمكن عدها بغض النظر عن أنها محدودة أو غير محدودة فإن المتغير متقطع، أو الذي يأخذ قيماً صحيحة ومنفصلة فقط مثل عدد الغرف في المنزل.

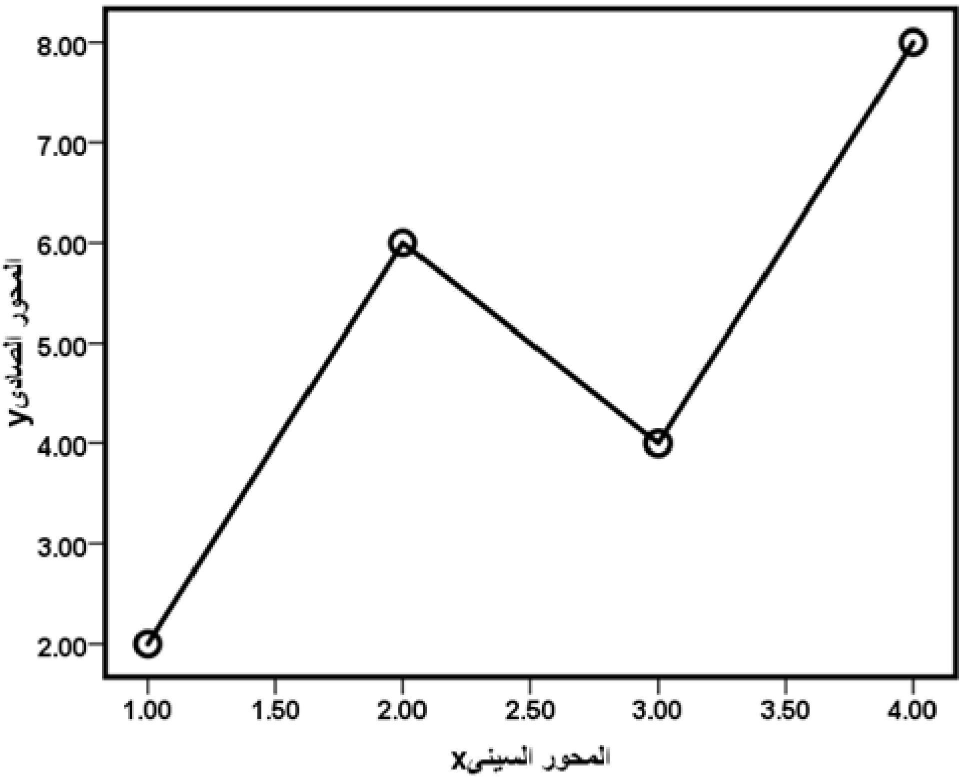
### 02. المتغير العشوائي المستمر: Continuous Random Variable

إذا كان المتغير العشوائي نحصل عليه بالقياس بمقياس مستمر، بمعنى أن جميع القيم التي يمكن أن يأخذها ذلك المتغير تكون رقمية أو قيماً متدرجة متصلة، أو الذي يأخذ قيماً لأي درجة من الدقة مثل أطوال مجموعة من الطلاب متغيراً عشوائياً مستمراً.

### (1- 6) شكل الانتشار: Scatter gram

رسم بياني يوضح مقدار (الارتباط) بين متغيرين، حيث يمثل الخط البياني الأفقي (المحور السيني) ورمزه (س "X") قيم متغير واحد (المتغير الأول)، والخط البياني الرأسي (المحور

الصادي) ورمزه (ص "Y") يمثل قيم (المتغير الثاني)، وتشير النقاط على الرسم البياني (الخط البياني) إلى قيمتي الحالة الواحدة مقاسه في ضوء المتغيرين كما يتضح من الشكل التالي:




شكل (10) يوضح شكل الانتشار

●**التوزيع التكراري:**

هو تبويب أو توزيع وحدات معينة في فئات القيم الخاصة بظاهرة معينة تكون موضوعا لدراسة التوزيع، وهو عبارة عن جداول مرتبة بشكل تصاعدي أو تنازلي تقسم إلى أصناف بحسب صفات مميزة ويسمى كل قسم أو صنف بالفئة، ويسمى هذا التوزيع بالتوزيع التكراري، والفئات إما متساوية وهو الشائع الاستخدام لكونها تسهل العمل الحسابي، أو فئات غير

متساوية. إن لكل فئة بداية تسمى بالحد الأدنى ونهاية تسمى بالحد الأعلى، وتسمى القيمة الواقعة عند منتصف الفئة مركز الفئة.

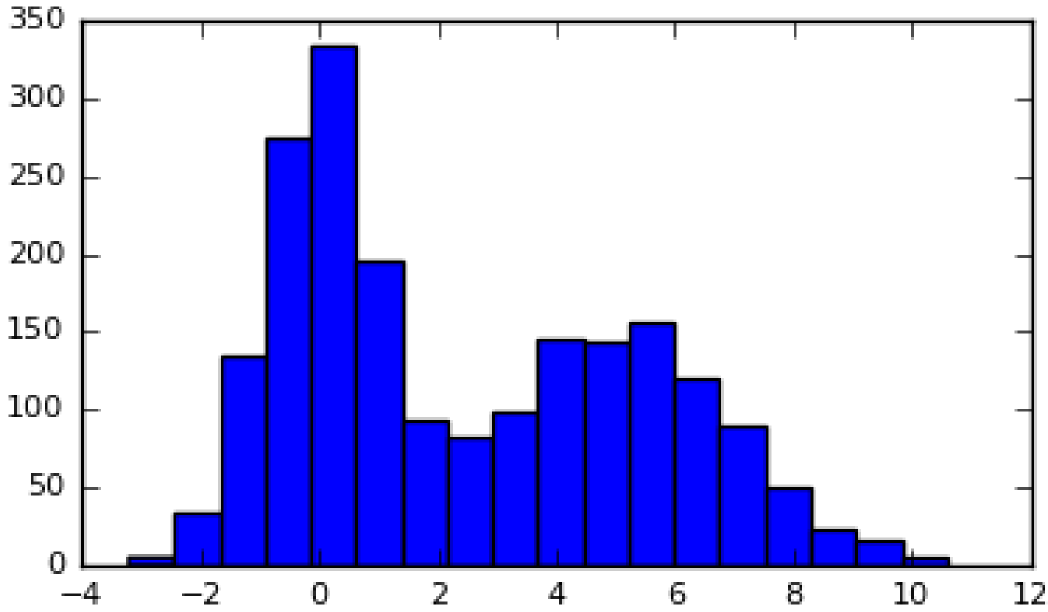
فهو عبارة عن جدول يحتوي على عمودين (المتغير ، والتكرار)، وقيم المتغير تعني القيمة المحتملة التي يأخذها المتغير في عينة البحث، ويشير التكرار إلى عدد الحالات البحثية التي ظهرت مقترنة بقيمة المتغير للمرة الأولى، زائداً عن الحالات التي تظهر بقيمة المتغير للمرة الثانية ... وهكذا.

- **المدرج التكراري: Histogram**

رسم بياني عمودي يستخدم لتوضيح التوزيع التكراري، حيث يشير فيه الخط البياني الأفقي إلى قيم المتغيرات، والخط البياني الرأسي إلى تكرار ظهور القيم في العينة.

وهو المحور الأفقي الذي يمثل العلاقة بين الفئات ويدرّج إلى أقسام متساوية بحيث يشمل جميع الحدود للفئات ويفضل ترك مسافة صغيرة بين نقطة الصفر والحد الأدنى للفئة الأولى (فيما إذا كانت بداية الفئة الأولى لا تساوي صفر)، والمحور العمودي يمثل التكرارات ويقسّم إلى أقسام متساوية بحيث تشمل على أكبر التكرارات.

ويرسم على كل فئة مستطيل رأسيًا تمثل قاعدته طول الفئة، وارتفاعه يمثل تكرار الفئة كما في الشكل التالي:

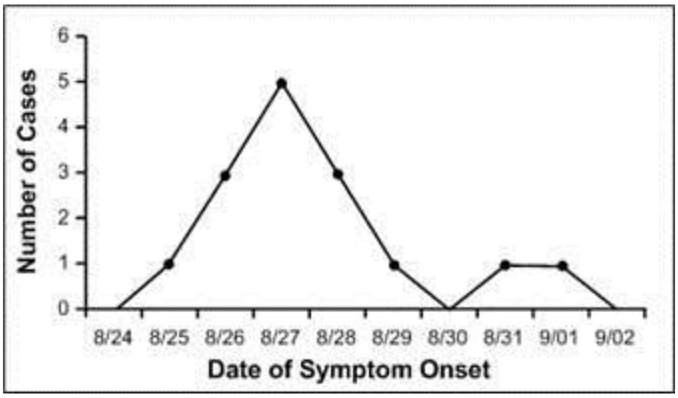



شكل (11) للمدرج التكراري

- **المضلع التكراري: Frequency Polygon**

هو خط بياني لتكرار الفئة المقابلة لمركز الفئة، ولرسم المضلع التكراري يجب استخراج مراكز الفئات كما يمكن رسمه من خلال توصيل القواعد العليا للمستطيلات المكونة للمدرج التكراري.

وهو يمثل العلاقة بين مراكز الفئات (المحور الأفقي)، والتكرارات (المحور العمودي) كما في الشكل التالي:



شكل (12) للمضلع التكراري

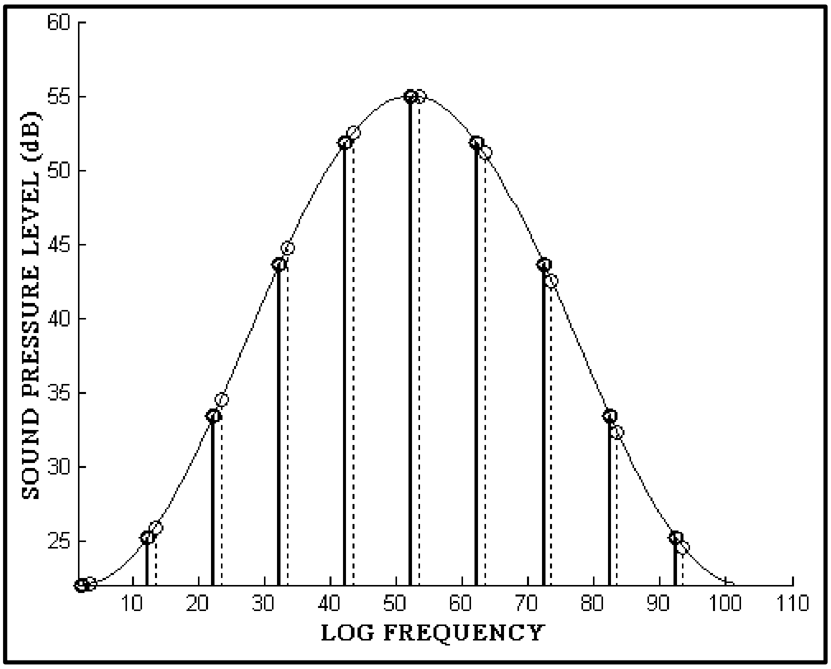
**ملاحظة:**

في كل من المضلع يتم إضافة فئة قبل الفئة الأولى، وأخرى بعد الفئة الأخيرة على الرسم فقط، تكرار كل منهما يساوي صفر وتمثل بخط متقطع.

● **المنحنى التكراري: Frequency Curve**

هو نفس المضلع التكراري بعد تهذيبه باليد بحيث يبدو على شكل منحنى ممهد Smooth Curve.

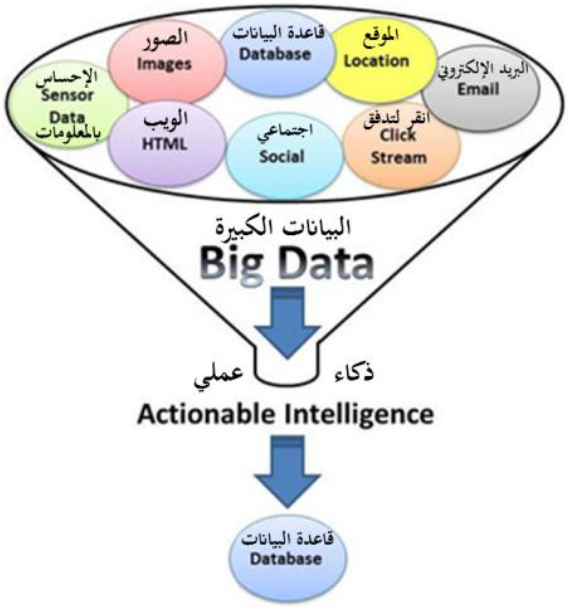
في المنحنى التكراري يتم إضافة فئة قبل الفئة الأولى وأخرى بعد الفئة الأخيرة على الرسم فقط تكرار كل منهما يساوي صفر وتمثل بخط متقطع كما في الشكل التالي:



شكل (13) للمنحنى التكراري

## الباب الثاني

## مصادر البيانات Sources of Data




- مقدمة.

(2 – 1) جمع البيانات.

(2– 2) مصادر البيانات.

(2 – 3) طرق جمع البيانات ميدانياً.

(2-4) الأخطاء الشائعة في جمع البيانات.

(2 -5) أساليب جمع البيانات الإحصائية.

(2 – 6) العينة الإحصائية.

**مقدمة:**

تتعدد طرق جمع البيانات بسبب تعدد طبيعة المجتمعات الإحصائية واختلاف البيانات التي نود جمعها، والإمكانيات المالية والبشرية والتنظيمية المتاحة للدراسة، وبصورة عامة يمكن تصنيف هذه الطرق إلى ما يلي:

1. الاستبانة.
2. طريقة الهاتف.
3. طريقة الإنترنت.
4. المقابلة الشخصية.
5. السجلات الرسمية.
6. النشرات والمجلات العلمية.
7. طريقة المشاهدة (التجارب).

**(2 – 1) جمع البيانات:**

المراحل الرئيسية للطريقة الإحصائية في البحوث العلمية والدراسات النفسية والاجتماعية لجمع البيانات هي:

1. تحديد المشكلة.
2. جمع البيانات عن الظاهرة.
3. تصنيف وتبويب البيانات وعرضها.
4. حساب المؤثرات الإحصائية.

5. تحليل النتائج.

6. تفسير النتائج.

**لذا يجب أن يراعي الباحث ما يلي عند تصميم بحثه:**

1. تحديد هدف البحث بشكل واضح ودقيق.
2. تحديد إمكانية التنفيذ الفعلي للبحث / المتطلبات المادية والبشرية.
3. تحديد إطار البحث (أي تحديد المجتمع الإحصائي).
4. تحديد أسلوب جمع البيانات والمعلومات.

**(2– 2) مصادر البيانات:**

1. **المصدر التاريخي: Historical Source (غير المباشر).**

هي البيانات المتوفرة فعلاً عن الظاهرة موضوع الدراسة والتي قد تتواجد في السجلات المحفوظة لدى أجهزة ومؤسسات ودوائر الدولة والمحفوظة لديها لسنوات سابقة وتنقسم إلى:

أ. **المصادر الأولية:**

وهي البيانات التي تفرغ وتبوب وتنشر من قبل الجهات التي قامت بجمعها مثل هيئة التخطيط أو الوزارات ومن أمثلتها:

التراجم والسير الشخصية التي تعطي فكرة عن كبار الشخصيات العلمية والسياسية والاجتماعية وإنجازاتها، وبراءات الاختراع التي تسجل اختراع شيء جديد لم يكن معروفا ولم

ينشر عنه شيء سابقا في الوثائق الرسمية الجارية التي تمثل مراسلات الدوائر والمؤسسات المعنية المختلفة.

**ب. المصادر الثانوية:**

وهي البيانات التي تقوم بتفريغها وتبويبها جهات أخرى غير الجهات التي قامت بجمعها وقد تكون منقولة أو مترجمة مثل: إحصاءات الأمم المتحدة، الجامعة العربية، منظمة التغذية والزراعة، ومن أهم المصادر الثانوية:

أ. **الكتب**: أكثر انتشارا وهي متخصصة في المعارف البشرية.

**ب. الأدلة**: تهتم بالمعلومات الخاصة بالمؤسسات العلمية.

**ج. الدوريات**: شكلها منتظم أو غير منتظم وتسمى مطبوعات مسلسلة.

**د. الموسوعات ودوائر المعارف**: (تجمع معلومات من مصادر أولية + ثانوية).

**هـ. الكتيبات والنشرات**: مطبوعات أصغر في حجمها من الكتاب الاعتيادي.

**و. المصادر السمعية والبصرية.**

**1.** سمعية مرئية: كالأفلام العلمية والوثائقية.

**2.** مرئية: كالصور والرسومات بأنواعها والخرائط الطبيعية.

**3.** سمعية: صوتية تعليمية وتسجيلات خاصة بالمقابلات ولقاءات صحفية وخطب لشخصيات مهمة.

**ز. المصادر الإلكترونية:**

هي المصادر التي أتاحتها تكنولوجيا المعلومات حيث أمكن تحويل المجموعات الورقية والمطبوعة إلى أشكال جديدة إلكترونية سهلة الاستخدام والتبادل مع المستفيدين في مواقع منتشرة جغرافيا على مستوى العالم مثل: الأقراص الصلبة وأقراص اقرأ ما في الذاكرة المكتنزة CD-ROOM والأقراص الليزرية المكتنزة DVD، والوسائط متعددة الأغراض.

**مزايا المصدر التاريخي:**

توفير الوقت والجهد والمال.

**عيوب المصدر التاريخي:**

**1.** قد تكون قديمة ( لم تحدث).

**2.** احتمالات الخطأ في نقل الأرقام.

**3.** قد لا تكون البيانات المنشورة تغطي جميع جوانب البحث.

**4.** احتمالات الخطأ في اختيار المفردات والمصطلحات المناسبة في حالة الترجمة.

**5.** احتمالات التحريف (التغيير المتعمد) في البيانات مما يؤدي إلى تشويه المعنى.

**6.** احتمالات الإضافة إلى البيانات الأصلية ومن ثم الوقوع في خطة تفسير البيانات.

**7.** قد تكون غير دقيقة، وقد تكون من جهة غير موثوق فيها (الغرض من البيانات الدعاية فقط).

**2. المصدر الميداني (المباشر)**

وفيه يتم جمع البيانات مباشرةً عن طريق اتصال الباحث بالوحدة محل الدراسة (شخص، هيئة، تجربة معملية، ...الخ) بأي وسيلة كانت سواء مراسلة أو مقابلة أو أخرى، ويتم ذلك بأسلوبين:

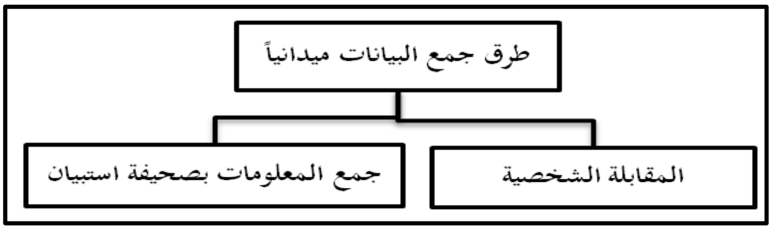
**أولهما:**

إذا كانت الظاهرة التي يريد الباحث الحصول على معلومات عنها تجربة معملية، مثل التجارب الزراعية أو الصيدلية ... الخ، فبعد وأثناء إجراء التجربة يستطيع الباحث بالملاحظة والمشاهدة أن يجمع البيانات المطلوبة.

**ثانيهما:**

هو اتصال الباحث بأشخاص أحياء لهم اتصال بالظاهرة محل الدراسة ليجيبوا عن تساؤلاته، ويتم خلالها توفير البيانات المطلوبة عن مفردات مجتمع الدراسة من خلال إجراء الاتصالات معها سواء مباشرة (المقابلة) أو بطريقة غير مباشرة (البريد – التليفون).

**(2 – 3) طرق جمع البيانات ميدانياً:**




**شكل (1.2)**: يوضح طرق جمع البيانات ميدانياً

**1. المقابلة الشخصية: Interview**

تحتم هذه الطريقة على الباحث أو مندوبة الاتصال المباشر بالوحدات البحثية أو الأشخاص محل الدراسة لتجميع البيانات المطلوبة منها، ويكون لدى الباحث كشف بحث Schedule يتم إعداده بعناية ويملؤه بنفسه (أو من ينوب عنه) من خلال اتصاله بالوحدة المبحوثة كما يحدث في التعداد السكاني.

فعن طريقة المقابلة يستطيع الباحث أن يتعرف على أفكار ومشاعر ووجهات نظر الآخرين. كما تمكن هذه الطريقة الباحث من إعادة بناء الأحداث الاجتماعية التي لم تلاحظ مباشرة.

- **أنواع المقابلة:**

**1. المقابلة المنظمة:**

وفيها يتم سؤال المشارك سلسلة من الأسئلة المعدة سلفًا، والتي سبق وحددت أنماط إجابتها، فهناك قدر ضئيل من التنوع في الأجوبة. وقد تستخدم هنا الأسئلة المفتوحة، وفي المقابلات المنظمة يتلقى جميع المشاركين الأسئلة نفسها وبنفس الترتيب والطريقة، ويكون دور الباحث محايدًا، وطبيعة هذا النوع من المقابلات يركز على الأجوبة العقلانية وليس على الأجوبة العاطفية.

**2. المقابلة غير المنظمة:**

وهي مقابلة غير مقننة، ذات أسئلة مفتوحة وعميقة ويكون دور الباحث أقرب لمدير الحوار أكثر منه مقابلاً، ويمكّن هذا النوع الباحث من فهم تفكير المشارك وسلوكه دون إسقاط فرضيات الباحث السابقة أو تصنيفاته عليه، والتي قد تحد من أقوال المشاركين.

**3. المقابلة الجماعية:**

هي المقابلة التي يعمل فيها الباحث مع مجموعة من الناس في وقت واحد، ويكون دور الباحث إدارة الحوار وتسهيل جريانه وانسيابيته، ومهمته تسجيل التفاعل الذي يدور بين المشاركين، وهذا يتطلب مهارات في إدارة الحوار وتوجيهه، وقد تكون المقابلة الجماعية منظمة،

أو غير منظمة. وقد تظهر جوانب من الحالة المدروسة ربما لا تظهر في أنواع المقابلات الأخرى، وذلك نتيجة لما يعطيه التفاعل بين آراء المشاركين ومشاعرهم وخبراتهم من إثراء للمقابلة وقدح لأفكار الآخرين من المشاركين.

- **تسجيل المقابلة:**

يجب القيام بالتسجيل الصوتي أو الفيديو صوتًا وصورة الأشياء للمقابلة، فلا يكفي أن يسجل الباحث ملاحظاته في أثناء المقابلة. فالتسجيل يساعد الباحث على إعادة النظر في المعلومات التي قيلت وتأملها مرة أخرى، وقد يكون من المفيد كتابة الملاحظات مع التسجيل لتنفيذ ما قد يلفت انتباه الباحث في أثناء المقابلة، ومن الضروري أن يفرغ نصها كتابة، ليسهل تحليله والتأمل فيه، وفي كثير من الأحيان يكون مفيدًا أن يعاد نص المقابلة لمن أجريت معه المقابلة ليعيد قراءة النص ويضيف ما يراه مناسبًا أو يوضح ما يحتاج إلى توضيح.

- **مميزات المقابلة:**

**1.** دقة المعلومات التي تجمع.

**2.** إمكانية معرفة طباع وردود أفعال الآخرين تفيد للأفراد الذين لا يعرفون القراءة والكتابة.

- **عيوب المقابلة:**

**1.** تحتاج إلى جهد ومال ووقت كثير.

**2.** قد يتحيز الباحث إلى نوع من الإجابات فيتسبب في خطأ التحيز.

## 2. جمع المعلومات بصحيفة استبيان: Questionnaire

- **الاستبيان:**

أداة لجمع البيانات المتعلِّقة بموضوع بحث محدَّد عن طريق استمارة يجري تعبئتها من قبل المستجيب، ويستخدم لجمع المعلومات بشأن معتقدات ورغبات المستجيبين، ولجمع حقائق هم على علمٍ بها؛ ولهذا يستخدم بشكلٍ رئيس في مجال الدراسات التي تهدف إلى استكشاف حقائق عن الممارسات الحاليَّة واستطلاعات الرأي العام وميول الأفراد، وإذا كان الأفرادُ الذين يرغب الباحث في الحصول على بيانات بشأنهم في أماكن متباعدة فإنَّ أداة الاستبيان تمكِّنه من الوصول إليهم جميعاً بوقت محدود وبتكاليف معقولة.

- **تعريف الاستبيان:**

أداة لجمع البيانات الأولية والميدانية المتعلِّقة بمشكلة أو بموضوع بحث محدَّد.

أو أحد وسائل البحث العلمي من أجل الحصول على بيانات أو معلومات تتعلق بأحوال الناس أو ميولهم أو اتجاهاتهم.

أو هي أداة أكثر استخداماً في الحصول على البيانات من المبحوثين مباشرة ومعرفة آرائهم واتجاهاتهم.

أو هي مجموعة من الأسئلة المكتوبة يقوم المجيب بالإجابة عنها.

- **تصميم الاستبيان:**

إعداد الشكل الأولي للاستبيان، ويتألف الاستبيان في صورته الأولية من صفحات:

1. غلاف الاستبيان.

**2.** الخطاب الذي يوجه للمبحوث.

**3.** البيانات الأولية للمبحوث.

**4.** فقرات وأسئلة الاستبيان.

وهي أن يقوم الشخص المستجيب بملء استبيان معد لجمع البيانات حول موضوع ما، ويصل إليه الاستبيان إما عن طريق البريد أو يسلم لهم شخصياً.

**ويجب عند تصميم الاستبيان مراعاة ما يلي:**

**1.** تحديد الهدف من الاستبيان.

**2.** مراجعة الدراسات والبحوث والمقالات العلمية ذات العلاقة بمشكلة البحث لصياغة الفقرات والأسئلة الفرعية.

**3.** الترتيب المنطقي للأسئلة.

**4.** البدء بالأسئلة البسيطة وتندرج إلى الأسئلة الأصعب.

**5.** تجنب الأسئلة الخاصة (المحرجة) والإيحائية التي توحي بإجابة محددة.

**6.** من الأفضل تكرار بعض الأسئلة المهمة في البحث بأكثر من صياغة.

**7.** تجنب الأسئلة التي تعتمد على ذاكرة الفرد لفترة زمنية سابقة بعيدة.

**8.** تجنب الأسئلة بدون داعي حتى لا يصاب الشخص محل الدراسة بالملل.

**9.** أن تكون الأسئلة واضحة ومحددة بحيث لا يختلف مفهوم السؤال من فرد إلى آخر (لا تحتمل أكثر من معنى).

**10.** التأكيد على سرية البيانات للشخص محل الدراسة حتى لا تكون إجاباته بعيدة عن الواقع.

- **أنواع الاستبيان:**

**1. الاستبيان المغلق أو (المقيدة):**

فيه يطلب من المبحوث اختيار الإجابة المناسبة من بين الإجابات المعطاة بوضع إشارة عليها كما هو الحال في الأسئلة الموضوعية ويتسم الاستبيان المغلق بسهولة الإجابة عن فقراته، ويساعد على الاحتفاظ بذهن المبحوث مرتبطاً بالموضوع وسهولة تبويب الإجابات وتحليلها. ويعاب عليه أنه لا يعطى معلومات كافية، وغموض موقف المبحوث، إذ لا يجد الباحث من بين الإجابات ما يعبر عن تردد المبحوث أو وضوح اتجاهاته.

وفي مثل هذا النوع ينصح الباحثون أن تكونَ هناك إجابةٌ أخرى مثل: غير ذلك، أو لا أعرف، ليحافظ الباحثُ على الموضوعيَّة.

- **مميزاته:**

**1.** سهولة الإجابة عن فقراته.

**2.** سهولة تبويب الإجابات وتحليلها.

**3.** يساعد على الاحتفاظ بذهن المبحوث مرتبطاً بالموضوع.

**4.** يشجع المشاركين على الإجابة عليه لأنه لا يطلب وقتا وجهدا كبيرين.

- **عيوب الاستبيان:**

**1.** لا يعطى معلومات كافية.

**2.** غموض موقف المبحوث.

**3.** المشارك (المبحوث) قد لا يجد بين الإجابات الجاهزة ما يريده.

**4.** لا يجد الباحث من بين الإجابات ما يعبر عن تردد المبحوث أو وضوح اتجاهاته.

**2. الاستبانة المفتوحة أو الحرة:**

تكون أسئلتها غير محددة الإجابة وفيه تكون الإجابة حرة مفتوحة، حيث يحتوي على عدد من الأسئلة يجيب عليها المشارك بطريقته ولغته الخاصة مثل: ما هي مقترحاتك لتطوير الجامعة، والأسئلة المقالية بهدف إعطاء المشارك فرصة لأن يكتب رأيه ويذكر تبريراته للإجابة بشكل كامل وصريح.

ويتميَّز بأنَّه أداة لجمع حقائق وبيانات ومعلومات كثيرة غير متوفِّرة في مصادر أخرى.

ومن عيوبه أنه يتطلب جهدا ووقتا وتفكيرا جادا من المشارك مما قد لا يشجعه على المشاركة بالإجابة.

**3. الاستبيان المصور:**

فيه تقدم أسئلة على شكل رسوم أو صور بدلا من العبارات المكتوبة ويقدم هذا النوع من الاستبيانات إلى الأطفال أو الأميّين، وقد تكون تعليمات شفهية.

- **مميزات الاستبيان المصور:**

**1.** يساعد في الحصول على بيانات حساسة أو محرجة لا يستطيع المشارك الحصول عليها في المقابلة.

**2.** يعطى المشارك فرصة كافية للتفكير دون ضغوط نفسية عليه كما هو الحال في المقابلة أو الاختبارات.

**3.** أكثر تمثيلاً للمشاركة المدروسة لأنه يمكن توزيع فقراته على جوانبها، كما هو الحال في استفتاءات الرأي العام.

**4.** تتوفر للاستبيان ظروف التقنين المناسب، فالألفاظ يمكن تخيرها والأسئلة يمكن ترتيبها والإجابات يمكن تسجيلها.

- **عيوب الاستبيان المصور**:

**1.** عدم جدية المشاركين في الإجابة أو اللجوء إلى الإجابة العشوائية.

**2.** قد يفسر المشارك بعض الأسئلة تفسيرا خاطئا فتأتي إجابته غير دقيقة.

**3.** يتأثر المشارك في الاستبيان بطريقة وضع الأسئلة، ويكتشف هدف الباحث فيميل إلى الإجابة التي ترضي الباحث.

**4.** التخلف عن إعادة الاستبيان إلى الباحث يقلل من تمثيل العينة لمجتمع الدراسة وينتج عن ذلك عدم صلاحية النتائج للتعميم.

**5.** يعتمد الاستبيان على القدرة اللفظية في الإجابة عليها لهذا فهو لا يصلح للأشخاص غير الملمّين بالقراءة والكتابة إلا إذا كان الاستبيان مصورا.

- **مواصفات الاستبيان الجيد:**

**1.** اللغة المفهومة والأسلوب الواضح الذي لا يتحمل التفسيرات المتعددة لأن ذلك يسبب إرباكا لدى المبحوثين مما يؤدي إلى إجابات غير دقيقة.

**2.** مراعاة الوقت المتوفر لدى المبحوثين وبالتالي يجب ألا تكون الأسئلة طويلة حتى لا تؤدي إلى رفض المبحوثين الإجابة على الاستبيان أو تقديم إجابات سريعة وغير دقيقة.

**3.** إعطاء عدد كافي من الخيارات المطروحة مما يمكن المبحوثين من التعبير عن آرائهم المختلفة تعبيرا دقيقًا.

**4.** استخدام العبارات الرقيقة واللائقة المؤثرة في نفوس الآخرين مما يشجعهم على التجاوب والتعاون في تعبئة الاستبيان مثل: (رجاء – شكرا ... الخ).

**5.** التأكد من الترابط بين أسئلة الاستبيان المختلفة وكذلك الترابط بينها وبين موضوع البحث ومشكلته.

**6.** الابتعاد عن الأسئلة المحرجة التي من شأنها عدم تشجيع المبحوثين على التجاوب في تعبئة الاستبيان.

**7.** الابتعاد عن الأسئلة المركبة التي تشتمل أكثر من فكرة واحدة عن الموضوع المراد الاستفسار عنه لأن في ذلك إرباك للمبحوثين.

**8.** تزويد المبحوثين بمجموعة من التعليمات والتوضيحات المطلوبة في الإجابة وبيان الغرض من الاستبيان ومجالات استخدام المعلومات التي سيحصل عليها الباحث مثال: بعض الاستفسارات تحتمل التأثير على أكثر من مربع واحد لذا يرجى التأشير على المربعات التي تعكس الإجابات الصحيحة.

- **تصميم الاستبيان وصياغته:**

بعد تحديد مشكلة الدراسة وتحديد أهدافها وصياغة فروضها وأسئلتها عقب استطلاع الدراسات السابقة وما كُتب من موضوعات تتَّصل بها فيتبيَّن للباحث أنَّ الاستبيان هو الأداة الأنسب لجمع البيانات والمعلومات اللازمة فإنَّ عليه تصميم استبيان فلابد من اتِّباع الخطوات التالية:

1. تحديد الموضوع العام للبحث وتقسيم موضوع البحث إلى عناصره الأوليَّة وترتيبها في ضوء علاقاتها وارتباطاتها وتحديد نوع البيانات والمعلومات المطلوبة لدراسة مشكلة البحث في ضوء أهداف البحث وفروضه وأسئلته حتى يتسنى للباحث تغطية كل عنصر بمجموعة من الأسئلة التي تشكل في مجموعها العام الأسئلة التي يتألف منها الاستبيان عند التطبيق.

2. تحديد عيِّنة الدراسة بنوعها ونسبتها وأفرادها أو مفرداتها بحيث تمثِّل مجتمعَ البحث.

3. تقويم الأسئلة ويتم ذلك بمراجعة أولية للأسئلة والتأكد من تغطية الأسئلة لكافة الموضوعات الفرعية والعامة وعرض الأسئلة على مجموعة من الأفراد لتلقي المزيد من الملاحظات.

4. طباعة الأسئلة بشكلها النهائي في نموذج خاص ثم توزيعها على المشاركين في البحث.

5. تحكيم استبانة الدراسة من قبل ذوي الخبرة في ذلك والمختصِّين بموضوع دراسته.

6. تجريب الاستبانة تجريباً تطبيقيّاً في مجتمع البحث لاستكشاف عيوبها أو قصورها.

7. صياغة استبانة الدراسة صياغة نهائيَّة وفق ملاحظات واقتراحات محكِّميها وفي ضوء تجربتها التطبيقيِّة.

ويجب على الباحث:

1. الإيجاز بقدر الإمكان.

2. تجنُّب الخلط بين إبداء الرأي وإعطاء الحقائق.

3. حسن الصياغة ووضوح الأسلوب والترتيب وتخطيط الوقت.

4. إيضاح أساليب إعادة نسخ الاستبانة وتسهيل ذلك ما أمكن.

5. الابتعاد عن الأسئلة الإيحائيَّة الهادفة إلى إثبات صحَّة فرضيَّات دراسته.

**6.** احتواء الاستبيان على أسئلة مرجعية للتأكُّد من صدق البيانات وانتظامها.

**7.** وَعْد المبحوثين بسريَّة إجاباتهم وأنَّها لن تستخدمَ إلاَّ لغرض البحث المشار إليه.

**8.** صياغة بدائل الإجابات المقترحة صياغة واضحة لا تتطلَّب إلاَّ اختياراً واحداً.

**9.** استخدام المصطلحات الواضحة البسيطة، وشرح المصطلحات غير الواضحة.

**10.** تزويد الاستبانة بما يشرح أهداف الدراسة وقيمتها التطبيقيَّة للأفراد أو المجتمع.

**11.** إعطاء المبحوثُ مساحةً حرَّة في نهاية الاستبانة لكتابة ما يراه من إضافة أو تعليق.

**12.** تجنُّب الأسئلة التي تستدعي تفكيراً عميقاً من المبحوثين أو المتعاونين مع الباحث.

**13.** إشارة الباحث إلى رقم هاتفه لتسهيل استفسار المبحوثين أو المتعاونين إن لزم ذلك.

**14.** مراجعة نسخ الاستبانة العائدة والتخطيط لتصنيف بياناتها وجدولتها وإعداد البرنامج الحاسوبيِّ الخاص بتفريغها.

**15.** البعد عن الأسئلة التي تتطلَّب معلومات وحقائق موجودة في مصادر أخرى؛ ممَّا يولِّد ضيقاً لدى المبحوث أو المتعاون مع الباحث.

- **مزايا وعيوب الاستبيان:**

تعرَّضت أداة الاستبيان إلى نقد شديد من المهتمِّين بأساليب البحث العلميِّ، ومعظم انتقاداتهم تركَّزت على مدى دقَّة وصحَّة البيانات والمعلومات التي يجمعها الباحث بهذه الأداة، وبرغم ذلك فإلى جانب عيوب أداة الاستبيان فلها مزايا تجعلها من أهمِّ أدوات جمع البيانات وأكثرها شيوعاً.

- **مزايا الاستبيان:**

**1.** تمكِّن أداة الاستبيان من حصول الباحثين على بيانات ومعلومات من وعن أفراد ومفردات يتباعدون تباعدًا جغرافيّاً بأقصر وقتٍ مقارنة مع الأدوات الأخرى.

**2.** يعدُّ الاستبيان من أقل أدوات جمع البيانات والمعلومات تكلفة سواءٌ أكان ذلك بالجهد المبذول من قبل الباحث أم كان ذلك بالمال المبذول لذلك.

**3.** تعدُّ البيانات والمعلومات التي تتوفَّر عن طريق أداة الاستبيان أكثر موضوعيَّة ممَّا يتوفَّر بالمقابلة أو بغيرها، بسبب أنَّ الاستبيان لا يشترط فيه أن يحملَ اسم المستجيب ممَّا يحفزه على إعطاء معلومات وبيانات موثوقة.

**4.** توفِّر طبيعة الاستبيان للباحث ظروف التقنين أكثر ممَّا توفِّره له أدواتٌ أخرى، وذلك بالتقنين اللفظيِّ وترتيب الأسئلة وتسجيل الإجابات.

**5.** يوفِّر الاستبيان وقتاً كافياً للمستجيب أو المتعاون مع الباحث للتفكير في إجاباته ممَّا يقلِّلُ من الضغط عليه ويدفعه إلى التدقيق فيما يدوِّنه من بيانات ومعلومات.

- **عيوب الاستبيان:**

**1.** لا يمكن استخدام الاستبيان في مجتمع لا يجيد معظمُ أفراده القراءة والكتابة.

**2.** قد لا تعود إلى الباحث جميعُ نسخ استبيانه؛ ممَّا يقلِّل من تمثيل العيِّنة لمجتمع البحث.

**3.** قد تكون الانفعالات من المعلومات المهمَّة في موضوع الدراسة، وبالاستبيان لا يتمكَّن الباحث من ملاحظة وتسجيل ردود فعل المستجيبين لفقدان الاتِّصال الشخصيِّ معهم.

**4.** قد يعطي المستجيبون أو يدوِّن المتعاونون مع الباحث إجابات غير صحيحة، وليس هناك من إمكانيَّة لتصحيح الفهم الخاطئ بسبب الصياغة أو غموض المصطلحات وتخصُّصها.

**5.** لا يمكن التوسُّع في أسئلة الاستبيان خوفاً من ملل المبحوث أو المتعاون مع الباحث حتى ولو احتاجت الدراسة إلى ذلك.

- **طرق أخرى لجمع البيانات ميدانياً:**

1. **الاتصال الهاتفي:**

حيث يتصل الباحث بالأفراد هاتفيا ويسجل إجاباتهم حول أسئلة البحث وهذه الطريقة تستخدم في بحوث التسويق أو استطلاع الرأي.

- **مميزات الاتصال الهاتفي:**

**1.** سرعة جمع البيانات.

**2.** قلة التكاليف وقصر الوقت.

**3.** السيطرة على طريقة توجيه الأسئلة.

- **عيوب الاتصال الهاتفي:**

**1.** عدم شمول الأفراد الذين ليس لديهم هاتف.

**2.** لا تصلح في حالة وجود عدد كبير من الأسئلة.

**3.** عدم إعطاء البيانات من قبل المشمولين لشخص غريب أو مجهول.

2. **المراسلة والبريد:**

ترسل استمارة (استمارة الأسئلة) مع تعليمات ملئها.

- **مميزات المراسلة والبريد:**

1. قلة التكاليف.
2. إعطاء الوقت الكافي للأفراد لدراسة الأسئلة والإجابة عليها.
3. تلافي التأثير الشخصي على الأفراد المشمولين بالبحث من قبل جامعي البيانات.

- **عيوب المراسلة والبريد:**

1. عدم اهتمام الأفراد بإعادة الأسئلة.
2. احتمال أن يرد على الأسئلة شخص آخر.
3. التضحية ببعض الأسئلة لإبعاد الملل لدى الأفراد.
4. نقص في عدد الردود عن العدد الذي حدده البحث وهذا مصدر من مصادر التحيز.

**(2–4) الأخطاء الشائعة في جمع البيانات:**

هناك بعض الأخطاء التي تحدث عندما يقوم الباحث بجمع البيانات والمعلومات التي تخص بحثه:

**1) خطأ التحيز:**

هذا الخطأ يرتكبه المصدر أو المفردة الإحصائية التي تزود الباحث بالمعلومات سواء بقصد أو بغير قصد، أو يحدث هذا الخطأ أحيانا عندما يستقي الباحث معلومات بحثه ليست من مصادرها الأصلية بل من مصادرها غير المباشرة.

**2) خطأ الصدفة:**

هذا الخطأ يرتكبه الباحث بنفسه سواء بتعمد أو بصورة غير متعمدة حيث يستقي معلومات بحثه بالاعتماد على ذاكرته بسبب بعد المفردة الإحصائية عنه، أو لأي سبب شخصي آخر وهذا سيؤدي إلى الحصول على نتائج واستنتاجات غير دقيقة وبعيدة عن الواقع.

**خطوات الدراسة الميدانية:**

**1.** تحديد مجتمع الدراسة.

**2.** تحديد البيانات اللازمة للدراسة.

أ. **بيانات ديموغرافية (شخصية):**

وهي بيانات عن عينة الدراسة مثل: (المكان، المؤهل العلمي، العمر، والجنس، والخبرة العملية، المستوى الوظيفي، والحالة الاجتماعية) وهي بيانات تفيد في التعرف على خصائص مجتمع الدراسة.

**ب. بيانات أساسية:**

وهي البيانات التي تم جمعها عن المشكلة موضوع الدراسة الميدانية باستخدام الاختبارات أو المقاييس أو أدوات آخري كالاستبانة.

**3.**تحديد أدوات جمع البيانات اللازمة للدراسة.

أ. الملاحظة.

**ب**. المقابلة الشخصية.

**ج**. استبانة.

**4.** بناء أدوات الدراسة.

أ. تحديد الأبعاد للمقياس أو أداة الدراسة.

**ب**. كتابة فقرات أداة المقياس.

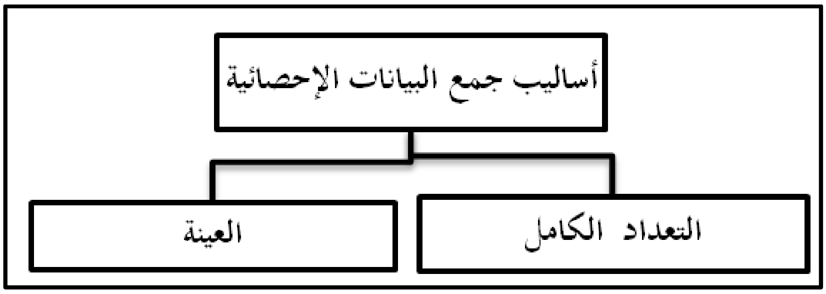
**ج**. تقدير صدق المحتوى للمقياس.

**د**. تطبيق الأداة على عينة استطلاعية لتقدير صدق وثبات الأداة.

**5.** مقاييس الدراسة.

**6.** معالجة البيانات وتحليلها.

**(2 – 5) أساليب جمع البيانات الإحصائية:**




**شكل (2.2)**: يوضح أساليب جمع البيانات الإحصائية

أ. **التعداد الكامل**: **Full Census**

هو جمع البيانات الخاصة بظواهر معينة في المجتمع باستخدام جميع الوحدات المكونة لذلك المجتمع.

أي الحصر الشامل لجميع مفردات المجتمع، وفي هذه الحالة يتم جمع بيانات عن كل مفردة من مفردات المجتمع بلا استثناء، كحصر جميع المزارع التي تنتج محصول معين، ويتميز

أسلوب الحصر الشامل بالشمول وعدم التحيز، ودقة النتائج، ولكن يعاب عليه أنه يحتاج إلى الوقت والمجهود، والتكلفة العالية.

**ب. العينة: Sample**

عبارة عن جمع بيانات معينة باستخدام جزء فقط من المجتمع الأصلي.

أو هي جزء من مفردات (الظاهرة أو الأفراد) يتم اختيارها من المجتمع محل الدراسة بحيث يمثل هذا الجزء مجتمع البحث.

- **العوامل المؤثرة في تحديد حجم العينة:**

**1. درجة التعميم:**

كلما ازدادت حاجة الباحث ورغبته بأن تكون نتائج بحثه قابلة للتعميم بشكل كبير على مجتمع الدراسة الأصلي كلما توجب عليه زيادة حجم العينة المختارة.

**2. مدى التجانس أو التباين في خصائص مجتمع الدراسة الأصلي:**

كلما كانت خصائص المجتمع الأصلي متجانسة كلما كان حجم العينة المطلوبة صغيرا نسبيا، وهناك ضرورة لزيادة حجم العينة حينما يوجد خلافات جوهرية هامة وعديدة بين أفراد أو مشاهدات مجتمع الدراسة الأصلي.

**3. حجم مجتمع الدراسة الأصلي:**

يجب الأخذ في الاعتبار أنه كلما ازدادت عناصر أو مشاهدات مجتمع الدراسة الأصلي، زاد حجم العينة المطلوبة والعكس صحيح، مع ملاحظة أن نسبة العينة إلى مجتمع الدراسة الأصلي تقل كلما زاد حجم المجتمع الأصلي.

ويجب مراعاة الآتي عند تحديد حجم العينة:

**1.** حجم العينة الذي يتراوح بين (30 – 500) مفردة يعتبر ملائما لمعظم أنواع البحوث.

**2.** عند استخدام العينة الطبقية، أي تقسيم المجتمع الأصلي إلى طبقات مثل: (ذكور – إناث) فإن حجم العينة لكل فئة يجب ألاّ يقل عن (30) مفردة.

**3.** في حالة استخدام الانحدار المتعدد أو الاختبارات المماثلة له فإن حجم العينة يجب أن يكون أضعاف متغيرات الدراسة، ويفضل أن يكون حجم العينة هنا (10) أضعاف متغيرات الدراسة.

**4.** في بعض أنواع البحوث التجريبية، التي يكون فيها حجم الضبط والرقابة عالياً، فإن حجم عينة مقداره (10) إلى (20) مفردة يكون مقبولاً.

- **الشروط الواجب مراعاتها عند اختيار العينة:**

**1.** أن تكون التكاليف اللازمة لاختيار ودراسة العينة صغيرة.

**2.** أن تكون ممثلة للمجتمع، أي تحتوي على جميع خصائص المجتمع المراد دراسته.

**3.** أن تكون النتائج التي سنحصل عليها من العينة قريبة جداً من النتائج الأصلية.

**4.** أن تكون الظاهرة المراد عمل معاينة لها سائدة ومنتشرة في المجتمع الأصلي ولا تكون نادرة الحدوث.

- **مميزات استخدام العينة:**

**1.** تقليل الوقت والجهد.

**2.** الحصول على بيانات أكثر تفصيلاً.

**3.** تستخدم لأنها أقل عرضة للأخطاء بالمقارنة بطريقة الحصر الشامل.

**4.** في حال عدم إمكانية إجراء حصر كامل لعناصر مجتمع الدراسة الأصلي، فهي جزء من كل، على أن تمثل الكل تمثيلاً صحيحاً وتحت شروط مضبوطة، مثل معاينة دم المريض.

- **عيوب استخدام العينة (أخطاء المعاينة):**

**1.** يزداد احتمال ورود خطأ الصدفة كلما صغر حجم العينة.

**2.** جمع بيانات ناقصة كنتيجة لإهمال العامل الجغرافي عند دراسة المستوى الاقتصادي للسكان.

**3.** التحيز الشخصي حينما يأخذ الباحث عينته المختارة من فئة معينة لها خصائص مميزة عن المجتمع الكلي.

**خطوات اختيار العينة:**

تتكون عملية اختيار العينة من عدة خطوات هي:

**1.** يحدد الباحث المجتمع الأصلي بدقة.

**2.** يعد قائمة كاملة ودقيقة بمفردات هذا المجتمع وتسمى (إطاراً) من خلال السجلات، ويجب أن تكون كاملة وحديثة.

**3.** يأخذ مفردات ممثلة من القائمة التي أعدها.

**4.** يحصل على عينة كافية لتمثيل المجتمع الأصلي بخصائصه التي يريد دراستها.

**(2 – 6) العينة الإحصائية:**

تصنف المعاينة بناء على كيفية سحب العينة إلى نوعين رئيسين هما: المعاينة العشوائية (الاحتمالية)، والمعاينة غير العشوائية (غير الاحتمالية).

**1. العينة الاحتمالية Probability Sample**

وتسمى أيضاً **المعاينة العشوائية** وهي أسلوب عن طريقه نختار وحدات العينة باستخدام الاحتمالات مثل: اختيار شخص من عدد كبير جداً من الأشخاص ليحصل على جائزة معينة في شهادة الاستثمار.

أي يتم فيها اختيار الأفراد بشكل عشوائي بحيث يعطى لكل عنصر من عناصر مجتمع الدراسة فرصة للظهور في العينة، وتكون هذه الفرصة معروفة ومحددة مسبقاً، ومن أهم أنواع العينات الاحتمالية، ما يلي:

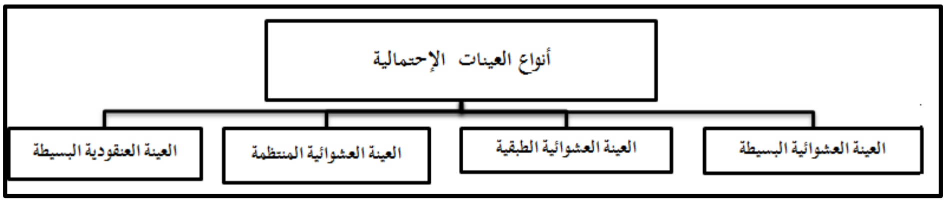
**1.** العينة العشوائية البسيطة Simple Random Sample.

**2.** العينة العشوائية الطبقية Stratified Random Sample.

**3.** العينة العشوائية المنتظمة Systematic Random Sample.

**4.** العينة العنقودية أو المتعددة المراحل Cluster Sample.

وتستخدم في اختبار الفرضيات السببية والعلائقية.

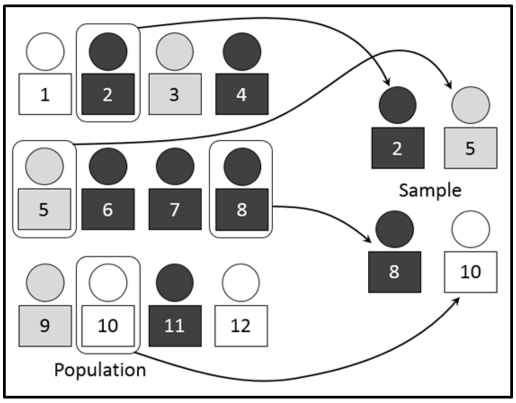



**شكل (3.2)**: يوضح تصنيف المعاينات الاحتمالية

**أولاً: أنواع المعاينات الاحتمالية:**

**أ. العينة العشوائية البسيطة: Simple Random sample**

هي العينة التي تؤخذ من المجتمع بطريقة تضمن أن يكون لكل فرد في هذا المجتمع فرصة متكافئة لاختياره ضمن هذه العينة مع توفر الوضوح في تعريف المجتمع، وعدم تكرار أي مفردة. أي يتم فيها حصر ومعرفة كامل العناصر التي يتكون منها مجتمع الدراسة الأصلي، ثم اختيار عينة مكونة من n وحدة من بين N وحدة من وحدات المجتمع محل الدراسة ويتم اختيار الوحدات الإحصائية على أساس تكافؤ الفرص بحيث يكون لكل عنصر من عناصر مجتمع الدراسة فرصة للظهور في العينة، والشكل التالي يوضح ذلك.




**شكل (4.2)**: يوضح طريقة اختيار عينة عشوائية

وتعتمد إجراءات سحب العينة العشوائية البسيطة على وجود إطار للعينة ونميز هنا بين نوعين من السحب:

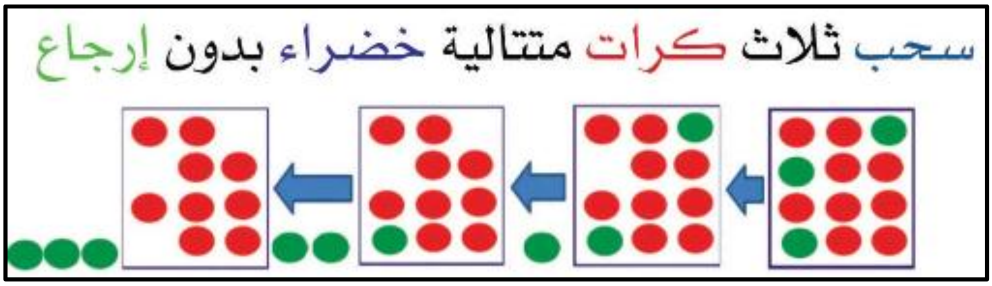
- **السحب مع الإرجاع:**

بمعنى سحب وحدات المعاينة مع إرجاع الوحدة المسحوبة في كل مرة، في هذه الحالة تسمى العينة بالعينة المستقلة، بمعنى أن الوحدة يمكن أن تظهر أكثر من مرة في العينة وبالتالي يكون عدد العينات الممكنة في هذه الحالة، مع مراعاة ترتيب الوحدات هو $N^n$.

- **السحب بدون إرجاع:**

يتم فيه سحب وحدات المعاينة دون إرجاع الوحدة المسحوبة، وتسمى العينة في هذه الحالة بالعينة غير المستقلة، بمعنى أن الوحدة المسحوبة لا تظهر إلا مرة واحدة في العينة وبالتالي عدد العينات الممكنة في هذه الحالة (مع عدم مراعاة الترتيب) هو $C_N^n$ حيث:

$$C_N^n = \frac{N!}{n!(N-n)!}$$




**شكل (5.2)**: يوضح طريقة السحب بدون إرجاع

وفي الواقع العملي إن المعاينة العشوائية البسيطة عن طريق السحب مع الإرجاع نادرة الاستعمال؛ لأنه لا معنى للحصول على الوحدة نفسها مرتين في العينة، كما أن السحب دون إرجاع يعطي تقديرات أكثر دقة، بالإضافة إلى أن التباين المرتبط بالعينة غير المستقلة يكون دائما أقل من المرتبط بالعينة المستقلة، وتستخدم عندما يكون حجم مجتمع الدراسة صغيراً أو محصورا أو معروف العدد.

**الخطوات المتبعة في سحب عينة عشوائية بسيطة:**

**1.** إعداد قوائم تتضمن جميع عناصر المجتمع (الإطار).

**2.** ترقيم جميع وحدات المجتمع بأرقام متسلسلة.

**3.** تحديد حجم العينة المطلوب سحبها.

**طرق المعاينة العشوائية:**

**1. القرعة:**

حيث يتم ترقيم أفراد المجتمع الأصلي ووضع الأرقام في صندوق خاص ويتم سحب الأرقام حتى نستكمل العدد المناسب للعينة.

وكمثال يمكن تطبيق هذه في المجتمعات الصغيرة فإذا كانت لدينا مجموعة من المفردات قدرها 50 مفردة، ومطلوب اختيار عينة مكونه من 5 مفردات منها، فما علينا إلا أن نأتي بخمسين ورقة صغيرة بنفس الشكل والحجم ونكتب على كل منهما رقماً يبدأ من 1 إلى 50 ثم نطويها بإتقان، ونضعها في صندوق مفتوح ونخلطها جيداً، ثم نسحب منها 5 ورقات بدون

تفضيل أي ورقة على أخرى، ونقرأ الأرقام الموجودة على هذه الورقات فتكون هي أرقام المفردات المطلوبة في العينة العشوائية.

وبهذه الطريقة البسيطة أمكننا أن نستخلص العينة الممثلة بدون دافع شخصي أو غرض خاص، وبدون أي تحيز ونكون قد أعطينا كل مفردة في المجتمع المراد فحصه نفس الفرصة في الاختيار.

**2. جداول الأعداد العشوائية: Tables of random numbers**

هي أعداد صحيحة مكونة من الأرقام ،8,9... 0,1,2,3، تم بناؤها على أساس عشوائي ووضعت ضمن جداول مكونة من صفوف وأعمدة، انظر الجدول الموضح أدناه.

وأفضل طريقة تلك التي تستخدم جداول الأعداد العشوائية كتلك التي أعدها "فيشر وييتس وكندال" Fisher, Yates, Kendall ولنفرض أن لدينا مجتمعا يتكون من 200 وحدة ونريد أن نختار عينة من 10 وحدات عن طريق استخدام الجداول العشوائية

ترقيم أفراد المجتمع من 1 – 200، ولابد أن يتكون كل عدد من ثلاث خانات مع ملاحظة أن عدد الخانات هنا يساوي عدد خانات أكبر عدد في المجتمع كالتالي:

| 2 خانة | |
|---|---|
| 0 | 1 |
| 5 | 0 |
| | |

| 3 خانات | | |
|---|---|---|
| 0 | 0 | 1 |
| 0 | 5 | 0 |
| 1 | 0 | 0 |

| 4 خانات | | | |
|---|---|---|---|
| 0 | 0 | 0 | 1 |
| 0 | 0 | 5 | 0 |
| 0 | 1 | 0 | 1 |
| 1 | 0 | 0 | 0 |
| | | | |

| 5 خانات | | | | |
|---|---|---|---|---|
| 0 | 0 | 0 | 0 | 1 |
| 0 | 0 | 0 | 5 | 0 |
| 0 | 0 | 1 | 0 | 1 |
| 0 | 1 | 0 | 0 | 0 |
| 1 | 0 | 0 | 0 | 0 |

**ولاستخدام الجداول العشوائية يجب اتباع الآتي:**

**1.** تحديد وتعريف المجتمع.

**2.** تحديد حجم العينة المرغوب فيه.

**3.** إعداد قائمة بأسماء أفراد المجتمع.

**4.** يتم وضع رقم مسلسل لكل فرد يبدأ من 0 بحيث يكون أول فرد فى القائمة الرقم أمامه 0، وتتزايد عدد الأصفار وفقاً لحجم العينة فإذا كان حجم العينة 400 يتم وضع 000 أما إذا كان حجم العينة 4000 يتم وضع 0000.

**5.** نضع إصبعنا على أي مكان فى الجدول وتكون هذه نقطة البداية فى اختيار العينة، أو نختار صفحة من الجداول العشوائية بطريقة عشوائية، وتختار الأعمدة الرأسية التي تعطينا أعدادا ذات ثلاثة أرقام وتقرأها من أسفل إلى أعلى أو من أعلى إلى أسفل أو من يمين إلى يسار أو العكس، وإذا اتبعنا نظاما في القراءة فلابد أن نلتزم به حتى يتم اختيار العينة.

وكمثال: لنفرض أننا قرأنا الأعمدة إلى أسفل فإننا ندوّن كل عدد أقل من 400 فإذا كان أول عدد يقرأ في الجدول هو 100 فان هذا يعني أن أول وحده يختار في العينة هي الوحدة رقم 100 وإذا كان العدد الثاني هو 375 اخترناه لهذه الوحدة، وإذا كان الثالث هو 084 اخترناه، ويليه 990 فإننا نهمله لأنه أكبر من 400 كما نهمل أي عدد يظهر لثاني مرة حيث أنه من غير الجائز سحب عدد مرتين حتى لا يسمح للوحدة الواحدة أن تختار أكثر من مرة.

وعندما نبدأ من أعلى ونستعمل الأعمدة الثلاثة التالية باستخدام هذه الطريقة قد نحصل على العينة التالية:

100، 375، 084، 128، 310، 118، 098، 125، 154، 235

044، 005، 369، 321، 195، 331، 186، 116، 226، 190

وذلك باستخدام جزء من جداول الأرقام العشوائية.

ومثال آخر: إذا كان حجم المجتمع 4500 أسرة، ونريد تقدير متوسط عدد أفراد الأسرة، من خلال سحب عينة عشوائية بسيطة حجمها 50 أسرة، نقوم باختيار وحدات العينة على النحو التالي:

يلاحظ أن عدد منازل حجم المجتمع 4 خانات، لذلك نحتاج إلى 4 أعمدة من جدول الأرقام العشوائية.

نختار أحد الأرقام عشوائيا من دون النظر إلى الجدول.

نفترض أن نقطة البداية كانت نقطة تقاطع السطر الرابع مع العمود الثاني، أي الرقم 4 نأخذ عن يمينه 3 خانات ونقرأ من الأعلى إلى الأسفل.

نحصل على الأرقام التالية والتي تمثل أرقاما للأسر المختارة:

152، 1678، 1027، 994، 1144، 4168، 652............

ونستمر بذلك حتى نحصل على 50 رقما عشوائيا مع إهمال أي أرقام تتجاوز حجم المجتمع (4500) أو تتكرر مرة أخرى وبذلك نحصل على أرقام 50 وحدة من وحدات المجتمع

والتي شكلت عينة عشوائية بسيطة، ثم نقوم بطرح السؤال المتعلق بعدد أفراد الأسرة على الوحدات المختارة.

والعينة العشوائية لا تمثل بالضرورة خصائص المجتمع الأصلي كله، ولكنها تترك اختيار الأفراد للصدفة، وبهذا تنقص إمكانية التحيز في اختيار العينة، ويمكن بطبيعة الحال بالصدفة أن يختار الباحث عينة لا تمثل المجتمع الأصلي الكلى بدقة.

**2. تحديد حجم العينة في حالة النسبة:**

يتحدد حجم العينة في حالة ما إذا كانت القيم محل القياس أو الدراسة عبارة عن نسبة كما يلي:

$$n = \frac{(1.96)^2 \times (p) \times (1-p)}{(e)^2}$$

حيث أن:

n: حجم العينة.

P: نسبة مفردات مجتمع الدراسة التي تتوافر فيها الخاصية محل الدراسة، وفي حالة عدم معرفة قيمة p ينصح بأن القيمة $p = 0.5$

e: الخطأ الأعظمي المقبول.

**3. تحديد حجم العينة في حالة المتوسط:**

يتحدد حجم العينة باستخدام الصيغة الرياضية التالية:

$$n = [\frac{1.96 \times 6}{e}]^2$$

حيث أن:

n: حجم العينة.

S: الانحراف المعياري للعينة.

e: الخطأ الأعظمي المقبول.

- **مميزات العينات العشوائية البسيطة:**

**1.** لا تتقيد بترتيب معين أو نظام مقصود.

**2.** لا تتطلب معرفة سابقة بخصائص مفردات المجتمع الأصلي.

**3.** تتفادى التحيز لاعتمادها على حد كبير على قانون الاحتمالات.

**4.** تعطي جميع مفردات المجتمع الأصلي نفس الفرصة المتكافئة في الاختيار.

- **عيوب العينات العشوائية البسيطة:**

**1.** استخدام جداول الأرقام العشوائية لتحديد كل مفردة عملية متعبة وخاصة إذا كانت العينة كبيرة وقد تأخذ جهداً كبيراً.

**2.** عدم توفر قوائم مسبقة لمجتمع الدراسة، وفى حالة توافرها فهي عادة ما تكون غير دقيقة.

**3.** في حالة عدم التأكد من تجانس مجتمع الدراسة يمكن أن تقود النتائج التي يتم التوصل إليها إلى أخطاء.

**ب. العينة العشوائية الطبقية: Stratified Random Sample**

يتم استخدامها اذا كان المجتمع غير متجانس ويمكن تقسيمه إلى طبقات أو فئات متجانسة وفقاً لخصائص معيّنة كالسنّ أو الجنس أو مستوى التعليم، وكتقسيم المدارس لدراسة وظيفتها في البيئة الخارجيَّة وفي المجتمع المحيط إلى مدارس حكوميَّة وأخرى مستأجرة، وبتقسيمها

بحسب مراحل التعليم، أو بحسب مجتمعها إلى مدارس في مجتمع حضريٍّ، ومجتمع قرويٍّ، ومجتمع بدويٍّ، ثمَّ يتم فيها اختيار أفراد العينة عن طريق تحديد نسبة كل طبقة إلي حجم المجتمع (بناء علي معرفة حجم كل طبقة في المجتمع) ويستخدم هذا النوع من العينات في المجتمعات التي تتكون من عدة طبقات حيث يمكن اختيار عينة تمثل المجتمع تمثيلاً حقيقياً.

ويتم هذا الأسلوب وفقا للخطوات التالية:

**1.** تحديد حجم العينة الإجمالي المطلوب اختياره من مجتمع البحث.

**2.** تقسيم المجتمع إلى طبقات متجانسة ويعتمد هذا التقسيم على التقدير الشخصي وخبرة الباحث.

**3.** توزيع مفردات العينة على طبقات مجتمع البحث لتمثيل هذه الطبقات.

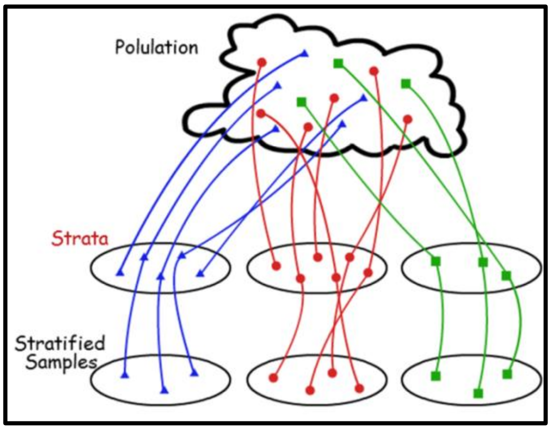
**4.** سحب عدد مفردات العينة من كل طبقة وفقا لطريقة العينة العشوائية البسيطة.

ونحصل على العدد الذي يمثل كل طبقة كالتالي:

العدد الذي يمثل الطبقة = (عدد أفراد الطبقة ÷ عدد أفراد المجتمع) × عدد أفراد العينة

العدد الذي يمثل الطبقة = (حجم الطبقة ÷ حجم المجتمع) × حجم العينة

والشكل التالي يوضح اختيار العينة العشوائية الطبقية:




**شكل (6.2)**: يوضح طريقة اختيار العينة العشوائية الطبقية

**مثال**: مجتمع يتكون من 2000 مفردة، مكون من أربع طبقات عدد المفردات بكل طبقة على التوالي 800، 600، 400، 200 مفردة، المطلوب سحب عينة طبقية، عدد مفرداتها 100 مفردة بحيث تمثل المجتمع تمثيلاً سليماً.

**الحل**:

∵ العدد الذي يمثل الطبقة = (عدد أفراد الطبقة ÷ عدد أفراد المجتمع) × عدد أفراد العينة.

العدد الممثل للطبقة = (حجم الطبقة ÷ حجم المجتمع) × حجم العينة

العدد الممثل للطبقة الأولى = (800 ÷ 2000) × 100= 40

العدد الممثل للطبقة الثانية = (600 ÷ 2000) × 100 = 30

العدد الممثل للطبقة الثالثة = (400 ÷ 2000) × 100= 20

العدد الممثل للطبقة الرابعة = (200 ÷ 2000) × 100= 10

إذن العينة الطبقية تتكون من (40، 30، 20، 10).

**مثال:** مجتمع حجمه 5000 مفردة، مكون من خمس طبقات حجم كل طبقة على التوالي 1500، 1400، 1300، 500، 300 مفردة، كون عينية طبقية حجمها 200 مفردة.

**الحل:**

العدد الممثل للطبقة = (حجم الطبقة ÷ حجم المجتمع) × حجم العينة

العدد الممثل للطبقة الأولى = (1500 ÷ 5000) × 200 = 60

العدد الممثل للطبقة الثانية = (1400 ÷ 5000) × 200 = 56

العدد الممثل للطبقة الثالثة = (1300 ÷ 5000) × 200 = 52

العدد الممثل للطبقة الرابعة = (500 ÷ 5000) × 200 = 20

العدد الممثل للطبقة الخامسة = (300 ÷ 5000) × 200 = 12

**مثال:** بفرض أن مجتمع مكون من 5000 مزرعة تمور منها 1000 حجم إنتاجها كبير، و1500 حجم إنتاجها صغير، و2500 حجم إنتاجها متوسط، والمطلوب سحب عينة طبقية عدد مفرداتها 1000 مزرعة بحيث تمثل المجتمع تمثيلاً سليماً.

**الحل:**

∵ العدد الذي يمثل الطبقة = (عدد المزارع ذات حجم محدد ÷ عدد المزارع) × عدد مزارع العينة.

العدد الذي يمثل الحجم الصغير = (1500 ÷ 5000) × 100 = 300 مزرعة.

العدد الذي يمثل الحجم المتوسط = (2500 ÷ 5000) × 100 = 500 مزرعة.

العدد الذي يمثل الحجم الكبير = (1000 ÷ 5000) × 100 = 200 مزرعة.

**أهم الطرق المستخدمة في تحديد حجم العينات المسحوبة من الطبقات:**

**(أ) طريقة التساوي Equal Method:**

وفيها يؤخذ عدد متساوي من كل طبقة، حتى ولو اختلف عدد الأفراد في كل منها، ويعاب عليها أنها تساوي بين الطبقات حتى في حالة الاختلاف.

**(ب) طريقة التناسب Proportional Method:**

ويؤخذ هنا عدد يتناسب مع النسبة التي تمثلها الطبقة من المجتمع الأصلي. فإذا كان لدينا مجتمعً حجمه (ن) ويمكن تقسيمه إلى عدة طبقات وأن حجم هذه الطبقات هو ط$_1$، ط$_2$، ط$_3$، .... ط$_ن$ ونفرض أننا اخترنا عينات من هذه الطبقات أحجامها ل$_1$، ل$_2$، ل$_3$، ...، ل$_ن$، وأن الحجم الكلى للعينات (ل).

**(ج) الطريقة المثلى Ideal Method:**

تعد هذه الطريقة من أدق الطرق، فهي لا تقصر تحديد العدد على نسبة كل طبقة للمجتمع الأصلي، بل تهتم بدرجة التباين داخل كل طبقة، فإذا كان كبيراً زاد العدد، وإذا كانت المجموعة متجانسة قل العدد.

- **مميزات العينات العشوائية الطبقية:**

**1.** يتحقق التمثيل، ليس فقط للمجتمع الأصلي، بل لكل طبقاته الفرعية مهما كان بعضها يشكل أقلية صغيرة.

**2.** أدق من العينة العشوائية البسيطة، لأنها تجمع العشوائية وبالتالي تحقق التكافؤ بين الأفراد، والحياد في الاختيار، والغرضية، فنضمن عدم خلوها من خصائص المجتمع الأصلي.

**3.** تتميز بالدقة الإحصائية وانخفاض نسبة حدوث الخطأ المعياري، خاصة كلما كانت المجموعات أو الطبقات متجانسة داخلياً.

- **عيوب العينات العشوائية الطبقية:**

**1.** تتطلب من الباحث التعرف وبشكل جيد على مجتمع دراسته لتحديد المجموعات التي يتكون منها.

**2.** تتطلب إجراءات كثيرة يجب على الباحث القيام بها قبل الشروع في استخدام أي من العينات العشوائية البسيطة أو المنتظمة.

**3.** يقوم الباحث بسحب عدد من العينات تبعاً لعدد مستويات المتغير الذي يتعامل معه مما يؤدي إلى مضاعفة الجهد الذي يقوم به.

**ج. العينة العشوائية المنتظمة: Systematic Sample**

يتم اختيار هذا النوع من العينات العشوائية في حالة تجانس مجتمع الدراسة الأصلي وتوافر إطاره، وسميت منتظمة لأن اختيار وحداتها بحيث تكون المسافة أو المدة بين كل وحدة وأخرى ثابتة لجميع وحدات العينة، وتقوم هذه العينة على تقسيم مجتمع الدراسة إلى عدد من المجموعات أو الفترات يتساوى عددها مع حجم العينة المطلوب سحبها، ويتحقق عنصر العشوائية في هذه العينة باختيار المفردة الأولى عشوائياً من المجموعة أو الفترة الأولى، بينما يتحقق عنصر الانتظام في اختيار بقية مفردات العينة من المجموعات (بنفس ترتيب المفردة التي تم اختيارها من المجموعة الأولى)، وتستخدم عندما يكون مجتمع الدراسة كبيرا وغير معروف العدد.

**الخطوات الواجب مراعاتها عند اختيار عينة عشوائية منتظمة:**

**1. تقسيم المجتمع إلى مجموعات:**

نقسم المجتمع إلى عدد (ن) من المجموعات أو الفترات المتساوية الطول وذلك باستخدام العلاقة التالية:

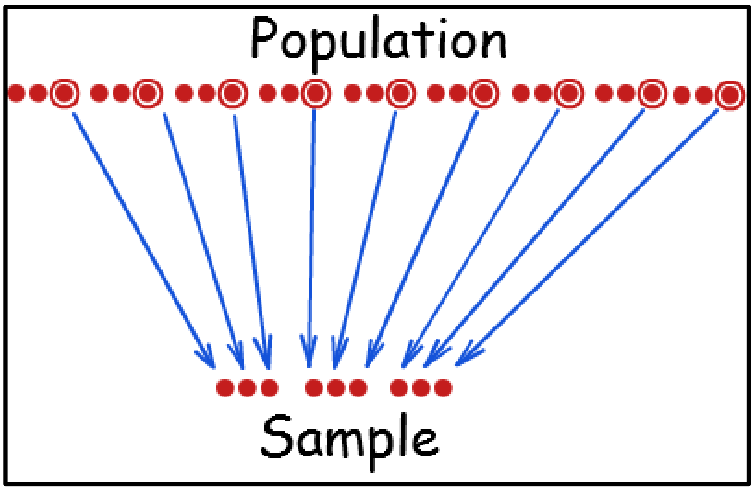
طول القسم (المجموعة) = مجموع وحدات مجتمع الدراسة ÷ عدد أفراد العينة

فمثلاً: إذا كان لدينا مجتمع مكون من 100 مفردة ونرغب في اختيار عينة عدد أفرادها 10 فإن عدد مفرداتها يصل إلى 10.

**2. اختيار العينة:**

يتم اختيار المفردة الأولى في العينة من المجموعة أو الفترة الأولى عشوائياً بطرق الاختيار.

ويتم اختيار هذا النوع من العينات العشوائية في حالة تجانس مجتمع الدراسة الأصلي وتوافر إطاره، وسميت منتظمة لأن المسافة بين كل رقم والذي يليه مسافة ثابتة، ويجب على الباحث أن يكون حذراً لئلا تكون القائمة مرتبة وفق ترتيب معين يجعل الاختيار غير عشوائي تماماً كما يتضح من الشكل التالي:




**شكل (7.2)**: يوضح طريقة اختيار العينة العشوائية المنتظمة

**مثال**: إذا كان المجتمع حجمه 50 مفردة، ويراد سحب عينة عشوائية منتظمة حجمها 10،

ولكي يتم ذلك يتبع الآتي:

طول القسم (المجموعة) = مجموع وحدات مجتمع الدراسة ÷ عدد أفراد العينة

طول القسم (المجموعة) = 50 ÷ 10 = 5

أي تكوين جدول يحتوي على عدد 5 صفوف و10 أعمدة بعدد مفردات العينة، ومن ثم يكون كل عمود به 5 أرقام متتالية كما يلي:

**الحل:**

تكوين عدد 10 عمود أي بعدد مفردات العينة، ومن ثم يكون كل عمود به 5 أرقام متتالية كما يلي:

| $C_1$ | $C_2$ | $C_3$ | $C_4$ | $C_5$ | $C_6$ | $C_7$ | $C_8$ | $C_9$ | $C_{10}$ |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 6 | 11 | 16 | 21 | 26 | 31 | 36 | 41 | 46 |
| 2 | 7 | 12 | 17 | 22 | 27 | 32 | 37 | 42 | 47 |
| 3 | 8 | 13 | 18 | 23 | 28 | 33 | 38 | 43 | 48 |
| 4 | 9 | 14 | 19 | 24 | 29 | 34 | 39 | 44 | 49 |
| 5 | 10 | 15 | 20 | 25 | 30 | 35 | 40 | 45 | 50 |

**1.** اختيار رقم عشوائي من العمود الأول، وبفرض أن الرقم العشوائي هو 3، تكون المفردة الأولى في العينة هي المفردة رقم 3، وتحدد أرقام المفردات الأخرى بشكل منتظم وفقا لمتوالية عددية (حسابية) حدها الأول هو 3، وأساسها يساوي عدد مفردات كل عمود "5".

**2.** بتطبيق الصيغة الرياضية للمتوالية العددية:

$$ح_ن = أ + (ن - 1) د$$

حيث (ن) رقم الحد، (د) الأساس، ($ح_ن$) قيمة الحد، وعليه:

فقيمة الحد الأول $ح_1=3$

فقيمة الحد الثاني $ح_2 = 3 + (2-1)\ 5 = 3 + 5 = 8$،

فقيمة الحد الثالث $ح_3= 3 + (3-1)\ 5 = 3 + 10 = 13$

**3**. مفردات العينة هي:

| $C_1$ | $C_2$ | $C_3$ | $C_4$ | $C_5$ | $C_6$ | $C_7$ | $C_8$ | $C_9$ | $C_{10}$ |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 6 | 11 | 16 | 21 | 26 | 31 | 36 | 41 | 46 |
| 2 | 7 | 12 | 17 | 22 | 27 | 32 | 37 | 42 | 47 |
| 3 | 8 | 13 | 18 | 23 | 28 | 33 | 38 | 43 | 48 |
| 4 | 9 | 14 | 19 | 24 | 29 | 34 | 39 | 44 | 49 |
| 5 | 10 | 15 | 20 | 25 | 30 | 35 | 40 | 45 | 50 |

العينة العشوائية المنتظمة هي:

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **3** | **8** | **13** | **18** | **23** | **28** | **33** | **38** | **43** | **48** |

- **مميزات العينات العشوائية المنتظمة:**

**1**. تعد من أسهل العينات العشوائية في التطبيق.

**2**. لا تحتاج إلى عملية إعداد مسبق لمفردات الدراسة خاصة إذا كانت مجموعات داخل مجتمع الدراسة.

**3**. لا تحتاج إلى الرجوع في كل مرة يتم فيها سحب المفردات إلى مرجع أو دليل فيكتفي بالمفردة الأولى أما باقي المفردات فتحدد تلقائياً عن طريق صيغة رياضية سهلة ومبسطة.

- **عيوب العينات العشوائية المنتظمة:**

1. اقتصار العشوائية فقط في تحديد الرقم الأول في بداية اختيار العينة.

**2.** تستلزم توفر قائمة حديثة تشمل كافة أسماء مفردات المجتمع الأصلي.

**3.** قد تكون العينة المختارة غير متجانسة، وذلك حينما تختار مفردات على أبعاد منتظمة يصادف أن يكونوا من طبقة معينة أو من ذوي خصائص وصفات مميزة وغير متشابهة مع بقية المفردات.

**4.** يشترط في المجتمع الأصلي أن يكون الأفراد في تسلسل منسق وتدرج من حيث التنوع.

**5.** لا تحدث احتمالية فرصة التمثيل لمفردات مجتمع الدراسة إلا مرة واحدة وهي عند اختيار المفردة الأولى.

**6.** في حالة كون طول الفئة كبيراً وهناك مجموعات داخل مجتمع الدراسة عددها أقل من طول الفئة فإن احتمال تمثيل هذه المجموعة في العينة يكون محدوداً.

**7.** عدم صلاحيتها إذا ما وجدت علاقة دورية مع ترتيب العناصر في القائمة وكان طول الفترة بين عناصر العينة مساوياً لطول الدورة أو إحدى مضاعفاتها.

### د. العينة العنقودية البسيطة: Cluster Sample

عندما يكون مجتمع الدراسة مؤلفا من طبقات عديدة ومتباينة أو متباعدة جغرافيا وكبيراً جداً ومتناثراً على مساحات شاسعة تكلف الكثير من الوقت والجهد في التنقل بينها عند جمع البيانات، وفي حالة عدم وجود إطار يضم جميع مفردات المجتمع فيستحيل الاختيار العشوائي مباشر من المجتمع. لهذا يلجأ الباحث إلى أخذ العينة على مراحل متعددة متتالية.

ويتم تقسيم مجتمع الدراسة الأصلي فيها إلى عدة فئات حسب معيار معين (مجموعات جزئية واضحة) نسمى كلا منها عنقودا، ثم يختار منها العينة المحتملة المطلوبة بطريقة عشوائية من كل عنقود.

يستعمل هذا النوع من العينات في الحالات التي يكون فيها مجتمع الدراسة مكوناً من طبقات مختلفة، وهنا يضع الباحث شروطاً معينة لاختيار أفراد العينة بحيث تمثل العينة جميع أفراد المجتمع المدروس، وبنفس نسبة وجودها في المجتمع. أي أن الباحث يختار لكل طبقة وبطريقة عشوائية عددا من المفردات، يتناسب مع حجمها الحقيقي في المجتمع الأصلي.

فإذا كان حجم المجتمع (عدد المفردات في المجتمع) = N، ويتكون المجتمع من عدة طبقات (n) أو عناقيد، فيتم تقدير عدد المفردات الواجب تمثيلها في العينة الممثلة للمجتمع الإحصائي باستخدام الصيغة التالية:

**تقدير العينة الممثلة للمجتمع الإحصائي داخل كل عنقود** (فئة، طبقة)

= (عدد مفردات الطبقة ÷ عدد مفردات المجتمع) × عدد مفردات العينة في المجتمع الإحصائي.

ثم يختار بطريقة عشوائية من داخل العنقود (فئة، طبقة) العدد الممثل للعينة الذي تم تقديره.

**مثال:** ندرس مجتمع الطلبة يتكون من 200 طالب موزعة كالتالي:

عدد الطلاب الذكور (100)، والإناث (40)، والمتفوقين (60) طالب.

**المطلوب:** اختيار عينة ممثلة لمجتمع الطلبة عددها = 50 طالب

**الحل:**

∵ تقدير العينة الممثلة للمجتمع الإحصائي =

(عدد مفردات الطبقة ÷ عدد مفردات المجتمع) × عدد مفردات العينة في المجتمع الإحصائي.

تقدير العينة الممثلة للذكور = (200÷100) × 50 = 25 طالب يتم اختيارهم بطريقة عشوائية.

تقدير العينة الممثلة للإناث = (200÷40) × 50 = 10 طالبات يتم اختيارهم بطريقة عشوائية.

تقدير العينة الممثلة للمتفوقين = ( 200 ÷ 60) × 50 = 15 طالب يتم اختيارهم بطريقة عشوائية.

- **مميزات العينات العشوائية العنقودية:**

**1.** تتعامل مع كل المجتمعات المتجانسة بغض النظر عن حجمها بشرط أن يكون مجتمع الدراسة موزعاً في أكثر من مكان جغرافي.

**2.** أن جميع المجتمعات الفرعية المكونة لمجتمع الدراسة الأصلي تتشابه في الخصائص العامة بصورة كبيرة.

**3.** تناسب المجتمعات الكبيرة المتناثرة التي تشغل حيزاً جغرافياً شاسعاً.

**4.** يمكن استخدام كل من العينة العشوائية البسيطة والمنتظمة عند الانتقال من مرحلة إلى أخرى.

- **عيوب العينات العشوائية العنقودية:**

**1.** تتطلب خطوات كثيرة تبعاً لعدد المراحل كما تتطلب سحب عينات كثيرة أيضاً "عينة في كل مرحلة".

**2.** احتمال كبير ألا تكون العينة ممثلة للمجتمع.

**3.** انخفاض مستوى تمثيلها لمجتمع الأصل.

**4.** تحليل بياناتها غير مناسب باستخدام معظم أساليب الإحصاء الاستدلالي.

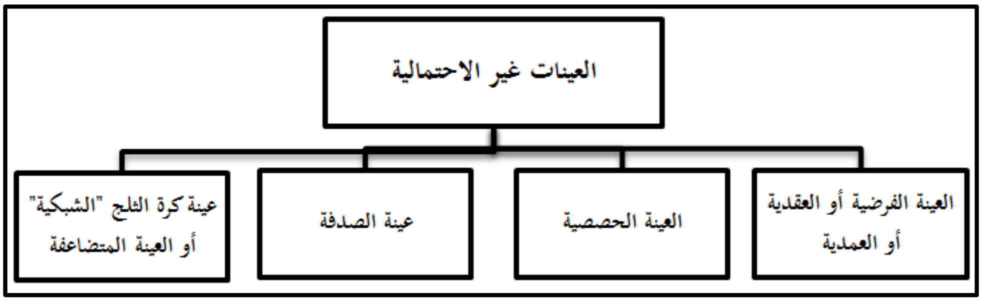
**ثانياً: العينات غير الاحتمالية: Non-probability samples**

هو أسلوب عن طريقه نختار وحدات العينة دون استخدام لعلم الاحتمالات. أو هي العينات التي يتم اختيارها بشكل غير عشوائي ولا تتم وفقا للأسس الاحتمالية المختلفة، وإنما تتم وفقا لأسس وتقديرات ومعايير معينة يضعها الباحث، فهنالك دراسات يصعب تحديد المجتمع الأصلي لها مثل: دراسة أحوال المدمنين، أو المنحرفين، أو المتهربين من الضرائب، تمثل هذه المجتمعات غير المحددة وأفرادها ليسوا معروفين فلا نستطيع أخذ عينة عشوائية منهم بحيث تمثلهم بدقة، فيعتمد الباحث إلى أسلوب العينة غير العشوائية ويختار عينة.

- **عيوب العينات غير الاحتمالية:**

**1)** لا تمثل المجتمع المأخوذ منه تمثيلاً صحيحاً، ولذلك فإن نتائجها لا تصلح للتصميم على المجتمع كله.

**2)** احتمال تحيز الباحث في الاختيار.




**شكل (8.2)**: يوضح تصنيف العينات غير الاحتمالية

- **أمثلة للعينات غير الاحتمالية:**

**1) العينة الفرضية أو العقدية أو العمدية: Purposive Sample**

يلجأ الباحث إلى هذه الطريقة فيما إذا كان مجتمع الدراسة كبيراً جداً وكانت إمكانياته لا تسمح له إلا بدراسة عينة حجمها صغير جداً بالنسبة لمجتمع الدراسة، في هذه الحالة يتعمد الباحث اختيار مفردات معينة كعينة لمجتمع الدراسة يرى بخبرته السابقة أن هذه العينة يمكن أن تعطي تمثيلاً مقبولاً لمجتمع الدراسة.

وهي العينة التي يعتمد الباحث اختيارها من وحدات معينة، لاعتقاده بأنها تمثل المجتمع الأصلي تمثيلاً صحيحاً.

أو سميت هذه العينة بهذا الاسم نظراً لان الباحث يقوم باختيارها طبقاً للغرض الذي يستهدف تحقيقه من خلال البحث، ويتم اختيارها على أساس توفر صفات محددة في مفردات العينة تكون هي الصفات التي تتصف بها مفردات المجتمع محل البحث.

فمثلا إذا أراد باحث دراسة آراء المستهلكين حول صنف من أصناف القهوة سريعة الذوبان (نس كافي) فعليه أن يختار عينة من الأفراد الذين لديهم بعض التجربة والمعرفة بهذا الصنف من القهوة، لأنه من غير المنطقي أن تتضمن العينة أفراداً لا يشربون هذا الصنف من القهوة.

- **عيوب العينة الفرضية أو العقدية أو العمدية:**

**1.** تقع في خطأ التحيز الذي يحدث عادة نتيجة اختيار مفردات البحث وفقا للرأي الشخصي للباحث.

**2.** عدم وجود أساس موضوعي للحكم على دقة نتائج البحث التي تم التوصل إليها وبالتالي مدى الاعتماد على هذه النتائج وتعميمها مستقبلا.

**3.** العينة العمدية تستلزم معرفة المعالم الإحصائية بالنسبة للمجتمع الأصلي وبالنسبة للوحدات التي يرغب الباحث في اختيارها، وهذا أمر قد لا يتيسر في جميع الأحوال.

**2) العينة الحصصية: Quota Sample**

هي نوع خاص من العينات غير العشوائية وتستخدم كثيراً في معاينة الرأي العام ويتم اختيار هذا النوع من العينات على أساس تقسيم مجتمع الدراسة إلى طبقات طبقا للخصائص التي ترتبط بالظاهرة محل البحث، ثم يختار الباحث عينة من كل طبقة من هذه الطبقات بحيث تتكون من عدد من المفردات يتناسب مع حجم الطبقة في المجتمع.

**مثال:** قد يسأل باحث المارة في أحد الشوارع عن رأيهم حول موضوع معين، ولكنه يختار من المارة أشخاصا من أعمار مختلفة لكي يمثل كل الفئات العمرية في مجتمع البحث. من الملاحظ أن هذه العينة تشبه إلى حد كبير العينة العشوائية الطبقية في تقسيم مجتمع الدراسة إلى طبقات، ثم يتم الاختيار من هذه الطبقات بما يتناسب مع وزنها النسبي في مجتمع الدراسة. إلا إن الفارق بينهما هو أسلوب اختيار أفراد كل طبقة، إذ لا يستعمل الأسلوب العشوائي في الاختيار في العينة الحصصية، بل يتم استعمال أسلوب الصدفة والقصد، ويستخدم هذا النوع من العينات في دراسة الرأي العام وفي الدراسات التربوية والاجتماعية.

**3) عينة الصدفة: Accidental Sample**

تتكون العينة من الأفراد الذين يقابلهم الباحث بالصدفة. فلو أراد الباحث إن يقيس الرأي العام للجمهور حول قضية ما فإنه يختار عدداً من الناس ممن يقابلهم بالصدفة سواء في الشارع أو في الباص.

**4) عينة كرة الثلج "الشبكية" أو العينة المتضاعفة: Snowball Sample**

يتعرف الباحث على فرد من المجتمع الأصلي، يقوده لفرد آخر وهكذا يتسع نطاق معرفة الباحث بهذا المجتمع، وتسمى بالعينة المتضاعفة.

تستخدم في حالة عدم توفر قائمة بكل أفراد المجتمع الأصلي.

**مثال**: يريد الباحث دراسة مجتمع المدمنين في مدينة ما، لا يجد أمامه إلا من هو في السجن أو مصحة علاجية، فعليه التعرف على أحدهم وتكوين علاقة معه لأنه سوف يقوده إلى مجموعة من زملائه المدمنين.

# الباب الثالث

## الاحتمالات

- مقدمة.

(3-1) بعض التعريفات الهامة.

(3-1.1) التجربة العشوائية.

(3-2.1) فضاء (فراغ) العينة.

(3-3.1) الحدث.

(3-4.1) تصنيف الأحداث.

(3-1.4.1) الأحداث المستقلة.

(3-2.4.1) الأحداث المتنافية (المانعة).

(3-3.4.1) الحدث المؤكد.

(3-4.4.1) الحدث المستحيل.

(3-2) العمليات على الأحداث.

(3-1.2) تقاطع الأحداث (أ ∩ ب).

(3-2.2) اتحاد الأحداث (أ ∪ ب).

(3-3.2) الفرق بين حدثين.

(3-4.2) الحدث المكمل.

(3-3) الاحتمال والحدث.

(3-1.3) احتمال حدث منفرد.

(3-2.3) الأحداث غير المنافية.

(3-3.3) الحوادث المتنافية.

(3-4) التعبير الرياضي عن قواعد حساب الاحتمالات.

**مقدمة:**

نظرية الاحتمالات تقدم قيما رقمية لتوقعاتنا الغير مؤكدة فهي تقدم القانون الرياضي الذي يساعدنا على التنبؤ بالنتائج غير المؤكدة.

فالاحتمالات تعتبر ذات أهمية لاستخدام الإحصاءات التحليلية وبالتالي تسهم في التخطيط والتنبؤ واتخاذ القرار.

**(3−1) بعض التعريفات الهامة:**

**(3−1.1) التجربة العشوائية:**

هي كل تجربة نستطيع أن نحدد مقدما (أي قبل إجرائها) جميع النواتج الممكنة الحدوث، ولكن لا يمكن تحديد أي من هذه النواتج سيتحقق فعلا عند إجراء هذه التجربة.

**من أمثلة التجارب العشوائية ما يأتي:**

تجربة إلقاء قطعة نقود وملاحظة الوجه الظاهر فنحن نعلم مقدما أن الناتجين لهذه التجربة هما وجه (الصورة) ووجه (الكتابة) ولكن لا يمكن تحديد أي الوجهين سيظهر فعلا عند إلقاء قطعة النقود.

تجربة إلقاء حجر نرد (زهرة الطاولة) وملاحظة العدد الذي سيظهر على الوجه العلوي له فنحن نعلم مقدما أن جميع النواتج التي يمكن أن تسفر عنها التجربة هي الأعداد (1،2،3،4،5،6) ولكن لا يمكن تحديد أي هذه الأعداد سيظهر فعلاً عند إلقاء الحجر.

**الأعداد الأولية:**

هي أعداد لا تقبل القسمة إلا على نفسها أو على الواحد الصحيح مثل: (2، 3، 5، 7. . . الخ).

**العدد المربـع الكامل:**

هو العدد الذي له جذر تربيعي (أي جذره يساوي عدد صحيح مثل: 1، 4، 9، 16 ، . . . الخ).

**(2.1–3) فضاء (فراغ) العينة:**

هو مجموعة لكل النواتج الممكنة الحدوث Set of all possible outcomes من تجربة عشوائية، ويرمز له بالرمز "ف" "S".

**Sample Space.** The set of all possible outcomes of a statistical experiment is called a *Sample space*. It is represented by the symbol S.

ويسمى كل عنصر من عناصر فضاء العينة عنصرا من عناصر فضاء العينة أو نقطة عينة.

Each element of a sample space is called an element of the sample space or a sample point.

- في التجربة عشوائية، **مجموعة النتائج الممكنة** تسمى **مجموعة الإمكانيات (عناصر، أحداث)** ويرمز لها بالرمز "ف" "S".
- ليكن أ **(A) جزءاً من** "ف" "S"، نقول عندئذ أن أ **(A) حادثة**.
- إذا احتوت المجموعة الجزئية أ **(A)** على **عنصر وحيد** فإنها تدعى **حادثة أولية**.

• ☜ "ف" "S" هي الحادثة الأكيدة و $\emptyset$ هي الحادثة المستحيلة. ($\emptyset$ الجزء الخالي).

**(3.1-3) الحدث:**

هو أي فئة جزيئية من فراغ العينة.

أي مجموعة من النتائج الواردة في فضاء العينة.

An event is any collection of outcomes contained in the sample space.

ويتكون الحدث البسيط بالضبط من عنصر واحد

A simple event consists of exactly one element

والحدث المركب يتكون من أكثر من عنصر واحد

A compound Event consists of more than one element.

**مثال (1):**

في تجربة إلقاء قطعة نقود مرة واحدة، اكتب فضاء (فراغ) العينة.

**الحل:** ف = {ص، ك} حيث {ص} ترمز لوجه الصورة، {ك} ترمز لوجه الكتابة.

**Example.** Consider the experiment of tossing a coin. The sample space S of possible outcomes may be written as
$S = \{H, T\}$.

**مثال (2):**

في تجربة إلقاء عملتان متوازنتان، اكتب فضاء (فراغ) العينة.

**الحل:** ف = {(ص، ص)، (ص، ك)، (ك، ص)، (ك، ك)}

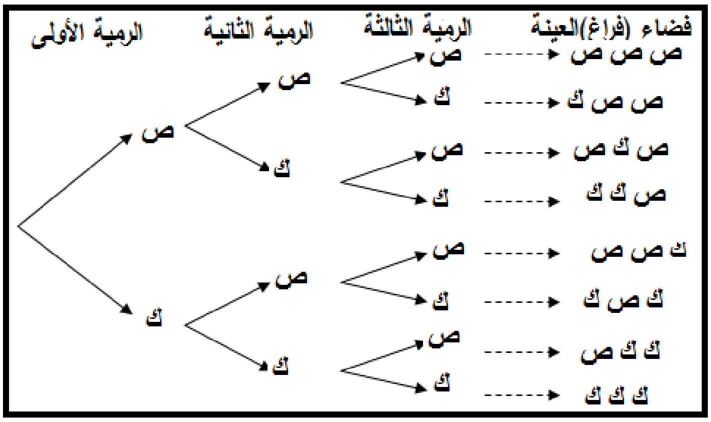
حيث {ص} ترمز لوجه الصورة، {ك} ترمز لوجه الكتابة.


**مثال (3):**

في تجربة إلقاء ثلاث عملات متوازنة (عملة ثلاث مرات)، اكتب فضاء (فراغ) العينة.

**الحل:** ف = {(ص، ص، ص)، (ص، ص، ك)، (ص، ك، ص)، (ص، ك، ك)، (ك، ص، ص)، (ك، ص، ك)، (ك، ك، ص)، (ك، ك، ك)}

حيث{ص} ترمز لوجه الصورة، {ك} ترمز لوجه الكتابة والشكل التالي يوضح ذلك.




**مثال (4):**

في تجربة إلقاء زهرة (حجر) نرد متوازنة، اكتب فضاء (فراغ) العينة.

**الحل**: **ف = {1، 2، 3، 4، 5، 6}**

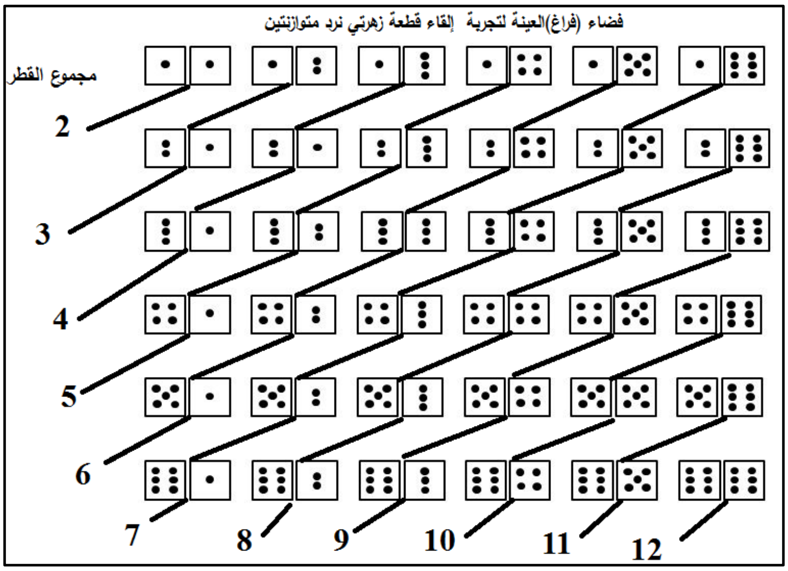
**Example.** Consider the experiment of flipping a die. Then the elements of the sample space S is listed as

$$S = \{1, 2, 3, 4, 5, 6\}.$$

**مثال** (**5**):

في تجربة إلقاء قطعة زهرتي نرد متوازنتين، اكتب فضاء (فراغ) العينة.

**الحل**: ف = {(1، 1)، (1، 2)، (1، 3)، (1، 4)، (1، 5)، (1، 6)، (2، 1)، (2، 2)، (2، 3)، (2، 4)، (2، 5)، (2، 6)، (3، 1)، (3، 2)، (3، 3)، (3، 4)، (3، 5)، (3، 6)، (4، 1)، (4، 2)، (4، 3)، (4، 4)، (4، 5)، (4، 6)، (5، 4)، (5، 5)، (5، 3)، (5، 4)، (5، 5)، (5، 6)، (6، 1)، (6، 2)، (6، 3)، (6، 4)، (6، 5)، (6، 6)} والشكل التالي يوضح ذلك.



**فضاء (فراغ)العينة لتجربة إلقاء قطعة زهرتي نرد متوازنتين**

| مجموع القطر | | | | | | |
|---|---|---|---|---|---|---|
| | (1،1) | (1،2) | (1،3) | (1،4) | (1،5) | (1،6) |
| 2 | (2،1) | (2،2) | (2،3) | (2،4) | (2،5) | (2،6) |
| 3 | (3،1) | (3،2) | (3،3) | (3،4) | (3،5) | (3،6) |
| 4 | (4،1) | (4،2) | (4،3) | (4،4) | (4،5) | (4،6) |
| 5 | (5،1) | (5،2) | (5،3) | (5،4) | (5،5) | (5،6) |
| 6 | (6،1) | (6،2) | (6،3) | (6،4) | (6،5) | (6،6) |
| 7 | 8 | 9 | 10 | 11 | 12 | |

رموز رياضية:

| الحرف الكبير | الحرف الصغير | الاسم بالعربي | الاسم بالانجليزي | الحرف الكبير | الحرف الصغير | الاسم بالعربي | الاسم بالانجليزي |
|---|---|---|---|---|---|---|---|
| Α | α | ألفا | Alpha | Ν | ν | نيو | Nu |
| Β | β | بيتا | Beta | Ξ | ξ | كي | Xi |
| Γ | γ | جاما | Gamma | Ο | ο | أوميكرون | Omicron |
| Δ | δ | دلتا | Delta | Π | π | باي | Pi |
| Ε | ε | إبسلون | Epsilon | Ρ | ρ | رو | Rho |
| Ζ | ζ | زيتا | Zeta | Σ | σ | سيجما | Sigma |
| Η | η | إيتا | Eta | Τ | τ | تاو | Tau |
| Θ | θ | ثيتا | Theta | Υ | υ | يبسلون | Upsilon |
| Ι | ι | أيوتا | Iota | Φ | φ | فاي | Phi |
| Κ | κ | كابا | Kappa | Χ | χ | كاي | Chi |
| Λ | λ | لمدا | Lambda | Ψ | ψ | إبساي | Psi |
| Μ | μ | ميو | Mu | Ω | ω | أوميغا | Omega |

### (3–4.1) تصنيف الأحداث:

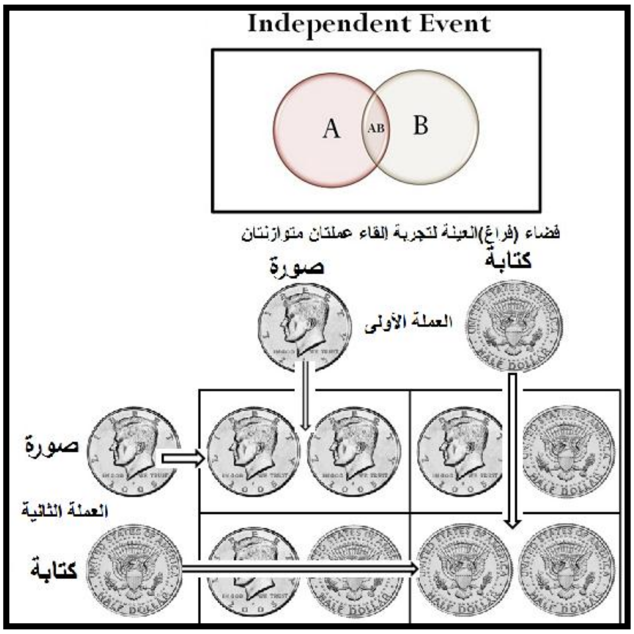
#### (3–1.4.1) الأحداث المستقلة: Independent events

يقال على حدثين أنهما مستقلين إذا كان وقوع أحدهما لا يمنع ولا يؤثر في وقوع الآخر.

**مثال:**

عند إلقاء عملتين فنتائج العملة الأولى لا علاقة لها بنتائج العملة الثانية، بمعنى أن ظهور الصورة من العملة الأولى لا يمنع ولا يؤثر ولا يعجل بظهور الصورة من العملة الثانية.

أي حادثة ظهور الصورة في الرمية الأولى لا تؤثر (مستقلة) على ظهور صورة الكتابة في الرمية الثانية.



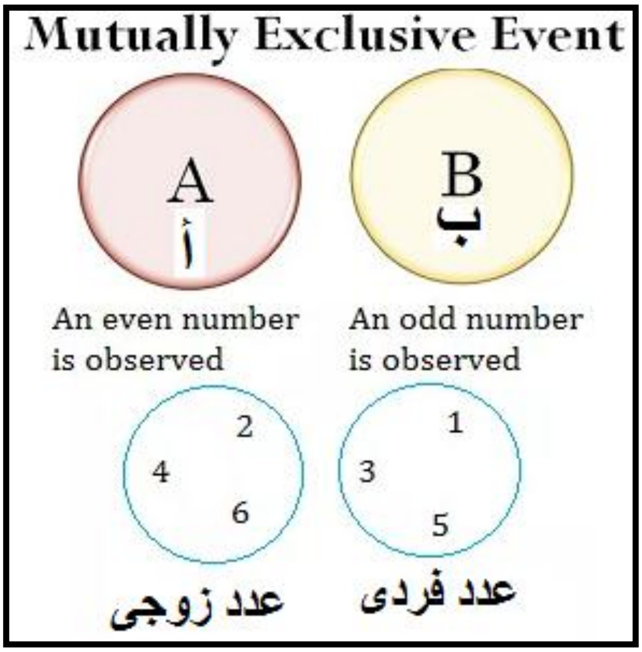
### (3–2.4.1) الأحداث المتنافية (المانعة): Mutually exclusive events

يقال على حدثين أنهما متنافيين (مانعين) إذا كان وقوع أحدهما يمنع وقوع الآخر، أي أن وقوعهما في نفس الوقت مستحيل.

أي لا يمكن أن يحدثا سوية في نفس الوقت.

مثلاً عند إلقاء عملة متوازنة فظهور الصورة والكتابة معاً وفى نفس الوقت مستحيلا، ولذلك يقال على الحدث الذي يمثل ظهور الصورة والآخر الذي يمثل ظهور الكتابة حدثين مانعين.

**الأحداث المتنافية (المانعة):**




### (3–3.4.1) الحدث المؤكد:

هو الحدث الذي سيقع بالتأكيد.

**مثال:** لو كان لدينا كيس به كرات سوداء فقط، وسحبنا كرة من ذلك الكيس فإنها سوف تكون من المؤكد سوداء.

### (3–4.4.1) الحدث المستحيل:

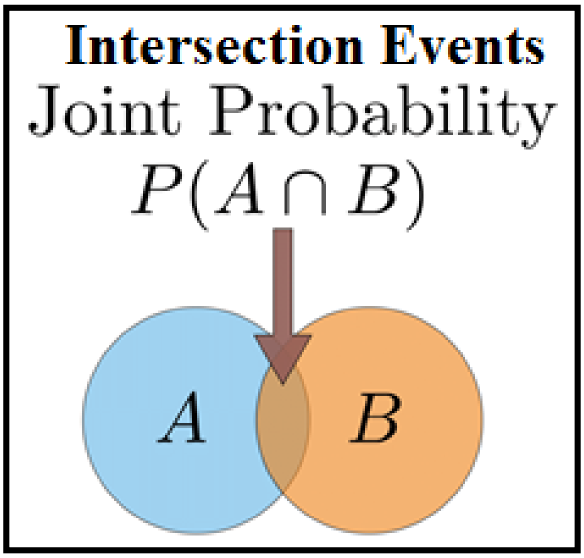
هو الحدث الذي لا يقع مطلقا، ونرمز له بالرمز $\{\varphi\}$ وينطق "فاي" أي "الفئة الخالية".

هي فئة لا تحتوي على أي عنصر ( ) = φ (باعتبارها فئة جزئية من أي فئة).

**مثال:** لو كان لدينا كيس به كرات سوداء فقط، وسحبنا كرة من ذلك الكيس فإن فرصة ظهور كرة بيضاء مستحيلة.

### (3–2) العمليات على الأحداث:

### (3–1.2) تقاطع الأحداث (أ ∩ ب): Intersection Events



إذا كان أ، ب حدثين من ف فإن تقاطع هذين الحدثين، والذي يرمز له بالرمز

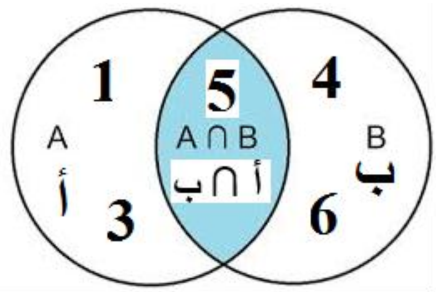
(أ ∩ ب)، وهو المجموعة التي عناصرها تنتمي إلى أ و ب معاً.

وعلى ذلك فإن أ ∩ ب تعنى حدث وقوع أ أو ب معاً، أي وقوع الحدثين معاً.

**The Intersection of Events.** The intersection of two events *A* and *B*, denoted by the symbol $A \cap B$ is the event containing all elements that are common to *A* and *B*.

**مثال:**

إذا كان الحدث أ = {1، 3، 5}، والحدث ب = {4، 5، 6} فإن أ ∩ **ب** = {5}




**(2.2−3) اتحاد الأحداث** (أ ∪ **ب**): **Union Events**

إذا كان أ، ب حدثين من ف فإن اتحاد هذين الحدثين، والذي يرمز له بالرمز

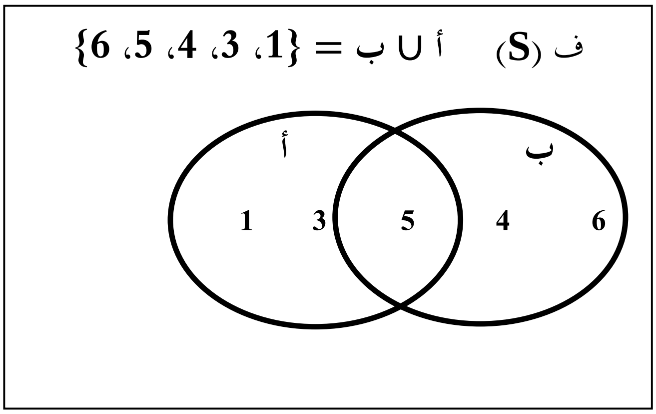
(أ ∪ ب)، وهو المجموعة التي عناصرها تنتمي إلى أ أو ب، أي وقوع أحدهما على الأقل.

وعلى ذلك فإن (أ ∪ ب) تعني حدث وقوع أ أو ب أو كلاهما أي وقوع أحدهما على الأقل.

**The Union of Events.** The union of two events *A* and *B*, denoted by the symbol $A \cup B$ is the event containing all elements that belong to *A* or *B* or both.

**مثال:**

أ = {1، 3، 5}، ب = {4، 5، 6} فإن:




**(3–3.2) الفرق بين حدثين":**

(أ – ب): إذا كان أ، ب حدثين من ف:

فإن:

(أ – ب) هي المجموعة التي عناصرها تنتمي إلى الحدث أ ولا تنتمي إلى الحدث ب.

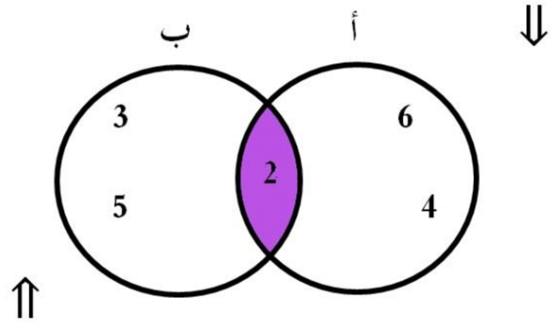
وعلى ذلك فإن الفرق بين الحدثين (أ – ب) يعني حدث وقوع الحدث (أ) فقط وعدم وقوع الحدث (ب).

بينما (ب – أ) هي المجموعة التي عناصرها تنتمي إلى الحدث ب ولا تنتمي إلى الحدث أ.

وعلى ذلك فإن الفرق بين الحدثين (ب – أ) يعنى حدث وقوع الحدث (ب) فقط وعدم وقوع أ.

∴ أ – ب = {4، 6}

هذه المنطقة تمثل مجموعة الفرق أي العناصر التي تنتمي إلى أ ولا تنتمي إلى **ب**.




هذه المنطقة تمثل مجموعة الفرق أي العناصر التي تنتمي إلى **ب** ولا تنتمي إلى أ.

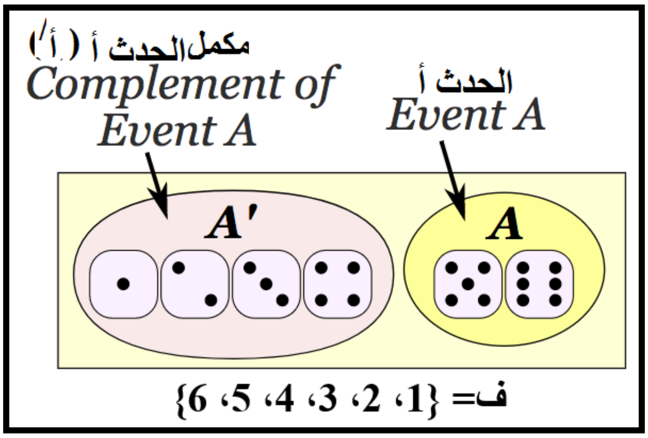
∴ ب – أ = {3، 5}

### (4.2–3) الحدث المكمل: Complement event

إذا كان (أ) حدثا من (ف) فإن مكملة المجموعة أ بالنسبة إلى ف وهي المجموعة (أ) التي عناصرها تنتمي إلى ف ولا تنتمي إلى (أ) تسمى بالحدث المكمل للحدث أ وهو الحدث الذي يقع إذا لم يقع.

*A'*, the complement of event A. It reads as "A ***complement***" or "***not*** A" (The area inside S but not covered by A)

**The Complement of an Event.** The complement of an event $A$ with respect to $S$ is the subset of all elements of $S$ that are not in $A$. The complement of $A$ is denoted by the symbol $A'$ or $A^c$.




**مثال:**

في تجربة إلقاء زهرة (حجر) نرد متوازنة، فإن: ف = {1، 2، 3، 4، 5، 6}.

بفرض أن الحدث أ = {1، 3، 5} فيكون الحدث المكمل أ` = {2، 4، 6} كما في الشكل التالي:

**(3−3) الاحتمال والحدث: Avengement et probability**

كثيراً ما يخلط الطلبة بين هذين المفهومين لارتباطهما ببعض. فالحدث العشوائي هو واقعة أو نتيجة ما، أما الاحتمال فهو عدد ينحصر بين الصفر أو يساوى الصفر والواحد أو يساوي الصفر.

$$0 \leq P(A) \leq 1$$

$$0 \leq \mathbf{P}\left[\text{أ}\right] \leq 1$$

ويعبر عن حظوظ وقوع الحدث (ليس شرطا أن يكون زمن وقوع الحدث هو المستقبل، فقد يكون الماضي أو الحاضر).

فعندما نرغب في التعبير بشكل دقيق على مدى إمكانية وقوع حدث معين فإننا عادة نستعمل عبارات مثل: 100% للحدث المؤكد أو 50% للحدث المحتمل و1% مثلا للحدث المستبعد، إذن نحن نستخدم الكسور في سلم تصاعدي من 0 إلى1، بحيث يرمز 0 للاستحالة و1 للتأكد.

فاحتمال حدوث الحدث (A) يساوى عدد مرات حدوث (التكرارات) الحدث (أ) مقسوما على عدد المشاهدات (فضاء العينة).

الحدث (A) هو حدث من مجموعة الأحداث (E) فالاحتمال يكون:

$$P(A) = \frac{\text{total outcomes in A}}{\text{total outcomes in S}}$$

$$P(A) = \frac{n_A}{N} = \frac{\text{عدد مرات ظهور الحدث (A)}}{\text{مجموع الأحداث } (E)}$$

$$\text{أحتمال} \left(\text{أ}\right) = \frac{\text{عدد مرات ظهور الحدث}\left(\text{أ}\right)}{\text{مجموع الأحداث}\left(\text{ن}\right)}$$

$$\text{ل} \left(\text{أ}\right) = \frac{\left[\text{عدد مرات حدوث} \left(\text{التكرارات}\right) \text{الحدث} \left(\text{أ}\right)\right]}{\left[\text{عدد المشاهدات}\left(\text{فضاء العينة}\right)\right]}$$

أي احتمال (أ) = عدد مرات ظهور الحدث (أ) ÷ مجموع الأحداث (ن).

ل(أ) = عدد مرات ظهور الحدث (أ) ÷ مجموع الأحداث (ن).

**(1.3-3) احتمال حدث منفرد:**

إذا كان الحدث $\boldsymbol{A}$ يمكن أن يحدث بطرق عددها $n_A$ من بين نواتج متساوية الفرصة في الوقوع عددها N، فإن احتمال وقوع الحدث $A$ يعرف كما يلي:

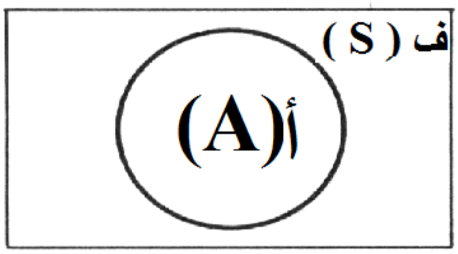
$$P(A) = \frac{n_A}{N} = \frac{\text{عدد مرات ظهور الحدث } (\text{A})}{\text{مجموع الأحدث } (E)}$$

حيث $P(A)$ = احتمال وقوع الحدث $\boldsymbol{A}$.

$A$ = عدد الطرق التي يمكن أن يقع بها $\boldsymbol{A}$.

$N$ = العدد الكلي للنواتج المتساوية الفرص في الوقوع.

ويمكن تصور الاحتمال باستخدام شكل فن. ففي الشكل التالي تمثل الدائرة الحدث $A$ بينما تمثل المساحة الكلية المستطيل كل النواتج الممكنة.




**مع ملاحظة أن:**

احتمال وقوع أي حدث هو عدد حقيقي غير سالب لا يقل عن صفر ولا يزيد عن 1.

احتمال وقوع فضاء النواتج (الفضاء العيني) أو الحادث المؤكد = 1.

احتمال الحادث المستحيل = صفر.

أي تتراوح $\mathbf{P}\left[\text{أ}\right]$ $P(A)$ بين $\mathbf{0\,,1}$ .

$$0 \leq P(A) \leq 1$$

$$0 \leq \mathbf{P}\left[\text{أ}\right] \leq 1$$

فإذا كانت $0 = \mathbf{P}\left[\text{أ}\right] = P(A)$، فإن الحدث $A$ لا يمكن أن يقع، وإذا كانت $1 = \mathbf{P}\left[\text{أ}\right] = P(A)$ ، فإن الحث $A$ مؤكد الوقوع.

وإذا استخدمنا P(À) لتمثل احتمال عدم وقوع الحدث $\boldsymbol{A}$ فإن:

$$\mathrm{P(A)} + \mathrm{P}\left(\grave{\mathrm{A}}\right) = 1$$

$$\mathrm{P}\left(\text{أ}\right) + \mathrm{P}\left(\grave{\text{أ}}\right) = 1$$

**مثال**: عند إلقاء قطعة نقود متوازنة فإن الصورة والكتابة يمثلان ناتجين لهما نفس فرصة الوقوع، أي أن:

$$P\left(\text{ص}\right) = \frac{n_A}{N} = \frac{\text{عدد مرات ظهور الحدث}\left(\text{ص}\right)}{\text{مجموع الأحدث}\left(\text{ص ك،}\right)} = \frac{1}{2}$$

$$P\left(\text{ك}\right) = \frac{n_A}{N} = \frac{\text{عدد مرات ظهور الحدث}\left(\text{ك}\right)}{\text{مجموع الأحدث}\left(\text{ص ك،}\right)} = \frac{1}{2}$$

$$\therefore P\left(\text{ص}\right) + P\left(\text{ك}\right) = \frac{1}{2} + \frac{1}{2} = 1$$

**مثال (2)**: عند إلقاء حجر نرد غير متحيز فإن هناك ستة نواتج متساوية الفرصة في الحدوث 1,2,3,4,5,6 ومن ثم فإن:

$$P\ (1) = P\ (2) = P\ (3) = P\ (4) = P\ (5) = P\ (6) = \frac{1}{6}$$

وإذا استخدمنا $\mathrm{P(\grave{A})}$ لتمثل احتمال عدم وقوع الحدث $\boldsymbol{A}$ فإن:

$$\mathrm{P(A) + P(\grave{A}) = 1}$$

$$\mathrm{P}\left(\text{أ}\right) + \mathrm{P}\left(\grave{\text{أ}}\right) = 1$$

ويكون احتمال عدم ظهور 1 هو:

حيث أن:

$$\mathrm{P}\left(\text{أ}\right) + \mathrm{P}\left(\grave{\text{أ}}\right) = 1$$

$$\mathrm{P}\left(\grave{\text{أ}}\right) = 1 - \mathrm{P}\left(\text{أ}\right) = 1 - (1/6) = 5/6$$

**مثال**:

في تجربة إلقاء حجر نرد منتظم. أوجد احتمال ظهور عدد فردي؟

**الحل**: عناصر الفضاء العيني ف (S) $=$ {1، 2، 3، 4، 5، 6}، أي أن ن $=$ 6 ف (S).

عناصر الحدث (أ) $=$ {1، 3، 5}، أي أن ن (أ) $=$ 3

وبالتالي ل (أ) $= \frac{3}{6} = \frac{1}{2}$

احتمال عدم وقوع حدث (ويسمى أيضا مكمل (متممة) الحدث أ) يرمز له بالرمز

ل$\left(\grave{أ}\right)$ حيث:

عناصر الحدث:

$$\left(\grave{أ}\right) = \{2، 4، 6\}$$

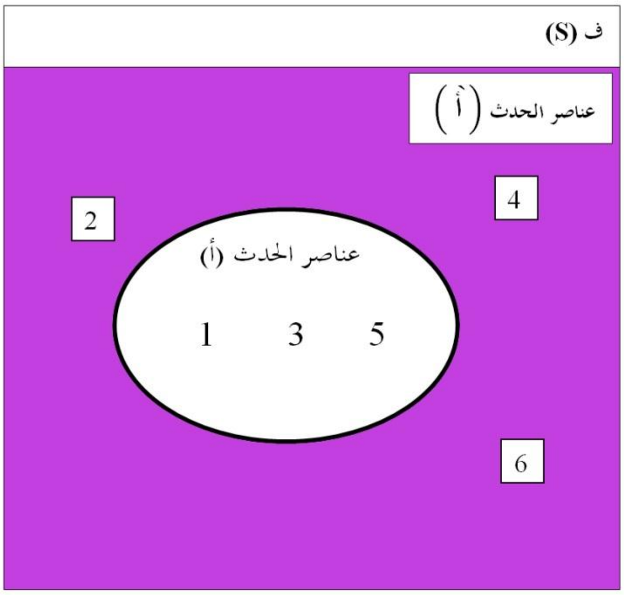
احتمال حدوث الحدث:

$$ل\left(\grave{أ}\right) = 1 - ل\left(أ\right)$$

**ويمكن توضيحه من خلال أشكال فن (المنطقة الملونة تمثل متممة الحدث).**

**مثال:**

في تجربة إلقاء حجر نرد منتظم. أوجد احتمال عدم ظهور عدد فردي؟



**الحل:**

حادث عدم ظهور عدد فردي هو متمم لحادث ظهور عدد زوجي وبمعنى آخر ظهور عدد زوجي.

أي أنّ : ل (أ̀) = 1 − ل(أ) = $1 - \frac{3}{6} = \frac{1}{2}$

**(3–2.3) الأحداث غير المنافية:**

**ويعني وقوع الحدث أ أو ب أو وقوع أحدهما على الأقل ويرمز له بالرمز ل (أ أو ب) أو ل (أ ∪ ب) .**

**قاعدة الجمع للأحداث غير المتنافية:**

يعتبر الحدثان **A , B** غير متنافيين إذا كان وقوع **A** لا يحجب وقوع **B** والعكس بالعكس فيكون:

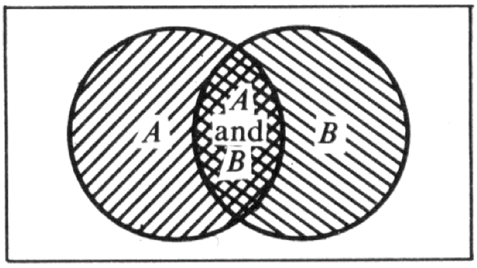
$$P(A \text{ or } B) = P(A) + P(B) - P(A, B)$$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

ل (أ أو ب) = ل (أ) + ل (ب) − ل (أ أو ب)

ل (أ ∪ ب) = ل (أ) + ل (ب) − ل (أ ∩ ب)

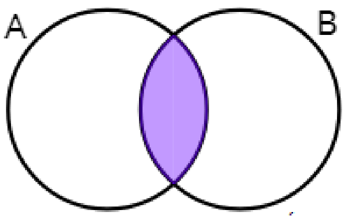
وتطرح قيمة **P (A , B)** حتى نتجنب حسابها مرتين. ويمكن إدراك ذلك من شكل فن الموضح بالشكل أدناه.




ويمكن حسابه بطريقة مباشرة عن طريق معرفة عدد العناصر الناتجة من اتحاد الحادثين وقسمتها على العدد الكلي للفضاء العيني. (الرسم يمثل توضيح باستخدام أشكال فن).

**الأحداث غير المتنافية: Non-Mutually Exclusive Events**

احتمال حصول أحد الحدثين أو كلاهما (احتمال الاتحاد)




$$P(A \text{ or } B) = P(A) + P(B) - P(A, B)$$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

ل (أ ∪ ب) = ل (أ) + ل (ب) − ل (أ ∩ ب)

**مثال:**

في تجربة إلقاء حجر نرد منتظم. أوجد احتمال ظهور عدد فردي أو أولي إذا كان الحدث أ = {1، 3، 5} ، والحدث ب = {2، 3 ،5}

**الحل:**

عناصر الفضاء العيني ف (S) = {1، 2، 3، 4، 5، 6}، أي أن: ن = 6 ف (S)

عناصر الحدث أ = {1، 3، 5}، بينما ب = {2، 3، 5}.

∴ أ $\cup$ ب = {1، 2، 3، 5}

حيث 4 عدد عناصر الاتحاد و6 عدد عناصر الفضاء العيني.

ل (أ أو ب) = ل (أ $\cup$ ب) = $\frac{4}{6} = \frac{2}{3}$

**حل آخر باستخدام القانون:**

ل (أ أو ب) = ل (أ $\cup$ ب) = (6/3) + (6/3) − ل(6/2)

ل(أ أو ب) = 6/[2 − 3 + 3] = (6/4) = 3/2

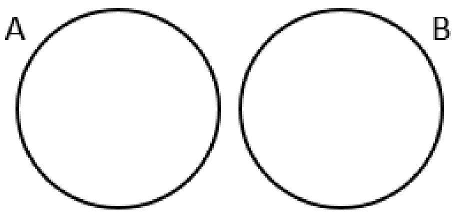
**(3–3.3) الحوادث المتنافية الحوادث المتنافية: أو (المنفصلة) أو (المتباعدة)**

إذا كان ف (S) فضاء عيني لتجربة عشوائية. وكان أ و ب حوادث في الفضاء العيني فإنه يقال إن الحادثين أ، ب متنافيان إذا كان أ $\cap$ ب = $\emptyset$

ولمزيد من التوضيح يمكن الاستعانة بالرسم:

**الأحداث المتنافية: Mutually Exclusive Events**

الأحداث التي لا يوجد بينها تقاطع.




$$P(A \text{ or } B) = P(A) + P(B)$$

$$P(A \cup B) = P(A) + P(B)$$

$$\text{ل}\left(\text{أ} \cup \text{ب}\right) = \text{ل}\left(\text{أ}\right) + \text{ل}\left(\text{ب}\right)$$

الحدثين أ، ب متنافيان إذا $(A \cap B) = \emptyset$

نلاحظ أن الحادثين لا يوجد بينهما منطقة مشتركة وبالتالي فالحادثين متنافيين.

ومن خلال التعريف إذن نستنتج شيئين مهمين لأي حادثين متنافيين:

ل (أ ∩ ب) = صفر

أي قاعدة الجمع للأحداث المتنافية: يعتبر الحدثان ***A*** **, B** متنافيان بالتبادل إذا كان وقوع ***A*** يحجب وقوع **B** والعكس بالعكس.

**Mutually Exclusive Events.** Two events *A* and *B* are mutually exclusive or disjoint if $A \cap B = \emptyset$, that is if *A* and *B* have no common elements in common.

عندئذ:

$$P(A \text{ أو } B) = P(A) + P(B)$$

$$P(A \cup B) = P(A) + P(B)$$

ل(أ ∪ ب) = ل (أ) + ل (ب)

**ويمكن التعميم لأكثر من حادثين:**

أي أن احتمال تقاطع عدة حوادث متنافية يساوي صفر.

قاعدة الضرب للأحداث المستقلة: يعتبر الحدثان A، B مستقلين إذا كان وقوع A غير مرتبط بأي طريقة بوقوع **B**. عندئذ الاحتمال المشترك للحدثين A، B هو:

$$P(A\ and\ B) = P(A)\,.P(B)$$

قاعدة الضرب للأحداث غير المستقلة: يعتبر الحدثان غير مستقلين إذا كان وقوع أحدهما مرتبطاً بطريقة ما بوقوع الآخر عندئذ:

$$P(A \text{ and } B) = P(A) . P(B/A)$$

وتقرأ كالآتي: "احتمال وقوع كل من الحدثين A، B يساوي احتمال وقوع الحدث ***A*** مضروباً في احتمال وقوع الحدث **B** إذا علم أن الحدث ***A*** قد وقع فعلا".

**P (B/*A*)=** الاحتمال الشرطي للحدث **B** علماً بأن الحدث قد وقع فعلا.

$$\mathbf{P (A \cap B) = P (B \cap A)}$$

**مثال:**

في رمية واحدة لحجر نرد، يمكن الحصول على واحد من ستة نواتج ممكنة: 2,3,4,5,61, وهذه الأحداث احتمال الحصول على **2** أو **3** في رمية واحدة:

ملاحظة: الرمز ∪ (ويقرأ "اتحاد") يحل محل "أو"، الرمز ∩ (ويقرأ "تقاطع") ويحل محل "و".

**مثال:**

تعتبر نواتج رميتين متتاليتين لقطعة نقود متوازنة أحداثاً **مستقلة**. فناتج الرمية الأولى لا يؤثر على أي نحو على ناتج الرمية الثانية فيكون:

$$\text{ل(ص و ص)} = \text{ل(ص} \cap \text{ص)} = 1/2 * 1/2 = 1/4 = 0.25$$

$$\text{ل(ص و ص)} = \text{ل(ص} \cap \text{ص)} = 1/2 * 1/2 = 1/4 = 0.25$$

$$P(H \text{ and } H) = P(H \cap H) = P(H).P(H) = \frac{1}{2}.\frac{1}{2} = \frac{1}{4}, \text{or } 0.25$$

وبالمثل:

ل(ص و ص و ص) = ل(ص ∩ ص ∩ ص) = 1/2 * 1/2 = 1/4 = 0.125

$$P(\text{H and H and H}) = P(H \cap H \cap H) = P(H).P(H).P(H) = \frac{1}{2}.\frac{1}{2}.\frac{1}{2} = \frac{1}{8}, \text{or } 0.125$$

### (3–4) التعبير الرياضي عن قواعد حساب الاحتمالات

**القاعدة رقم 1:**

**الحدث المكمل أو التعبير الرياضي عن:**

نعبر عن الحدث المعاكس (المكمل) ل $A$ ب $\bar{A}$ أو $A'$ واحتماله هو احتمال عدم تحقق الحدث $A$، ونكتب:

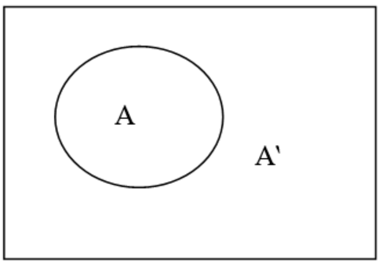
$$P(\bar{A}) + P(A) = 1 \Leftrightarrow P(\bar{A}) = 1 - P(A)$$

ل(ص) = ل(صَ) = 1

أو ل(صَ) = 1 − ل(ص) = 1

أو ل(ص) = 1 − ل(صَ)

ف (S)




**مثال**: نرمي قطعة نقدية ونرمز للصور (ص) وللكتابة (ك) نلاحظ أن:

ل(ص) + ل(صَ) = 1 أو − ل(صَ) = 1 − ل(ص) أو ل(صَ) = 1 − ل(صَ)

**مثال:**

عند رمي حجر نرد فإن احتمال الحصول على العدد 5 هو: P(5) = 1/6، فما هو الحدث المكمل في هذه الحالة وما احتماله؟

الحدث المكمل (المعاكس) هو الحصول على عدد غير 5، واحتماله هو:

$$6/1 = \text{ل}(5) \;\therefore\; 6 = \text{ف}(S) \;\because$$

$$6/5 = (6/1) - (6/6) = (6/1) - 1 = \text{ل}(5) - 1 = \text{ل}(\bar{5}) \;\therefore$$

**مثال:**

نرمي حجر نرد، ما هو احتمال الحصول على عدد زوجي، ما هو الحدث المعاكس(المكمل) وما هو احتماله؟

$$6/3 = \text{ل}(2،4،6) = \text{ل}\left(\text{عدد زوجي}\right) \;\therefore\; 6 = \text{ف}(S) \;\because$$

$$\text{ل}\left(\text{عدد فردي}\right) \therefore 6/3 = 6/3 - (6/6) = 6/3 - 1 = \text{ل}\left(\text{عدد زوجي}\right) - 1 =$$

$$P\,(\text{even numbers}) = P\,(2 \text{ or } 4 \text{ or } 6) = 3/6$$

الحدث المعاكس (المكمل) هو الحصول على **عدد غير زوجي**، واحتماله:

$$P\,(\text{odol numbers}) = 1 - P(\text{even}) = 1 - (3/6) = 3/6$$

**قاعدة رقم 2:**

**احتمال وقوع الحدث "A" و "B" أيا كانت.**

$$P\,(A \cap B) = P\,(A) * P(B/A)$$

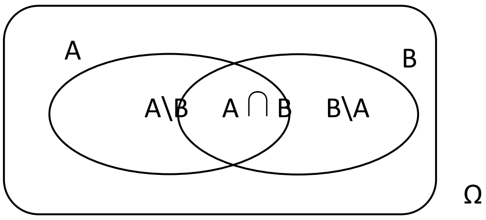
$$P(A \cap B \cap C) = P(A) * P(B/A) * P(C/(A \cap B))$$

A, B, C أحداث ما (مستقلة أو لا، متنافية أو لا)، P(B/A) يسمى الاحتمال الشرطي لـB علما أن A محقق.

ومن المعادلة الأولى نحصل على:

لما $P(A) > 0$ $\qquad P(B/A) = P(A \cap B) \,/\, P(A)$

حيث A تصبح فضاء المعاينة بما أن A محقق.




رسم 1: الحدث B/A غير الحدث

**مثال: 1**. أحسب عند إلقاء حجر نرد احتمال الحصول على قيم أقل من 4 (حدثB).

**2**. أحسب احتمال الحصول على نتيجة أقل من 4 إذا علمت أن الوجه المحصل لمكعب النرد عدد فردي (حدثA).

**3**. أحسب احتمال الحصول على قيمة أكبر أو يساوي 4 إذا علمت أن النتيجة عدد فردي.

$$P(B) = P(1 \text{ or } 2 \text{ or } 3) = P(1) + P(2) + P(3) = 3/6$$

$$P(B/A) = P(B \cap A)/P(A)$$

$$P(B \cap A) = P(\text{ood} \leq 4) = P(1 \text{ or } 3) = P(1) + P(3) = 1/6 + 1/6 = 2/6$$

$$P(B/A) = P(B \cap A)/P(A) = 2/6 \;/\; 3/6 = 2/3$$

**احتمال وقوع الحدث "A" و "B" لما "A" و "B" مستقلان (قاعدة رقم 3)**.

$$P(A \cap B) = P(A) * P(B) \qquad (P(B/A) = P(B))$$

وهو تعريف استقلال حدثين، أي أن وقوع B لا يتأثر بوقوع A أو عدم وقوعه نقول إن A وB مستقلان:

$$P(A \cap B \cap C) = P(A) * P(B) * P(C) \qquad P(C/(A \cap B)) = P(C)$$

**مثال**: نرمي حجر نرد وقطعة نقدية معا. ما هو احتمال الحصول على الصورة والعدد 6؟ (نتيجة مكعب النرد مستقلة عن نتيجة القطعة النقدية).

$$P(A \cap B) = P(A).P(B) = 0.5 * 1/6 = 1/12$$

**مثال 2**. نلقي قطعة نقدية مرتين. أحسب احتمال الحصول على صورة في الرمية الأولى وفي الرمية الثانية.

$$P(FF) = P(A \cap B) = P(A)\,P(B) = 0.5 * 0.5 = 0.25$$

**مثال 3**: صندوق به 5 كريات 2 حمراء و3 بيضاء. نسحب كرية نسجل لونها ثم نعيدها للصندوق ونكرر العملية 3 مرات.

- أحسب احتمال الحصول على 2 كريات حمراء، 3 كريات حمراء (أحداث مستقلة).
- كيف يكون الاحتمال في حالة كون السحب بدون إرجاع الكرية (أحداث غير مستقلة)؟

$$P(RR) = P(R_1 \cap R_2) = P(R_1)\,P(R_2) = 2/5 * 2/5 = 8/25$$

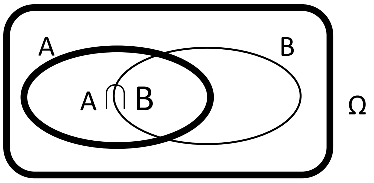
$$P(RRR) = P(R_1 \cap R_2 \cap R_3) = P(R_1)\,P(R_2)\,P(R_3) = 2/5 * 2/5 * 2/5 = 8/125$$

$$P(RR) = P(R_1 \cap R_2) = P(R_1)\,P(R_2/R_1) = 2/5 * 1/4 = 2/20$$

$$P(RRR) = P(R_1 \cap R_2 \cap R_3) = P(R_1)\,P(R_2/R_1)\,P(R_3/(R_1 \cap R_2)) = 2/5 * 1/4 * 0 = 0$$

احتمال وقوع حدث "A" أو "B" (القاعدة رقم 4).

**$P(AUB) = P(A) + P(B) - P(A\cap B)$**

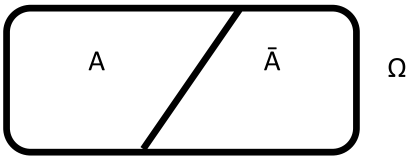



**احتمال وقوع حدث "أ" أو "ب" لما "أ" و "ب" متنافيان (القاعدة رقم 5).**

لتكن الأحداث المتنافية A, B

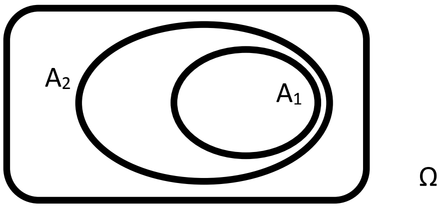
$$\mathbf{P(A\cup B) = P(A) + P(B); \quad (P(A\cap B) = 0)}$$

$$P(AUBUC) = P(A) + P(B) + P(C); \quad (P(A\cap B\cap C) = 0)$$

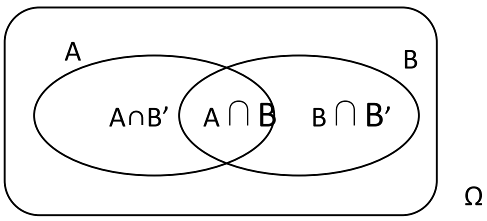



**قواعد إضافية مهمة:**

– من أجل $A_1 \subset A_2$ فإن: $P(A_1) \le P(A_2)$ و $P(A_2 - A_1) = P(A_2) - P(A_1)$

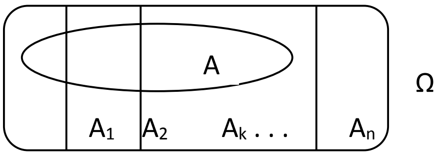



– من أجل A وB أحداث أيا كانت: $P(A \cap B) + P(A\cap B') = P(A)$



– إذا كان A هو نتيجة أحد أو بعض الأحداث المتنافية: $A_1, A_2, A_3, A_4, \ldots\ldots A_n$

$$P(A) = P(A\cap A_1) + P(A\cap A_2) + P(A\cap A_3) + \ldots\ldots + P(A\cap A_n)$$

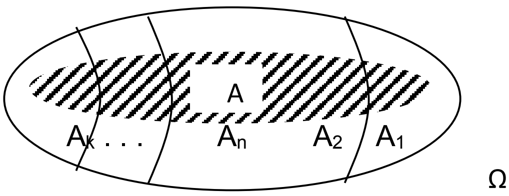



**نظرية الاحتمال السببي أو نظرية بايز**: Théorème ou règle de BAYES

لتكن $A_1$، $A_2$، $A_3$، . . . $A_k$، . . .، $A_n$ أحداث متنافية فيما بينها حيث اتحادها يشكل المجموعة الكلية (الأساسية) Ω، وA حدث ما يتحقق عن طريق واحد أو أكثر من الأحداث $A_k$، إذا علمنا أن A تحقق، نحسب احتمال تحققه عن طريق الحدث $A_k$ كما يلي:

$$p = \left(\frac{A_k}{A}\right) = \frac{P(A_K)P\left(\frac{A}{A_k}\right)}{\sum_{k=1}^{n} P(A_k)P(\frac{A}{A_k})} = \frac{P(A\cap A_k)}{P(A)}$$

تسمى هذه النظرية نظرية الاحتمال السببي لأنها تمكن من حساب احتمال أن يكون حدث ما ($A_k$) هو المسبب لوقوع حدث آخر (A).




رسم 2 : يوضح نظرية بايز.

**مثال**: وظفت أمينة مكتب ($A_1$) بمكتب للمحاسبة حيث تولت طبع 20% من الفواتير. يشغل المكتب عاملتين أخريين إحداهما ($A_2$) تطبع 30% من الفواتير والأخرى ($A_3$) 50%. ترتكب الموظفة الجديدة أخطاء في 5% من الفواتير، بينما نسبة الخطأ لدى الثانية ($A_2$) 2% ولدى الثالثة ($A_3$) 1%.

أخذت فاتورة بشكل عشوائي فتبين أن بها أخطاء. استبعدت الأولى أن تكون هي من أنجزت الفاتورة بحجة أنها لا تنجز إلا 20% من الفواتير، وردت عليها العاملات الأخريات بأن نسبة الأخطاء لديها هي الأكبر (5%).

**1.** أحسب احتمال أن تكون الموظفة الجديدة ($A_1$) هي التي حررت الفاتورة وقارن مع احتمال أن يكون مصدر الخطأ هو $A_2$ أو $A_3$.

**2.** أحسب مجموع الاحتمالات الثلاث.

**3.** أحسب احتمال أن تكون فاتورة مختارة عشوائيا من مجموع المراسلات، أن تكون بها أخطاء.

$$P(A_1/A)=\frac{P(A_1)P(A/A_1)}{\sum_{k=1}^{3}P(A_k)P(A/A_k)}=\frac{0.2*0.05}{(0.2*0.05)+(0.3*0.02)+(0.5*0.01)}=0.238$$

$$P(A_2/A)=\frac{P(A_2)P(A/A_2)}{\sum_{k=1}^{3}P(A_k)P(A/A_k)}=\frac{0.3*0.02}{(0.2*0.05)+(0.3*0.02)+(0.5*0.01)}=0.2857$$

$$P(A_3/A)=\frac{P(A_3)P(A/A_3)}{\sum_{k=1}^{3}P(A_k)P(A/A_k)}=\frac{0.5*0.01}{(0.2*0.05)+(0.3*0.02)+(0.5*0.01)}=0.476$$

يظهر من الحساب أن الاحتمال الأكبر هو أن تكون $A_3$ هي التي حررت الفاتورة.

**4.** مجموع الاحتمالات $P(A_1/A) + P(A_2/A) + P(A_3/A) = 1$ لأنها تمثل احتمالات الأحداث المتنافية الثلاث.

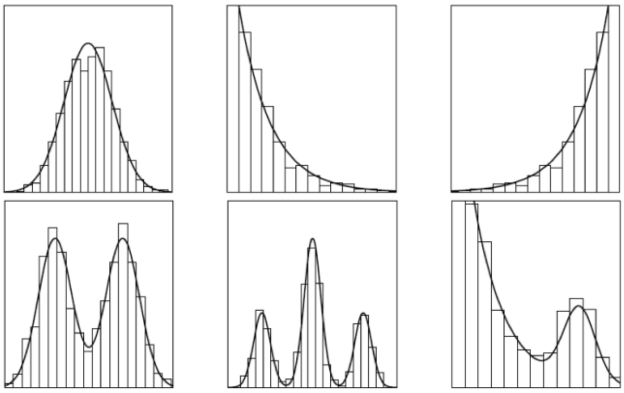
**5.** احتمال وجود خطأ في مراسلة ما:

$$P(A) = \sum P(A_k)P(A/A_k) = (0.2*0.005)+(0.3*0.02)+(0.5*0.01) = 0.012$$

# الباب الرابع

## التوزيعات التكرارية

## Frequency Distributions



- مقدمة.

(4-1) التصنيف والعرض للبيانات الإحصائية.

(4– 2) طرق كتابة الفئات.

(4-3) الجدول التكراري.

(4-4) التوزيعات التكرارية المتجمعة.

(4-5) التوزيعات التكرارية المزدوجة.

**مقدمة:**

البيانات التي تم جمعها والتي تخص مجتمع البحث غالباً ما تكون بيانات أولية غير مصنفة وغير منتظمة بحيث يتعذر على الباحث تكوين فكرة عن المشكلة قيد البحث وبالتالي لا يمكن إجراء التحليل الإحصائي للوصول إلى نتائج البحث.

**(4–1) التصنيف والعرض للبيانات الإحصائية:**

العرض البياني للبيانات، هو إحدى الطرق التي يمكن استخدامها في وصف البيانات، من حيث شكل التوزيع ومدى تمركز البيانات، وتتم عملية التبويب (عرض البيانات) في شكل جدول تكراري على مرحلتين متتابعتين.

- **تكوين جدول تفريغ البيانات:**

هو جدول يحتوي على علامات إحصائية، كل علامة تعبر عن تكرار للمجموعة التي ينتمي إليها مثلاً: نوع التمر الذي تنتجه المزرعة، وكل خمس علامات تكون حزمة إحصائية، كما هو مبين بالجدول التالي:

**1. مرحلة تكوين جدول تفريغ البيانات:**

قراءة وتسجيل كل بيان أمام الصفة أو الفئة أو المعيار المتفق مع فقرتها، وذلك بتمثيله بشرطة مائلة من أعلى اليمين إلى أدنى اليسار كما يلي (/) حتى تبلغ الشرطة الخامسة تكون كخط أو شرطة تقطع الأربعة السابقة في صورة عكسية.

- كما يلي ( ////)، والخمس قراءات في الصورة السابقة يطلق عليها حزمة.

**2. مرحلة تصنيف أو تبويب البيانات:**

يتم ترجمة الحزم أو المفردات الموجودة بعمود التفريغ في جدول التفريغ أمام كل صفة أو وجه أو معيار بنفس الجدول السابق إلى ما "يتم تكراره" بعدد المفردات أو مفردات الحزم.

**3. إنشاء جدول تكراري:**

تقوم فكرة تبويب البيانات الكمية على أساس بسيط مؤداه تقسيم مدى القيم الأصلية للظاهرة إلى مجموعات جزيئية وذلك بضم بعض القيم المتقاربة إلى بعضها البعض في مدى بسيط نسبياً في تتابع يطلق عليه فئات، فالفئة هي مجموعة من البيانات متشابهة إلى حد كبير جداً في الصفات، ويفضل أن تكون هذه الفئات متساوية الطول ويتم ذلك عمليا وفقا للخطوات التالية:

**1.** تحديد مدى التغير في البيانات الأصلية وهو عبارة عن الفرق بين أكبر قيمة وأصغر قيمة في مفردات الظاهرة الكمية موضوع التبويب أي أن:

المدى = (أكبر قيمة – أصغر قيمة)

**Range = highest value − lowest value**

**2.** تحديد عدد الفئات وأطوالها أخذ جدلاً كبيراً بين المتخصصين، إلا أن الأمر لا يتطلب كل هذا العناء لأن وضع قواعد محددة في مثل هذه الأمور يضر أكثر مما يفيد، حيث يتوقف هذا الأمر على طبيعة الدراسة والقائمين عليها والظروف التي يعملون من خلالها، ومع ذلك نشير إلى بعض القواعد التي توصل إليها المهتمون بهذا الأمر فقد وضع ستيرجس قاعدة لتحديد عدد فئات التوزيع التكراري كما يلي:

عدد فئات التوزيع = أ + 3.3 لو ن

حيث ن عدد المفردات، ويفضل استخدام هذه القاعدة في الحالات التي تتراوح فيها (ن) بين 100 — 1000 إلا أن ذلك لا يمنع الاسترشاد بها إذا كانت (ن > 100).

**1) تحديد عرض الفئة:**

عرض الفئة = المدى ÷ عدد الفئات = (أكبر قيمة — أصغر قيمة) ÷ عدد الفئات

**Class width = (Maximum entry − Minimum entry) / Number of classes**

**2)** اختيار بداية الفئة الأولى أي الحد الأدنى لها مساوٍ لأقل قيمة موجودة بالبيانات أو أقل بقليل منها فمثلاً تكون من الأرقام الصفرية لتسهيل الحسابات بعد ذلك.

**3) كتابة الفئات:**

**(4— 2) طرق كتابة الفئات:**

- **الفئات المنفصلة (المتباعدة):**

يذكر الحدان الأدنى والأعلى لكل فئة كما بالجدول التالي:

| الفئات (Calls) | قراءاتها |
|---|---|
| 15 – 24 | تقرأ من 15 وحتى 24 |
| 25 – 34 | تقرأ من 25 وحتى 34 |
| 35 – 44 | تقرأ من 35 وحتى 44 |

ويعاب على هذه الطريقة أنها تترك فجوات بين كل فئتين، وعلى الرغم من أن هذه الطريقة تصلح لكتابة الفئات في حالة المتغيرات المنفصلة (المتقطعة) وذلك لأنه بطبيعة المتغير المنفصل (المتقطع) فإنه لابد وأن يكون من المفهوم أن كل فئة تنتهي بوحدة واحدة أقل من بداية الوحدة التالية وهكذا إلا أنها لا تصلح في حالة البيانات التي تحتوي على كسور.

فإذا أردنا التعبير عن توزيع عدد أفراد الأسرة فإنه ممكن أن تكتب الفئات على النحو التالي:

1–3 وتقرأ من فرد إلى ثلاثة أفراد.

4–6 وتقرأ من أربعة إلى ستة أفراد.

- **الفئات المتصلة (المتداخلة) مثل:**

نذكر كلا من الحد الأدنى والحد الأعلى للفئة كما بالجدول التالي:

| الفئات | قراءتها |
|---|---|
| 15 — 25 | تقرأ من 15 إلى 25 |
| 25 — 35 | تقرأ من 25 إلى 35 |
| 35 — 45 | تقرأ من 35 إلى 45 |

ويعاب عليها أن بها تداخلاً فمثلاً 25 أين توضع؟ هل في الفئة الأولى أم في الفئة الثانية وفي هذه الحالة لا نعرف إلى أي فئة ينتمي هذا الرقم، وإن كانت تستخدم كثيراً سواء للمتغيرات المنفصلة أو المتصلة، ويكون مفهوما أن الفئة الأولى تبدأ من 15 وحتى أقل من 25، والثانية من 25 وحتى أقل من 35 وهكذا.

- **الفئات المتلاحقة:**

يذكر الحد الأدنى فقط ويترك الحد الأعلى ليتحدد ضمنا من بداية الفئة التالية وهي الأكثر استخداماً كما بالجدول التالي:

| Class Boundaries حدود الفئة | | | you read تقرأ |
|---|---|---|---|
| تكتب على إحدى هذه الصور | | | |
| -15 | 25 > 15 | 15 < 25 | تقرأ من 15 وحتى أقل من 25 |
| -25 | 35 > 25 | 25 < 35 | تقرأ من 25 وحتى أقل من 35 |

| | | | |
|---|---|---|---|
| تقرأ من 35 وحتى 45 | 35 to 45 | 45 - 35 | 45 - 35 |

**4) بناء الجدول ووضع العلامات التي تمثل التكرار.**

**(4–3) الجدول التكراري:**

هو الجدول الذي يبين عدد تكرار المفردات في كل فئة.

**أنواع الجداول التكرارية:**

**1. الجدول التكراري المنتظم:**

هو الجدول الذي عرض فئاته متساوية (عرض الفئة ثابت وعرض الفئة هو الفرق بين الحد الأقصى والأدنى للفئة)، ويأخذ الصورة التالية:

| عرض الفئة<br>Class Width | التكرار<br>Frequency | الفئات<br>Class |
|---|---|---|
| 9 = 24 – 15 | 12 | 24 15– |
| 9 = 34 – 25 | 14 | 34 25– |
| 8 = 44 – 35 | 8 | 35 – 44 |

**2. الجدول التكراري غير المنتظم**

هو الجدول الذي أطوال فئاته غير متساوية ولو فئة واحدة أو هو الجدول الذي فيه فئة واحدة على الأقل مختلفة الطول عن باقي أطوال الفئات ويأخذ الصورة التالية:

| عرض الفئة<br>Class Width | التكرار<br>Frequency | الفئات<br>Class |
|---|---|---|
| 9 = 24 – 15 | 12 | 24 15– |
| 10 = 35 – 25 | 14 | 35 25– |
| 89 = 44 – 36 | 8 | 36 – 44 |

**3.الجدول التكراري المغلق:**

هو الجدول المحدد الحد الأدنى للفئة الأولى والحد الأعلى للفئة الأخيرة ويأخذ الصورة التالية:

| جدول تكراري مغلق فيه عرض الفئة ثابت | | | جدول تكراري مغلق فيه عرض الفئة متغير | | |
|---|---|---|---|---|---|
| **الفئات** | **التكرار** | **طول الفئة** | **الفئات** | **التكرار** | **طول الفئة** |
| 24 – 15 | 12 | 9 15= – 24 | 24 – 15 | 12 | 10 = 25 – 24 |
| 34 – 25 | 14 | 9 = 25 – 34 | 35 – 25 | 14 | 10 = 25 – 35 |
| 44 – 35 | 8 | 9 = 35 – 44 | 44 – 36 | 8 | 8 = 36 – 44 |

**4. الجدول التكراري المفتوح:**

هو الجدول الذي يصعب تحديد الحد الأدنى للفئة الأولى أو الحد الأعلى للفئة الأخيرة أو الاثنين معاً ويأخذ إحدى الصور التالية:

| جدول يصعب تحديد الحد الأدنى للفئة الأولى | | جدول يصعب تحديد الحد الأعلى للفئة الأخيرة | | جدول غير محدد الحدين للفئتين الأدنى والأعلى | |
|---|---|---|---|---|---|
| **الفئات** | **التكرار** | **الفئات** | **التكرار** | **الفئات** | **التكرار** |
| أقل من 15 – 24 | 12 | 24 – 15 | 12 | أقل من 15 – 24 | 12 |
| 34 – 25 | 14 | 34 – 25 | 14 | 34 – 25 | 14 |
| 44 – 35 | 8 | 35 فأكثر | 8 | 35 فأكثر | 8 |

**ملاحظة:**

الجداول المفتوحة يصعب تحديد مركز الفئة المفتوحة، أو تمثيله بيانياً وحساب بعض المقاييس.

**(4–4) التوزيعات التكرارية المتجمعة:**

**1. التوزيعات التكرارية المتجمعة لبيانات كمية مستمرة:**

جدول التوزيع المتجمع الصاعد والهابط لدرجات 40 طالباً في مادة الرياضيات:

| التوزيع التكراري المتجمع الصاعد (يبدأ بصفر وينتهي بمجموع التكرارات) | | التوزيع التكراري المتجمع الهابط (يبدأ بمجموع التكرارات وينتهي بصفر) | |
|---|---|---|---|
| **الحدود العليا للفئات** | **التكرار الصاعد** | **الحدود الدنيا للفئات** | **التكرار الهابط** |
| أقل من صفر | صفر | صفر فأكثر | 40 |
| أقل من 10 | 12 | 10 فأكثر | 28 |
| أقل من 20 | 26 | 20 فأكثر | 14 |
| أقل من 30 | 32 | 30 فأكثر | 8 |
| أقل من أو يساوي 40 | 40 | أكثر من 40 | صفر |

**2.التوزيعات التكرارية المتجمعة لبيانات وصفية (متقطعة):**

**مثال (1)**: نفرض أننا ألقينا زهرة نرد أربعين مرة، وكانت التي ظهرت في كل مرة كالآتي:

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **1** | **2** | **3** | **4** | **1** | **5** | **3** | **6** | **2** | **3** |
| **1** | **2** | **3** | **4** | **1** | **5** | **3** | **6** | **2** | **3** |
| **6** | **3** | **5** | **2** | **3** | **4** | **2** | **6** | **1** | **4** |
| **6** | **3** | **5** | **2** | **3** | **4** | **2** | **6** | **1** | **4** |

**المطلوب**: ضع هذه البيانات في صورة جدول توزيع تكراري ثم كون جدول تكراري متجمع صاعد وهابط.

**الحل**: الجدول التكراري

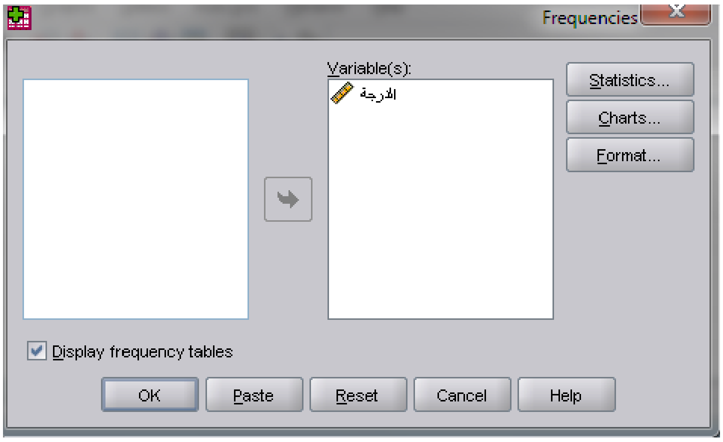
| **Score occurs النتائج** | **Tally العلامات** | **Frequency التكرارات** |
|---|---|---|
| 1 | \| 卌 | 6 |
| 2 | \|\|\| 卌 | 8 |
| 3 | 卌 卌 | 10 |
| 4 | \| 卌 | 6 |
| 5 | \|\|\|\| | 4 |
| 6 | \| 卌 | 6 |

**باستخدام الحزمة الإحصائية Spss**

ندخل البيانات وباتباع المسار التالي:

**Analyze** ⇨ **Descriptive Statistics** ⇨ **Frequencies**

يظهر الشكل التالي:




وبالضغط على Format لنحدد شكل البيانات الناتجة فنحصل على الجدول التالي:

| الدرجة | | | | | |
|---|---|---|---|---|---|
| | | Frequency | Percent | Valid Percent | Cumulative Percent |
| Valid | 1 | 6 | 15.0 | 15.0 | 15.0 |
| | 2 | 8 | 20.0 | 20.0 | 35.0 |
| | 3 | 10 | 25.0 | 25.0 | 60.0 |
| | 4 | 6 | 15.0 | 15.0 | 75.0 |
| | 5 | 4 | 10.0 | 10.0 | 85.0 |
| | 6 | 6 | 15.0 | 15.0 | 100.0 |
| | Total | 40 | 100.0 | 100.0 | |

التوزيع التكراري المتجمع (الصاعد والهابط):

| التوزيع التكراري المتجمع الصاعد | | التوزيع التكراري المتجمع الهابط | |
|---|---|---|---|
| الحدود العليا للفئات | التكرار الصاعد | الحدود الدنيا للفئات | التكرار الهابط |
| أقل من 1 | صفر | أكثر من صفر | 40 |
| أقل من أويساوي 1 | 6 | أكثر من 1 | 34 |
| أقل من أويساوي 2 | 14 | أكثر من 2 | 26 |
| أقل من أويساوي 3 | 24 | أكثر من 3 | 16 |
| أقل من أويساوي 4 | 30 | أكثر من 4 | 10 |
| أقل من أويساوي 5 | 34 | أكثر من 5 | 6 |
| أقل من أو يساوي 6 | 40 | أكثر من 6 | صفر |

**مثال (2)**: إذا كان لدينا تقديرات أربعين طالباً في مادة ما وكانت النتائج كالتالي:

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| ضعيف جداً | ضعيف جداً | ضعيف جداً | ضعيف | جيد | مقبول | ضعيف | ممتاز |
| جيد جداً | ضعيف جداً | جيد جداً | ضعيف | جيد | مقبول | مقبول | ضعيف |
| جيد جداً | مقبول | ضعيف | مقبول | مقبول | جيد | ضعيف | مقبول |
| مقبول | ضعيف جداً | جيد | مقبول | ممتاز | ضعيف | جيد | جيد |
| ممتاز | ضعيف جداً | ضعيف | ممتاز | ممتاز | مقبول | مقبول | ممتاز |

ضع هذه البيانات في صورة جدول توزيع تكراري ثم كون جدولاً تكرارياً متجمعاً صاعداً وهابطاً؟

**الحل**: الجدول التكراري

| التقدير | العلامات | التكرارات |
|---|---|---|
| ضعيف جداً | $\cancel{////}$ / | 6 |
| ضعيف | $\cancel{////}$ /// | 8 |
| مقبول | $\cancel{////}$ $\cancel{////}$ | 10 |
| جيد | $\cancel{////}$ / | 6 |
| جيد جداً | //// | 4 |
| ممتاز | $\cancel{////}$ / | 6 |

### 1. التوزيع التكراري المتجمع الصاعد:

Ascending cumulative frequency distribution

ويسمى أيضا أكثر من التوزيع التكراري المتجمع:

Less than Cumulative Frequency: Distribution

| التوزيع التكراري المتجمع الصاعد<br>Ascending cumulative frequency distribution | |
|---|---|
| الحدود العليا للفئات | التكرار الصاعد |
| أقل من ضعيف جداً | صفر |
| أقل من أو يساوى ضعيف جداً | 6 |
| أقل من أو يساوى ضعيف | 14 |
| أقل من أو يساوى مقبول | 24 |
| أقل من أو يساوى جيد | 30 |
| أقل من أو يساوى جيد جداً | 34 |
| أقل من أو يساوى ممتاز | 40 |

**وباستخدام Spss**

| التقدير | | | | | |
|---|---|---|---|---|---|
| | | Frequency | Percent | Valid Percent | Cumulative Percent |
| Valid | ضعيف جداً | 6 | 15.0 | 15.0 | 15.0 |
| | ضعيف | 8 | 20.0 | 20.0 | 35.0 |
| | مقبول | 10 | 25.0 | 25.0 | 60.0 |
| | جيد | 6 | 15.0 | 15.0 | 75.0 |
| | جيد جداً | 4 | 10.0 | 10.0 | 85.0 |
| | ممتاز | 6 | 15.0 | 15.0 | 100.0 |
| | Total | 40 | 100.0 | 100.0 | |

**مثال2:**

'Less than' cumulative frequency table

| Marks | Frequency | Less than c.f. |
|---|---|---|
| 20 - 30 | 2 | 2 |
| 30 - 40 | 8 | 10 |
| 40 - 50 | 10 | 20 |
| 50 - 60 | 20 | 40 |
| 60 - 70 | 12 | 52 |
| 70 - 80 | 8 | 60 |

**2. التوزيع التكراري المتجمع التنازلي:**

**Descending cumulative frequency distribution:**

ويسمى أيضا أكثر من التوزيع التكراري المتجمع:

More than Cumulative Frequency: Distribution

| التوزيع التكراري المتجمع الهابط (التنازلي)<br>Descending cumulative frequency distribution | |
|---|---|
| الحدود الدنيا للفئات | التكرار الهابط |
| أكثر من أو يساوى ضعيف جداً | **40** |
| أكثر من ضعيف جداً | **34** |
| أكثر من ضعيف | **26** |
| أكثر من مقبول | **16** |
| أكثر من جيد | **10** |
| أكثر من جيد جداً | **6** |
| أكثر من ممتاز | صفر |

**مثال2:**

'More than' cumulative frequency table

| Marks | Frequency | More than c.f. |
|---|---|---|
| 20-30 | 2 | 60 |
| 30-40 | 8 | 58 |
| 40-50 | 10 | 50 |
| 50-60 | 20 | 40 |
| 60-70 | 10 | 20 |
| 70-80 | 8 | 8 |
| Total | 60 | |

### (4–5) التوزيعات التكرارية المزدوجة:

هي التوزيعات الخاصة بظاهرتين متلازمتين، وقد تكون ظاهرتين وصفيتين، أو كميتين، أو ظاهرة وصفية وأخرى كمية.

**مثال**: فيما يلي درجات 40 طالباً في الصف الأول الثانوي في مادتي الرياضيات والإحصاء.

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| الرياضيات | 3 | 6 | 8 | 6 | 9 | 6 | 13 | 6 | 4 | 11 |
| الإحصاء | 6 | 6 | 9 | 14 | 11 | 16 | 26 | 16 | 26 | 20 |
| الرياضيات | 4 | 7 | 6 | 7 | 10 | 8 | 14 | 8 | 5 | 10 |
| الإحصاء | 7 | 7 | 5 | 12 | 14 | 16 | 29 | 16 | 28 | 22 |
| الرياضيات | 5 | 8 | 3 | 7 | 12 | 7 | 13 | 7 | 9 | 11 |
| الإحصاء | 8 | 9 | 11 | 13 | 15 | 19 | 22 | 19 | 21 | 23 |
| الرياضيات | 3 | 6 | 5 | 8 | 13 | 8 | 14 | 8 | 10 | 9 |
| الإحصاء | 9 | 9 | 12 | 12 | 14 | 18 | 23 | 18 | 20 | 24 |

**المطلوب**: وضع هذه البيانات في صورة جدول توزيع تكراري.

**الحل**:

جدول التوزيع التكراري المزدوج.

| الإحصاء / الرياضيات | -5 | -10 | -15 | -20 | -25 30 | مجموع |
|---|---|---|---|---|---|---|
| -3 | 4 | 2 | | | 2 | 8 |
| -6 | 6 | 4 | 8 | | | 18 |
| -9 | | 2 | | 6 | | 8 |
| 15-12 | | | 2 | 2 | 2 | 6 |
| مجموع | 10 | 8 | 10 | 8 | 4 | 40 |

التوزيع الهامشي لدرجات الطلبة في الرياضيات فقط.

| الفئات | -3 | -6 | -9 | 15-12 | مجموع |
|---|---|---|---|---|---|
| التكرارات | 8 | 18 | 8 | 6 | 40 |

التوزيع الهامشي لدرجات الطلبة في الإحصاء فقط.

| الفئات | -5 | -10 | -15 | -20 | 30- 25 | مجموع |
|---|---|---|---|---|---|---|
| التكرارات | 10 | 8 | 10 | 8 | 4 | 40 |

- **ظاهرتين وصفتين:**

فيما يلي نتائج الإجابة على استبيان لدراسة العلاقة بين الوفاة والتدخين كالتالي:

| الشخص | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| الوفاة | نعم | نعم | لا | لا | لا | نعم | لا | لا | نعم | لا |
| التدخين | نعم | لا | نعم | لا | لا | لا | نعم | نعم | نعم | لا |
| الشخص | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
| الوفاة | نعم | نعم | لا | لا | لا | نعم | لا | لا | نعم | لا |
| التدخين | نعم | لا | نعم | لا | لا | لا | نعم | نعم | نعم | لا |

**المطلوب**: وضع هذه البيانات في صورة جدول توزيع تكراري.

**الحل**: الجدول التكراري.

| التدخين \ الوفاة | نعم | لا | مجموع |
|---|---|---|---|
| نعم | 4 | 6 | 10 |
| لا | 4 | 6 | 10 |
| مجموع | 8 | 12 | 20 |

**ظاهرتين أحداهما وصفية والأخرى كمية:**

في أحد الأبحاث تمت دراسة العلاقة بين حضور المحاضرة ودرجة الطالب في امتحان إحدى المواد التعليمية.

| رقم الطالب | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| الحضور | ح | غ | ح | غ | ح | غ | ح | غ | ح | غ |
| الدرجة | 8 | 7 | 4 | 5 | 7 | 6 | 10 | 10 | 9 | 2 |
| رقم الطالب | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
| الحضور | غ | ح | ح | ح | غ | ح | ح | ح | ح | غ |
| الدرجة | 12 | 14 | 15 | 16 | 17 | 18 | 19 | 18 | 19 | 15 |

**المطلوب**: وضع هذه البيانات في صورة جدول توزيع تكراري.

**الحل**:

جدول تكراري مزدوج لحضور المحاضرة ودرجة الطالب في امتحان إحدى المواد التعليمية.

| درجة الطالب \ الحضور | ح | غ | مجموع |
|---|---|---|---|
| -1 | 1 | 2 | 3 |
| -6 | 4 | 3 | 7 |
| -11 | 2 | 2 | 4 |
| 20 – 16 | 5 | 1 | 6 |

| | | | |
|---|---|---|---|
| 20 | 8 | 12 | مجموع |

<table>
<tr><th>Cores</th><th>Tally</th><th>Frequency</th></tr>
<tr><td>100 – 96</td><td>|||</td><td>3</td></tr>
<tr><td>95 – 91</td><td>||||</td><td>4</td></tr>
<tr><td>90 – 86</td><td>||| |||||</td><td>8</td></tr>
<tr><td>85 – 81</td><td>|| |||||</td><td>7</td></tr>
<tr><td>80 – 76</td><td>| |||||</td><td>6</td></tr>
<tr><td>75 – 71</td><td>|||||</td><td>5</td></tr>
<tr><td>70 – 66</td><td>|||</td><td>3</td></tr>
<tr><td>65 – 61</td><td>||</td><td>2</td></tr>
<tr><td>60 – 56</td><td>||</td><td>2</td></tr>
<tr><td>$\sum$</td><td></td><td>40</td></tr>
</table>

**مثال**: درجات 40 طالب في مادة الرياضيات، المطلوب عمل جدول تكراري.

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 76 | 88 | 86 | 72 | 89 | 87 | 73 | 56 | 82 | 86 |
| 93 | 84 | 78 | 70 | 97 | 62 | 85 | 84 | 69 | 72 |
| 81 | 92 | 72 | 65 | 94 | 83 | 76 | 91 | 60 | 70 |
| 80 | 74 | 83 | 90 | 89 | 96 | 76 | 88 | 78 | 98 |

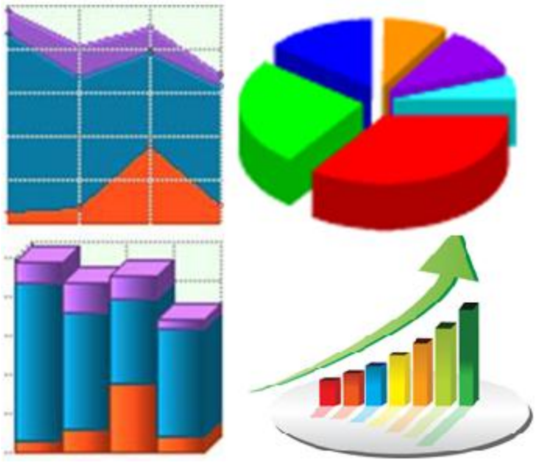
**المطلوب**: تكوين 10 فترات.

**الحل**: المدى = 98 − 56 = 4.2، مدى الفترة تقريباً = 4.2 تقريباً 5 فنحصل على الجدول التالي:

# الباب الخامس

# العرض البياني للبيانات الإحصائية

# Graphical Presentation



- مقدمة.

(5-1) العرض البياني لبيانات غير مبوبة.

(5– 2) العرض البياني للبيانات المبوبة.

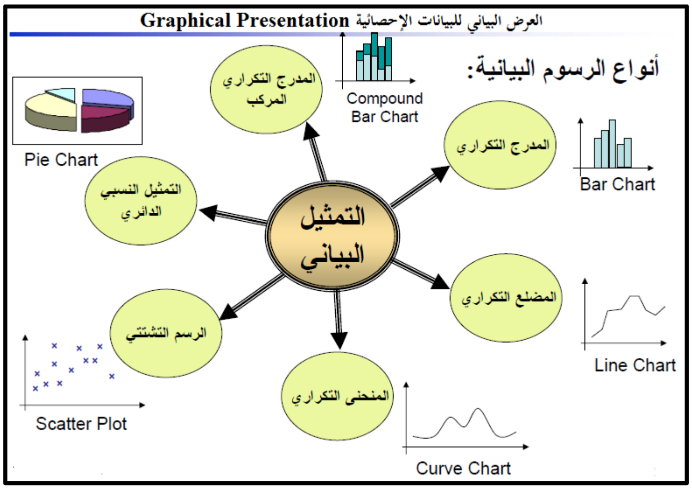
(5– 3) التوزيعات التكرارية.

(5-4) التمثيل البياني للبيانات المتقطعة.

## مقدمة:

كما نعلم أن الهدف الأساسي لعملية جمع البيانات هو التعرف على خصائص الظاهرة محل البحث أو الدراسة، ويتم ذلك من خلال تبويب وعرض هذه البيانات إما في صورة جداول تكرارية أو أشكال هندسية تسهم في عملية المقارنة وتجعل عملية تفسير البيانات سهلة، والعرض البياني هو تعبير تصويري للعلاقة بين المتغيرات وله عدة صور تتوقف على طبيعة المتغيرات وتعتبر الرسوم البيانية أكثر الأساليب الإحصائية استخداماً لوصف وتلخيص البيانات، منها:

1. الأعمدة (Bar Charts).
2. الدوائر المجزأة (Pie Chart).
3. الخطوط البيانية (Graphs).
4. الأعمدة البسيطة (Simple Bar Charts).
5. الأعمدة المتلاصقة (Multiple Bar Charts).
6. الأعمدة المركبة (Component Bar Charts).



**(1–5) العرض البياني لبيانات غير مبوبة:**

يمكن عرض البيانات غير المبوبة، أي المفردات لدراسة ظاهرة معينة في صورة بيانية توضح المقارنة بين المفردات، كما يمكن أن تكون ظروف الدراسة والمقارنة تتطلب دراسة ظاهرتين أو أكثر تتوافر عن كل منهما بيانات في صورة مفردة.

- **عرض البيانات في صورة أعمدة (مستطيلات) (Bar chart) لظاهرة واحدة.**

إذا توافرت بيانات مقارنة لظاهرة واحدة بحيث يكون عدد المفردات محدداً فإنه يمكن التعبير عن هذه البيانات في صورة أعمدة (مستطيلات) تكون منفصلة عن بعضها مع مراعاة أن يكون عرض هذه المستطيلات والأعمدة ثابتاً لأن عملية المقارنة تعتمد على (أطوال الأعمدة أو المستطيلات.

**مثال:** فيما يلي عدد الطلاب بإحدى المدارس والمطلوب تمثيل ذلك بيانياً في شكل أعمدة بسيطة.

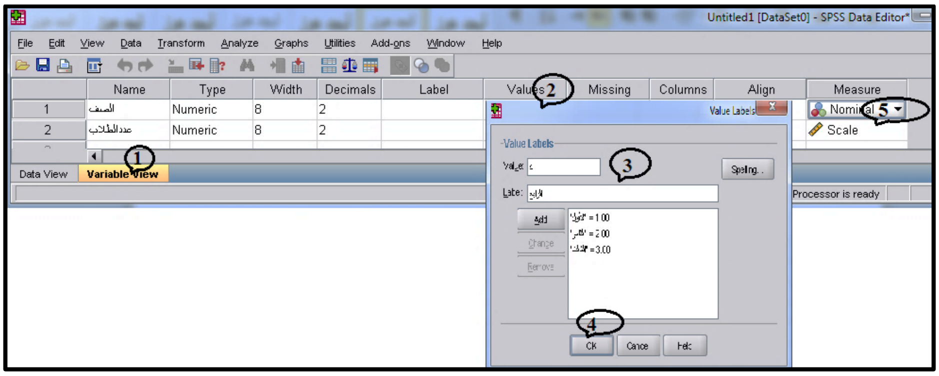
| **الصف** | **الأول** | **الثاني** | **الثالث** | **الرابع** | **المجموع** |
|---|---|---|---|---|---|
| **عدد الطلاب** | 60 | 80 | 40 | 20 | 200 |

**الحل:**

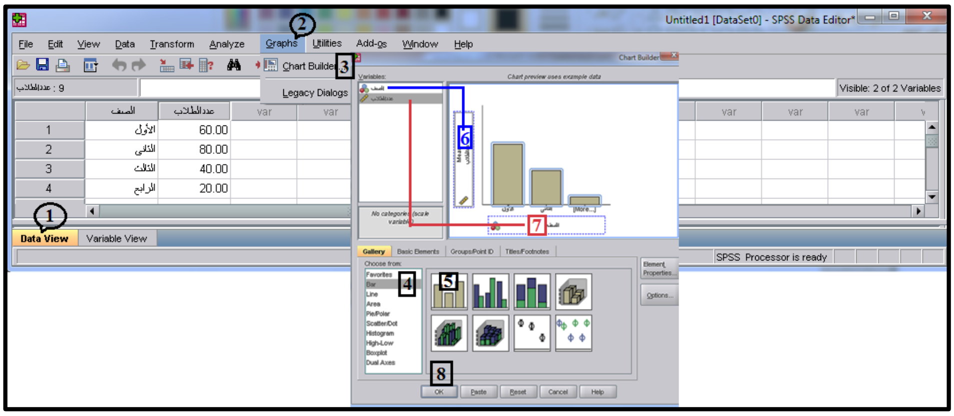
إدخال البيانات إلى برنامج Spss بكتابة المتغيرات وبالنسبة للمتغير الفترة نضغط على values ونسمى الصف ثم إدخال البيانات تحت كل متغير (Data View)، ولرسم الأعمدة البيانية نتتبع الخطوات التالية:

Graphs $\Rightarrow$ Chart Builder

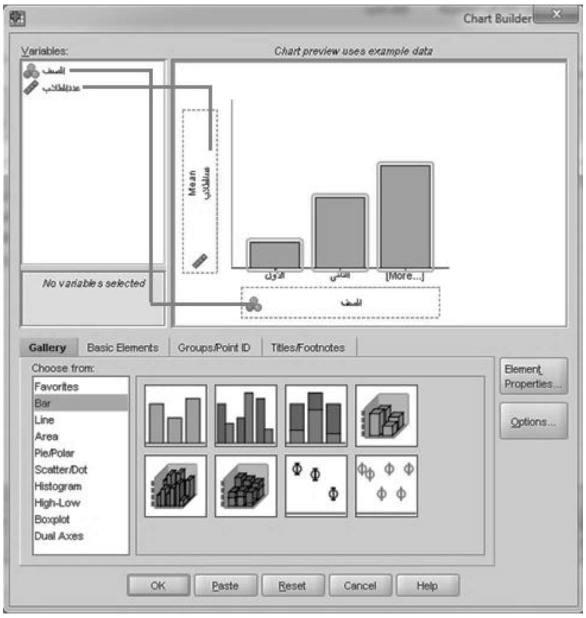
وبالضغط على Chart Builder نحصل على الشكل الحواري ومن قائمة Choose from نختار Bar ثم الضغط على أيقونة لبدء رسم الأعمدة ثم يتم إدراج الفئات والتكرار باستخدام الفأرة كما في الأشكال التالية:



**شكل (1.5)**: يوضح ترميز متغير الصف.

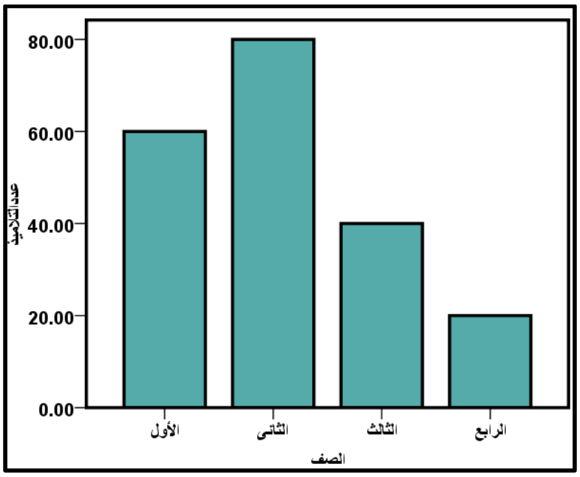


**شكل (2.5)**: يوضح إدراج المتغيرين لرسم أعمدة مستطيلة.



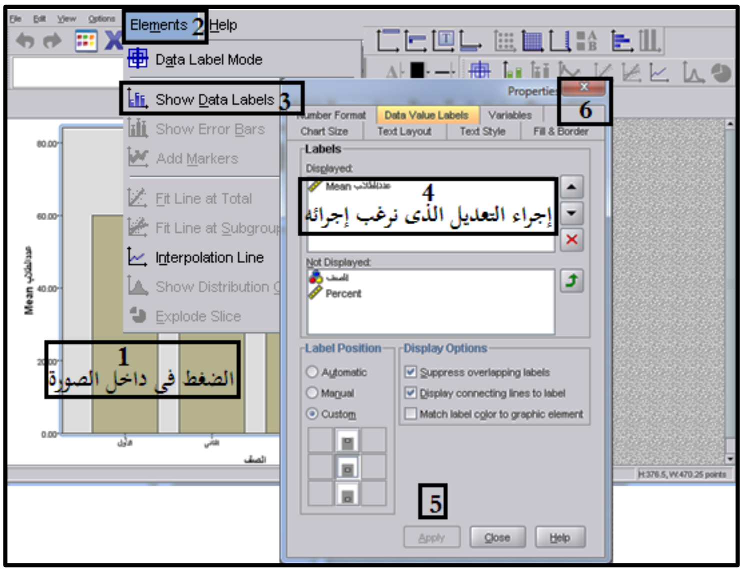

**شكل (3.5)**: يوضح تمثيل البيانات على المحورين الأفقي والرأسي.

وبالضغط على ok نحصل على الشكل التالي:

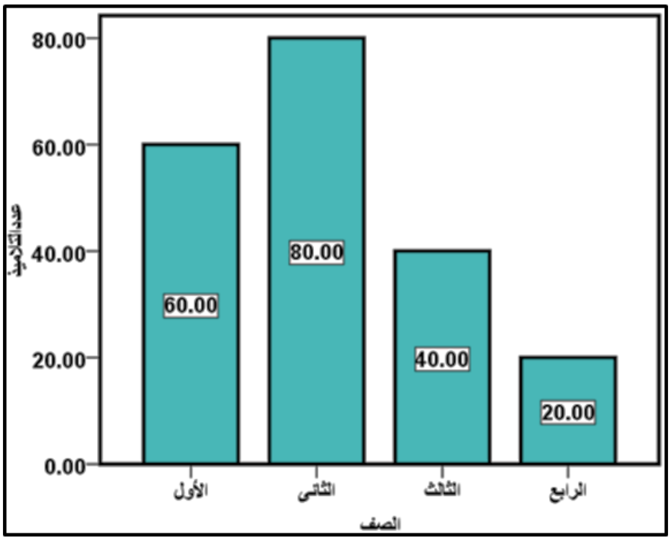



**شكل (4.5)**: يوضح تمثيل البيانات في صورة أعمدة (مستطيلات).

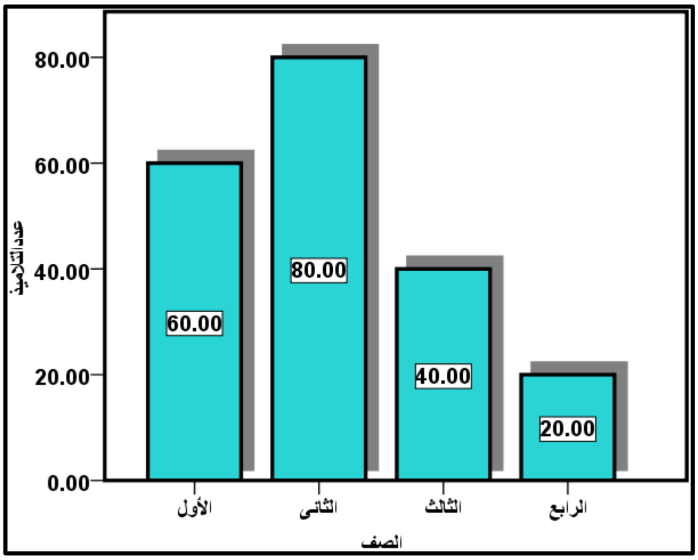
ويمكن إجراء تعديلات بالضغط في داخل الصورة كما في الشكل التالي:



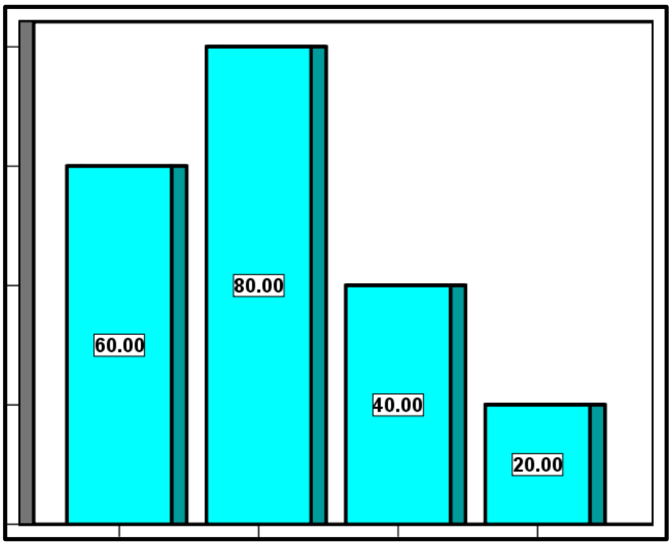

فنحصل على العديد من الأشكال حسب ما نريد إضافته على الصورة كما يلي:




**شكل (5.5)**: يوضح تمثيل البيانات في صورة أعمدة (مستطيلات) مدون عليها البيانات.




**شكل (6.5)**: يوضح تمثيل البيانات في صورة أعمدة (مستطيلات) مدون عليها البيانات مع إظهار ظل الأعمدة وتغيير اللون




**شكل (7.5)**: يوضح تمثيل البيانات في صورة أعمدة (مستطيلات) ثلاثية الأبعاد.

**مثال**: لدينا درجات 50 طالب في إحدى المواد وهي:

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 22 | 44 | 28 | 21 | 19 | 23 | 37 | 51 | 42 | 88 |
| 33 | 46 | 31 | 39 | 20 | 50 | 40 | 41 | 11 | 7 |
| 56 | 72 | 56 | 17 | 7 | 69 | 30 | 80 | 56 | 29 |
| 39 | 36 | 77 | 73 | 59 | 34 | 29 | 18 | 41 | 78 |
| 17 | 54 | 44 | 53 | 31 | 39 | 67 | 54 | 62 | 30 |

**المطلوب**: ضع هذه البيانات في صورة جدول توزيع تكراري ثم تمثيلها بالأعمدة البيانية باستخدام برنامج Spss.

**الحل**:

**1.** يتم ذكر عدد من الفئات (7).

تحديد طول الفئة = (أكبر قيمة − أصغر قيمة) ÷ عدد الفئات (المقترحة)

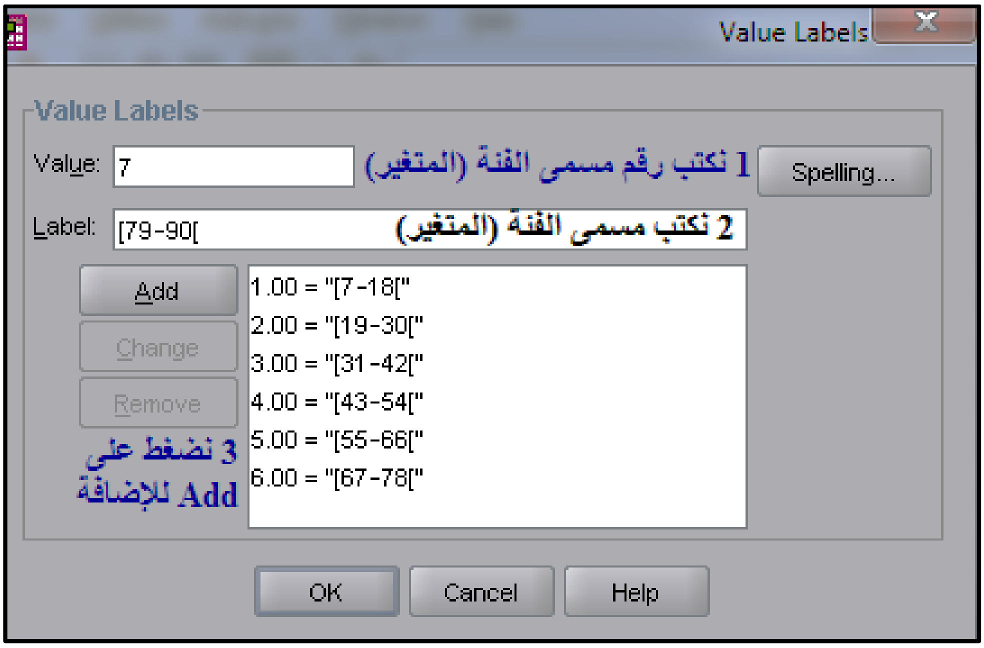
**Class width** = (Maximum entry − Minimum entry)/Number of classes

Class width = (88 − 7) / 7 = 81/7 ≈ 11.57 Round up to 12**.**

**2.** إدخال البيانات الحد الأدنى هو 7 وإدخال البيانات الحد الأقصى هو 88، ونكون الجدول التالي:

| **Class** | Tally | *f* | Midpoint | Relative frequency | Cumulative frequency |
|---|---|---|---|---|---|
| [7-18] | 𝍸 𝍷 | 6 | (7+18) ÷ 2 =12.5 | 6÷50 =0.12 | 6 |
| [19-30] | 𝍸 𝍸 | 10 | 24.5 | 0.2 | 6+10 =16 |
| [31-42] | 𝍸𝍸 | 13 | 36.5 | 0.26 | 29 |
| [43-54] | 𝍸 𝍷𝍷𝍷 | 8 | 48.5 | 0.16 | 37 |
| [55-66] | 𝍸 | 5 | 60.5 | 0.1 | 42 |
| [67-78] | 𝍸 𝍷 | 6 | 72.5 | 0.12 | 48 |
| [79-90] | 𝍷𝍷 | 2 | 84.5 | 0.04 | 50 |
| $\sum$ | | 50 | − | 1 | − |

**3.** إدخال البيانات إلى برنامج Spss بكتابة المتغيرات، وبالنسبة للمتغير الفترة نضغط على values ونسمى الفترات كما في الشكل التالي:

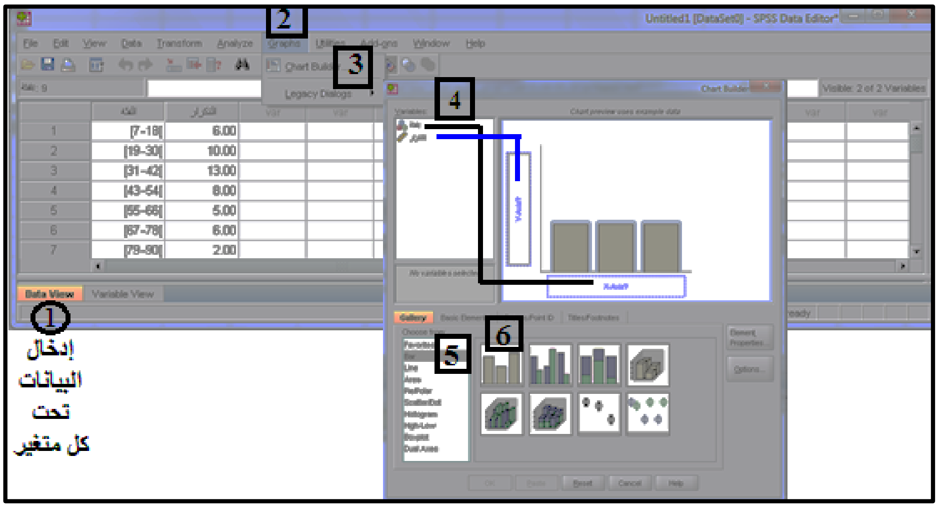



**شكل (8.5)**: يوضح ترميز متغير الفئة.

ثم إدخال البيانات تحت كل متغير (Data View)، ولرسم الأعمدة البيانية نتتبع الخطوات التالية:

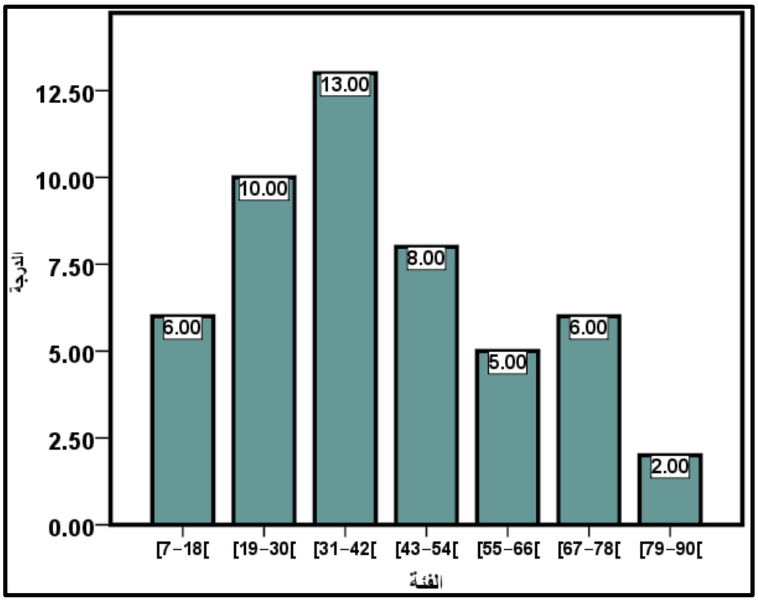
Graphs ⇒ Chart Builder

و بالضغط على Chart Builder نحصل على الشكل الحوارى ومن قائمة Choose from نختار Bar ثم الضغط على أيقونة  لبدء رسم الأعمدة ثم يتم إدراج الفئات والتكرار باستخدام الفارة كما في الشكل التالي:



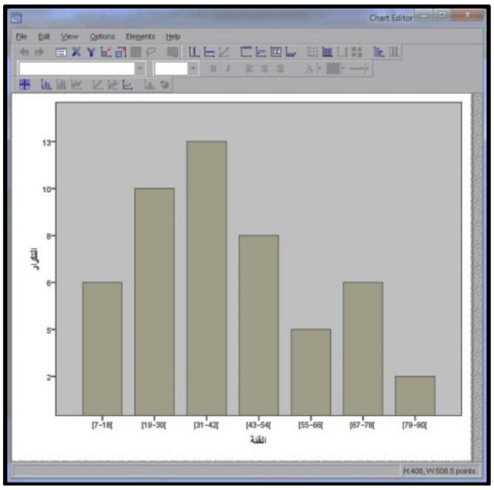

**شكل (9.5)**: يوضح إدراج المحورين السيني والصادي.

وبالضغط على Ok نحصل على الأعمدة البيانية كما في الشكل التالي:

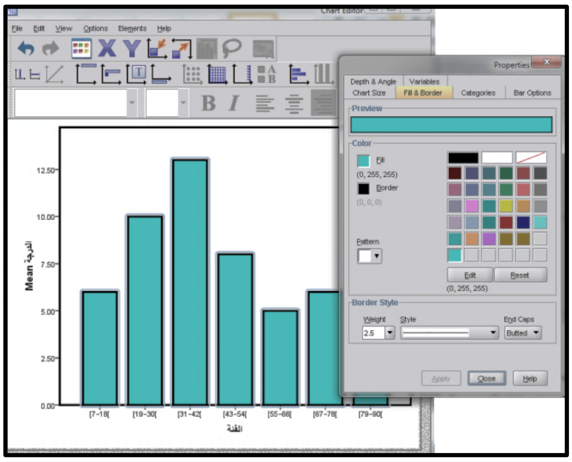



**شكل (10.5)**: يوضح الأعمدة البسيطة للفئات.

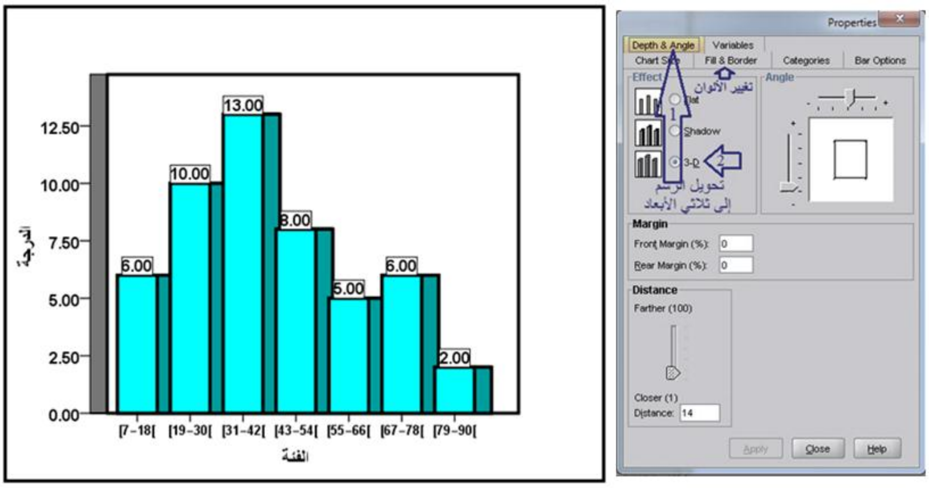
وبالضغط على الرسم بالفارة ضغطتين نحصل مدقق الرسم Chart Editor كما في الصندوق الحواري التالي:



ويمكن إجراء تعديلات على الرسم فبالضغط على (Elements) نحصل على الصندوق الحواري التالي:



ومن الممكن إضافة النسب المئوية أو تغيير الألوان أو تحويل الرسم إلى ثلاثي الأبعاد بالضغط على الأعمدة فتظهر النافذة:




**شكل (11.5)**: يوضح الأعمدة البسيطة ثلاثية الأبعاد للفئات.

**ب. عرض البيانات في صورة الأعمدة المزدوجة (المتلاصقة):**

**Double Bar Graph**

تستخدم إذا كانت هناك سلسلتان أو أكثر من القيم لظاهرتين أو أكثر أو لظاهرة ذات عدة أوجه مختلفة في عدد السنوات أو الأماكن المختلفة ... الخ.

ولسهولة إجراء المقارنات يمكن تظليل أيهما أو إعطاء كل منهما لوناً مختلفاً عن الآخر، أي وضع مفتاح في صورة مربع يبين دلالة العمود حسب اللون أو الظل المستخدم.

**مثال**: فيما يلي بيان بأعداد الناجحين في مقرر دراسي في دور مايو من كل عام.

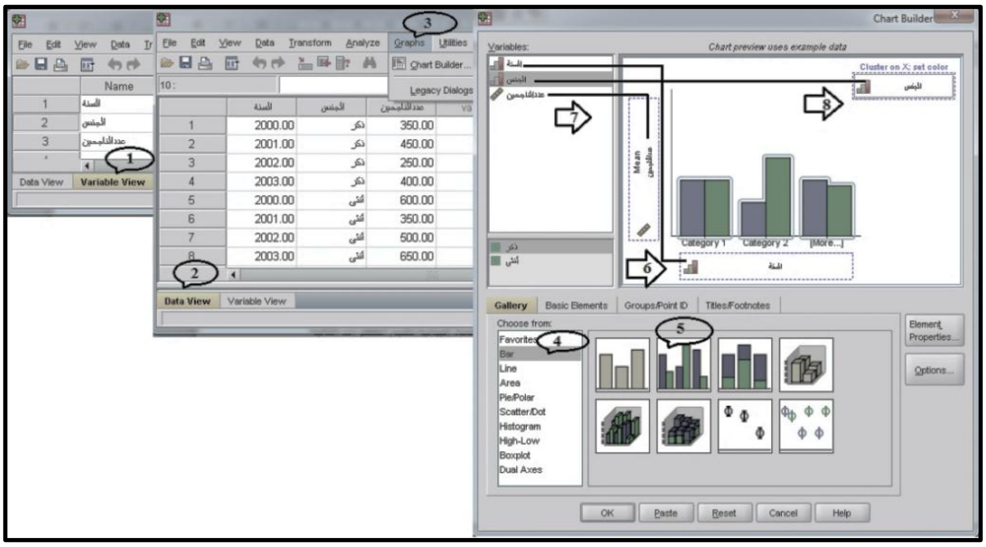
| السنة | 2000 | | 2001 | | 2002 | | 2003 | |
|---|---|---|---|---|---|---|---|---|
| | ذكور | إناث | ذكور | إناث | ذكور | إناث | ذكور | إناث |
| عدد الناجحين | 350 | 600 | 450 | 350 | 250 | 500 | 400 | 650 |

**والمطلوب**: عبّر بيانياً باستخدام الأعمدة المزدوجة:

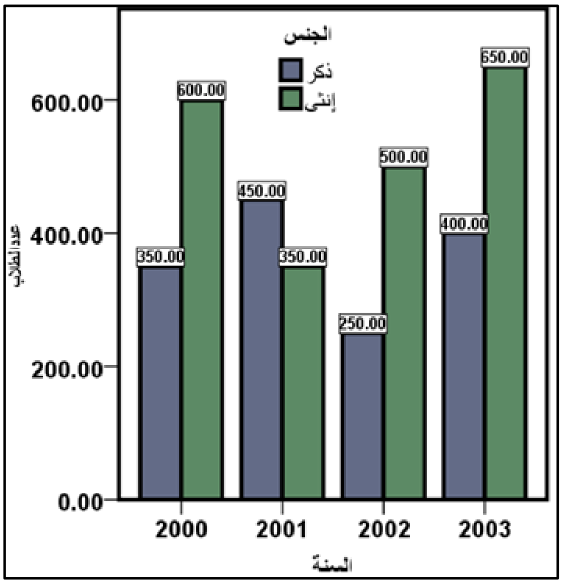
**الحل:**

إدخال البيانات إلى برنامج Spss بكتابة المتغيرات، وبالنسبة للمتغير الجنس نضغط على values ونسمى الجنس، ثم إدخال البيانات تحت كل متغير (Data View)، ولرسم الأعمدة البيانية نتتبع الخطوات التالية:

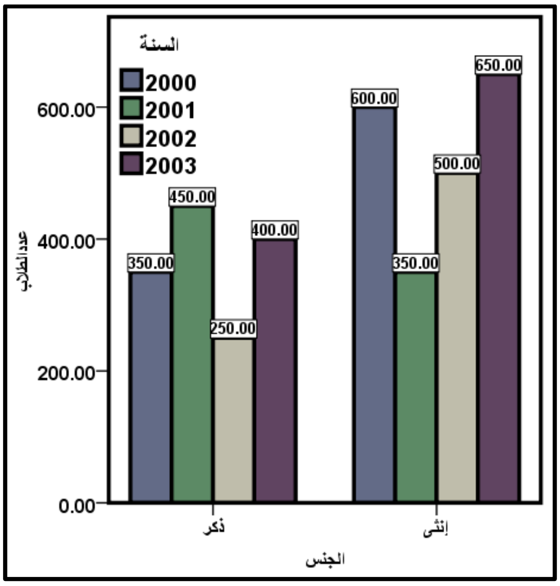
Graphs ثم Chart Builder، وبالضغط على Chart Builder نحصل على الشكل الحواري ومن قائمة Choose from نختار Bar ثم الضغط على أيقونة  لبدء رسم الأعمدة ثم يتم إدراج الفئات والتكرار باستخدام الفارة، وبالضغط على Ok كما في الصندوق الحواري التالي:




فنحصل على الأعمدة البيانية كما في الشكلين التاليين:




**شكل (12.5 – أ)**: يوضح تمثيل عدد الناجحين بالأعمدة البسيطة المزدوجة.




**شكل (12.5 – ب)**: يوضح تمثيل عدد الناجحين بالأعمدة البسيطة المزدوجة

**ج) الأعمدة المركبة (المجزأة):**

تستخدم إذا كانت هناك ظاهرة ما تتكون جملتها من عدة أجزاء من نوعيات مختلفة فمثلاً إجمالي عدد الطلاب الناجحين في فرقة ما جزء يتكون من الذكور والجزء الآخر من الإناث ... الخ.

**مثال:**

| السنة | 2000 | | 2001 | | 2002 | | 2003 | |
|---|---|---|---|---|---|---|---|---|
| | ذكور | إناث | ذكور | إناث | ذكور | إناث | ذكور | إناث |
| عدد الناجحين | 350 | 600 | 450 | 350 | 250 | 500 | 400 | 650 |

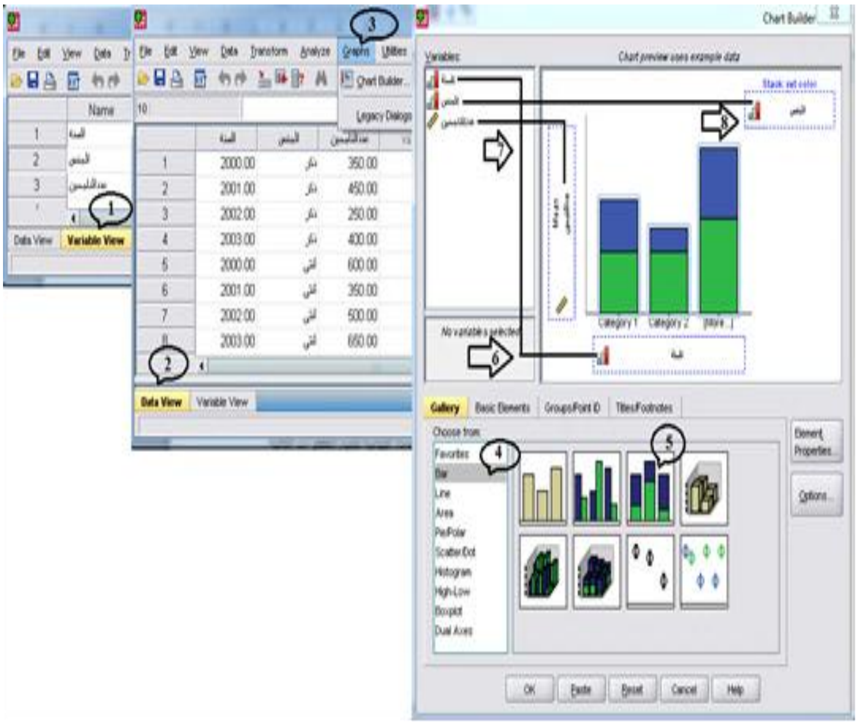
**والمطلوب:** عبر بيانياً باستخدام الأعمدة المركبة (المجزأة):

**الحل:** يراعى عند استخدام الأعمدة المركبة (المجزأة) حساب إجمالي كل عمود حتى يمكنك تقسيم المحور الرأسي الذي يعبر عن المتغير.

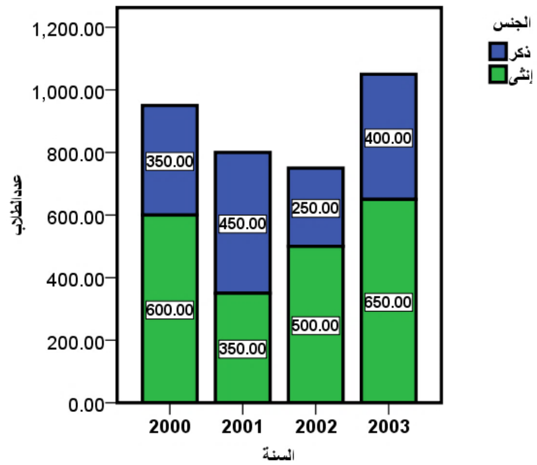
إدخال البيانات إلى برنامج Spss بكتابة المتغيرات، وبالنسبة للمتغير الجنس نضغط على values ونسمى الجنس، ثم إدخال البيانات تحت كل متغير (Data View)، ولرسم الأعمدة البيانية نتتبع الخطوات التالية:

Graphs $\Rightarrow$ Chart Builder

وبالضغط على Chart Builder نحصل على الشكل الحواري ومن قائمة Choose from نختار Bar ثم الضغط على أيقونة لبدء رسم الأعمدة ثم يتم إدراج الفئات والتكرار باستخدام الفارة، وبالضغط على Ok كما في الشكل التالي:



فنحصل على الأعمدة البيانية كما في الشكل التالي:




**شكل (13.5)**: يوضح تمثيل عدد الناجحين بالأعمدة المركبة (المجزأة).

### د. عرض البيانات في صورة دوائر (الدوائر المجزأة): Pie graph or pie chart

حيث تمثل الدائرة مجموع القيم الكلية للظاهرة، فيتم تقسيمها إلى قطاعات جزئية وتميز تلك القطاعات عن بعضها إما بألوان مختلفة أو بظلال مختلفة من أجل ضمان الإيضاح.

ويستخدم هذا النوع من الرسوم البيانية في الحالات التالية:

**1.** عندما تكون الأجزاء المقارنة قليلة العدد نسبياً.

**2.** لتوضيح التطور النسبي لأجزاء الظاهرة لفترات زمنية مختلفة.

**3.** لمقارنة الأجزاء المختلفة بالنسبة للمجموع الكلي لبيانات وصفية (غير كمية).

ويحتاج استخدام الدوائر لعرض هذه البيانات الإلمام ببعض الحقائق الرياضية عن الدائرة.

**1.** مساحة الدائرة = ط نق$^2$ ط: النسبة التقريبية = 22/7=3.14 نق: نصف قطر الدائرة.

**2.** الزاوية المركزية = $360^\circ$.

ولتحديد زاوية كل جزء:

الزاوية المركزية المناظرة لكل جزء = (قيمة الجزء ÷ مجموع الأجزاء) $360x^\circ$

**مثال:** الجدول التالي يبين توزيع تلاميذ إحدى المدارس:

| الصف | الأول | الثاني | الثالث | الرابع | المجموع |
|---|---|---|---|---|---|
| عدد الطلاب | 60 | 80 | 40 | 20 | 200 |

**المطلوب:** عبر بيانياً باستخدام الدائرة.

**الحل:**

**أولاً**: تحديد الزاوية المركزية المناظرة لكل قطاع (جزء).

الزاوية المركزية المناظرة لكل جزء = (قيمة الجزء ÷ مجموع الأجزاء) x °360

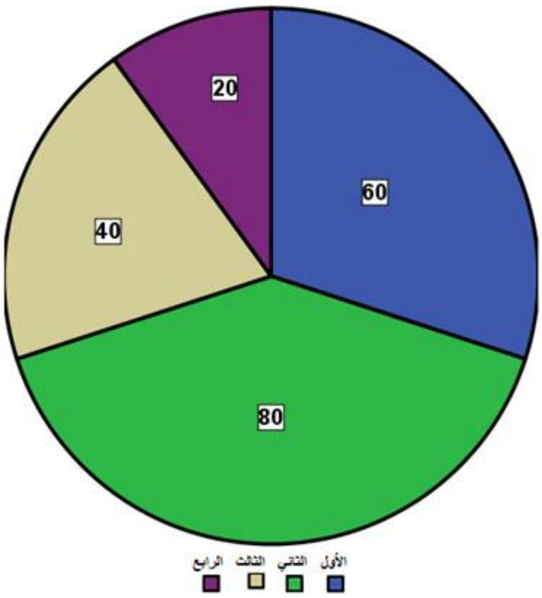
الزاوية المركزية المناظرة للصف الأول = $(60 \div 200) \times 360° = 108°$

الزاوية المركزية المناظرة للصف الثاني = $(80 \div 200) \times 360° = 144°$

الزاوية المركزية المناظرة للصف الثالث = $(40 \div 200) \times 360° = 72°$

الزاوية المركزية المناظرة للصف الرابع = $(20 \div 200) \times 360° = 36°$

**ثانياً**: نرسم دائرة مناسبة ليكن نصف قطرها = 4 سم.



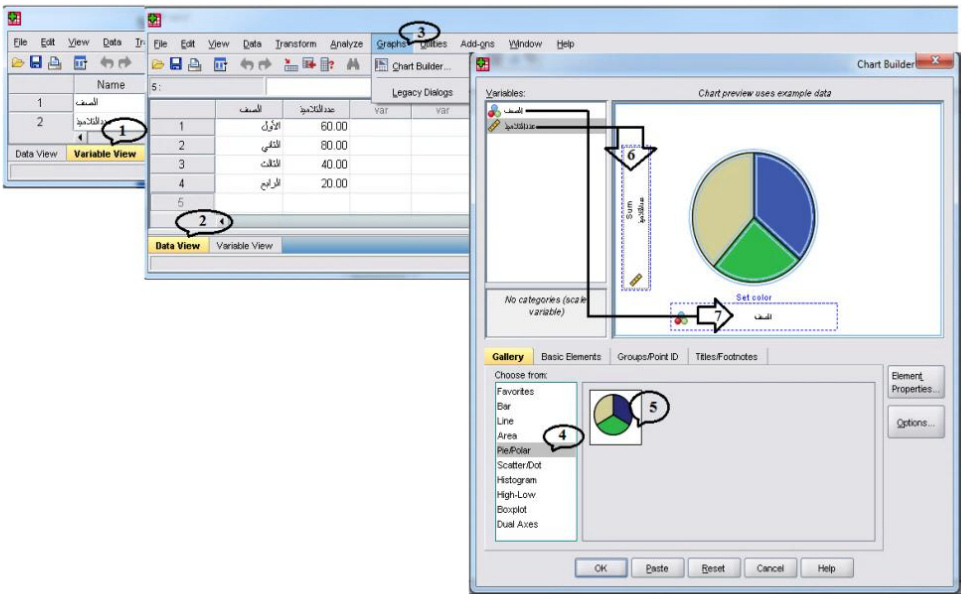

**شكل (14.5 – أ)**: يوضح تمثيل عدد التلاميذ باستخدام الدائرة المجزأة.

**الحل:**

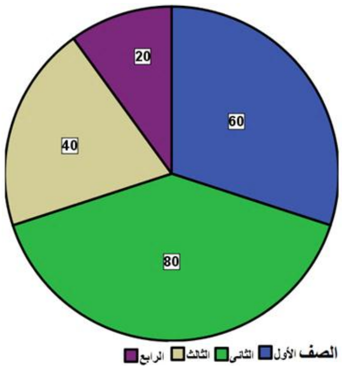
إدخال البيانات إلى برنامج Spss بكتابة المتغيرات، وبالنسبة للمتغير الصف نضغط على values ونسمى الصفوف، ثم إدخال البيانات تحت كل متغير (Data View)، ولرسم الأعمدة البيانية نتتبع الخطوات التالية:

**Graphs ⇒ Chart Builder**

وبالضغط على Chart Builder نحصل على الشكل الحواري ومن قائمة Choose from نختار Pie/Polar ثم الضغط على أيقونة  لبدء رسم الدائرة، وبالضغط على Ok كما في الصندوق الحواري التالي:

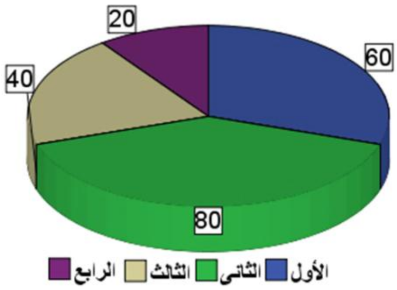


فنحصل على الدائرة البيانية كما في الشكل التالي:

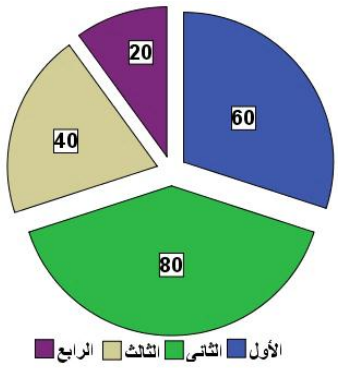



**شكل (14.5 – ب)**: يوضح تمثيل عدد التلاميذ باستخدام الدائرة المجزأة.

ويمكن عرضها باستخدام الأبعاد الثلاثية كما في الشكل التالي:




**شكل (14.5 – ج)**: يوضح تمثيل عدد التلاميذ باستخدام الدائرة ثلاثية الأبعاد.




**شكل (14.5 – د)**: يوضح تمثيل عدد التلاميذ باستخدام الدائرة المتباعدة الأجزاء.

**الخط البياني: Line Graph**

المخطط البياني الخطي Line Chart أو المخطط الخطي هو أحد أنواع المخططات البيانية التي تعرض المعلومات كسلسلة من نقاط البيانات المتصلة بواسطة خطوط مستقيمة. ويتم إنشاء هذا المخطط بتوصيل سلسلة من النقاط التي تمثّل مقاييس مفردة بواسطة قطاعات خطية.

غالباً ما يستخدم مخطط البيانات الخطي لعرض توجّه البيانات خلال فترات زمنية.

ويتم رسم المخطط البياني الخطي برسم خطين متعامدين يسميان "محورين". يسمى المحور الأفقي "محور السينات" ويسمى المحور العمودي "محور الصادات".

غالباً ما يستخدم محور الصادات العمودي لتمثيل المتغير التابع، ومحور السينات الأفقي (الإحداثي السيني) لتمثيل المتغير المستقل.

**مثال**: الجدول التالي يوضح درجات 30 طالباً في مادة علم النفس:

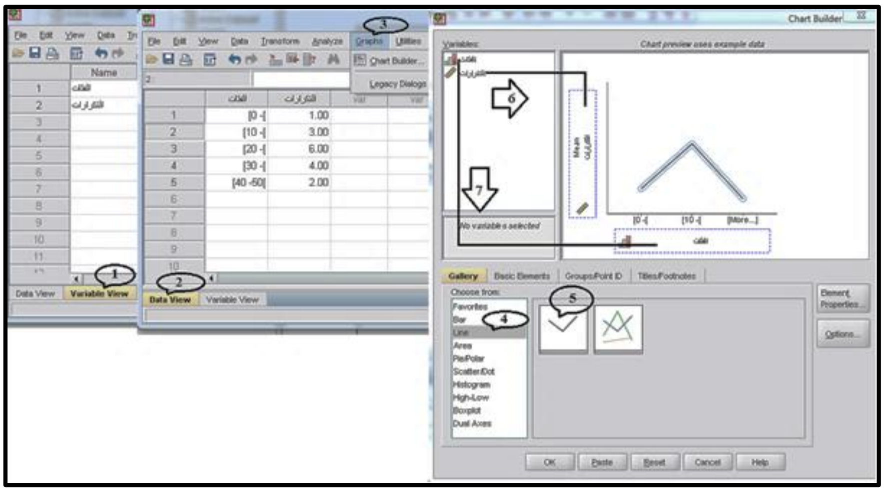
| الفئات | [-0] | [-10] | [-20] | [-30] | [50 -40] |
|---|---|---|---|---|---|
| التكرارات | 1 | 3 | 6 | 4 | 2 |

**والمطلوب**: رسم الخط البياني.

**الحل**: إدخال البيانات إلى برنامج Spss بكتابة المتغيرات، ثم إدخال البيانات تحت كل متغير (Data View)، ولرسم الخط البياني نتتبع الخطوات التالية:

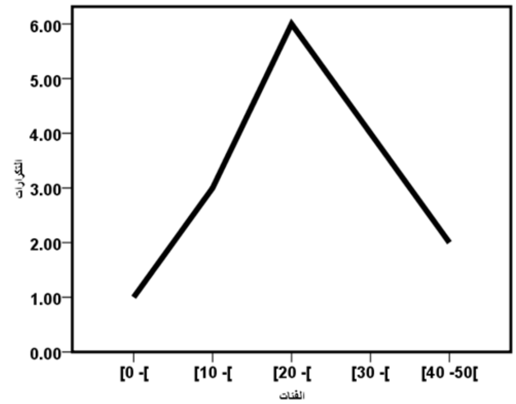
**Graphs ⇨ Chart Builder ⇨ Line ⇨ √ أيقونة**

ثم نبدأ رسم الخط البياني ثم يتم إدراج الفئات والتكرار باستخدام Mouse، وبالضغط على Ok كما في الشكل التالي:



**شكل (15.5)**: يوضح إدراج البيانات لرسم خط بياني.

فنحصل على الخط البياني كما في الشكل التالي:




**شكل (16.5)**: يوضح تمثيل البيانات باستخدام خط بياني.

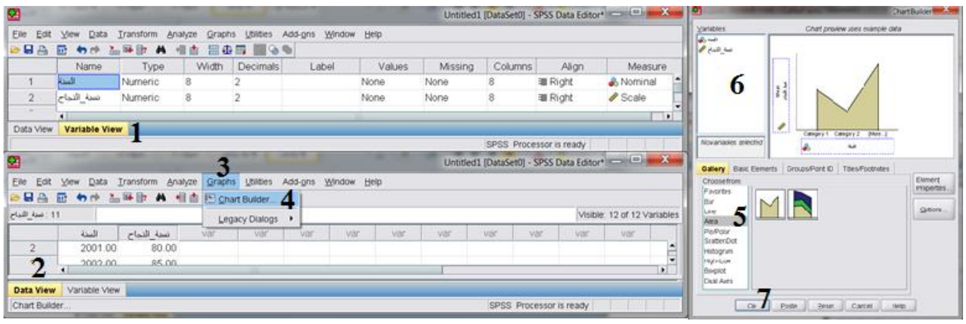
**التمثيل بمخطط المساحة البسيطة نتبع المسار التالي:**

Graph → Chart builder → Choose Form → Area

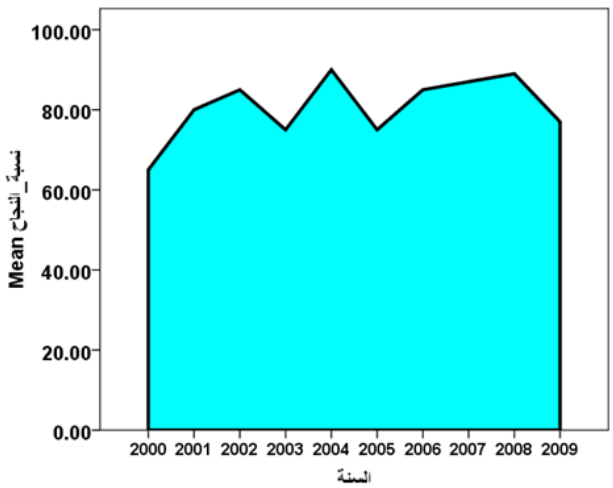
ثم نختار الأيقونة: 

**كمثال:**

| النسبة | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 |
|---|---|---|---|---|---|---|---|---|---|---|
| نسبة النجاح | 65% | 80% | 85% | 75% | 90% | 75% | 85% | 87% | 89% | 77% |




وبالضغط على OK نحصل النتائج التالية:

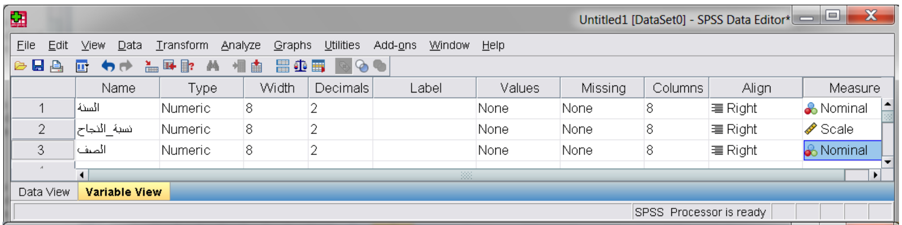



**التمثيل بمخطط المساحة المتراكمة:**

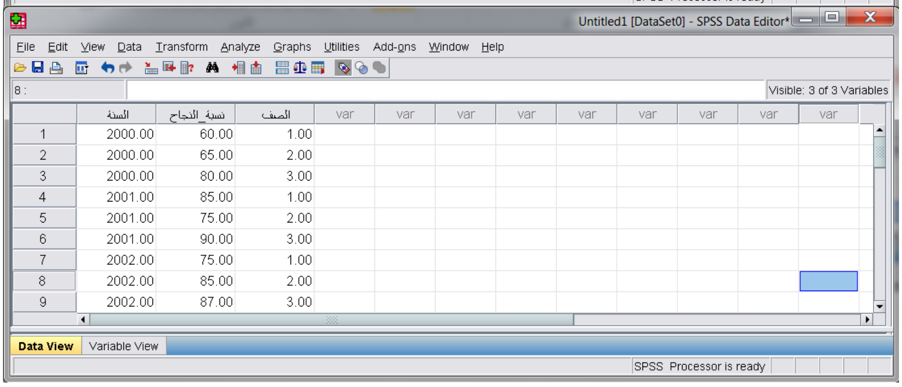
نتبع المسار التالي:

**Graph** ⇨ **Chart builder** ⇨ **Choose Form** ⇨ **Area**

ثم نختار الأيقونة: 

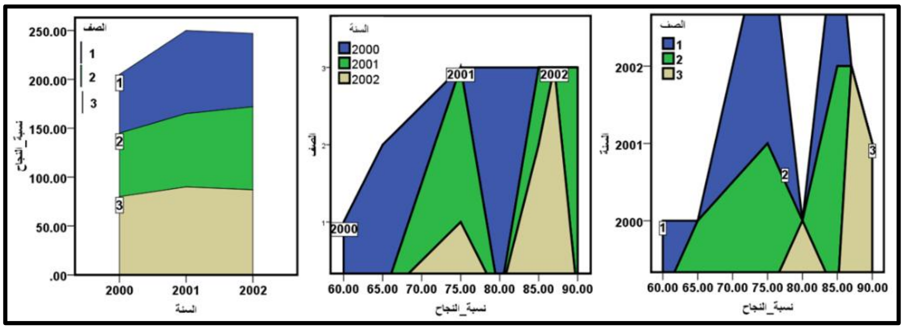


| | Name | Type | Width | Decimals | Label | Values | Missing | Columns | Align | Measure |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | السنة | Numeric | 8 | 2 | | None | None | 8 | Right | Nominal |
| 2 | نسبة_النجاح | Numeric | 8 | 2 | | None | None | 8 | Right | Scale |
| 3 | الصف | Numeric | 8 | 2 | | None | None | 8 | Right | Nominal |



| | السنة | نسبة_النجاح | الصف |
|---|---|---|---|
| 1 | 2000.00 | 60.00 | 1.00 |
| 2 | 2000.00 | 65.00 | 2.00 |
| 3 | 2000.00 | 80.00 | 3.00 |
| 4 | 2001.00 | 85.00 | 1.00 |
| 5 | 2001.00 | 75.00 | 2.00 |
| 6 | 2001.00 | 90.00 | 3.00 |
| 7 | 2002.00 | 75.00 | 1.00 |
| 8 | 2002.00 | 85.00 | 2.00 |
| 9 | 2002.00 | 87.00 | 3.00 |

يمكن الحصول على إحدى الأشكال التالية:

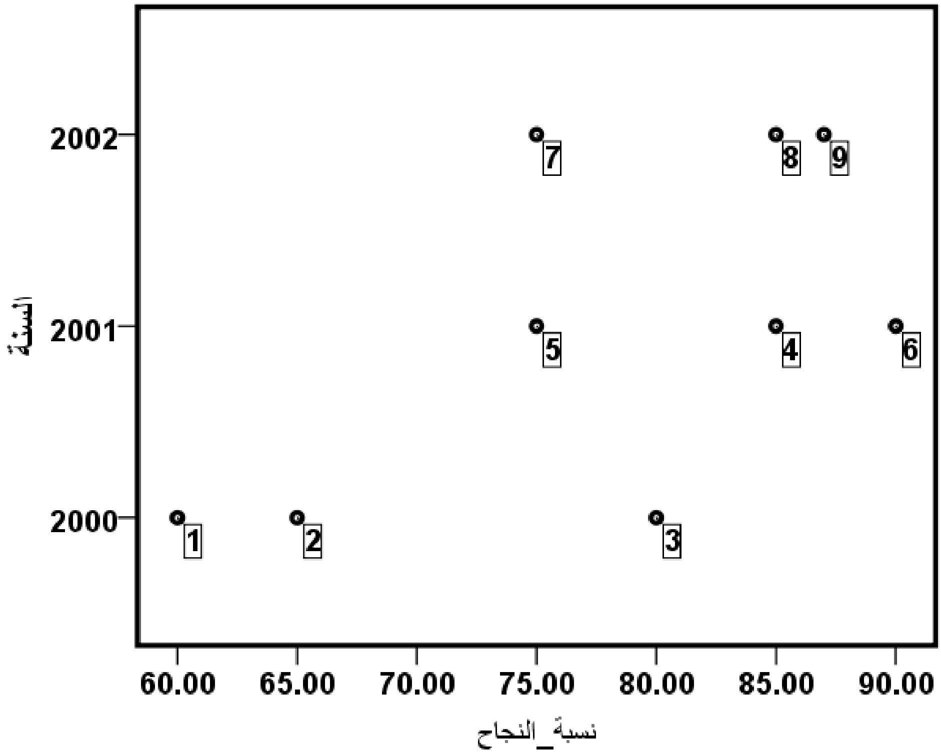



**لوحة الانتشار: Scatter**

نتبع المسار التالي للبيانات التالية:

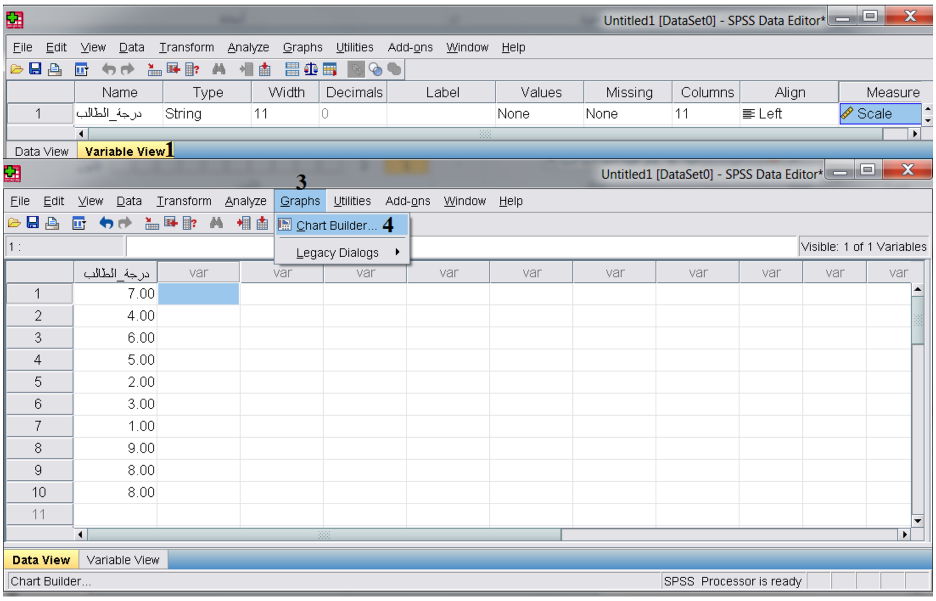
**Graph** ⇨ **Chart builder** ⇨ **Scatter**

ثم نختار الأيقونة: 

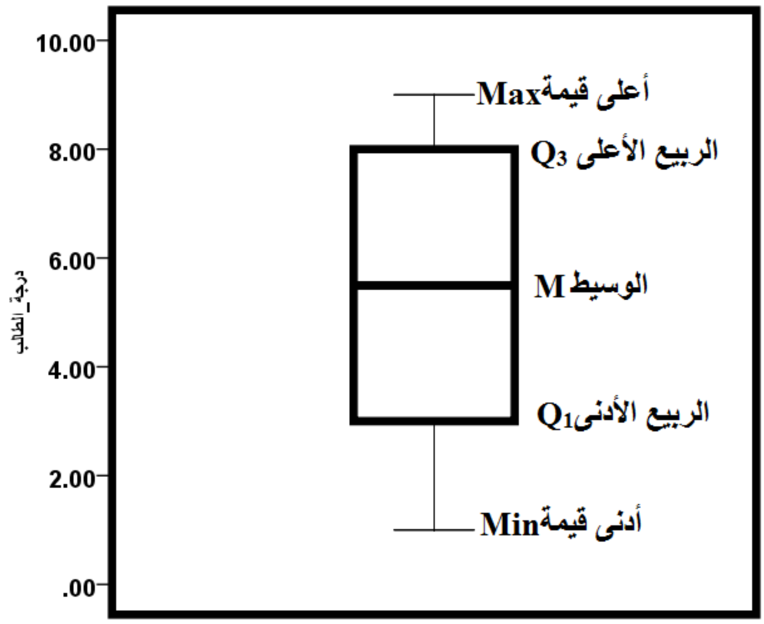



**التمثيل بالبناء الصندوقي: Boxplots**

يستخدم هذا التمثيل لتخليص بيانات كمية مسترة كمثال:




فنحصل على النتائج التالية:



**مثال:**

الجدول التالي يعطي عدد سكان إحدى الدول بالمليون.

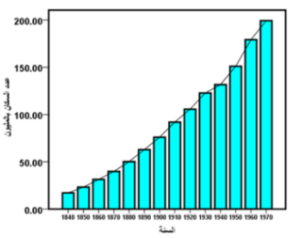
| | | | | | | | |
|---|---|---|---|---|---|---|---|
| السنة | 1840 | 1850 | 1860 | 1870 | 1880 | 1890 | 1900 |
| عدد السكان (بالمليون) | 17.1 | 23.2 | 31.4 | 39.8 | 50.2 | 62.9 | 76.0 |
| السنة | 1910 | 1920 | 1930 | 1940 | 1950 | 1960 | 1970 |
| عدد السكان (بالمليون) | 92.0 | 105.7 | 122.8 | 131.7 | 151.1 | 179.3 | 199.3 |

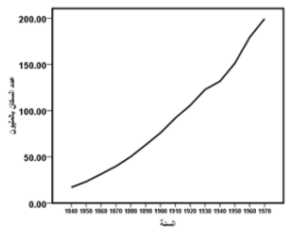
**المطلوب**: تمثيل البيانات بيانياً.

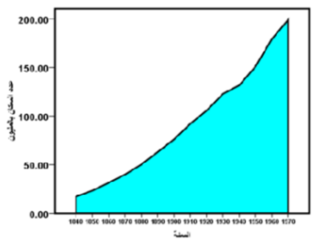
**الحل**: بعد فتح صفحة spss وإدخال البيانات نتبع المسار التالي:

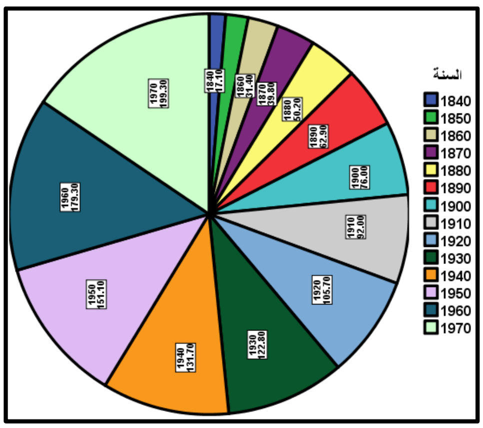
Graphs ⇒ Chart Builder

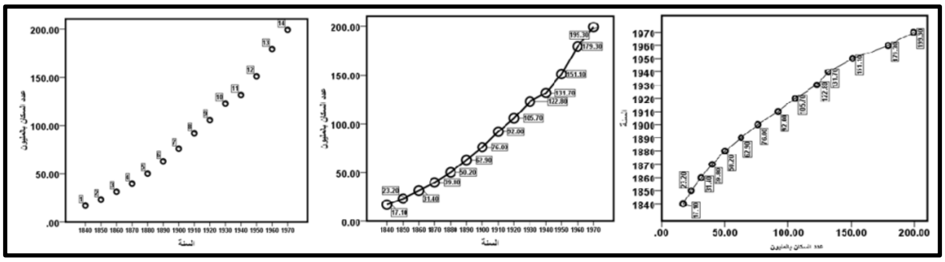
يمكنا تمثيل البيانات السابقة بإحدى الرسوم البيانية.





**(5 – 2) العرض البياني للبيانات المبوبة:**

**المدرج التكراري: Histogram**

هو يمثل العلاقة بين الفئات (المحور الأفقي)، والتكرارات (المحور العمودي).

**أ) المدرج التكراري لجداول منتظمة:**

لرسم المدرج التكراري في حالة الجداول المنتظمة.

**أ.** نرسم أولاً مستقيمين متعامدين نسميهما محورين.

**ب.** نمثل الفئات على المحور الأفقي ويقسم إلى أقسام متساوية بحيث عددها يساوي عدد الفئات.

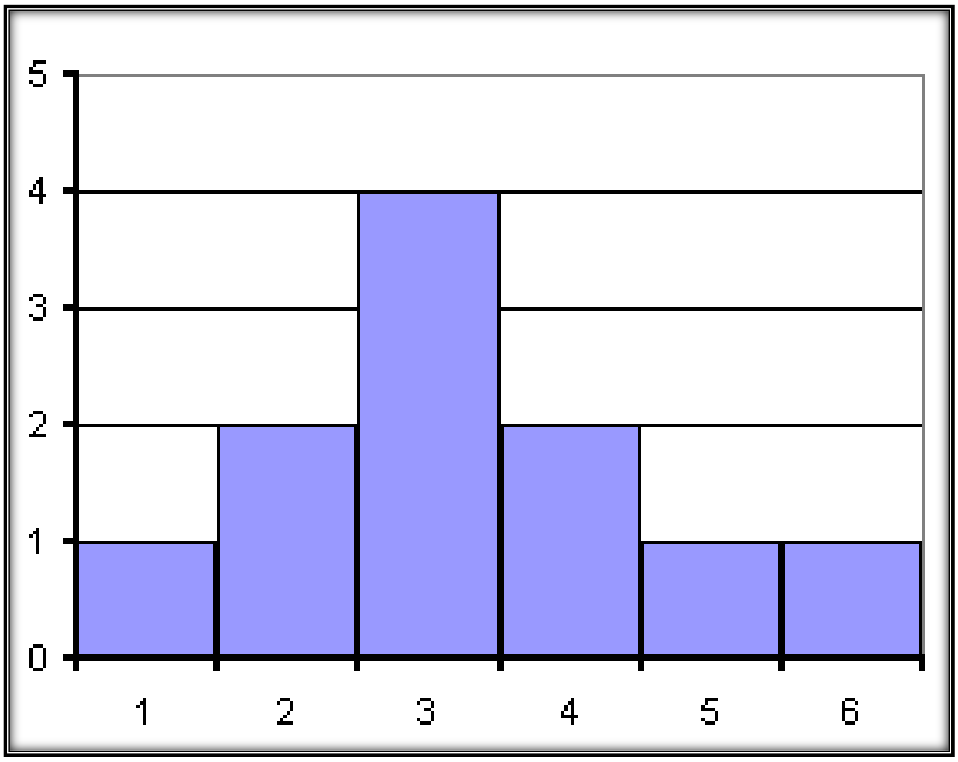
**ج.** نقيم على كل فئة مستطيلاً ارتفاعه يساوي التكرار المقابل لكل فئة بحيث تكون كل المستطيلات متلاصقة، وذلك لاستمرار الظاهرة.

**مثال:** مثل البيانات التالية (6،5،4،4،3،3،3،3،2،2،1) بالمدرج التكراري.

**الحل:** نكون جدول تكراري:

| **الرقم** | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| **التكرار** | 1 | 2 | 4 | 2 | 1 | 1 |

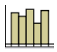
ثم نرسم المدرج التكراري:

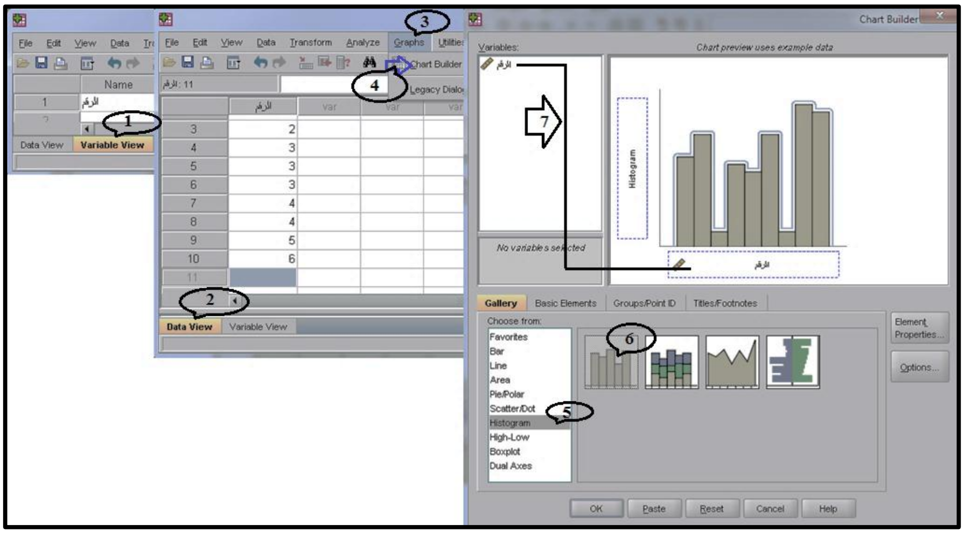



**شكل (17.5)**: يوضح تمثيل الأرقام بجدول تكراري.

وباستخدام الحزمة الإحصائية Spss.

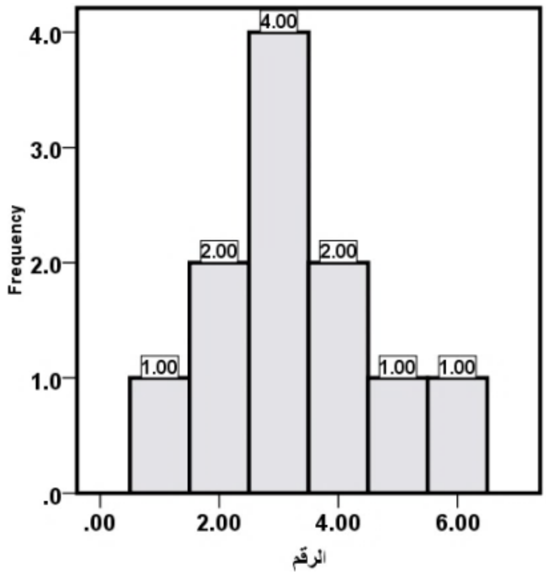
إدخال البيانات إلى برنامج Spss بكتابة المتغيرات، ثم إدخال البيانات تحت كل متغير (Data View)، ولرسم الخط البياني نتتبع الخطوات التالية:

Graphs ثم بالضغط على Chart Builder نحصل على الشكل الحواري ومن قائمة Choose from نختار Line ثم الضغط على أيقونة  لبدء رسم الخط البياني ثم يتم إدراج الفئات والتكرار باستخدام الفارة، وبالضغط على Ok كما في الشكل التالي:



**شكل (18.5)**: يوضح إدراج البيانات لرسم مدرج تكراري.

فنحصل على المدرج التكراري كما في الشكل التالي:




**شكل (19.5)**: يوضح تمثيل البيانات بمدرج تكراري

## (5– 3) التوزيعات التكرارية

| Class الفئة | Frequency (f) التكرار |
|---|---|
| 90 − 99 | 4 |
| 80 − 89 | 6 |
| 70 − 79 | 4 |
| 60 − 69 | 3 |
| 50 − 59 | 2 |
| 40 − 49 | 1 |
| Σ | 20 |

**عرض الفئة (فترة الفئة): Interval Width(Class Interval)**

هو الفرق بين (الحد الأعلى للفئة – الحد الأدنى للفئة).

Class Width – The difference between the upper (or lower) class limits of consecutive classes.

Class Midpoint = upper – lower

| الفئة **Class** | التكرار **Frequency (f)** | التكرار المتجمع الصاعد **Cumulative Frequency** |
|---|---|---|
| 90 - 99 | 4 | 4 |
| 80 – 89 | 6 | 10 |
| 70 – 79 | 4 | 14 |
| 60 – 69 | 3 | 17 |
| 50 – 59 | 2 | 19 |
| 40 - 49 | 1 | 20 |

**مثال:**

قام باحث بدراسة قدرة خريج الجامعة على حل المسائل الرياضية وأعطى كل واحد 10 مسائل والجدول التالي تقييم الطلاب.

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 7 | 4 | 7 | 3 | 8 | 5 | 6 | 3 | 8 | 5 |
| 6 | 10 | 7 | 8 | 5 | 6 | 4 | 7 | 8 | 6 |
| 10 | 7 | 5 | 8 | 6 | 8 | 1 | 6 | 3 | 5 |
| 3 | 6 | 9 | 9 | 5 | 6 | 9 | 6 | 9 | 8 |
| 5 | 10 | 4 | 7 | 6 | 8 | 4 | 8 | 5 | 8 |
| 10 | 6 | 9 | 6 | 7 | 5 | 8 | 6 | 7 | 8 |
| 6 | 9 | 7 | 4 | 5 | 6 | 8 | 6 | 8 | 4 |
| 5 | 4 | 7 | 6 | 6 | 9 | 5 | 5 | 4 | 7 |
| 6 | 6 | 7 | 6 | 10 | 9 | 7 | 5 | 3 | 10 |
| 8 | 6 | 4 | 7 | 5 | 8 | 7 | 8 | 9 | 6 |

**المطلوب:** تكوين جدول تكراري وتمثيل البيانات بالمدرج التكراري.

| Score(X) | Frequency(f) |
|---|---|
| 10 | 6 |
| 9 | 9 |
| 8 | 17 |
| 7 | 15 |
| 6 | 23 |
| 5 | 15 |
| 4 | 9 |
| 3 | 5 |
| 2 | 0 |
| 1 | 1 |
| $\sum$ | 100 |

**الحل:** باستخدام spss

**Analyze** ⇨ **Descriptive Statistics** ⇨ **Frequencies**

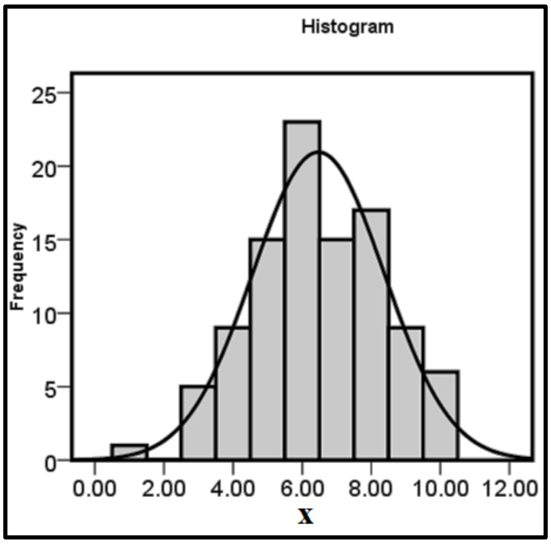
فنحصل على الوصف الإحصائي للمتغيرات كما في الجدول التالي:

| x | | | | | |
|---|---|---|---|---|---|
| | | Frequency<br>عدد التكرارات | Percent<br>النسبة | Valid Percent<br>النسبة المئوية للتكرارات | Cumulative Percent<br>النسبة التراكمية |
| Valid | 1 | 1 | 1.0 | 1.0 | 1.0 |
| | 3 | 5 | 5.0 | 5.0 | 6.0 |
| | 4 | 9 | 9.0 | 9.0 | 15.0 |
| | 5 | 15 | 15.0 | 15.0 | 30.0 |
| | 6 | 23 | 23.0 | 23.0 | 53.0 |
| | 7 | 15 | 15.0 | 15.0 | 68.0 |
| | 8 | 17 | 17.0 | 17.0 | 85.0 |
| | 9 | 9 | 9.0 | 9.0 | 94.0 |
| | 10 | 6 | 6.0 | 6.0 | 100.0 |
| | Total | 100 | 100.0 | 100.0 | |

ولرسم المدرج التكرارى نتبع المسار التالى:

Charts $\Longrightarrow$ Histograms

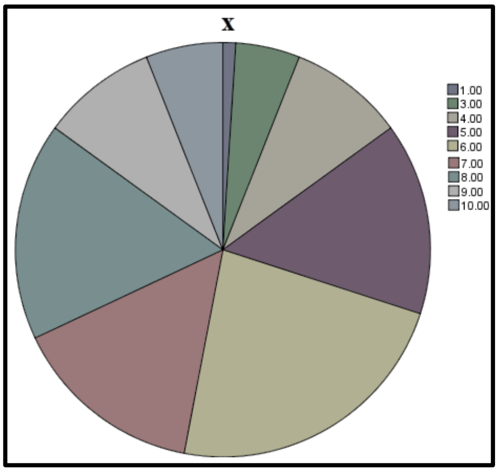
فنحصل على الرسم التالي.




**شكل (20.5)**: يوضح البيانات بالمدرج والمنحنى التكراري

ولتمثيل البيانات بالدائرة المجزأة نتبع المسار التالى:

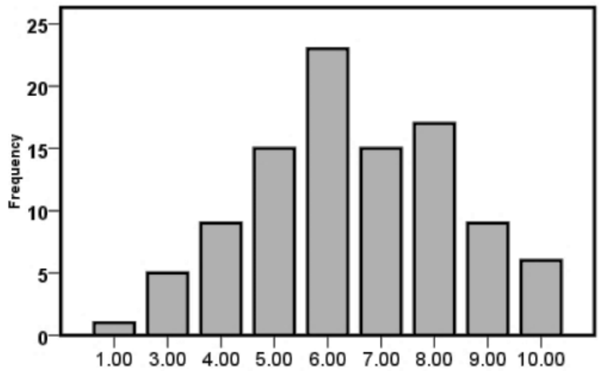
Charts ⟹ Pie charts




**شكل (21.5)**: يوضح البيانات بالدائرة المجزأة.

ولتمثيل البيانات بالأعمدة البسيطة نتبع المسار التالى:

Charts ⟹ Bar charts




**شكل (22.5)**: يوضح البيانات بالأعمدة البسيطة.

**مثال 2**: فيما يلي درجات 40 طالباً في الصف الأول الثانوي في مادة الرياضيات:

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 14 | 12 | 11 | 10 | 6 | 5 | 4 | 3 | 2 | 1 |
| 12 | 19 | 18 | 13 | 9 | 7 | 4 | 9 | 8 | 7 |
| 39 | 38 | 35 | 33 | 22 | 21 | 20 | 17 | 16 | 15 |
| 31 | 37 | 36 | 32 | 26 | 24 | 22 | 11 | 18 | 14 |

**المطلوب:** كون الجدول التكراري المناسب: Construct a frequency table

**الحل**:

| Class(الفئات) | Tally العلامات | Frequency (f)(التكرارات ) |
|---|---|---|
| 0- | 𝍸 𝍸 // | 12 |
| 10- | 𝍸 𝍸 //// | 14 |
| 20- | 𝍸 / | 6 |
| 30 - 40 | 𝍸 | 8 |
| مجموع ( مجـ ) Σ | | 40 |

**مثال 3**:

لدينا درجات 25 طالب في إحدى المواد في الجدول التالي:

| | | | | |
|---|---|---|---|---|
| 19.55 | 27.57 | 26.32 | 25.19 | 22.55 |
| 20.75 | 27.79 | 26.33 | 25.7 | 23.75 |
| 21.28 | 28.17 | 27.01 | 25.91 | 24.03 |
| 22.02 | 30.46 | 27.13 | 26.13 | 24.24 |
| 22.51 | 30.91 | 27.55 | 26.13 | 25.17 |

**المطلوب**: تكوين جدول تكراري مناسب construct a frequency table

**الحل**:

عرض الفئة = (أكبر قيمة − أصغر قيمة) ÷ عدد الفئات المقترحة

$= (30.91 - 19.55) \div 2.272 = 5$

الجدول التكراري:

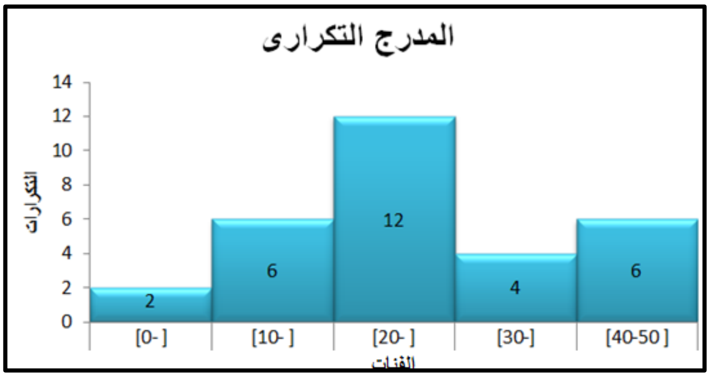
| الفئة Class Limits | التكرار Frequency | مركز الفئة Midpoint |
|---|---|---|
| 19.55 – 21.82 | 3 | 20.685 |
| 21.83 – 24.10 | 5 | 22.965 |
| 24.11 – 26.38 | 9 | 25.245 |
| 26.39 – 28.66 | 6 | 27.525 |
| 28.67 – 30.94 | 2 | 28.665 - 30.945 |
| $\sum$ | 25 | |

**مثال2**: الجدول التالي يوضح درجات 30 طالباً في مادة علم النفس

| الفئة Class | التكرار Frequency ($f$) |
|---|---|
| [0-] | 2 |
| [10-] | 6 |
| [20-] | 12 |
| [30-] | 4 |
| [40-50] | 6 |

**والمطلوب**: رسم المدرج التكراري.

**الحل**:




**شكل (23.5)**: يوضح المدرج التكراري لجدول تكراري منتظم.

### ب) المدرج التكراري لجداول غير منتظمة:

إذا كان الجدول التكراري غير منتظم فيجب تعديل التكرارات للحصول على تكرار معدل.

**التكرار المعدل للفئة = التكرار الأصلي للفئة ÷ عرض الفئة**

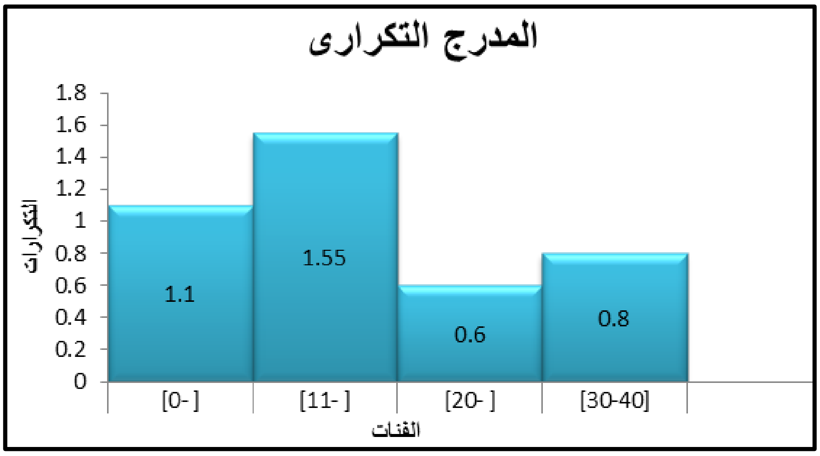
والمقارنة تعتمد على ارتفاعات المستطيلات بدلاً من مقارنة المساحات.

**مثال**: الجدول التالي يوضح درجات 40 طالباً في مادة علم النفس

| الفئات | 0 - | 11 - | 20 - | 30 - 40 |
|---|---|---|---|---|
| التكرارات | 12 | 14 | 6 | 8 |

**والمطلوب**: رسم المدرج التكراري:

| الفئات | 0- | 11- | 20- | 30- 40 |
|---|---|---|---|---|
| التكرار الأصلي | 12 | 14 | 6 | 8 |
| طول الفئة | 11 | 9 | 10 | 10 |
| التكرار المعدل | 1.1=11÷12 | 1.55=9÷14 | 0.6=10÷6 | 0.8=10÷8 |




**شكل (24.5)**: يوضح المدرج التكراري لجدول تكراري غير منتظم.

**المضلع التكراري:Frequency Polygon**

هو الخط المنكسر الواصل بين مراكز الفئات (المحور الأفقي) عند قمة المستطيلات للمدرج التكراري، والتكرارات (المحور العمودي)، ويمكن رسمه بطريقتين:

**الطريقة الأولى:**

**1**. تحديد نقطة التكرار عند مركز الفئة، بشرط مراعاة نوع التوزيع.

**2**. التوصيل بين النقاط بمستقيمات أي بخط منكسر.

**الطريقة الثانية:**

**1**. نرسم المدرج التكراري، مع مراعاة التوزيع المنتظم وغير المنتظم.

**2**. تحديد مراكز الفئات عند قمة المستطيلات.

**3**. نصل بين مراكز الفئات بعضها بالبعض محافظين على شيئين وهو أن المساحة تحت المضلع التكراري يجب أن تساوي المساحة تحت المدرج التكراري وبذلك يجب أن يكون المضلع التكراري متصلاً بالمحور الأفقي عند طرفيه ويتم ذلك كالآتي:

بأن نحدد مركز الفئة السابقة للفئة الأولى ونحدد مركز الفئة اللاحقة للفئة الأخيرة على المحور الأفقي باعتبار أن طول الفئة السابقة يساوي طول الأولى واللاحقة يساوي الأخيرة (جدول منتظم).

## أولاً: الجدول التكراري البسيط.

**مثال:** مثل بيانات الجدول التالي بالمضلع التكراري.

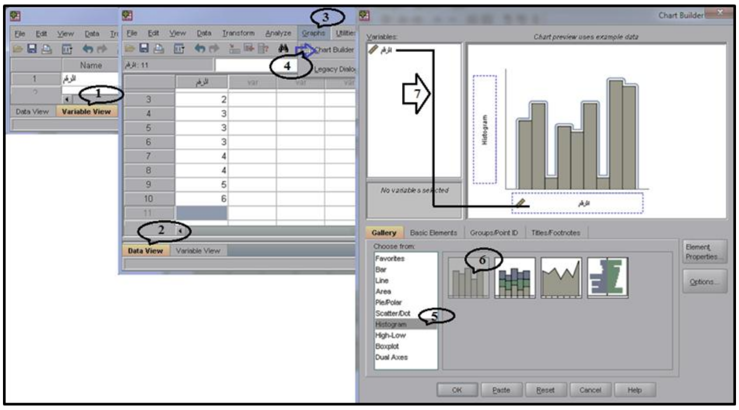
| **الرقم** | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| **التكرار** | 1 | 2 | 4 | 2 | 1 | 1 |

**الحل:** باستخدام الحزمة الإحصائية Spss.

إدخال البيانات إلى برنامج Spss بكتابة المتغيرات، ثم إدخال البيانات تحت كل متغير (Data View)، ولرسم الخط البياني نتتبع الخطوات التالية:

**Graphs ⇨ Chart Builder ⇨ Histograms ⇨ أيقونة**

لبدء رسم الخط البياني ثم يتم إدراج الفئات والتكرار باستخدام الفارة، وبالضغط على Ok كما في الشكل التالي:

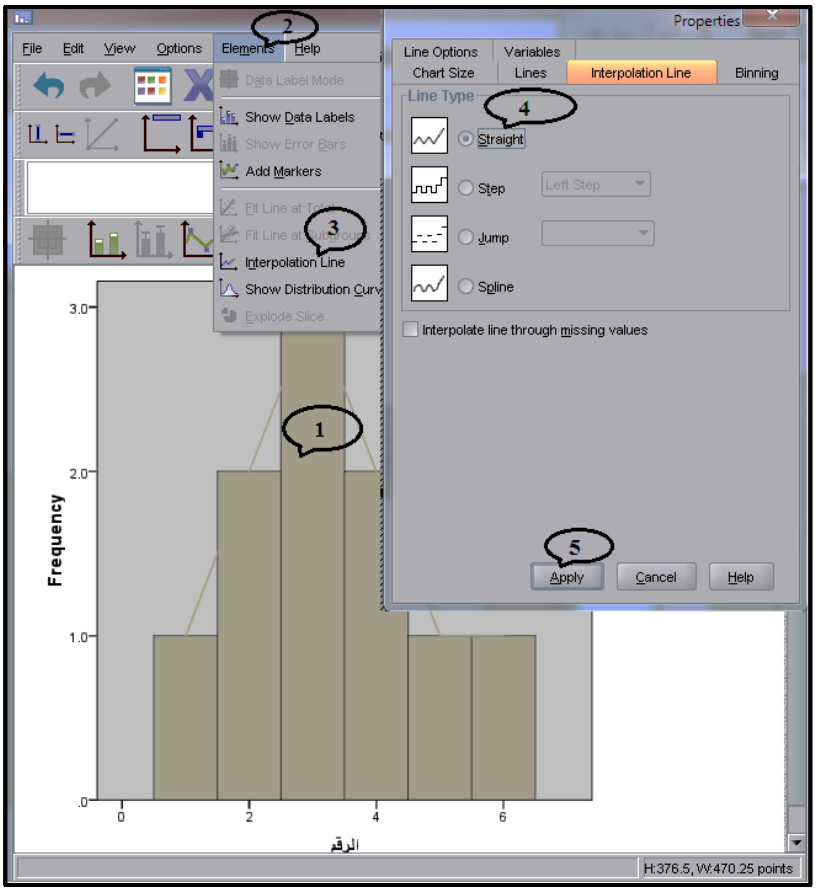


**شكل (25.5)**: يوضح إدراج البيانات لرسم مضلع تكراري.

فنحصل على المدرج التكراري ثم بالضغط على الصورة مرتين نحصل على صندوق حواري باتباع المسار التالي لنتمكن من رسم المضلع التكراري

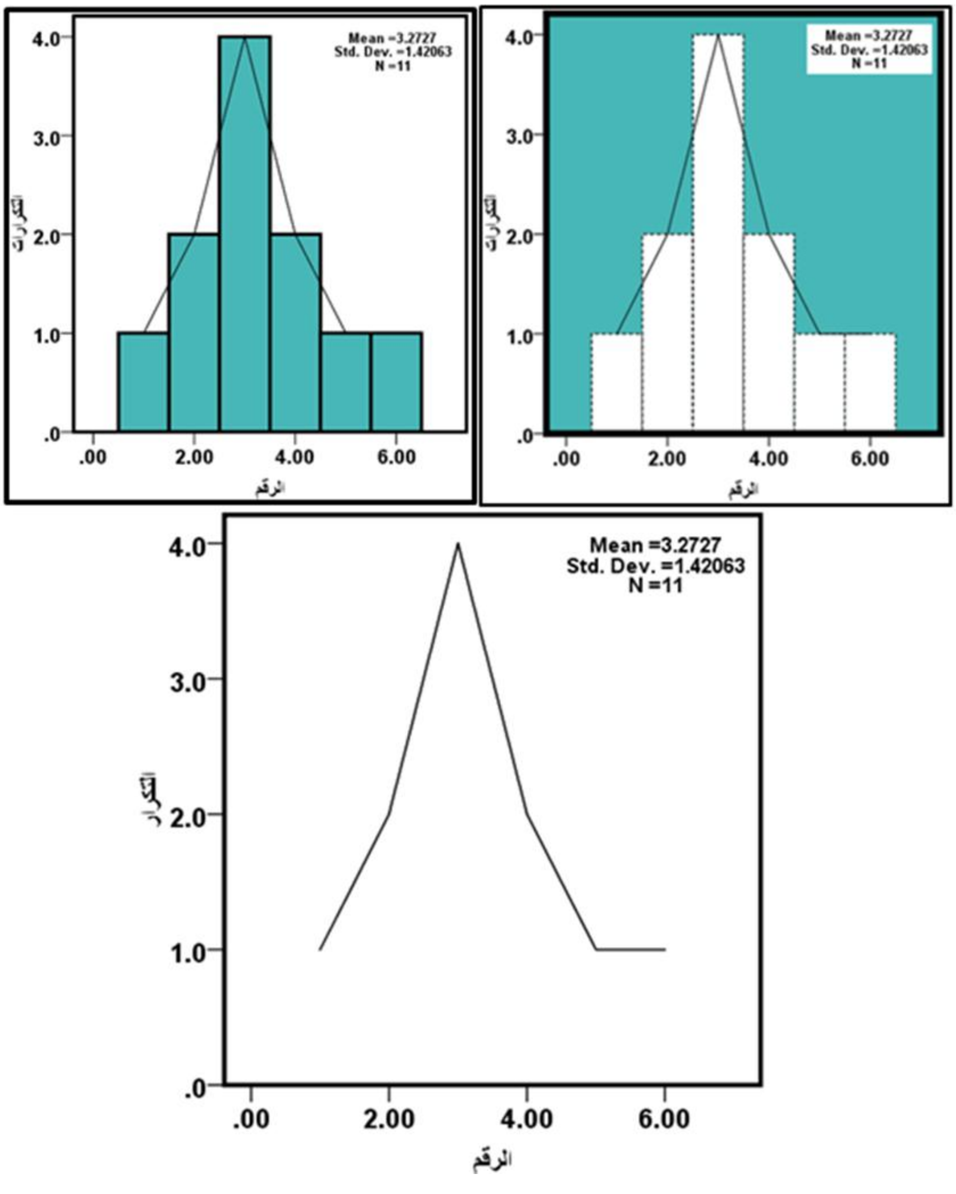
**Elemants ⇨ Interpolation Lin ⇨ Straight**

ويمكنا تعديل الشكل النهائي بتحديد الجزء المراد إيضاحه بالضغط عليه مرتين متتاليتين على الجانب الأيسر ل Mouse.




**شكل (26.5)**: يوضح إدخال تعديلات على المدرج التكراري لرسم المضلع التكراري.

فنحصل على الشكل التالي:




**شكل (27.5)**: يوضح المضلع التكراري.

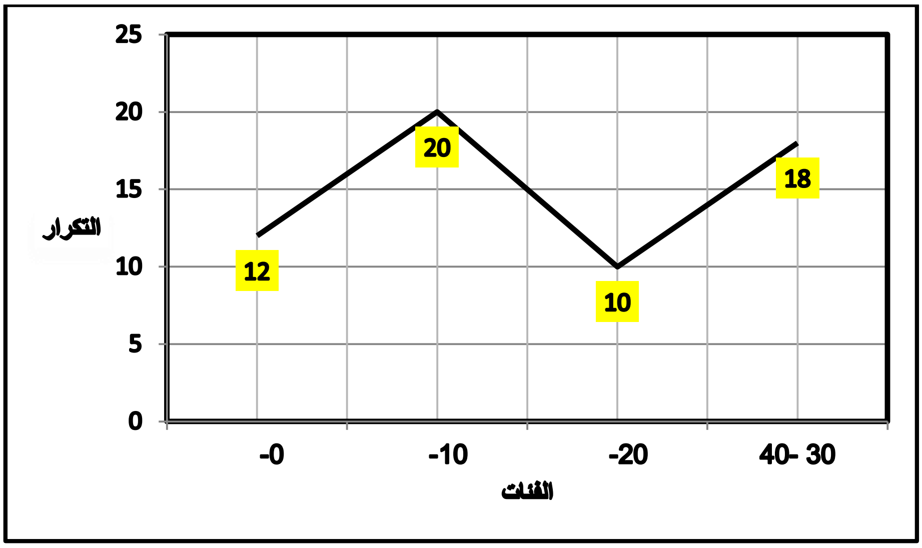
**ثانياً**: **جدول تكراري منتظم.**

**مثال**: الجدول التالي يوضح درجات 60 طالباً في مادة علم النفس

| **الفئات** | 0- | 10- | 20- | 30-40 | **المجموع** |
| --- | --- | --- | --- | --- | --- |
| **التكرارات** | 12 | 20 | 10 | 18 | 60 |

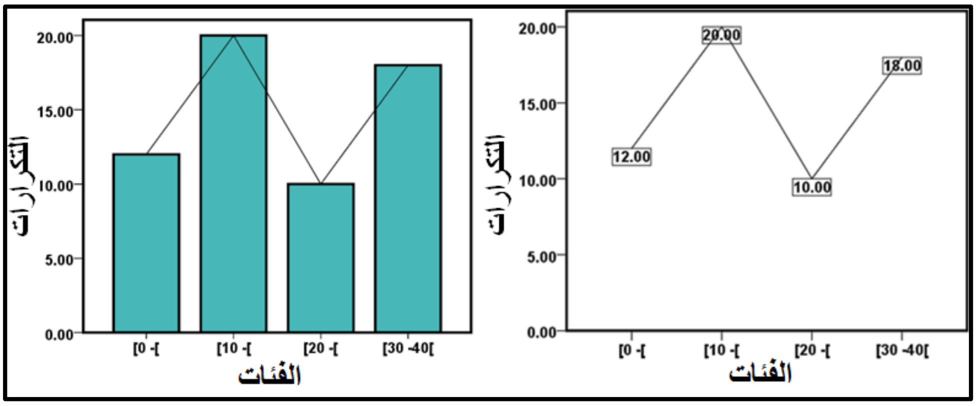
**والمطلوب**: رسم المضلع التكراري:

**الحل**:




**شكل (28.5 – أ)**: يوضح المضلع التكراري.

وباستخدام الحزمة الإحصائية Spss.




**شكل (28.5 – ب)**: يوضح المضلع التكراري باستخدام الحزمة الإحصائية Spss.

**ثالثاً: جدول غير منتظم:**

نحدد مركز الفئة السابقة واللاحقة للفئة الأولى والأخيرة على المحور الأفقي ولكي تكون المساحة تحت المضلع التكراري متساوية مع المساحة تحت المدرج فيجب أن يكون طول الفئة السابقة للفئة الأولى متساوياً مع طول الفئة الأولى (جدول غير منتظم) وطول الفئة اللاحقة للفئة الأخيرة مساوٍ لطول الفئة الأخيرة وذلك كما في المثال التالي:

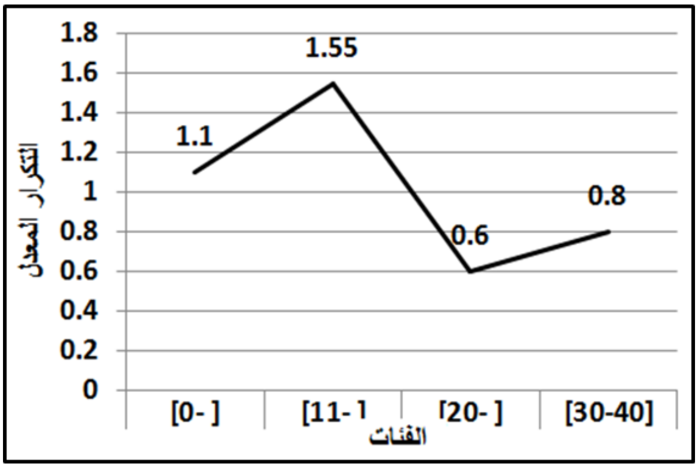
**مثال**: الجدول التالي يوضح درجات 40 طالباً في مادة علم النفس

| الفئات | 0 – | 11 – | 20 – | 30 – 40 | المجموع |
|---|---|---|---|---|---|
| التكرارات | 12 | 14 | 6 | 8 | 40 |

**والمطلوب**: رسم المضلع التكراري:

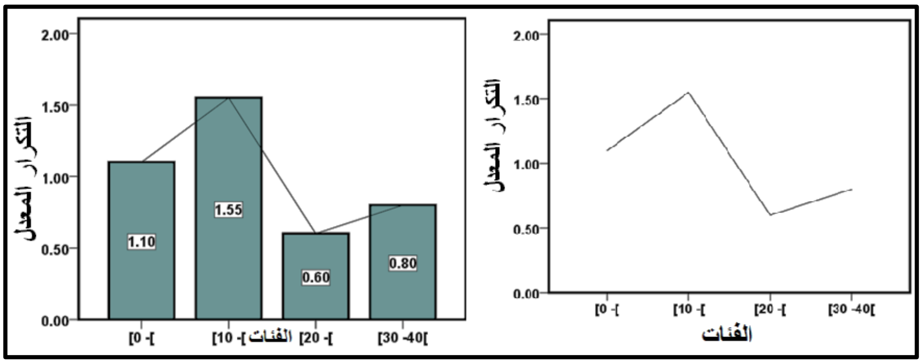
**الحل**: الجدول غير منتظم لابد من إيجاد التكرار المعدل.

| الفئات | 0- | 11- | 20- | 30- 40 |
|---|---|---|---|---|
| التكرار الأصلي | 12 | 14 | 6 | 8 |
| طول الفئة | 11 | 9 | 10 | 10 |
| التكرار المعدل | 1.1=11÷12 | 1.55=9÷14 | 0.6=10÷6 | 0.8=10÷8 |




**شكل (29.5)**: يوضح المضلع التكراري لجدول غير منتظم.

وباستخدام الحزمة الإحصائية Spss.




**شكل (30.5)**: يوضح المضلع التكراري باستخدام الحزمة الإحصائية Spss لجدول غير منتظم.

### 4. المنحنى التكراري Frequency Curve

هو الخط الممهد الذي يصل بين مراكز الفئات عند قمة المستطيلات، على أن يتم التوصيل بدون مسطرة ويمكن استخدام أدوات هندسية أخرى أو باليد.

أو إدخال البيانات إلى برنامج Spss بكتابة المتغيرات، ثم إدخال البيانات تحت كل متغير (Data View)، ولرسم الخط البياني نتتبع الخطوات التالية:

**Graphs ⇨ Chart Builder ⇨ Histograms ⇨ أيقونة**

لبدء رسم الخط البياني ثم يتم إدراج الفئات والتكرار باستخدام Mouse، وبالضغط على Ok وبعد الحصول على المدرج التكراري كما في المثال السابق ثم بالضغط على الصورة مرتين نحصل على صندوق حواري باتباع المسار التالي لنتمكن من رسم المنحنى التكراري.

**Elemants ⇨ Interpolation Lin ⇨ Spline**

ويمكننا تعديل الشكل النهائي بتحديد الجزء المراد إيضاحه بالضغط عليه مرتين متتاليتين على الجانب الأيسر لل Mouse.

**مثال**: الجدول التالي يوضح درجات 60 طالباً في مادة علم النفس:

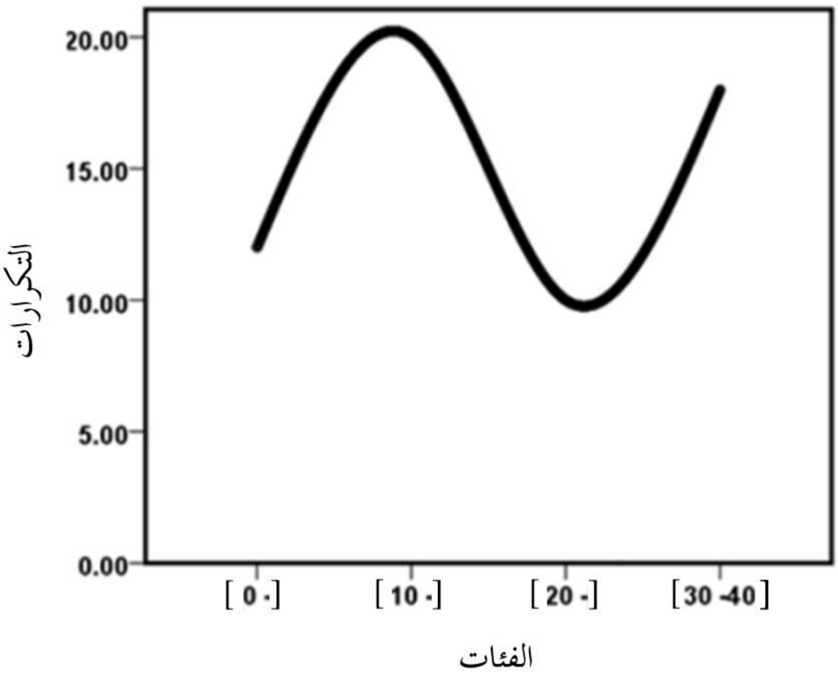
| الفئات | 0 – | 10 – | 20 – | 30 – 40 | المجموع |
|---|---|---|---|---|---|
| التكرارات | 12 | 20 | 10 | 18 | 60 |

**والمطلوب**: رسم المنحنى التكراري والذي سبق رسم المدرج والمضلع التكراري:

**الحل**: باستخدام الحزمة الإحصائية Spss.

**Graphs ⇨ Chart Builder ⇨ Histograms ⇨ أيقونة**

**Elemants ⇨ Interpolation Lin ⇨ Spline**



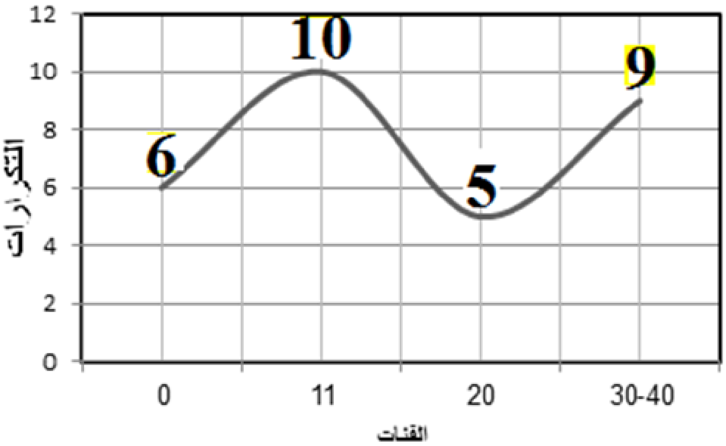

**شكل (31.5)**: يوضح المنحنى التكراري.

**مثال**: الجدول التالي يوضح درجات 30 طالباً في مادة علم النفس

| **الفئات** | 0 - | 11 - | 20 - | 30 - 40 | المجموع |
|---|---|---|---|---|---|
| **التكرارات** | 6 | 10 | 5 | 9 | 30 |

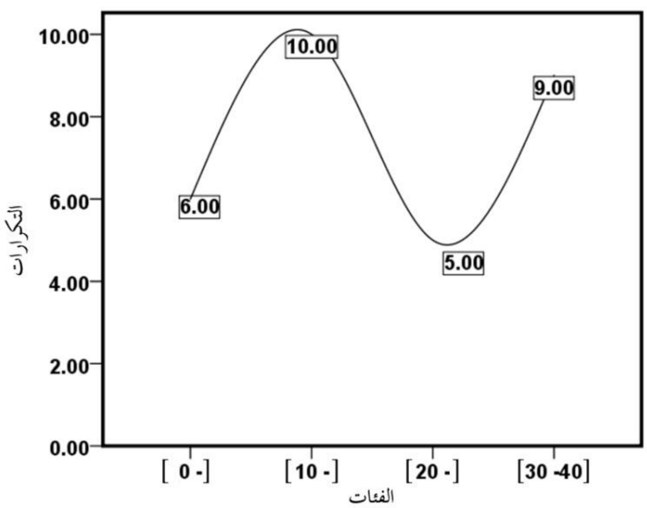
**والمطلوب**: رسم المنحنى التكراري:

**الحل**:




**شكل (32.5)**: يوضح المنحنى التكراري لدرجات 30 طالباً في مادة علم النفس.

وباستخدام الحزمة الإحصائية Spss.




**شكل (33.5)**: يوضح المنحنى التكراري باستخدام spss لدرجات 30 طالباً في مادة علم النفس

وتوجد أشكال مختلفة للمنحنى التكراري حيث أنه لا يأخذ شكلاً ثابتاً إلا أنه يختلف من مجموعة إلى أخرى ولكن هناك بعض المنحنيات الشائعة سنحاول الآن عرضها على سبيل المثال وليس الحصر، ومن خلالها سنتعرف على بعض التعريفات الهامة في علم الإحصاء الوصفي ويمكن تقسيم هذه المنحنيات إلى المجموعات الآتية:

**المنحنى التكراري المتجمع Cumulative Frequency Curve**

يمكن رسم المنحنى المتجمع الصاعد والمنحنى المتجمع الهابط بيانياً.

**مثال:**

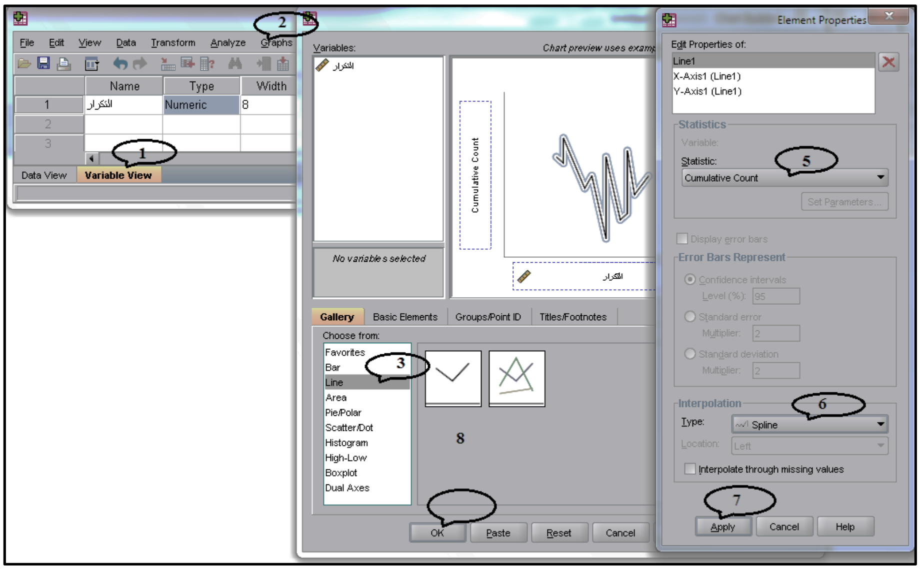
مثل البيانات التالية (6،5،4،4،3،3،3،3،2،2،1) بالمضلع التكراري.

**الحل:**

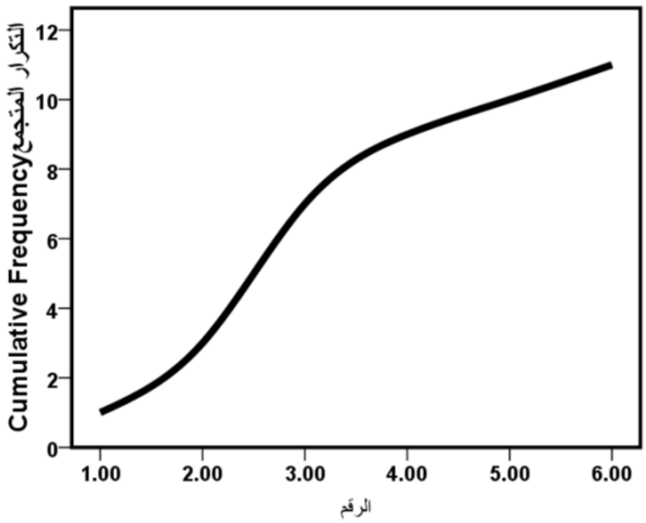
إدخال البيانات إلى برنامج Spss بكتابة المتغيرات، ثم إدخال البيانات تحت كل متغير (Data View)، ولرسم الخط البياني نتتبع الخطوات التالية:

**Graphs ⇨ Chart Builder ⇨ Line ⇨ ✓ أيقونة**

ثم نبدأ رسم الخط البياني ثم يتم إدراج الفئات والتكرار باستخدام Mouse، وبالضغط على Ok كما في الشكل التالي:




**شكل (34.5)**: يوضح إدراج البيانات لرسم المنحنى التكراري المتجمع.




**شكل (35.5)**: يوضح إدراج البيانات لرسم المنحنى التكراري المتجمع.

**مثال**: البيانات الآتية تمثل التوزيع التكراري لدرجات مائة طالباً في مادة الإحصاء.

| **الفئات** | 30 - | 40 - | 50 - | 60 - | 70 - | 80 - | 90 - 100 | **المجموع** |
|---|---|---|---|---|---|---|---|---|
| **التكرارات** | 2 | 3 | 11 | 20 | 32 | 25 | 7 | 100 |

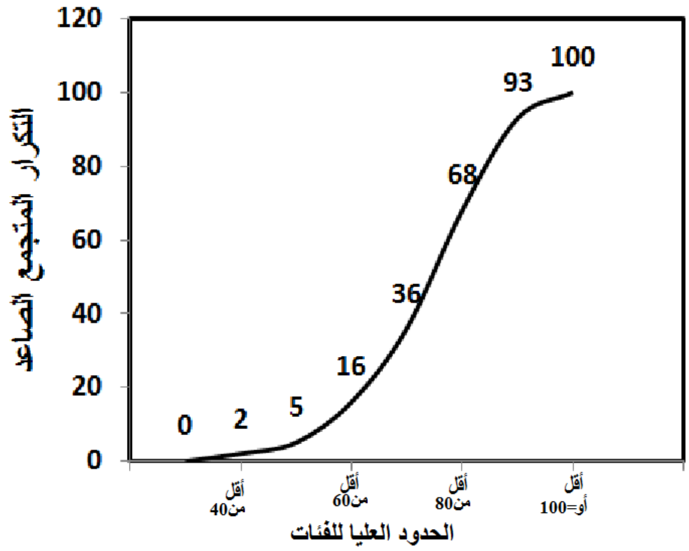
**والمطلوب**: رسم المنحنى المتجمع الصاعد والهابط.

**الحل**:

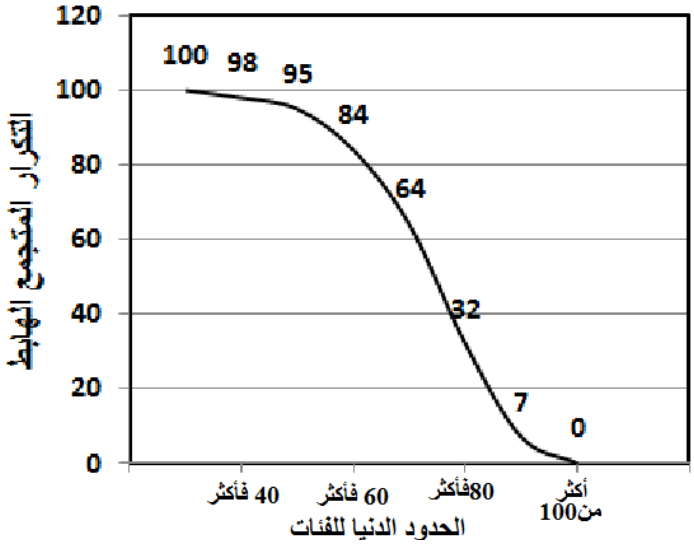
للحصول على المنحنى المتجمع الصاعد والهابط لابد أن نحصل أولاً على جدول التوزيع المتجمع للظاهرة.

التوزيع المتجمع لدرجات مائة طالب في الإحصاء:

| **التكرار المتجمع الصاعد** | | **التكرار المتجمع الهابط** | |
|---|---|---|---|
| الفئة الأولى تساوي الصفر<br>الفئة الأخيرة تساوي مجموع التكرارات | | الفئة الأولى تساوي مجموع التكرارات<br>الفئة الأخيرة تساوي الصفر | |
| **الحدود العليا للفئات** | **التكرار المتجمع الصاعد** | **الحدود الدنيا للفئات** | **التكرار المتجمع الهابط** |
| أقل من 30 | صفر | 30 فأكثر | 100 |
| أقل من 40 | 2 | 40 فأكثر | 98 |
| أقل من 50 | 5 | 50 فأكثر | 95 |
| أقل من 60 | 16 | 60 فأكثر | 84 |
| أقل من 70 | 36 | 70 فأكثر | 64 |
| أقل من 80 | 68 | 80 فأكثر | 32 |
| أقل من 90 | 93 | 90 فأكثر | 7 |
| أقل من 70 أو يساوي 100 | 100 | أكثر من 100 | صفر |



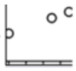

**شكل (36.5 – أ)**: يوضح المنحنى المتجمع الصاعد لدرجات مائة طالب في الإحصاء




**شكل (36.5 – ب)**: يوضح المنحنى المتجمع الهابط لدرجات مائة طالب في الإحصاء

## شكل الانتشار: Scatter Plot

إدخال البيانات إلى برنامج Spss بكتابة المتغيرات، ثم إدخال البيانات تحت كل متغير (Data View)، ولرسم الخط البياني نتتبع الخطوات التالية:

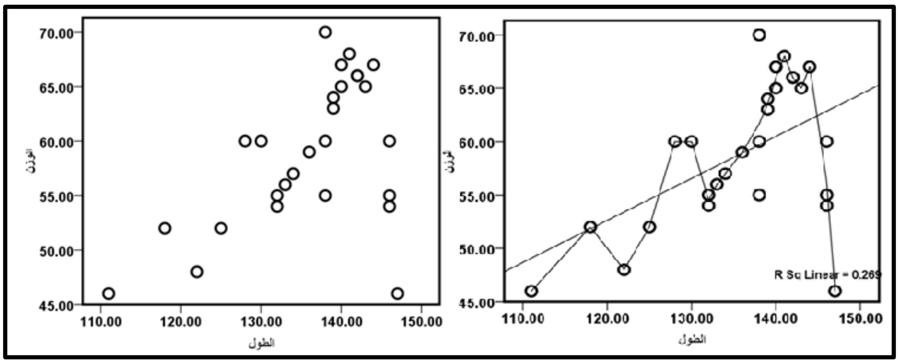
Graphs ⇨ Chart Builder ⇨ Scatter Plot ⇨  أيقونة

ثم نبدأ رسم ثم يتم إدراج الفئات والتكرار باستخدام Mouse، وبالضغط على Ok كما في الشكل التالي:

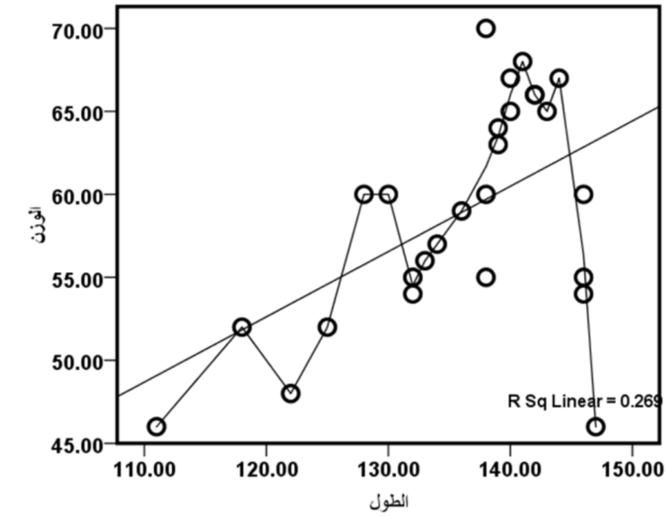
**مثال**: البيانات التالية تمثل أطوال وأوزان 27 من الطلاب.

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| الطول | 111 | 118 | 122 | 125 | 128 | 130 | 132 | 132 | 133 |
| الوزن | 46 | 52 | 48 | 52 | 60 | 60 | 55 | 54 | 56 |
| الطول | 134 | 136 | 138 | 138 | 138 | 139 | 139 | 140 | 140 |
| الوزن | 57 | 59 | 55 | 60 | 70 | 63 | 64 | 65 | 67 |
| الطول | 141 | 142 | 142 | 143 | 144 | 146 | 146 | 146 | 147 |
| الوزن | 68 | 66 | 66 | 65 | 67 | 60 | 55 | 54 | 46 |

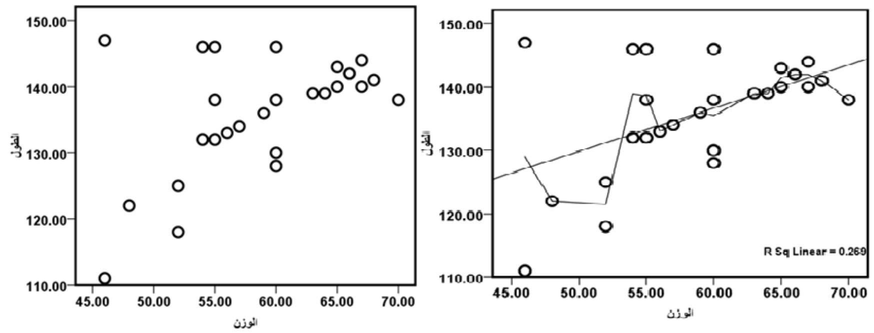
فنحصل على الأشكال التالية:




**شكل (37.5)**: يوضح شكل الانتشار للمغترين الطول والوزن لمائة طالب.




**شكل (38.5)**: يوضح شكل الانتشار للمغترين الطول والوزن والعلاقة بينهما لمائة طالب.




**شكل (39.5)**: يوضح شكل الانتشار للمغترين الوزن والطول

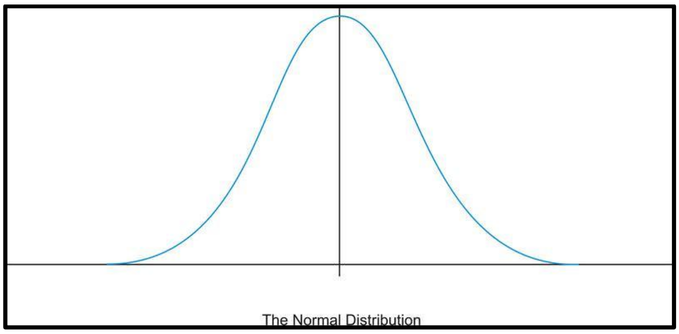
## (4–3) المنحنيات التكرارية المتصلة التي تظهر في الناحية العملية أشكالاً مميزة منها:

### 1. المنحنى المتماثل: Symmetric Curve

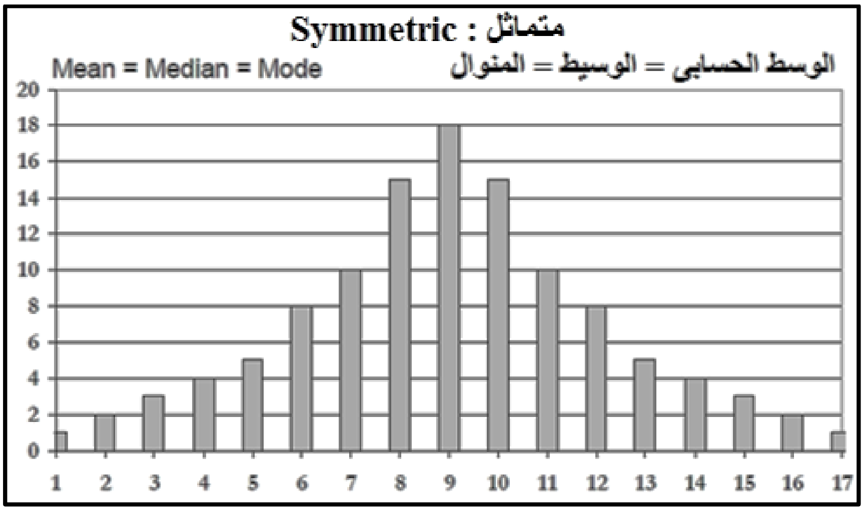
يعرف بأنه المنحنى الذي لو أسقط من قمته عموداً لقسم المساحة تحت المنحنى إلى جزأين متكافئين ومتماثلين.

ويتميز المنحنى التكراري المتماثل أو ذو الشكل الناقوسي بأن المشاهدات المتساوية البعد عن مركز النهاية العظمى لها نفس التكرار. ومن الأمثلة الهامة له المنحنى الاعتدالي.

وهو منحنى ذو نهاية عظمى في منتصفه ثم يقترن من المحور الأفقي تدريجياً على كل من جانبي هذه النهاية تقارباً متساوياً من الجانبين.




**شكل (40.5)**: يوضح المنحنى المتماثل

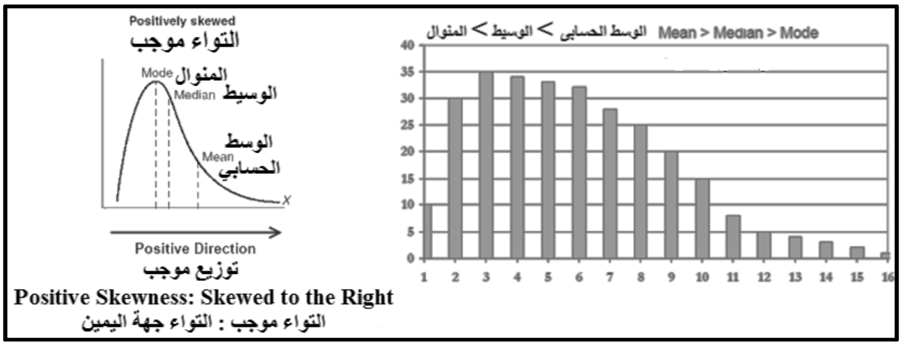



**شكل (41.5)**: يوضح المنحنى المتماثل

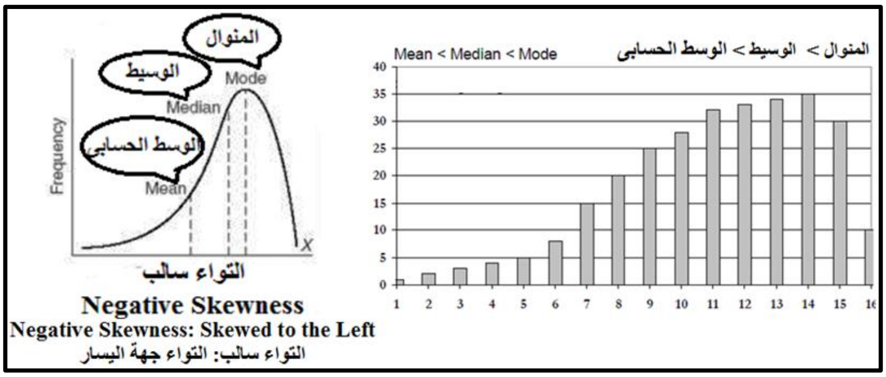
## 2.المنحنيات الملتوية: Skewed Curves

إذا بَعُد المنحنى عن التماثل فإننا نطلق عليه اسم المنحنى الملتوي ويسمى عدم التماثل بالالتواء Skewness.

وهذه المنحنيات التكرارية متوسطة عدم التماثل أو الالتواء تتميز بأن أحد طرفيها يمتد أكثر من الآخر على جانبي مركز النهاية العظمى، وإذا كان الطرف (الأيمن) أطول فيكون المنحنى في هذه الحالة ملتوياً إلى اليمين أو ملتوياً التواء موجباً. بينما لو كان العكس صحيحاً فإن المنحنى يكون ملتوياً إلى اليسار أو ملتوياً التواء سالباً.

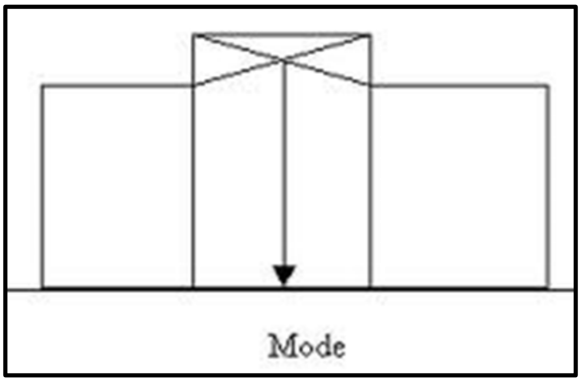



**شكل (42.5)**: يوضح منحنى التواء موجب.




**شكل (43.5)**: يوضح منحنى التواء سالب.

**3. المنوال: Mode**




**شكل (44.5)**: يوضح المنوال.

**(5–4) التمثيل البياني للبيانات المتقطعة:**

يمكن تمثيل البيانات المتقطعة بيانياً بأن نحصل على شكل يمثل التوزيع التكراري البسيط وكذلك التوزيع المتجمع الصاعد والمتجمع الهابط للظاهرة.

**مثال:**

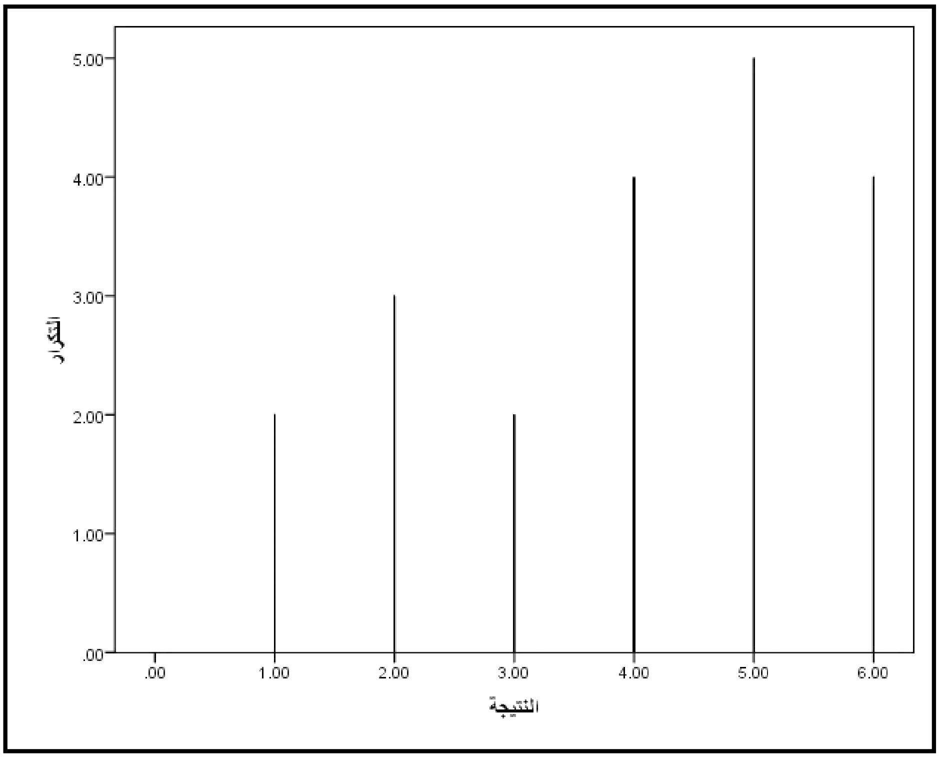
البيانات الآتية تمثل التوزيع التكراري الناتج عن إلقاء زهرة نرد عشرين مرة.

| النتيجة | 1 | 2 | 3 | 4 | 5 | 6 | مجموع |
|---|---|---|---|---|---|---|---|
| التكرارات | 2 | 3 | 2 | 4 | 5 | 4 | 20 |

**والمطلوب**: التمثيل البياني للتوزيع التكراري وأيضاً التمثيل البياني للتكرار المتجمع الصاعد لإلقاء زهرة نرد عشرين مرة.

**الحل:**

**أولاً**: التمثيل البياني للتوزيع التكراري.



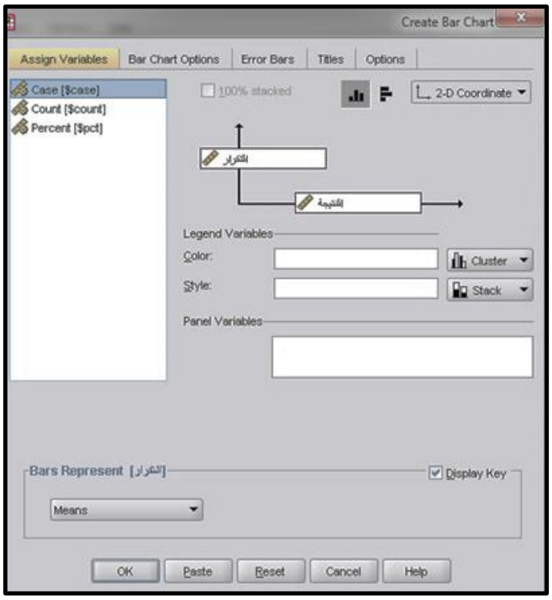

**شكل (45.5)**: التمثيل البياني للتوزيع التكراري إلقاء زهرة نرد عشرين مرة.

وباستخدام الحزمة الإحصائية Spss باتباع المسار التالي.

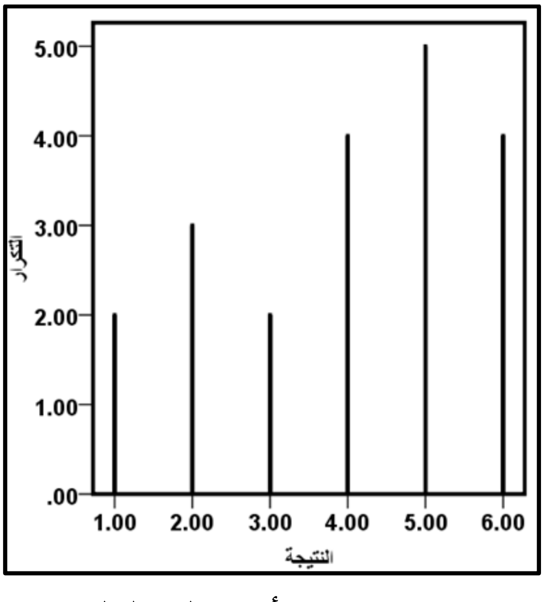
**Graphs ⇨ Interactive ⇨ Bar**

واستكمال الصندوق الحواري التالي:

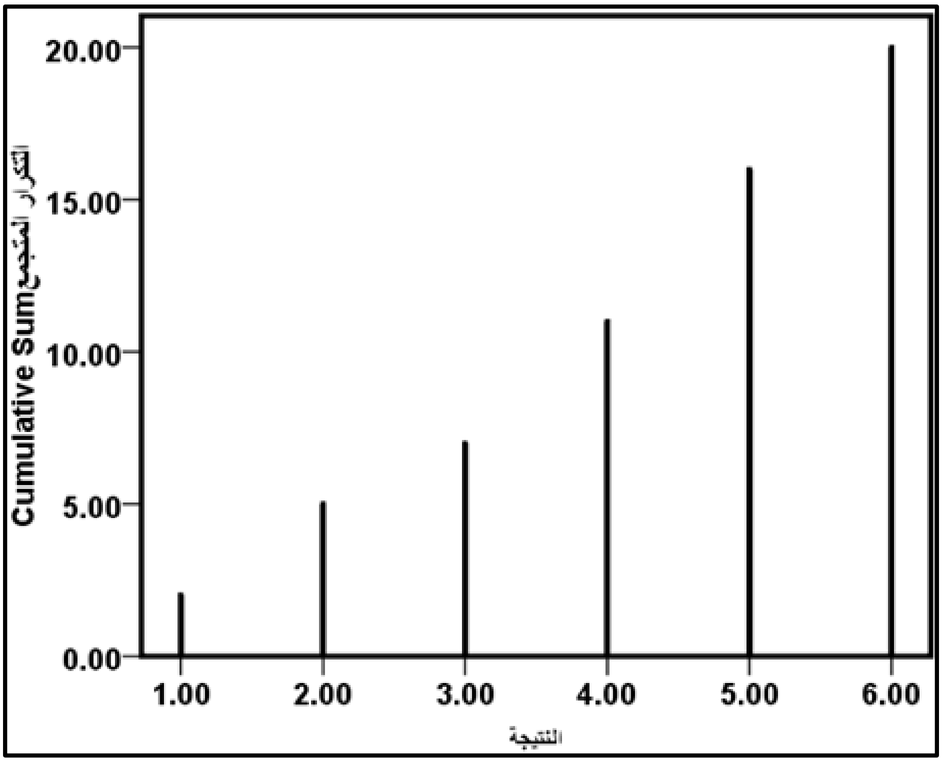
ثم يتم إدراج الفئات والتكرار باستخدام Mouse، وبالضغط على Ok نحصل على الرسم البياني وبالضغط مرتين متتاليتين يمكننا إجراء تعديلات على الرسم البياني من حيث اللون والخط ... إلخ كما يلي:




**شكل (46.5)**: يوضح إدراج المتغيرين لرسم الأعمدة البسيطة.




**شكل (47.5)**: يوضح رسم الأعمدة البسيطة لرمية زهرة النرد.




**شكل (48.5)**: يوضح رسم التكرار المتجمع لرمية زهرة النرد.

**(4–5) المنحنيات النسبية: Relative Curves**

من المهم خصوصاً في الإحصاء التحليلي (الاستدلالي) أن نتعامل مع منحنيات تحتها تساوي الواحد الصحيح، ومن أجل عمل ذلك فإنه من الضروري تحويل التكرارات العادية إلى تكرارات نسبية ويتم ذلك بتحويل التكرار المقابل لكل فئة إلى جزء من التكرار الكلي وبذلك يمكن إيجاد التكرار النسبي في التوزيعات الوصفية والكمية كالآتي:

التكرار النسبي = التكرار الأصلي ÷ مجموع التكرارات

### أ) المنحنيات النسبية في حالة الجدول المنتظم

**مثال:**

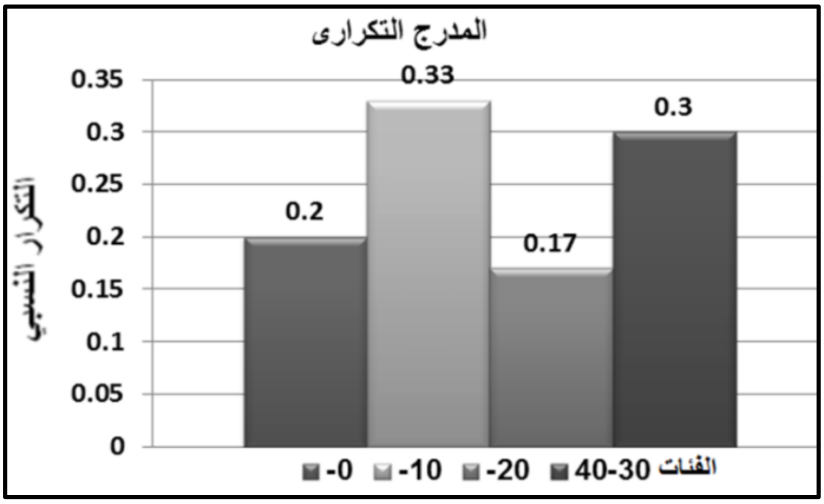
الجدول التالي يوضح درجات 30 طالباً في مادة علم النفس

| الفئات | 0 - | 10 - | 20 - | 30 - 40 | المجموع |
|---|---|---|---|---|---|
| التكرارات | 6 | 10 | 5 | 9 | 30 |

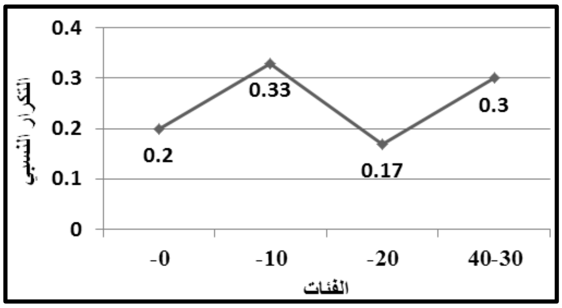
**والمطلوب:** رسم المدرج والمضلع والمنحنى التكراري النسبي

**الحل:**

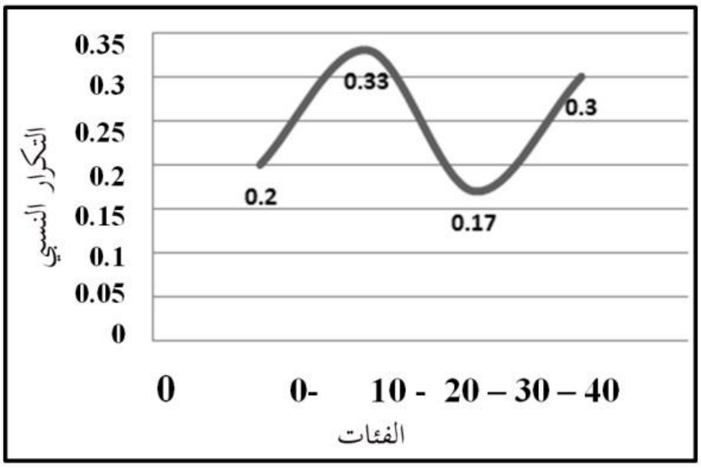
| الفئات | 0 - | 10- | 20 - | 30- 40 | المجموع |
|---|---|---|---|---|---|
| التكرارات | 6 | 10 | 5 | 9 | 30 |
| التكرار النسبي | 0.2 | 0.33 | 0.17 | 0.3 | 1 |




**شكل (49.5)**: يوضح المدرج التكراري النسبي لتوزيع درجات 30 طالباً في علم النفس




**شكل (50.5)**: يوضح المضلع التكراري النسبي لتوزيع درجات 30 طالباً في علم النفس




**شكل (51.5)**: يوضح المنحنى التكراري النسبي لتوزيع درجات 30 طالباً في علم النفس

**ب) المنحنيات النسبية في حالة الجدول غير المنتظم:**

عندما يكون الجدول غير منتظماً فإن التكرارات النسبية تمثل مساحة المستطيل والتي يكون مجموعها مساوياً للواحد الصحيح إلا أنه يجب أن نعدل التكرارات النسبية حتى يتسنى لنا رسم المدرج والمضلع والمنحنى. بأن نحسب التكرارات النسبية مقاسة بوحدات طول الفئة.

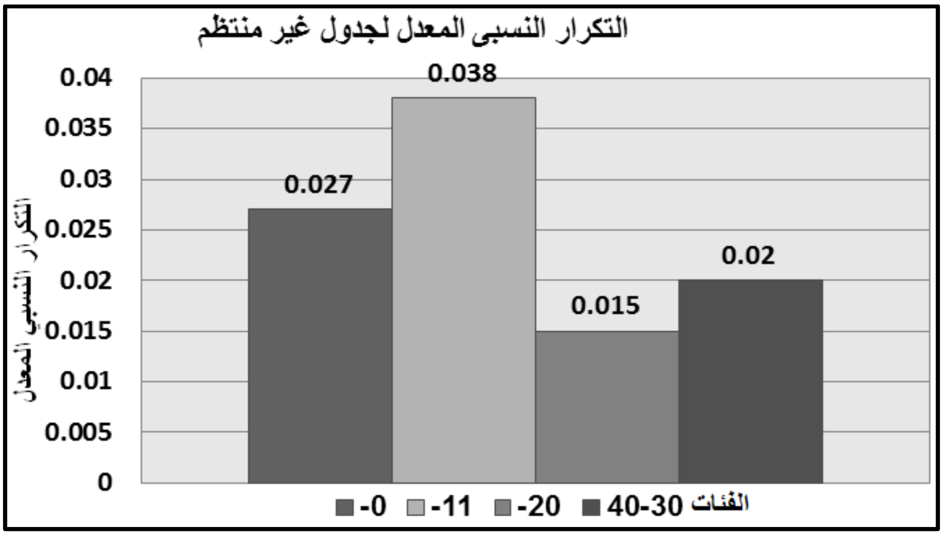
التكرار النسبي المعدل = التكرار الأصلي ÷ ( مجموع التكرارات × طول الفئة المناظر للتكرار الأصلي)

**مثال:** الجدول التالي يوضح درجات 20 طالباً في مادة علم النفس.

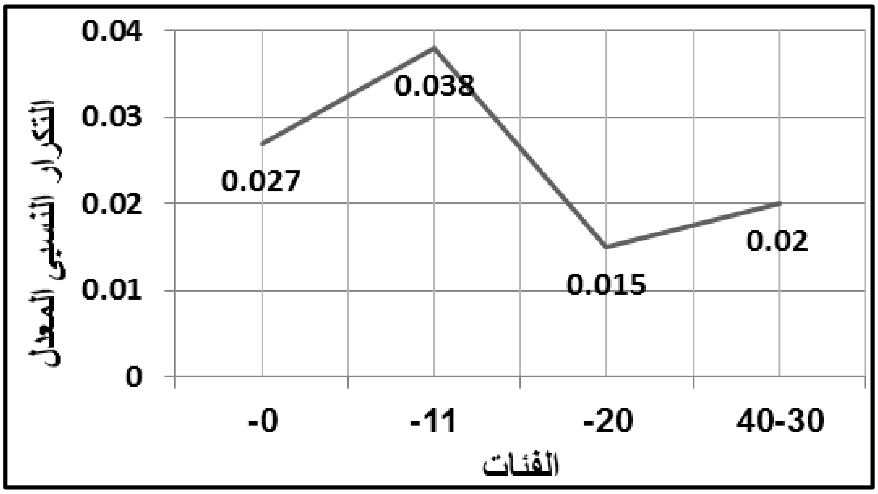
| الفئات | 0 - | 11 - | 20 - | 30 - 40 | المجموع |
|---|---|---|---|---|---|
| التكرارات | 6 | 7 | 3 | 4 | 20 |

**والمطلوب:** رسم المدرج والمضلع والمنحنى التكراري النسبي المعدل:

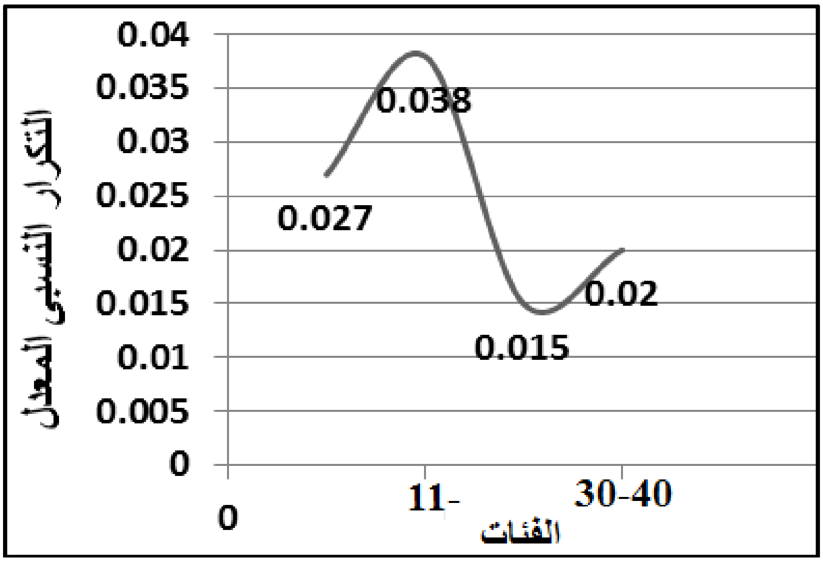
| الفئات | 0 - | 11 - | 20 - | 30 – 40 | المجموع |
|---|---|---|---|---|---|
| التكرارات | 6 | 7 | 3 | 4 | 20 |
| طول الفئة | 11 | 9 | 10 | 10 | - |
| التكرار النسبي | 6 ÷ 20 = 0.3 | 7 ÷ 20 = 0.35 | 3 ÷ 20 = 0.15 | 4 ÷ 20 = 0.2 | 1 |
| التكرار النسبي المعدل | 6 ÷ (8× 20) = 0.027 | 7 ÷ (9× 20) = 0.038 | 3 ÷ (10× 20) = 0.015 | 4 ÷ (10× 20) = 0.02 | - |




**شكل (52.5):** يوضح المدرج التكراري النسبي لتوزيع درجات 20 طالباً في علم النفس.




**شكل (53.5)**: يوضح المضلع التكراري النسبي المعدل لتوزيع درجات 20 طالباً في علم النفس

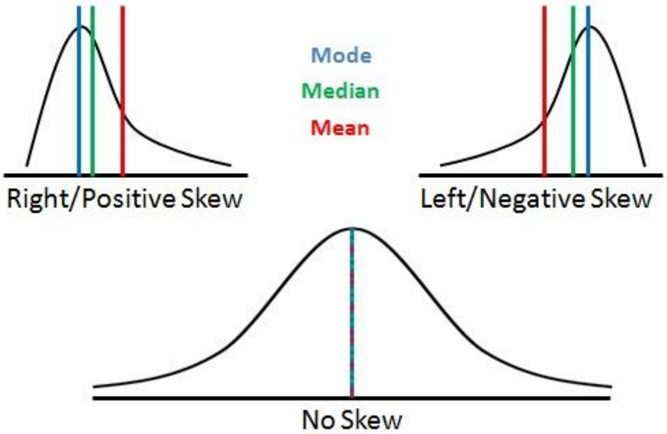



**شكل (54.5)**: يوضح المنحنى التكراري النسبي المعدل لتوزيع درجات 20 طالباً في علم النفس

## الباب السادس

## مقاييس النزعة المركزية

## Central Tendency Measures




- مقدمة.

(6 – 1) الوسط (المتوسط) الحسابي.

(6-2) الوسيط.

(6-3) الرُبيعيات.

(6-4) المنوال.

(6 – 5) العلاقة الاعتبارية بين الوسط والوسيط والمنوال.

(6 – 6) أنواع المنحنيات.

**مقدمة:**

تشير مقاييس النزعة المركزية إلى توزيع البيانات وتشير إلى قيمة المتغير بالقرب من مركز التوزيع وهي تدرس تراكم أو التمركز حول قيمة معينة يسمى بالنزعة المركزية للتوزيع.

وحيث أن المتوسط هو القيمة النموذجية أو الممثلة لمجموعة من البيانات، وحيث إن مثل هذه القيمة النموذجية تميل إلى الوقوع في المركز داخل مجموعة بيانات مرتبة حسب قيمتها، فإن المتوسطات تسمى أيضاً بمقاييس النزعة المركزية.

**تذكر:**

الحزمة الإحصائية Spss تتعامل مع الدرجات الخام (المفردات)، وأن البيانات المبوبة (الجداول التكرارية) هي وسيلة تقريبية للتعامل اليدوي مع الدرجات الخام بصورة مبسطة ولتوضيح ذلك كمثال الفئة (4–2) لنفرض التكرار المكتوب أمامها 3 أي 3 أرقام فاحتمال (3؛2؛2) أو (3؛3؛2) أو (3؛3؛3) أو (4؛2؛2) أو (4؛3؛2) أو (4؛4؛4) فأي الدرجات الخام (المفردات) ستكون معبرة عن الدرجات الفعلية للبيانات (المفردات) التي أساس تكوين هذه الفئة، لذا فالبرنامج مصمم للتعامل مع الدرجات الخام (المفردات) والكتاب سوف يشرح الأساليب الإحصائية معتمداً على الدرجات الخام(البيانات الغير مبوبة؛ المفردات).

**(6 – 1) الوسط (المتوسط) الحسابي: Arithmetic Mean or Average ($\bar{x}$)**

يطلق على الوسط الحسابي اسم المتوسط الحسابي وهو يعتبر من أكثر المتوسطات انتشاراً ويأخذ جميع قيم المفردات في الحسبان عند حسابه.

ويرمز للوسط الحسابي للمجتمع The population mean بالرمز μ (ميو Mu)، أما الوسط الحسابي للعينة The sample فيرمز له بالرمز $\bar{x}$ (X-bar) ويقرأ إكس بارX bar.

**تعريف الوسط الحسابي:**

**1. حالة البيانات غير المبوبة (المفردة):**

هو القيمة التي لو أعطيت لكل مفردة من المفردات لكان مجموع القيم الجديدة مساويا لمجموع القيم الأصلية أو هو حاصل جمع مفردات القيم مقسوماً على عددها.

$$\text{Mean} = \text{sum of data values}/\text{number of data values}$$

$$\text{Mean} = \frac{\text{sum of data values}}{\text{number of data values}}$$

**2. في حالة البيانات المبوبة (الجداول التكرارية):**

هو الوسط الحسابي المرجح بالتكرارات.

**حساب الوسط الحسابي:**

**أ. الوسط الحسابي من المجتمع:**

بفرض أن لدينا قيم (مشاهدات المتغير observations of a variable) المتغير x هي ($x_1, x_2,\dots x_N$) حيث (N) تمثل عدد المشاهدات في المجتمع (number of observations in the population) فإن الوسط الحسابي (μ) لمجموعة البيانات هو مجموع هذه البيانات مقسوماً

The sum of all of the values of the variable in the data set divided by the number of

$$\mu = [X_1 + X_2 \ldots\ldots\ldots +X_N]/N = \sum X_i /N$$

$$\mu = \frac{[X_1+X_2 \ldots\ldots\ldots+X_N]}{N}$$

$$\mu = \frac{\sum X_i}{N}$$

**حساب الوسط الحسابي للعينة:**

بفرض أن لدينا قيم مشاهدات المتغير (observations of a variable ) المتغير x هي ($x_1$ $x_2$, … $x_n$) حيث (n) تمثل عدد المشاهدات في العينة ( number of observations in the sample) فإن الوسط الحسابي لمجموعة البيانات هو مجموع هذه البيانات مقسوماً على عددها The sum of all of the values of the variable in the data set divided by the number of ويرمز لذلك بالرمز $\bar{x}$ ويعطى بالعلاقة التالية:

$$\bar{x} = \sum(x_1 + x_2 + x_3 \ldots\ldots\ldots +x_n) \div n$$

$$\overline{X} = \frac{[X_1+X_2 \ldots\ldots\ldots+X_n]}{n}$$

$$\overline{X} = \frac{\sum X_i}{n}$$

حيث أن: $\sum x$: مجموع القيم (المفردات)، n: عدد القيم (المفردات).

**مثال:**

احسب الوسط الحسابي للأرقام 15، 12، 8، 7، 6.

**الحل:**

الوسط الحسابي:

$$\bar{x} = \sum(x_1 + x_2 + x_3 \ldots\ldots\ldots + x_n) \div n$$

$$\bar{x} = (6 + 7 + 8 \ + 12 + 15)\ ) \div 5 = 9.6$$

أو

$$\bar{x} = \frac{\sum(x_1+x_2+x_3 \ldots\ldots\ldots+x_n)}{\text{n}}$$

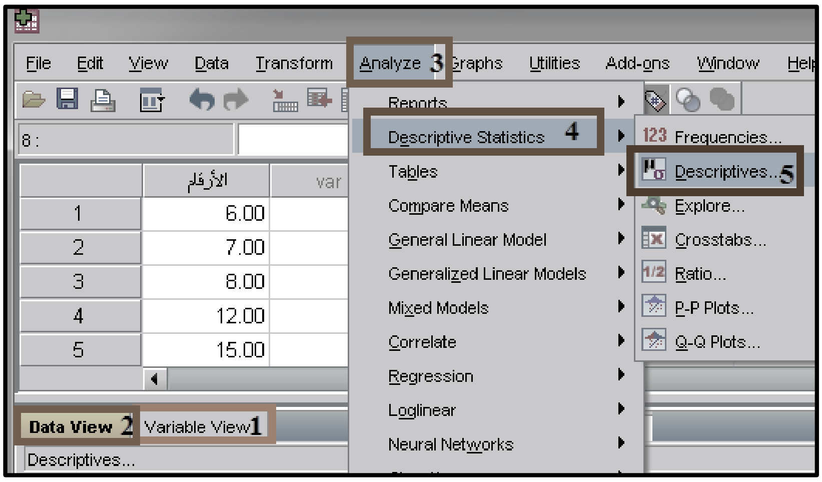
$$\bar{x} = \frac{\sum(6+7+8+12+15)}{5} = 9.6$$

**الحل:**

باستخدام الحزمة الإحصائية Spss.

من قائمة Analyze نختار الأمر Descriptive Statistics ثم نختار الأمر Descriptives.

**Analyze** ⇨ **Descriptive Statistics** ⇨ **Descriptives**




**شكل (1.6):** يوضح إدراج البيانات واختيار المسار لحساب الوسط الحسابي.


**شكل (2.6)**: يوضح اختيار الأسلوب لحساب الوسط الحسابي.

بالضغط على ok نحصل على النتائج التالية:

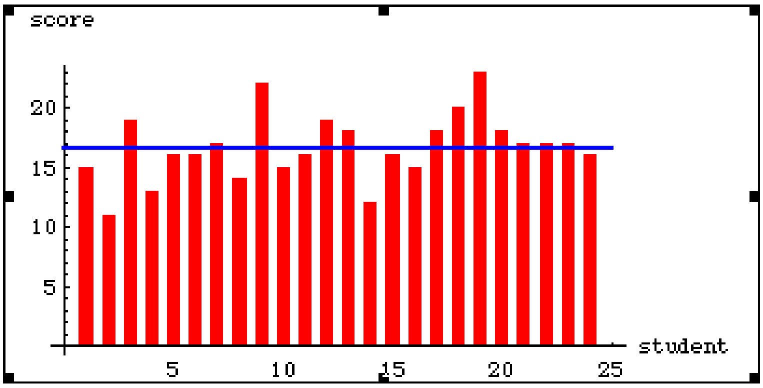
| Descriptive Statistics | | |
|---|---|---|
| | N | Mean |
| الأرقام | 5 | 9.6000 |
| Valid (List Wist) | 5 | |

**س**: الجدول التالي نتائج 24 طالباً في إحدى الامتحانات:

| رقم الطالب | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| الدرجة | 15 | 11 | 19 | 13 | 16 | 16 | 17 | 14 | 22 | 15 | 16 | 19 |
| رقم الطالب | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 |
| الدرجة | 18 | 12 | 16 | 15 | 18 | 20 | 23 | 18 | 17 | 17 | 17 | 16 |

**المطلوب**: مثل البيانات التالية بالأعمدة البسيطة، واحسب الوسط الحسابي.

**الحل**: تمثيل البيانات بالأعمدة البسيطة.




**شكل (3.6)**: يوضح تمثيل البيانات بالأعمدة البسيطة.

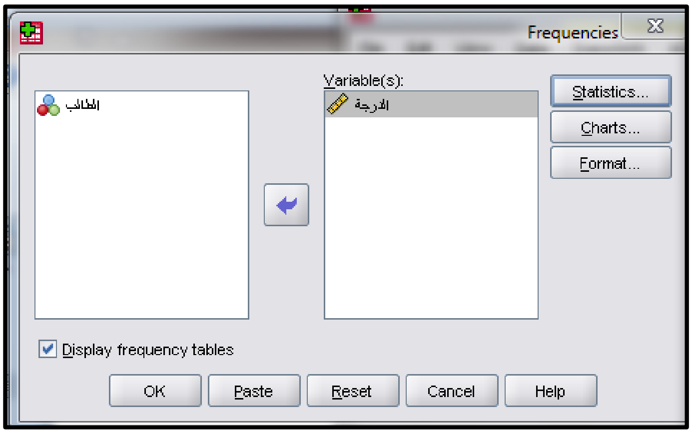
$$\overline{X} = 400 \div 24 = 16.667 \cong 16.67$$

**باستخدام الحزمة الإحصائية Spss.**

ندخل البيانات وباتباع المسار التالي:

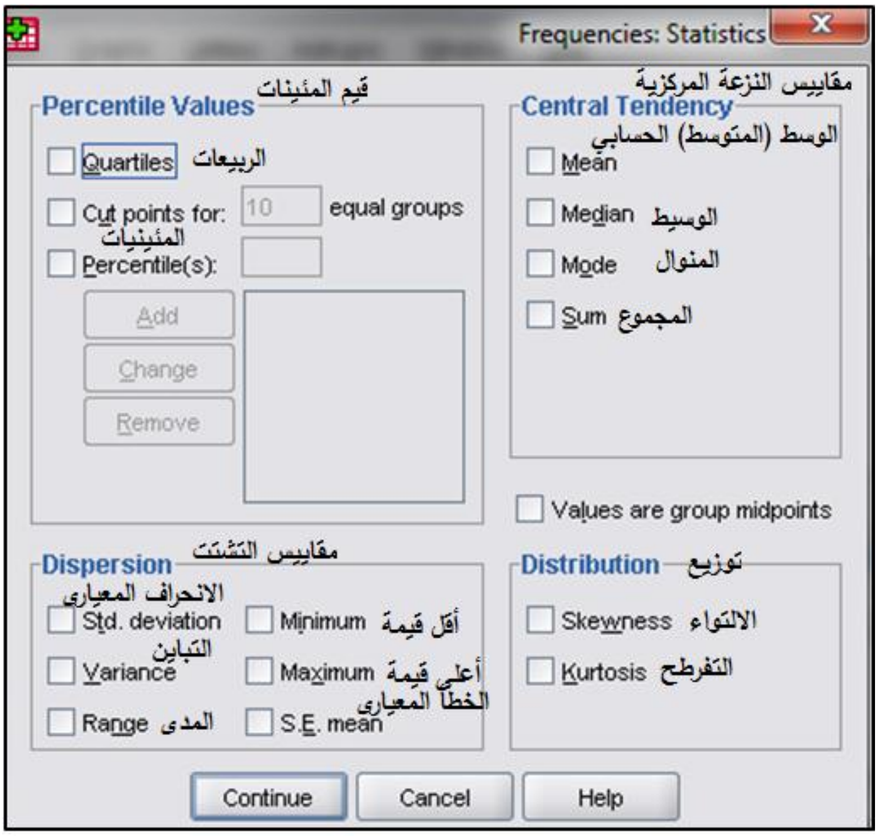
**Analyze** ⇨ **Descriptive Statistics** ⇨ **Frequencies**

**يظهر الشكل التالي:**



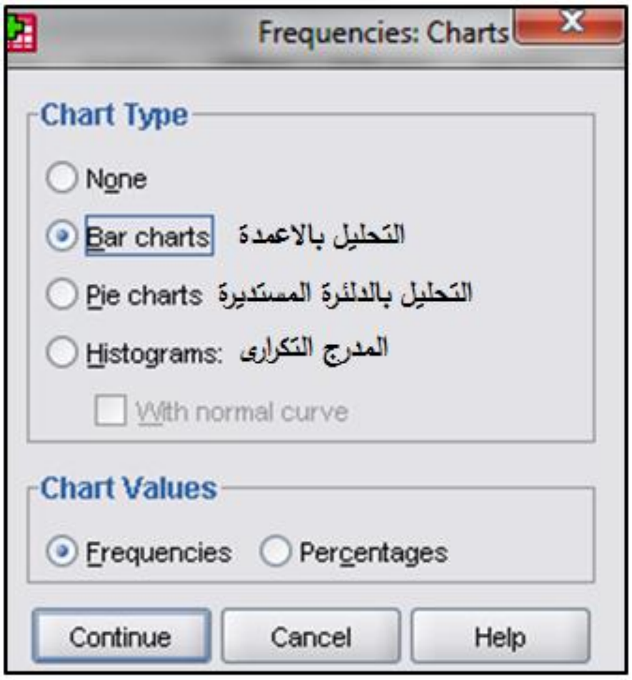

**شكل (4.6)**: يوضح إدراج المتغير المراد إجراء الإحصاء عليه.

بالضغط على Statistics لتحديد الأسلوب الإحصائي المراد تقديره نحصل على الشكل التالي:




**شكل (5.6)**: يوضح إدراج الأسلوب الإحصائي المطلوب إجرائه.

وبالضغط على Continue نعود للشكل السابق ونضغط على Charts لتحديد الرسم البياني الذي نرغب تمثيل البيانات به فنحصل على الشكل التالي:

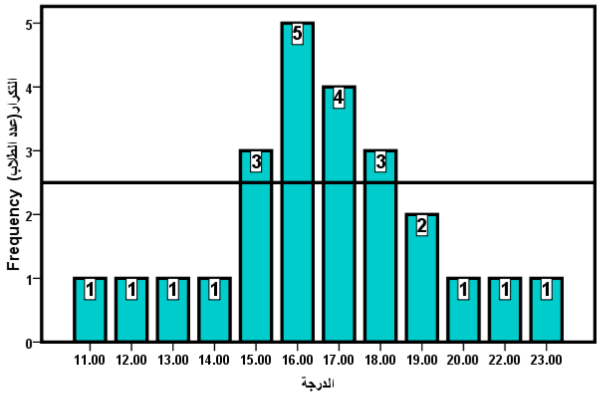



**شكل (6.6)**: يوضح إدراج الرسم البياني المطلوب إجراؤه.

وبالضغط على Continue نعود للشكل الأوّلي والضغط على Ok نحصل على النتائج التالية:

الوسط الحسابي:

| Statistics | | |
|---|---|---|
| الدرجة | | |
| N | Valid | 24 |
| | Missing | 0 |
| Mean | | 16.6667 |

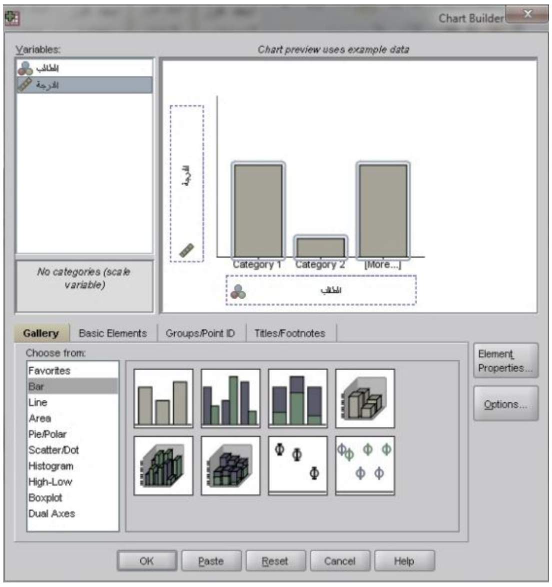



**شكل (7.6)**: يوضح الرسم البياني بالأعمدة لتوزيع الدرجات.

**ندخل البيانات ثم نتبع المسار التالي:**

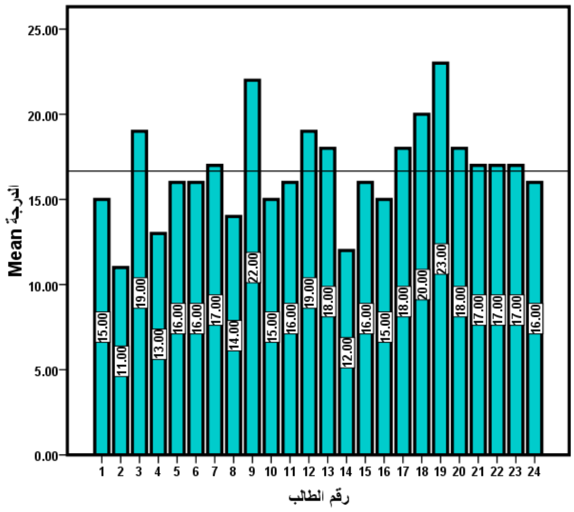
**Graphs** ⟹ **Chart Builder**

فنحصل على الشكل التالي نختار منه الرسم البياني الذي نرغب الحصول عليه.




**شكل (8.6)**: يوضح إدراج الرسم البياني المطلوب إجراؤه.

فنحصل على الرسم التالي:




**شكل (9.6)**: يوضح الرسم البياني بالأعمدة البسيطة لدرجات الطلاب.

**مثال**:

لدينا وزن لمجموعة من البرتقال بالجرام 0.903؛ 0.920؛ 0.861؛ 1.009؛ 0.971؛ 0.898؛ 0.942؛ 0.897

**المطلوب**: حساب الوسط الحسابي للعينة Find the mean of this sample .

**الحل**:

ندخل البيانات وباتباع المسار التالي:

Analyze → Descriptive Statistics → Frequencies

| Descriptive Statistics | | | |
|---|---|---|---|
| | N | Sum | Mean |
| N | 8 | 7.40 | 0.9251 |
| Valid (List Wist) | 8 | | |

**مزايا الوسط الحسابي:**

**1**. يستخدم للمقارنة بين المجموعات.

**2**. المجموع الجبري لانحرافات (فروق القيم عن وسطها الحسابي يساوي صفراً.

مثال:

فيما يلي درجات 8 طلاب في مقرر إحصاء تطبيقي 1، 2، 3، 4، 5

**بحساب الوسط الحسابي:**

$$\bar{x} = \frac{x_1+x_2 \ldots\ldots +x_n}{n}$$

$$\bar{X} = [1+2+3+4+5] \div 5 = 3$$

**بما أن عدد الطلاب 5 والوسط الحسابي3** $\bar{X} = 3$ **فإن:**

| X | 1 | 2 | 3 | 4 | 5 | $\Sigma$ |
|---|---|---|---|---|---|---|
| $x - \bar{x}$ | $1-3=$ | $2-3=$ | $3-3=$ | $4-3=$ | $5-3=$ | **0** |
| | -2 | -1 | 0 | 1 | 2 | |

**3**. يمكن حسابه دون معرفة القيم مفردة بل يكفي لحسابه معرفة مجموع هذه القيم وعددها.

بما أن مجموع درجات الطلاب 15 وعددهم 5 إذن الوسط الحسابي

$$\overline{\mathrm{X}} = 15 \div 5 = 3$$

**4.** لا يمكن إهمال أي مفردة من المفردات عند حسابه مما يؤدي إلى ارتفاع كفاءة الوسط الحسابي بشكل عام كمتوسط للقيم.

**5.** إذا أضيف مقدار ثابت إلى كل قيمة من القيم، فإن الوسط الحسابي للقيم المعدلة (بعد الإضافة) يساوي الوسط الحسابي للقيم الأصلية (قبل الإضافة) مضافا إليها هذا المقدار الثابت. فإذا كانت القيم هي: $x_1 + x_2 \ldots\ldots x_{\mathrm{n}}$، وتم إضافة مقدار ثابت (a) إلى كل قيمة من القيم، ونرمز للقيم الجديدة بالرمز $y$، أي أن $y = x + a$ ، فإن: الوسط الحسابي لقيم $y$ (القيم بعد الإضافة) هو:

$$\bar{y} = \bar{x} + a$$

حيث أن $\bar{y}$ هو الوسط الحسابي للقيم الجديدة

**6.** إذا ضرب مقدار ثابت (a) في كل قيمة من القيم، فإن الوسط الحسابي للقيم المعدلة (القيم الناتجة بعد الضرب) يساوي الوسط الحسابي للقيم الأصلية (القيم بعد التعديل) مضروبا في هذا المقدار الثابت. أي أنه إذا كان: $y = a\,x$، ويكون الوسط الحسابي $\bar{y}$ للقيم الجديدة.

$$\bar{y} = a\,\bar{x}$$

**عيوب الوسط الحسابي:**

**1.** يهضم حق القيم المعتدلة باعتباره يتأثر بالقيم المتطرفة، فيتأثر المتوسط بالدرجات القريبة منه تأثرا قليلاً، ويتأثر بالدرجات البعيدة عنه تأثراً كبيراً.

**مثال**: القيم (5، 13، 17، 133) متوسطها 42 وهو يزيد على قيمتين ولا يقل عن قيمة واحدة.

**2**. لا يمكن حسابه للتوزيعات التكرارية المفتوحة من أحد طرفيها أو من الطرفين معاً لاعتماده على مراكز الفئات.

**3**. لا يمكن حسابه بيانياً (بالرسم).

**4**. لا يمكن حسابه لبيانات غير كمية (وصفي) سواء أكانت وصفية اسمية أو ترتيبية، إذ لا معنى له في هذه الحالات.

**5**. لا يصلح كمقياس للنزعة المركزية إذا كان التوزيع ملتوياً.

**الأهمية العملية للوسط (المتوسط) الحسابي فيما يلي**:

تعتمد المعايير المختلفة على الوسط (المتوسط) الحسابي، ولهذا فإن مقياس ذكاء الفرد بالنسبة لمتوسط ذكاء جيله وأقرانه، ومدى انحرافه عن هذا المعيار زيادة أو نقصان، وينسب وزنه وطوله وحجمه إلى معايير أقرانه أيضا.

ولهذا تصنع الملابس المختلفة لتناسب متوسطات أطوال وأحجام كل عمر من أعمار الإنسان، وبما أن هذه المعايير تختلف في بعض نواحيها من بيئة لأخرى، فلكل بيئة معاييرها الخاصة بها، ومن هذا نرى خطأ نسبة الفرد إلى معايير عبر معايير بيئة أخرى.

وتستخدم المتوسطات (الأوساط) الحسابية أحيانا لمقارنة مجموعة من الأفراد بمجموعة أخرى مثل مقارنة متوسط درجات فصل دراسي ما في اختبار ما بمتوسط درجات فصل دراسي

آخر لنفس الاختبار، ولا تصح للمقارنة إلا إذا كانت المجموعات متجانسة وتقبل خواصها مثل تلك المقارنات.

ومن أمثلة المقارنات الخاطئة ما يقوم منها على مقارنة متوسط أعمار الناس في بيئة صناعية أغلبها من الشباب بمتوسط أعمار الناس في بيئة زراعية قد يكون أغلبها من الأطفال والشيوخ، ولهذا تعتمد شركات التأمين على دراسة متوسطات الأعمار بالنسبة لكل مهنة، وكل عمر، حتى لا تخسر كثيراً.

### (6–2) الوسيط: Median (Middle Quartile)

هو القيمة التي تقع في منتصف البيانات عند ترتيبها في ترتيب تصاعدي أو تنازلي.

The value that lies in the middle of the data when arranged in ascending or descending order.

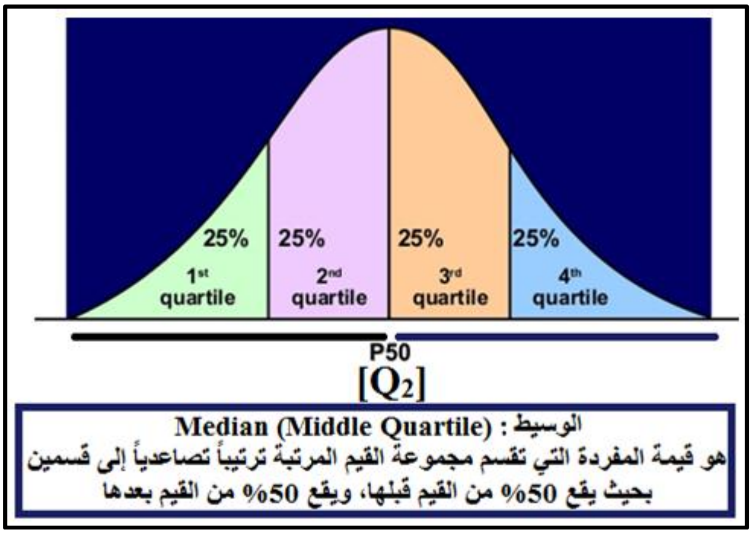
أو نصف البيانات أقل من المتوسط ونصف فوق المتوسط. الوسيط يشير إلى M

Half the data are below the median and half are above the median. The median is denoted M.

أو هو النقطة التي تقع تماما في منتصف توزيع الدرجات بحيث يسبقها نصف عدد الدرجات ويتلوها النصف الآخر.

The score that is larger than 50% and smaller than 50% of scores.

أو القيمة التي تقسم المجموعة إلى مجموعتين عدد المفردات الأقل منها يساوي عدد المفردات الأكثر منها.




**شكل (10.6)**: يوضح مفهوم الوسيط.

**حساب الوسيط**:

لحساب الوسيط من البيانات غير المبوبة فلابد من ترتيبها تصاعدياً أو تنازلياً أولاً ثم يتم تقدير قيمة الوسيط كالآتي:

Order the sample to get the Order Statistics:

$$x_1, x_2, \ldots\ldots, x_n \Rightarrow x_{(1)} \leq x_{(2)} \leq \ldots\ldots \leq x_{(n)}$$

**Median ($M_e$)**

**أ. إذا كان عدد المفردات (ن) فردي**:

تحديد ترتيب الوسيط (رتبة الوسيط) = [عدد المشاهدات +1] ÷ 2

If n is odd, the median is the $(n + 1)/2$ observation in the position

**If N is odd, the median = middle score**

فإن:

ترتيب الوسيط من العلاقة التالية:

$$\text{Median rank} = (n + 1)/2$$

$$\text{Median rank} = \frac{(n+1)}{2}$$

وعليه فإن قيمة الوسيط هي قيمة المفردة التي ترتيبها هذا الموضع.

**مثال**: أوجد الوسيط لمجموعة الأرقام 6، 10، 8، 3، 5

**الحل**:

n = 5 فردي

الترتيب التصاعدي: 3، 5، <u>6</u>، 8، 10. أو الترتيب التنازلي: 10، 8، <u>6</u>، 5، 3.

$$\text{Median rank} = (n + 1)/2$$
$$\text{Median rank} = (5 + 1)/2 = 3$$

∴ قيمة الوسيط (قيمة المفردة التي ترتيبها رقم 3) = 6

$$\boldsymbol{Me = 6}$$

باستخدام الحزمة الإحصائية Spss ندخل البيانات وباتباع المسار التالي:

**Analyze** ⇨ **Descriptive Statistics** ⇨ **Frequencies**

| Statistics | | |
|---|---|---|
| المشاهدات | | |
| N | Valid | 5 |
| | Missing | 5 |
| Median الوسيط | | 6.0000 |

**ب) إذا كان عدد المفردات (ن) زوجياً:**

فإن قيمة الوسيط لن تساوي قيمة مفردة محددة من المفردات الفعلية حيث لا توجد مفردة تقع في منتصف البيانات تماماً، ولذلك نفترض أن قيمة الوسيط تساوي القيمة المتوسطة للقيمتين المركزيتين كالتالي:

رتبتا الوسيط:

Median rank = n/2; [n/2]+1

If n is even, the median is the mean of the two middle observations

وعليه قيمة الوسيط = متوسط قيمتي المشاهدتين (المفردتين) اللتين يحتل ترتيبهما الموضعين.

If N is even median = average of two middle scores

**مثال**: أوجد الوسيط لمجموعة الأرقام: 6، 8، 2، 5، 9، 5، 6، 2، 5، 3

**الحل**:

**أولاً**: ترتيب المفردات: 2، 2، 3، 5، 5، 5، 6، 6، 8، 9

$\mathbf{n} = 10$ زوجي.

$$\text{Median rank} = 10/2;\ [10/2]+1$$

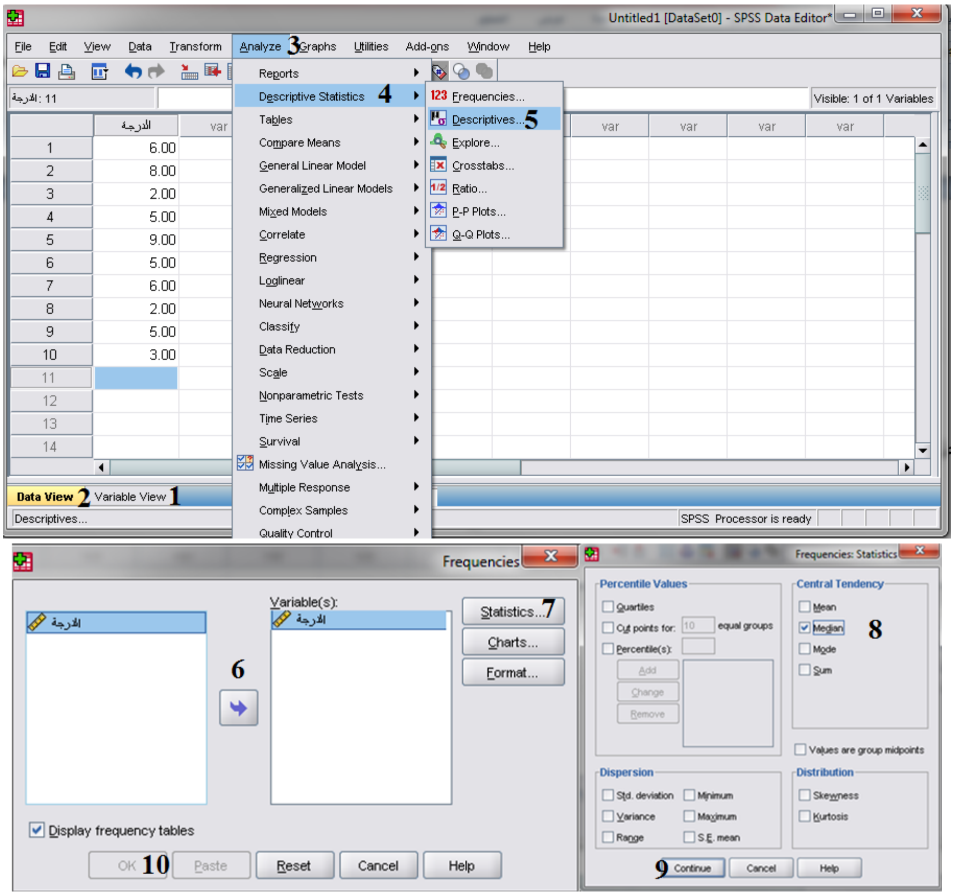
$$\text{Median rank} = 5;\ 6$$

∴ قيمة الوسيط = متوسط قيمتي المفردتين اللتين يحتل ترتيبهما الموضعين (5، 6)

$$M_e = [5+5]/2 = 5$$

باستخدام الحزمة الإحصائية Spss ندخل البيانات وباتباع المسار التالي:

**Analyze** ⇨ **Descriptive Statistics** ⇨ **Frequencies**




فنحصل على قيمة الوسيط:

| Statistics | | |
|---|---|---|
| المشاهدات | | |
| N | Valid | 10 |
| | Missing | 0 |
| Median الوسيط | | 5.0000 |

**الوسيط بالرسم:**

يمكن التعبير عن الوسيط هندسياً بأنه القيمة (س) على الإحداثي السيني الذي إذا رسم عنده عمود رأسي فإنه يقسم المدرج التكراري إلى جزأين متساويين.

**مزايا الوسيط:**

**1.** يمكن حسابه للتوزيعات التكرارية المغلقة والمفتوحة.

**2.** لا تتأثر قيمته بالقيم المتطرفة وإنما بتغير عدد المفردات.

**3.** متوسط موضعي: لأن قيمته تتحدد بقيمة المفردة التي تحتل موضعاً معيناً داخل التوزيع.

**4.** يمكن حسابه عن طريق الرسم من المدرج التكراري أو المنحنى التكراري الصاعد أو الهابط أو الاثنين معاً.

**5.** يمكن حسابه للتوزيعات التكرارية التي تختلف فيها أطوال الفئات دون الحاجة إلى تعديل هذه الأطوال.

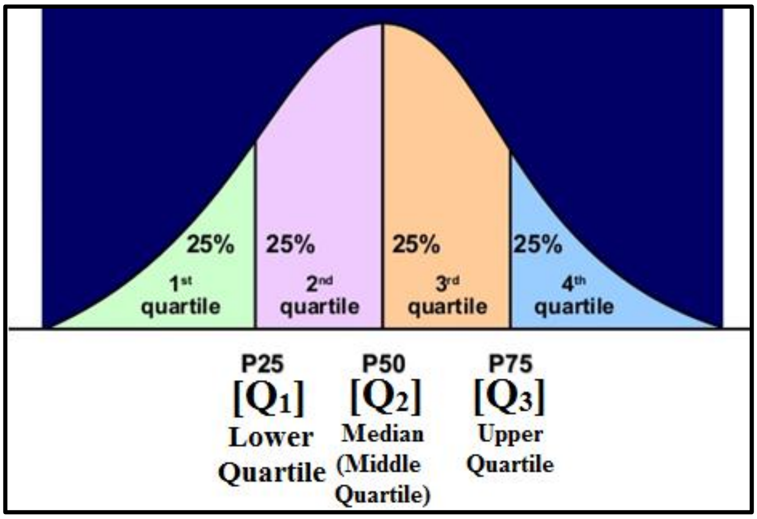
**عيوب الوسيط:**

**1.** ليس له نفس شيوع استخدام الوسط الحسابي.

**2.** تحدد قيمته مفردة أو مفردتان على الأكثر بمعنى لا تسهم جميع القيم في تحديد قيمته.

**(3–6) الرُبيعات: *Quartiles***

يمكن تقسم المساحة تحت المضلع التكراري إلى أربعة أقسام متساوية تسمى الرُبيعات وعددها ثلاثة هي من اليسار إلى اليمين:

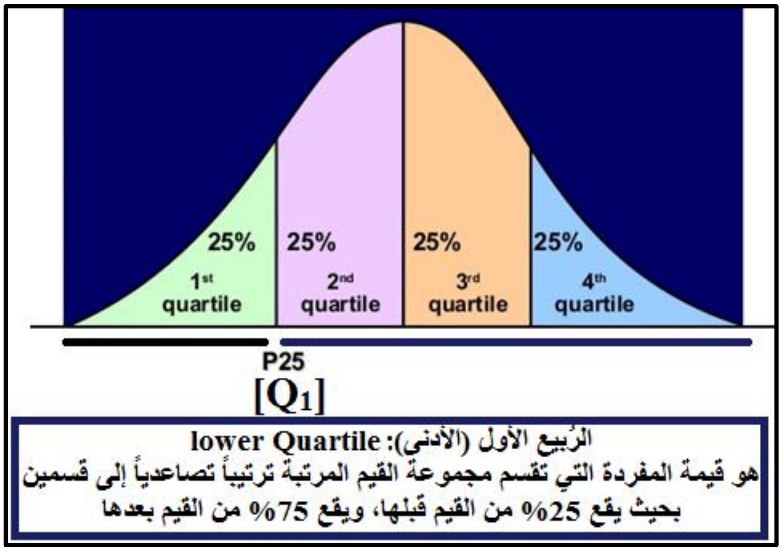



**شكل (11.6)**: يوضح موضع الربيع الأول والوسيط والربيع الأعلى.

**الرُبيع الأول (الأدنى)**: [Q1] **lower Quartile**

وهو القيمة التي تقسم مجموعة القراءات (بعد ترتيبها تصاعدياً) إلى قسمين بحيث يسبقها ربع البيانات ويليها ثلاثة أرباع البيانات.

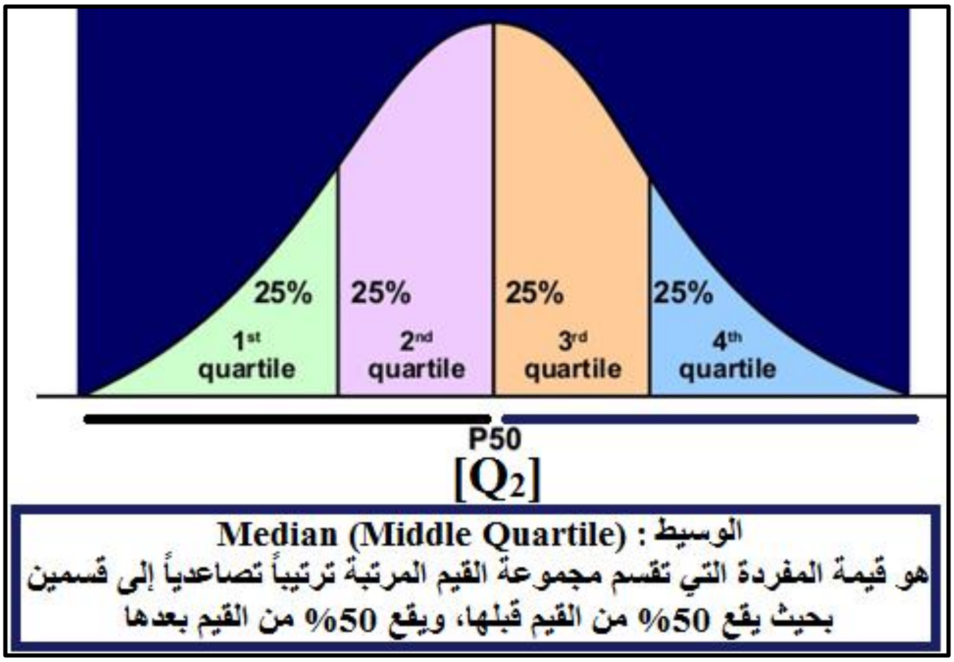
أو هو قيمة المفردة التي تقسم مجموعة القيم المرتبة ترتيباً تصاعدياً إلى قسمين بحيث يقع 25% من القيم قبلها، ويقع 75% من القيم بعدها، أي أنه قيمة المفردة التي تقع في نهاية الربع الأول من القيم المرتبة.




**شكل (12.6)**: يوضح موضع الربيع الأول.

**الرُبيع الثاني (الوسيط)**: $[Q_2]$ **Median(Middle Quartile)**

هو القيمة التي تقسم مجموعة القراءات (بعد ترتيبها تصاعدياً) إلى قسمين بحيث يسبقها نصف البيانات ويليها نصف البيانات أيضاً.

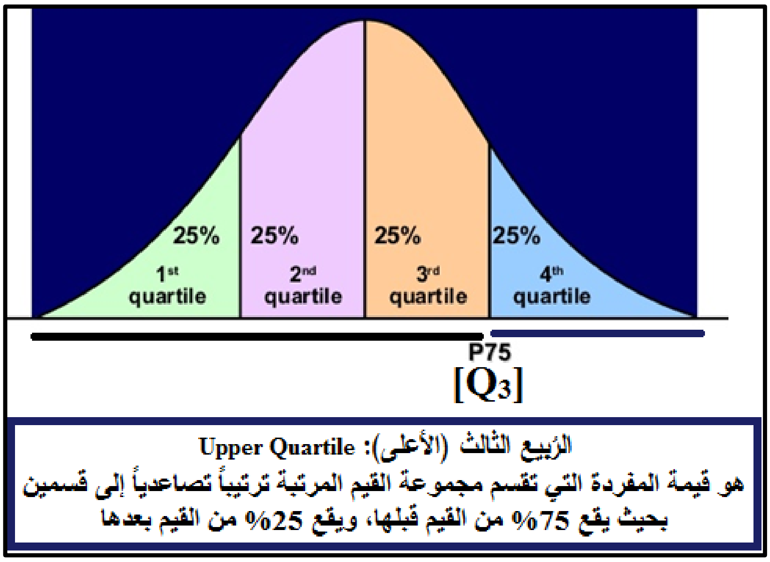



**شكل (13.6)**: يوضح موضع الوسيط (الربيع الثاني).

**الرُبيع الثالث** (**الأعلى**): **[Q3] Upper Quartile**

هو القيمة التي تقسم مجموعة القراءات (بعد ترتيبها تصاعدياً) إلى قسمين بحيث يسبقها ثلاثة أرباع البيانات ويليها ربع البيانات.

أو هو قيمة المفردة التي تقسم مجموعة القيم المرتبة ترتيباً تصاعدياً إلى قسمين بحيث يقع 75% من القيم قبلها، ويقع 25% من القيم بعدها، أي أنه قيمة المفردة التي تقع في نهاية الربع الثالث من القيم المرتبة.




**شكل (14.6)**: يوضح موضع الربيع الثالث (الأعلى).

**حساب الربيعين من البيانات غير المبوبة:**

**أ) حساب الرُبيع الأول (الأدنى) $Q_1$**

**1**. إذا كانت القيم مرتبة تصاعدياً فإن:

If the values are ordered ascending

Quartile location ($Q_1$) = (N+1) / 4

**2**. إذا كانت القيم مرتبة تنازلياً فإن:

If the values are ordered descending

Quartile location ($Q_1$) = 3(N+1) / 4

وذلك مهما كان عدد القيم (n) فردي أو زوجي

لدينا درجات مجموعة من الطلاب 413 ,8 ,5 ,11 ,15 ,9 ,7 , – المطلوب تقدير الربيع الأدنى والربيع الأعلى:

**الحل**: ترتيب الدرجات ترتيب تصاعدي.

| الترتيب | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| الدرجات | 4 | 5 | 7 | 8 | 9 | 11 | 13 | 15 |
| موضع الربيع | | Q1 | | | | | Q3 | |

تعيين موضع الربيع الأدنى:

$$\text{Location } Q_1 = (n+1)/4 = (8+1)/4 = 2.25$$

قيمة الربيع الأدنى:

$$Q_1 = 5 + [0.25 \times (7-5)] = 5.5$$

تعيين موضع الربيع الأعلى:

$$\text{Location } Q_3 = 3(n+1)/4 = 3(9)/4 = 6.75$$

قيمة الربيع الأعلى:

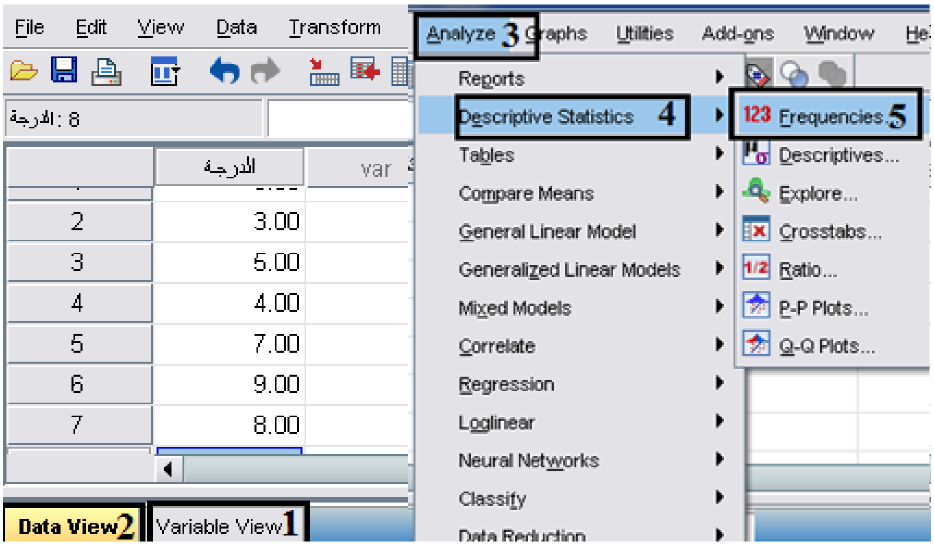
$$Q_3 = 11 + [0.75\,(13-11)] = 11 + [3/4(2)] = 12.5$$

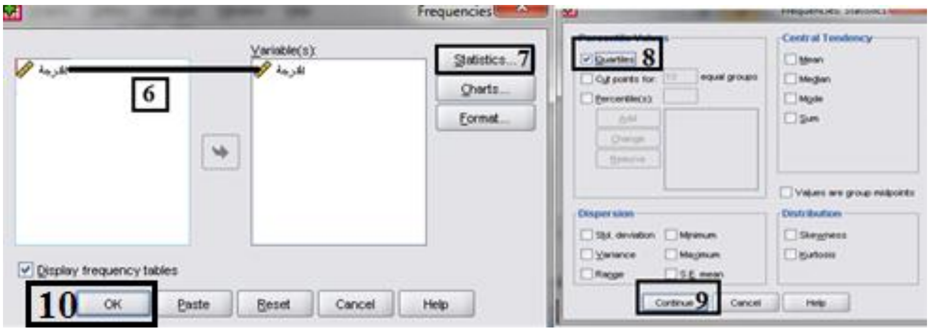
باستخدام الحزمة الإحصائية Spss.

من قائمة Analyze نختار الأمر Descriptive Statistics ثم نختار الأمر Frequencies:

**Analyze** ⇨ **Descriptive Statistics** ⇨ **Frequencies**

فنحصل على النتائج التالية:




| Statistics | | |
|---|---|---|
| الدرجة | | |
| N | Valid | 8 |
| | Missing | 0 |
| Percentiles | 25 | 5.5000 |
| | 50 | 8.5000 |
| | 75 | 12.5000 |

حساب الربيعات والعشيرات والمئينات:

$$\text{الربيع أو العشيري أو المئيني} = \text{الحد الأدنى لفئة الربيع أو العشيري أو المئيني} +$$

$$\left[ \frac{\left[\text{ترتيب}\left(\text{الربيع أو العشيري أو المئيني}\right) - \text{التكرار المتجمع الصاعد السابق}\right]}{\left(\text{التكرار المتجمع الصاعد اللاحق} - \text{التكرار المتجمع الصاعد السابق}\right)} \times \text{عرض فئة} \left(\text{الربيع أو العشيري أو المئيني}\right) \right]$$

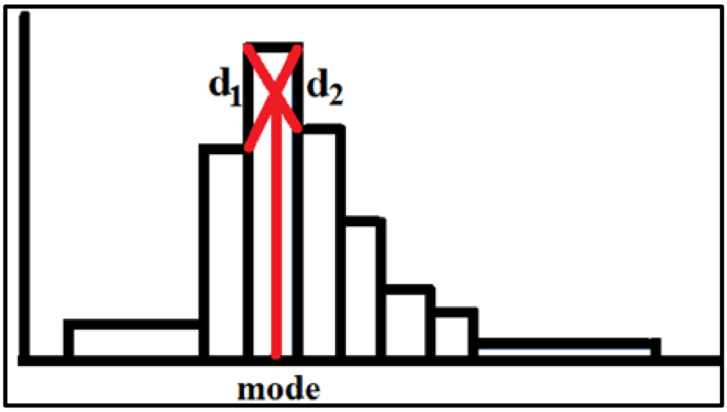
## (4−6) المنوال Mode (THE MODE (Mo)

هو القيمة التي تتكرر أكثر من غيرها من قيم الظاهرة. أو هو القيمة الأكثر شيوعاً أو الأكثر تكراراً بالنسبة لغيرها من قيم المتغير، ويرمز له بالرمز $M_0$.

ويفضل استخدام المنوال في حالة البيانات الوصفية والترتيبية.

The most frequent observation that occurs in the data set.

الملاحظة الأكثر شيوعا التي تحدث في مجموعة البيانات.




**شكل (15.6)**: يوضح موضع المنوال بالرسم.

**ملاحظة**:

يمكن أن تحدد المعنى الدقيق للمنوال من خلال تحديده من مجموعة القيم التالية:

**1)** 1، 5، 7، 9 لا يوجد لها منوال حيث لا توجد أية قيمة تكررت عدداً من المرات يزيد عن تكرار غيرها.

**2)** التوزيع به منوال واحد وهو الأكثر تكراراً

Unimodal: only one value that occurs with the greatest frequency.

كمثال:

27، 9، 7، 12، 15، 9، 10، 4، 7، المنوال= 7 (لأنه تكرر عدد = 2)

**3)** التوزيع به منوالين وهما رقمين أو صفتين كل منهم مكرر مرتين

Bimodal: Two values that occur with the greatest frequency, both values are considered to be the mode.

كمثال:

1، 5، 7، 9، 5، 7 (يوجد منوالان هما 5، 7 لأن كلاً منهما تكرر مرتين)

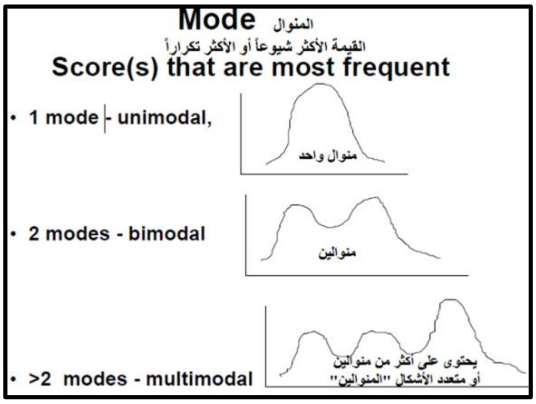
التوزيع متعدد المناويل: به لأكثر من رقم أو صفة لهما نفس التكرار

Multimodal: More than two values that occur with the greatest frequency, each value is used as the model.

كمثال:

7، 5، 6، 12، 8، 10، 1، 4، 6، 10، 5، 2، 3، 4، 11، 8

بما أن كلاً من الأرقام 10، 8، 6، 5، 4 يتكرر مرتين فمن الممكن اعتبار أن هناك خمسة مناويل أو يحتوي على أكثر من منوالين أو متعدد الأشكال "المنوالين". وقد يكون من الأصوب الانتهاء في مثل هذه الحالة إلى القول بعدم وجود منوال كما يتضح من الشكل التالي:




**شكل (17.6)**: يوضح أشكال المنوال.

**أولا: حساب المنوال من البيانات غير المبوبة**

**أ. البيانات الوصفية:**

**مثال**: فيما يلي عينة من 8 أشخاص طبقاً للحالة الاجتماعية:

(متزوج، أرمل، أعزب، مطلق، أعزب، متزوج، أعزب، مطلق) حدد المنوال للبيانات السابقة.

**الحل**:

الأعزب هو الحالة الأكثر تكراراً أو شيوعاً في الحالات الاجتماعية المدروسة وتساوي = 3

**ب. البيانات الوصفية الترتيبية:**

**مثال**: فيما يلي التقديرات في مادة الإحصاء لعينة مكونة من 7 طلاب في السنة الثانية وكانت بالترتيب كما يلي:

| الطلاب | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| **التقدير** | ممتاز | مقبول | ممتاز | جيد | ضعيف | جيد جداً |

**الحل:**

تقدير **ممتاز** هي الحالة الأكثر تكراراً أو شيوعاً في التقديرات المختلفة وتساوي

= 2

**مثال:**

أوجد المنوال للقيم التالية:

14، 14، 13، 15، 11، 15؛13، 10، 12، 13، 14، 13، 14؛15 ؛17؛14، 14، 15.

**الحل:**

| Score | 17 | 15 | 14 | 13 | 12 | 11 | 10 |
|---|---|---|---|---|---|---|---|
| f | 1 | 4 | 6 | 4 | 1 | 1 | 1 |

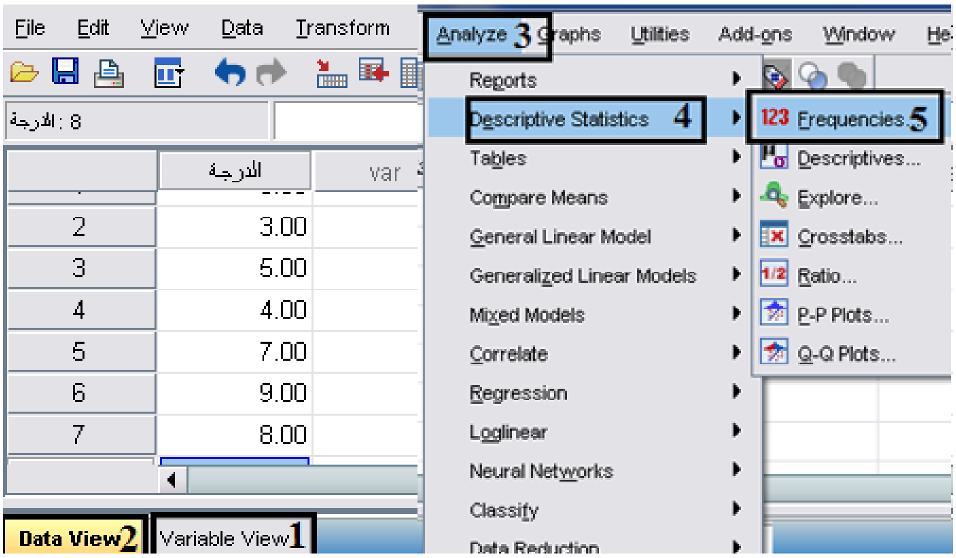
إذن المنوال = 14

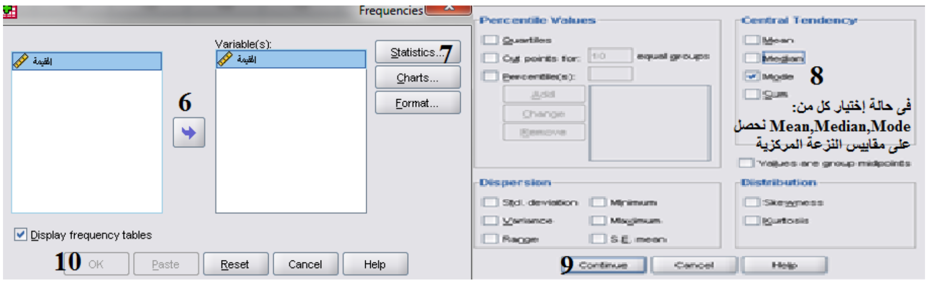
باستخدام الحزمة الإحصائية Spss.

من قائمة Analyze نختار الأمر Descriptive Statistics ثم نختار الأمر Frequencies.

**Analyze** ⇨ **Descriptive Statistics** ⇨ **Frequencies**

فنحصل على النتائج التالية:




فنحصل على النتائج التالية:

| Statistics | | |
|---|---|---|
| القيمة | | |
| N | Valid | 18 |
| | Missing | 0 |
| Mode | | 14.00 |

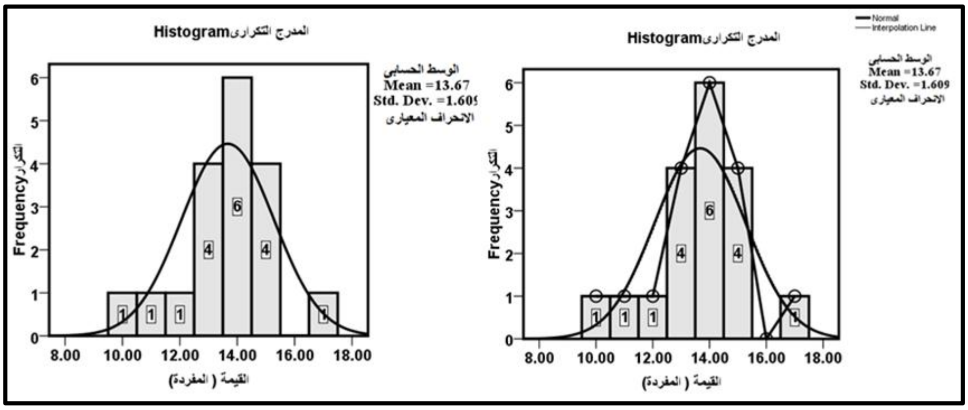
وفي حالة اختيار مقاييس النزعة المركزية نحصل على النتائج التالية:

| Statistics |
|---|

| القيمة | | |
|---|---|---|
| N | Valid | 18 |
| | Missing | 0 |
| Mean | | 13.6667 |
| Median | | 14.0000 |
| Mode | | 14.00 |
| Sum | | 246.00 |

وجدول يحتوي على بيانات توضيحه للقيم (المفردات)

| القيمة | | | | | |
|---|---|---|---|---|---|
| | | Frequency | Percent | Valid Percent | Cumulative Percent |
| Valid | 10 | 1 | 5.6 | 5.6 | 5.6 |
| | 11 | 1 | 5.6 | 5.6 | 11.1 |
| | 12 | 1 | 5.6 | 5.6 | 16.7 |
| | 13 | 4 | 22.2 | 22.2 | 38.9 |
| | 14 | 6 | 33.3 | 33.3 | 72.2 |
| | 15 | 4 | 22.2 | 22.2 | 94.4 |
| | 17 | 1 | 5.6 | 5.6 | 100.0 |
| | Total | 18 | 100.0 | 100.0 | |




**(6 – 5) العلاقة الاعتبارية بين الوسط والوسيط والمنوال:**

**الوسط الحسابي – المنوال = 3 (الوسط الحسابي – الوسيط).**

وهذه العلاقة الاعتبارية تحققها المنحنيات التكرارية وحيدة المنوال.

**مزايا المنوال:**

**1.** هو متوسط موضعي تحدده قيمة المتغير التي لها أكبر تكرار، ولا تدخل في حسابه جميع القيم، ولذلك لا يتأثر بالقيم المتطرفة أو الشاذة.

**2.** لا غبار عليه كأحد مقاييس الموضع (الموقع) خصوصاً إذا كان التوزيع الذي يمثله توزيعاً متماثلاً أما إذا لم يكن كذلك فإن قيمة المنوال تبدو بعيدة عن مركز التوزيع أي بعيدة عن وسطه، ويفقد المنوال بذلك جودته كأحد مقاييس الموقع.

**3.** يمكن حسابه من الجداول التكرارية المفتوحة لعدم اعتماده على مراكز الفئات.

**عيوب المنوال:**

**1.** عادة يصعب تقديره إذا زادت المفردات زيادة كبيرة وتساوت التكرارات الكبيرة في فئات متلاصقة.

**2.** طرق حسابه تقريبية خصوصاً في حالة التوزيعات التكرارية.

**3.** يحسن عدم استخدامه إذا كانت التوزيعات التكرارية مفتوحة لأن الحكم على أكبر تكرار أو صغره يستلزم معرفة طول الفئة.

**4.** تتأثر قيمته بطريقة اختيار فئات التوزيع التكراري وبتغير عدد الفئات ومن ثم يمكن أن تختلف قيمته بين باحث وآخر للتوزيع الواحد تبعاً لاختلاف طول الفئة المختارة.

**5.** لا يصلح استخدامه كممثل للقيم في حالة التوزيعات التكرارية حادة الالتواء.

**6.** لا يعتمد على جميع قيم المتغير موضع البحث ولذلك فهو قليل الحساسية والثبات.

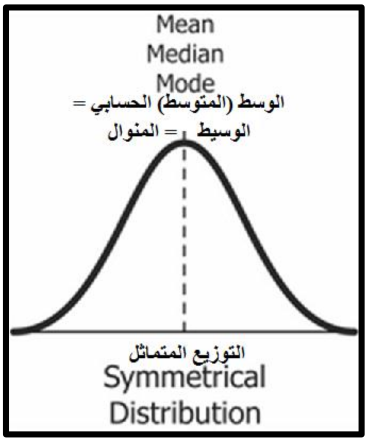
**7.** يقتصر استخدامه في التحليل الوصفي للبيانات.

**(6 – 6) أنواع المنحنيات:**

**أ. حالة المنحنيات المتماثلة:**

تتطابق أي تتساوى قيمة المتوسطات الثلاثة (الوسط" المتوسط" الحسابي = الوسيط = المنوال). وشكل التوزيع يشبه الجرس أو الناقوس المقلوب كما هو موضح في الشكل التالي:
تطابق المتوسطات الثلاث (الوسط = الوسيط = المنوال) حال تماثل (Symmetry)

وتجانس شكل المنحنى تماماً.

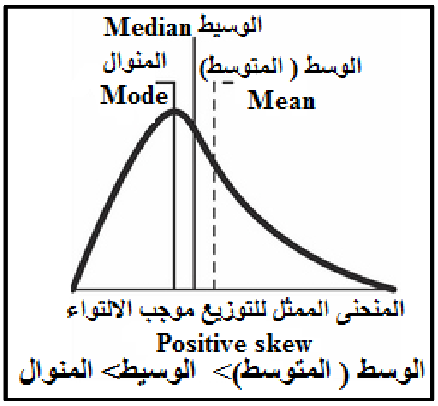



**شكل (18.6)**: يوضح المنحنى المتماثل وفيه (الوسط الحسابي = الوسيط = المنوال).

**ب. حالة المنحنى الممثل للتوزيع موجب الالتواء:**

عندما يكون المنحنى غير متماثل أي مفرطحاً (Skewness) والتفرطح جهة اليمين (+) فإن:

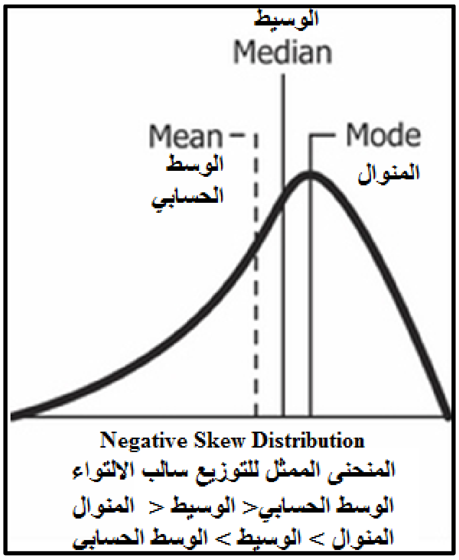
قيمة المنوال (تكون أقل من) < الوسيط (أقل من) < الوسط الحسابي أي الوسط الحسابي > الوسيط > المنوال، كما هو موضح في الشكل التالي:




**شكل (19.6)**: يوضح منحنى الالتواء الموجب وفيه (الوسط" المتوسط" الحسابي > الوسيط > المنوال).

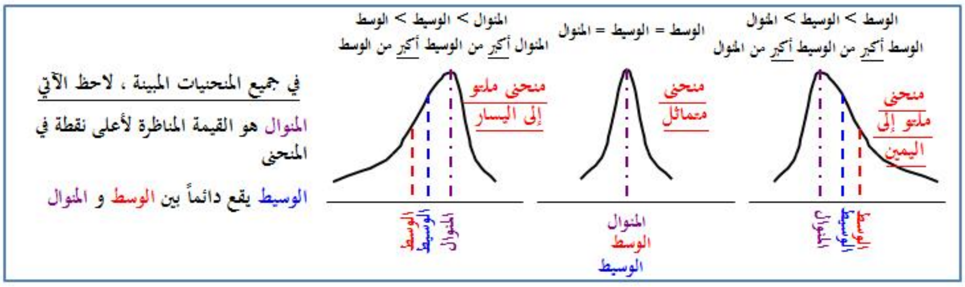
### ج. حالة المنحنى الممثل للتوزيع سالب الالتواء:

عندما يكون المنحنى غير متماثل أي مفرطحاً (Skewness) والتفرطح جهة اليسار (—) فإن قيمة المنوال تكون أكبر من > الوسيط أكبر من > الوسط الحسابي أي (الوسط الحسابي< الوسيط < المنوال) كما هو موضح بالشكل التالي:




**شكل (20.6)**: يوضح منحنى الالتواء السالب وفيه (المنوال > الوسيط > الوسط" المتوسط" الحسابي).

والشكل التالي يلخص الحالات السابقة:



# الباب السابع

## مقاييس التشتت

## Measures of Dispersion


– مقدمة.

(7–1) المدى المطلق.

(7 – 1.1) المدى للبيانات غير المبوبة.

(7–2.1) المدى للبيانات المبوبة.

(7–2) المدى الربيعي.

(7–3) نصف المدى الرُبيعي (الانحراف الربيعي).

(7–4) التباين والانحراف المعياري.

(7–5) معامل الاختلاف.

(7–6) الدرجة المعيارية.

**مقدمة:**

عند مقارنة مجموعتين من البيانات، يمكن استخدام شكل التوزيع التكراري، أو المنحنى التكراري، وكذلك بعض مقاييس الترعة المركزية، مثل الوسط الحسابي والوسيط، والمنوال، والإحصاءات الترتيبية، ولكن استخدام هذه الطرق وحدها لا يكفي عند المقارنة، فقد يكون مقياس النزعة المركزية للمجموعتين متساوياً، وربما يوجد اختلاف كبير بين المجموعتين من حيث مدى تقارب وتباعد البيانات من بعضها البعض، أو مدى تباعد أو تقارب القيم عن مقياس النزعة المركزية.

ومثال على ذلك، أوجد الوسط الحسابي والوسيط لدرجات مجموعة من الطلاب في مادتي الرياضيات واللغة الإنجليزية:

درجات الرياضيات: 40، 58، 72، 80، 100.

درجات اللغة الإنجليزية: 64، 67، 72، 71، 76.

وباستخدام Spss نحصل على النتائج التالية:

| Statistics | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| المادة | الرياضيات | | | | | الإنجليزي | | | | |
| الدرجة | 100 | 80 | 72 | 58 | 40 | 76 | 71 | 72 | 67 | 64 |
| N | 5 | | | | | 5 | | | | |
| Meanالوسط الحسابي | 70.0000 | | | | | 70.0000 | | | | |
| Medianالوسيط | 72.0000 | | | | | 71.0000 | | | | |
| Range المدى | 60.00 | | | | | 12.00 | | | | |
| Minimum | 40.00 | | | | | 64.00 | | | | |
| Maximum | 100.00 | | | | | 76.00 | | | | |
| Sum المجموع | 350.00 | | | | | 350.00 | | | | |

ومما سبق يتضح أن الوسط الحسابي لكل من الظاهرتين هو 70 درجة وأن الوسيط لكل منهما هو71 درجة، فإذا اكتفينا بمقارنة الوسط والوسيط نجد أن مستوى الطلاب هو نفسه في المادتين وهذا يخالف الواقع حيث إن درجات اللغة الإنجليزية متقاربة من بعضها وتتركز حول وسطها (مثلاً) بينما درجات الرياضيات متباعدة ومبعثرة في مدى كبير، وعلى ذلك لا يمكننا اقتصار المقارنة بين الظواهر على متوسطاتها فقط.

وأيضاً لنفرض أن لدينا شريحتين من شرائح المجتمع تعيشان في منطقتين مختلفتين وكانت دخولهم الأسبوعية (بالجنية) هي كما يلي:

المنطقة الأولى A:

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 70 | 75 | 71 | 75 | 74 | 76 | 73 | 78 |

المنطقة الثانية B:

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 99 | 56 | 80 | 100 | 29 | 70 | 65 | 93 |

وبحساب الوسط الحسابي للشريحتين المذكورتين في كل من المجموعتين – حسب ما أوضحناه في الفصل السابق – فإن الوسط الحسابي لدخول المجموعة الأولى:

$$\overline{X}_A = \frac{70+75+71+75+74+76+73+78}{8}$$

$$\overline{X}_A = \frac{592}{8} = 74 \text{ جنيه}$$

أي أن الوسط الحسابي لدخل المجموعة A هو 74 جنية.

وكذلك بالنسبة للمجموعة B فإن الوسط الحسابي لدخولها هو:

$$\overline{X}_B = \frac{99 + 56 + 80 + 100 + 29 + 70 + 65 + 93}{8}$$

$$\overline{X}_B = \frac{592}{8} = 74 \text{ جنيه}$$

والوسط الحسابي لدخل المجموعة B هو أيضاً 74 جنيه.

وباستخدام Spss نحصل على نفس النتائج:

| Statistics | | | |
|---|---|---|---|
| المنطقة | | A | B |
| عدد القيم N | Valid الصحيحة | 8 | 8 |
| | Missing القيم المفقودة | 0 | 0 |
| Mean الوسط الحسابي | | 74.0000 | 74.0000 |
| Range المدى | | 8.00 | 71.00 |
| Minimum أصغر قيمة | | 70.00 | 29.00 |
| Maximum أكبر قيمة | | 78.00 | 100.00 |
| Maximum أكبر قيمة | | 78.00 | 100.00 |
| Percentiles | 25 $Q_1$ | 71.5000 | 58.2500 |
| | 50 $Q_2$ | 74.5000 | 75.0000 |
| | 75 $Q_3$ | 75.7500 | 97.5000 |

ومعنى هذا أن المستوى العام لدخل الأفراد في المجموعتين واحد ويساوي 74 جنيه.

ولكن بإمعان النظر في دخول المجموعة الأولى نجد أنها متجانسة إلى حدٍ كبير أي أنها قريبة جداً من بعضها أو من الوسط الحسابي والذي يساوي 74 جنية، وهنا نقول إن تشتت الدخول قليل أو صغير حيث أن الفرق بين أكبر دخل وأصغر دخل في المجموعة الأولى: 78 – 70 = 8 جنيه.

بينما دخول المجموعة الثانية غير متجانسة فهي بعيدة عن بعضها، أو عن الوسط الحسابي بشكل كبير حيث أن الفرق بين أكبر دخل وأصغر دخل في المجموعة الثانية: 100 – 29 = 71 جنيه.

وهنا نقول إن تشتت دخول المجموعة الثانية كبير فهي أقل تجانساً (أو أكثر تشتتاً) من المجموعة الأولى.

ونستنتج مما سبق بأن الوسط الحسابي ليس كافياً لتوصيف البيانات أو تحليلها كمياً.

وهنا تبرز الحاجة إلى مقاييس كمية أو إحصائية أخرى تبين مدى تقارب أو تباعد مفردات الظواهر بعضها عن بعض.

من أجل ذلك لجأ الإحصائيين إلى استخدام مقاييس أخرى لقياس مدى تجانس البيانات، أو مدى انتشار البيانات حول مقياس النزعة المركزية، ويمكن استخدامها في المقارنة بين مجموعتين أو أكثر من البيانات، ومن هذه المقاييس: مقاييس التشتت، والالتواء، والتفرطح.

فالتشتت لأي مجموعة من القيم يقصد به التباعد أو الاختلاف بين مفرداتها وهذا التشتت يكون صغيراً إذا كان الاختلاف بين قيم المفردات قليلاً (وإذا تساوت جميع القيم فإن التشتت = صفراً) ويكون التشتت كبيراً إذا كان الاختلاف بينها كثيراً أي إذا كانت الفروق بين القيم كبيرة، وعلى ذلك فيمكننا اتخاذ مقدار تشتت القيم كمقياس لتركيز القيم وقربها من بعضها أو لتبعثرها وتباعدها بعضها عن بعض – أي كمقياس لتجانس المجموعات ولا شك أن تجانس المجموعة صفة مهمة يجب دراستها بالإضافة إلى المتوسط إذ لا تغني معرفة المتوسط عنها.

لهذا فإن مقدار التشتت يعتبر مقياساً لقياس تجانس أو تشتت البيانات الإحصائية أو عدم تجانسها في ظاهرة ما.

ويمكن تقسيم مقاييس التشتت إلى نوعين:

**أ. مقاييس التشتت المطلق:**

هي مقاييس لتشتت البيانات ويأخذ المقياس وحدة قياس هي نفس وحدة قياس القيم الأصلية منها (المدى المطلق، نصف المدى الربيعي (الانحراف الربيعي)، الانحراف المعياري، الانحراف المتوسط).

**ب. مقاييس التشتت النسبية:**

هي مقاييس ليس لها تمييز وتمتاز عن المطلق في أنها تصلح للمقارنة بين مجموعتين إذا كانت وحدات القياس في المجموعتين مختلفتين، وهناك عديد من مقاييس التشتت النسبية يمكن استخدامها وإن كان الأكثر شيوعاً هو معامل الاختلاف.

$$\text{معامل الاختلاف}\left(\text{التشتت النسبي}\right) = \frac{\text{مقاييس للتشتت المطلق}}{\text{مقاييس للنزعة المركزية}} \times 100$$

**مقاييس التشتت:**

هو رقم يقيس مدى تباعد وتشتت البيانات بعضها عن بعض أو عن مقاييس النزعة المركزية أي تحديد درجة انحراف البيانات عن القيمة الوسطية.

**(1–7) المدى المطلق Range Absolute**

الفرق بين الحد الأقصى والحد الأدنى.

The difference between maximum and minimum

المدى = أكبر قيمة — أصغر قيمة

Range = highest value − lowest value = Max value – Min value

$$\text{Range} = x\ \max - x\ \min$$

**(7 – 1.1) المدى للبيانات غير المبوبة:**

هو الفرق بين أكبر قيمة وأصغر قيمة في مجموعة من المفردات.

المدى = أكبر قيمة لمجموعة من المفردات — أصغر قيمة لمجموعة من المفردات

**مثال:** احسب المدى للأرقام التالية:

**1.**المجموعة الأولى: 20، 12، **5**، 35، **41**، 18، 30، 22، 8

**2.** المجموعة الثانية: 20، 12، **5**، 35، 41، 18، 30، 22، **589**

**الحل:**

المدى المطلق للمجموعة الأولى = 5 — 41 = 36

المدى المطلق للمجموعة الثانية: = 589 — 5 = 584

**(7–1.2) المدى للبيانات المبوبة:**

يمكن تقديره بإحدى الصيغ الرياضية التالية:

**1.** المدى المطلق =

قيمة الحد الأعلى للفئة الأخيرة في التوزيع — قيمة الحد الأدنى للفئة الأولى في التوزيع.

**2.** المدى المطلق = قيمة مركز الفئة الأخيرة في التوزيع — قيمة مركز الفئة الأولى في التوزيع.

**3**. المدى المطلق = قيمة الحد الحقيقي (الفعلي) للفئة الأخيرة للتوزيع — الحد الحقيقي (الفعلي) للفئة الأولى للتوزيع.

**مثال**: احسب المدى المطلق لهذا التوزيع

| **الفئات** | صفر – | 10 – | 20 - | 30 – 40 | **مجموع** |
|---|---|---|---|---|---|
| **التكرارات** | 5 | 8 | 3 | 4 | 20 |

المدى = الحد الأعلى للفئة الأخيرة في التوزيع — الحد الأدنى للفئة الأولى في التوزيع

$= 40 - 0 = 40$

**خصائصه**:

رغم سهولة حسابه إلا أنه له عيوب منها:

**1**. يعتمد على قيمتين فقط في حسابه use only 2 scores.

**2**. لا يمكن حسابه من جداول مفتوحة لاعتماده على الحد الأعلى للفئة الأخيرة في التوزيع والحد الأدنى للفئة الأولى في التوزيع.

**3**. يتأثر بالقيم المتطرفة uses only 2 most extreme scores مما يؤدي إلى إعطائه صورة غير صحيحة عن درجة تجانس وتوزيع الظاهرة.

والمثال يوضح ذلك:

إذا كانت أعمار أعضاء السلطة التشريعية في بلد ما هي:

| | | | | | | |
|---|---|---|---|---|---|---|
| 72 | 78 | 25 | 74 | 73 | 75 | 70 |

فإن المدى في هذه الحالة هو: 5378 = 25 –

وهنا نلاحظ وجود قيمة شاذة بالنسبة لباقي القيم وهي 25 وهي أصغر قيمة، وإذا أهملت هذه القيمة (أو لم تكن موجودة أصلاً) لكان المدى:

$$78 - 70 = 8$$

وهذا يعني أن وجود قيمة شاذة (25) رفعت قيمة المدى من 8 سنوات إلى 53 سنة. وهذا يوضح مدى حساسية هذا المقياس للقيم الشاذة (أو المتطرفة).

ولكل هذه الخصائص فإن كثيراً من الإحصائيين والباحثين لا يعتمدون كثيراً على المدى كمقياس للتشتت، ويستخدم فقط إذا كان المطلوب فكرة سريعة أو عامة (وليست دقيقة) عن مدى تشتت البيانات.

**(7–2) المدى الربيعي: (IQR) Interquartile range**

المدى الربيعي = الربيع الأعلى − الربيع الأدنى

$$\text{Interquartile range} = Q3 - Q1$$

حيث أن:

**الرُبيع الأول (الأدنى)**: [Q1] **lower Quartile**

**الرُبيع الثالث (الأعلى)**: **Upper Quartile [Q3]**

**(7–3) نصف المدى الرُبيعي (الانحراف الربيعي): ( Quartile Deviation (QD**

يمكن التخلص من العيب الذي يسببه المدى وهو تأثره بالقيم المتطرفة وذلك بأن نستبعد الرُبع الأول من القراءات والرُبع الأخير منها ويُحسب المدى للقراءات الباقية، ونستخدم نصف المسافة بين الرُبيعيين الأدنى والأعلى كمقياس للتشتت في حالة وجود قيم متطرفة ويسمى هذا المقياس بنصف المدى الرُبيعي أو الانحراف الرُبيعي.

هو نصف الفرق بين المدى الربيعي الثالث والمدى الربيعي الأول، ويرمز له بالرمز QD.

يفضل للباحث استخدام الوسيط ويسمى الانحراف الربيعي أيضاً بنصف المدى الربيعي للقانون أعلاه، ويسمى كذلك الربيع الثاني أسوة بالربيع الأول والثالث، وهو أفضل من المدى لأنه لا يتأثر بالقيم المتطرفة مستبعد القيم المتطرفة من الأعلى والأسفل.

نصف المدى الربيعي = [ الربيع الأعلى – الربيع الأدنى ] ÷ **2**

**هو نصف الفرق بين المدى الربيعي الثالث والمدى الربيعي الأول، ويرمز له بالرمز** QD .

$$QD = (Q3 - Q1) \div 2$$

**نصف المدى الرُبيعي (الانحراف الربيعي) من البيانات غير مبوبة:**

**مثال 1:**

لدينا درجات مجموعة من الطلاب 413 ,8 ,5 ,11 ,15 ,9 ,7 , – المطلوب تقدير نصف المدى الربيعي:

**الحل**: ترتيب الدرجات ترتيب تصاعدي.

| **الترتيب** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| **الدرجات** | 4 | 5 | 7 | 8 | 9 | 11 | 13 | 15 |
| **موضع الربيع** | | Q1 | | | | Q3 | | |

تعيين موضع الربيع الأدنى:

$$\text{Location } Q_1 = (n+1)/4 = (8+1)/4 = 2.25$$

قيمة الربيع الأدنى:

$$Q_1 = 5 + [0.25 \times (7 - 5)] = 5.5$$

تعيين موضع الربيع الأعلى:

$$\text{Location } Q_3 = 3(n + 1)/4 = 3(9)/4 = 6.75$$

قيمة الربيع الأعلى:

$$Q_3 = 11 + [0.75(13 - 11)] = 11 + [3/4(2)] = 12.5$$

$$\text{The Interquartile Range (IQR)} = Q_3 - Q_1 = 12.5 - 5.5 = 7$$

**إذن: نصف المدى الرُبيعي (الانحراف الربيعي):** ***Quartile Deviation***

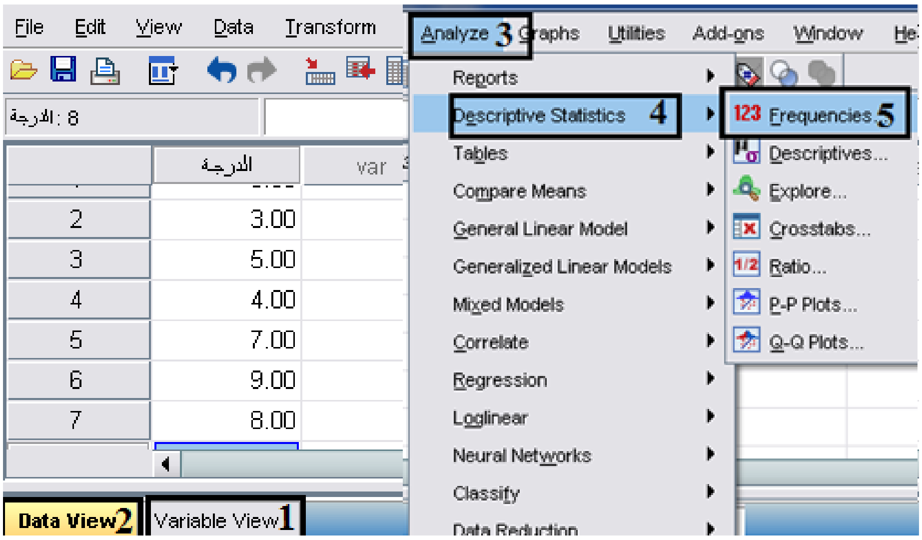
$$QD = (Q_3 - Q_1) \div 2 = (12.5 - 5.5) \div 2 = 3.5$$

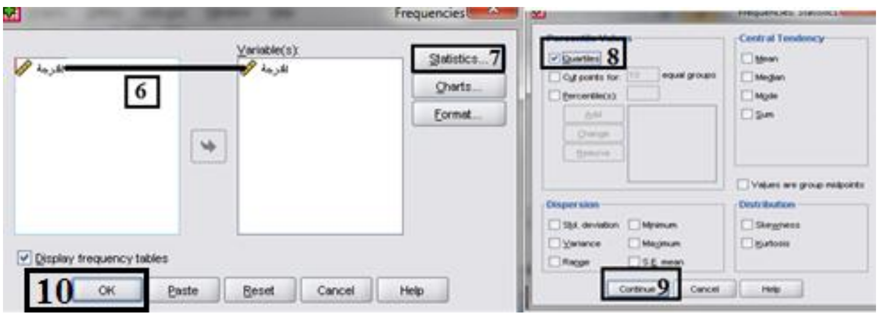
باستخدام الحزمة الإحصائية Spss.

من قائمة Analyze نختار الأمر Descriptive Statistics ثم نختار الأمر Frequencies

**Analyze** ⇨ **Descriptive Statistics** ⇨ **Frequencies**

فنحصل على النتائج التالية:


File
Edit
View
Data
Transform
Analyze 3
Graphs
Utilities
Add-ons
Window
الدرجة : 8
الدرجة
var
2
3.00
3
5.00
4
4.00
5
7.00
6
9.00
7
8.00
Reports
Descriptive Statistics 4
Tables
Compare Means
General Linear Model
Generalized Linear Models
Mixed Models
Correlate
Regression
Loglinear
Neural Networks
Classify
Data Reduction
Frequencies 5
Descriptives...
Explore...
Crosstabs...
Ratio...
P-P Plots...
Q-Q Plots...
Data View 2
Variable View 1


Frequencies
Variable(s):
6
Statistics... 7
Charts...
Format...
Display frequency tables
10
OK
Paste
Reset
Cancel
Help
Quartiles 8
Central Tendency
Continue 9

| Statistics | | |
|---|---|---|
| الدرجة | | |
| N | Valid | 8 |
| | Missing | 0 |
| Percentiles | 25 | 5.5000 |
| | 50 | 8.5000 |
| | 75 | 12.5000 |

**إذن: نصف المدى الرُبيعي** (**الانحراف الربيعي**): **Quartile Deviation**

$$QD = (Q_3 - Q_1) \div 2 = (12.5 - 5.5) \div 2 = 3.5$$

**مثال 2:**

أوجد نصف المدى الرُبيعي (الانحراف الربيعي) لكل من درجات المجموعتين الآتيتين.

المجموعة الأولى: 22، 24، 36، 21، 25، 30، 20، 28.

| 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | الترتيب Serial Number |
|---|---|---|---|---|---|---|---|---|
| 36 | 30 | 28 | 25 | 24 | 22 | 21 | 20 | الدرجات |
| | | $Q_3$ | | | | $Q_1$ | | موضع الربيع Quartile Location |

تعيين موضع الربيع الأدنى:

$$\text{Location } Q_1 = (n + 1)/4 = (8 + 1)/4 = 2.25$$

قيمة الربيع الأدنى:

$$Q_1 = 21 + [0.25 \times (22 - 21)] = 21.25$$

تعيين موضع الربيع الأعلى:

$$\text{Location } Q_3 = 3(n+1)/4 = 3(9)/4 = 6.75$$

قيمة الربيع الأعلى:

$$Q_1 = 28 + [0.75 \times (30 - 28)] = 29.5$$

$$\text{The Interquartile Range (IQR)} = Q_3 - Q_1 = 29.5 - 21.25 = 8.25$$

**إذن: نصف المدى الرُبيعي (الانحراف الربيعي):** ***Quartile Deviation***

$$QD = (Q_3 - Q_1) \div 2 = (29.5 - 21.25) \div 2 = 4.125$$

باستخدام الحزمة الإحصائية Spss.

من قائمة Analyze نختار الأمر Descriptive Statistics ثم نختار الأمر Frequencies

**Analyze** $\Rightarrow$ **Descriptive Statistics** $\Rightarrow$ **Frequencies**

فنحصل على النتائج التالية:

| Statistics | | |
|---|---|---|
| الدرجة | | |
| N | Valid | 8 |
| | Missing | 0 |
| Percentiles | 25 | 21.2500 |
| | 50 | 24.5000 |
| | 75 | 29.5000 |

**إذن: نصف المدى الرُبيعي (الانحراف الربيعي):** ***Quartile Deviation***

$$QD = (Q_3 - Q_1) \div 2 = (29.5 - 21.25) \div 2 = 4.125$$

المجموعة الثانية: 21، 20، 25، 17، 19، 15، 22، 18، 20.

| الترتيب Serial Number | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|
| الدرجات | 15 | 17 | 18 | 19 | 20 | 20 | 21 | 22 | 25 |
| موضع الربيع Quartile Location | | | $Q_1$ | | | | $Q_3$ | | |

$$\text{Location } Q_1 = (n+1)/4 = (9+1)/4 = 2.5$$

$$Q_1 = 17 + [0.5 \times (18 - 17)] = 17.5$$

$$\text{Location } Q_3 = 3(n+1)/4 = 3 * (9+1)/4 = 30/4 = 7.5$$

$$Q_3 = 21 + [0.5 \times (22 - 21)] = 21.5$$

**إذن: نصف المدى الرُبيعي (الانحراف الربيعي):** ***Quartile Deviation***

$$QD = (Q_3 - Q_1) \div 2 = (21.25 - 17.5) \div 2 = 4$$

باستخدام الحزمة الإحصائية Spss.

من قائمة Analyze نختار الأمر Descriptive Statistics ثم نختار الأمر Frequencies

**Analyze** ⇨ **Descriptive Statistics** ⇨ **Frequencies**

فنحصل على النتائج التالية:

| Statistics | | |
|---|---|---|
| الدرجات | | |
| N | Valid | 9 |
| | Missing | 0 |
| Percentiles | 25 | 17.5000 |
| | 50 | 20.0000 |
| | 75 | 21.5000 |

**إذن: نصف المدى الرُبيعي** (**الانحراف الربيعي**): ***Quartile Deviation***

$$QD = (Q_3 - Q_1) \div 2 = (21.25 - 17.5) \div 2 = 4$$

**مميزاته:**

**1.** لا يتأثر نوعاً ما بالقيم المتطرفة.

**2.** نصف المدى الربيعي يمكن حسابه من الجداول المفتوحة والمغلقة.

**3.** نستفيد من الربيعين الأدنى والأعلى والوسيط في التعرف على تماثل التوزيع أو التوائه، ففي حالة التوزيع متماثلاً يكون:

**الربيع الأعلى – الوسيط = الوسيط – الربيع الأدنى.**

$$(Q_3 - Q_2) = (Q_2 - Q_1)$$

**ب**. حالة الالتواء موجباً ناحية اليمين يكون:

**الربيع الأعلى – الوسيط > الوسيط – الربيع الأدنى.**

$$(Q_3 - Q_2) > (Q_2 - Q_1)$$

ج. حالة الالتواء سالباً ناحية اليسار يكون:

**الربيع الأعلى – الوسيط < الوسيط – الربيع الأدنى.**

$$(Q_3 - Q_2) < (Q_2 - Q_1)$$

**عيوبه:**

يعتمد في حسابه على قيمتين فقط ويهمل باقي القيم.

**(7–4) التباين والانحراف المعياري: Variance and standard deviation**

**الانحراف المعياري: standard deviation**

هو الجذر التربيعي الموجب لمتوسط مجموع مربعات انحرافات القيم عن وسطها الحسابي، ويرمز له بالرمز σ للمجتمع، وبالرمز S للعينة.

أي هو الجذر التربيعي الموجب للتباين Standard deviation S is the square root of the variance.

**التباين: Variance**

هو متوسط مجموع مربعات انحرافات هذه القيم عن وسطها الحسابي، ويرمز له بالرمز للمجتمع، وبالرمز $S^2$ للعينة.

ويمكن حساب كل من التباين والانحراف المعياري كما يلي:

| وجه المقارنة | المجتمع | العينة |
|---|---|---|
| التباين<br>$\sigma(Sigma)$ | $\sigma^2 = \frac{\sum_{i=1}^{n}[X_1 - \mu]^2}{N}$ | $S^2 = \frac{\sum_{i=1}^{n}[X_i - \mu]^2}{N}$<br>$S^2 = \frac{\sum_{i=1}^{n}[X_i - \mu]^2}{N}$<br>والصورة المبسطة المعدلة (المصححة) كما في برنامج Spss:<br>$S^2 = \frac{\sum(x - \bar{x})^2}{n-1}$ |
| | | $S^2 = \frac{\sum x^2}{n} - \left[\frac{\sum X}{n}\right]^2$<br>والصورة المبسطة المعدلة (المصححة) كما في برنامج Spss:<br>$S^2 = \frac{\sum x^2 - \left[\frac{\sum X}{n}\right]^2}{n-1}$ |

| وجه المقارنة | المجتمع | العينة |
|---|---|---|
| الانحراف المعياري | $\sigma = \sqrt{\frac{\sum_{i=1}^{n}[X_i - \mu]^2}{N}}$ | $S = \sqrt{\frac{\sum_{i=1}^{n}[X_i - \bar{X}]^2}{n}}$<br>والصورة المبسطة المعدلة (المصححة) كما في برنامج Spss:<br>$S = \sqrt{\frac{\sum(x - \bar{x})^2}{n-1}}$ |
| | | $S = \sqrt{\frac{\sum x^2}{n} - \left[\frac{\sum X}{n}\right]^2}$<br>والصورة المبسطة المعدلة (المصححة) كما في برنامج Spss:<br>$S = \sqrt{\frac{\sum x^2 - \left[\frac{\sum X}{n}\right]^2}{n-1}}$<br>$S = \sqrt{\frac{\sum X^2}{n-1} - \frac{\left[\frac{\sum X}{n}\right]^2}{n-1}}$<br>$S = \sqrt{\frac{\sum X^2}{n-1} - \frac{(\sum X)^2}{n(n-1)}}$ |

**تقدير الانحراف المعياري لبيانات غير مبوبة:**

**حساب التباين والانحراف المعياري من البيانات غير المبوبة:**

إذا فرضنا أن قيم العينة هي: $X_1 , X_2 , \dots , X_n$ (أي عددها n) فإن الخطوات تكون

كما يلي:

1. حساب الوسط الحسابي أي حساب $\bar{X}$ حيث: $\bar{X} = \frac{\sum X}{n}$
2. حساب انحرافات القيم عن الوسط الحسابي أي:

$$X_1\ \bar{X}\ ,\ \ X_n\ \text{-}\bar{X}\ \dots\dots,\ \ X_n\ \text{-}\bar{X}$$

3. تربيع هذه الانحرافات، أي:

$$(X_1 - \bar{X})^2, (X_2 - \bar{X})^2\ , \dots\dots, (X_n - \bar{X})^2$$

4. التباين هو الوسط الحسابي لهذه المربعات، أي:

$$S^2 = \frac{(X_1 - \bar{X})^2 +,(X_2 - \bar{X})^2\ ,+ \cdots,+(X_n - \bar{X})^2}{n}$$

والصورة المبسطة المعدلة(المصححة) المطبقة في برنامج Spss:

$$S^2 = \frac{(X_1 - \bar{X})^2 +,(X_2 - \bar{X})^2\ ,+ \cdots,+(X_n - \bar{X})^2}{n+1}$$

وعليه يمكن حساب كل من التباين والانحراف المعياري لعينة من البيانات غير مبوبة

كما يلي:

| العينة | **التباين** | **الانحراف المعياري** |
|---|---|---|
| | $S^2 = \frac{\sum_{i=1}^{n} [X_i - \bar{X}]^2}{N}$ | $= \sqrt{\frac{\sum_{i=1}^{n} [X_i - \bar{X}]^2}{n}}$ |

| | | |
|---|---|---|
| | والصورة المبسطة المعدلة (المصححة) كما في برنامج Spss: $S^2 = \frac{\sum(x-\bar{x})^2}{n-1}$ | والصورة المبسطة المعدلة (المصححة) كما في برنامج Spss: $S = \sqrt{\frac{\sum(x-\bar{x})^2}{n-1}}$ |
| | $S^2 = \frac{\sum x^2}{n} - \left[\frac{\sum X}{n}\right]^2$ والصورة المبسطة المعدلة (المصححة) كما في برنامج Spss: $S^2 = \frac{\sum x^2 - \left[\frac{\sum X}{n}\right]^2}{n-1}$ | $S = \sqrt{\frac{\sum x^2}{n} - \left[\frac{\sum X}{n}\right]^2}$ والصورة المبسطة المعدلة (المصححة) كما في برنامج Spss: $S = \sqrt{\frac{\sum x^2 - \left[\frac{\sum X}{n}\right]^2}{n-1}}$ $S = \sqrt{\frac{\sum X^2}{n-1} - \frac{\left[\frac{\sum X}{n}\right]^2}{n-1}}$ $S = \sqrt{\frac{\sum X^2}{n-1} - \frac{(\sum X)^2}{n(n-1)}}$ |

**ملاحظات هامة:**

**1)** إن السبب في القسمة على $(n-1)$ عوضاً عن n لأن هناك $(n-1)$ انحرافا مستقلاً من الشكل $x_i - \bar{x}$، ولأن مجموع هذه الانحرافات يساوي الصفر دوماً فإن أياً منها يساوي مجموع كل البقية بإشارة سالبة وأي منها يعطى بدلالة مجموع القيم الأخرى وبإشارة معاكسة، ولتوضيح هذه الفكرة تصور أن لدينا ثلاث بيانات:

$$x_1, x_2, x_3 \quad \text{ولدينا} \quad \bar{x} = \frac{1}{3}(x_1 + x_2 + x_3)$$

نعلم أن: $(x_1 - \bar{x}) + (x_2 - \bar{x}) + (x_3 - \bar{x}) = 0$

وبالتالي يمكن التعبير عن أي منهم وليكن الأول بـ :

$$(x_1 - \bar{x}) = -[(x_2 - \bar{x}) + (x_3 - \bar{x})]$$

أو

$$x_2 - \bar{x} = -[(x_1 - \bar{x}) + (x_3 - \bar{x})]$$

أو

$$x_3 - \bar{x} = -[(x_1 - \bar{x}) + (x_2 - \bar{x})]$$

أي يمكن تمثيل أي انحراف بدلالة الاثنين الآخرين وحيث لدينا $n = 3$ فسنقسم على $3 - 1 = 2$.

**(2)** عندما تكون البيانات كبيرة وغالباً ما تكون كذلك يمكن استخدام علاقة بديلة عن علاقتي التباين، وتسمى العلاقتان البديلتان بالعلاقتين الحسابيتين حيث يمكن حساب التباين (التشتت) منهما بسهولة.

$$s^2 = \frac{1}{n-1}[\sum_{i-1}^{n} x_i^{\,2} - n\bar{x}^2] = \frac{n\sum_{i=1}^{n} x_i^{\,2} - (\sum_{i=1}^{n} x_i)^2}{n(n-1)}$$

$$s^2 = \frac{1}{n-1}[\sum_{1}^{k} x_i^2 f_i - n\bar{x}^2] = \frac{n\sum_{i=1}^{k} x_i^2 f_i - (\sum_{i=1}^{k} x_i f_i)^2}{n(n-1)}$$

وبالطبع العلاقة الأولى للقيم المفردة أما العلاقة الثانية فهي للبيانات المبوبة في جداول توزيع تكرارية ذو k فئة.

يقيس الانحراف المعياري والتباين كمية التباعد الحاصلة في مجموعة بيانات وهذا التباعد يعتمد في الدرجة الأولى على وحدة القياس.

فلمقارنة التباين في عدة مجموعات من البيانات غالباً ما يستخدم التباين النسبي relative variation.

لهذا الغرض أو معامل التباين coefficient of variation الذي يعطي الانحراف المعياري كنسبة مئوية للمتوسط أي:

$$v = \frac{s}{\bar{x}} 100\,\%$$

حيث $\bar{x}$ ، $s$ هما المتوسط والانحراف المعياري على الترتيب لمجموعة من البيانات المراد دراستها.

**مثال:** أوجد الانحراف المعياري لمجموعة الأرقام: 2، 4، 2، 8، 6

**الحل:**

| n | 1 | 2 | 3 | 4 | 5 | $\sum$ |
|---|---|---|---|---|---|---|
| X | 2 | 4 | 2 | 8 | 6 | 22 |
| $X^2$ | 4 | 16 | 4 | 64 | 36 | 124 |

$$S = \sqrt{\frac{\sum x^2}{n} - \left[\frac{\sum X}{n}\right]^2}$$

$$S = \sqrt{\frac{124}{5} - \left[\frac{22}{5}\right]^2}$$

$$S = \sqrt{24.8 - 19.36} = \sqrt{5.44} = 2.33$$

**أو الصورة المعدلة(المصححة) المطبقة في برنامج Spss**

$$S = \sqrt{\frac{\sum X^2}{n-1} - \frac{\left[\frac{\sum X}{n}\right]^2}{n-1}}$$

$$S = \sqrt{\frac{124}{4} - \frac{(22)^2}{5*4}} = \sqrt{31 - 24.2} = \sqrt{6.8} = 2.60768$$

أو

| **n** | X | $\bar{X}$ | $(X - \bar{X})$ | $(X - \bar{X})^2$ |
|---|---|---|---|---|
| 1 | 2 | 22/5 = 4.4 | 2 − 4.4 = −2.4 | −2.4* − 2.4 = 5.76 |
| 2 | 4 | | 2 − 4.4 = −2.4 | 0.16 |
| 3 | 2 | | 4 − 4.4 = −.4 | 5.76 |
| 4 | 8 | | 8 − 4.4 = 3.6 | 12.96 |
| 5 | 6 | | 6 − 4.4 = 1.4 | 2.56 |
| $\sum$ | 22 | | 0 | 27.2 |

$$S = \sqrt{\frac{\sum_{i=1}^{n}[X_i - \bar{X}]^2}{n}}$$

$$S = \sqrt{\frac{27.2}{5}} = \sqrt{5.44} = 2.33$$

**أو الصورة المعدلة(المصححة) المطبقة في برنامج Spss**

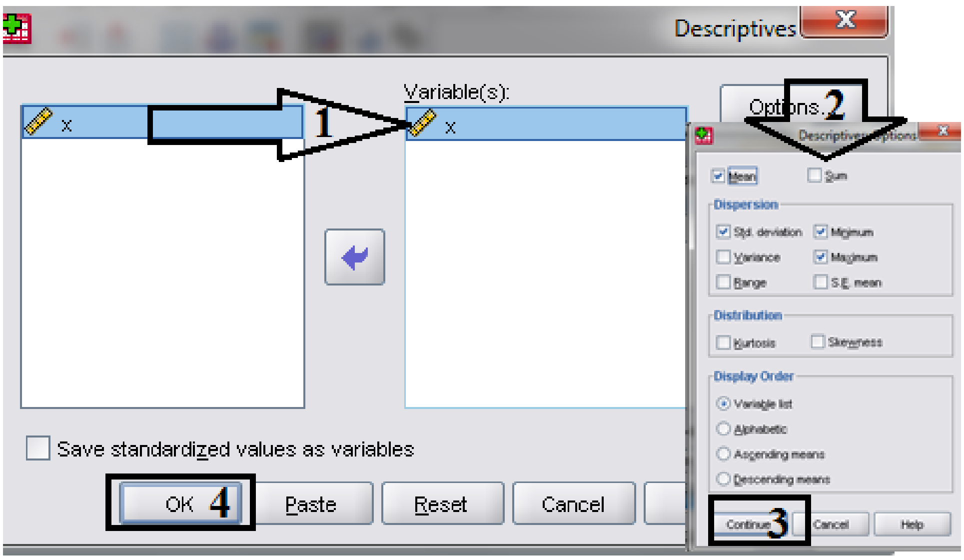
$$S = \sqrt{\frac{\sum(x - \bar{x})^2}{n-1}}$$

$$S = \sqrt{\frac{27.2}{4}} = \sqrt{6.8} = 2.60768$$

وباستخدام الحزمة الإحصائية نتبع المسار التالي:

**Analyze ⇒ Descriptive ⇒ Descriptive**

فيظهر الصندوق الحواري التالي:

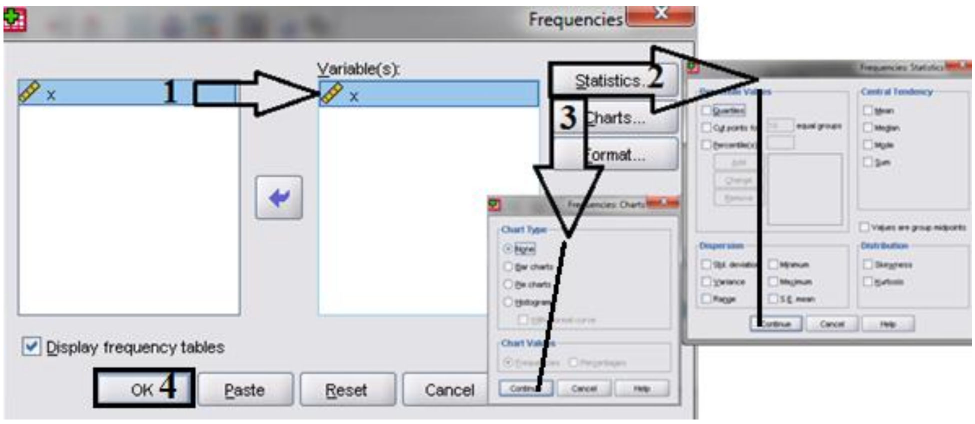



فنحصل على النتائج التالية:

| Descriptive Statistics | | | | | | | |
|---|---|---|---|---|---|---|---|
| | N | Range | Minimum | Maximum | Mean | Std. Deviation | Variance |
| X | 5 | 6.00 | 2.00 | 8.00 | 4.4000 | 2.60768 | 6.800 |
| Valid N (List Wise) | 5 | | | | | | |

أو نتبع المسار التالي:

**Analyze ⟹ Descriptive statistics ⟹ Frequencies**



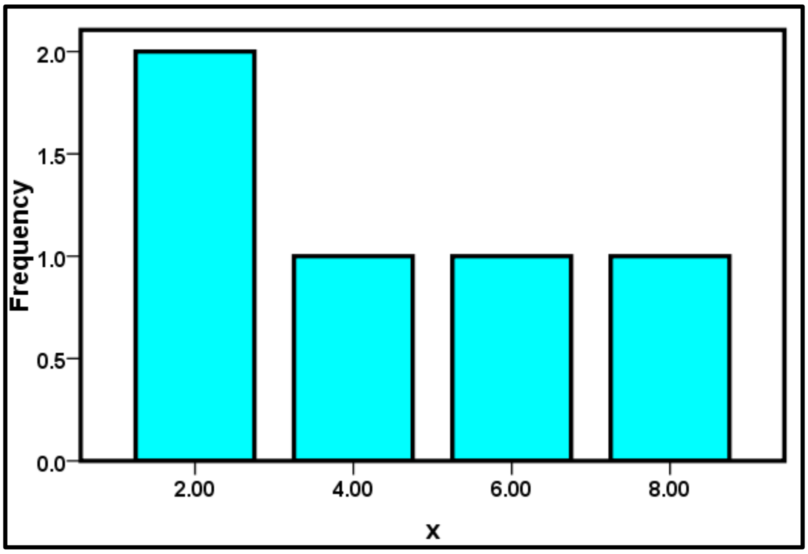

نحصل على النتائج التالية:

| Statistics | | |
|---|---|---|
| x | | |
| N | Valid | 5 |
| | Missing | 0 |
| Std. Error of mean | | 1.16619 |
| Std. Deviation | الانحراف المعياري | 2.60768 |
| Variance | التباين | 6.800 |

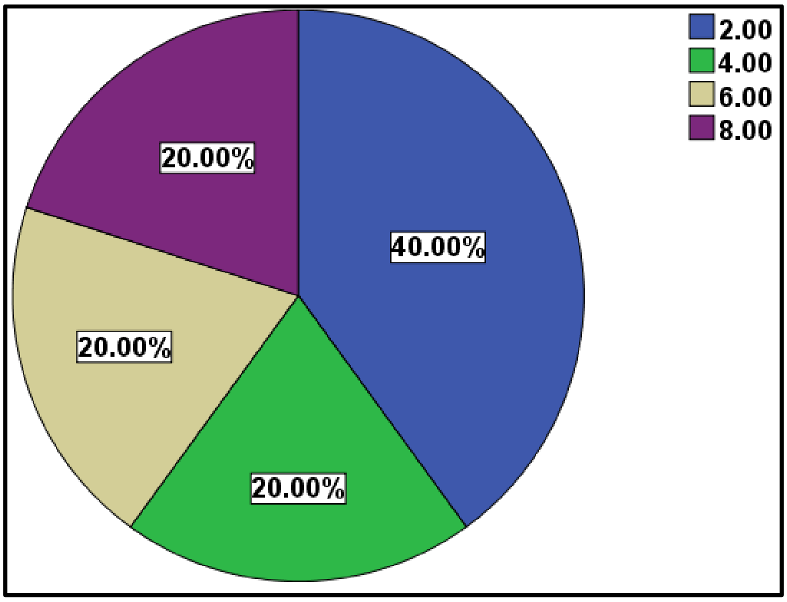
| | | |
|---|---|---|
| Range | المدى | 6.00 |
| | الحد الأدنى<br>Minimum | 2.00 |
| | الحد الأعلى<br>Maximum | 8.00 |

| X | | | | | |
|---|---|---|---|---|---|
| | | **Frequency** | **Percent** | **Valid Percent** | **Cumulative Percent** |
| Valid | 2 | 2 | 40.0 | 40.0 | 40.0 |
| | 4 | 1 | 20.0 | 20.0 | 60.0 |
| | 6 | 1 | 20.0 | 20.0 | 80.0 |
| | 8 | 1 | 20.0 | 20.0 | 100.0 |
| | Total | 5 | 100.0 | 100.0 | |

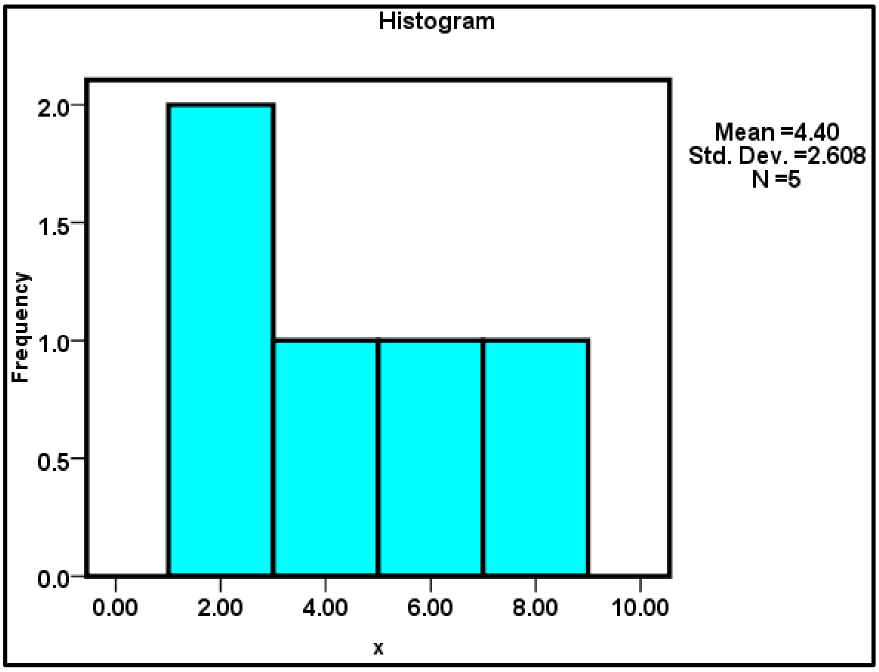
ويمكن تمثيل البيانات بالأعمدة البسيطة كما في الشكل التالي:




أو الدائرة المجزئة للبيانات:




أو المدرج التكراري:




**(7–5) معامل الاختلاف: Coefficient of variation**

كانت مقاييس التشتت التي ذكرت سابقا كلها مقاييس مطلقة، تقدر بدلالة وحدات القياس المستعملة في قياس المتغير الموضوع تحت البحث والدراسة، سواء أكانت هذه الوحدات مقاسه بالسنتيمتر أو المتر أو الكيلو جرام وغيره، وعلى ذلك فإذا أردنا مقارنة عينتين أو مجتمعين فقد يحول دون ذلك اختلاف وحدات القياس المستعملة في كل منهما، كمثال التغير الفعلي أو التشتت كما نحصل عليه من الانحراف المعياري أو غيره من مقاييس التشتت يسمى بالتشتت المطلق. ولكن تغير أو تشتت 1 متر عن مسافة 1000 متر يختلف في تأثيره عن تغير 1 متر في مسافة 20 متر.

فالمقارنة تشتت مجموعتين (أو أكثر) من البيانات وكانت البيانات تختلف في مستواها العام (أي في أوساطها الحسابية) و / أو تختلف في وحدات القياس (مثلاً مقارنة بيانات الدخل حيث تقاس بالريال ببيانات العمر حيث تقاس بالسنوات) فإن المقارنة لا تتم مباشرة بمقارنة الانحراف المعياري لكل منهما بل تتم من خلال مقياس آخر هو "معامل الاختلاف" أو ما يسمى أحياناً مقياس التشتت النسبي حيث ينسب الانحراف المعياري لكل مجموعة إلى وسطها الحسابي والضرب في 100 فنحصل على مقياس نسبي أو مئوي (وبدون تمييز) أي تتم المقارنة بحساب معامل الاختلاف لكل منهما، والمجموعة التي لها معامل اختلاف أكبر تكون أكبر تشتتاً والعكس صحيح ومقياس هذا التأثير للمقارنة بين التغير في الحالتين نحصل عليه بالتشتت النسبي ويعرف بما يلي:

التشتت النسبي = التشتت المطلق ÷ المتوسط (الوسط) الحسابي

**تعريف معامل الاختلاف**:

هو مقياس لمقارنة التشتت لمجموعتين معتمداً على الانحراف المعياري والمتوسط (الوسط) الحسابي.

فإذا كان التشتت المطلق هو الانحراف المعياري $\sigma$، والمتوسط (الوسط) الحسابي هو $\bar{X}$ فإن التشتت النسبي يسمى بمعامل الاختلاف أو معامل التشتت ويتم تقديره بالصيغة الرياضية التالية:

$$\text{معامل الاختلاف} = \frac{\text{الانحراف المعياري}}{\text{الوسط الحسابي}} \times 100$$

$$CV = \frac{\sigma}{x} \times 100$$

وهو معامل ليس له تميز ويصلح للمقارنة بين مجموعات فيها وحدات القياس مختلفة.

والخطوات المتبعة لحسابه:

$$\overline{X} = \frac{\sum X_i}{n}$$

حساب الوسط (المتوسط) الحسابي:

$$S = \sqrt{\frac{\sum X^2}{n} - \left[\frac{\sum X}{n}\right]^2}$$

حساب الانحراف المعياري:

$$CV = \frac{\sigma}{x} \times 100$$

ثم حساب معامل الاختلاف:

**مثال:**

تم اختيار مجموعتين من الطلاب في إحدى المدارس، وتم استخدام برنامج تعليمي معين لتعليم المجموعة الأولى، بينما تم استخدام برنامج تعليمي آخر لتعليم المجموعة الثانية وبعد فترة زمنية تم جمع بيانات عن نتائج المجموعتين، وتم الحصول على المقاييس التالية:

**والمطلوب**: مقارنة درجة تشتت المجموعتين:

| المجموعة الثانية | المجموعة الأولى | البيان |
|---|---|---|
| 198 | 173 | المتوسط (الوسط) الحسابي |
| 25 | 23 | الانحراف المعياري |

**الحل:**

- معامل الاختلاف النسبي للمجموعة الأولى:

$$CV = \frac{\sigma}{X} \times 100 = \frac{23}{173} \times 100 = 13.3$$

- معامل الاختلاف النسبي للمجموعة الثانية:

$$CV = \frac{\sigma}{X} \times 100 = \frac{25}{195} \times 100 = 12.3$$

**الاستنتاج:**

درجة التشتت النسبي للمجموعة الثانية أقل من درجة التشتت النسبي للمجموعة الأولى.

**(6–7) المعايير:**

القيمة الخامة لا تستعمل عادة في المقارنات ويستخدم لهذا الغرض الدرجة المعيارية والمئينية.

**الدرجة المعيارية:standard score**

للمقارنة بين درجتين مختلفتين أُتفق على أن نقيس بُعد كل درجة منهما عن وسطها الحسابي بانحرافها المعياري فنحصل على درجة جديدة قابلة للمقارنة تُعرف بالدرجة المعيارية.

فإذا كان لدينا متغير طبيعي **x** وسطه الحسابي $\overline{X}$، وانحرافه المعياري S فإنه يمكن تحويله إلى متغير طبيعي معياري ص حسب القاعدة:

**الدرجة المعيارية = (الدرجة الأصلية – الوسط "المتوسط" الحسابي) ÷ الانحراف المعياري**

$$Z - \text{Score} = [\text{Raw Score} - \text{Mean}] \div \text{Standard Deviation}$$

$$Z = \frac{X\text{−Mean}}{\text{Standard Deviation}}$$

وتقدر للمجتمع بـ:

$$Z = (X - \mu)/\sigma \quad \text{or} \quad Z = \frac{(X - \mu)}{\sigma}$$

من العينة:

$$Z = (X - \mu)/S \quad \text{or} \quad Z = \frac{(X - \overline{X})}{\sigma}$$

**ملاحظات:**

**1**. الدرجة المعيارية تساوي صفراً في حالة تساوي القيمة بالمتوسط.

2. الدرجة المعيارية موجبة الإشارة إذا كانت القيمة أعلى من المتوسط "الوسط" الحسابي.

**3**. الدرجة المعيارية سالبة الإشارة إذا كانت القيمة أقل من المتوسط "الوسط" الحسابي.

**مثال:**

إذا كانت لدينا درجتان لطالب في مادتي اللغة العربية واللغة الإنجليزية هما 70، و75 وكان متوسط درجات الطلبة في اللغة العربية واللغة الإنجليزية هو 60، 65 والانحراف المعياري لهما هو 2، 4 على التوالي. المطلوب تحديد مستوى الطالب في اللغة العربية واللغة الإنجليزية.

**الحل:**

الدرجة المعيارية للغة العربية = (الدرجة الأصلية — المتوسط "الوسط" الحسابي) ÷ الانحراف المعياري

الدرجة المعيارية للغة العربية:

$$Z = \frac{X-\overline{X}}{S} = \frac{(70-60)}{2} = 5$$

الدرجة المعيارية للغة الإنجليزية = (الدرجة الأصلية — المتوسط "الوسط" الحسابي) ÷ الانحراف المعياري.

الدرجة المعيارية للغة الإنجليزية:

$$Z = \frac{X-\overline{X}}{S} = \frac{(75-65)}{4} = 2.5$$

- إذن مستوى الطالب في اللغة العربية أكبر من مستواه في اللغة الإنجليزية

**مثال:**

إذا كانت لدينا درجتان لطالب في مادتي اللغة العربية واللغة الإنجليزية هما 65، و75 وكان متوسط درجات الطلبة في اللغة العربية واللغة الإنجليزية هو 60، 65 والانحراف المعياري لهما هو 2، 4 على التوالي. المطلوب تحديد مستوى الطالب في اللغة العربية واللغة الإنجليزية.

**الحل:**

الدرجة المعيارية للغة العربية = (الدرجة الأصلية – المتوسط "الوسط" الحسابي) ÷ الانحراف المعياري

الدرجة المعيارية للغة العربية:

$$Z = \frac{X-\overline{X}}{S} = \frac{(65-60)}{2} = 2.5$$

الدرجة المعيارية للغة الإنجليزية = (الدرجة الأصلية — المتوسط "الوسط" الحسابي) ÷ الانحراف المعياري

الدرجة المعيارية للغة الإنجليزية:

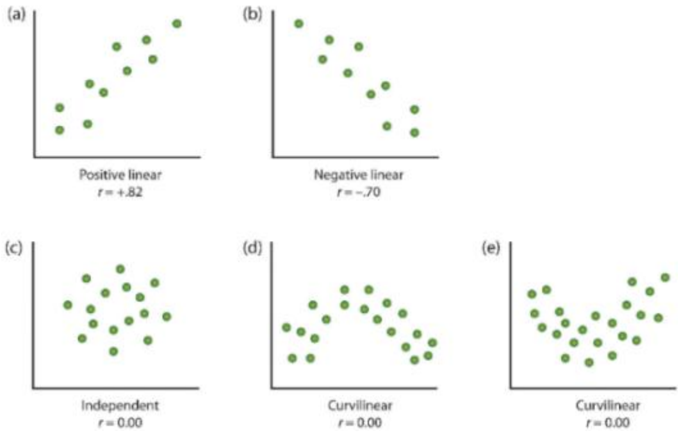
$$Z = \frac{X-\overline{X}}{S} = \frac{(75-65)}{4} = 2.5$$

إذن مستوى الطالب في اللغة العربية يتساوى مع مستواه في اللغة الإنجليزية.

# الباب الثامن

## الارتباط الخطي بين متغيرين

## Linear correlation between two variables




- مقدمة.

(8-1) تحديد العلاقة بين المتغير X، Y.

(8 –1.1) العلاقة الخطية.

(8 – 2.1) العلاقة الخطية (الموجبة).

(8 – 3.1) العلاقة الطردية غير التامة.

(8 – 4.1) العلاقة الخطية الطردية التامة.

(8-5.1) العلاقة الطردية (الموجبة) بين الظاهرتين ولكنها ليست قوية.

(8-6.1) العلاقة الخطية العكسية التامة.

(8-7.1) العلاقة العكسية غير التامة.

(8-8.1) صور متعددة للانتشار الخطي.

(8-2) معامل الارتباط الخطي البسيط.

(8-1.2) معامل ارتباط بيرسون.

(8-2.2) حساب معامل ارتباط بيرسون للبيانات غير المبوبة.

(8-3) الارتباط المتعدد والجزئي.

(8-1.3) معامل الارتباط المتعدد.

(8-2.3) الارتباط الجزئي.

(8-4) معامل ارتباط الرتب لسبيرمان.

(8-5) ارتباط الصفات غير القابلة للترتيب.

(8-1.5) مقياس معامل الاقتران.

(8-2.5) معامل التوافق.

**مقدمة:**

في الفصول السابقة تم عرض بعض المقاييس الوصفية، مثل مقاييس النزعة المركزية، والتشتت، ومقاييس الالتواء والتفرطح، وغيرها من المقاييس الأخرى والتي يمكن من خلالها وصف شكل توزيع البيانات التي تم جمعها عن متغير واحد، وننتقل من التعامل مع متغير واحد إلى التعامل مع متغيرين أو أكثر، ويتناول هذا الفصل دراسة وتحليل العلاقة بين متغيرين، وذلك باستخدام بعض طرق التحليل الإحصائي مثل تحليل الارتباط، والانحدار الخطي البسيط، فإذا كان اهتمام الباحث هو دراسة العلاقة بين متغيرين استخدم لذلك أسلوب تحليل الارتباط، وإذا كان اهتمامه بدراسة أثر أحد المتغيرين على الآخر استخدم لذلك أسلوب تحليل الانحدار، ومن الأمثلة على ذلك:

**1.** وزن الجسم، وضغط الدم.

**2.** الإنفاق، والدخل العائلي.

**3.** سعر السلعة، والكمية المطلوبة منها.

**4.** الفترة الزمنية لتخزين الخبز، وعمق طراوة الخبز.

**5.** تقديرات الطلاب في مقرر الإحصاء، وتقديراتهم في مقرر الرياضيات.

**6.** عدد مرات ممارسة نوع معين من الرياضة البدنية، ومستوى الكولسترول في الدم.

**7.** كميات السماد المستخدمة، وكمية الإنتاج من محصول معين تم تسميده بهذا النوع من السماد.

فإذا كان لدينا متغيران (X، Y) وإذا كان تغير المتغير (X) يؤثر في المتغير (Y)، فإننا نطلق على المتغير (X) اسم المتغير المستقل، والمتغير (Y) اسم المتغير التابع.

أي أن المتغير التابع دالة في المتغير المستقل $Y = f(X)$ .

**طرق دراسة الارتباط بين متغيرين:**

أ) مخطط مبعثر Scatter diagram.

**ب)** معامل كارل بيرسون للارتباط.Karl Pearson's coefficient of correlation

**ج)** سبيرمان الرتبة معامل الارتباط Spearman’s Rank correlation coefficient.

**د)** طريقة المربعات الصغرى Method of least squares.

**ومن مقاييس الارتباط:**

**1.** معامل بيرسون: لقياس درجة الارتباط بين المتغيرات الكمية (الرقمية)، والتي لها توزيع طبيعي.

**2.** معامل سبيرمان لقياس درجة الارتباط بين المتغيرات الرتبية، والمتغيرات الكمية التي ليس لها توزيع طبيعي.

**3.** معامل كندال: لقياس ارتباط الرتب بين المتغيرات التصنيفية.

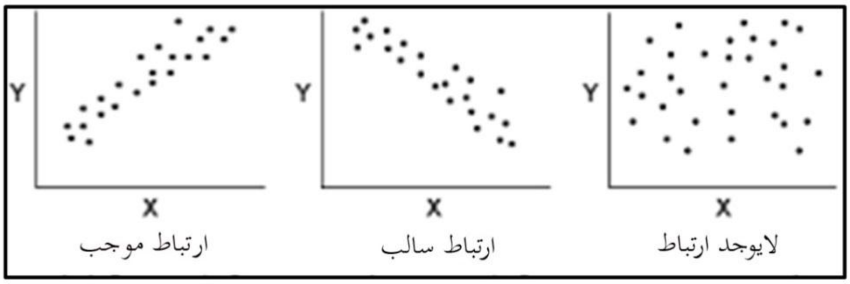
**a.** يصلح لقياس الارتباط من العينات الصغيرة التي يقل حجم العينة عن 10 مفردات.

**b.** يستخدم لقياس الارتباط بين البيانات المزدوجة الرتبية.

**4.** الارتباط بين المتغيرات الوصفية: مربع كاي Chi-Square.

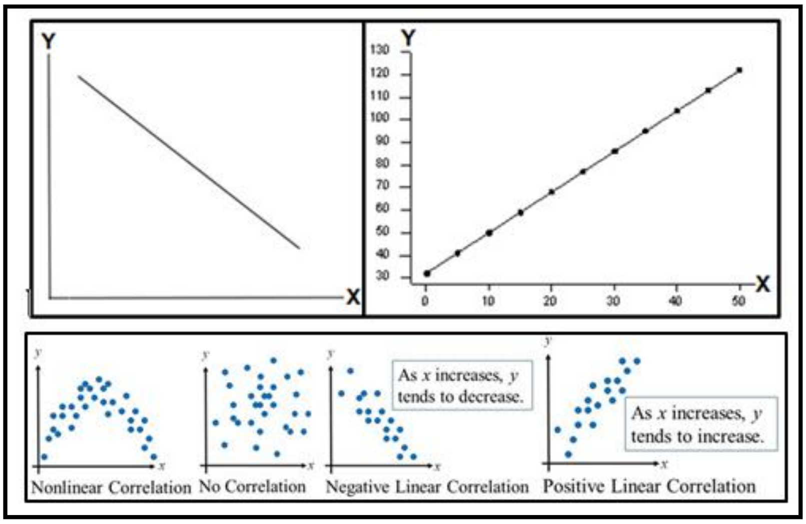
### (8–1) تحديد العلاقة بين المتغير X، Y:

عندما نرسم العلاقة بين المتغيرين (X، Y) على محورين متعامدين أحداهما أفقي ويمثل قيم المتغير (X)، والثاني رأسي ويمثل قيم المتغير (Y) ثم نوقع بتوقيع النقاط على الرسم فنحصل على شكل معين (شكل الانتشار) كما في الشكل التالي:



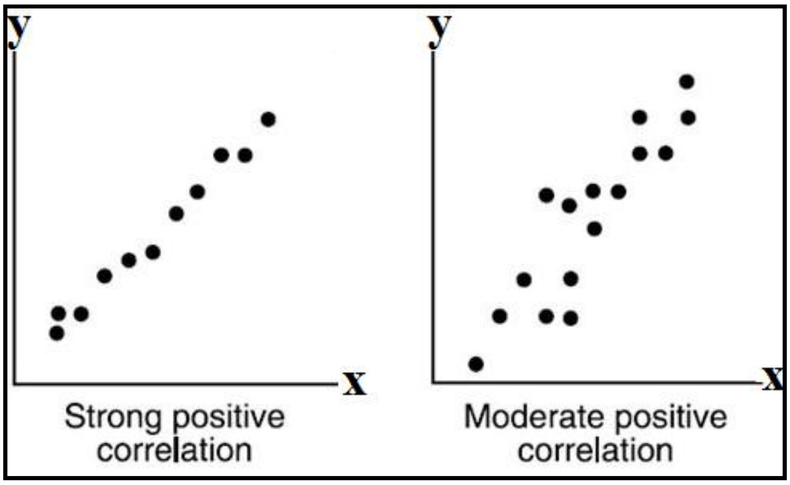

### (8 –1.1) العلاقة الخطية:

إذا تجمعت النقاط حول اتجاه معين فيقال: إن المتغيرين بينهما علاقة خطية كما في الشكل التالي:



**(8 – 2.1) العلاقة الخطية (الموجبة):**

إذا كان تغير أحد المتغيرين يؤدي إلى تغير الآخر في نفس الاتجاه زيادة (نقصان) المتغير المستقل يؤدي لزيادة (نقصان) المتغير التابع كما يلي:




**(8 – 3.1) العلاقة الطردية غير التامة:**

أزواج القيم تتركز في الاتجاه الموجب مما يعني أن هناك علاقة ارتباطية خطية موجبة بين Y، X ولكن غير تامة كمثال البيانات التالية:

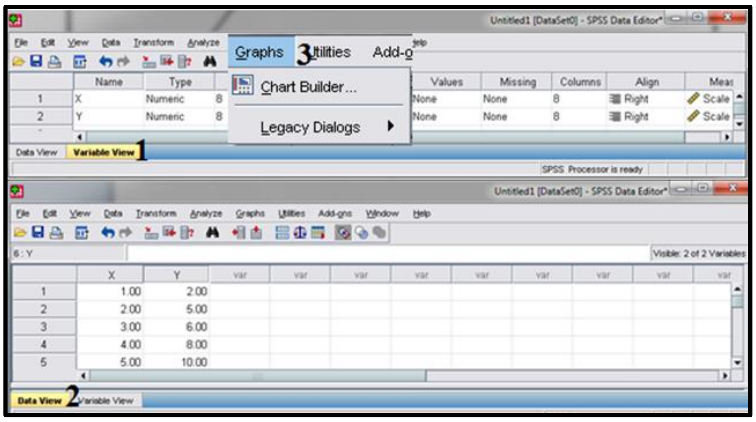
| | | | | | |
|---|---|---|---|---|---|
| x | 1 | 2 | 3 | 4 | 5 |
| y | 2 | 5 | 6 | 8 | 10 |

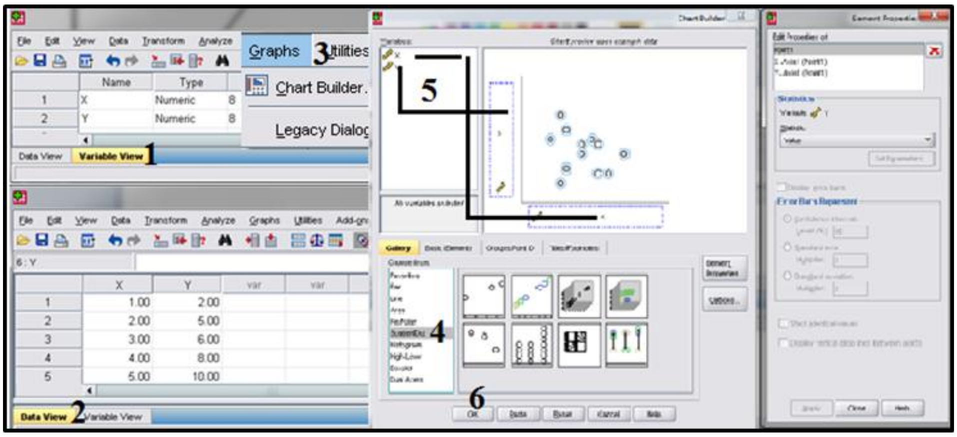
وباستخدام برنامج spss:

نحدد المتغيرات وندحل البيانات:

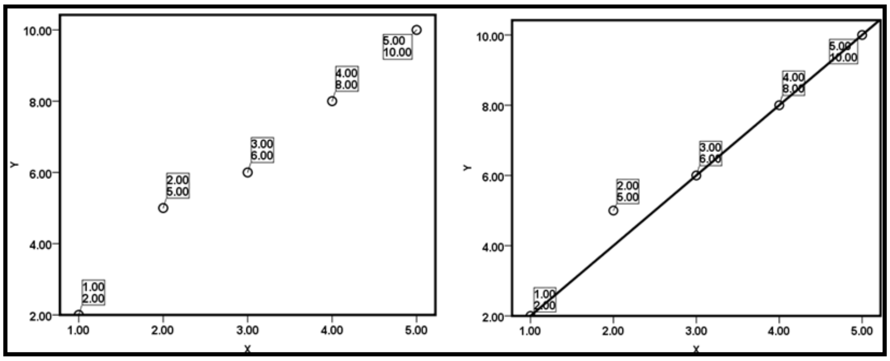
**Graphs ⇨ Chart Builder ⇨ Line**

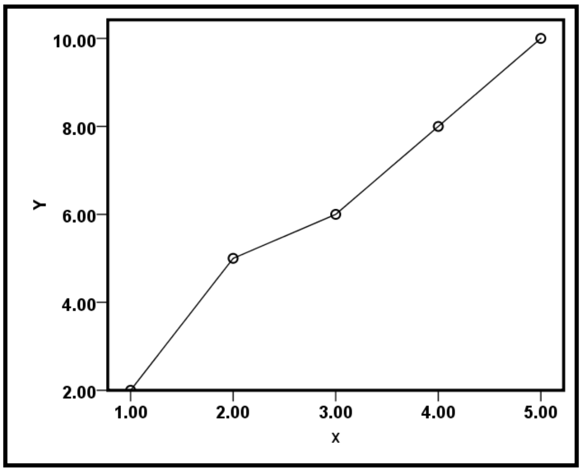
**Graphs ⇨ Chart Builder ⇨ Scatter/Dot**



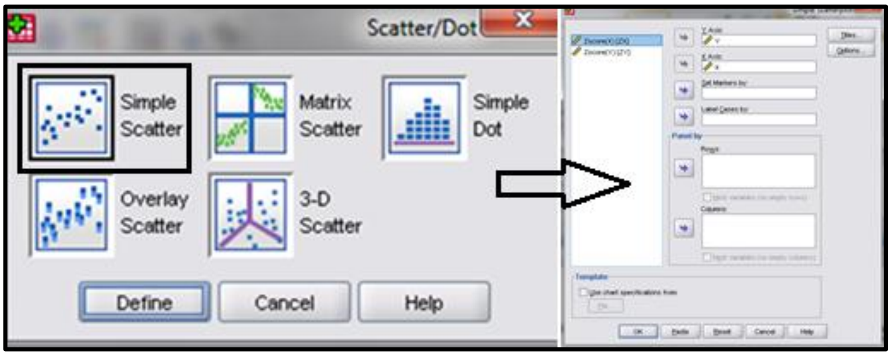

فنحصل على الاشكال التالية:






**Graphs ⇨ Legacy Dialogs ⇨ Interactive ⇨ Scetter/Dot**

فيظهر المربع الحواري التالي ونختار منه Simple Scatter ثم المربع الحواري لتحديد المتغيرات.




فنحصل شكل الانتشار السابق:

من الاشكال السابقة يتضح أن ازدواج القيم تتركز في الاتجاه الموجب مما يعني أن هناك علاقة ارتباطية خطية موجبة بين x، y ولكن لا تقع كل القيم على خط مستقيم مما يدل على وجود علاقة طرية موجبة قوية.

**(8 − 4.1) العلاقة الخطية الطردية التامة:**

أزواج القيم تتركز في الاتجاه الموجب مما يعني أن هناك علاقة ارتباطية خطية موجبة بين Y، X وتقع جميع نقاط أزواج القيم على خط مستقيم واحد في الاتجاه الموجب مما يعني أن هناك علاقة خطية طردية تامة بين x، y، وقيمة معمل ارتباط بيرسون (ر) = + 1؛ (r = 1+) كمثال البيانات التالية:

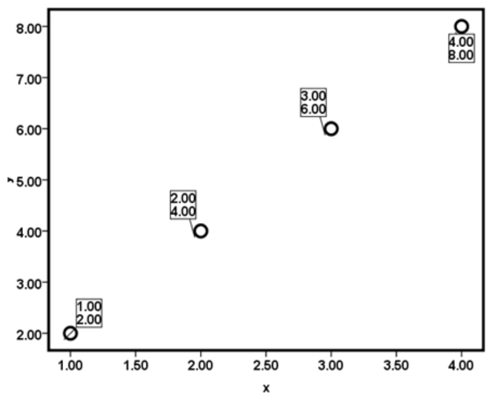
| | | | | |
|---|---|---|---|---|
| x | 1 | 2 | 3 | 4 |
| y | 2 | 4 | 6 | 8 |

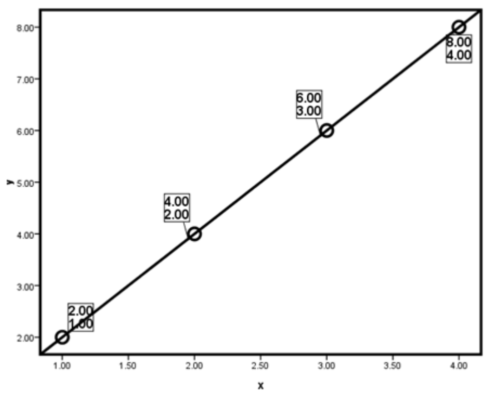
**Graphs ⇨ Chart Builder ⇨ Line**

**Graphs ⇨ Chart Builder ⇨ Scatter/Dot**

**Graphs ⇨ Legacy Dialogs ⇨ Interactive ⇨ Scetter/Dot**

فنحصل على الشكلين التاليين:




**(8−5.1) العلاقة الطردية (الموجبة) بين الظاهرتين ولكنها ليست قوية**

أزواج القيم تتركز في الاتجاه الموجب مما يعني أن هناك علاقة ارتباطية خطية موجبة بين Y، X ولكن لا تقع جميع نقاط أزواج القيم على خط مستقيم واحد مما يعني أن هناك علاقة خطية طردية بين x، y كمثال البيانات التالية:

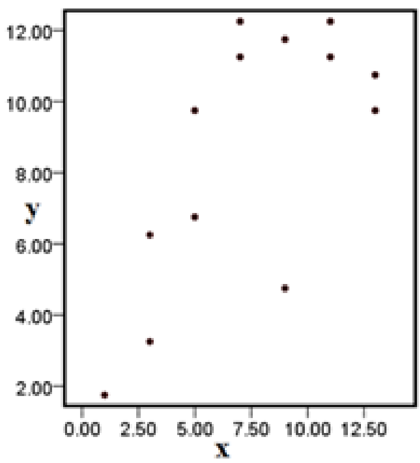
| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| x | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
| y | 2 | 3 | 6 | 7 | 10 | 12 | 11 | 5 | 12 | 11 | 12 | 10 | 12 |

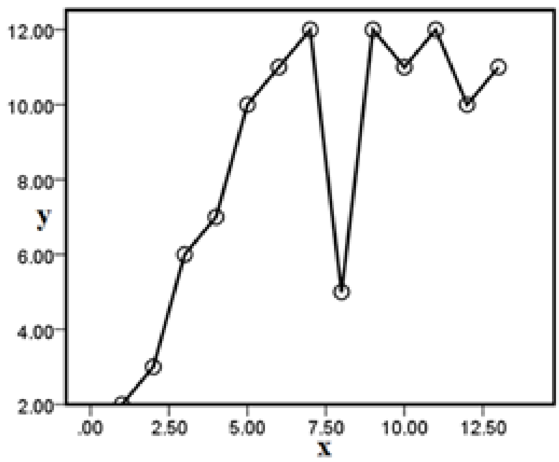
**Graphs ⇨ Chart Builder ⇨ Line**

**Graphs ⇨ Chart Builder ⇨ Scatter/Dot**

**Graphs ⇨ Legacy Dialogs ⇨ Interactive ⇨ Scetter/Dot**

فنحصل على الشكلين التاليين:




يتضح من الشكل السابق أن هناك علاقة خطية موجبة بين الظاهرتين ولكنها ليست قوية نظراً لانتشار نقاط ازدواج قيم x، y.

**(8–6.1) العلاقة الخطية العكسية التامة:**

أزواج القيم تتركز في الاتجاه السالب مما يعني أن هناك علاقة ارتباطية خطية سالبة بين Y، X وتقع جميع نقاط أزواج القيم على خط مستقيم واحد في الاتجاه السالب مما يعني أن هناك علاقة خطية عكسية تامة بين x، y، وقيمة معمل ارتباط بيرسون (ر) = + 1؛ (r = + 1) كمثال البيانات التالية:

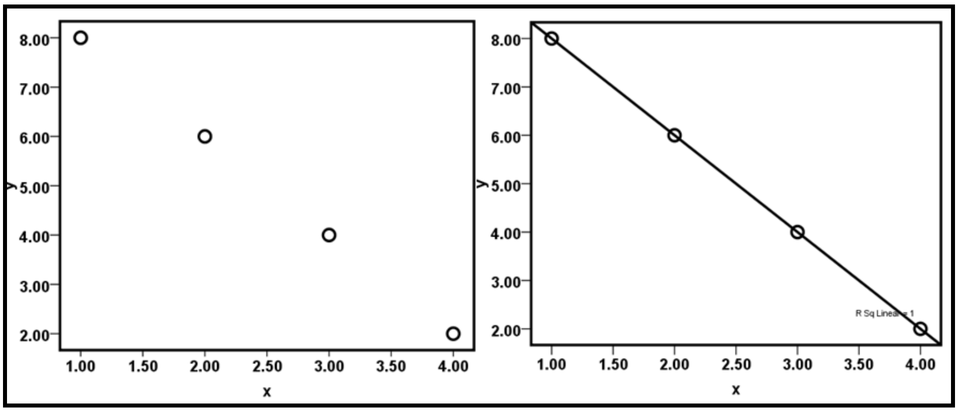
بمعنى إذا كان تغير أحد المتغيرين يؤدي إلى تغير الآخر في عكس الاتجاه فإذا تغير المتغير المستقل زيادة (نقصان) يتبعه تغير التابع إلى نقصان (زيادة) وقيمة معمل ارتباط بيرسون (ر) = – 1؛ (r = 1–) كمثال البيانات التالية:

| | | | | |
|---|---|---|---|---|
| x | 1 | 2 | 3 | 4 |
| y | 8 | 6 | 4 | 2 |

**Graphs ⇨ Chart Builder ⇨ Line**

**Graphs ⇨ Chart Builder ⇨ Scatter/Dot**

كما في الشكل التالي.



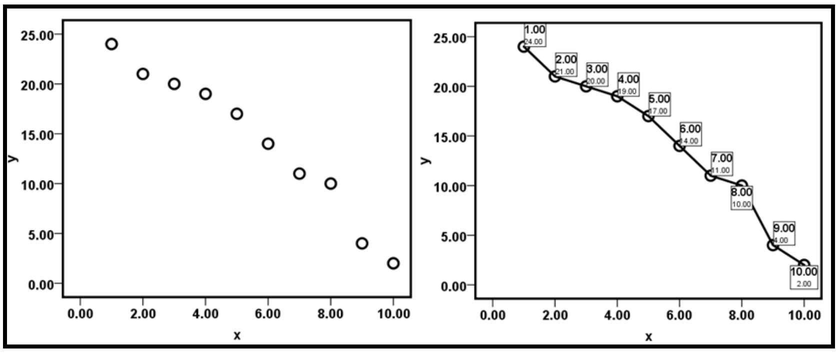

**(7.1–8) العلاقة العكسية غير التامة:**

أزواج القيم تتركز في الاتجاه الموجب مما يعني أن هناك علاقة ارتباطية خطية سالبة بين x، y.

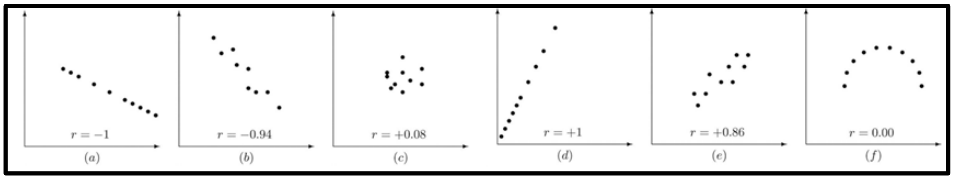
أزواج القيم تتركز في الاتجاه السالب مما يعني أن هناك علاقة ارتباطية خطية سالبة بين X ،Y ولكن لا تقع جميع نقاط أزواج القيم على خط مستقيم واحد مما يعني أن هناك علاقة خطية عكسية بين x، y كمثال البيانات التالية:

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| x | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| y | 24 | 21 | 20 | 19 | 17 | 14 | 11 | 10 | 4 | 2 |

**Graphs ⇨ Chart Builder ⇨ Line**
**Graphs ⇨ Chart Builder ⇨ Scatter/Dot**




**(8.1–8) صور متعددة للانتشار الخطي:**




**الارتباط الخطي البسيط Simple Linear Regression**

في معظم التطبيقات العملية نجد أن هناك علاقة بين متغيرين (أو أكثر)، فمثلاً نجد أن هناك علاقة وارتباط بين درجة الطالب وعدد ساعات الدراسة. يوجد نوعان من المتغيرات هما:

**المتغير التابع: Dependent (Response) Variable**

هو المتغير الذي يقيس نتيجة دراسة ما، وعادة يرمز له بالرمز Y.

**المتغير المستقل: Independent (Explanatory) Variable**

هو المتغير الذي يُفسِّــر أو يســبب التغيرات في المتغير التابع أي هو الذي يؤثر في تقدير قيمة المتغير التابع، وعادة يرمز له بالرمز X. فمثلاً عدد أيام الغياب X ودرجة الطالب في الإحصاء Y، العُمر X والإصابة بضغط الدم Y.

وفي بعض التطبيقات العملية يكون لدينا أكثر من متغيرين تحت الدراسة، فمثلاً قد توجد علاقة خطية بين ضغط الدم وكل من العُمر والوزن، ويسمى الارتباط في هذه الحالة الارتباط الخطي المتعدد.

وعند دراسة العلاقة بين متغيرين X, Y فإن شكل الانتشار Scatter plot يمكن أن يوضح طبيعة هذه العلاقة، وتكون العلاقة بين X, Y قوية جداً إذا وقعت معظم نقاط شكل الانتشار على منحنى أو خط مستقيم، وتكون ضعيفة كلما تناثرت نقاط شكل الانتشار حول منحنى أو خط مستقيم يمر بتلك النقاط.

**(8–2) معامل الارتباط الخطي البسيط: Simple Linear Correlation Coefficient**

**(8–2.1) معامل ارتباط بيرسون : Pearson Correlation Coefficient(r)**

هو مؤشر إحصائي يمكن بواسطته أخذ فكرة عامة عن البيانات من ناحية ما إذا كان هناك ترابط بين المتغيرين (x ،y) أم لا ويرمز له بالرمز r وقيمته $-1 \leq r \leq +1$ .

ومعامل ارتباط بيرسون له شروط لابد من توافرها في توزيع قيم المتغيرين حتى يعطى المدلول العلمي له وهي:

**1.** أن تكون العلاقة بين المتغيرين علاقة خطية، وأن تكون القراءات المسجلة عن المتغيرين المتناظرين (x ،y) هى قراءات عشوائية.

**2.** إذا تغير أحد المتغيرين بمقدار ثابت وتبعه الآخر بمقدار ثابت في نفس الاتجاه يكون معامل الارتباط $+1$ وتكون العلاقة طردية تامة .

**3.** إذا تغير أحد المتغيرين بمقدار ثابت وتبعه الآخر بمقدار ثابت في عكس الاتجاه يكون معامل الارتباط $-1$ وتكون العلاقة عكسية تامة.

**4.** كلما بعدت النقط عن الخط المستقيم كلما نقصت قيمة معامل الارتباط .

**5.** إذا لم تكن هناك علاقة بين المتغيرين فإن معامل الارتباط يكون مساوياً للصفر.

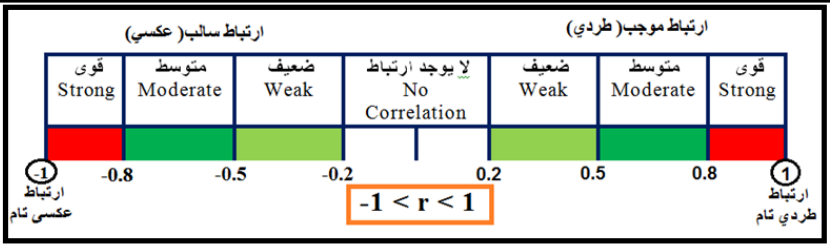
**6.** إشارة معامل الارتباط تشير لاتجاه العلاقة بين المتغيرين، وهي موجبة حين تكون العلاقة طردية، وسالبة حين تكون العلاقة عكسية، وتتوقف إشارة معامل الارتباط على إشارة البسط لأن المقام موجب الإشارة دائماً لأنه حاصل ضرب $\sigma x \times \sigma y$ وكلاهما موجب، أما البسط فهو تغاير x، y وتكون سالبة حين يكون التغير في اتجاهين متضادين.

**7.** يحدد معامل الارتباط الخطي البسيط درجة جودة مطابقة العلاقة بين المتغيرين للعلاقة الخطية وهو على هذا الأساس يحدد نسبة الاختلاف أو التباين في أحد المتغيرين التي يمكن تفسيرها على أساس وجود هذه العلاقة الخطية.

**8.** العلاقة السببية بين (x ،y) لا تؤثر على قيمة معامل الارتباط الخطي أو بعبارة أخرى قيمة معامل الارتباط لا تؤخذ في الاعتبار أي المتغيرات مستقل وأيهما تابع.

**9.** قيمة معامل الارتباط ($\pm 1$) حيث تدل القيمة على قوة الارتباط، وتدل الاشارة على اتجاه العلاقة الارتباطية.

ويمكن تصنيف قوة الارتباط (r) إلى مستويات كما في الجدول التالي:

<table dir="rtl">
<tr><th>قيمة معامل الارتباط</th><th colspan="2">المعنى</th></tr>
<tr><td colspan="3">
</td></tr>
<tr><td colspan="3">اذا كانت قيمة معامل الارتباط صفرا لا يعنى بالضرورة عدم وجود ارتباط بين المتغيرين ولكن قد يوجد ارتباط غير خطى</td></tr>
<tr><td>+1 ارتباط طردي تام</td><td rowspan="6">0<br>لا يوجد<br>ارتباط<br>أو<br>يوجد ارتباط<br>غير خطى</td><td>(من − 0.01 إلى −0.29 ) ارتباط عكسي ضعيف جداً</td></tr>
<tr><td>(من0.90 إلى 0.99) ارتباط طردي قوي جداً</td><td>(من −0.30 إلى 0.49 −) ارتباط عكسي ضعيف</td></tr>
<tr><td>(من 0.70 إلى 0.89) ارتباط طردي قوي</td><td>(من − 0.50 إلى − 0.69) ارتباط عكسي متوسط</td></tr>
<tr><td>(من 0.50 إلى 0.69) ارتباط طردي متوسط</td><td>(من − 0.70 إلى 0.89 −) ارتباط عكسي قوي</td></tr>
<tr><td>(من 00.30 إلى 00.49) ارتباط طردي ضعيف</td><td>(من − 0.90 إلى − 0.99) ارتباط عكسي قوي جداً</td></tr>
<tr><td>(من 00.01 إلى 0.290) ارتباط طردي ضعيف جداً</td><td>−1 ارتباط عكسي تام</td></tr>
</table>

### (8−2.2) حساب معامل ارتباط بيرسون للبيانات غير المبوبة:

يمكن حساب معامل ارتباط بيرسون من باستخدام إحدى الصيغ التالية:

$$r = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2 \sum_{i=1}^{n}(y_i-\bar{y})^2}}$$

$$\mathrm{r}_{XY} = \frac{n\sum XY-\sum X\sum Y}{\sqrt{n\sum X^2-(\sum X)^2}\sqrt{n\sum Y^2-(\sum Y)^2}}$$

**مثال:** أحسب معامل الارتباط الخطي (معامل ارتباط بيرسون) للبيانات التالية:

| | | | | |
|---|---|---|---|---|
| x | 1 | 2 | 3 | 4 |
| y | 2 | 4 | 6 | 8 |

**الحل:**

| n | x | y | xy | $X^2$ | $Y^2$ |
|---|---|---|---|---|---|
| 1 | 1 | 2 | $1 \times 2 = 2$ | $1 \times 1 = 1$ | $2 \times 2 = 4$ |
| 2 | 2 | 4 | $2 \times 4 = 8$ | $2 \times 2 = 4$ | $4 \times 4 = 16$ |
| 3 | 3 | 6 | $3 \times 6 = 18$ | $3 \times 3 = 9$ | $6 \times 6 = 4$ |
| 4 | 4 | 8 | $4 \times 8 = 32$ | $4 \times 4 = 16$ | $8 \times 8 = 4$ |
| $\sum$ | 10 | 20 | 60 | 30 | 120 |

$$\mathrm{r_{XY}} = \frac{n\sum XY - \sum X \sum Y}{\sqrt{n\sum X^2 - \left(\sum X\right)^2}\sqrt{n\sum Y^2 - \left(\sum Y\right)^2}}$$

$$\mathrm{r_{XY}} = \frac{(4 \times 60) - (10 \times 20)}{\sqrt{(4 \times 30) - (10 \times 10)}\sqrt{(4 \times 120) - (20 \times 20)}}$$

$$\mathrm{r_{XY}} = \frac{240 - 200}{\sqrt{120 - 100}\sqrt{480 - 400}}$$

$$\mathrm{r_{XY}} = \frac{40}{\sqrt{20 - 80}} = \frac{40}{40} = 1$$

أو باستخدام:

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}}$$

نكون الجداول التالية:

| المتغير $X$ | | | |
|---|---|---|---|
| n | $X$ | $(x - \bar{x})$ | $(x - \bar{x})^2$ |
| 1 | 1 | 1 – 2.5 = 1.5 | 1.5 × 1.5 = 2.25 |
| 2 | 2 | 2 – 2.5 = -0.5 | -0.5 × -0.5 = 0.25 |
| 3 | 3 | 3 – 2.5 = 0.5 | 0.5 × 0.5 = 0.25 |
| 4 | 4 | 4 – 2.5 = 1.5 | -1.5 × -1.5 = 2.25 |
| $\sum$ | 10 | | 5 |
| $\bar{x}$ | 10 ÷ 4 = 2.5 | | |

| المتغير $y$ | | | |
|---|---|---|---|
| n | y | $(y - \bar{y})$ | $(y - \bar{y})^2$ |
| 1 | 2 | 2 −5 = −3 | -3 × -3 = 9 |
| 2 | 4 | 4 – 5 = -1 | 1 × 1 = 1 |
| 3 | 6 | 6 – 5 = 1 | 3 × 3 = 1 |
| 4 | 8 | 8 – 5 = 3 | 3 × 3 = 9 |
| $\sum$ | 20 | | 20 |
| $\bar{y}$ | 20 ÷ 5 = 20 | | |

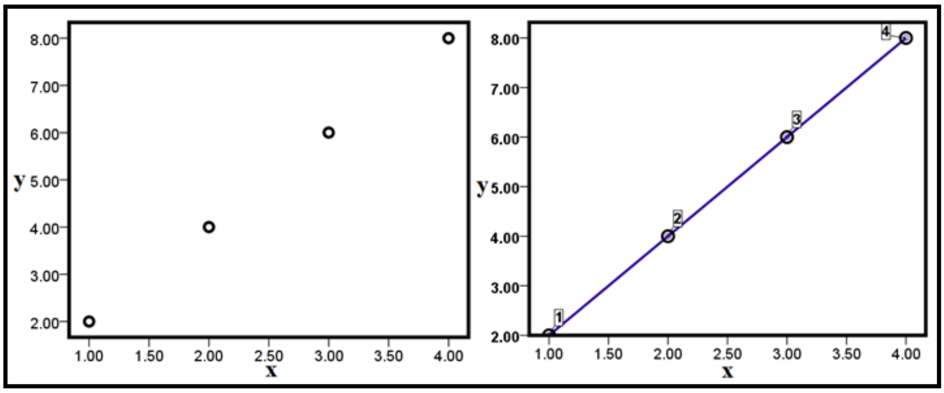
| | |
|---|---|
| n | $(x - \bar{x})(y - \bar{y}))$ |
| 1 | -1.5 × -3 = 4.5 |
| 2 | -0.5 × -1 = 4.5 |
| 3 | 0.5 × 1 = 0.5 |
| 4 | 1.5 × 3 = 4.5 |
| $\sum$ | 10 |

وبتطبيق الصيغة التالية:

$$r = \frac{\sum_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\overline{x})^2 \sum_{i=1}^{n}(y_i-\overline{y})^2}} = \frac{10}{\sqrt{5\times 20}} = \frac{10}{10} = 1$$

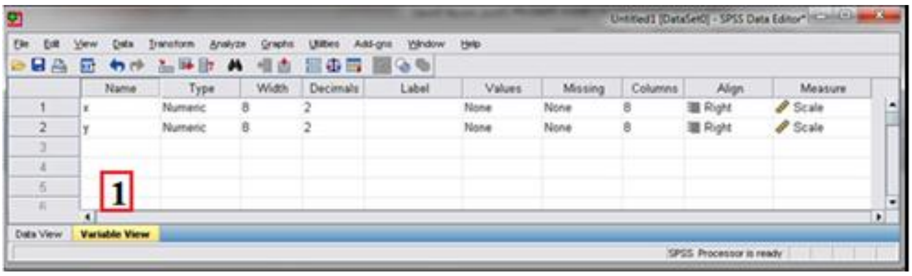
معنى ذلك أن الارتباط طردي تام وهذا يعني أن معدل التغير في قيم س ثابت وأن معدل التغير في قيم ص ثابت.

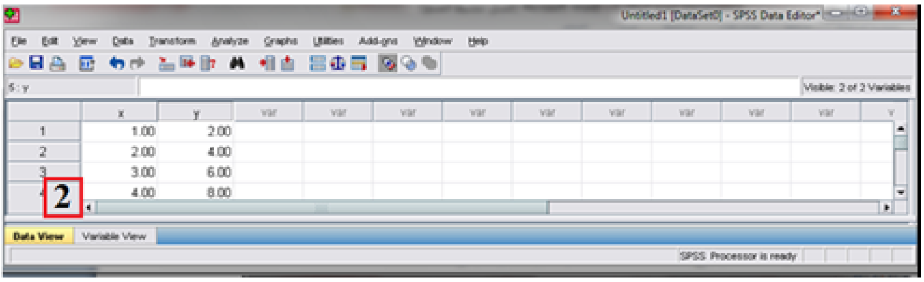
ويتضح ذلك لو رسمنا هذه البيانات فيكون شكل الانتشار لهذه النقاط نجد أنها تقع كلها على خط مستقيم يصنع زاوية حادة مع المحور الأفقي كما في الشكل التالي:




وباستخدام برنامج spss:

نحدد المتغيرات وندحل البيانات:

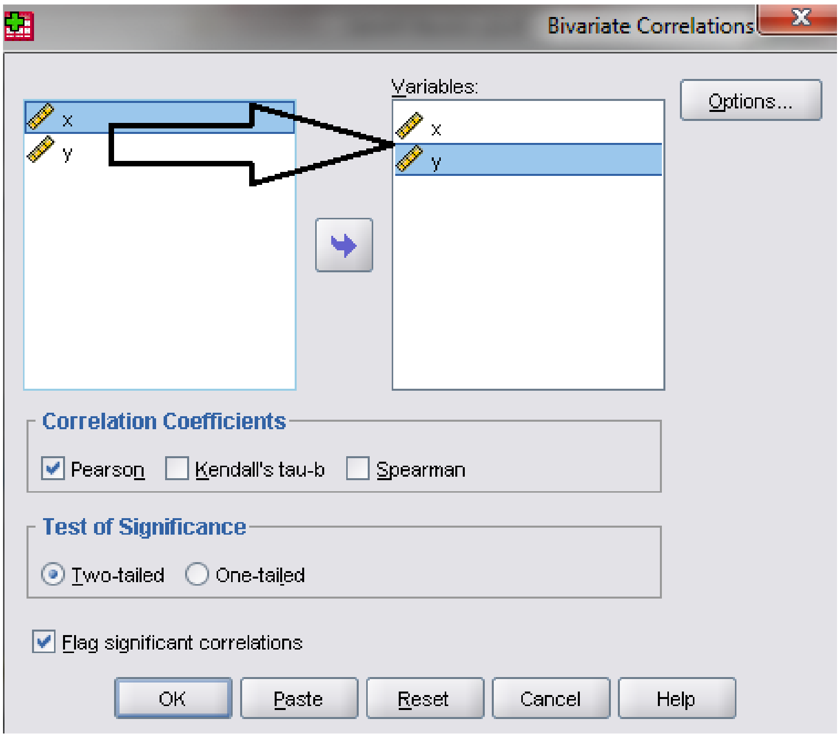





وباتباع المسار التالي:

**Analyze** ⇒ **Correlate** ⇒ **Bivariate**

يظهر الحوار التالي:




وبعد تحديد المتغيرات والضغط على ok نحصل النتائج التالية:

| Correlations | | | |
|---|---|---|---|
| | | x | y |
| x | Pearson Correlation | 1 | 1.000** |
| | Sig. (2-tailed) | | .000 |
| | N | 4 | 4 |
| y | Pearson Correlation | 1.000** | 1 |
| | Sig. (2-tailed) | .000 | |
| | N | 4 | 4 |
| **. Correlation is significant at the 0.01 level (2-tailed). | | | |

من النتائج السابقة يمكن استنتاج ما يلي:

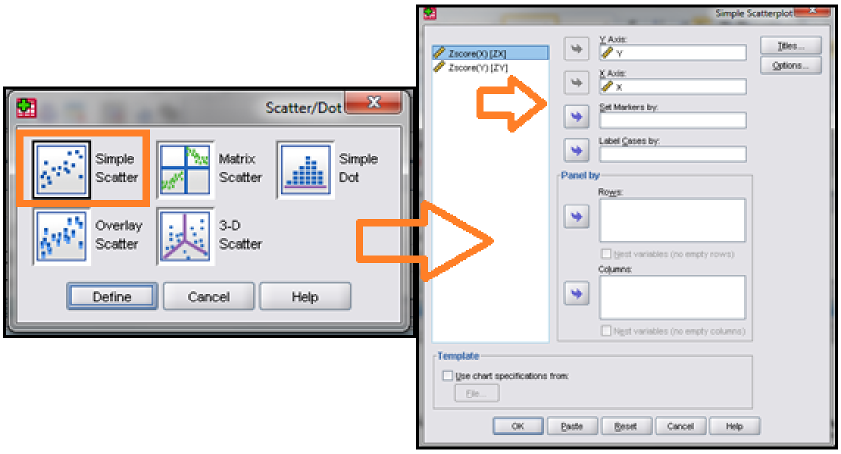
قيمة معامل الارتباط بين (x,y) تساوي 1 وهذا يدل على أن الارتباط بينهما طردي،

Sig.=.000 وهي أقل من 0.05 مما يدل على وجود ارتباط معنوي بين المتغيرين.

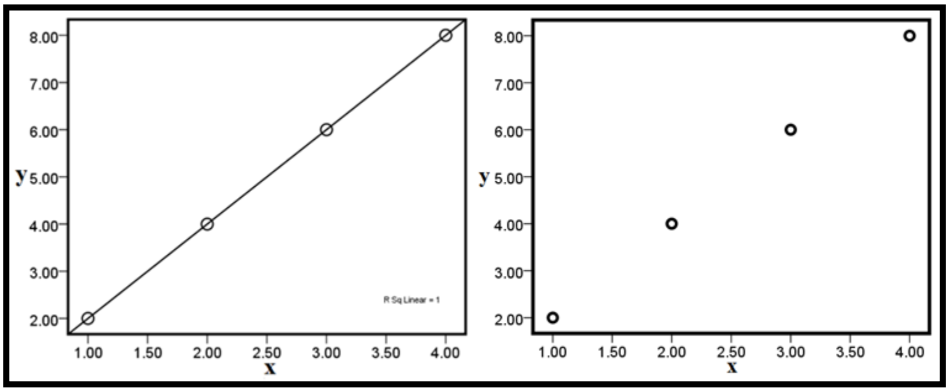
ولرسم شكل الانتشار باتباع الخطوات التالية:

**Graphs ⇨ Legacy Dialogs ⇨ Interactive ⇨ Scetter/Dot**

فيظهر المربع الحوارى التالي ونختار منه Simple Scatter




فنحصل شكل الانتشار التالي:

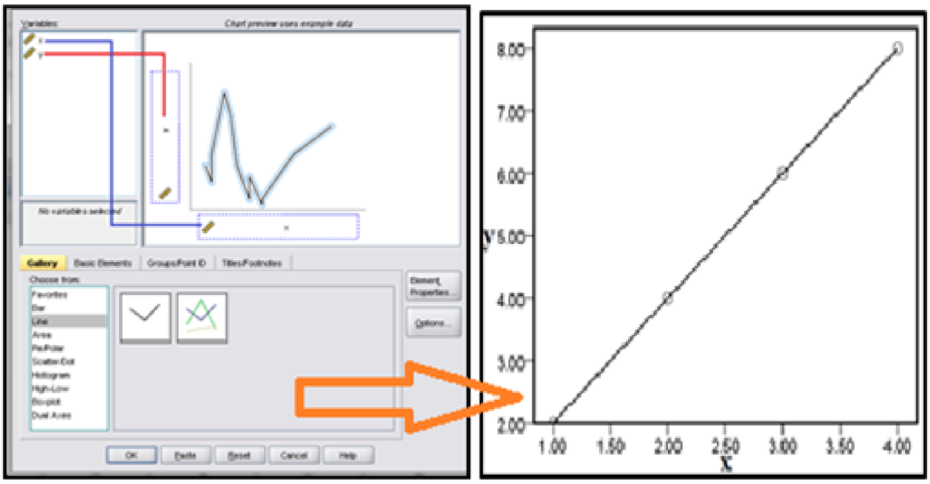



أو:

Graphs ⇨ Chart Builder

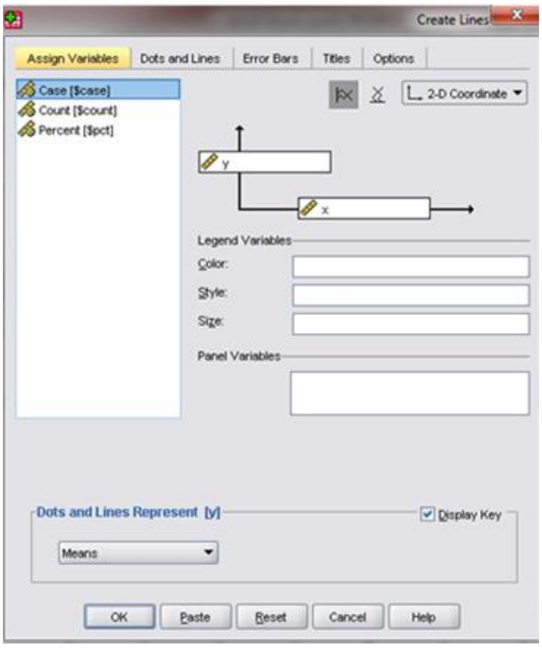
ليفتح مربع الحوار ونختار Line ونحدد المتغيرات على الرسم كما في الشكل التالي:

**Graphs ⇨ Chart Builder**



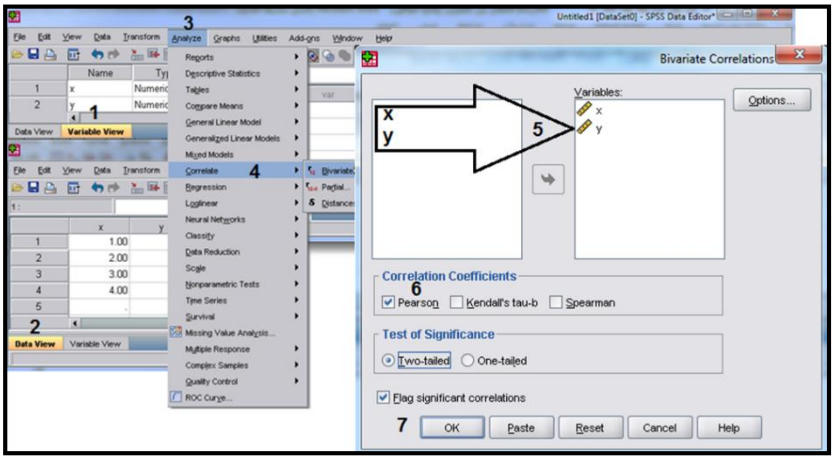

أو:

**Graphs ⇨ Legacy Dialogs ⇨ Interactive ⇨ Line**




فنحصل على نفس المنحى في الشكل الموضح أعلاه.

والشكل التالي يوضح خطوات استخدام الحزمة الاحصائية Spss لحساب معامل الارتباط.



**مثال**: أحسب معامل الارتباط الخطي (معامل ارتباط بيرسون) للبيانات التالية:

| | | | | | |
|---|---|---|---|---|---|
| x | 32 | 25 | 19 | 15 | 12 |
| y | 37 | 37 | 35 | 31 | 33 |

**الحل**:

| n | x | y | $X^2$ | $Y^2$ | xy |
|---|---|---|---|---|---|
| 1 | 32 | 37 | $32 \times 32 = 1024$ | $37 \times 37 = 1369$ | $23 \times 37 = 1184$ |
| 2 | 25 | 37 | $25 \times 25 = 625$ | $37 \times 37 = 1369$ | $25 \times 37 = 925$ |
| 3 | 19 | 35 | $19 \times 19 = 361$ | $35 \times 35 = 1225$ | $19 \times 35 = 665$ |
| 4 | 15 | 31 | $15 \times 15 = 225$ | $31 \times 31 = 961$ | $15 \times 31 = 465$ |
| 5 | 12 | 33 | $12 \times 12 = 144$ | $33 \times 33 = 1089$ | $12 \times 33 = 396$ |
| $\sum$ | 103 | 173 | 2379 | 6013 | 3635 |

$$r_{XY} = \frac{n\sum XY - \sum X \sum Y}{\sqrt{n\sum X^2 - (\sum X)^2}\sqrt{n\sum Y^2 - (\sum Y)^2}}$$

$$r_{XY} = \frac{(5 \times 3635) - (103 \times 173)}{\sqrt{(5 \times 2379) - (103 \times 103)}\sqrt{(5 \times 6013) - (173 \times 173)}}$$

$$r_{XY} = \frac{18175 - 17819}{\sqrt{11895 - 10609}\sqrt{30065 - 29929}} \quad r_{XY} = \frac{356}{\sqrt{1286 - 136}} = \frac{356}{418.2} = 0.851$$

وباستخدام برنامج Spss بنفس الخطوات السابقة نحصل على النتائج التالية:

**Correlations**

| | | x | y |
|---|---|---|---|
| x | Pearson Correlation | 1 | .851* |
| | Sig. (1-tailed) | | .034 |
| | N | 5 | 5 |
| y | Pearson Correlation | .851* | 1 |
| | Sig. (1-tailed) | .034 | |
| | N | 5 | 5 |

*. Correlation is significant at the 0.05 level (1-tailed).

**Correlations**

| | | x | y |
|---|---|---|---|
| x | Pearson Correlation | 1 | .851 |
| | Sig. (2-tailed) | | .067 |
| | N | 5 | 5 |
| y | Pearson Correlation | .851 | 1 |
| | Sig. (2-tailed) | .067 | |
| | N | 5 | 5 |

**مثال**: أحسب معامل الارتباط الخطي (معامل ارتباط بيرسون) للبيانات التالية:

| x | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| y | 8 | 6 | 4 | 2 |

**الحل:**

| n | x | y | xy | $X^2$ | $Y^2$ |
|---|---|---|---|---|---|
| 1 | 1 | 8 | $1 \times 8 = 8$ | $1 \times 1 = 1$ | $2 \times 2 = 4$ |
| 2 | 2 | 2 | $2 \times 6 = 12$ | $2 \times 2 = 4$ | $4 \times 4 = 16$ |
| 3 | 3 | 3 | $3 \times 4 = 12$ | $3 \times 3 = 9$ | $6 \times 6 = 36$ |
| 4 | 4 | 4 | $4 \times 2 = 8$ | $4 \times 4 = 16$ | $8 \times 8 = 64$ |
| $\sum$ | 10 | 20 | 40 | 30 | 120 |

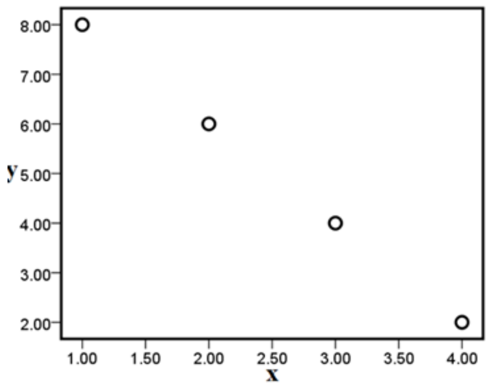
$$r_{XY} = \frac{n\sum XY - \sum X \sum Y}{\sqrt{n\sum X^2 - \left(\sum X\right)^2}\sqrt{n\sum Y^2 - \left(\sum Y\right)^2}}$$

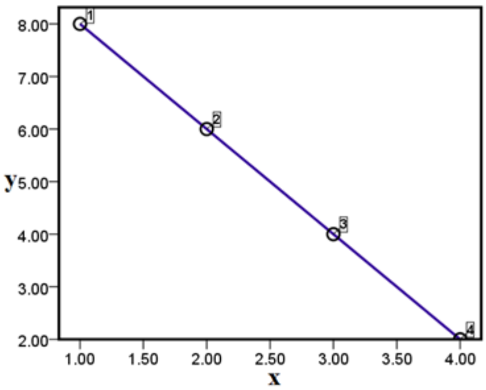
$$r_{XY} = \frac{(4\times40)-(10\times20)}{\sqrt{(4\times30)-(10\times10)}\sqrt{(4\times120-(20\times20)}}$$

$$r_{XY} = \frac{160-200}{\sqrt{120-100}\sqrt{480-400}} \quad r_{XY} = \frac{-40}{\sqrt{20\times80}} = \frac{-40}{40} = -1$$

وباستخدام برنامج spss

وشكل الانتشار:



ويتضح من شكل الانتشار لهذه النقاط أنها تقع كلها على خط مستقيم يصنع زاوية منفرجة مع المحور الأفقي، ومعنى ذلك أن الارتباط عكسى تام وهذا يعني أن معدل التغير في قيم المتغير x ثابت وأن معدل التغير في قيم y ثابت.

**مثال:**

في دراسة للعلاقة بين احترام الذات وطول القامة فكانت النتائج كما في الجدول التالي:

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **احترام الذات** | 4.1 | 4.6 | 3.8 | 4.4 | 3.2 | 3.1 | 3.8 | 4.1 | 4.3 | 3.7 |
| **طول القامة** | 68 | 71 | 62 | 75 | 58 | 60 | 67 | 68 | 71 | 69 |
| **احترام الذات** | 3.5 | 3.2 | 3.7 | 3.3 | 3.4 | 4.0 | 4.1 | 3.8 | 3.4 | 3.6 |
| **طول القامة** | 68 | 67 | 63 | 62 | 60 | 63 | 65 | 67 | 63 | 61 |

**المطلوب:** حساب معامل الارتباط لبيرسون.

**الحل:**

نكون الجدول التالي ونطبق المعادلة التالية:

| رقم الشخص | طول القامة (x) | احترام الذات (y) | $y * y$ | $x * x$ | $x * y$ |
|---|---|---|---|---|---|
| 1 | 68 | 4.1 | 16.81 | 4624 | 278.2 |
| 2 | 71 | 4.6 | 21.16 | 5041 | 326.6 |
| 3 | 62 | 3.8 | 14.44 | 3844 | 235.6 |
| 4 | 75 | 4.4 | 19.36 | 5625 | 330 |
| 5 | 58 | 3.2 | 10.24 | 3364 | 185.6 |
| 6 | 60 | 3.1 | 9.61 | 3600 | 186 |
| 7 | 67 | 3.8 | 14.44 | 4489 | 254.6 |
| 8 | 68 | 4.1 | 16.81 | 4624 | 278.8 |
| 9 | 71 | 4.3 | 18.49 | 5041 | 305.3 |
| 10 | 69 | 3.7 | 13.69 | 4761 | 255.3 |
| 11 | 68 | 3.5 | 12.25 | 4624 | 238 |

| رقم الشخص | طول القامة (x) | احترام الذات (y) | $y*y$ | $x*x$ | $x*y$ |
|---|---|---|---|---|---|
| 12 | 67 | 3.2 | 10.24 | 4489 | 214.4 |
| 13 | 63 | 3.7 | 13.69 | 3969 | 233.1 |
| 14 | 62 | 3.3 | 10.89 | 3844 | 204.6 |
| 15 | 60 | 3.4 | 11.56 | 3600 | 204 |
| 16 | 63 | 4 | 16 | 3969 | 252 |
| 17 | 65 | 4.1 | 16.81 | 4225 | 266.5 |
| 18 | 67 | 3.8 | 14.44 | 4489 | 254.6 |

بالطريقة الحسابية:

$$r = \frac{(\sum f*\sum \mathrm{xy}f)-(\sum \mathrm{x}f*\sum \mathrm{y}f)}{\sqrt{[(\sum f*\sum \mathrm{x}^2 f)-(\sum \mathrm{x}f)^2]*[(\sum f*\sum \mathrm{y}^2 f)-(\sum \mathrm{x}f)^2]}}$$

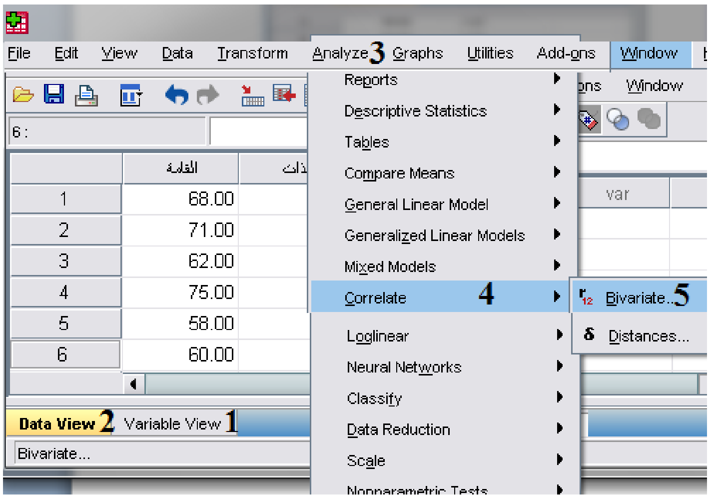
$$r = \frac{(20*4937.6)-(1308*75.1)}{\sqrt{[(20*85912)-(1308)^2]*[(20*285.45)-(75.1)^2]}}$$

$$r = \frac{98752-98230.8}{\sqrt{[1718240-1710864]*[5709-5640.01]}}$$

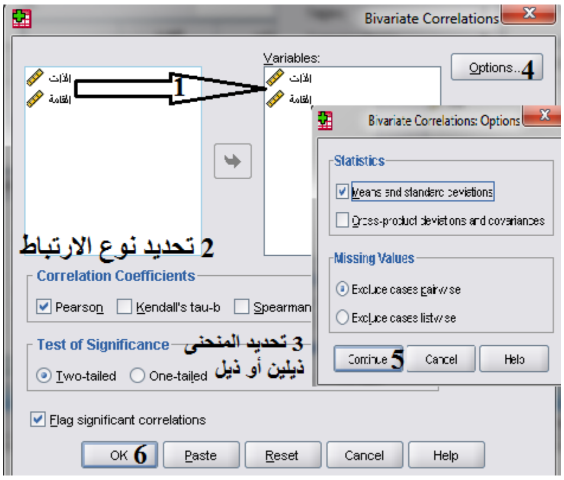
$$r = \frac{521.2}{\sqrt{7376*68.99}}$$

$$r = \frac{521.2}{713.3514} = 0.7306$$

باستخدام Spss: بإدخال البيانات كما في الشكل التالي:



فنحصل على المربع التالي:



فنحصل على الجداول التالية:

| Descriptive Statistics الإحصاء الوصفي للمتغيرات | | | |
|---|---|---|---|
| | Mean | Std. Deviation | N |
| احترام الذات | 3.7550 | .42609 | 20 |
| طول القامة | 65.4000 | 4.40574 | 20 |

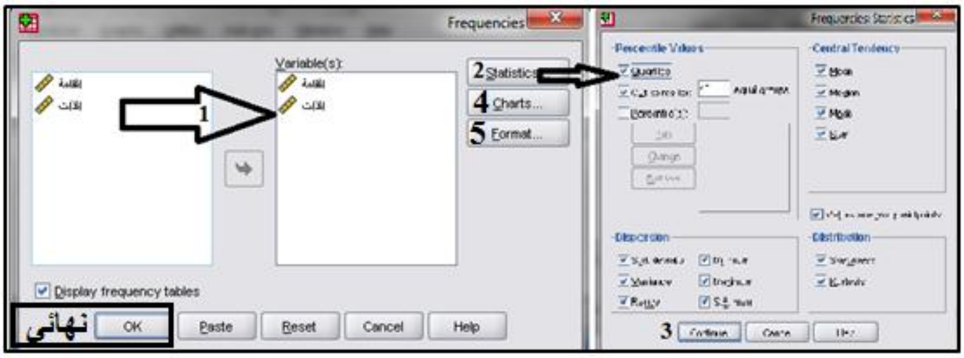
| Correlations الارتباط | | | |
|---|---|---|---|
| | | احترام الذات | طول القامة |
| احترام الذات | Pearson Correlation ارتباط بيرسون | 1 | .731** |
| | Sig. (2-tailed) | | .000 |
| | N عدد المتغيرات | 20 | 20 |
| طول القامة | Pearson Correlation | .731** | 1 |
| | Sig. (2-tailed) | .000 | |
| | N | 20 | 20 |
| ** . Correlation is significant at the 0.01 level (2-tailed) | | | |

ويمكن الحصول على شكل الانتشار كالتالي:

**Analyze ⟹ Descriptive statistics ⟹ Frequencies**

بالضغط على Statistics يمكن تحديد الإحصاء الوصفي، والضغط على Charts

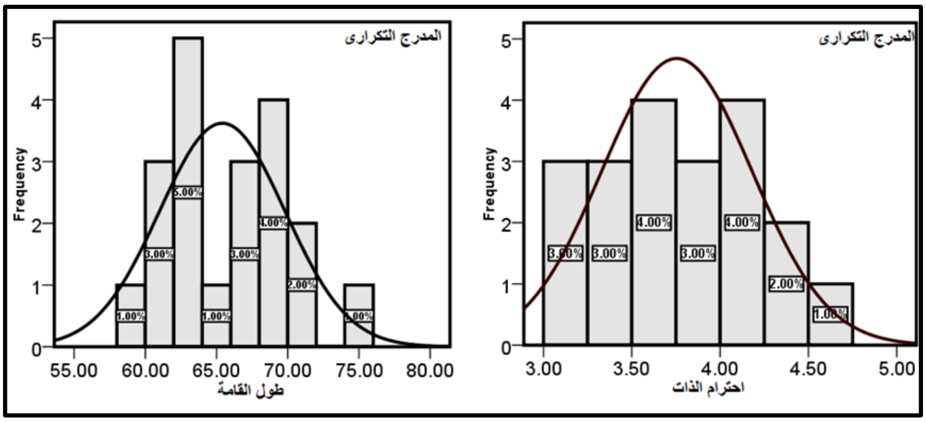
يمكن تحديد الرسم الذى نرغب تمثيل البيانات به، وبالضغط على Format يمكن شكل التكرار.



وبالضغط على Ok نحصل على الجدول التالي:

| Statistics | | | |
|---|---|---|---|
| | | طول القامة | احترام الذات |
| N | Valid | 20 | 20 |
| | Missing | 0 | 0 |
| Mean | | 65.4000 | 3.7550 |
| Std. Error of Mean | | .98515 | .09528 |
| Median | | 6.5500E1[a] | 3.7400[a] |
| Mode | | 63.00[b] | 3.80[b] |
| Std. Deviation | | 4.40574 | .42609 |
| Variance | | 19.411 | .182 |
| Skewness | | .289 | .260 |
| Std. Error of Skewness | | .512 | .512 |
| Kurtosis | | -.428- | -.755- |
| Std. Error of Kurtosis | | .992 | .992 |
| Range | | 17.00 | .501 |
| Minimum | | 58.00 | 3.10 |
| Maximum | | 75.00 | 4.60 |
| Sum | | 1308.00 | 75.10 |
| a. Calculated from grouped data. | | | |
| b. Multiple modes exist. The smallest value is shown | | | |

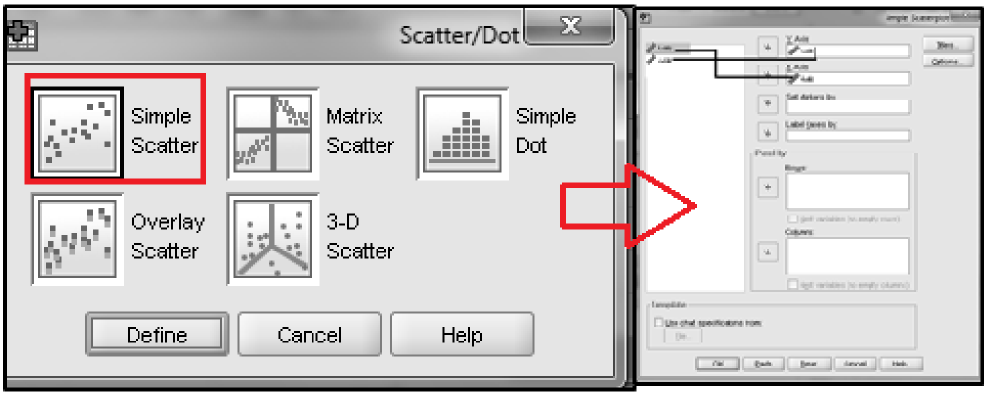
ونحصل على الرسم البياني التالي:



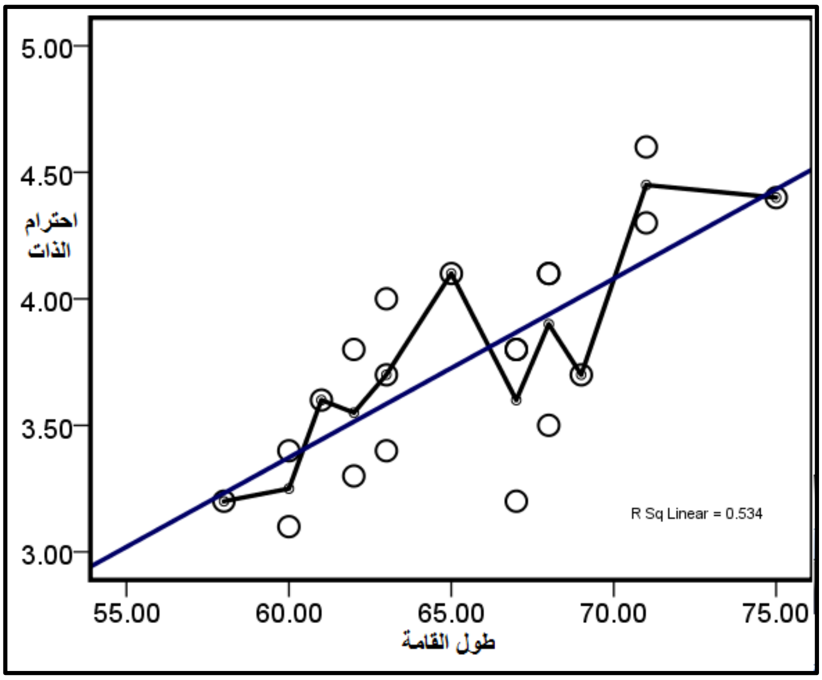

وإذا قمنا باختيار:

**Graphs** ⇨ **Legacy Dialogs** ⇨ **Scatter/Dot**

وباختيار :Simple Scatter




نحصل على الرسم البياني التالي:




**(8–3) الارتباط المتعدد والجزئي "Multi and Partial Correlation"**

**(8–3.1) معامل الارتباط المتعدد:**

إن حساب معامل الارتباط بين أى متغيرين إنما نفترض فيه ثبات باقى العوامل، فالباحث التجريبي يحاول تثبيت أكبر عدد ممكن من العوامل التجريبية، ولكنه لا ينجح إلا نجاحًا محدودًا جزيئًا سواء فى العلوم النفسية والتربوية، فمثلا عند دراسة العلاقة بين الذكاء والذاكرة يجد العديد من العوامل المتداخلة فى هذه العلاقة، الميل والاتجاه والاتزان الانفعالى والقدرات الخاصة والمعلم ونظام المدرسة وطرق التدريس ..الخ، وفى البحوث العلمية كمثال دراسة إنتاجية الفدان لأربع أنواع البذور، نجد العديد من العوامل المتداخلة، نوع التربة ومكوناتها، ونوع الاسمدة وطرق الري وعدد مرات الري، وتنظيف الحقل من الأعشاب العشوائية ..الخ، وفى الأبحاث الطبية لدراسة إحدى مركبات علاج الضغط نجد أيضا العديد من العوامل المتداخلة، عمر المريض وكذلك وظائف الأجهزة الحيوية والغدد الصماء والقنوية ونوعية الغذاء ...الخ، وهذا مانراه من انتشار أنواع مختلفة لعلاج ضغط الدم المرتفع مما يدل على صعوبة السيطرة على هذه العوامل المتداخلة.

لذا معامل الارتباط المتعدد يفيدنا فى معرفة درجة توقف أحد المتغيرات على متغيرين آخرين أو أكثر.

فهو يقيس قوة العلاقة بين أكثر من متغيرين.

أي العلاقة بين المتغير التابع Dependent Variable (y)، والمتغيرات المستقلة Independent Variable ($x_1, x_2, \ldots\ldots\ldots, x_n$).

أو بعبارة أخرى يقيس مدى الترابط بين قيم الظاهرة (المتغير) y، والظواهر (المتغيرات) ($x_1, x_2$ , ………., $x_n$.)

### أ) معامل الارتباط المتعدد لثلاث متغيرات:

يقال عن درجة العلاقة بين ثلاثة متغيرات فأكثر بالارتباط المتعدد، والمبادئ الأساسية في حل مسائل الارتباط المتعدد وهي في إيجاد الارتباط البسيط بين هذه المتغيرات ($x_2$، $x_3$، $x_1$)، وتطبيق الصيغة الرياضية التالية لمعامل الارتباط المتعدد:

$$r_{1.23} = \frac{\sqrt{(r^2{}_{12} + r^2{}_{13}) - 2(r_{12}\, r_{13}\, r_{23})}}{1 - r^2{}_{23}}$$

المتغير التابع (v=1)

المتغيرات المستقلة (x=2.3)

حيث يرمز رقم (1) للمتغير التابع، والرقمان (2،3) يرمزان للمتغيرين المستقلين ويمكن تقدير معاملات الارتباط من الصيغ الرياضية التالية:

معامل الارتباط بين ( **$x_1x_2$** ) مع تثبيت **$x_3$** من العلاقة الرياضية:

$$r_{12} = \frac{n\sum x_1x_2 - \sum x_1 \sum x_2}{\sqrt{n\sum x_1{}^2 - (\sum x_1)^2}\sqrt{n\sum x_2{}^2 - (\sum x_2)^2}}$$

معامل الارتباط بين (**$x_1x_3$**) مع تثبيت $x_2$ من العلاقة الرياضية:

$$r_{13} = \frac{n\sum x_1x_3 - \sum x_1 \sum x_3}{\sqrt{n\sum x_1{}^2 - (\sum x_1)^2}\sqrt{n\sum x_3{}^2 - (\sum x_3)^2}}$$

معامل الارتباط بين (**$x_2x_3$**) مع تثبيت **$x_1$** من العلاقة الرياضية:

$$r_{23} = \frac{n\sum x_2x_3 - \sum x_2 \sum x_3}{\sqrt{n \sum {x_2}^2 - (\sum x_2)^2}\sqrt{n \sum {x_3}^2 - (\sum x_3)^2}}$$

**مثال:**

يبين الجدول التالي أوزان ($X_1$) وأطوال ($X_2$) وأعمار ($X_3$)أفراد عينة مكونة من 12 فرداً. احسب معامل الارتباط الخطي المتعدد للأوزان والأطوال والأعمار لأفراد العينة.

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **الوزن $x_1$** | 64 | 71 | 53 | 67 | 55 | 58 | 77 | 57 | 56 | 51 | 76 | 68 |
| **الطول $x_2$** | 57 | 59 | 49 | 62 | 51 | 50 | 55 | 48 | 52 | 42 | 61 | 57 |
| **العمر $x_3$** | 8 | 10 | 6 | 11 | 8 | 7 | 10 | 9 | 10 | 6 | 12 | 9 |

**الحل:**

| ن | الوزن $x_1$ | الطول $x_2$ | العمر $x_3$ | $x_1x_2$ | $x_1x_3$ | $x_2x_3$ | ${x_1}^2$ | ${x_2}^2$ | ${x_3}^2$ |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 64 | 57 | 8 | 3648 | 512 | 456 | 4096 | 3249 | 64 |
| 2 | 71 | 59 | 10 | 4189 | 710 | 590 | 5041 | 3481 | 100 |
| 3 | 53 | 49 | 6 | 2597 | 318 | 294 | 2809 | 2401 | 36 |
| 4 | 67 | 62 | 11 | 4154 | 704 | 682 | 4489 | 3844 | 121 |
| 5 | 55 | 51 | 8 | 2805 | 357 | 408 | 3025 | 2601 | 64 |
| 6 | 58 | 50 | 7 | 2900 | 406 | 350 | 3364 | 2500 | 49 |
| 7 | 77 | 55 | 10 | 4235 | 770 | 550 | 5929 | 3025 | 100 |
| 8 | 57 | 48 | 9 | 2736 | 513 | 432 | 3249 | 2304 | 81 |
| 9 | 56 | 52 | 10 | 2912 | 560 | 520 | 3136 | 2704 | 100 |
| 10 | 51 | 42 | 6 | 2142 | 306 | 252 | 2601 | 1764 | 36 |
| 11 | 76 | 61 | 12 | 4636 | 912 | 732 | 5776 | 3721 | 144 |
| 12 | 68 | 57 | 9 | 3876 | 612 | 413 | 4624 | 3249 | 81 |
| $\sum$ | 753 | 643 | 106 | 40830 | 6680 | 5679 | 48139 | 34843 | 976 |

**الوزن والطول:**

$$r_{12} = \frac{n\sum x_1x_2 - \sum x_1 \sum x_2}{\sqrt{n \sum {x_1}^2 - (\sum x_1)^2}\sqrt{n \sum {x_2}^2 - (\sum x_2)^2}}$$

$$r_{12} = \frac{(12*40830)-(754*643)}{\sqrt{(12*48139)-(753*753)}\sqrt{(12*34843)-(643*643)}}$$

$$r_{12} = \frac{5781}{\sqrt{10659*4667}} = \frac{5781}{7053} = 0.819 \cong 0.82$$

**الطول والعمر:**

$$r_{23} = \frac{n\sum x_2x_3 - \sum x_2 \sum x_3}{\sqrt{n\sum x_2{}^2 - (\sum x_2)^2}\sqrt{n\sum x_3{}^2 - (\sum x_3)^2}}$$

$$r_{23} = \frac{(12*5779)-(643*106)}{\sqrt{(12*34843)-(643*643)}\sqrt{(12*976)-(106*106)}}$$

$$r_{23} = \frac{1190}{\sqrt{4667*467}} = \frac{1190}{1490} = 0.798$$

**الوزن والعمر:**

$$r_{13} = \frac{n\sum x_1x_3 - \sum x_1 \sum x_3}{\sqrt{n\sum x_1{}^2 - (\sum x_1)^2}\sqrt{n\sum x_3{}^2 - (\sum x_3)^2}}$$

$$r_{13} = \frac{(12*6796)-(753*106)}{\sqrt{(12*48139)-(753*753)}\sqrt{(12*976)-(106*106)}}$$

$$r_{13} = \frac{1734}{\sqrt{10659*467}} = \frac{1734}{2252} = 0.7699 \cong 0.77$$

لحساب معامل الارتباط المتعدد عندما يكون المتغير الطول والعمر مستقلان والتابع الوزن:

| الوزن (1)؛ الطول (2)؛ العمر (3) |
|---|

$$r_{1.23} = \frac{\sqrt{(r^2{}_{12} + r^2{}_{13}) - (2r_{12}\, r_{13}\, r_{23})}}{1 - r^2{}_{23}}$$

$$r_{1.23} = \frac{\sqrt{[(0.82)^2 + (0.77)^2] - [0.8 * 0.77 * 0.8]}}{1 - (0.8^2)}$$

$$r_{1.23} = \frac{\sqrt{0.25506}}{0.36} = 0.84$$

حيث أن:

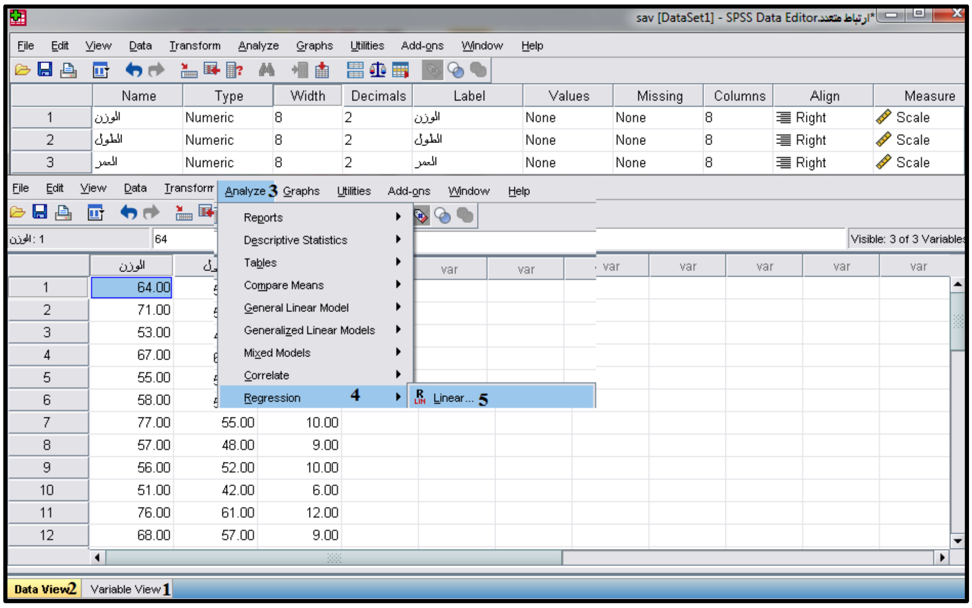
$r_{1.23}$ = معامل الارتباط المتعدد بين $x_1,\ x_2, x_3$ معاً.

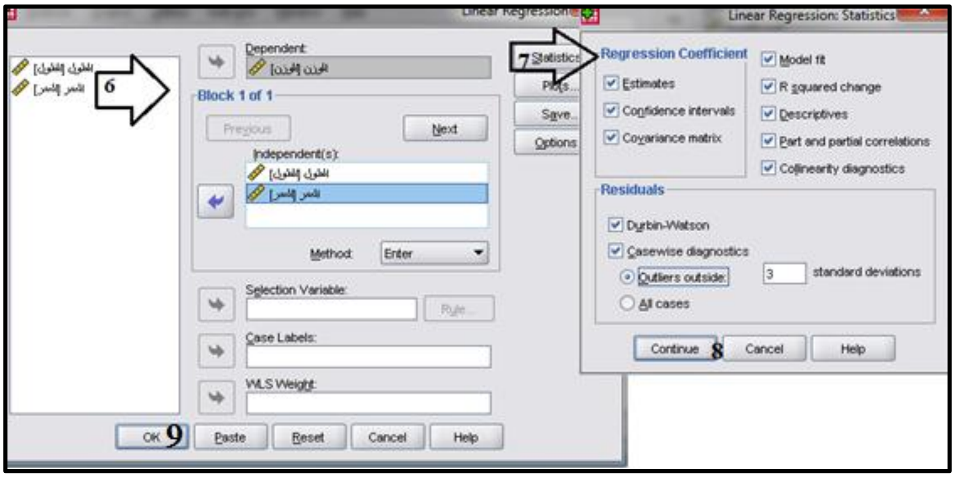
$r_{12}$ = معامل الارتباط البسيط بين $x_1, x_2$.

$r_{13}$ = معامل الارتباط البسيط بين $x_1,\ x_3$.

$r_{23}$ = معامل الارتباط البسيط بين $x_2, x_3$ .

## حساب الارتباط المتعدد باستخدام Spss:






| Descriptive Statistics الإحصاء الوصفي | | | |
|---|---|---|---|
| | Mean | Std. Deviation | N |
| الوزن | 62.7500 | 8.98610 | 12 |
| الطول | 53.5833 | 5.94610 | 12 |
| العمر | 8.8333 | 1.89896 | 12 |

**معامل ارتباط بيرسون بين كل متغيرين:**

| Correlations | | | | |
|---|---|---|---|---|
| | | **الوزن** | **الطول** | **العمر** |
| Pearson Correlation | الوزن | 1.000 | .820 | .770 |
| | الطول | .820 | 1.000 | .798 |
| | العمر | .770 | .798 | 1.000 |

معامل الارتباط المتعدد على أساس الطول والعمر متغيران ثابتان والزمن متغير تابع.

| Model Summary | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Model | معامل الارتباط المتعدد R | معامل التحديد R Square | معامل التحديد المعدل Adjusted Square | Std. Error of Estimate | Change Statistics | | | | | Durbin Watson |
| | | | | | R Square Change | F Change | df1 | Df2 | Sig. F Change | |
| 1 | .842[a] | .709 | .644 | 5.36321 | .709 | 10.940 | 2 | 9 | .004 | 1.771 |

a. Predictors: (Constant), الطول، العمر (ثابتان)

b. Dependent Variable: الوزن (متغير ثابت)

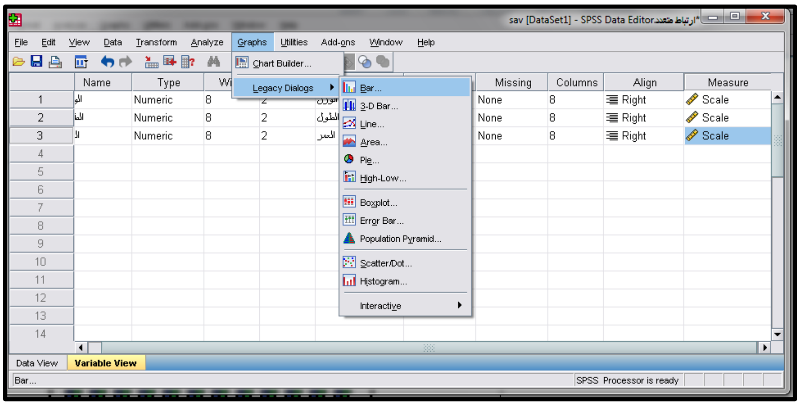
**مما سبق يتضح أن:**

**1.** معامل الارتباط الخطي المتعدد $r_{1.23} = 0.84$ بقيم معاملات الارتباط الخطية البسيطة التي دخلت في تكوينه وهي $r_{12} = 0.82$ ، $r_{13} = 0.77$، $r_{23} = .7980$، فنجد قيمة ذلك العامل أكبر من أي من قيم المعاملات البسيطة الثلاثة، وهذه القيمة صحيحة ومتوقعة دائماً إذ أن تقديرنا للعلاقة الكلية بين المتغيرات الثلاث يكون أفضل دائماً باستخدام معلومات إضافية عن العلاقات الثنائية بين أزواج المتغيرات.

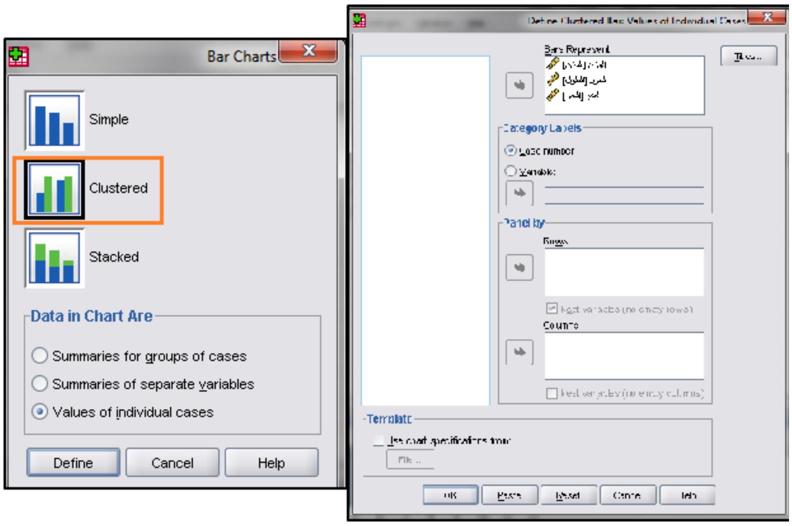
**2.** هذا المعامل موجب القيمة دائماً.

**ويمكن تمثيل البيانات بيانياً.**

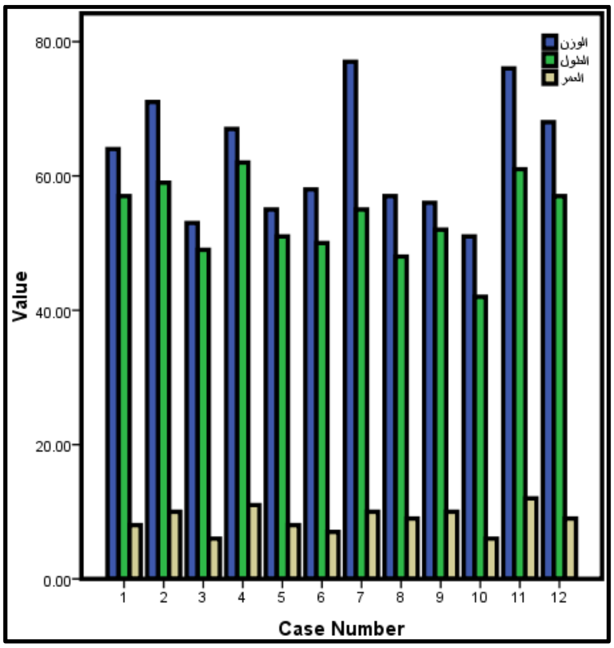
**Graphy ⇨ Lejacy ⇨ Dialogs ⇨ Bar**



يظهر مستطيل الحوار التالي ونؤشر على Clustered:




وبالضغط على OK يظهر الرسم البيانى التالى:



معامل الارتباط المتعدد على أساس الوزن والعمر متغيران ثابتان والطول متغير تابع:

| Variables Entered/Removed[b] | | | |
|---|---|---|---|
| Model | Variables Entered | Variables Removed | Method |
| 1 | الوزن، العمر[a] | | Enter |
| a. All requested Variables Entered | | | |
| b. Dependent Variable: الطول | | | |

| Model Summary | | | | |
|---|---|---|---|---|
| Mode | معامل الارتباط المتعدد R | معامل التحديد R Square | معامل التحديد المعدل Adjusted Square | Std. Error of Estimate |
| 1 | .861[a] | .741 | .683 | 3.34785 |
| a. Predictors: (Constant), الوزن، العمر (ثابتان) | | | | |

معامل الارتباط المتعدد على أساس الوزن والطول متغيران ثابتان والعمر متغير تابع:

| Variables Entered/Removed[b] | | | |
|---|---|---|---|
| Model | Variables Entered | Variables Removed | Method |
| 1 | الوزن، العمر[a] | | Enter |
| a. All requested Variables Entered | | | |
| b. Dependent Variable: العمر | | | |

| Model Summary | | | | |
|---|---|---|---|---|
| Mode | R معامل الارتباط المتعدد | معامل التحديد R Square | معامل التحديد المعدل Adjusted Square | Std. Error of Estimate |
| 1 | .823[a] | .678 | .606 | 1.19123 |
| a. Predictors: (Constant), الطول، الوزن | | | | |

### (8–3 ب) الارتباط الجزئي: Partial Correlation

في بعض الظواهر والدراسات يوجد هناك عدد من المتغيرات (ثلاثة فأكثر) مرتبطة بعلاقة رياضية فيما بينها مثل إنفاق أسرة يكون مرتبط بدخلها الشهري وعدد أفرادها وكذلك حجم مبيعات سلعة معينة يرتبط بسعرها وحجم الدعاية لها وكذلك الفترة الزمنية للبيع ففي هذه الحالة ولغرض حساب معامل الارتباط بين متغيرين اثنين في دراسة معينة مع وجود متغيرات أخرى نلجأ إلى حساب ما يسمى بالارتباط الجزئي.

فالارتباط الجزئي هو العلاقة الرياضية الصافية بين متغيرين اثنين فقط مع وجود متغيرات أخرى قيد الدراسة ويمكن حساب هذه العلاقة الرياضية من خلال معامل الارتباط الجزئي.

فالفرق بينه وبين معامل الارتباط البسيط هو أن معامل بيرسون يستخرج العلاقة بين متغيرين اثنين لأي ظاهرة بدون أن يأخذ بعين الاعتبار وجود متغيرات أخرى تؤثر في الظاهرة أو لا، بينما معامل الارتباط الجزئي لقياس الارتباط بين متغيرين بمعزل عن تأثير المتغيرات الأخرى لكي يستخرج الارتباط الصافي بين أي متغيرين.

خصائص معامل الارتباط الجزئي تفسر قيمته كما تفسر قيمة معامل الارتباط البسيط.

إن قيمة معامل الارتباط الجزئي تتراوح بين (1−،1)

إن معامل الارتباط الجزئي لأي متغيرين تكون إشارته مماثلة لإشارة معامل الارتباط البسيط بينهما.

**معادلة الارتباط الجزئي:**

يقيس هذا العامل العلاقة بين متغيرين اثنين بعد عزل (تثبيت) أثر المتغير الثالث:

فعلى فرض أن معاملات الارتباط الخطية البسيطة بين ($x_1$،$x_2$)، ($x_1$،$x_3$)، ($x_2$،$x_3$) ف تكوينه وهي $r_{12} = 0.82$ ، $r_{13} = 0.77$، .7980 = $r_{23}$، فيمكن حساب معامل الارتباط الجزئي كالتالى:

معامل الارتباط الجزئي بين ($x_1$،$x_2$) مع تثبيت $x_3$ من العلاقة الرياضية التالية:

$$r_{12.3} = \frac{r_{12} - [r_{13} * r_{23}]}{\sqrt{[1-r^2\ _{13}][1-r^2\ _{23}]}}$$

معامل الارتباط الجزئي ($x_1$،$x_3$) مع تثبيت $x_2$ من العلاقة الرياضية:

$$r_{13.2} = \frac{r_{13} - [r_{12} * r_{32}]}{\sqrt{[1-r^2\ _{12}][1-r^2\ _{32}]}}$$

معامل الارتباط الجزئي ($x_2$،$x_3$) مع تثبيت $x_1$ من العلاقة الرياضية:

$$r_{23.1} = \frac{r_{23} - [r_{12} * r_{13}]}{\sqrt{[1-r^2\ _{12}][1-r^2\ _{13}]}}$$

**تقدير معامل الارتباط الجزئي:**

باستخدام قيم $r_{12}$، $r_{13}$، $r_{32}$ فى المثال السابق فإن:

**معامل الارتباط الجزئى بين ($x_1$،$x_2$) مع تثبيت $x_3$**

أي $r_{12}$: أي معامل الارتباط بين المتغيرين 1، 2 مع حذف أثر المتغير 3 من العلاقة الرياضية التالية:

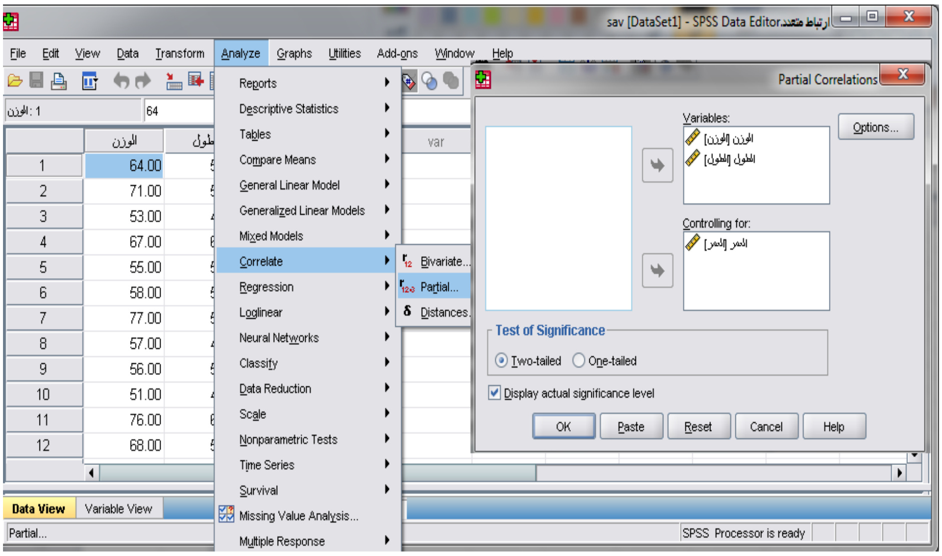
$$r_{12.3} = \frac{r_{12} - [r_{13} * r_{23}]}{\sqrt{[1-r^2\ _{13}][1-r^2\ _{23}]}}$$

$$r_{12.3} = \frac{0.82 - [0.77 * 0.798]}{\sqrt{[1-(0.77)^2][1-(0.798)^2]}}$$

$$r_{12.3} = \frac{0.82 - 0.61446}{\sqrt{0.4071 \times 0.363196}}$$

$$r_{12.3} = \frac{0.20554}{\sqrt{0.147857091}} = \frac{0.20554}{0.38452189} = 0.53$$

وباستخدام Spss ندخل البيانات كما في الشكل التالي:



فنحصل على النتائج التالية:

| Correlations | | | | |
|---|---|---|---|---|
| Control Variables | | | الوزن | الطول |
| العمر | الوزن | Correlation | 1.000 | .533 |
| | | Significance (2- tailed) | . | .091 |
| | | df | 0 | 9 |
| | الطول | Correlation | .533 | 1.000 |
| | | Significance (2- tailed) | .091 | . |
| | | df | 9 | 0 |

**معامل الارتباط الجزئى بين ($x_1$،$x_3$) مع تثبيت** $x_2$

أي $r_{13.2}$ معامل الارتباط بين المتغيرين 1، 3 مع حذف أثر المتغير 2 من العلاقة الرياضية التالية:

$$r_{13.2} = \frac{r_{13} - [r_{12} * r_{32}]}{\sqrt{[1-r^2{}_{12}][1-r^2{}_{32}]}} = \frac{0.77-[0.798*0.82]}{\sqrt{[1-(0.798)^2][1-(0.82)^2]}}$$

$$r_{13.2} = \frac{0.77-0.65436}{\sqrt{0.363196 \times 0.3276}} = \frac{0.11564}{\sqrt{0.118983}} = \frac{0.11564}{0.344939} = 0.3352476$$

وباستخدام Spss نحصل على النتائج التالية:

| Correlations | | | | |
|---|---|---|---|---|
| Control Variables | | | الوزن | العمر |
| الطول | الوزن | Correlation | 1.000 | .533 |
| | | Significance (2- tailed) | . | .315 |
| | | df | 0 | 9 |
| | العمر | Correlation | .533 | 1.000 |
| | | Significance (2- tailed) | .315 | . |
| | | df | 9 | 0 |

معامل الارتباط الجزئى بين ($x_2$،$x_3$) مع تثبيت $x_1$.

$r_{13.2}$ أي معامل الارتباط بين المتغيرين 2، 3 مع حذف أثر المتغير1 من العلاقة الرياضية التالية:

$$r_{23.1} = \frac{r_{23} - [r_{12} * r_{13}]}{\sqrt{[1-r^2\ _{12}][1-r^2\ _{13}]}} = \frac{0.798 - [0.77 * 0.82]}{\sqrt{[1-(0.77)^2][1-(0.82)^2]}}$$

$$r_{23.1} = \frac{r_{23} - [r_{12} * r_{13}]}{\sqrt{[1-r^2\ _{12}][1-r^2\ _{13}]}} = \frac{0.798 - [0.77 * 0.82]}{\sqrt{[1-(0.77)^2][1-(0.82)^2]}}$$

$$r_{23.1} = \frac{0.798 - 0.6314}{\sqrt{0.4071 \times 0.3276}} = \frac{0.1666}{\sqrt{0.13336596}} = \frac{0.1666}{0.36519} = 0.4562$$

وباستخدام Spss نحصل على النتائج التالية:

| Correlations | | | | |
|---|---|---|---|---|
| Control Variables | | | الطول | العمر |
| الوزن | الطول | Correlation | 1.000 | .458 |
| | | Significance (2- tailed) | . | .157 |
| | | df | 0 | 9 |
| | العمر | Correlation | .458 | 1.000 |
| | | Significance (2- tailed) | .157 | . |
| | | df | 9 | 0 |

**(4–8) معامل ارتباط الرتب لسبيرمان:**

**Spearman's rank correlation coefficient**

إذا كانت الظاهرة محل الدراسة تحتوي على متغيرين وصفيين ترتيبيين، ومثال على ذلك: قياس العلاقة بين تقديرات الطلبة في مادتين، أو العلاقة بين درجة تفضيل المستهلك لسلعة معينة، ومستوى الدخل، فإنه يمكن استخدام طريقة "بيرسون" السابقة في حساب معامل ارتباط يعتمد على رتب مستويات المتغيرين كبديل للقيم الأصلية، ويطلق على هذا المعامل "معامل ارتباط اسبيرمان" Spearman ويعبر عنه بالمعادلة التالية:

$$r_s = 1 - \frac{6\sum d^2}{n(n^2-1)}$$

**حيث أن $d$ هي الفرق بين رتب مستويات المتغير الأول $x$، ورتب مستويات المتغير الثاني $y$.**

**أي أن**: $d = R_x - R_y$.

**معامل ارتباط الرتب لسبيرمان يستخدم في حالة:**

**1**. البيانات الوصفية.

**2**. يفضل استخدامه عندما تكون عدد قيم المشاهدات للمتغيرين (X، Y) أقل من ثلاثين.

**مميزاته:**

**أ**. يمتاز هذا المعامل بالسهولة والسرعة.

**ب**. يصلح للبيانات الكمية والوصفية.

**عيوبه:**

أننا لا نجد معنى طبيعياً للفرق بين رتبتين وكذلك لتربيع ذلك الفرق.

**حساب معامل ارتباط لسبيرمان من البيانات غير المبوبة:**

**1**. نرتب المتغير الأول، وعادة يرمز للمتغير الأول بـ (x)، ويكون هذا الترتيب تنازلياً بإعطاء الرتبة الأولى لأكبر درجة والرتبة الثانية للدرجة التي تليها... وهكذا ترتيب تنازلي أو العكس، ويكون ذلك في العمود المسمى (رتبة x).

**2**. نرتب المتغير الثاني، وعادة يرمز للمتغير الثاني بـ (y)، ويكون هذا الترتيب بنفس الأسلوب المتبع في ترتيب المتغير (x)، ويكون ذلك في العمود المسمى (رتبة y).

**3**. حساب الفرق بين رتبة (x) ورتبة (y)، بطرح رتبة y من رتبة x، ويوضع ذلك في العمود المسمى (d) أي الفرق.

**4**. تربيع الفرق ويضع الناتج في العمود المسمى ($d^2$) أي مربع الفرق.

**5**. جمع العمود الأخير ليحصل على ($\Sigma d^2$).

**6**. نطبق المعادلة التي توصل إليها سبيرمان لحساب معامل الارتباط.

$$r_s = 1 - \frac{6\sum d^2}{n(n^2-1)} \quad \textbf{or} \quad r_s = 1 - \frac{6\sum d^2}{n^3-n}$$

حيث n: عدد أزواج المشاهدات للمتغيرين y، x.

d: الفرق بين رتب (x,y).

**ملاحظة:**

عندما تتساوى قيمتان أو أكثر فإن هذه القيم تشترك في ترتيب واحد، ثم نعطي رتبة متساوية هي الوسط الحسابي للرتب التي تأخذها هذه القيم لو أنها كانت مختلفة.

مثال: احسب معامل ارتباط الرتب لسبيرمان بين المتغيرين (x ، y).

| | | | | | | |
|---|---|---|---|---|---|---|
| x | 5 | 9 | 10 | 7 | 6 | 8 |
| y | 5 | 10 | 9 | 13 | 7 | 8 |

**الحل**: نرتب البيانات تصاعدياً ثم نوجد الفرق:

| n | x | y | Rank x | Rank y | d = x − y | $d^2$ |
|---|---|---|---|---|---|---|
| 1 | 8 | 8 | 4 | 3 | $4-3=1$ | 1 |
| 2 | 6 | 7 | 2 | 2 | $2-2=0$ | 0 |
| 3 | 7 | 13 | 3 | 6 | $3-6=-3$ | 9 |
| 4 | 10 | 9 | 6 | 4 | $6-4=2$ | 4 |
| 5 | 9 | 10 | 5 | 5 | $5-5=0$ | 0 |
| 6 | 5 | 5 | 1 | 1 | $1-1=0$ | 0 |
| $\sum$ | | | − | − | 0 | 14 |

**أو** نرتب البيانات تنازلياً ثم نوجد الفرق:

| n | x | y | Rank x | Rank y | d = rank x − rank y | $d^2$ |
|---|---|---|---|---|---|---|
| 1 | 8 | 8 | 3 | 4 | $3-4=1$ | 1 |
| 2 | 6 | 7 | 5 | 5 | $0-0=0$ | 0 |
| 3 | 7 | 13 | 4 | 1 | $4-1=-3$ | 9 |
| 4 | 10 | 9 | 1 | 3 | $1-3=2$ | 4 |
| 5 | 9 | 10 | 2 | 2 | $0-0=0$ | 0 |
| 6 | 5 | 5 | 6 | 6 | $6-6=0$ | 0 |
| $\sum$ | | | – | – | 0 | 14 |

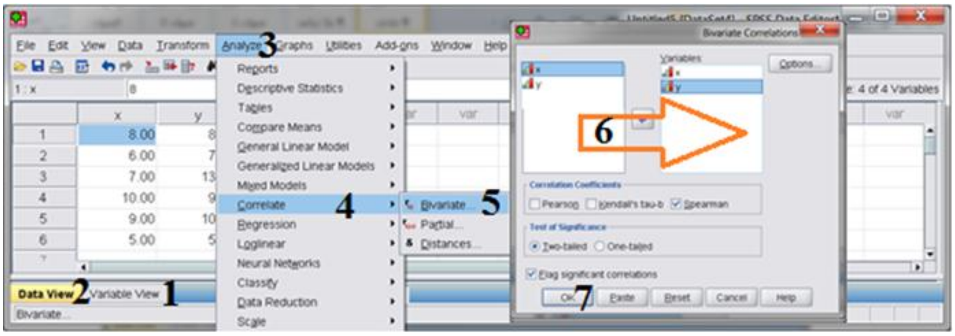
$$r_s = 1 - \frac{6\sum d^2}{n(n^2-1)} = 1 - \frac{6\times 14}{6(36-1)} = 1 - \frac{84}{6\times 35} = 1 - 0.4 = 0.6$$

**الحل : باستخدام الحزمة الإحصائية spss باتباع المسار:**

**1**. ندخل رتب كل متغير ثم نتبع المسار التالي:

أو 3. ندخل البيانات كما هي ونتبع المسار التالي:

**Analyze** → **Correlate** → **Bivariate**




والجدول التالي يلخص النتائج:

| Correlations | | | | |
|---|---|---|---|---|
| | | | x | y |
| Spearman's rho | x | Correlation Corefficient | 1.000 | .600 |
| | | Sig. (2- tailed) | . | .208 |
| | | N | 6 | 6 |
| | y | Correlation Corefficient | .600 | 1.000 |
| | | Sig. (2- tailed) | .208 | . |
| | | N | 6 | 6 |

**(8–5) ارتباط الصفات غير القابلة للترتيب:**

نجد بعض الصفات غير القابلة للترتيب مثل الجنس، نوع التعليم، لون الزهرة ... الخ

في مثل هذه الصفات يمكننا استخدام مقاييس الاقتران ومقياس معامل التوافق.

**(8–1.5) مقياس معامل الاقتران** Coefficient of Association

معامل الارتباط للرتب (سبيرمان) هو للبيانات الوصفية التي يمكن ترتيبها وفي الحالة التي لا يمكن ترتيب البيانات أو تلك البيانات وإن رتبت فلا أهمية لترتيبها نستخدم ما يعرف بمعامل الاقتران Coefficient of Association لمتغيرين كل منهم لحالتين فقط، ويستخدم لقياس العلاقة بين ظاهرتين تنقسم كل منهما إلى قسمين (أو صفتين) فقط، وتكون البيانات موضوعة في جدول مزدوج يتكون من أربع خلايا 2X2 فقط، ويسمى الجدول في هذه الحالة

"جدول الاقتران" (2X2) حيث توضع إحدى الظاهرتين أفقياً والأخرى رأسياً ويكون الشكل العام لجدول الاقتران 2X2.

ولتوضيح ذلك نقول لدينا المتغيرين y،x وهناك صفتان للمتغير x هما a، b وصفتان للمتغير y هما d، c فيكون لدينا الجدول التالي:

| المتغير | $y_1$ | $y_2$ |
|---|---|---|
| $x_1$ | a | b |
| $x_2$ | c | d |

حيث:

**عناصر الصف الأول:**

(**a**) عبارة عن التكرارات المشتركة في الصفة الأولى $x_1$ للمتغيرx، والصفة الأولى $y_1$ للمتغيرy.

(**b**) عبارة عن التكرارات المشتركة في الصفة الأولى $x_1$ للمتغيرx، والصفة الثانية $y_2$ للمتغيرy.

**عناصر الصف الثاني:**

(**c**) عبارة عن التكرارات المشتركة في الصفة الثانية لمتغير x، والصفة الأولى $y_1$ للمتغير y.

(**d**) عبارة عن التكرارات المشتركة في الصفة الثانية $x_2$ للمتغير x، والصفة الثانية $y_2$ للمتغيرy.

**قياس معامل الاقتران:**

قام ييل (yule) بوضع تعريف معامل الاقتران حسب العلاقة الرياضية التالية:

$$R_A = (ad - bc)/\,(ad + bc)$$

إن قيمة معامل الاقتران تقع بين الصفر والواحد الصحيح أي أن: 1 ≥ rA ≥ صفر.

**مثال:**

من الجدول التالي والذي يبين بيانات الاقتران بين العمل والتعلم احسب معامل الاقتران.

| | متعلم | غير متعلم |
|---|---|---|
| يعمل | 8 | 6 |
| لا يعمل | 3 | 7 |

**الحل**:

بتطبيق الصيغة السابقة نجد أن:

$$R_A = (ad - bc) / (ad + bc)$$
$$R_A = [(8 * 7) - (6 * 3)] / [(8 * 7) + (6 * 3)]$$
$$R_A = 38 / 72 = 0.514$$

**الحل**: باستخدام الحزمة الإحصائية spss.

نقوم بإدخال البيانات كالتالي:

أسفل العمود الأول (X):

ندخل بيانات الصف الأول على التوالي ثم يليهم إدخال بيانات الصف الثاني على التوالي.

أسفل العمود الثاني (Y):

نرمز بيانات الصف الأول بالرقم 1 ثم نرمز بيانات الصف الثاني بالرقم 2.

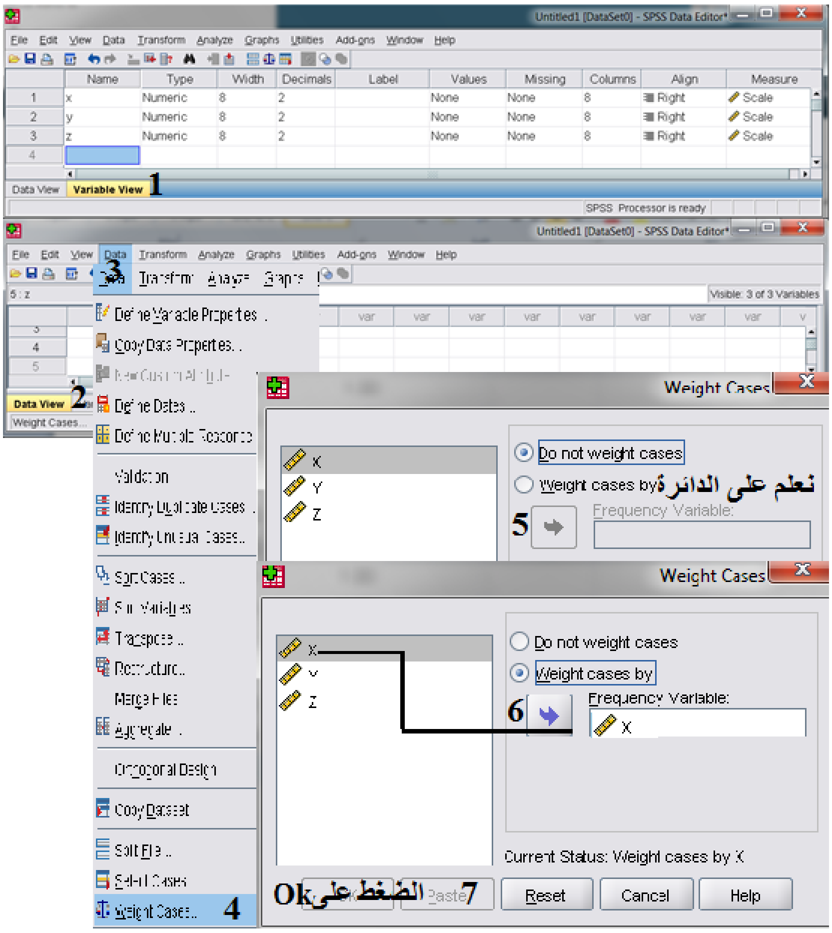
أسفل العمود الثالث (Z):

نرمز بيانات العمود الأول بالرقم 1 ثم نرمز بيانات العمود الثاني بالرقم 2، ثم نرمز بيانات العمود الثالث بالرقم 3.

ثم اتباع المسار:

**Data ⇨ Weight Cases**

بعد اختيار Weight Cases نحصل على الصندوق المبين بالشكل أدناه، ونعلم على الدائرة الواقعة على يسار Weight Cases by وننقل المتغير X للخانة Frequency Variable ثم OK.




ثم نتبع المسار التالي:

**Analyze** → **Descriptive Statistics** → **Crosstabs**

ثم ننقل المتغير Y لخانة Row والمتغير Z لخانة Column.

وباختيار Statistics نختار معامل فاي وكآي تربيع ثم نضغط OK فنحصل على النتائج كما موضح بالشكل التالي:

فنحصل على النتائج التالية:

| Case Processing Summary | | | | | | |
|---|---|---|---|---|---|---|
| | Cases | | | | | |
| | Valid | | Missing | | Total | |
| | N | Percent | N | Percent | N | Percent |
| y * z | 24 | 100.0% | 0 | .0% | 24 | 100.0% |

| y*z Cross tabulation | | | | |
|---|---|---|---|---|
| Count | | | | |
| | | Z التعلم | | Total |
| | | متعلم 1 | غير متعلم 2 | |
| y العمل | يعمل 1 | 8 | 6 | 14 |
| | لا يعمل 2 | 3 | 7 | 10 |
| Total | | 11 | 13 | 24 |

| Chi-Square Tests | | | | | |
|---|---|---|---|---|---|
| | Value | df | Asymp. Sig. (2-sided) | Exact Sig. (2-sided) | Exact Sig. (1-sided) |
| Pearson Chi-Square | 1.731a | 1 | .188 | | |
| Continuity Correctionb | .810 | 1 | .368 | | |
| Likelihood Ratio | 1.765 | 1 | .184 | | |
| Fisher's Exact Test | | | | .240 | .185 |
| Linear-by-Linear Association | .6591 | 1 | .198 | | |
| N of Valid Casesb | 24 | | | | |
| a. 1 cells (25.0%) have expected count less than 5. The minimum expected count is 4.58. | | | | | |
| b. Computer only for a 2x2 table | | | | | |

| Symmetric Measures | | | |
|---|---|---|---|
| | | Value | Approx. Sig |
| Nominal by Nominal | Phi | .269 | .188 |
| | Cramer's V | .269 | .188 |
| N of Valid Cases | | 24 | |

**(2.5–8) معامل التوافق Coefficient of contingency**

يستخدم هذا المقياس لإيجاد معامل الارتباط لبيانات وصفية من جداول التوافق كما أنه يستخدم أيضاً عندما تكون إحدى الظاهرتين وصفية والأخرى كمية.

هو مقياس لا معلمي Non parametric فهو لا يضع أي قيد على المتغيرين ولا على شكل توزيع المتغيرين، وهو يستخدم لقياس العلاقة بين مجموعتين من الصفات.

فإذا كان للمتغيرين (أحدهم على الأقل) أكثر من صفتين كلون العيون (أسود – أزرق – عسلي – ...) فيعرف معامل الاقتران في هذه الحالة بمعامل التوافق ويرمز له

بالرمز rc ويقاس الارتباط من الصيغة الآتية والتي تعتمد على حساب معامل ($x2$)، فنكون جدول البيانات ونعوض في الصيغ الرياضية والتي نبينها هنا بين المتغيرين x , y.

| $y_i \backslash x_i$ | $x_1$ | $x_2$ | ……….. | $x_k$ | total |
|---|---|---|---|---|---|
| $y_1$ | $n_{y_1}x_1$ | $n_{y_1}x_2$ | ……….. | $n_{y_{1xk}}$ | $n_{y_1}$ |
| $y_2$ | $n_{y_1}x_1$ | $n_{y_1}x_2$ | ……….. | $n_{y_{2xk}}$ | $n_{y_2}$ |
| : | : | : | ……….. | : | : |
| $y_r$ | $n_{yr}x_1$ | $n_{yr}x_2$ | ……….. | $n_{yrxk}$ | $n_{yr}$ |
| total | $n_{x_1}$ | $n_{x_2}$ | ……….. | $n_{xk}$ | $n_n$ |

وحين تكون قياسات أحد المتغيرين أو كلاهما من النوع الوصفي أو النوعي أ، التي يفرق بين مفردة وأخرى على أساس وصفي محض فيستخدم معامل خاص هو معامل التوافق لقياس العلاقة بين المتغيرين وله عدة تعاريف منها:

وهذه الصيغة الرياضية لكل من: $r_c$، $x^2$

$$r_c = \sqrt{\frac{x^2}{x^2+n}}$$

$$x^2 = n\left[\frac{n_{y_1x_1}^2}{n_{y_1}n_{x_1}} + \frac{n_{y_1x_2}^2}{n_{y_1}n_{x_2}} + \cdots + \frac{n_{y_1x_c}^2}{n_{y_1}n_{x_c}} + \frac{n_{y_2x_1}^2}{n_{y_2}n_{x_1}} + \cdots + \frac{n_{y_xx_1}^2}{n_{y_x}n_{x_c}}\right] - n$$

**مثال:**

الجدول الآتي يبين بيانات متغيري المهنة والتدخين والمطلوب حساب معامل الارتباط التوافقي.

| Work / Smoke | $x_1$ | $x_2$ | $x_3$ | Total |
|---|---|---|---|---|
| Smoke | 32 | 75 | 25 | 132 |
| No Smoke | 28 | 25 | 15 | 68 |
| Total | 60 | 100 | 40 | 200 |

**الحل**: نطبق القانون الخاص بحساب $x_2$ السابق ونحسب معامل التوافق باستخدام صيغته السابق ذكرها أعلاه.

$$x^2 = n\left[\frac{n_{y_1x_1}^2}{n_{y_1}n_{x_1}} + \frac{n_{y_1x_2}^2}{n_{y_1}n_{x_2}} + \cdots + \frac{n_{y_1x_c}^2}{n_{y_1}n_{x_c}} + \frac{n_{y_2x_1}^2}{n_{y_2}n_{x_1}} + \cdots + \frac{n_{y_rx_1}^2}{n_{y_r}n_{x_c}}\right] - n$$

$$x^2 = 200\left[\frac{(32)^2}{60\times132} + \frac{(75)^2}{100\times132} + \frac{(25)^2}{40\times132} + \frac{(28)^2}{60\times68} + \frac{(25)^2}{100\times68} + \frac{(15)^2}{40\times68}\right] - 200$$

$$x^2 = 200\left[\frac{1024}{7920} + \frac{5625}{13200} + \frac{625}{5280} + \frac{784}{4080} + \frac{625}{6800} + \frac{225}{2720}\right] - 200$$

$$x^2 = 200 \times 1.04059 - 200 = 208.118 - 200 = 8.118$$

$$r_c = \sqrt{\frac{8.118}{208.118}}$$

وهذا يشير لضعف القوة بين التدخين والمهنة مع التنبيه على أن زيادة الأعمدة والصفوف يزيد من ارتفاع معنوية معامل التوافق ولكن لن تتجاوز الواحد الصحيح.

**الحل**: باستخدام SPSS:

نقوم إدخال البيانات كالتالي:

أسفل العمود الأول (X):

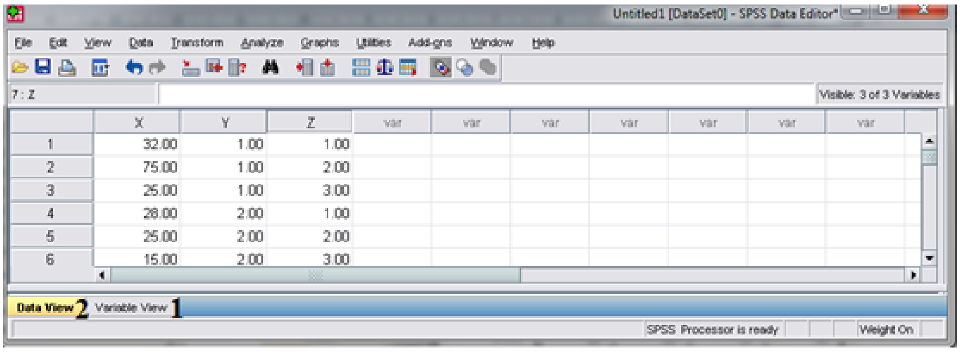
ندخل بيانات الصف الأول على التوالي ثم يليهم إدخال بيانات الصف الثاني على التوالي.

أسفل العمود الثاني (Y):

نرمز بيانات الصف الأول بالرقم 1 ثم نرمز بيانات الصف الثاني بالرقم 2.

أسفل العمود الثالث (Z):

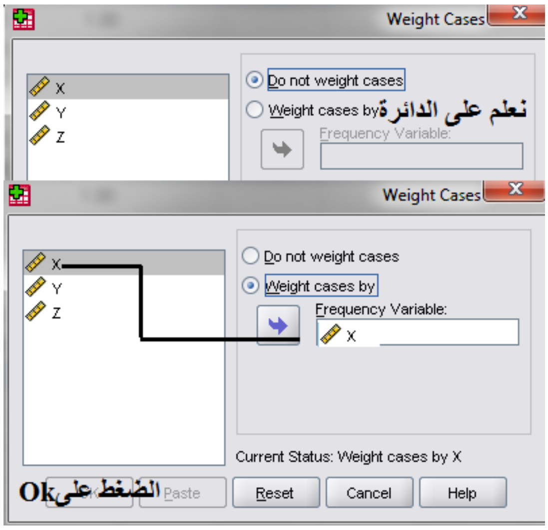
نرمز بيانات العمود الأول بالرقم 1 ثم نرمز بيانات العمود الثاني بالرقم 2، ثم نرمز بيانات العمود الثالث بالرقم 3.




ثم اتباع المسار:

**Data ⟹ Weight Cases**

بعد اختيار Weight Cases نحصل على الصندوق المبين بالشكل أدناه، ونعلم على الدائرة الواقعة على يسار Weight Cases by وننقل المتغير X للخانة Frequency Variable ثم OK.

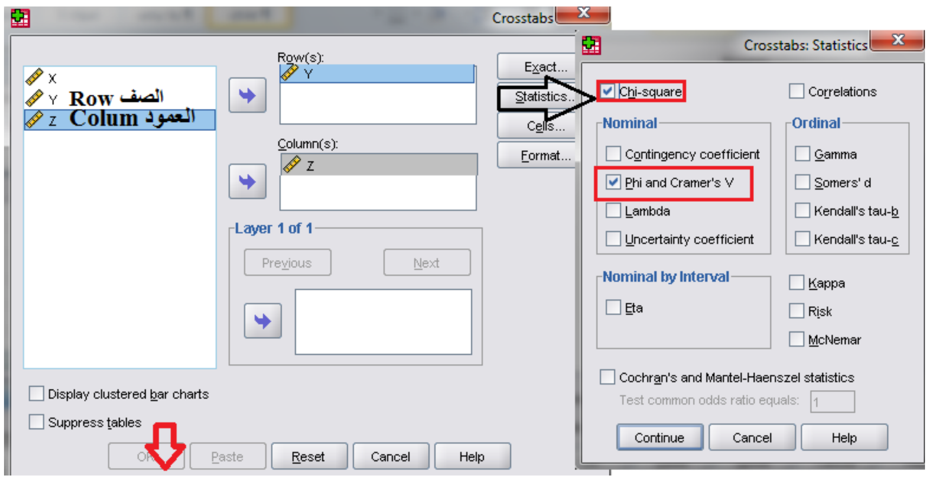



**Analyze** → **Descriptive Statistics** → **Crosstabs**

ثم ننقل المتغير Y لخانة Row والمتغير Z لخانة Column

وباختيار Statistics نختار معامل فاي وكآي تربيع ثم نضغط OK فنحصل على

النتائج كما موضح بالشكل التالي:



| Case Processing Summary | | | | | | |
|---|---|---|---|---|---|---|
| | Cases | | | | | |
| | Valid | | Missing | | Total | |
| | N | Percent | N | Percent | N | Percent |
| Y * Z | 200 | 100.0% | 0 | .0% | 200 | 100.0% |

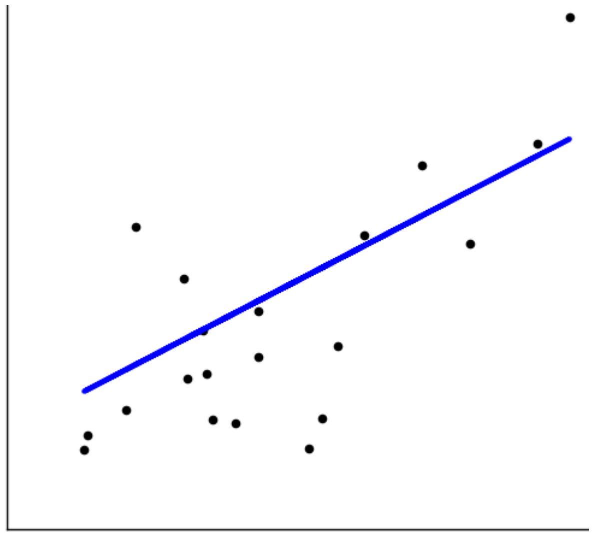
| **Y * Z Cross tabulation** | | | | | |
|---|---|---|---|---|---|
| **Count** | | | | | |
| | | **Z(work)** | | | **Total** |
| | | **1** | **2** | **3** | |
| **Y(smoke)** | **1** | **32** | **75** | **25** | **132** |
| | **2** | **28** | **25** | **15** | **68** |
| **Total** | | **60** | **100** | **40** | **200** |

| **Chi-Square Tests** | | | |
|---|---|---|---|
| | Value | df | Asymp. Sig. (2-sided) |
| Pearson Chi-Square | 8.118a | 2 | .017 |
| Likelihood Ratio | 8.111 | 2 | .017 |
| Linear-by-Linear Association | 1.739 | 1 | .187 |
| N of Valid Cases | 200 | | |
| a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 13.60. | | | |

| **Symmetric Measures** | | | |
|---|---|---|---|
| | | Value | Approx. Sig. |
| Nominal by Nominal | Phi | .201 | .017 |
| | Cramer's V | .201 | .017 |
| N of Valid Cases | | 200 | |

# الانحدار الخطي البسيط

# Linear Regression



– مقدمة.

(9-1) تقدير خط الانحدار للبيانات غير المبوبة.

(9-1.1) خط انحدار y على x.

**مقدمـة:**

تهدف دراسة الانحدار التنبؤ بقيمة متغير يعرف بالمتغير التابع (dependent) ويرمز له (y)، ويقاس دون خطأ بمعرفة متغير آخر يعرف بالمتغير المستقل (Independent) ويرمز له (x).

فخط الانحدار يترجم العلاقة بين المتغيرين في صورة رياضية والغرض من ذلك هو التنبؤ بقيمة إحدى الظاهرتين إذا علمت قيمة الظاهرة الأخرى فيمكن بواسطته تقدير قيمة أحد المتغيرين بمعلومية قيمة المتغير الآخر عن طريق معادلة الانحدار.

والغرض من استخدام أسلوب تحليل الانحدار الخطي البسيط، هو دراسة وتحليل أثر متغير كمي على متغير كمي آخر، ومن الأمثلة على ذلك ما يلي:

- دراسة أثر كمية السماد على إنتاجية الدونم.
- دراسة أثر الإنتاج على التكلفة.
- دراسة أثر كمية البروتين التي تتناولها الأبقار على الزيادة في الوزن.
- أثر الدخل على الإنفاق الاستهلاكي.

**أنواع الانحدار:**

1. **الانحدار الخطي البسيط**: وفيه المتغير التابع y يعتمد على متغير مستقل واحد x في صورة خطية أي تقترب من الخط المستقيم.
2. **الانحدار المتعدد**: وفيه المتغير التابع y يعتمد على أكثر من متغير مستقل.

**3. الانحدار غير الخطي**: إذا كانت العلاقة بين المتغير y والمتغيرات المستقلة غير الخطية كأن تكون المعادلة من الدرجة الثانية أو الأسية.

**(9–1) تقدير خط الانحدار للبيانات غير المبوبة:**

**(9–1.1) خط انحدار y على x**

لو فرضنا (X) متغير مستقل(Independent variable)، (Y) متغير تابع (dependent variable) وأن العلاقة بينهما علاقة خطية فإنه يمكن وصف العلاقة بالمعادلة الآتية:

$$Y = ax + b$$

حيث a: ميل الخط المستقيم ويسمى معامل انحدار y على x [ميل الخط المستقيم أو معامل انحدار yعلى x(y/x)] أي معدل تغير y إلى تغير x وهي تشير لمعدل الزيادة أو النقصان في الظاهرة.

ويتم تقديرها بالعادلة التالية:

$$a = \frac{n\sum xy-(\sum x\sum y)}{n\sum X^2-(\sum x)^2}$$

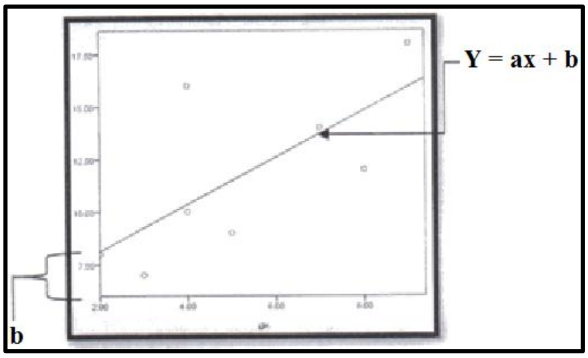
**b**: ثابت الانحدار أو طول الجزء المقطوع من محور الصادات (الرأسي) محور y.

ويتم تقديره بالعادلة التالية:

$$b = \frac{\sum y-(a\sum x)}{n}$$

ويتضح ذلك من البيانات التى يمثلها الشكل التالى:

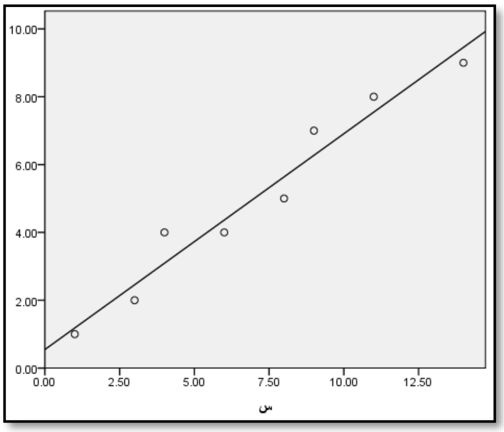
| X | 4 | 5 | 2 | 3 | 4 | 8 | 7 | 9 | 4 | 5 | 2 | 3 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| y | 16 | 9 | 8 | 7 | 10 | 12 | 14 | 18 | 16 | 9 | 8 | 7 |




**شكل (1.9)**: يوضح خط انحدار y على x

**مثال**: أوجد خط انحدار y على x.

| X | 1 | 3 | 4 | 6 | 8 | 9 | 11 | 14 |
|---|---|---|---|---|---|---|---|---|
| y | 1 | 2 | 4 | 4 | 5 | 7 | 8 | 9 |




**شكل(2.9)**: يوضح محاولة توفيق خط مستقيم للمتغيرين (x ، y) للمثال السابق.

**الحل**: تكوين الجدول التالي:

| n | x | y | xy | $X^2$ | $Y^2$ |
|---|---|---|---|---|---|
| 4 | 9 | 6 | 2 | 3 | 2 |
| 1 | 1 | 1 | 1 | 1 | 1 |
| 2 | 3 | 2 | 5 | 9 | 4 |
| 3 | 4 | 4 | 16 | 16 | 16 |
| 4 | 6 | 4 | 24 | 36 | 16 |
| 5 | 8 | 5 | 40 | 64 | 25 |
| 6 | 9 | 7 | 63 | 81 | 49 |
| 7 | 11 | 8 | 88 | 121 | 64 |
| 8 | 14 | 9 | 126 | 196 | 81 |
| $\sum$ | 56 | 40 | 364 | 524 | 256 |

$$a = \frac{n\sum xy - (\sum x \sum y)}{n\sum X^2 - (\sum x)^2}$$

$$a = \frac{(8\times364)-(56\times40)}{(8\times524)-(56\times56)} = \frac{2912-2240}{4192-3136} = \frac{672}{1056} = 0.636$$

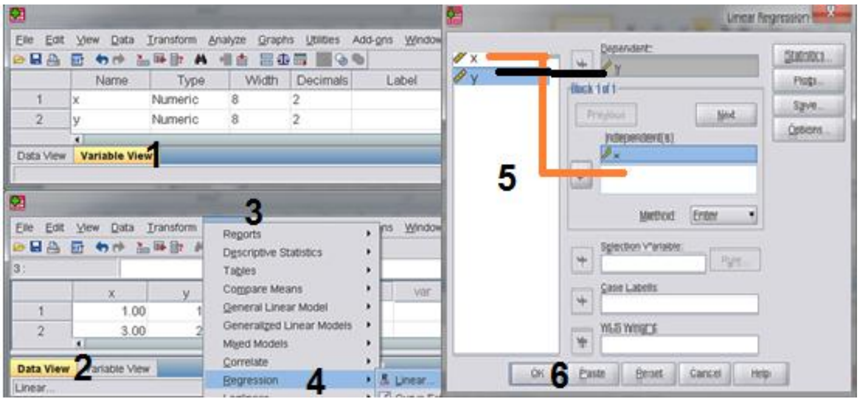
$$b = \frac{\sum y - (a\sum x)}{n}$$

$$b = \frac{40-(0.636\times56)}{8} = \frac{40-35.616}{8} = \frac{4.384}{8} = 0.548$$

$$\because Y = b + ax$$

$$Y = 0.548 + 0.636x$$

وباستخدام spss نتبع الخطوات التالية:




نحصل على النتائج التالية:

**Variables Entered/Removed[b]**

| Model | Variables Entered | Variables Removed | Method |
|---|---|---|---|
| 1 | x[a] | . | Enter |
| a. All requested variables entered. | | | |
| b. Dependent Variable: y | | | |

**Model Summary**

| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate |
|---|---|---|---|---|
| 1 | .977[a] | .955 | .947 | .65134 |
| a. Predictors: (Constant), x | | | | |

**ANOVA[b]**

| Model | | Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|---|
| 1 | Regression | 53.455 | 1 | 53.455 | 126.000 | .000[a] |
| | Residual | 2.545 | 6 | .424 | | |
| | Total | 56.000 | 7 | | | |
| a. Predictors: (Constant), x | | | | | | |
| b. Dependent Variable: y | | | | | | |

وتحتوي النتائج على جدول ANOVA والذي يوضح مقدار ما يفسر النموذج الخطي من البيانات، وتشير القيمة الكبيرة لـF أن نموذج الانحدار الخطي يفسر جزء كبير من البيانات وأن الاختلافات العشوائية قليلة. ويختبر جدول ANOVA معنوية النموذج باستخدام وتوزيع F، حيث يختبر الفرضية التالية:

$$H_0: a = 0$$
$$H_1: a \neq 0$$

وبناءً على الجدول السابق فإن فرضية العدم مرفوضة. وفي الحقيقة فإن اختبار F يختبر معنوية العلاقة بين المتغيرين ولا يختبر معنوية المعلمة **a** كمعلمة للنموذج. وتبرز أهمية اختبار F عندما يوجد في النموذج أكثر من متغير مستقل واحد، فإن اختبار F في هذه الحالة يختبر معنوية جميع معالم النموذج دفعة واحدة.

وفي حالة رفض فرضية العدم ننتقل إلى اختبار معالم النموذج بحيث يتم اختبار معنوية كل معلم من معالم النموذج بصورة منفصلة عن الآخر، أما إذا لم يتم رفض فرضية العدم، فليس هناك حاجة لاختيار معالم النموذج مما يشير إلى أن النموذج المستخدم غير مناسب. كذلك فإنه يمكن حساب معامل التحديد للنموذج .

وبذلك فإن النموذج يفسر 95.4% من الاختلافات في قيم المتغير التابع، في حين 4.6 من الاختلافات ناتجة من عوامل عشوائية. ولمعرفة مدى تشتت الخطأ العشوائي حول خط الانحدار، يستخدم متوسط مجموع مربعات الفروق للبواقي Residuals والذي يساوي:

$$R^2 = \frac{SSR}{SST} = \frac{53.455}{56.000} = 0.95455$$

$$MSE = \frac{2.545}{6} = 0.424$$

| Coefficients[a] | | | | | | |
|---|---|---|---|---|---|---|
| Model | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. |
| | | B | Std. Error | Beta | | |
| 1 | (Constant) | .545 | .459 | | 1.189 | .279 |
| | x | .636 | .057 | .977 | 11.225 | .000 |
| a. Dependent Variable: y | | | | | | |

وتشير قيم MSE الصغيرة إلى تركز البيانات حول خط الانحدار، إلا أنه يجب ألا نهمل تأثير قيم المتغير التابع الكبيرة على تباين الخطأ العشوائي.

ويحتوي الجدول الثاني على تقديرات لمعالم النموذج مع اختبارات T لمعنوية معالم النموذج، حيث كانت قيمة (b = 0.545) (constant) وقيمة (a = 0.636)، وبناءً على ذلك فإن النموذج المقدر هو: Y = 0.545 + 0.636x .

**(9-1. 2) خصائص خط الانحدار ومعامل الانحدار:**

أ. **خصائص خط الانحدار له خاصيتان رئيسيتان هما:**

**1.** مجموع انحرافات نقط الانتشار حوله هو نهاية صغرى.

**2.** خط الانحدار يمر بنقطتي الوسط الحسابي لكل من المتغيرين س، ص أي يمر بالنقطة س، ص.

**ب. خصائص معامل الانحدار:**

**1.** لا يتأثر بالجمع أو الطرح ولكن يتأثر بالضرب والقسمة.

**2**. يمكن أن يأخذ قيماً أكبر من (+ 1) أو أقل من (− 1) فهو معدل تغير كل ما هنالك أنه إذا كان موجباً فإن العلاقة ستكون طردية أو أن التقاطع تتجمع حول خط يصنع زاوية حادة مع المحور الأفقي، وإذا كان سالباً فإن النقاط تتجمع حول خط يصنع زاوية منفرجة مع المحور الأفقي والعلاقة عكسية، أما إذا كان مساوياً للصفر فهذا يعني أنه لا توجد علاقة بين المتغيرين.

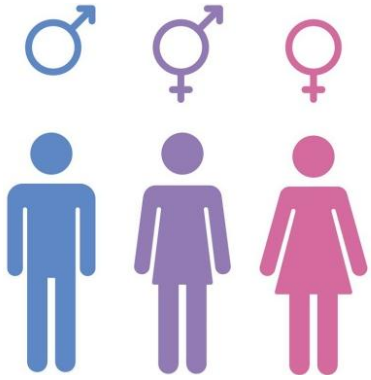
**تعليق**:

يتضح أن معادلة الانحدار هى معادلة تنبئية لتوضيح العلاقة بين المتغيرين، بحيث يمكن عن طريقها التنبؤ بقيمة لا تكون معروفة لدى الباحث ولا يفهم مطلقاً أن القيمة المقدرة تقديراً تنبئياً لابد وأن تنطبق تماماً على القيمة الحقيقية التي تنتج عن البحث الواقعي، ولكن المقصود هو أنه لو أجري البحث على حالات كثيرة العدد فإن متوسط القيم يكون قريباً جداً من القيمة النظرية الناتجة من المعادلة الانحدارية.

# الباب العاشر

# الإحصاءات الحيوية

# Vital Statistics




مقدمة

(10-1) أهمية تعداد السكان.

(10-2) أهمية دراسة السكان.

(10-3) مفاهيم ومصطلحات.

(10-4) الهرم السكاني.

(10-1.4) خصائص الهرم السكاني.

(10-2.4) أهمية الهرم السكاني.

(10-3.4) العلاقة بين جغرافيا السكان والديموغرافيا.

(10-5) العوامل المؤثرة في توزيع السكان.

(10-6) البيانات السكانية.

(10-7) مصادر البيانات الديموغرافية.

(10-1.7) التعداد السكاني.

(10-2.7) أساليب التعداد السكاني.

(10-3.7) خصائص التعداد.

(10-4.7) أهمية التعداد السكاني.

(10-5.7) التسجيل الحيوي.

(10-6.7) المسح الميداني.

(10-8) تقدير السكان.

(10-1.8) التغير المطلق في عدد السكان بين تعدادين.

(10-2.8) عدد السكان في نهاية فترة زمنية.

(10-3.8) الطريقة الحسابية.

(10-9) تقدير العوامل الديموغرافية.

(10-1.9) التغير المئوي في عدد السكان بين تعدادين.

(10-2.9) معدل الزيادة السنوية في عدد السكان.

(10-3.9) الكثافة السكانية.

(10-4.9) العوامل المؤثرة على كثافة السكان.

(10-9.9) تقدير كثافة السكان وتوزيعهم

(10-6.9) كثافة السكان.

(10-7.9) الخصوبة.

(10-8.9) إحصائيات الخصوبة.

(10-9.9) إحصائيات المواليد.

(10-10.9) إحصائيات الوفاة.

(9-11.9) إحصائيات الزواج.

(10-12.9) إحصائيات الهجرة.

**مقدمة:**

تعتبر الإحصاءات الحيوية من الإحصاءات الضرورية لأنها تستخدم في التقديرات السكانية التي تستند أساسا إلى أعداد المواليد والوفيات وتشمل أيضا عقود الزواج وواقعات الطلاق التي لها أهميتها في التزايد السكاني إلى جانب حجم الهجرة الصافية.

كما تعتبر هذه الإحصاءات جزءا لحساب العديد من المؤشرات الديموغرافية في عمليات التخطيط الاقتصادي والاجتماعي والتي تشتمل أيضا على الاتجاهات في النمو السكاني والتوزيع السكاني، ويتوقف تحديث البيانات المتعلقة بالسكان خلال الفترات الفاصلة بين التعدادات على توافر الإحصاءات الحيوية، فدراسة السكان هي خطوة أساسية للتخطيط لمستقبل أفضل ولتوفير احتياجات السكان.

فإحصاء السكان أو الإحصاء السكاني (Census) هو عملية منهجية موحدة وغالباً ما تكون رسمية أو حكومية، تذهب أبعد من تعداد السكان فحسب لتشمل جمع البيانات الديموغرافية والاقتصادية والاجتماعية التي تنطبق في فترة زمنية محدودة على كافة الأشخاص في بلد ما أو جزء محدد منه، وتجميع هذه البيانات وتحليلها ونشرها، وهو عملية دورية يتم خلالها عد السكان رسمياً.

فتعداد السكان هو العملية الشاملة لجمع وتصنيف وتحليل وتقييم ونشر المعلومات المتعلقة بأعداد السكان، أو بنظرة أخرى هي توفير البيانات الديموغرافية والاقتصادية والاجتماعية لجميع الأفراد في الدولة أو لجزء منها خلال مدة زمنية محددة.

**(10–1) أهمية تعداد السكان:**

**1.** تحقيق وتوفير البيانات والمعلومات المطلوبة لأغراض التخطيط والإحصاء.

**2.** المصدر الأكثر شمولاً في الغالب للحصول على معلومات عن السكان والمساكن.

**3.** توضح حجم السكان، وخصائصه وكيفية توزيعه في مختلف أنحاء المناطق داخل البلد.

**4.** تعتبر كمصدر مرجعي للبيانات ومرجعاً لجميع الإحصاءات السكانية والاجتماعية المرتبطة بعدد السكان.

**5.** يعتبر التعداد ممارسةً ذات رابط وطني ولا يتوقف عند اعتباره نشاط روتيني حكومي.

**(10–2) أهمية دراسة السكان:**

**1.** معرفة اتجاهات النمو السكاني من حيث الزيادة والنقصان وتحديد الإجراءات والسياسة المطلوبة.

**2.** تحديد السكان ممن هم في سن النشاط الاقتصادي وتقدير إمكانيات المجتمع من الموارد البشرية ومن القوة العاملة المنتجة.

**3.** إعداد قاعدة عريضة من البيانات، واستخدامها كأساس موثوق به في إجراء الدراسات والبحوث التي تتطلبها برامج التنمية.

**4.** مؤشر أساسي لمعرفة احتياجات المجتمع من التعليم والصحة والقوى العاملة في مجالات العمل الزراعي الصناعي التجاري.

**5.** إعطاء صورة دقيقة عن أوضاع المساكن وخصائصها بما يساعد على وضع خطط الإسكان وخطط البناء والتشييد المستقبلية.

**6.** رصد التغيرات الديموغرافية والاجتماعية والاقتصادية التي طرأت على السكان خلال تلك الفترة في مختلف التقسيمات الإدارية.

**7.** توفير بيانات عن خصائص المباني سواء التي تتبع القطاع العام أو الخاص وحالة إشغالها حتى يمكن تحديد الاحتياجات المختلفة في المستقبل.

**8.** أساسية للتخطيط لمستقبل أفضل ولتوفير احتياجات السكان، وتحديد التزايد السكاني في المستقبل، وبالتالي تقدير عدد السكان في السنوات المقبلة.

**9.** توفير إطار حديث لكافة الأبحاث الإحصائية المتخصصة التي تجرى بأسلوب العينة مثل بحث القوى العاملة، ومسح إنفاق ودخل الأسرة، والبحوث الديموغرافية كالخصوبة والوفيات والهجرة، والبحوث الصحية والتعليمية، والبحوث الخاصة بالمساكن.

**10.** توفير البيانات والمؤشرات السكانية دورياً لقياس التغير الحادث في الخصائص السكانية مع مرور فترة من الزمن، وإجراء المقارنات المحلية والإقليمية والدولية، ومراجعة وتقييم التقديرات السكانية المستقبلية.

**11.** جمع ونشر البيانات الديموغرافية والاجتماعية والاقتصادية للسكان بهدف توفير متطلبات الدولة من البيانات الأساسية عن السكان والمساكن لإعداد خطط التنمية، واحتياجات المخططين والباحثين.

**12.** دراسة التغيرات السكانية المتعلقة بالخصوبة والمواليد والوفيات والهجرة وعلاقتها بمتغيرات النوع (ذكور وإناث) والعمر وتحديد الأساليب والتوجهات الملائمة لكل هذه التغيرات حتى لا تقود إلى خلل ديموغرافي.

**(10–3) مفاهيم ومصطلحات**:

- الإنجاب: Fertility هو العدد الفعلي للولادات.
- الإحصاء السكاني: قيام الدولة بحصر عدد السكان.
- الحضر: هو جميع المدن والأقسام والشياخات في أي محافظة.
- الواحات: مناطق منخفضة في الصحراء تتوافر فيها مياه جوفية.
- الريف: هو جميع القرى والعزب والكفور والنجوع في أي محافظة.
- توزيع السكان: الأماكن التي يعيش فيها السكان ويعملون بها.
- الكثافة السكانية: قسمة عدد السكان على المساحة الكلية للدولة.
- الخصوبة: مفهوم عام يوضح دور المواليد في التغير السكاني (بالزيادة أو النقصان) والإحلال.
- التعداد: هو حصر شامل لكل الأفراد في الدولة مع جمع معلومات عن أهم خصائصهم في فترة زمنية محددة.
- القوة البشرية بمنطقة: هو عدد الأشخاص الموجودين على قيد الحياة داخل حدود هذه المنطقة في الوقت المعين.
- القدرة على الإنجاب: Fecundity هي المقدرة الطبيعية على الإنجاب أي الحد الأقصى للإنجاب الذي يمكن أن يتحقق.

- المواليد (المولود الحي): نتاج الخروج أو الاستخراج التام من رحم الأم بعد الحمل وظهور أي علامة من علامات الحياة عليه مثل: التنفس أو النبض بالحبل السري أو نبض القلب بغض النظر عن مدة الحمل أو قطع الحبل السري أو انفصال المشيمة.
- مجتمع الدراسات السكانية: وهو يدرس السكان حسب الخصائص الاجتماعية والاقتصادية واللغوية وغيرها، وبمعنى آخر هو علم الدراسات السكانية لذلك يجب الربط فيما بين العلمين عند عمل دراسات أو أبحاث سكانية.
- الجغرافية السكانية: تهتم بدراسة الاختلافات المكانية في توزيع الظاهرات السكانية، ومدى تأثير الخصائص المكانية على توزيع السكان وتفسير سلوكهم وفق المكان فسلوكيات الإنسان تتأثر بالمكان، فالإنسان المقيم في البيئة البدوية تختلف سلوكياته عمن يقيم في البيئة الزراعية أو البيئة الصناعية. الخ.
- الديموغرافي: Demography هو العلم الذي يبحث في خصائص السكان وهو يتكون من الكلمتين Demoss وهي كلمة إغريقية ومعناها البشر، وكلمة Graph ومعناها تصوير وحرف y نهاية الكلمة لتدل على علم من العلوم، وبذا يصير معناها علم تصوير البشر أو بالمعنى الأدق علم خصائص السكان.

فالديموغرافيا تعتمد على لغة الأرقام والتحليل الديموغرافي؛ لذا يجب على الدارس لجغرافيا السكان أن يكون ملماً ببعض أساليب الإحصاء الوصفي والسكاني ليستطيع وصف وتفسير الظواهر السكانية، والجدول التالي يوضح أهم البيانات الديموغرافية.

جدول يوضح أمثلة للمؤشرات الإحصائية الديموغرافية

| | |
|---|---|
| – الكثافة السكانية بالنسبة للمساحة الكلية | ... نسمة / كم 2 |
| – الكثافة السكانية بالنسبة للمساحة المأهولة | ... نسمة / كم 2 |
| – نسبة الزيادة الطبيعية | ... % |
| – نسبة قوة العمل إلى إجمالي عدد السكان | ... % |
| –كثافة الفصل بالتعليم قبل الجامعي | ... تلميذ/ فصل |
| – نصيب الفصل من المدرسين | ... مدرس / فصل |
| – نسبة الأمية | ... % |
| – نصيب السكان من الأطباء | ... نسمة / طبيب |
| – نصيب السكان من الممرضين | ... نسمة / ممرض |
| – نصيب السكان من الأسرة | ... نسمة / سرير |
| – نصيب السكان من سيارات الإسعاف | ... نسمة / سيارة |
| – نصيب السكان من نقاط الإسعاف | ... نسمة / نقطة |
| – نصيب السكان من وحدات تنظيم الأسرة | ... نسمة / وحدة |
| – نسبة المترددات على مراكز تنظيم الأسرة | ... % |
| –نصيب السكان من المجمعات الاستهلاكية | ... نسمة / مجمع استهلاكي |
| – نصيب السكان من قصور الثقافة | ... نسمة / قصر وبيت |
| – نصيب الفرد من إجمالي كمية مياه الشرب المنتجة | ... لتر / يوم |
| – نصيب الفرد من إجمالي كمية مياه الشرب المستهلكة | ... لتر / يوم |
| – نسبة الأسر المتصلة بشبكة مياه الشرب | ... % |
| – نصيب الفرد من إجمالي طاقة الصرف الصحي | ... لتر / يوم |
| – نسبة الأسر المتصلة بشبكة مياه الصرف الصحي | ... % |
| – نصيب الفرد من الكهرباء المستهلكة للإنارة | ... ك . و . س سنويا / فرد |
| – الكثافة التليفونية | ... خط / 100 نسمة |
| – نصيب السكان من إجمالي مكاتب البريد الحكومية | ... نسمة / مكتب |
| – نصيب السكان من مراكز التدريب المهني | ... نسمة / مركز |
| – نسبة إشغال الغرف بالفنادق السياحية | ... % |
| – نسبة إشغال الأسرة بالفنادق السياحية | ... % |
| – نصيب السكان من الهيئات الشبابية | ... نسمة / مركز ونادى ولجنة رياضية |
| – نصيب السكان من وحدات الخدمة الاجتماعية | ... نسمة/ وحدة اجتماعية |

**(10–4) الهرم السكاني**:

عرض بياني للسكان حسب العمر والنوع.

**(10–4 أ) خصائص الهرم السكاني:**

**1.** التعبير عن الأعمار إما سنة سنة، أو كل خمس أعوام.

**2.** دائما ما يتم رسم الهرم السكاني بوضع السكان الذكور على اليسار والسكان الإناث على اليمين.

**3.** الفئات العمرية التي تزيد عن سن معين (85 سنة مثلا) عادة ما تهمل عند رسم الهرم السكاني.

يوضع صغار السن في قاع الهرم السكاني وكبار السن في قمة الهرم السكاني، ويمكن تصنيفه كالتالي صغار السن (من 0 إلى 14)، والشباب (القوى العاملة) (من 15 إلى 60)، وكبار السن (من 60 فما فوق).

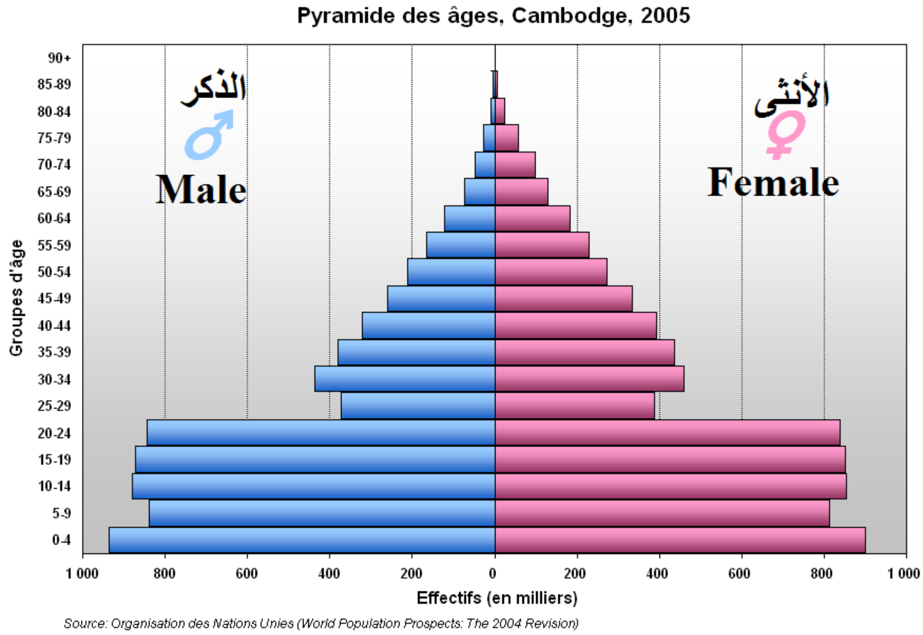
**4.** مقياس الرسم في قاعدة الهرم يمكن أن يعبر عن أعداد السكان في الفئات العمرية المختلفة، أو عن نسبة السكان في الفئات العمرية المختلفة إلى مجموع السكان، واستخدام أي من الأسلوبين لن يؤثر على شكل الهرم السكاني، ولكن من الضروري عند حساب نسبة السكان في المجموعات العمرية المختلفة أن يتم نسبة كل من الذكور والإناث في المجموعة العمرية إلى مجموع السكان أما إذا نسب الذكور إلى مجموع السكان والإناث إلى مجموع السكان بشكل منفرد فإن ذلك من شأنه أن يعطي صورة مزيفة للهرم السكاني في المجتمع لأن ذلك لن يعكس

الأعداد المختلفة لكل من الذكور والإناث في المجتمع لأن المساحة في أي من جانبي الهرم ستكون واحدة.

**(10–4 ب) أهمية الهرم السكاني:**

وضع تخطيط شامل لكافة الجوانب مثل التعليم والصحة وغيرها.

والشكل التالي يوضح مثال للهرم السكاني.




**(10–4 ج) العلاقة بين جغرافيا السكان والديموغرافيا:**

تتناول الجغرافيا السكانية ثلاثة جوانب: نمو السكان، وتوزيع السكان على سطح الأرض، وتركيب السكان بينما الديموغرافي تهتم بالأرقام والطرق الإحصائية، وباحث جغرافيا السكان يربط هذه الأرقام بالبيئة، ويعتمد في تحليله على الخرائط المتعددة.

فالعلاقات المكانية تميز جغرافيا السكان عن الديموغرافيا التي تتناول السكان رقمياً كموضوع مستقل عن البيئة. ولا يمكن الفصل بين جغرافيا السكان والديموغرافيا، فهناك علاقة تكاملية بينهما حيث أن الديموغرافيا تتناول الجانب الرقمي والجغرافيا الجانب التحليلي لتحديد الإطار المكاني الصحيح وتوضيح مختلف العوامل التي تحكم علاقات السكان ببيئتهم داخل حدود هذه الإطار.

**(10–5) العوامل المؤثرة في توزيع السكان:**

**1.** عوامل طبيعية: التربة والماء، وأشكال التضاريس، والمناخ.

**2.** عوامل بشرية: الحرفة وفرص العمل، والمواصلات، والخدمات المتوفرة.

**3.** الهجرة الداخلية: هجرة سكان الريف إلى العواصم والمدن الكبرى.

**4.** الهجرة الخارجية: الهجرة إلى الدول الخارجية بحثا عن ظروف معيشية أفضل.

**(10–6) البيانات السكانية:**

**1.**البيانات الثابتة: وهي تشمل توزيع السكان وخصائصهم مثل النوع والعمر والحالة الزوجية والأنشطة الاقتصادية ... الخ، مثل: التعدادات والمسح بالعينة.

**2.**البيانات المتغيرة: وهي تشمل سجلات المواليد والوفيات وحالات الزواج والطلاق وسجلات الهجرة مثل:

**1.** سجلات المواليد الأحياء: وتتضمن بيانات عن المولود وجنسه واسمه وتاريخ ولادته ومكانها وتاريخ التسجيل. كذلك يتم الحصول على أهم خصائص الوالدين كمكان الإقامة وتاريخ الزواج والمهنة، والحالة التعليمية والدينية والجنسية والعمر لكل منهما.

**2.** سجلات الوفاة: وتتضمن بيانات عن المتوفى عمره – جنسه – مكان الإقامة المعتاد – الحالة الزوجية – عدد الأطفال – الحالة الدينية – بيانات عن حادثة الوفاة تاريخها – مكانها – سببها وتاريخ تسجيلها.

**3.** سجلات وفيات الأجنة: وتتضمن بيانات مشابهة التي تجمع عن المولود الحي بالإضافة إلى بعض البيانات عن حادثة وفاة الجنين.

**4.** سجلات الزواج: وتتضمن بيانات عن مكان الزواج وتاريخه وكذلك البيانات المتعلقة بكل من الزوجين مكان الإقامة – العمر– الحالة التعليمية – الدينية.

**5.** سجلات الطلاق: وتتضمن بيانات مشابهة لتلك التي تجمع في الزواج بالإضافة إلى تاريخ الزواج كذلك في العديد من البلدان تجمع بيانات عن أحداث حيوية مثل: الانفصال بين الزوجين.

**6.** سجلات الهجرة: يتزايد السكان نتيجة الزيادة الطبيعية (الفرق بين الولادات والوفيات) أو غير الطبيعية (الهجرة) فاختلاف الرأي حول ما هو مهاجر لا يسمح بضبط هذه السجلات فبيانات هذه الأخيرة أقل قيمة من الإحصاءات الحيوية وذلك لأنها لا تحتوي على كل الحقائق المرتبطة بالمهاجرين من ناحية، فبعض الدول تعتمد في إحصاء المهاجرين على تصاريح السفر، والبعض الآخر يأخذ بالبيانات التي تسجلها مصالح الموانئ والحدود إلى غير ذلك من الطرق.

### (10–7) مصادر البيانات الديموغرافية:

تعتمد الدراسات السكانية على مجموعة من المصادر الإحصائية ذلك لأنها تتناول دراسة أحوال السكان في وقت معين بما في ذلك توزيعهم الجغرافي وتركيبهم المتعدد الجوانب كما تدرس حركة السكان الطبيعية وغير الطبيعية *الهجرة* وما ينتج عنها من زيادة أو نقصان في حجم السكان ويمكن تقسيمها إلى مجموعتين رئيسيتين: مصادر البيانات الثابتة ويمثلها التعداد، المسح بالعينة، مصادر البيانات الغير الثابتة ويمثلها التسجيل الحيوي، وسجلات الهجرة.

### (10–1.7) التعداد السكاني:

توضح البيانات الإحصائية (Statistical Date) كم من الناس أو الأحداث في تاريخ معين أو في فترة زمنية معينة فهي تمدنا بالأعداد الكمية الكافية لتحقيق بعض أغراض التحليل السكاني مثل: (أعداد المواليد التي وقعت في الأعوام السابقة، والفئات العمرية للسكان)، والواقع أن ترجمة كلمة التعداد بالإنكليزية (Census) مستمدة أصلا من الكلمة اللاتينية (Censure) أي بمعنى القيمة، أو الضريبة، وتعداد السكان هو أهم مصدر من المصادر لدراسة السكان وخصائصهم وتغيراتهم، ولعل تعدادات السكان هي أولى العمليات الإحصائية المهمة التي فكر فيها الإنسان منذ القدم لأغراض متعددة (الخدمة العسكرية، أو لجمع الضرائب، أو للقوة العاملة)، وتجدر الإشارة إلى أن أغلبية دول العالم تجري تعدادات لسكانها منذ القدم وإلى وقتنا الحاضر، والتعداد هو: "مجموع العمليات لجمع وإعداد ونشر المعلومات الديموغرافية عن كافة الأشخاص في القطر أو في منطقة معينة في زمن ما".

فالتعداد السكاني هو العملية الكلية لجمع عدد السكان ووضعها في جداول وتوزيعات تكرارية والعرض البياني للتعداد وتجهيز وتقييم وتحليل ونشر البيانات الديموغرافية والاقتصادية والاجتماعية المتعلقة بالسكان الموجودين على قيد الحياة لكافة سكان الدولة أو جزء منها، مثل: الإقليم أو المحافظة في فترة زمنية معينة أو فترات زمنية معينة.

وتعتبر الرسوم البيانية أكثر الأساليب الإحصائية استخداماً لوصف وتلخيص بيانات التعداد، منها:

**1.** الأعمدة.

**2.** الدوائر المجزأة.

**3.** الخطوط البيانية.

**4.** الأعمدة البسيطة.

**5.** الأعمدة المتلاصقة.

**6.** الأعمدة المركبة.

**(2.7–10) أساليب التعداد السكاني:**

**1.** التعداد النظري: حصر السكان بحسب أماكن إقامتهم، أي أن أفراد الأسرة الغائبون أو المهاجرون يدخلون في التعداد مع الأفراد الموجودين، وهذا يساعد على وضع السياسات المستقبلية المتعلقة بأغراض تعليمية أو اجتماعية أو صحية ... الخ.

**2.** التعداد الفعلي: حصر الأشخاص حسب أماكن تواجدهم وقت التعداد بغض النظر عن كونهم من سكان هذا المكان بصفة دائمة أو زائرين له بصفة مؤقتة. وهذه باختيار ليلة التعداد

في فترة تقل فيها الزيارات والتنقلات الرسمية نسبياً. المقيمين في كل منطقة بصفة دائمة؛ ولذلك فهي تصور السكان على غير الحقيقة، خصوصاً إذا كانوا في حركة مستمرة. ويمكن تقليل حركة السكان إلى حدها الأدنى باختيار ليلة التعداد في فترة تقل فيها الزيارات والتنقلات الرسمية نسبياً.

**3**. التعداد الفعلي – النظري: وهو يجمع بين التعدادين السابقين أي يسجل السكان حسب مناطق إقامتهم المعتادة والموجودين في آن واحد.

**(3.7–10) خصائص التعداد**:

**التعداد**: هو العملية الكلية لجمع وتجهيز وتقويم وتحليل ونشر البيانات الديموغرافية والاقتصادية والاجتماعية المتعلقة بكل الأفراد في قطر أو جزء محدد المعالم وفي زمن محدد ويحتوي هذا التحليل على عدد من العناصر الهامة:

**1**. الخصوبة. **2**. التركيب الأسري

**3**. النشاط الاقتصادي. **4**. مجموع عدد السكان

**5**. نمط العمران (حضر– ريف). **6**. النوع والسن والحالة المدنية.

**7**.مكان الميلاد والجنسية ومحل الإقامة. **8**.اللغة الأصلية والحالة التعليمية والدينية.

ويجب أن يتسم بالعديد من النقاط التالية:

**1**. الصفة الرسمية: يتم تحت الإشراف الحكومي بناء على التشريعات والقوانين، ولذلك يتم تخصيص جهاز قائم بذاته للقيام بعملية التعداد على أن تشرع الدولة القوانين التي تكفل سرية البيانات وتحديد السلطات المسؤولة والخاصة بعملية إجراء التعداد.

**2.** الشمولية: يجب أن يشمل التعداد كل فرد في الدولة أو الإقليم سواء كان مواطناً أو أجنبياً وكما يشمل جميع رعايا الدولة في الخارج.

**3.** الآنية: يجب جمع البيانات من كل سكان الأقاليم الإدارية التابعة للدولة في آن واحد وعادة يحدد يوم واحد لهذا الغرض تكون فيه الحياة طبيعية وكما أن السكان من أكثر الظواهر الجغرافية تغيرا ففي كل لحظة يولد إنسان ويموت آخر ويهاجر آخر.

**4.** الفردية: يجب أن يدلي كل فرد بنفسه عن المعلومات الشخصية الواردة في استمارة التعداد، ويجب أن تتضمن البيانات كل الأحوال الشخصية لكل فرد يعيش في المجتمع حتى يمكن تصنيف السكان فيما بعد حسب هذه الخصائص مثل التصنيف حسب العمر، الحالة التعليمية، النشاط الاقتصادي وغير ذلك تصنيفا متقاطعا في جداول التعداد النهائية، وتحديد المناطق المشمولة بالتعداد تحديدا جغرافيا كاملا لكي يمكن بعد ذلك إعداد البيانات حسب الأقاليم الإدارية (المحافظات) في الدولة.

**5.** زمن التطبيق: يجب أن يكتمل العد في زمن محدد فالقاعدة العامة هي تخصيص يوم كامل لهذا الغرض.

**6.** الدورية: يجب أن يجري التعداد بصورة دورية في أوقات منتظمة كل خمس أو عشر سنوات لغرض تسهيل مهمة الباحثين في عملية المقارنة بين الدول أو الأقاليم الإدارية التابعة للدولة (المحافظات)، ومن الأفضل أن تأخذ التعدادات على فترات زمنية متساوية تجري عموما كل عشر سنوات.

**(10–4.7) أهمية التعداد السكاني**:

**1.** تعيين الالتزامات العسكرية والضريبية والعملية للأفراد في المجتمع.

**2.** التعرف على عوامل كالهجرة والخصوبة والخصائص الاقتصادية، ومحددات الأمن الاجتماعي التي تصاحب عملية التنمية الاجتماعية والاقتصادية في المجتمع.

**3.** توفير البيانات الإحصائية حول الخصائص الهامة للسكان والتي تحتاجها الحكومات والمصالح والتعليم وهيئات البحث وجمهور المواطنين لرسم الخطط العلمية ومواجهة المشاكل الحياتية.

ويشمل التعداد موضوعات منها:

**1.** أحوال التعليم.
**2.** دراسة القوة البشرية.
**3.** نمو وتركيب السكان.
**4.** دراسة مستوى المعيشة.
**5.** توزيعهم والهجرة الداخلية.
**6.** مشكلات الغذاء والزراعة.
**7.** دراسة احتياجات السكان.
**8.** مشكلاته والخدمات الصحية ومرفقاتها.

**استمارة التعداد**:

تتضمن العديد من المعلومات منها:

بيانات عن الموقع الجغرافي للأسرة: المحافظة، المدينة، القسم أو المركز، الشياخة أو القرية أو العزبة أو التجمع.

البيانات الأساسية للفرد: الاسم، ترتيبه داخل الأسرة، النوع (ذكر/ أنثى) الصلة، الديانة، الجنسية، تاريخ الميلاد، جهة الميلاد، محل الإقامة السابق، محل الإقامة الحالي، سبب تغير محل الإقامة.

الحالة الزوجية: لكل من بلغ سن الزواج (18 للذكر، 16 للأنثى)، متزوج، غير متزوج، مطلق، أرمل، عدد الزوجات بالعصمة للذكور المسلمين المتزوجين.

أما للنساء المتزوجات والمطلقات والأرامل مدة الحياة الزوجية، عدد المواليد أحياء "ذكور وإناث"، عدد المواليد الباقين على قيد الحياة، الحالة التعليمية للزوج الحالي أو آخر زوج بالنسبة للمطلقات والأرامل.

الحالة التعليمية لمن 10 سنوات فأكثر: أمي، يقرأ أو يكتب، اسم أعلى مؤهل دراسي في حالة الحصول على مؤهل دراسي.

المرحلة التعليمية: نوع المرحلة، منتظم، منتسب.

الحالة العملية: يعمل، لا يعمل.

بيانات جهة العمل: حكومي، عام، خاص، تعاوني، أجنبي دولي، دبلوماسي.

المهنة الحالية والتخصص:

الإعاقات الظاهرة: أعمى، أصم وأبكم، أصم، أبكم، ... الخ.

بيانات عن أفراد الأسرة الذين يعملون خارج الوطن:

بيانات السكن: نوع السكن "فيلا، شقة، حجرة، عشة، بيت ريفي... الخ"، "ملك، إيجار غير مفروش، إيجار مفروش"، قيمة السكن، عدد الحجرات التي تشغلها الأسرة، الإضاءة، المياه.

**(10–5.7) التسجيل الحيوي:**

تسجيل الأحداث الحيوية التي تقع خلال سنة ميلادية، وهي عملية إجبارية تتم عن طريق مشروعات التسجيل المصممة لقيد جميع هذه الأحداث من (مواليد، ووفيات، وهجرة، وحالات الزواج والطلاق) أثناء وقوعها.

**(10–6.7) المسح الميداني:**

وهو أسلوب علمي للحصول على البيانات السكانية، وفيه يتم أخذ عينة من المجتمع بهدف معرفة الخصائص السكانية في فترة معينة فهو يشبه من ناحية التعداد السكاني لأنه يعتمد على سؤال المبحوثين عن خصائصهم، وقد يشبه من ناحية أخرى أسلوب التسجيل الحيوي إذ يسأل الناس عن الواقعات التي حدثت لأفراد أسرهم أو لجيرانهم في شهر أو سنة سابقة، وعادة ما تستخدمه البحوث الاجتماعية للحصول على عينة من السكان لدراسة ظاهرة اجتماعية ما، وتكمن أهميتها في تسجيل البيانات اللازمة للوقوف على بعض جوانب الأحوال السكانية لمجتمع منعزل لم يتعرض لعمليات التسجيل الدورية وبشكل منتظم، فضلاً عن أن طريقة المسح محاولة تجريب نموذج مبتكر لتسجيل الوقائع الحيوية للتأكد من مدى فاعليته.

**(10–8) تقدير السكان:**

**(10–1.8) التغير المطلق في عدد السكان بين تعدادين:**

بمقارنة نتائج تعدادين متتالين تفصلهما فترة زمنية قدرها 10 سنوات في كثير من الأحوال. وأبسط مقياس هو قياس التغير، وهو كالتالي:

**(10 – 2.8) عدد السكان في نهاية فترة زمنية:**

التغير في عدد السكان = تعداد لاحق = تعداد سابق

عدد السكان في نهاية فترة زمنية = عدد السكان في بداية الفترة الزمنية + الزيادة الطبيعية + صافي الهجرة

حيث أن:

الزيادة الطبيعية = عدد المواليد خلال الفترة الزمنية – عدد الوفيات خلال نفس الفترة

صافي الهجرة = عدد المهاجرين للداخل خلال الفترة الزمنية – عدد المهاجرين خلال نفس الفترة

**(10 – 3.8) الطريقة الحسابية:**

**أولاً**: الطريقة الأولى على أساس المتوالية الحسابية (العددية) :

تفترض هذه الطريقة أن مقدار (متوسط) الزيادة السنوية للسكان ثابت ولا يتغير من فترة زمنية لأخرى، أي إنها تمثل متوالية عددية (حسابية):

$$ك_{ن} = ك_{0} + \left(ن \times د\right)$$

حيث $ك_{ن}$ = تقديرعدد السكان بعد ن سنة.

$ك_{0}$ = عدد السكان فى بداية السنة التى يبدأ بعدها التعداد.

ن = عدد السنوات

د = مقدار الزيادة الطبيعية فى السنة الواحدة (متوسط الزيادة السنوية للسكان).

أي:

عدد السكان في سنة معينة = عدد السكان في سنة سابقة + [مقدار الزيادة السنوية × عدد السنوات]

**مثال:**

قدر عدد السكان لدولة ما في 2000/7/1 بـ 35220 ألف نسمة ووصل هذا العدد فى 2010/7/1 إلى 45755 ألف نسمة والمطلوب اتباع طريقة المتوالية العددية لحساب:

**أ**. متوسط الزيادة السنوية فى عدد السكان خلال الفترة.

**ب**. تقدير عدد السكان فى 2014/1/1.

**الحل:**

ن = عدد السنوات = الفرق بين 2000/7/1 و 2010/7/1 = 10 سنوات

$$\because \text{ك}_{\text{ن}} = \text{ك}_{0} + (\text{ن} \times \text{د})$$

$$45755 = 35220 + (10\text{د})$$

$$10535 = (10\text{د})$$

د (متوسط الزيادة السنوية في عدد السكان خلال الفترة) = 10535 ÷ 10 = 1053.5 ألف نسمة

**ب**. تقدير عدد السكان فى 1/ 1/ 1987.

ن = 2014/1/1 − 2010/7/1 = 3.5 سنة

$$\because \text{ك}_{\text{ن}} = \text{ك}_{0} + (\text{ن} \times \text{د})$$

$$\text{ك}_{2014.1.1} = \text{ك}_{2010.7.1} + (\text{ن} \times \text{د})$$

$$ك_{2014.1.1} = 45755 + (1053.5 \times 3.5)$$

$$ك_{2014.1.1} = 45755 + 3687.25 = 49442.25 \text{ ألف نسمة}$$

$$ك_{2014.1.1} = 49442.25 \times 1000 = 49442250 \text{ نسمة}$$

**مثال:**

كان تعداد السكان في مصر سنة 1960 هو 26 مليون نسمة وسنة 1966 هو 30 مليون نسمة فما هو تقدير عدد السكان في سنة 1967، 1971 مستخدما طريقة المتوالية العددية ؟

**الحل:**

الزيادة في عدد السكان 6 سنوات $(1966 - 1960) = 30 - 26 = 4$ مليون نسمة

الزيادة في سنة واحدة = 4 مليون نسمة ÷ 6 = 666.676 نسمة

المدة من 1966 إلى 1967 هي سنة واحدة

$$ك_{ن} = ك_{0} + (ن \times د)$$

$$ك_{67} = ك_{66} + (ن \times د)$$

$$\because ك_{67} = 30000000 + 1 \times 666.676 = 30666676 \text{ نسمة}$$

المدة من 1966 إلى 1971 = 5 سنوات.

$$ك_{71} = ك_{66} + \left(ن \times د\right)$$

$$ك_{71} = 30000000 + 5 \times 666667 = 30000000 + 3333335 = 33333335 \text{ نسمة}$$

**ثانياً: على أساس نظام المتوالية الهندسية:**

عدد السكان في سنة معينة = عدد السكان في سنة سابقة $\times$ $\left(1 + \text{معدل الزيادة السنوية}\right)^{ن}$ أي:

$$ك_{ن} = ك_{0} + \left(1 + ر\right)^{ن}$$

حيث :

ن: عدد السنوات ر: معدل الزيادة السنوية $ك_{ن}$: تقديرعدد السكان بعد ن سنة.

$ك_{0}$: عدد السكان فى بداية السنة التى يبدأ بعدها التعداد.

ويمكن الحل بأخذ اللوغاريتمات للمعادلة السابقة للسهولة.

**مثال:**

قدر عدد السكان لدولة ما فى 2000/7/1 بـ 35220 ألف نسمة ووصل هذا العدد فى 2010/7/1 إلى 45755 ألف نسمة والمطلوب اتباع طريقة المتوالية الهندسية لحساب:

**أ**. متوسط الزيادة السنوية فى عدد السكان خلال الفترة.

**ب**. تقدير عدد السكان فى 2014/1/1.

**الحل:**

ن = عدد السنوات = الفرق بين 2000/7/1 و 2010/7/1 = 10 سنوات

$$ك_ن = ك_0 + \left(1 + ر\right)^ن$$

$\because$ $ك_{2010/7/1}$ = 45755 ألف نسمة؛ $ك_{2000/7/1}$ = 35220 ألف نسمة

$$\therefore\ 45755 = 35220\ (1 + ر)^{10}$$

$$(1 + ر)^{10} = \frac{45755}{35220}$$

بأخذ لوغاريتم الطرفين:

$$10\ لو\left(1 + ر\right) = لو\ 45755 - لو\ 35220$$

$$لو\left(1 + ر\right) = \frac{لو\ 45755 - لو\ 35220}{10}$$

$$لو\left(1 + ر\right) = \frac{4.66044 - 4.54679}{10}$$

$$لو\left(1 + ر\right) = \frac{0.11365}{10} = 0.011365$$

بايجاد اللوغاؤيتم المقابل لـ (0.011365)

$$\therefore\ 1 + ر = 1.0265142$$

$$ر = 1.0265142 - 1 = 0.0265142$$

عدد السكان في 2014/1/1 حيث ن = 3.5

$$\text{ك}_{\text{ن}} = \text{ك}_{0} + \left(1 + \text{ر}\right)^{\text{ن}}$$

$$\text{ك}_{(1.1.2014)} = \text{ك}_{(1.7.2010)} + (1 + 0.0265)^{3.5}$$

$$\text{ك}_{(1.1.2014)} = 45755 + (1 + 0.0265)^{3.5}$$

$$\text{لو ك}_{(1.1.2014)} = \text{لو } 45755 + \left(3.5 \times \text{لو } 1.265\right)$$

$$\text{لو ك}_{(1.1.2014)} = 4.66044 + (3.5 \times 0.01114)$$

$$\text{لو ك}_{(1.1.2014)} = 4.66044 + 0.03899$$

$$\text{لو ك}_{(1.1.2014)} = 4.69943$$

بإيجاد اللوغاريتم المقابل لـ(4.69943)

= 50052.987 ألف نسمة = 50052987 نسمة

**(10–9) تقدير العوامل الديموغرافية:**

**(10–1.9) التغير المئوي فى عدد السكان بين تعدادين:**

**مثال:** احسب التغير فى عدد السكان فى كلا من لبنان وليبيا.

| عدد سكان ليبيا | السنة | عدد سكان لبنان | السنة | البيان |
|---|---|---|---|---|
| 93418000 | 1960 | 43480700 | 1961 | تعداد لاحق |
| 83199000 | 1950 | 356879000 | 1956 | تعداد سابق |
| 10219000 | | 77928000 | | الفرق (التغير المطلق بين تعدادين متتاليين) |

**الحل:**

قياس التغير المئوي فى عدد السكان بين تعدادين:

$$\text{التغير المئوي لعدد سكان بلد} = \frac{\text{التغير المطلق بين تعدادين متتاليين}}{\text{عدد السكان في التعداد السابق}} \times 100$$

$$\text{التغير المئوي لعدد سكان لبنان} = \frac{77928000}{356879000} \times 100 = 21.28\%$$

$$\text{التغير المئوي لعدد سكان ليبيا} = \frac{10219000}{83199000} \times 100 = 21.28\%$$

بمعنى أن عدد سكان لبنان يتزايدون بنسبة 21.83 % لكل 10 سنوات بينما عدد سكان ليبيا يتزايدون بنسبة 12.28% لكل 10 سنوات.

**(10–2.9) معدل الزيادة السنوية في عدد السكان:**

$$\text{معدل الزيادة السنوية في عدد السكان} = \frac{\text{عدد السكان في سنة المقارنة} - \text{عدد السكان في سنة الأساس}}{\text{عدد السنوات}} = \cdots \text{ نسمة}$$

**(10–3.9) الكثافة السكانية:**

هي تعبير عن وجود الإنسان في المكان في وقت محدد، وتتحدد صفة الكثافة السكانية بأربعة عوامل هي:

**1.** المواليد

**2.** الوفيات

**3.** الهجرة الخارجية

**4.** الهجرة الداخلية.

وتعرف بأنها: عدد السكان في الكيلومتر المربع من المساحة المأهولة في سنة محددة.

**(4.9–10) العوامل المؤثرة على كثافة السكان:**

**1.** العوامل الجغرافية:

تزداد كثافة السكان في السهول والوديان الخصبة التي يسهل زراعتها والمعروفة باعتدال مناخها ووفرة مياهها العذبة ومواردها الطبيعية، وتقل في الجبال والأراضي الصحراوية أو المغطاة بالغابات والثلوج حيث يواجه الإنسان مصاعب لاستصلاحها خصوصاً إذا كان طقسها منحرفاً عن حدود الاعتدال وتندر فيها المياه العذبة والموارد الطبيعية.

فالقرب والبعد من مصادرِ المياهِ، والثّروات المعدنيّة، والأنشطة الإنسانيّة الهامّة تُسهم مثل هذه العوامل بشكلٍ كبيرٍ في تغيير الكثافات السّكانيّة بين مناطقِ العالمِ المختلفةِ، فالنّاس إجمالاً يميلون إلى الاستقرار في المناطقِ التي يُمكنهم فيها البدء بأنشطتِهم المختلفة، والتي تُوفّر لهم حاجاتِهم الأساسيّة دون عناءٍ يُذكر.

**ب. العوامل الاقتصادية:**

ترتفع كثافة السكان في البيئات الصناعية المتقدمة اقتصادياً، وتنخفض في المجتمعات الزراعية المختلفة.

**ج. العوامل الاجتماعية:**

ترتفع كثافة السكان في الوحدات الاجتماعية الكبيرة كعواصم المحافظات والمدن حيث تستقر فيها من ناحية الأمن والكسب، وتتوافر سبل الراحة، وتضعف في القرى والكفور والعزب حيث يصعب وجود الظروف المعيشية الملائمة.

**(10–5.9) تقدير كثافة السكان وتوزيعهم:**

الكثافة الحسابية (الخام): هي مجموع السكان في وحدة مساحية معينة مقسوما على المساحة الكلية للوحدة ذاتها.

وهذا النوع من الكثافة لا يعطي إلا فكرة بسيطة عن مدى تركز السكان، وتتناسب فائدته عكسياً مع حجم المساحة الكلية الأرضية، فكلما كبرت المساحة كلما كان مدلول الكثافة الخام سطحياً، ومن ثم فهذه الكثافة ذات أهمية قليلة في دراسة العلاقة بين السكان والموارد.

$$\text{كثافة السكان الخام} = \frac{\text{عدد السكان في دولة (منطقة؛ مدينة ؛ قرية) ما}}{\text{المساحة الكلية لهذه الدولة (المنطقة؛ المدينة ؛ القرية) ما}}$$

**مثال**: إذا كان تعداد أحد البلدان هو 20 مليون نسمة وكانت مساحته 1 مليون كلم$^2$ احسب كثافة السكان.

**الحل**:

كثافة السكان = 20 مليون نسمة ÷ 1 مليون كلم $^2$ = 20 نسمة / كلم$^2$

**ملاحظة:**

لا يصلح هذا المقياس لمقارنة درجة الازدحام في بلدين أحدهما به جزء كبير عبارة بحيرات وصحاري أو جبال، والآخر به أراضي خصبة ومسكونة، لذلك فعند استخدامه للمقارنة نحاول أن نستبعد الأجزاء غير المسكونة أو الغير صالحة للسكن.

**(6.9–10) كثافة السكن:**

هي النسبة بين عدد السكان وعدد الغرف بالبلد جميعها، ويمكن حساب ذلك بقياس متوسط عدد الأشخاص لكل حجرة بالمسكن، أي أن:

$$\text{كثافة السكن} = \frac{\text{إجمالي عدد السكان في البلد (نسمة)}}{\text{عدد الغرف في كل مساكن البلد}}$$

**مثال:**

بفرض أن تعداد السكان في إحدى الدول 50 مليون نسمة في منتصف عام 2000م وكانت مساحة هذه الدولة 4 مليون كم$^2$ وعدد حجرات المساكن 25 مليون حجرة.

**أ**. احسب كلاً من كثافة السكان وكثافة السكن.

**ب**. بفرض أن تعداد السكان لهذه الدولة في منتصف عام 2005م هو 60 مليون نسمة فما هو معدل الزيادة السنوية للسكان.

**الحل:**

$$\text{كثافة السكان الخام} = \frac{\text{عدد السكان في دولة (منطقة؛ مدينة ؛ قرية) ما}}{\text{المساحة الكلية لهذه الدولة (المنطقة؛ المدينة ؛ القرية) ما}}$$

كثافة السكان = 50 مليون نسمة ÷ 4 مليون كلم $^2$ = 12.5 نسمة / كلم$^2$

كثافة السكن = 50 مليون نسمة ÷ 25 مليون حجرة = 2 شخص لكل حجرة

**ب–** تسمى سنة 2000م بسنة الأساس وسنة 2005 م بسنة المقارنة وبالتالي فإن:

معدل الزيادة السكانية = (60 – 50) مليون نسمة ÷ 5 سنوات = 2 مليون نسمة

**(10–7.9) الخصوبة:**

تعد الخصوبة من الظواهر الرئيسة في دراسة السكان فهي المحدد الرئيسي لنمو السكان، وخصوبة السكان لفظ يطلق للدلالة على ظاهرة الإنجاب في أي مجتمع سكاني التي يعبر عنها بعدد المواليد الإحياء، وإن عدد المواليد في أي دولة لا يرتبط ارتباطاً مباشرا بمجموع عدد سكان هذه البلد ولكن يتوقف أساسا على العدد الذي تتوفر لديه مقومات الخصوبة أو الإنجاب.

وللخصوبة جانبان: بيولوجي وهو القدرة على الإنجاب أو الطاقة الإنجابية، وجانب اجتماعي يرتبط بالقرارات المتعلقة بما إذا كان الطفل سيولد أم لا (أخذا في الاعتبار وجود القدرة على الإنجاب) وإذا كان هناك نية لإنجاب الأطفال كم سيكون عددهم داخل الأسرة، كل هذه تخضع للمحيط الاجتماعي الذي يعيش الأفراد فيه.

فالخصوبة هي الأداء الإنجابي للشخص، أي العدد الحقيقي للأطفال المولودين لكل شخص وليس طاقة أو قدرة هذا الشخص على الإنجاب.

ولها مصطلحان:

**1.** الخصوبة الفعلية أو الإنجاب (fertility) هي عملية الإنجاب الفعلي للأطفال.

**2.** الخصـوبة الكامنة أو البيولوجية أو الفسـيولوجية (Fecundity) هي قدرة المرأة على الولادة في سن الحمل متزوجة أو غير متزوجة.

**(8.9–10) إحصائيات الخصوبة:**

معدل الخصوبة العام: General fertility rate

$$\text{معدل الخصوبة العام} = \frac{\text{عدد المواليد أحياء خلال السنة}}{\text{عدد الإناث في سن الحمل } (15-45) \text{ سنة في منتصف السنة}} \times 1000$$

**مثال:**

بفرض أن عدد المواليد أحياء 490000 عام 2000م، وكان عدد الإناث في سن الحمل (15–45) في منتصف السنة 2200000 في إحدى الدول. احسب معدل الخصوبة العام.

**الحل:**

$$\text{معدل الخصوبة العام} = \frac{\text{عدد المواليد أحياء خلال السنة}}{\text{عدد الإناث في سن الحمل } (15-45) \text{ سنة في منتصف السنة}} \times 1000$$

معدل الخصوبة العام $= [490000 \div 2200000] \times 1000 =$ 222.72 مولود حياً لكل

معدل الخصوبة لفئة العمر.

معدل الخصوبة العام $= [490000 \div 2200000] \times 1000 =$ 222.72 مولود حياً لكل

معدل الخصوبة لفئة العمر.

$$\text{معدل الخصوبة لفئة العمر}(-) = \frac{\text{عدد المواليد أحياء في فئة العمر (–) خلال السنة}}{\text{عدد الإناث اللاتي في منتصف السنة للفئة العمرية نفسها (–)}}$$

**مثال**: بفرض أن عدد المواليد أحياء 17308 عام 2000م، وكان عدد الإناث في سن الحمل (من 15–أقل من 20 سنة) 45570 في إحدى الدول. احسب معدل الخصوبة للفئة العمرية (من 15 – أقل من 20 سنة).

**الحل**:

$$\text{معدل الخصوبة لفئة العمر}(-) = \frac{\text{عدد المواليد أحياء في فئة العمر (–) خلال السنة}}{\text{عدد الإناث اللاتي في منتصف السنة للفئة العمرية نفسها (–)}}$$

للفئة العمرية (من 15– أقل من 20 سنة)

$= [17308 \div 45570] \times 1000 =$ 38.8 مولودًا حياً لكل ألف من الإناث للفئة العمرية

(من 15– أقل من 20 سنة).

معدل الأطفال لكل امرأة.

$$\text{معدل الأطفال لكل امرأة} = \frac{\text{عدد الأطفال بين سن صفر – 4 سنوات}}{\text{عدد النساء في سن الحمل (15–45) سنة}} \times 1000$$

**(10–9.9) إحصائيات المواليد**:

تعتبر إحصائيات المواليد عنصراً أساسياً في الإحصائيات الحيوية، ويجب أن تحتوي على هذه البيانات (اسم المولود، تاريخ الميلاد، اسم الوالد، اسم الوالدة، ديانة الأب والأم، جنسية الأب والأم، مهنة الأب، الأم).

$$\text{معدل الولادة العام}\left(\text{معدل المواليد الخام}\right) = \frac{\text{عدد المواليد أحياء في عام}}{\text{عدد السكان عند منتصف هذا العام}} \times 1000$$

$$\text{معدل التوالد} = \frac{\text{عدد المواليد أحياء خلال عام}}{\text{عدد النساء المتزوجات في سن الحمل في منتصف السنة}} \times 1000$$

$$\text{المعدل الإجمالي التوالد} = \frac{\text{عدد المواليد أحياء من الإناث في البلد أثناء السنة}}{\text{عدد النساء المتزوجات في سن الحمل في السنة } (15-49) \text{ سنة}} \times 1000$$

$$\text{المعدل الصافي للتوالد في الفترة من}\cdots\text{إلى}\cdots = \frac{\text{عدد المواليد الإناث اللاتي يبلغن فترة الحمل من}\cdots\text{إلى}\cdots}{\text{عدد النساء في السن من}\cdots\text{إلى}\cdots} \times 1000$$

$$\text{معدل الذكورة في منطقة ما في سنة ما} = \frac{\text{جملة المواليد الذكور الأحياء داخل حدود المنطقة خلال السنة}}{\text{جملة المواليد الإناث الأحياء داخل حدود المنطقة خلال السنة}} \times 1000$$

$$\text{معدل الأطفال لكل امرأة} = \frac{\text{عدد الأطفال بين سن}\left(\text{صفر سنة – أربع سنوات}\right)}{\text{عدد النساء بين سن } (15-45) \text{ سنة}} \times 1000$$

**مثال**: إذا كان عدد الأطفال حتى سن 4 سنوات هو 14649000، وعدد النساء بين سن 15 – 49 هو 14902000، احسب معدل الأطفال لكل امرأة.

**الحل**:

$$\text{معدل الأطفال لكل امرأة} = \frac{\text{عدد الأطفال بين سن}\left(\text{صفر سنة – أربع سنوات}\right)}{\text{عدد النساء بين سن } (15-45) \text{ سنة}} \times 1000$$

$$\text{معدل الأطفال لكل امرأة} = [14649000 \div 14902000] \times 1000 = 983.02$$

أي حوالي 980 طفلاً لكل ألف امرأة في هذا العام.

**مثال:**

كان عدد المواليد أحياء خلال عام 1985 في إحدى الدول هو 2233000 مولود وأن عدد النساء في سن الحمل في منتصف السنة 24845000 منهن 2621000 غير متزوجات، احسب كل من معدل الخصوبة العام ومعدل التوالد لهذه البلد في هذه السنة.

**الحل:**

**أولاً:**

$$\text{معدل الخصوبة العام} = \frac{\text{عدد المواليد أحياء خلال السنة}}{\text{عدد الإناث في سن الحمل (15–45) سنة في منتصف السنة}} \times 1000$$

معدل الخصوبة العام $= [24845000 - 2233000] \times 1000 = 89.877$

أي أن عدد المواليد الأحياء التي حدثت خلال السنة هو حوالي 90 مولود لكل 1000 أنثى في سن الحمل.

**ثانياً:**

$$\text{معدل التوالد} = \frac{\text{عدد المواليد أحياء خلال عام}}{\text{عدد النساء المتزوجات في سن الحمل في منتصف السنة}} \times 1000$$

عدد النساء المتزوجات في منتصف السنة $= 24845000 - 2621000 = 22224000$ امرأة

معدل التوالد $= [22224000 \div 2233000] \times = 1000\ 100.47$

أي أن عدد المواليد الأحياء التي حدثت خلال السنة هو 101 مولودًا لكل ألف أنثى متزوجة في سن الحمل.

**مثال**: بفرض أن عدد المواليد الذكور هو 253329 نسمة وعدد المواليد إناث 237901 نسمة في إحدى المناطق عام 2000م. احسب معدل الذكورة.

**الحل**:

$$\text{معدل الذكورة} = [253329 \div 237901] \times 100 = 106.48$$

**(10-10.9) إحصائيات الوفاة**:

يعبر عنها علميا بأنها: الاختفاء الدائم لكل مظهر من مظاهر الحياة للفرد (مثل التنفس أو النبض أو الحركات اللا إرادية) في أي وقت بعد مولده (أي يجب أن يكون مولودا حيا) ولا يتم إدراج فقد الأجنة فيها.

الوفاة إكلينيكيا: هي الحالات التي تكون بعض أجهزة الجسم تعمل من خلال أجهزة صناعية وهي حالات قليلة مثل وفاة جذع المخ.

حالات فقد الأجنة: تشمل حالات الإجهاض سواء كان إجهاضًا عمديًا أو إجهاضًا غير عمدي كما تشمل حالات المولود الميت.

المولود الميت: هو أي جنين مكتمل أو غير مكتمل في عمر من 28 أسبوع يخرج من رحم الأم وليس به أي مظهر من مظاهر الحياة وذلك بغرض تسهيل الإحصاءات.

شهادة الوفاة: يوجد في شهادة الوفاة بيانات عديدة بهدف تسجيل هذه البيانات للاستفادة منها فيما بعد وأهمها:

**1.** اسم المتوفى.

**2.** النوع (ذكر – أنثى).

**3.** مكان الوفاة وجهة الوفاة.

**4.** تاريخ الوفاة والسن عند الوفاة.

**5.** تاريخ الميلاد وجهة الميلاد.

**6.** الحالة العائلية للمتوفى.

**7.** المهنة.

**8.** أسباب الوفاة وهي تقسم إلى أربع خانات (أسباب):

**9.** السبب المباشر للوفاة.

**10.** الحالة أو نوع الإصابة التي نشأ عنها السبب المباشر.

**أ**. المرض الأصلي أو الظرف الذي أدى إلى هذه الحالة.

**ب**. الحالة التي ربما تكون ساعدت على حدوث الوفاة ولا علاقة لها بالمرض الأصلي.

**المعدل**: هو علاقة بين بسط (وهو جزء من المقام) ومقام (يشتمل على البسط أي أن المقام جزء منه) مع أخذ الزمن.

والصيغ الرياضية لتقديرها:

**معدل الوفيات الخام**:

هو عبارة عن عدد وفيات مجتمع ما (بلد أو منطقة) خلال سنة معينة مقسوما على عدد أفراد نفس المجتمع سنوات معاشه (ولصعوبة حسابه - لعدم تسجيل جميع الأحداث الحيوية وكذلك لتسهيل الإحصاءات كما أن الفرق بينهما لا يؤثر في حجم المجتمع بشكل كبير - يستخدم عدد السكان لنفس المجتمع في نفس السنة) مضروباً في 1000 ويعبر عنه

حسابياً بالشكل التالي:

$$\text{معدل الوفيات الخام} = \frac{\text{عدد الوفيات أثناء السنة}}{\text{عدد السكان في منتصف نفس السنة}} \times 1000$$

معدل وفيات الأطفال حديثي الولادة:

$$\text{معدل المواليد الأموات في منطقة ما في سنة ما} = \frac{\text{عدد المواليد الأموات داخل حدود المنطقة في سنة ما}}{\text{عدد المواليد الأحياء داخل حدود المنطقة في نفس السنة}} \times 1000$$

**مثال**:

بفرض أن عدد المواليد أموات 7000 عام 2000م، وكان عدد المواليد أحياء في نفس السنة 490000 في إحدى المدن. احسب معدل المواليد أموات.

**الحل**:

$$\text{معدل المواليد الأموات في منطقة ما في سنة ما} = \frac{\text{عدد المواليد الأموات داخل حدود المنطقة في سنة ما}}{\text{عدد المواليد الأحياء داخل حدود المنطقة في نفس السنة}} \times 1000$$

$\text{معدل المواليد الأموات} = [490000 - 7000] \times 1000 = 14.28$ مولوداً ميتاً لكل ألف من المواليد الأحياء.

معدل وفيات الأطفال حديثي الولادة:

$$\text{معدل وفيات الأطفال حديثي الولادة} = \frac{\text{عدد الوفيات الذين تقل أعمارهم عن } 28 \text{ يوماً خلال العام}}{\text{عدد الأطفال المولودين أحياء في العام نفسه}} \times 1000$$

معدل وفيات الرضع:

$$\text{معدل وفيات الرضع} \left(\text{أقل من سنة}\right) = \frac{\text{عدد الوفيات التي حدثت بين الأطفال} \left(\text{المواليد أحياء}\right) \text{ خلال السنة}}{\text{عدد المواليد الأحياء في نفس السنة}} \times 1000$$

متوسط معدل الوفيات الخام لفترة:

وذلك يقضي على تذبذب البيانات من سنة لأخرى ويتم حسابه بعدة طرق:

$$\text{معدل الوفاة لفئة العمر}\ (\quad) = \frac{\text{عدد الوفيات لفئة عمرية}\ (\quad)\ \text{خلال سنة معينة}}{\text{عدد السكان لهذه الفئة العمرية}\ (\quad)\ \text{في منتصف السنة}} \times 1000$$

$$\text{معدل وفيات المواليد المتأخرة} = \frac{\text{عدد وفيات المواليد التي تتراوح أعمارها ما بين}\ (4\ \text{أسابيع} - \text{عام})\ (\text{في منطقة معينة في سنة معينة})}{\text{عدد المواليد الأحياء في نفس المنطقة في نفس السنة}} \times 1000$$

معدل الوفيات النوعي (حسب النوع أو الجنس):

وهو المؤشـــر الذي يعتمد على تقســيم الســكان والوفيات إلى ذكور وإناث ويتم حسابه بالصورة التالية (إذا ما كان للذكور):

$$\text{معدل وفيات الذكور} = \frac{\text{عدد وفيات الذكور خلال سنة}}{\text{عدد الذكور في منتصف نفس السنة}} \times 1000$$

$$\text{معدل وفيات الإناث} = \frac{\text{عدد وفيات الإناث خلال سنة}}{\text{عدد الإناث في منتصف نفس السنة}} \times 1000$$

**معدل الوفاة السببي (بسبب مرض معين):**

هو عبارة عن عدد وفيات مجتمع ما (بلد أو منطقة) خلال سنة معينة بسبب مرض معين مقسوما على عدد أفراد نفس المجتمع سنوات معاشه (ولصعوبة حسابه – لعدم تسجيل جميع الأحداث الحيوية وكذلك لتسهيل الإحصاءات كما أن الفرق بينهما لا يؤثر في حجم المجتمع بشكل كبير – يستخدم عدد السكان لنفس المجتمع في نفس السنة) مضروبا في 100000 وذلك بسبب صغر عدد الوفيات بسبب مرض معين ويعبر عنه حسابيا بالشكل التالي:

$$\text{معدل الوفيات الخام بسبب مرض معين} = \frac{\text{عدد الوفيات بسبب المرض خلال سنة}}{\text{عدد السكان في منتصف نفس السنة}} \times 1000$$

**(9–11.9) إحصائيات الزواج:**

$$\text{معدل الزواج الخام في منطقة ما في سنة} \cdots = \frac{\text{عدد الزيجات التي تم عقدها في المنطقة خلال السنة}}{\text{عدد سكان المنطقة في منتصف السنة}} \times 1000$$

**مثال:**

إذا كان عدد عقود الزواج التي تمت في بلد ما سنة 2000م هي 260000، وكان عدد السكان 19000000 في نفس السنة. احسب معدل الزواج الخام.

**الحل:**

$$\text{معدل الزواج الخام في منطقة ما في سنة} \cdots = \frac{\text{عدد الزيجات التي تم عقدها في المنطقة خلال السنة}}{\text{عدد سكان المنطقة في منتصف السنة}} \times 1000$$

$$\text{معدل الزواج الخام} = \frac{260000}{19000000} \times 1000 = 13.68 \text{ في الألف}$$

**(9–12.9) إحصائيات الطلاق:**

$$\text{معدل الطلاق الخام في منطقة ما في سنة} \cdots = \frac{\text{عدد حالات الطلاق التي تمت في المنطقة خلال السنة}}{\text{عدد سكان المنطقة في منتصف السنة}} \times 1000$$

**مثال:**

إذا كان عدد حالات الطلاق التي تمت في بلد ما سنة 2000م هي 26000، وكان عدد السكان 19000000 في نفس السنة. احسب معدل الطلاق الخام.

**الحل:**

$$\text{معدل الطلاق الخام في منطقة ما في سنة} \cdots = \frac{\text{عدد حالات الطلاق التي تمت في المنطقة خلال السنة}}{\text{عدد سكان المنطقة في منتصف السنة}} \times 1000$$

$$\text{معدل الطلاق الخام في منطقة ما في سنة} \cdots = \frac{26000}{19000000} \times 1000 = 1.368 \text{ في الألف}$$

**(13.9−10) إحصائيات الهجرة:**

تعد الهجرة أحد العناصر الثلاثة المسؤولة عن التغير السكاني في مجتمع ما:

**1.** الخصوبة.

**2.** الوفيات.

**3.** الهجرة.

فإن الهجرة من الظواهر السكانية الكفيلة بتغيير الهيكل السكاني لأي مجتمع بصورة سريعة جدا بالمقارنة بآثار كل من الخصوبة والوفيات على الهيكل السكاني والتي تحتاج إلى فترة طويلة جدا من الزمن لكي تؤدي إلى تغيير الهيكل السكاني.

**الهجرة:**

هي انتقال الأفراد من منطقة ما إلى منطقة أخرى سواء كان ذلك داخل حدود الدولة أو خارج حدود الدولة.

**أنواع الهجرة:**

**هجرة داخلية:** الانتقال داخل حدود الدولة، وتتم أساسا من المناطق التي يقل فيها الطلب على العمل إلى المناطق التي تتوافر فيها فرص التوظف، أو تتوافر فيها فرص أفضل للمعيشة،

ومن ثم فإن النمط الغالب للهجرة الداخلية هو من المناطق الريفية إلى المدن، ويلاحظ أن الهجرة الداخلية يكون الدافع من وراءها اقتصاديا بالدرجة الأولى.

**هجرة خارجية**: الانتقال خارج حدود الدولة، فقد تكون الدوافع اقتصادية، أو سياسية مثال ذلك حالة اللاجئين والهاربين والمطاردين من قبل النظم الحاكمة في دولهم، أو قد يكون الدافع علميا، من خلال سعي الفرد إلى فرص تعليمية أفضل أو فرص للبحث أفضل من تلك المتوافرة له في دولته، وغالبا ما يطلق على الهجرة من هذا النوع الأخير لفظ "نزيف العقول".

والهجرة قد تكون بشكل مؤقت وذلك حينما ينوي المهاجر الإقامة في المهجر لمدة مؤقتة ثم العودة مرة أخرى إلى الوطن، أو قد تكون الهجرة دائمة حينما لا ينوي المهاجر العودة مرة أخرى إلى بلده الأصلي.

$$\text{معدل الهجرة للخارج الخام} = \frac{\text{عدد المهاجرين إلى خارج البلد خلال العام}}{\text{تقدير عدد السكان في منتصف نفس العام}} \times 1000$$

$$\text{معدل الهجرة للداخل الخام} = \frac{\text{عدد المهاجرين إلى داخل البلد خلال العام}}{\text{تقدير عدد السكان في منتصف نفس العام}} \times 1000$$

$$\text{معدل صافي الهجرة الخام} = \frac{\text{عدد المهاجرين إلى داخل البلد} - \text{عدد المهاجرين خارج البلد}}{\text{تقدير عدد السكان في منتصف نفس السنة}} \times 1000$$

معدل صافي الهجرة الخام = معدل الهجرة داخل البلد – معدل الهجرة خارج البلد

# المراجع

**المراجع العربية:**


**1**. إبراهيم على وناشد محمود: مبادئ الأساليب الإحصائية. القاهرة، أفراد، 2005.

**2**. أحمد عبادة سرحان: مذكرات في الإحصاء البيولوجي. القاهرة، دار المعارف، 1965

**3**. أحمد عبد السميع طبيه: مبادئ الإحصاء. عمان، دار البداية، 2008.

**4**. السيد محمد خيري: الإحصاء في البحوث النفسية والتربوية والاجتماعية. القاهرة، مطبعة دار التأليف، 1963م.

**5**. أماني موسى محمد: التحليل الإحصائي للبيانات. مشروع الطرق المؤدية للتعليم العالي، مركز تطوير الدراسات العليا والبحوث، كلية الهندسة، جامعة القاهرة، 2007.

**6**. تشارلز هكس: المفاهيم الأساسية في تصميم التجارب. ترجمة قيس سبع خماس، الموصل: وزارة التعليم العالي والبحث العلمي، الجامعة المستنصرية، 1984.

**7**. تيسير حامد أبو سنينة: مفهوم التعداد السكاني، موجود في
http://forum.nooor.com/t40785.html

**8**. دوجلاس ماكنتوش: الإحصاء للمعلمين. ترجمة إبراهيم بسيوني عميرة، القاهرة: دار المعارف،1989.

**9**. زكريا الشربيني: الإحصاء وتصميم التجارب في البحوث النفسية والتربوية والاجتماعية. القاهرة: الأنجلو المصرية، 1995.

**10**. سمير كامل عاشور، سامية أبو الفتوح سالم: مقدمة في الإحصاء الوصفي. أفراد، 1986.

**11.** سمير كامل عاشور، سامية أبو الفتوح سالم: العرض والتحليل الإحصائي باستخدام spsswin. أفراد، 2003.

**12.** ضياء القاضي، لطفي هندي: الإحصاء التطبيقي. مراجعة أحمد على عبد الحليم، القاهرة، جامعة القاهرة، التعليم المفتوح، 1991.

**13.** عبد الرحمن عيسوي: الإحصاء السيكولوجي التطبيقي. دار المعرفة الجامعية، 2000.

**14.** عبد اللطيف عبد الفتاح، أحمد محمد عمر: مقدمة الطرق الإحصائية. القاهرة. مكتبة الجلاء الجديدة، 1978.

**15.** محمد جاسم الياسري ومروان عبد المجيد، الأساليب الإحصائية في مجالات البحوث التربوية، ط1: (عمان، مؤسسة الوراق للنشر والتوزيع، 2001.

**16.** عزت عبد الحميد محمد حسن: الإحصاء النفسي والتربوي تطبيقات باستخدام spss18، دار الفكر العربي، 2011.

**17.** محمد خيري سليم: التحليل الإحصائي لبيانات باستخدام برمجية SPSS، دار صفاء للطباعة والنشر، 2009.

**18.** محمد صلاح الدين صدقي، منى محمد عمار: مبادئ النظرية الإحصائية. القاهرة، مكتبة عين شمس، 1996.

**19.** محمود السيد أبو النيل: الإحصاء النفسي والاجتماعي والتربوي. دار النهضة العربية، 1987.

**20.** مصطفى جلال مصطفى: مقدمة في الإحصاء، ج 1. مكتبة عين شمس،1994.

**21.** موراي ر.شبيجل: ملخصات شوم نظريات ومسائل في الإحصاء. ترجمة شعبان عبد الحميد شعبان، القاهرة، الدار الدولية للنشر والتوزيع، 1989.

## المراجع الأجنبية:

**1.** Altman DG: Practical statistics for medical research. Boca Raton, London, New York, Washington D.C.: Chapman & Hall/CRC 1999.

**2.** Arriaga , Eduardo E. , "population Analysis with Microcomputers" ,VOI. I , New York , 1994.

**3.** Andyfied: Factor Analysis Using SPSS - University of Sussex, http://users.sussex.ac.uk/~andyf/factor.pdf

**4.** ASHOK SHARMA: Discovering Statistics Using SPSS for Windows ANDY FIELD, Turkish Online Journal of Distance Education-TOJDE July 2002 ISSN 1302-6488 Volume:3 Number:3.

**5.** Bortz J: Statistik für Sozialwissenschaftler. Berlin Heidelberg New York: Springer 1999; 5. Auflage: 17– 47.

**6.** David Brink : Essentials of statistics: Exercises. © 2010 David Brink & Ventus Publishing ApS ISBN 978-87-7681-409-0

**7.** Cobb, G. and Moore, D. (2000). "Statistics and Mathematics: Tension and Cooperation," American Mathematical Monthly, pp. 615-630.

**8.** College Entrance Examination Board. Course Description: Statistics. (2004) New York: College Board.

**9.** Conference Board of the Mathematical Sciences. The Mathematical Education of Teachers. Providence, RI,and Washington, DC: American Mathematical Society and Mathematical Association of America,(2001).

**10.** D G Rossiter :An introduction to statistical analysis Overheads, 2006,, available at: http://www.itc.nl/~rossiter/teach/stats/sintro.pdf

**11.** Greenfield MLVH, Kuhn JE, Wojtys EM: A statistics primer: descriptive measures for continuous data. Am J Sports Med 1997; 25: 720–3.

**12.** Holmes, Peter. "Correlation: From Picture toFormula," Teaching Statistics, (2001), 23(3):67–70.
**13.** Joshua M. Tebbs: Introduction to Descriptive Statistics .2006, The University of South Carolina.

**14.** J. Robert Buchanan: Measures of Central Tendency. 2008 available at:

http://banach.millersville.edu/~BobBuchanan/math130/CentralTendency/main.pdf
**15.** Kader, Gary and Perry, Mike. "Learning Statistics with Technology," Mathematics Teaching in the Middle School, 1984 1(2):130–136
**16.** Lawrence D. Brown. Fundamentals of statistical exponential families with applications in statistical decision theory. Institute of Mathematical Statistics Lecture Notes Monograph Series, 9. Institute of Mathematical Statistics, Hayward, CA, 1986. ISBN 0-940600-10-2.
**17.** L. J. Wei, D. Y. Lin, and L. Weissfeld. Regression analysis of multivariate incomplete failure time data by modeling marginal distributions. Journal of the American Statistical Association, 84:1065–1073, 1989.
**18.** Lenth, R. V. (2001) Some Practical Guidelines for Effective Sample Size Determination. The American Statistician, 55, 187-193.
**19.** Douglas C. Montgomery& George C. Runger: Applied Statistics and Probability for Engineers, 2002, Printed in the United States of America.
**20.** Karl Mosler, Friedrich Schmid: Wahrscheinlichkeitsrechnung und schließende Statistik, Springer 2010, Verlag C.H. Beck im Internet:

www.beck.de ISBN 978 3 642 15009 8
**21.** Manfred Piok: Vorgeschmack auf schließende Statistik,**available at** https://btmdx1.mat.unibayreuth.de/kcm/data/Materials/FertigeLernumgebungen/Schliessende_Statistik.pdf
**22.** Marija J. Norusis: Spss for Windows Advanced Statistics Release 6.0, https://aavmqxaxn.updog.co/YWF2bXF4YXhuMDEzMTc4ODIzWA.pdf

**23.** Martin Becker: Schließende Statistik,2017, available at http://www.lehrstab-statistik.de/download/SchlStat/WS201718/SchlStat-WS201718-A01.pdf
**24.** McHugh ML: Descriptive statistics, part I: level of measurement. a. JSPN 2003; 8: 35–7.
**25.** Michael Lavine: Introduction to Statistical Thought, 2013, Overholser BR, Sowinski KM: Biostatistics primer: part I. Nutr ClinPract 2007; 22: 629–35.
**26.** Phil Crewson : Applied Statistics Handbook , Version 1.2, Copyright 2006, AcaStat Software. All rights Reserved, http://www.acastat.com.
**27.** Richard J. Bolton and David J. Hand. Statistical fraud detection: A review. Statistical Science, 17:235–255, 1992
28. Sabine Landau: Statistical Analyses using SPSS, © 2004 by Chapman & Hall/CRC Press LLC.
**29.** Sachs L: Angewandte Statistik: Anwendung statistischer Methoden. a. Berlin, Heidelberg, New York: Springer 2004; 11. Auflage: 1–177.
**30.** Seema Jaggi: Descriptive Statistics and Exploratory Data Analysis, Indian Agricultural Statistics Research Institute, Library Avenue, New Delhi - 110 012 seema@iasri.re
**31.** Scott, D. (1985). Averaged shifted histograms: Eective nonparametric density estimationin several dimensions, Annals of Statistics13: 1024-1040
**32.** SPSS Incorporated: SPSS 16.0 Schneller Einstieg. Dublin: SPSS Inc. 2007; 55–62.
**33.** Stan Brown: Measures of Shape: Skewness and Kurtosis, 2008 available at :http://www.tc3.edu/instruct/sbrown/stat/shape.htm
**34.** Stefan Keppeler: Mathematik II f ̈ur Biologen Schließende Statistik: Hypothesentests, 2008
**35.** Trampisch HJ, Windeler J: Medizinische Statistik. Berlin, eidelberg, New York: Springer 2000; 2. Auflage: 52–82.
**36.** T.S. Tsou and R.M. Royall. Robust likelihoods. Journal of the American Statistical Association, 90:316–320, 1995.

**37.** Tufte, E. R. (2001). The Visual Display of Quantitative Information (2nd ed.) (p. 178). Cheshire, CT: Graphics Press Wolfgang Hardle& Leopold Simar: Applied Multivariate Statistical Analysis, Version, 2003, XploRe is not sold anymore. You can download your free copy of XploRe 4.8 from here, the serial number required during installation is de-0001-48-015.

**38.**Yamani, Taro (1967) ***Statistics, An Introductory Analysis,*** Harper & Row Publishers, New York, USA.

39. Wolfgang Ludwig-Mayerhofer : Willkommen zur Vorlesung Statistik, Universität Siegen – Philosophische Fakultät, Seminar für Sozialwissenschaften, available at, https://www.uni-siegen.de/phil/sozialwissenschaften/soziologie/mitarbeiter/ludwig-mayerhofer/statistik/statistik_downloads/statistik_i_1_online.pdf.

منشورات جامعة عمر المختار 2021