



جامعة دمشق

كلية التربية

تأليف

الدكتور رمضان محمد درويش

الدكتورة عزيزة عبد العال رحمة

الأستاذ المساعد في قسم القياس

المدرسة في قسم القياس والتقويم

والتقويم التربوي والنفسي

التربوي والنفسي

الإحصاء الوصفي

سنة أولى - جميع الاختصاصات

العام الدراسي 2010-2011



الفهرس

الفصل الأول

مقدمة في الإحصاء

Introduction to statistics

19.....	معنى الإحصاء و تعريفه.
19.....	أهمية علم الإحصاء واستخداماته
20.....	أقسام الإحصاء
20.....	1- الإحصاء الوصفي Descriptive statistics
21.....	2- الإحصاء الاستدلالي Inferential statistics
21.....	المتغيرات Variables
22	تصنيف المتغيرات.
26.....	مستويات القياس
26.....	1- المستوى الاسمي أو التصنيفي Nominal Scales
27.....	2- المستوى الترتيبي ordinal scales
29.....	3- المستوى الفئوي Interval Scales

30..... -4 مستوى النسبة Ratio Scales

33..... أسئلة وتمارين

الفصل الثاني

البحث العلمي والإحصاء

Scientific Research and Statistics

37..... خطوات البحث العلمي

37..... 1- تحديد مشكلة البحث

39..... 2- صياغة الفرضيات

40..... 3- منهج البحث

41..... 4- أساليب جمع البيانات الميدانية:

41..... أولاً : أسلوب الحصر (المسح) الشامل

41..... ثانياً : أسلوب العينة

42.....	5- تنظيم البيانات وعرضها .
42.....	6- تحليل البيانات .
42.....	المجتمع الإحصائي والعينة Statistical Population and sample
43.....	العينة تعريفها وأنواعها .
44.....	العينات العشوائية أو الاحتمالية Probabilistic samples:
45.....	1- العينة العشوائية البسيطة Simple Random sample
46.....	2- العينة العشوائية المنتظمة Systematic Random Sample
48.....	3- العينة العشوائية الطبقية Satisfied random sample .
50.....	4- العينة العشوائية العنقودية Cluster Random Sample .
52.....	5- العينة العشوائية المتعددة المراحل Multi-stage sample .
53.....	العينات غير الاحتمالية وغير العشوائية Non probability sample
53.....	1- العينة العرضية Accidental sample .
54.....	2- العينة المقصودة Purposeful .
54.....	3- عينة الحصص Quota Sample
54.....	تقدير حجم العينة .
56	أخطاء العينات .

أسئلة وتمارين. 58

الفصل الثالث

أساليب تنظيم البيانات وعرضها

Methods of data organization and Presentation

أساليب تبويب البيانات (العرض الجدولي) Tabulation Presentation.... 60

إنشاء جداول تكرارية بسيطة 60

أ- التوزيع التكراري لمتغيرات نوعية (وصفية). 60

ب- التوزيع التكراري لمتغير كمي منفصل 62

ج- التوزيع التكراري لمتغير كمي متصل. 62

الجدول التكراري للفئات 63

قواعد تحديد الفئات 65

افتراضات خاصة بالفئات 66

إنشاء جداول تكرارية بسيطة متجمعة 68

أساليب العرض البياني للمتغيرات :Graphical Presentation :72.....

أ- التمثيل البياني للبيانات النوعية . 72.....

ب- التمثيل البياني لمتغيرات كمية متصلة . 76.....

ج- التمثيل البياني لمتغيرات في مستوى الرتبة . 80.....

أسئلة وتمارين . 82.....

الفصل الرابع

مقاييس النزعة المركزية

Measures of central tendency

مقدمة . 88.....

1- المتوسط الحسابي Mean..... 89.....

- حساب المتوسط الحسابي..... 90.....

أولاً حساب المتوسط الحسابي من الدرجات مباشرة للبيانات غير المبوبة.... 90.....

ثانياً: حساب المتوسط من تكرار الدرجات غير المبوبة في فئات 90.....

ثالثاً: حساب المتوسط من فئات القيم . 92.....

رابعاً: حساب المتوسط في حالة القيم المنفصلة. 98.....

100.....	- المتوسط الموزون <i>Weighted Mean</i>
101	- خواص المتوسط الحسابي
103	2- الوسيط <i>Median</i>
103	1- حساب الوسيط للبيانات غير المبوبة
105	2- حساب الوسيط من فئات القيم :
105	أ- حساب الوسيط من الفئات المغلقة
107	ب: حساب الوسيط من فئات مفتوحة
108	خواص الوسيط
109	3- المنوال <i>Mode</i>
110.....	1- حساب المنوال للبيانات غير المبوبة
111.....	2- حساب المنوال من فئات القيم
111	خواص المنوال
112.....	فوائد مقاييس النزعة المركزية
113.....	مقاييس النزعة المركزية ومستويات القياس
114	حساب مقاييس النزعة المركزية من علاقة بيرسون
115	العلاقة بين المتوسط الحسابي والوسيط والمنوال

118..... Coefficients of skewness مقاييس الالتواء

121..... أسئلة وتمارين

الفصل الخامس

مقاييس التشتت

Measures of Dispersion

126..... Range المدى 1-

126..... أولاً: حساب المدى لبيانات غير مبوبة.

126..... ثانياً: حساب المدى لفئات القيم.

127..... خواص المدى.

127..... Quartile Deviation الانحراف الربيعي 2-

128..... مفهوم الربيعات.

129..... حساب نصف المدى الربيعي لبيانات غير مبوبة.

131..... الانحراف الربيعي النسبي

132..... حساب الانحراف الربيعي (نصف المدى الربيعي) لبيانات مبوبة.

136.....	بعض خواص الانحراف الربيعي.
137.....	العشریات Deciles
137.....	مفهوم العشریات Deciles .
137.....	طرائق حساب العشریات.
138.....	- حساب العشریات.
138.....	- الفوائد العلمية والتطبيقية للعشریات.
140.....	3- الانحراف المعياري <i>Standard deviation</i>
140.....	أولاً: حساب الانحراف المعياري للبيانات غير المبوبة :
145.....	ثانياً: حساب الانحراف المعياري من فئات الدرجات (بيانات مبوبة):
154.....	معامل الاختلاف النسبي Coefficient of Variation :
160.....	الدرجات المعيارية Z Scores
162.....	الدرجة المعيارية T Scores T .
164.....	التوزيع الطبيعي المعياري Standard Normal Distribution .
165.....	خصائص التوزيع الطبيعي.

المساحات تحت التوزيع الطبيعي . 167.....

تمارين وأسئلة . 170.....

الفصل السادس

الارتباط *Correlation*

معنى الارتباط: 173.....

تعريف الارتباط: 174.....

أنواع الارتباط: 174.....

مؤشرات الارتباط : 176.....

شروط حساب معامل الارتباط..... 184.....

طرائق حساب معامل الارتباط. 184.....

معامل ارتباط بيرسون Pearson's coefficient : 184.....

أولاً : حساب معامل الارتباط باستخدام المعادلة الأساسية . 185.....

ثانياً : حساب معامل الارتباط باستخدام الانحرافات المعيارية للمتغيرين..... 186.....

ثالثاً : حساب معامل الارتباط باستخدام انحرافات المتغيرين عن متوسطيهما 186.....

186.....	رابعاً: الطريقة العامة.
189.....	معامل ارتباط سبيرمان (للرتب) Spearman's coefficient
197.....	أسئلة وتمارين محلولة

الفصل السابع

الانحدار Regression.....

205.....	مفاهيم إحصائية متعلقة بتحليل الانحدار
205.....	1. خط الانحدار Line-Regression
	2. الخطأ المعياري للتقدير Standard Error of
205.....	estimation
205.....	3. معامل التحديد Coefficient of determination
206	4. التباين المفسر و التباين غير المفسر
207.....	الافتراضات حول تحليل الانحدار
207.....	استخدامات تحليل الانحدار
209.....	نماذج الانحدار The Regression Models
211.....	تفسير معامل الانحدار
211.....	اعتبارات في تطبيق تحليل الانحدار
213.....	شكل الانتشار Scatter plot

215.....	نموذج الانحدار البسيط .
215.....	معادلة مستقيم الانحدار .
216.....	طريقة المربعات الصغرى Least-squares
220.....	الأخطاء المعيارية للتقدير .
233.....	أسئلة وتمارين .
253.....	أسئلة وتمارين عامة.
295.....	المصطلحات المستخدمة
301.....	المراجع.
307.....	الملحق



المقدمة

نظراً إلى التطورات العلمية المتسارعة في مجالات عديدة، منها المجالات التربوية والنفسية، أصبح للبيانات والمعلومات دوراً أكثر حيوية من ذي قبل إذ تتدفق المعلومات بصورة هائلة من مصادر عديدة، الأمر الذي يقتضي اتخاذ قرارات سليمة قائمة على أسس علمية رشيدة.

وتعد الطرائق الإحصائية من أهم الأساليب من أهم الوسائل العلمية المستخدمة في الميادين المختلفة للبحث العلمي بوجه عام وفي ميادين العلوم الإنسانية بوجه الخصوص. واستخدام الإحصاء في الميادين النفسية التربوية هو وسيلة وليس غاية بحد ذاته، فمن دون الإحصاء لا يستطيع الباحث الإجابة عن فرضيات بحثه وأسئلته، ومن ثم لا يستطيع استنتاج معلومات معينة عن مجتمع معين بناء على النتائج التي استخلصها من عينة مختارة من مجتمع ما لاتخاذ القرارات المناسبة.

والإحصاء علم يهتم باستخدام وتطبيق أكثر الطرائق والأساليب فاعلية في جمع البيانات الخاصة بالظواهر التربوية والاجتماعية والاقتصادية والسياسية وغيرها، وترتيبها وتحليلها في محاولة للتوصل إلى استنتاجات علمية تصف خصائص المجتمع الإحصائي من خلال مجموعة جزئية منه.

ولذلك تقضي الضرورة العلمية تدريس مبادئ الإحصاء بفرعيه الوصفي والتحليلي في جميع التخصصات المختلفة، بما فيها التخصصات النفسية والتربوية وتصلح منهجاً إحصائياً يمكن الطلبة مستقبلاً من الدخول إلى مضمار البحث العلمي بأدوات علمية إحصائية وصفية وتحليلية .

ويقتصر الكتاب الحالي على الإحصاء الوصفي متضمناً فصول عديدة: اشتمل الفصل الأول على مقدمة في الإحصاء، كما يتناول مفهوم الإحصاء وأقسامه وأهميته. إضافة إلى أنواع المتغيرات وتصنيفاته، و مستويات القياس وتدرجاته.

أما الفصل الثاني فيتناول مفهوم البحث العلمي وخطوات البحث العلمي وموقع الإحصاء في البحث العلمي. ومفهوم العينة وأنواعها والأخطاء التي قد تقع أثناء عملية السحب. أما الفصل الثالث فيتناول أساليب تنظيم البيانات وعرضها.

وأُفردَ الفصل الرابع لمقاييس النزعة المركزية (المتوسط الحسابي والوسيط والمنوال) وكيفية حسابها بطرائق عديدة، والعلاقة بينها، وموقع كل منها من مستويات القياس. والفصل الخامس تناول مقاييس التشتت وكيفية حسابها، وموقع كل مقياس من هذه المقاييس من مستويات القياس، إضافة إلى مفهوم التوزيع الطبيعي وكيفية حساب المساحات تحت المنحنى الطبيعي. أما الفصل السادس فيتناول مفهوم الارتباط، وأنواعه وكيفية حساب معاملات الارتباط إضافة للأهمية الإحصائية لمعامل الارتباط. أما الفصل السابع فيتناول مفهوم الانحدار إضافة إلى تعريف ببعض المفاهيم المتعلقة بالانحدار: خط الانحدار، الخطأ المعياري للتقدير، معامل التحديد، التباين المفسر والتباين غير المفسر، ثم الافتراضات حول تحليل الانحدار، ثم يستعرض الكتاب استخدامات تحليل الانحدار ونماذج الانحدار وتفسير الانحدار واعتبارات في تطبيق تحليل الانحدار، وشكل الانتشار ومعادلة مستقيم الانحدار، وطريقة المربعات الصغرى المستخدمة في الوصول إلى معادلة مستقيم الانحدار.

الفصل الأول

مقدمة في الإحصاء

Introduction to statistics



معنى الإحصاء و تعريفه:

تعد كلمة الإحصاء من الكلمات الشائعة الاستخدام، سواء في الحياة اليومية، أم في ميادين العلوم المختلفة، حيث لا يمكن فهم البيانات الإحصائية إلا من خلال الإطار النظري، ومجموعة الطرائق التي يتم تطبيقها ضمن هذا الإطار. فالحديث عن الإحصاء في التربية يختلف عنه في الاقتصاد عنه في البيولوجية أوفي الرياضيات بالرغم من وجود قاسم مشترك بينها . وقد اقتصر استخدام الإحصاء في البداية على الأعمال الخاصة بشؤون الدولة، كما يدل على ذلك الأصل اللغوي في اسم هذا العلم، وهو بالإنجليزية "Statistics" وهو مشتق من كلمة "State" أي الدولة ومعناه " مجموعة الحقائق الخاصة بشؤون الدولة " (أبو صالح، 2000، ص120) .

وبهذا يشار للإحصاء إلى أنه " مجموعة من الطرائق العلمية القياسية التي يمكن توظيفها لجمع المعطيات (البيانات والمعلومات) عن الظواهر، وتبويبها وتلخيصها وتقييمها، والخروج منها باستنتاجات حول مجموع وحدات المجتمع من خلال اعتماد جزء صغير من هذا المجتمع (البلداوي، 2004، ص17) .

أهمية علم الإحصاء واستخداماته :

تطور استخدام علم الإحصاء في العقود الأخيرة تطوراً ملحوظاً نتيجة التغيرات السريعة والكبيرة التي طرأت على حياة الإنسان ونشاطاته في مختلف الميادين العلمية والاقتصادية والاجتماعية والإنسانية. وقد كان لإسهامات جالتون "Galton" في نظريته عن أخطاء القياس إسهاماً هاماً في تطور الإحصاء، كما أن لنظريات وابتكارات فيشر "Fisher" دوراً مهماً في هذا التطور، وقد ساعد تطور الحاسوب والبرامج الإحصائية المتخصصة في معالجة العمليات الإحصائية المعقدة (أبو النيل، 1987، ص17) .

----- الفصل الأول : مقدمة في الإحصاء

وأسهل استخدام علم الإحصاء في ميادين العلوم المختلفة في تطور هذه العلوم تطوراً كبيراً سواء في مستوى التخطيط أم في قياس النوعية ومعالجة المشكلات. وقد أصبح الإحصاء الركيزة العلمية الهامة من ركائز البحث العلمي، و يعتمد عليه في دراسة البيانات التي يتم جمعها عن ظاهرة معينة وتبويبها وتنظيمها وعرضها للتعرف إليها بوصفها أو بتحليلها، واستقراء النتائج منها (أبو زينة وآخرون، 2006، ص 17). وبهذا يساهم علم الإحصاء في مساعدتنا على فهم أعمق للسلوك الإنساني، واستنباط البراهين التجريبية لقبول الفرضيات أو رفضها، كما أنه يساهم في عملية التخطيط واتخاذ القرارات. ومن الأهمية بمكان أن تكون التحليلات الإحصائية ذات نتائج صادقة إذ يتوقف اتخاذ القرار السليم على سلامة النتائج الإحصائية التي جرى التوصل إليها وتزداد أهمية اختيار الإحصائيات المناسبة كلما زادت أهمية القرارات المترتبة على النتائج إذ أن كثيراً من الدراسات والبحوث ليست ذات قيمة نظرية أو تطبيقية نتيجة لضعف في خطوات البحث العلمي، وترجع معظم مصادر الضعف إلى نوعية التحليلات الإحصائية المستخدمة (عودة والخليلي، 1988، ص 12).

أقسام الإحصاء :

ينقسم علم الإحصاء إلى قسمين أساسيين هما :

1- الإحصاء الوصفي Descriptive statistics:

هو ذلك الفرع من علم الإحصاء الذي يهتم بجمع البيانات وتنظيمها وعرضها عن طريق الجداول أو الرسوم البيانية وغيرها، بهدف التعرف إلى خصائصها وصفاتها ، كما أنه الخطوة الأولى في عملية تحليل البيانات باستخدام طرائق قياسية تحدد قيمة المتوسط أو المدى حول المتوسط و الوسيط،و المنوال، و الانحراف المعياري وغيره بغية الوصول للنتائج دون محاولة

تعميمها من أفراد العينة التي درست خصائصها على عينة أخرى (Streiner, 2003,p. 11)
(Norman&).

2- الإحصاء الاستدلالي Inferential statistics:

وهو الفرع من الإحصاء الذي يهتم بتحليل البيانات المتوفرة حول عينة مسحوبة من مجتمع إحصائي محدد بغرض التنبؤ والاستقراء، واستنتاج البيانات غير المتوفرة لدينا عن المجتمع الإحصائي المدروس بدرجة محددة من الاحتمال، بمعنى أننا نستدل على بيانات المجتمع من بيانات العينة .

وعليه فإن الإحصاء الاستدلالي هو الذي يتعامل مع التنبؤ والتقدير والتعميم ، وبعتماد التقدير الإحصائي و اختبار الفرضيات المتعلقة بالفروق بين المتوسطات أو النسب لعينة واحدة أو عينتين أو أكثر .

ويقدم الإحصاء الاستدلالي استناداً إلى نظرية الاحتمالات "Probability theory" ونظرية القرارات "Decision Theory" العديد من الطرائق والأساليب الكمية التي تمكن الباحث من التوصل إلى الأدلة التي يستند إليها في عملية التعميم، والتمييز بين العوامل والتغيرات التي ترجع إلى الصدفة، والتأثيرات التي ترجع إلى عوامل جوهرية، لذلك فإن هذه الأساليب الكمية لا تؤدي إلى نتائج حتمية مؤكدة، وإنما تكون نتائجها احتمالية.(علام، 2005،
(15)

المتغيرات الإحصائية Variables:

تسمى المعلومات الإحصائية المتعلقة بالظواهر الاجتماعية والاقتصادية والثقافية والصحية والتربوية والمالية والسكانية بالبيانات، وتختلف هذه البيانات من حيث نوعها وطبيعتها باختلاف الظاهرة المدروسة . وفي المجال التربوي والنفسي يتركز الاهتمام على خصائص

----- الفصل الأول : مقدمة في الإحصاء

الأفراد وسماتهم الجسمية والشخصية والعقلية، وهذه الخصائص تأخذ قيماً مختلفة بالنسبة للأفراد المختلفين، وهذه الخصائص تدعى بالمتغير "Variable"، وبالمقابل إذا كانت هذه الخصائص لها القيمة نفسها عند جميع الأفراد، فإن هذه الخصائص تدعى بالثابت "Constant".

يمكن توضيح الفرق بين الثابت والمتغير من خلال المثال، إذا قمنا بقياس ذكاء مجموعة من طلبة الصف الثالث الثانوي بأحد اختبارات الذكاء فإنه من الطبيعي أن يأخذ الطلبة درجات متفاوتة في اختبار الذكاء، هنا تسمى درجة الذكاء المتغير، أما الثابت بين جميع هؤلاء الطلبة هو أنهم من طلبة الصف الثالث الثانوي، أي أن المستوى الصفّي يعد ثابتاً .

وبهذا يمكن أن نعرف المتغير أنه : " الخاصية التي تأخذ قيماً مختلفة للأفراد المختلفين في المجموعة قيد الدراسة " أما الثابت : فهو الخاصية التي لها القيمة نفسها عند جميع أفراد المجموعة قيد الدراسة (المنزل، 2000، ص18) .

والتعريفات الإجرائية للمتغيرات تصف أو تعرف المتغير في حدود الإجراءات أو التقنيات التي تجعله يحدث أو يقاس به، في حين يصف الباحثون المتغيرات الكمية في أبحاثهم بالمتغيرات الديمغرافية من مثل الجنس، النوع، العمر وهي متغيرات تقاس عادة ببساطة عن طريق سؤال المفحوصين باختيار الفئة التي ينتمون إليها من قائمة تعطى لهم . (Leech & Barrett, 2005, p1)

تصنيف المتغيرات:

تصنف المتغيرات الإحصائية في العلوم النفسية والتربوية تصنيفات عديدة من هذه التصنيفات:

- متغيرات إحصائية كمية Quantitative Variables:
- متغيرات إحصائية نوعية Qualitative Variables :

----- الفصل الأول : مقدمة في الإحصاء

المتغيرات الكمية: يستند هذا التصنيف على مدلول القيمة الممثلة للخاصية المقاسة، فإذا كانت هذه القيمة تشير إلى مقدار ما في الفرد من خاصية مقارناً بأفراد مجموعته، فإن هذه القيمة تحمل معنى كمياً و المتغير متغير كمي، وإذا كانت القيمة لا تعبر عن مقدار الخاصية عند فرد معين، وإنما تعبر فقط عما إذا كان يمتلك تلك الخاصية أم لا، أو أنها تشير إلى فئة أو مجموعة مثل الجنس، المرحلة الدراسية، المستوى التحصيلي، فإن هذه المتغيرات هي متغيرات نوعية (عودة، 1988، ص12-13) فالمتغيرات الكمية هي المتغيرات التي تتعامل مع قيم رقمية من مثل التحصيل، الذكاء .

أما المتغيرات النوعية، فتسمى أيضاً متغيرات تصنيفية "Categorized Variable" فهي تصنف الأفراد أو الأشياء ضمن أصناف أو فئات معينة دون أن تتضمن هذه المتغيرات قيماً رقمية، وتتميز المتغيرات النوعية أنه لا يوجد تداخل بين تلك الفئات أو التصنيفات وهذه التصنيفات قد تكون ترتيبية أو غير ترتيبية .

وتصنف المتغيرات الكمية إلى :

أ- متغيرات كمية متصلة "Continuous Variables":

وهي المتغيرات التي تأخذ أي قيمة بين ضمن مجال و تأخذ قيماً لا نهائية من القيم المحتملة، أو هي المتغيرات التي يمكن أن تأخذ قيمها أرقاماً صحيحة أو كسرية، مثل: الطول والوزن والعمر والزمن .

ب- متغيرات كمية منفصلة Discrete Variables :

هي المتغيرات التي تأخذ قيماً محددة، أي المتغير الذي تأخذ قيمه أرقاماً عددية صحيحة قابلة للعد مثل : عدد الطلاب في المدرسة أو الكلية، عدد الغرف في الفندق، عدد الأبنية في أحد الأحياء .

وهناك تصنيف آخر يقسم المتغيرات في مجال البحوث إلى ما يلي :

أ- المتغيرات المستقلة "Independent Variable":

ب- المتغيرات التابعة "Dependent Variable":

يقوم هذا التصنيف على أساس وجود علاقة بين متغيرين وهذه العلاقة تمكن الإحصائي من التنبؤ بقيمة أحد المتغيرين (المتغير التابع أو المتنبأ به) من معرفته لقيمة المتغير الآخر وهو المتغير المستقل أو المتنبأ .

والمتغير المستقل : هو المتغير الذي يستطيع الباحث أن يتحكم به ويغيره وفقاً لطبيعة البحث. فعلى سبيل المثال إذا أراد باحث أن يدرس أثر الضوضاء في زيادة شدة القلق لدى مجموعة من الأفراد، فإنه يعتمد إلى زيادة شدة القلق لدى مجموعة من الأفراد عن طريق زيادة المستويات المختلفة للضوضاء وملاحظة أثرها في المتغير التابع وهو القلق . والمتغير المستقل في بعض الدراسات هو متغير تصنيفي، فعلى سبيل المثال إذا كان الباحث يبحث في تأثير طرائق التدريس المختلفة في التعليم: (طريقة المحاضرة، طريقة النقاش، طريقة التعليم المبرمج) على التحصيل عند عينة من طلبة الصف السابع، فإن طريقة التعليم تعد متغيراً مستقلاً واحداً، والطرائق الثلاث تمثل مستويات المتغير المستقل .

أما المتغير التابع هو : الذي يتأثر بالمتغير المستقل، وعلى الباحث ملاحظة تأثيرات التغير في قيم المتغير المستقل على قيم المتغير التابع، ولا يكون ذلك إلا في حال وجود علاقة ارتباطية بين المتغيرين المستقل والتابع .

وفي بعض التجارب هناك أكثر من متغير مستقل واحد أو أكثر من متغير تابع واحد، والمتغير المستقل في بعض التجارب يمكن أن يكون على شكل مجموعات "Categorical" أو منفصل "Discrete" ويأخذ شكل وجود أو عدم وجود معالجة معينة، أو على شكل الفروق بين معالجتين . وفي تجارب أخرى يمكن أن يكون المتغير المستقل متصلاً وملاحظة الباحث له يمكن أن يعبر عنه بقيمة رقمية مشيرة إلى الدرجة. وفي الدراسات الارتباطية عندما يكون

----- الفصل الأول : مقدمة في الإحصاء

المتغيران مستمرين، فإن عملية تحديد أي منهما المتغير المستقل وأي منهما المتغير التابع مسألة اعتباطية، وفي بعض الحالات لا تتم الإشارة إلى أي منهما المتغير المستقل وأيها المتغير التابع .

والمتغيرات المستقلة تدعى بالعوامل وتبايناتها تدعى بالمستويات ففي المثال السابق طريقة التعليم عبارة عن عامل والتصنيفات الواقعة ضمن طرائق التدريس (طريقة المحاضرة، طريقة النقاش، طريقة التعليم المبرمج) تدعى بالمستويات. والجدول التالي يوضح أنواع المتغيرات

الجدول (1-1)

أنواع المتغيرات

نوع المتغير	الوصف
متغير متصل	متغير نقيسه باستخدام وسائل القياس من المستوى القوي، من أمثلته: القلق و الذكاء والتحصيل، ويتصف بأنه لا يوجد فجوات بين قيم المتغير
متغير منفصل	قيمه غير متصلة، ولذلك لا يمكن استخدام الكسور في هذا المتغير بل إن جميع قيمه صحيحة، مثل عدد أفراد الأسرة.
متغير نوعي (تصنيفي)	متغير من المستوى الاسمي، ولذلك تحل أقسامه محل الأسماء ووظيفة هذا المتغير الأساسية هي تصنيف المفهوم في فئات، مثل متغير الجنس ومتغير المنطقة والكلية. والأرقام في هذا المتغير لا تعبر عن كميات من الخصائص أو السمات.
متغير مستقل	في البحوث التجريبية هو المتغير التجريبي الذي يعالجه الباحث ليرى أثره على المتغير التابع. والمتغير المستقل هو متغير تصنيفي (منفصل) غالباً
متغير تابع	هو المتغير الذي يظهر أثر المتغير المستقل فيه، والمتغير التابع متغير متصل غالباً

(نقلاً عن أبو علام، 2004، ص 57)

مستويات القياس:

تختلف أشكال القياس باختلاف طبيعة وخصائص السمات المقاسة، كما تختلف هذه المقاييس في مستوى دقتها. ولابد للباحث في العلوم السلوكية من أن يتعرف على مستويات القياس، كي يستطيع استخدام الطرائق الإحصائية المناسبة لتحليل البيانات الخاصة بدراساته وبحوثه استخداماً صحيحاً الأمر الذي يقود إلى نتائج على درجة عالية من الوثوقية.

وتتدرج مستويات القياس وفق تصنيف ستيفنز "Stevens" في أربعة مستويات وهي من الأدنى إلى الأعلى على النحو التالي :

1- المستوى الاسمي أو التصنيفي Nominal Scales:

وهو أدنى مستويات القياس و يستخدم مقياساً لتحديد هوية الأفراد أو العناصر، فهو يصنف عناصر الظاهرة التي تختلف في النوعية لا في الكمية، ويجري في هذا المستوى تصنيف أفراد المجموعة إلى فئات استناداً إلى متغير أو سمة معينة تعتمد أساساً للتصنيف، مثل تصنيف الطلاب إلى ذكور وإناث، أو طلاب علمي وطلاب أدبي وتصنيف الناس إلى مدخنين وغير مدخنين أو مرضى وأسوياء، ومن الممكن أن تكون التصنيفات حسب الفئات مثل: (الجنس، التخصص، الجنسية، الوضع العائلي) ويمكن الاستعاضة عن الأسماء بأرقام أو رموز، فقد نرسم للذكور برقم 1 وللإناث بالرقم 2 . والأرقام هنا لا تحمل أي معنى كمياً، أي أنها لا تشير إلى كمية معينة للخاصية أو المتغير، ولا تتضمن معنى للأفضلية (الأكبر والأصغر)، فمثلاً إذا أعطي الذكور رقم (1) والإناث رقم (2) فهذا لا يعني أن الذكور أقل من الإناث أو الإناث أفضل من الذكور .

والتصنيف عملية جوهرية لأي علم من العلوم ولعلنا هنا نشير إلى أن جميع المستويات الأخرى من القياس مهما بلغت درجة دقتها، تتضمن عملية التصنيف على نحو أو آخر . ولعل

----- الفصل الأول : مقدمة في الإحصاء

هذا ما يبرر لنا اعتبار هذه العملية المستوى الأدنى للقياس بمعناه الواسع، ويتطلب ذلك توافر شرطين في الفئات : أولهما الشمول حيث تشتمل الفئة جميع الحالات الفردية المحتملة، وثانيهما عدم التداخل حيث لا يجوز بحالة ما أن تضمناها فئتين من النظام نفسه التصنيفي (أبو حطب، آمال صادق، 1991، ص36) "ويرى كومبس أن هذا النوع من المقاييس تحكمه علاقة التساوي، وهذا يعني أن أي زوج من الأشياء يجب أن ينتهي بوضوح إلى الفئة نفسها أو لا ينتهي إليها (أبو حطب، صادق، 1991، ص36)

ولا يمكن في هذا المستوى استخدام العمليات الحسابية الأربع، إلا الجمع ضمن الفئة الواحدة، من مثل جمع عدد الذكور أو عدد الإناث، وحساب النسب في كل فئة .
أما العمليات الإحصائية التي يمكن القيام بها في هذا المستوى فهي حساب المنوال الذي يعني القيمة الأكثر تكراراً .

2- المستوى الترتيبي ordinal scales:

يعد القياس في هذا المستوى أعلى من المستوى السابق، فهو بالإضافة إلى تصنيف الأفراد والأشياء في مجموعات متباعدة، فهو يرتب الأفراد أو الأشياء ترتيباً تصاعدياً أو تنازلياً وفقاً لصفة معينة، "ويرى كومبس أنه في الفئة الواحدة من فئات المقاييس الاسمية قد نجد ما هو أكثر من مجرد أن الوحدات التي تتألف منها الفئة متساوية فيما بينها ومختلفة عن الوحدات التي تتألف منها فئة أخرى، فقد وجد بعض العلاقات بين بعض وحدات الفئات المختلفة، ومن هذه العلاقات أن وحدات إحدى الفئات قد تكون أكبر (أو أصغر) من وحدات فئة أخرى " (أبو حطب، 1990، ص37).

وبهذا يكون للعدد مدلول كمي بصورة تمكن الإحصائي من ترتيب الأفراد تصاعدياً أو تنازلياً حسب مستوى الخاصية المقاسة، فيمكن على سبيل المثال ترتيب الطلاب ترتيباً تنازلياً

----- الفصل الأول : مقدمة في الإحصاء

وفقاً لمستوى تحصيلهم إلى ممتاز جيد جداً وجيد وضعيف أو ترتيب الموظفين وفقاً لسنوات الخدمة وغير ذلك.

ولابد من الإشارة هنا أن الفروق بين الرتب ليست فروقاً منتظمة، فالفرق بين الرتبة الأولى والثانية لا يساوي بالضرورة الفرق بين الرتبة الثانية والثالثة، فعلى سبيل المثال إذا كانت درجات الطلاب مرتبة تنازلياً كما يلي :

65	70	80	90	95
الخامس	الرابع	الثالث	الثاني	الأول

يلاحظ من الترتيب السابق أن الفرق بين الأول والثاني هو خمس درجات وبين الثاني والثالث عشر درجات كذلك الفرق بين الثالث والرابع عشر درجات، ثم يعود الفرق بين الرابع والخامس إلى خمس درجات .

ثم أن الترتيب نسبي وليس مطلقاً أو ثابتاً، بمعنى أنه من الممكن أن ينضم فرد إلى الفئة أو يخرج فرد من الفئة فيختلف الترتيب فيها، مما يؤدي إلى تعديل الموقع النسبي للفرد، وتغيير رتبته صعوداً أو هبوطاً . ولا يمكن في هذا المستوى القيام بأي عملية من العمليات الحسابية الأربع، فالرتب غير قابلة للجمع والطرح والقسمة والضرب، أما العمليات الإحصائية التي يمكن القيام بها في هذا المستوى فهي تحويل الرتب إلى رتب مئوية و معامل الارتباط الرتبي، والوسيط.

3- المستوى الفئوي Interval Scales :

يعد هذا المستوى أعلى وأدق من المستويين السابقين وهو يمتلك خاصية الفواصل أو المسافات المتساوية التي تفصل بين كل درجتين متجاورتين، كما أن الأرقام فيه تحمل معنى كمياً وتكون قابلة لعمليات الجمع والطرح، فإذا كانت درجات أربعة طلاب هي

----- الفصل الأول : مقدمة في الإحصاء

50	55	60	70	75
----	----	----	----	----

فإنه يمكن أن نقول إن الفرق بين (50) و (55) هو 5 ويساوي الفرق بين (55) و(60)، كما أنه يجوز أن نقول بأن الفرق بين (60) و (70) هو ضعف الفرق بين (50) و(55) .

و يجدر بالذكر أن الصفر في هذا المستوى هو صفر افتراضي وليس صفرًا حقيقيًا، إذ أنه لا يعني انعدام السمة، ذلك أن درجة الصفر التي يحصل عليها طالب في مادة من المواد لا تعني أنه لا يعرف شيئاً أو أن تحصيله منعدم، كذلك درجة الحرارة الصفر لا تعني انعدام الحرارة، ولكن تعني أن حرارة هذا الجسم أعلى من أي درجة سالبة، وفي هذا المستوى ليس للنسبة معنى، إذ لا يمكننا القول أن شخصاً ما متفوق عقلياً وحصل على حاصل ذكاء (IQ) قدره (150) درجة يعادل حاصل ذكاء ثلاثة أشخاص متخلفين عقلياً حاصل ذكاء كل واحد منهم (50) درجة . والأمر نفسه ينطبق على درجات التحصيل .

ويمكن في هذا المستوى إجراء العديد من العمليات الإحصائية من مثل حساب المتوسط الحسابي والانحراف المعياري، ومعامل ارتباط بيرسون، إضافة إلى العديد من اختبارات الدلالة الإحصائية استناداً إلى إمكان تحويل الدرجات الخام إلى درجات موزونة أو معيارية والتي يمكن معالجتها حسابياً عن طريق الجمع والطرح وحدهما .

4- مستوى النسبة Ratio Scales:

وهذا المستوى هو أعلى المستويات في القياس ويتميز بوجود الصفر الحقيقي الذي يعني انعدام السمة المقاسة، كما أنه يتصف بجميع صفات المستويات السابقة، بمعنى أن وحداته متساوية ومنتظمة، كما أن النسبة في هذا المستوى لها معنى، إذ إن الوزن 80 كغ هو ضعف الوزن 40 كغ، والطول 30 سم هو نصف الطول 60 سم . ويمكن في هذا المستوى القيام

----- الفصل الأول : مقدمة في الإحصاء

بجميع العمليات الحسابية من جمع وطرح وضرب وقسمة . ومن الأمثلة على مقاييس النسبة الطول، الوزن، الحجم

وهذه المقاييس تتميز بها العلوم الدقيقة، إذ يندر استخدامها في التربية وعلم النفس لافتقار المقاييس التربوية والنفسية لخاصية الصفر الحقيقي، كذلك تشكو المقاييس التربوية والنفسية من صعوبة الحصول على وحدات أو فئات متساوية. إذ قد يكون الاختبار التحصيلي يحتوي على أسئلة متفاوتة في درجة صعوبتها مما يجعل وحدات القياس غير متساوية. وفي هذه المقاييس يمكن إجراء جميع العمليات الإحصائية .

وجدير بالذكر أن معظم المتغيرات التربوية والنفسية تقع بين المستوى الفئوي والمستوى الترتيبي، وقد يكون المستوى الترتيبي أكثر فائدة من المستوى الفئوي في حالات شائعة، مثال قولنا أن درجة الطالب هو أكبر من 80% من الطلاب الذين في مثل عمره وهذا النوع أفضل من القول إن الطالب حصل على 95 درجة.

والجدير بالذكر بأنه يتم عادة تصنيف المتغيرات وفقاً لهذه المقاييس فهي بمجموعها ذات صفة تراكمية، بمعنى أن مستويات القياس متضمنة فيما بينها.

ومن المهم تذكر أن مستويات تدرج ستيفنز للقياس هو تصنيف مفيد لكنه ليس الطريقة الوحيدة لتصنيف أو وصف المستويات .فعلى سبيل المثال تحدث تورغرسن (Torgerson,1958,p.19) عن مستويين للتدرج الرتبي، وقدم كومبس (Coombs,1950) نوع من التدرج يقع بين الرتبي والفاصلي. (كروكر والجينا،2009،ص76)

والجدول يبين التالي وصفاً لمستويات القياس الأربعة كما صنفها ستيفنز :

الجدول (1-2)

مستويات القياس

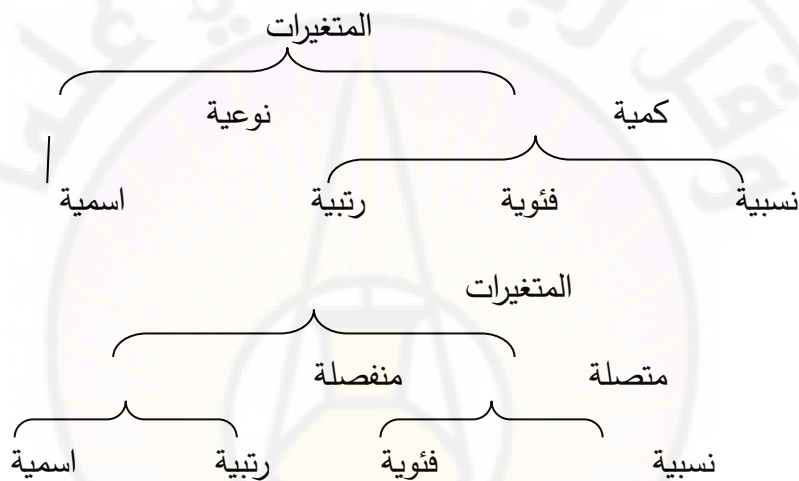
مستوى القياس	خصائص المقياس	الأمثلة	العمليات الحسابية والإحصائية
النسبي	يوجد صفر حقيقي ويعني غياب الصفة، وتمثل الأرقام كميات متساوية عن موقع الصفر المطلق	العمر، الدخل، السرعة، الزمن، حجم الصف، عدد أفراد الأسرة	يمكن إجراء جميع العمليات الحسابية وجميع مقاييس النزعة المركزية والتشتت، وللنسبة معنى
الفئوي	الفروق بين القيم تساوي الفروق بين قيم المتغير أو قيم السمة التي تخضع للقياس، الصفر افتراضي غير مطلق، ولا يعني غياب السمة.	درجة الحرارة المئوية، التقويم السنوي، الدرجات على اختبار تحصيلي أو اختبار ذكاء	يمكن إجراء عمليات حسابية كالجمع والطرح وجميع مقاييس النزعة المركزية والتشتت وليس للنسبة معنى.
الرتبي	تمثل الأرقام رتباً للأفراد حسب درجة امتلاكهم للسمة موضع البحث. تمثل الأرقام صفة كمية تبين مفهوم أكبر وأصغر	المستوى الاجتماعي والاقتصادي، ترتيب الطلبة داخل الصف	لا يمكن إجراء عمليات حسابية، يمكن التعامل مع الوسيط) التعبير عن القيم بالمئينيات فقط)
الاسمي	تستخدم الأرقام لأغراض التصنيف، ولا تمثل أي صفة كمية أو ترتيب معين، ولا يخضع استخدام الأرقام في التصنيف إلى قواعد أو قوانين	الجنس، الجنسية، نمط الشخصية، اللغة، الديانة، الرقم الجامعي.	لا يمكن إجراء أي عمليات حسابية، ويمكن التعامل مع المنوال فقط.

(نقلا عن النبهان، 2004، ص 32)

----- الفصل الأول : مقدمة في الإحصاء

لاشك أن المتغير الواحد يمكن تصنيفه في أكثر من فئة، فالجنس هو متغير من المستوى الاسمي، وهو متغير نوعي ومتغير منفصل من ناحية أخرى، وهو قد يكون متغيراً مستقلاً في تصميم تجريبي معين .

والجدير بالذكر أن معرفة نوع المقياس المستخدم مهم جداً للإحصائي الذي سيستخدم الأساليب الإحصائية المختلفة، كي يتسنى له اختيار الأسلوب الإحصائي المناسب لمتغيراته.



(فهيم، 2005، ص31)

أسئلة وتمارين خاصة بالفصل الأول:

1- ميّز بين المتغيرات الكمية والمتغيرات النوعية الآتية:

أ- أطوال مجموعة من طلاب الصف الأول الابتدائي.

ب- الرتبة العلمية لأستاذة الجامعة.

ج- التخصص العلمي لمجموعة من طلبة الجامعة.

د- عدد الأسئلة التي أجاب عنها الطلاب بصورة صحيحة.

هـ - انتماء الفرد إلى نقابة معينة.

و - عدد الفائزين في مسابقة معينة.

ز - الدرجات التي يحصل عليها الطالب في مادة الرياضيات.

2- صنف المتغيرات المذكورة في المثال السابق وفق مستويات القياس (الاسمي، الترتيبي،

الفئوي، النسبي)



الفصل الثاني

البحث العلمي والإحصاء

Scientific Research and Statistics



خطوات البحث العلمي :

يستخدم الإحصاء جنباً إلى جنب مع خطوات البحث العلمي والتي تتلخص فيما يلي:

1- تحديد مشكلة البحث :

هناك العديد من المشكلات الملحة التي تنتظر الدراسة في كل مجال من مجالات التربية والعلوم الإنسانية. ويمكن لأي فرد يعمل في هذا المجال أن يضع قائمة بالكثير من الموضوعات التي تحتاج إلى تعميق معرفته بها.

وهناك مجموعة من الأسئلة التي ينبغي على الباحث أن يجيب عنها عند اختيار مشكلة البحث. ولعل أول هذه الأسئلة التي ينبغي على الباحث أن يجيب عنها عند اختيار مشكلة البحث. ولعل أول هذه الأسئلة يتعلق بمدى كون المشكلة التي جرى اختيارها قابلة للبحث.

وتتنوع المصادر التي يأخذ منها الباحث مشكلته. فقد يتطوع الباحث للبحث في مشكلة جرى تحديدها من قبل أستاذه، أو من قبل المؤسسة التي يعمل فيها.

ويمثل الأدب النظري المنشور مصدراً آخر لمشكلة البحث. إذ تتفتح للباحث إمكانيات جديدة وأفاق واسعة بخصوص موضوعات البحث، عندما يصبح أكثر ألفة بالدراسات والبحوث المنشورة في مجال تخصصه، سواء في الدوريات المتخصصة أو في المراجع العامة. ومن المهم التأكيد على أن المعرفة الكافية بالأدب السابق متطلب ضروري يسبق قرار مواصلة العمل في بحث مشكلة معينة.

أما البحث فيعرف أنه : " عملية منظمة للتوصل إلى حلول لمشكلات أو إجابات عن أسئلة، تستخدم فيها أساليب الاستقصاء والملاحظة بصورة مقبولة ومعروفة بين الباحثين في مجال معين، ويمكن أن تؤدي إلى معرفة جديدة". (الكيلاني، الشريفين، 2004، ص35)

ويتناول هذا التعريف الجوانب التالية:

الأول أن هناك مشكلة تحتاج إلى حل. والثاني أن هناك أساليب وإجراءات يفترض في الباحث إتباعها حتى يمكن الوثوق بنتائج البحث وقبولها والاعتراف بها لدى الباحثين الآخرين في المجال العلمي الذي يريته الباحث. أما الجانب الثالث: هو أن البحث يُولد معرفة جديدة إذ أن هدف الباحث هو الوصول إلى معرفة تفسر من خلالها ظاهرة ما، أو يتعرف على علاقة أو استخلاص نظرية أو يصف ظاهرة ما وصفاً دقيقاً .

هذا ويعد الإحساس بالمشكلة واكتشافها الخطوة الأولى من خطوات البحث العلمي، ولا بد للباحث من أن يبذل جهداً في تحديد المشكلة تحديداً دقيقاً وصياغتها صياغة واضحة، ويمكن أن تصاغ المشكلة بصورة إخبارية أو صيغة استفهامية، وقد تصاغ المشكلة بعبارة واحدة أو عدة عبارات وفقاً لبساطة أو تعقد طبيعة المشكلة .

ولكي تصاغ المشكلة صياغة علمية دقيقة يجب أن تتوافر المعايير التالية:

- 1- يجب أن تكون صياغة المشكلة في عبارة محددة أو سؤال واضح.
- 2- يجب أن توضح المشكلة علاقة بين متغيرين أو أكثر مع تحديد المجتمع الذي تشمله الدراسة.
- 3- يجب أن تكون المتغيرات التي تحدها المشكلة متفقة مع المتغيرات التي تعالجها أدوات الدراسة، كما يجب أن يكون المجتمع كما حددته المشكلة متفقاً مع عينة البحث.
- 4- يجب أن تكون المشكلة قابلة للبحث . (أبو علام، 2004، ص 67)

2- صياغة الفرضيات :

من المتفق عليه أنه عندما يواجه المرء بمشكلات أثناء ممارساته الحياتية، فإنه يبدأ بجمع المعلومات، ويسعى لإيجاد الحلول ووضع عدد من التخمينات أو التفسيرات للمشكلة موضع الاهتمام. وتتركز جهود الفرد أثناء عملية جمع المعلومات نحو التأكد من مصداقية هذه التخمينات أو التفسيرات. وعادة تدور هذه التخمينات والتفسيرات حول مسببات المشكلة والعلاقات المرتبطة بها، ويطلق عليها لفظ الفرضيات" فالفرضية ينظر إليها بأنها عبارة عن جملة تصريحية توضح العلاقة بين متغيرين أو أكثر. وحتى تكون الفرضية ملائمة يجب أن تتضمن متغيرات قابلة للملاحظة والقياس، وأن تحدد بدقة كيف ترتبط المتغيرات الواردة فيها بعضها مع بعض.

إن الصياغة الجيدة للفرضيات تعد خطوة أولى على طريق الدراسة الفعالة للمشكلة. فالفرضيات هي الأدوات العاملة للنظرية التي يمكن اختبارها والتأكد من معقوليتها، وهي أدوات قوية لتقدم المعرفة لأنها خاضعة لعملية التجريب. كما أن الفرضية الجيدة تلعب دوراً في توجيه البحث انطلاقاً من العلاقات التي تتضمنها الفرضية إذ أنها تخبر الباحث ما الذي عليه عمله، وهي جسر بين النظرية والمنهجية التجريبية.

و عندما نتحدث عن الفرضيات فإننا في الحقيقة نتحدث عن نمطين من الفرضيات ؛ الفرضية العلمية هي حل مقترح لمشكلة تصاغ بصورة استنتاجية كتخمين ذكي يستند على معلومات علمية سابقة (عودة والخليلي، 1988، ص270).

أما الفرضية الإحصائية فهي عبارة محددة حول سمات المجتمع أو صفاته، يخضعها الباحث للاختبار، فإما أن يقبلها أو يرفضها .

----- الفصل الثاني: البحث العلمي والإحصائي

والفرضية الإحصائية التي تخضع للاختبار الإحصائي تسمى بالفرضية الصفرية، وتصاغ عادة بصيغة النفي، من مثل الفرضية التالية : لا توجد علاقة ذات دلالة إحصائية ($\alpha = 0.05$) بين الذكاء الانفعالي والتحصيل الدراسي.

منهج البحث : يشتمل منهج البحث على العناصر التالية :

- 1- عينة البحث : وهي جزء من المجتمع الإحصائي وتسحب وفق أسس وضوابط علمية معينة، كي تكون ممثلة تماماً في خصائصها لخصائص المجتمع الذي سحبت منه ويجري الاعتماد على عينة البحث نظراً لأنه في معظم الحالات الواقعية لا يمكن دراسة المجتمع بأكمله .
- 2- إجراءات جمع المعلومات وفيها تتحدد خطة البحث، وكيفية تنفيذه، نوع البيانات اللازمة، ولابد هنا من اختيار المنهج المناسب لنوع البيانات وهدف البحث .
- وتعد عملية جمع البيانات هي العملية الأولى التي ستنبنى عليها كل العمليات التالية، وبالتالي كان لابد من أن تكون هذه البيانات صحيحة ودقيقة ومفصلة وشاملة لكل جوانب الموضوع حتى يجري التوصل إلى القرارات السليمة. ومصادر جمع المعلومات هي:

1- المصادر الثانوية لجمع المعلومات secondary sources:

- أ- وتسمى أيضاً بالمصادر التاريخية "Historical sources" وتتضمن البيانات المخوذة من النشرات الإحصائية التي تنشرها الوزارات والمؤسسات الحكومية والخاصة والهيئات والشركات المختلفة كل في مجال نشاطها، وما يتوفر في المكتبات من مؤلفات ومطبوعات تضم معطيات إحصائية . من مثل سجلات الولادات والوفيات، أو السجلات الخاصة بطلبة الجامعات والمدارس وغيرها (الفهارس، الدوريات) .

2- المصادر الميدانية لجمع البيانات Prisoner Field sources:

وتسمى أيضاً بالمصادر الأولية "Field sources" ويجري هنا الحصول على البيانات من وحدات المجتمع بصورة مباشرة عن طريق الاستبيانات والاستمارات والاختبارات المعدة لهذا الغرض. وتتميز هذه المصادر بالدقة والثقة، ما يساعد على الحصول على نتائج موثوقة بدرجة كبيرة، ولا بد لهذه الأدوات من أن تتوفر فيها شروط الصلاحية وهي الصدق والثبات، هذا وتستخدم الأساليب الإحصائية من أجل التأكد من صدق المقاييس وثباتها. إلا أن هذه المصادر تحتاج إلى وقت وجهد وتكلفة وجهد كبير حتى يستطيع الباحث الحصول على بيانات واقية ودقيقة .

أساليب جمع البيانات الميدانية:

هناك أسلوبان لجمع البيانات الميدانية :

أولاً : أسلوب الحصر (المسح) الشامل :

يجري في هذا الأسلوب جمع البيانات من جميع عناصر أو مفردات المجتمع الإحصائي، وبالتالي تكون النتائج على درجة عالية من الدقة والتفصيل والوضوح، وهي تحتاج إلى تكلفة ووقت كبير، وتحتاج أيضاً إلى وجود العديد من الباحثين، وفي حال المجتمع الكبير جداً قد يصعب أو يستحيل جمع البيانات بهذا الشكل.

ثانياً : أسلوب العينة :

يجري وفق هذا الأسلوب جمع البيانات من جزء من المجتمع الإحصائي يسمى بالعينة، التي يجب أن تكون ممثلة في خصائصها لخصائص المجتمع قيد الدراسة تمثيلاً صادقاً

----- الفصل الثاني: البحث العلمي والإحصائي

وسليماً وصحيحاً، ومن ثم يجري تعميم النتائج التي تم الحصول عليها من العينة على المجتمع التي سحبت منه العينة. (فهيم، 2005 ص34-35-36؛ البلداوي، 2004، ص19)

ويتميز المسح بالعينة كما لخص كوهن (1977) Cochran على المسح الشامل بأنها: تقلل تكاليف المسح، وسرعة الحصول على النتائج، ودقة أكبر في النتائج ومدى أوسع، وكذلك سهولة تعديل وتبديل المسح بالعينة (علوان، 1993، ص6)

3- **تنظيم البيانات وعرضها** : يقوم الباحث بتبويب البيانات وعرضها باستخدام الجداول والرسوم البيانية، والغرض من تنظيم البيانات في جداول ورسوم هو تسهيل فهمها وقراءتها، كذلك جعلها جاهزة للتحليل الإحصائي المناسب.

4- **تحليل البيانات** : هنا يبدأ الباحث باستخدام الأساليب والطرائق الإحصائية المناسبة لأهداف بحثه، سواء أكانت أساليب وصفية مثل مقاييس النزعة المركزية أم مقاييس التشتت أو معاملات ارتباط أم كانت أساليب تحليلية باستخدام اختبارات الدلالة الإحصائية : وبعد الحصول على النتائج الإحصائية يعتمد الباحث إلى تفسيرها، والتوصل إلى تعميمات وتوصيات . (أبو زينة وآخرون، 2006، ص18-19؛ أبو علام، ص81-85)

المجتمع الإحصائي والعينة Statistical Population and sample:

يعرف المجتمع الإحصائي : بأنه مجموعة من الأفراد أو الأشياء أو الحوادث التي يتناولها البحث، أي أنه المجموعة الشاملة التي يجري اختيار العينة منها .

ولذلك لابد من تعريف المجتمع وتحديدّه تحديداً دقيقاً، أي تحديد ما يشمله المجتمع من مفردات (أفراد، أشياء ...) على سبيل المثال : تحديد المجتمع بأنه جميع طلبة كلية التربية، أو

----- الفصل الثاني: البحث العلمي والإحصائي

قد يكون المجتمع أوسع من ذلك فنقول جميع طلبة جامعة دمشق، أو جميع طلبة الجامعات السورية الحكومية . ولابد عند تحديد المجتمع وتعريفه من تحديد خصائص المجتمع من حيث الجنس، والعمر، والمنطقة الجغرافية، والوضع الاقتصادي الاجتماعي.

وتجدر الإشارة إلى أنه إذا كان المجتمع صغيراً فإنه يمكن دراسته وتناول جميع عناصره وأفراده، و في هذه الحالة نقوم بالمسح الشامل. أما إذا كان المجتمع كبيراً ومكون من مجموعة كبيرة جداً من العناصر والأفراد، ولا يمكن تناول جميع عناصره بعملية المسح الشامل، فإنه لابد من الاختصار على عينة منهم ونصف المجتمع على أنه مجتمع غير محدود (أبو زينة، 2007، ص19-20؛ أبو علام، 2004، ص151-153)

ويتصف المجتمع بصفتين رئيسيتين :

- 1- المرونة: ويقصد به إمكانية التحكم بالإطار الزمني والمكاني.
- 2- التجانس: ويقصد به أن جميع الأفراد متجانسين من حيث الخاصية

العينة تعريفها وأنواعها :

العينة : مجموعة جزئية من المجتمع الإحصائي يجري اختيارها بطريقة وحجم نستطيع من خلالها أن نقوم بعملية التعميم .

ويجري اختيار العينة من المجتمع الإحصائي نظراً لضخامة المجتمع وصعوبة الوصول إلى جميع أفراده، لهذا تستخدم النتائج المستخلصة من العينة للاستدلال على معالم المجتمع أي تعميم نتائج العينة على المجتمع الذي سحبت منه.

----- الفصل الثاني: البحث العلمي والإحصائي

ولهذا يلزم أن تكون العينة ممثلة تماماً في خصائصها لخصائص المجتمع حتى نستطيع القيام بعملية التعميم. هذا و تسهم عملية سحب العينة في توفير الوقت والتكاليف التي يمكن أن تتفق في حال دراسة المجتمع بأكمله، وكلما كان اختيار العينة دقيقاً وسليماً تكون الفروق بين قيم العينة وقيم المجتمع أقل و نسبة الخطأ أقل.

ويسمى كل مقياس إحصائي محسوب من العينة **بالتابع الإحصائي** (والتابع قد يكون متوسط، نسبة، انحراف معياري، معامل ارتباط) ويسمى كل مقياس إحصائي محسوب من المجتمع **الثابت الإحصائي**. ومقدار الفرق بين التابع والثابت يعرف **بخطأ المعاينة**.

وتشير كلمة المعاينة إلى عملية تحديد وسحب أفراد الدراسة، وتتحدد الخطوات المنطقية للمعاينة بالخطوات التالية: 1- تحديد المجتمع الكلي للأفراد الذين تعمم عليهم نتائج البحث. 2- تحديد مجتمع البحث الذين يجري سحب العينة من قبلهم. 3- سحب عينة الدراسة وهم الأفراد الذين يكونون موضع الدراسة. والعينة قد تحدد بصورة ضيقة محدودة من مثل (النساء المتزوجات من عمر 35- 45 سنة اللائي يقطن في مدينة دمشق) أو يمكن أن يحدد بصورة واسعة من مثل (النساء السوريات من عمر 18 سنة فما فوق).

وتقسم العينات إلى نوعين رئيسيين:

العينات العشوائية أو الاحتمالية Probabilistic samples:

وهي العينات التي يتم اختيارها باتباع طرائق علمية غير متحيزة تستند إلى مبادئ الاحتمالات بغرض تمثيل المجتمع. (عودة و الخليلي، 1988، ص173)

والعينات العشوائية لها عدة أنواع هي :

1- العينة العشوائية البسيطة Simple Random sample:

تتميز هذه العينة بأنها تلك العينة التي تتاح فيها الفرصة لكل مفردة من مفردات المجتمع بالظهور في العينة و تكون الفرص متكافئة لا انحياز فيها .

أي أن احتمال سحب أي مفردة من مفردات المجتمع يكون ثابتاً ويساوي

عدد مرات تكرار المفردة

عدد أفراد المجتمع

فمثلاً إذا أردنا سحب 50 طالباً من أصل 1000 طالب فإن احتمال سحب كل طالب =

$$5\% = \frac{5}{100} = \frac{50}{1000}$$

وهناك طرائق عديدة لاختيار العينة العشوائية البسيطة منها طريقة القرعة وفي هذه

الطريقة تكتب أسماء جميع أفراد المجتمع في بطاقات صغيرة وتخلط مع بعضها، ويجري سحب المطلوب من هذه البطاقات دون تحيز .

كما أنه من الممكن اللجوء إلى جداول الأعداد العشوائية بحيث يجري اختيار أي رقم

من تلك الجداول بصورة مستقلة عن الأرقام الأخرى، مع كون الفرصة المتاحة لجميع الأرقام بالظهور في العينة .

ويمكن أيضاً الاستفادة من دواليب اليانصيب، وهنا نرقم أفراد المجتمع ابتداءً من

الصفراً، ونستخدم عدداً من الدواليب بعدد منازل الرقم الذي يمثل حجم العينة -1 فإذا كانت

الحجم المطلوب للعينة 1500 يلزمنا ثلاث دواليب والرقم الذي يظهر في هذه الدواليب يمثل رقم

الفرد في المجتمع والذي اختير كأحد أفراد العينة، كذلك يمكن الاستعانة بالحاسبات لاختيار العينات

2- العينة العشوائية المنتظمة : Systematic Random Sample

إذا توافرت قوائم بأسماء أفراد جميع أفراد المجتمع وكان هؤلاء مرتبين بصورة عشوائية فإن الباحث يمكن أن يختار عينة دراسته بصورة دورية، فإذا أردنا اختيار عينة منتظمة مؤلفة من 50 فرداً من أصل 500 فرداً فإن فترة السحب سوف تتحدد بقسمة حجم المجتمع إلى حجم العينة أي $K = \frac{N}{n}$

من المجتمع $\frac{500}{50} = 10$ ويسمى حاصل قسمة $\frac{\text{حجم المجتمع}}{\text{حجم العينة}}$ بفترة السحب وهذا يعني أننا سوف نأخذ فرداً من بين كل عشر أفراد متتاليين، على أن نحدد رقم البداية، فإذا حددنا رقم البداية بالرقم 5 فإن الأرقام التالية التي يجب سحبها هي

45	35	25	15	5
----	----	----	----	---

وعند عدم توفر قوائم بأسماء الأفراد وليكن (طلبة المدارس) فإن الباحث يمكن له الوقوف أمام المدرسة واختيار كل عاشر طالب يدخل المدرسة إذا كانت فترة السحب 10 وكل خامس طالب أو ثاني طالب حسب فترة السحب .

ومن مزايا العينة العشوائية المنتظمة أنها تقلل من احتمالات الخطأ، كما أنها عينة سهلة الاستخدام في مجالات العلوم التربوية والنفسية، كما أن تباين وحدات العينة المنتظمة قد يكون أقل من تباين وحدات العينات الاحتمالية الأخرى .

----- الفصل الثاني: البحث العلمي والإحصائي

على الرغم من أن هذا الأسلوب في اختيار العينة قد يعطي عينة ممثلة للمجتمع الإحصائي يكون فيها خطأ المعاينة أقل من خطأ المعاينة في العينة العشوائية البسيطة إلا أن فيه بعض المخاطر (عودة و الخليلي، 1988، ص176). وتأتي هذه المخاطر من أن نقطة البداية تحدد بقية عناصر العينة، أي ليس هناك استقلال بين عناصر العينة لذلك لا يمكن استخدام هذه العينة إذا شعر الباحث بأن هنالك أية ملامح لترتيب دوري في خاتمة عناصر المجتمع، فإذا أراد باحث أن يختار عينة من 100 طالباً وطالبة من أصل 1000 طالباً وطالبة من قسمي معلم الصف وقسم التربية وكان أسماء الطلبة في المجتمع مرتبين معلم الصف أولاً والتربية ثانياً، فإن هذا يعني أن جميع أفراد العينة سيكونون من طلبة معلم الصف، لهذا لا بد من التأكد من أسماء الطلبة من القسمين مرتبة بصورة مختلطة وعشوائية.

مزايا العينة المنتظمة :

تمتاز هذه العينة بأنه إذا كان المجتمع مكوناً من مجموعات، وكانت مفردات كل مجموعة مرتبة مع بعضها في الإطار، فإننا نتوقع أن تمثل كل مجموعة منها في العينة يكون بنفس وزنها في المجتمع، خاصة إذا كان عدد مفردات المجموعة لا يقل عن طول الفترة فمثلاً في حالة 100 فرد منها 60 من الذكور و 40 من الإناث وكانت القائمة تشمل الـ 60 ذكراً ثم الـ 40 أنثى وأخذنا عينة عشوائية منتظمة منهم بواقع 10% فإننا نرى أن العينة سوف تشمل على 6 ذكور و 4 إناث (فهمي، 2005، ص106)

عيوبها :

1- عدم صلاحيتها إذا ما وجدت علاقة دورية مع ترتيب العناصر في القائمة، لأنه إذا كان ترتيب المفردات داخل الفئات يتم بطريقة مرتبطة مع البيانات المراد جمعها، فإن العينة المنتظمة تكون متحيزة بدرجة كبيرة. فمثلاً إذا كان المجتمع مكوناً من مجموعة

طلاب الابتدائي منتظمة وفق تقدير متوسط تحصيلهم، فإن العينة ستكون متحيزة، إذا أخذنا أول طالب من الأوائل لأن جميع أفراد العينة سيكونون من الطلاب المتفوقين.

2- إن اختيار الوحدة الأولى عشوائياً في العينة المنتظمة لا يعني أن تكون العينة المنتظمة مختارة كلية على أساس عشوائي، هذا خلاف العينة العشوائية البسيطة التي تختار جميع وحداتها عشوائياً (فهمي، 2005، ص106)

3- العينة العشوائية الطبقية Satisfied random sample :

عندما يكون المجتمع مقسماً إلى مجموعات غير متجانسة في الخصائص التي يقوم الباحث بدراستها ومن أجل ضمان التمثيل الجيد لكل مجموعة من هذه المجموعات ضمن العينة، نقوم بتقسيم المجتمع إلى طبقات على أساس متغير واحد مثل الجنس أو العمر، المستوى الاقتصادي الاجتماعي، المهنة ...) أو أكثر من متغير وذلك حسب موضوع الدراسة على أن تكون كل طبقة متجانسة فيما بينها قدر الإمكان، ومن ثم يسحب عينة من كل طبقة إما بطريقة عشوائية بسيطة أم بطريقة عشوائية منتظمة أما حجم العينة المسحوبة في كل طبقة فيتوقف على :

1- حجم الطبقة : إذ يسحب من كل طبقة بما يتناسب مع نسبة هذه الطبقة إلى المجتمع الأصلي الإحصائي.

2- مدى التجانس داخل الطبقة : فكلما زادت درجة التجانس بين مفردات الطبقة، قل حجم العينة المسحوبة في هذه الطبقة (فهمي، 2005، ص107) وتوفر العينة التي يتم اختيارها بهذه الطريقة إمكانية إجراء المقارنة فيما بين الطبقات .

وتتميز هذه العينة على العينة العشوائية البسيطة، في أن العينة العشوائية البسيطة قد تركز على إحدى الطبقات مهمة طبقات المجتمع الأخرى مما يضعف تمثيلها لهذا تعد العينة الطبقية من أفضل العينات تمثيلاً للمجتمعات غير المتجانسة .

مزايا العينة العشوائية الطبقية :

- تساعد في الحصول على عينة أكثر تمثيلاً لمجتمع البحث الذي اختيرت منه .
- نظراً إلى قدرة العينة الطبقية التمثيلية العالية، فإننا نحصل من خلال هذه العينة على قيم للمقاييس تقترب من القيم الحقيقية للمجتمع .
- الدقة العالية في النتائج المتحصلة في العينة العشوائية الطبقية .
- تستخدم هذه العينة في المجتمعات الكبيرة الحجم .
- وتتلخص الطريقة بتحديد حجم العينات الجزئية متناسبة من كل طبقة على أساس المعادلة
- $\text{حجم العينة الطبقية} = (\text{حجم الطبقة} \div \text{حجم المجتمع}) \times \text{حجم العينة}$
- (بركات، 2007، ص 6،7)

مثال : إذا أراد باحث أن يستطلع آراء طلبة كلية التربية حول جدول دوام المحاضرات، فلا بد له من أخذ آراء الطلاب من جميع التخصصات وكان أعداد الطلاب في التخصصات المختلفة كما يلي :

الجدول (1-2)

الاختصاص	عدد الطلبة
تربية	2000
علم النفس	1500
رياض أطفال	200
إرشاد نفسي	1000
معلم صف	2300
المجموع	7000

----- الفصل الثاني: البحث العلمي والإحصائي

ولتحقيق هذا الغرض اختار عينة مكونة من 100 طالب، وأخذ في الحسبان عدد الطلبة في كل اختصاص فيكون الاختيار كما يلي :

$$\text{من طلبة التربية} = 100 \times \frac{2000}{7000} = 28.57 \approx 29 = \text{طالباً}$$

$$\text{علم نفس} = 21.42 \approx 21 \text{ طالباً}$$

$$\text{معلم صف} \approx 33 \text{ طالباً}$$

$$\text{إرشاد نفسي} \approx 14 \text{ طالباً}$$

$$\text{رياض أطفال} \approx 3$$

4- العينة العشوائية العنقودية Cluster Random Sample:

تستند طريقة سحب هذه العينة العنقودية إلى تقسيم مجتمع البحث إلى عدد من العناقيد (مجموعات) وفق معيار معين غالباً ما يكون جغرافياً وهذه المجموعات (ليست طبقات ذات خصائص مشتركة) تسمى كل مجموعة عنقوداً وتكون وحدة الاختيار في هذا النوع من العينات هي العنقود، ويتم اختيار هذا العنقود عشوائياً وبهذا تكون وحدة السحب في هذه العينة هي المجموعة وليس الفرد كما هو الحال في العينات السابقة. وفي بعض الأحيان تختار العينة من جميع أفراد هذه الوحدات وتسمى حينها عنقودية ذات المرحلة الواحدة وفي حالات أخرى تقسم هذه الوحدات إلى وحدات جزئية، ثم تسحب العينة بمرحلة إضافية وتسمى عملية سحب العينة هنا بالعينة العنقودية ذات المراحل المتعددة "Multi-stage cluster sample".

والمثال التالي يوضح كيفية سحب عينة عنقودية :

لنفرض أن مدينة دمشق مقسمة إلى عدة مناطق جغرافية ولتكن ثمان مناطق (عناقيد) وجرى اختيار أربع مناطق منها عشوائياً منها : ولتكن هذه المناطق

----- الفصل الثاني: البحث العلمي والإحصائي

الجنوب	الشمال	الشرق	الغرب
--------	--------	-------	-------

يتم اختيار عدد من المدارس من كل عنقود (منطقة) باستخدام إحدى الطرائق العشوائية، ولنفرض أن حجم العينات الجزئية كما يلي :

الجدول (2-2)

عدد المدارس المسحوبة من كل منطقة

المنطقة	عدد المدارس
الشرق	3
الجنوب	5
الشمال	5
الغرب	4
المجموع	17

وفي مرحلة لاحقة يمكن أن يختار عدد من الصفوف من كل مدرسة ولتكن على النحو التالي:

الجدول (2-3)

عدد المدارس والصفوف المسحوبة من كل منطقة

المنطقة	عدد المدارس	عدد الصفوف
الشرق	3	2
الجنوب	5	3
الشمال	5	3
الغرب	4	2
المجموع	17	10

وبالتالي يصبح عدد الوحدات الجزئية (العناقيد) الجزئية عشر صفوف ثم يتبع أسلوب الحصر الشامل أي يختار جميع الطلبة في تلك الصفوف .

ولابد من التنويه أن اختيار العينات في كل مرحلة يجري بصورة عشوائية إما بالطريقة العشوائية البسيطة أو الطبقية أو المنتظمة .

5- العينة العشوائية المتعددة المراحل Multi-stage sample:

عندما تكون المجتمعات كبيرة الحجم يصعب سحب عينة عشوائية بسيطة أو منتظمة أو طبقية، ولكن بدلاً منها تستخدم العينة العشوائية متعددة المراحل . ويناسب هذا النوع من العينات الدراسات السكانية أو الدراسات في مجال الجغرافيا الاقتصادية . ويتم الاختيار العشوائي لعدد من المفردات لكل قسم بحيث يتم ذلك تتابعياً، يتم الاختيار العشوائي من القسم الأول كمرحلة أولى، ثم يتم الاختيار العشوائي من القسم الثاني كمرحلة ثانية، وهكذا حتى يصل إلى الاختيار في المرحلة النهائية ولذا أطلق على هذا النوع من العينات بأنه متعدد المراحل (فهومي، 2005، ص110) .

فإذا أردنا سحب عينة لدراسة حالة الوضع الاجتماعي للأسر القروية على مستوى القطر، فإنه يمكن البدء بسحب عينة من المحافظات، وفي مرحلة لاحقة يجري اختيار عينة من النواحي التابعة للمحافظات، وفي مرحلة ثالثة تسحب عينة من القرى، ومن ثم يجري اختيار عينة عشوائية من أسر القرى المختارة، وفي هذه العينة يكون خطأ المعاينة أكبر. أما عيوب هذه العينة فهو كثرة عدد المراحل التي قد تتضمنها العينة، مما يضعف العلاقة بين معالم المجتمع الأصلي وخصائص العينة، مما يؤثر في تقدير معالم المجتمع من بيانات العينة المتحصل عليها في آخر مرحلة .

يتطلب هذا النوع من العينات جهداً كبيراً في تحديد حدود أو إطار كل مرحلة وتحديد حجم العينة الفرعية المطلوبة من كل منها، وذلك في ضوء الاعتبارات الخاصة بالاختيار في المعاينة العشوائية . (فهومي، 2005، ص111)

العينات غير الاحتمالية غير العشوائية Non probability sample:

يضطر الباحث أحياناً لاتباع التحيز الشخصي في اختيار مفردات العينة، أو ما يعرف بأسلوب المعاينة غير الاحتمالية، إما لعدم توفر الإمكانيات المادية المناسبة أو لوجود صعوبة في الحصول على مفردات المجتمع، وفي حالة العينات غير الاحتمالية يجري السحب دون مراعاة التكافؤ في الفرص لمفردات المجتمع للظهور في العينة .

وعادة تستخدم المعاينة غير العشوائية، في الأبحاث الاستطلاعية، بهدف التأكد من وضوح وسلامة الأدوات المستخدمة في الأبحاث، أو في مرحلة تجريب الاختبارات القبلية Prior "tests" للتأكد من وضوح التعليمات والأسئلة حتى يمكن إجراء التعديلات اللازمة في الأسئلة قبل سحب العينة الأساسية.

والعينات غير الاحتمالية على أنواع هي :

1- العينة العرضية Accidental sample: في بعض الظروف قد تتوافر للباحث عينة عن طريق الصدفة ممن يصادفهم من الأشخاص، مثل الإعلامي الذي يستطلع آراء بعض الناس الذين يصادفهم في طريقه. أو يأخذ الباحث عينته من طلبة المدرسة القريبة من بيته ويجري الدراسة عليهم، وتقيد هذه العينة في تكوين فكرة أولية عن موضوع البحث، وبالتالي الاستناد إلى هذه الفكرة في صياغة أهداف البحث وفرضياته، إلا أن نتائج هذه العينة لا يمكن تعميمها وبناء أحكام عليها لأن نتائجها لا تتعدى الإطار الضيق الذي خضع له الباحث .

2- العينة المقصودة Purposeful: في بعض الأحيان يلزم الحصول على معلومات من فئات محددة، بمعنى أنه لا يمكن الحصول على المعلومات المطلوبة إلا من هؤلاء الأشخاص أو أنهم أناس تتوافر لديهم المعايير المحددة من قبل الباحث من مثل سحب عينة من مدرء المدارس في خاصية معينة للكشف عن مشكلات التعليم في هذه الناحية، أو دراسة واقع التعليم الجامعي من

وجهة نظر المشرفين عليه في الجامعات الرسمية في بلد ما. الأمر الذي يستلزم سحب عينة من رؤساء الجامعات ونوابهم إضافة إلى عمداء الكليات ونوابهم .

3- عينة الحصص Quota Sample:

في هذه العينة يقسم المجتمع إلى حصص تشبه طبقات المجتمع في العينة الطبقية، يهدف الباحث في هذه العينة إلى الحصول على عينة ممثلة في خصائصها لخصائص المجتمع، وذلك باختيار عينة من كل حصة بالنسبة نفسها تمثل هذه الحصة إلى المجتمع إلا أن اختيار العينة من كل حصة لا يكون اختياراً عشوائياً، وإنما تختار بطريقة غير عشوائية قد تكون قصدية .

تقدير حجم العينة :

هناك عوامل عديدة تحدد حجم العينة وهي :

- أن التباين بين أفراد المجتمع وليس حجم المجتمع هو العامل الحاسم في تقدير حجم العينة، فكلما كبر حجم التباين بين أفراد المجتمع، كلما استلزم ذلك أن يكون حجم العينة كبيراً بغض النظر عن حجم ذلك المجتمع .
- عند تقدير حجم العينة لابد من الأخذ في الحسبان التكلفة المادية والزمن لإنجاز الدراسة . (عودة والخليلي، 1988، ص179)
- مدى التفصيل المطلوب في نتائج العينة كتقديرات لخصائص المجتمع فكلما زادت درجة التفصيل المطلوبة زاد حجم العينة المسحوبة.
- مدى الخطأ الذي يسمح به في نتائج العينة كتقديرات لخصائص المجتمع، فكلما قل مدى الخطأ الذي يمكن السماح به زاد حجم العينة .
- درجة الثقة المطلوب توفرها في تحقيق السمات السابقة . فكلما زادت درجة الثقة المطلوبة، زاد حجم العينة اللازم (فهيم، 2005، ص117)

----- الفصل الثاني: البحث العلمي والإحصائي

- حجم المجتمع : يجب أن يتناسب حجم العينة مع حجم المجتمع الأصلي.
- عندما يكون هناك توقع أن للمعالجة التجريبية أثراً قليلاً على المتغير التابع، أو عندما يتوقع أن تكون العلاقة بين المتغيرات قيد الدراسة فيها ضعيفة فإنه من المرغوب فيه أن تكون العينة كبيرة الحجم.
- عندما يكون هناك توقع لإهدار (فقدان) كبير في عدد المفحوصين، أو هناك عدداً من الأفراد لن يرد الاستبيانات، لابد من اختيار عينة كبيرة الحجم، حتى لا تتأثر النتائج بهذا الفقدان أو الهدر.

مع الإشارة إلى استخدام عينات ذات حجوم صغيرة في بعض البحوث ربما كان أفضل من استخدام عينات ذات أحجام كبيرة مثل الدراسات التي تعتمد على التحليل النفسي وأدوات القياس الإسقاطي (الشرييني، 1995، ص72). كما أن عينة كبيرة الحجم في التصميم التجريبي تعد ضعيفة لأن هناك عدد من المتغيرات غير المضبوطة.

وباختصار هناك اتجاهان لتحديد حجم العينة، **الاتجاه الأول** : يعتمد على خبرات الباحث وخبرات الباحثين الآخرين في هذا المجال، ووفقاً لهذا الأسلوب فإننا نجد أن الإحصائيين ينصحون بأن يكون الحد الأدنى للأفراد كما يلي (50-100) في الدراسات الارتباطية Co "relational studies" يفضل أن يكون حجم العينة في الدراسات التجريبية "experimental studies" ذات المعالجة الواحدة 30 فرداً، أما في الدراسات التجريبية ذات المعالجتين أو أكثر فيفضل ألا يقل عدد الأفراد في الخلية الواحدة عن 5 أفراد . أما في الدراسات المسحية "Survey studies" : فالعدد المناسب 20% من أفراد المجتمع الكلي إذا كان عدد المجتمع (500 إلى 1000) وتنخفض النسبة إلى 5% في المجتمعات الكبيرة جداً (عودة والخليلي، 1988، ص177-178; البطش وأبو زينة، 2007، ص105-106)

والاتجاه الثاني : يعتمد على القواعد الاحتمالية الرياضية لتحديد حجم العينة كاحتمال الوقوع في الخطأ من النوع الأول أو الخطأ من النوع الثاني .

أخطاء العينات :

1- أخطاء التحيز : ويقصد به اختيار ناتجة من التقصير في خطوات البحث العلمي العينات مع انحياز الباحث إلى فئة دون أخرى ويعطيها وزناً أكبر مما تستحق، مثل اختيار الطلاب المتفوقين في الصف ليختبر ذكائهم الأمر الذي يؤدي إلى رفع متوسطات الذكاء، أو أن يختار الباحث عينته ممن هم في متناول يده، كأن يختار أفراد العينة في مدرسة أو مدرستين قريبتان من مكان سكنه أو ممن يرتبطون بصلة أو علاقة ما بالقائمين على هذه المدارس وفي هذه الحالة ينطوي تعميم النتائج على مخاطر كثيرة . لأنه لم تؤخذ صفة العشوائية في الاختيار ولم تتوفر فرص متكافئة أمام جميع أفراد المجتمع للظهور في العينة . واختيار بعض الأفراد ليسوا من مجتمع الدراسة .

2- الأخطاء الاحتمالية: هذه الأخطاء لا يد للباحث فيها ناتجة عن دراسة جزء من المجتمع، وهي تنشأ في الاتجاهين الموجب والسالب وبالحجم نفسه، ولذلك فهي تلغي بعضها بعضاً وخصوصاً مع زيادة حجم العينة وذلك وفق نظرية النزعة المركزية "central limit Theorem " إذ أنه مهما كان شكل توزيع المجتمع الأصلي، فإن توزيع المعاينة للمتوسط الحسابي يؤول إلى التوزيع الطبيعي تدريجياً مع زيادة حجم العينة (حجم العينة $n \geq 30$ يمكن اعتباره شرطاً كافياً حتى يؤول توزيع المعاينة للمتوسط الحسابي إلى التوزيع الطبيعي). (فهمي، 2005، ص 293) أي أن الفرق بين قيمة متوسط العينة ومتوسط المجتمع يصغر مع زيادة حجم العينة، بمعنى أن مقدار الخطأ المعياري يقلص مع زيادة حجم العينة، لأن سحب عينة كبيرة يعني الحصول على معلومات أكثر، مما يقود إلى دقة أكبر في تلك المعلومات. (Rumsey, 2007, P.67)

أسئلة وتمارين خاصة بالفصل الثاني:

- 1- قارن بين العينة العشوائية والعينة غير العشوائية من حيث استخدام كل منها؟
- 2- قارن بين كل من العينة العشوائية البسيطة، والعشوائية المنتظمة، والعشوائية الطبقية والعشوائية العنقودية من حيث:
 - وحدة الاختيار في كل منها.
 - المواقف التي يفضل استخدام كل منها فيه.
- 3- اذكر الخطوات اللازم اتباعها لسحب عينة عشوائية عنقودية
- 4- اختر مشكلة بحثية في مجال تخصصك وبين أي نوع من العينات يناسب هذه المشكلة البحثية.

الفصل الثالث

أساليب تنظيم البيانات وعرضها

Methods of data organization and Presentation



مقدمة:

بعد الحصول على البيانات من المصادر المختلفة لابد من تنظيمها وعرضها بصورة ملخصة يسهل فهمها وقراءتها ومن ثم تحليلها واستخراج النتائج.

وتنظيم البيانات هذه يأخذ أشكالاً عديدة منها الجداول أو التوزيعات التكرارية ومنها بصورة الأشكال والرسوم البيانية .

1- أساليب تبويب البيانات (العرض الجدولي):

يعني العرض الجدولي للبيانات تنظيم هذه البيانات في صورة جداول تسمى بالجدول التكرارية، وهذه الجداول قد تكون جداول بسيطة إذا تناولت ظاهرة واحدة، وقد تكون جداول "مزدوجة" إذا تناولت دراسة العلاقة بين ظاهرتين. أو قد تكون الجداول "مركبة" إذا تناولت أكثر من ظاهرتين، وتفيد هذه الجداول (البسيطة أو المزدوجة أو المركبة) في التعرف على خصائص البيانات وتكون المنطلق والأساس للعملية الإحصائية المختلفة للبيانات .

إنشاء جداول تكرارية بسيطة :

يستخدم هذا النوع من الجداول التكرارية لوصف وتلخيص بيانات تتعلق بظاهرة (متغير) واحد فقط، سواء كان هذا المتغير كمياً أم نوعياً، وهو أسهل وأبسط أنواع الجداول تركيباً ومفهوماً .

أ- التوزيع التكراري لمتغيرات نوعية (وصفية):

يتألف الجدول التكراري للبيانات النوعية من عمودين: العمود الأول قائمة بتصنيفات المتغير، والثاني بتكرار كل تصنيف من تصنيفات المتغير، ويمكن إضافة عمود ثالث يمثل النسبة المئوية للأوجه المختلفة للبيانات .

----- الفصل الثالث :أساليب تنظيم البيانات وعرضها

ب- التوزيع التكراري لمتغير كمي منفصل :

ينشأ الجدول التكراري لمتغير كمي منفصل بالطريقة نفسها التي ينشأ فيها الجدول التكراري لمتغير نوعي

مثال :

فيما يلي جدول تكراري بسيط لمتغير كمي منفصل وهو عدد أيام الغياب لمجموعة من الموظفين.

الجدول (3-3)

توزيع عينة من الطلاب وفق عدد أيام الغياب اليومي

النسبة % Percentage	عدد الأفراد (التكرار) Frequency	عدد أيام الغياب
37.5	30	1
25	20	2
18.75	15	3
12.5	10	4
6.25	50	5
%100	80	المجموع

ج- التوزيع التكراري لمتغير كمي متصل :

يمكن إنشاء جدول تكراري لمتغير كمي متصل بالطريقة نفسها التي يتم فيها إنشاء جدول تكراري لمتغير نوعي أو كمي منفصل عندما يكون عدد البيانات قليل، أما إذا كان عدد البيانات كبيراً

----- الفصل الثالث :أساليب تنظيم البيانات وعرضها

فإنه يلزم تجميع المفردات ضمن فئات والتعامل معها بدلاً من التعامل معها مفردة وذلك لتسهيل العمليات الإحصائية، ومن أجل تبسيط عرض البيانات .

مثال :

فيما يلي جدول لبيانات كمية متصلة بصورة جدول تكراري بسيط:

الجدول (3-4)

الجدول تكراري بسيط لبيانات كمية متصلة

الوزن	عدد الأفراد (التكرار) Frequency	النسبة % Percent
50	20	40
53	12	24
60	8	16
65	6	12
67	4	8
المجموع	50	100%

الجدول التكراري للفئات :

ويتكون الجدول التكراري من فئات القيم إضافة إلى عمود التكرارات

وله عدة أساليب والأسلوب الأسهل هو أن نبدأ بالحد الأدنى ثم نضيف طول الفئة.

تفريغ البيانات : ويمثل الأرقام الواقعة ضمن الفئة ونرمز لها ب(//)

التكرارات : نحول الإشارات في جدول تفريغ البيانات إلى أرقام .

يجب أن يكون حاصل التكرارات مساوياً حجم العينة وإلا كان هناك خلل .

----- الفصل الثالث: أساليب تنظيم البيانات وعرضها

وعند كتابة الفئات نبدأ من الصفر في حال إذا كانت أصغر قيمة قريبة من الصفر أي إذا كانت أقل من طول الفئة.

يترك الباحث الفئة مفتوحة لأحد السببين:

1- المحافظة على سرية البيانات .

2- هو أن الدراسة لم تنته بعد.

وفيما يلي تلخيص لخطوات إنشاء الجدول التكراري للفئات

1- حساب المدى الكلي وهو الفرق بين أعلى قيمة وأصغر قيمة للبيانات، وذلك وفق

العلاقة التالية :

المدى = أكبر قيمة - أصغر قيمة

2- تحديد عدد الفئات المناسبة لعرض البيانات، ويفضل أن لا يقل هذا العدد عن (5)

فئات وأن لا يزيد عن (15) فئة.

3- تحديد طول الفئة وفق العلاقة

$$\text{طول الفئة} = \frac{\text{المدى الكلي}}{\text{عدد الفئات}}$$

ويرمز لطول الفئة بـ C

ويجري التقريب إلى الأعلى عندما يكون الناتج عدداً كسرياً .

ويحسب طول الفئة استناداً إلى علاقة سترجس وفق العلاقة التالية :

طول الفئة = الفرق بين أكبر قيمة وأصغر قيمة (المدى الكلي)

+1 3.322 لغ عدد القيم (n)

$$= \frac{\text{Range}}{1 + 3.322 \log(n)}$$

4- تحديد عدد الفئات المناسبة لعرض البيانات، ويفضل أن لا يقل هذا العدد عن (5) فئات وأن لا يزيد عن (15) فئة.

5- بعد تحديد الحد الأدنى للفئة الأولى والذي يمثل أصغر قيمة من قيم المفردات، نحدد الحد الأعلى للفئة الأعلى وذلك بإضافة طول الفئة إلى الحد الأدنى، والحد الأعلى للفئة الأولى هو حد أدنى للفئة التي تليها وهكذا لباقي الفئات وبهذا يمكن تعريف طول الفئة أنه الفرق بين الحدين الأعلى والأدنى.

6- إيجاد عدد القيم في كل فئة من فئات الجدول التكراري ويسمى هذا العدد بالتكرارات (Frequencies) ومجموع التكرارات يساوي عدد البيانات ويرمز للتكرار بـ f ومجموع التكرارات $\sum f$. نقوم بعملية العد، وفي كل مرة نصادف درجة نضع إشارة هكذا (/) في العمود الخاص بالإشارات ونسجل عدد الإشارات في عمود التكرار . والأرقام في عمود التكرار تظهر لنا كم مرة تكررت المفردات داخل كل فئة .

قواعد تحديد الفئات :

- تحديد عدد الفئات المناسبة لعرض البيانات، ويفضل أن لا يقل هذا العدد عن (5) فئات وأن لا يزيد عن (15) فئة.
- يجب أن تكون الفئات شاملة وغير متداخلة ولا توجد بينها فجوات .
- يستحسن أن يكون أطوال الفئات متساوية .
- ينبغي الابتعاد عن استخدام الفئات المفتوحة وتعني الفئات المفتوحة الفئات الحدية التي يكون حدها الأعلى أو الأدنى مجهولاً لأن في استخدام الفئات المفتوحة صعوبة في حساب بعض المقاييس الإحصائية

----- الفصل الثالث :أساليب تنظيم البيانات وعرضها

مثل القول عن الفئة الأولى أقل من 18 أو أكثر من (42) للفئة الأخيرة .

(فهيمى، 2005، ص159)

افتراضات خاصة بالفئات :

الافتراض الأول :

أن مركز الفئة يمثل تلك الفئة، ومركز الفئة يحسب بجمع الحد الأدنى للفئة والحد الأعلى للفئة نفسها وقسمة المجموع على 2

$$\text{مركز الفئة} = \frac{\text{الحد الأدنى} + \text{الحد الأعلى}}{2}$$

الافتراض الثاني :

المفردات في أي فئة تتوزع بالتساوي بين حدود الفئة ففي الجدول (3-5) نفترض أن المفردات الثمانية في الفئة تتوزع على عناصر الفئة بواقع تكرار واحد لكل مفردة .(أبو زينة، الشايب، 2006، ص78)

مثال:

فيما يلي بيانات لدرجات طلبة كلية التربية في مادة الإحصاء، والمطلوب إنشاء جدول توزيع تكراري ذو فئات متساوية عددها (5)

10	15	20	25	32	32	35	36
37	39	40	40	40	45	49	50

أولاً : نحسب المدى الكلي = أكبر قيمة - أصغر قيمة

----- الفصل الثالث: أساليب تنظيم البيانات وعرضها

$$50 - 10 = 40$$

ثانياً : نحسب طول الفئة = $\frac{\text{المدى الكلي}}{\text{عدد الفئات}}$

$$C = \frac{40}{5} = 8$$

ثالثاً: نحدد الحد الأدنى للفئة الأولى وهو يساوي (10) وهي أصغر قيمة موجودة في البيانات.

رابعاً : نحسب الحد الأعلى للفئة الأولى وذلك بإضافة طول الفئة (8) إلى الحد الأدنى ليصبح الحد الأعلى (18)

خامساً : يبدأ الحد الأدنى للفئة الثانية بالحد الأعلى للفئة الأولى وهكذا لبقية الفئات.

سادساً : نحسب التكرارات، عن طريق وضع إشارة (/) لكل درجة في العمود الخاص بالإشارات ونسجل عدد الإشارات في العمود الخاص بالتكرار.

سابعاً : هذا يمكن حساب مراكز الفئات من خلال العلاقة التالية :

$$\text{مركز الفئة} = \frac{\text{الحد الأدنى} + \text{الحد الأعلى}}{2}$$

$$14 = \frac{10+18}{2} \quad \text{وبذلك يكون مركز الفئة الأولى}$$

مثال :

الجدول (3-5)

جدول تكراري للفئات

مركز الفئة	التكرار	الإشارات	الفئات
14	2	//	10-18
22	2	//	18-26
30	2	//	26-34
38	7	// ###	34-42
46	3	///	42-50
	16		

رموز الجدول التكراري : التكرار: f مركز الفئة: xi

إنشاء جداول تكرارية بسيطة متجمعة :

هناك جداول تكرارية متجمعة صاعدة وجداول تكرارية متجمعة هابطة، ومنها نستنتج ما يسمى بالجداول التكرارية النسبية المتجمعة الصاعدة أو الهابطة والجدول (3-6) يبين الجدول التكراري المجتمع الصاعد والجدول التكراري المتجمع الهابط لفئات تمثل أعمار مجموعة من الطلاب.

----- الفصل الثالث: أساليب تنظيم البيانات وعرضها

الجدول (3-6)

الجدول التكراري المتجمع الصاعد

فئات القيم	f التكرار	$f \uparrow$ التكرار المتجمع الصاعد	التكرار المتجمع الصاعد النسبي $f \downarrow \%$
18-20	3	3	10
20-22	4	7	23.33
22-24	6	13	43.33
24-26	8	21	70
26-28	4	25	83.33
28-30	3	28	93.33
30-32	2	30	%100
المجموع	30		

الملاحظ من الجدول أنّ التكرارات المتجمعة الصاعدة تبدأ بتكرار أول فئة وتنتهي بالمجموع الكلي للتكرارات .

الجدول (3-7)

الجدول التكراري المتجمع الهابط

فئات القيم	f التكرار	$f \downarrow$ التكرار المتجمع الهابط	التكرار المتجمع الهابط النسبي $f \downarrow \%$
18-20	3	30	%100
20-22	4	27	90
22-24	6	23	76.66
24-26	8	17	56.66
26-28	4	9	30
28-30	3	5	16.66
30-32	2	2	6.66
المجموع	30		

----- الفصل الثالث: أساليب تنظيم البيانات وعرضها

تبدأ التكرارات المتجمعة الهابطة دائماً بالمجموع الكلي للتكرارات (كما تبدأ النسب المئوية الهابطة بـ 100%).

أما التكرارات المتجمعة النسبية (الصاعدة أو الهابطة) فتحسب بقسمة التكرار المتجمع لكل فئة على مجموع التكرارات وضرب الناتج بـ 100 فمثلاً التكرار المتجمع النسبي لفئة (20-18) هو:

$$\frac{3}{30} \times 100 = 10$$

وتفيد التكرارات المتجمعة النسبية الصاعدة في تحديد النسبة التي تقل عن الحد الأعلى للفئة .

أما التكرارات المتجمعة النسبية الهابطة فتفيد في تحديد النسبة التي تزيد عن الحد الأدنى للفئة.

مثال:

لدينا البيانات التالية:

50	40	30	90	44	53	61	74	88
15	24	40	50	62	54	63	55	48
70	30	60	50	40	20	22	33	22
			33	55	66	50	80	20

$$\frac{\text{الفرق بين أكبر قيمة و أصغر قيمة}}{\text{عدد الفئات}} = \frac{\text{المدى}}{\text{عدد الفئات}} = \text{طول الفئة}$$

بفرض أن عدد الفئات يساوي 8 بهذا يكون طول الفئة

$$C = \frac{90 - 15}{8} = 9.3 = 9$$

----- الفصل الثالث: أساليب تنظيم البيانات وعرضها

الجدول التكراري :

جدول (3-8)

جدول تكراري

الفئات	الإشارات	التكرارات f	تكرار متجمع صاعد $f \uparrow$	تكرار متجمع هابط $f \downarrow$	التكرار النسبي $f \%$	تكرار متجمع نسبي صاعد $f \%$ $\uparrow$	تكرار متجمع هابط $f \%$ $\downarrow$
15-24	////	5	5	33	$\frac{5}{33} \approx 0.15$	$\frac{5}{33} \times 100 = 15.15$	$\frac{33}{33} \times 100 = 100$
24-33	///	3	8	28	$\frac{3}{33} \approx 0.90$	$\frac{8}{33} \times 100 = 24.24$	$\frac{28}{33} \times 100 = 84.84$
33-42	////	5	13	25	$\frac{5}{33} \approx 0.15$	$\frac{13}{33} \times 100 = 39.39$	$\frac{25}{33} \times 100 = 75.75$
42-51	///// /	6	19	20	$\frac{6}{33} \approx 0.18$	$\frac{19}{33} \times 100 = 57.57$	$\frac{20}{33} \times 100 = 60.60$
51-60	///	4	23	14	$\frac{4}{33} \approx 0.12$	$\frac{23}{33} \times 100 = 69.69$	$\frac{14}{33} \times 100 = 42.42$
60-69	////	5	28	10	$\frac{5}{33} \approx 0.15$	$\frac{28}{33} \times 100 = 84.84$	$\frac{10}{33} \times 100 = 30.30$
69-78	//	2	30	5	$\frac{2}{33} \approx 0.06$	$\frac{30}{33} \times 100 = 90.90$	$\frac{5}{33} \times 100 = 15.15$
78 فأكثر أي فئة مفتوحة	///	3	33	3	$\frac{3}{33} \approx 0.90$	$\frac{33}{33} \times 100 = 100$	$\frac{3}{33} \times 100 = 9.09$
المجموع		33					

----- الفصل الثالث: أساليب تنظيم البيانات وعرضها

نكتب التكرار الأول ثم نضيف إليه التكرار الثاني فنحصل على التكرار الصاعد الثاني وهكذا
 $3+5=8$ ، $8+5=13$

التكرار التجميعي الهابط $f \downarrow$:

نكتب أولاً مجموع التكرارات وهنا هو 33 ثم نطرح منه التكرار الذي يقابله ثم الجدول $28-5=23$ ،
33، $28-3=25$ وهكذا :

التكرار النسبي % f : نحصل عليه من قسمة $\frac{\text{التكرار}}{\text{مجموع التكرارات}}$

التكرار النسبي الصاعد % $f \uparrow$: نحصل عليه من قسمة $\frac{\text{التكرار الصاعد}}{\text{مجموع التكرارات}}$

التكرار النسبي الهابط % $f \downarrow$: نحصل عليه من قسمة $\frac{\text{التكرار الهابط}}{\text{مجموع التكرارات}}$

أساليب العرض البياني للمتغيرات Graphical Presentation:

تعد أساليب العرض البياني وسيلة إضافية في وصف المتغيرات إلى جانب الجداول التكرارية إذ يساعد العرض البياني في إعطاء صورة سريعة أقرب إلى الذهن والتذكر عن الظاهرة المدروسة، كما أنه يمكن من إجراء المقارنات بين التوزيعات التكرارية المختلفة، إضافة إلى إعطائه بعض المؤشرات الإحصائية بسرعة وبدون استخدام الصيغ الرياضية .

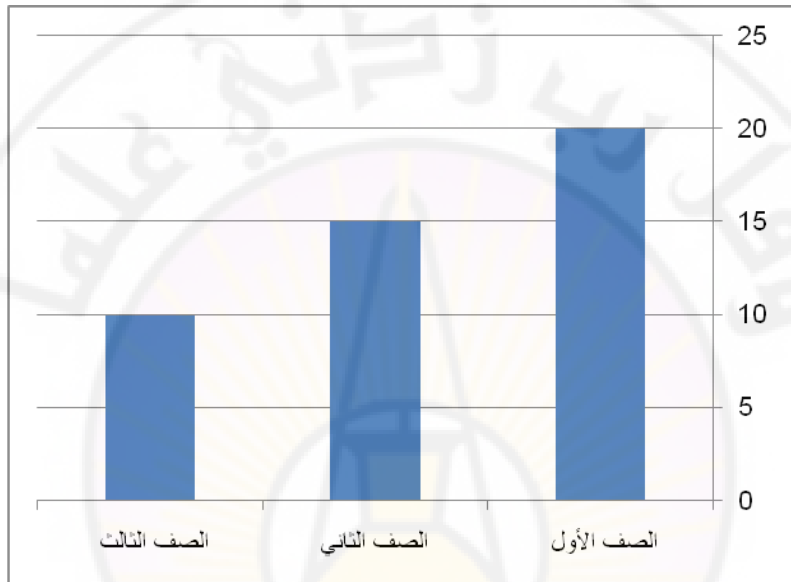
أ- التمثيل البياني للبيانات النوعية :

1- طريقة الأعمدة Bar Chart:

تمثل البيانات باستخدام الأعمدة برسم محورين متعامدين المحور الأفقي والمحور العمودي، وتوضع فئات المتغير على المحور الأفقي والتكرارات على المحور العمودي .

----- الفصل الثالث: أساليب تنظيم البيانات وعرضها

والشكل التالي يمثل رسماً بيانياً لمتغير الصف الدراسي لمجموعة من الطلبة، وهو متغير نوعي.



الشكل (1-3) : المدرج التكراري لبيانات نوعية

من الشكل البياني السابق يلاحظ أن عدد تلاميذ الصف الأول أكبر من عدد تلاميذ الصف الثاني والثالث ونظراً إلى أن البيانات نوعية فهناك فجوات بين الأعمدة .

2- طريقة الدوائر:

تستخدم طريقة الدوائر للعرض البياني من أجل إظهار نسب توزيع الظاهرة والمقارنة بينها، فمثلاً إذا كان لدينا بيانات عن أعداد الطلبة في كل اختصاص من الاختصاصات في كلية التربية على النحو التالي:

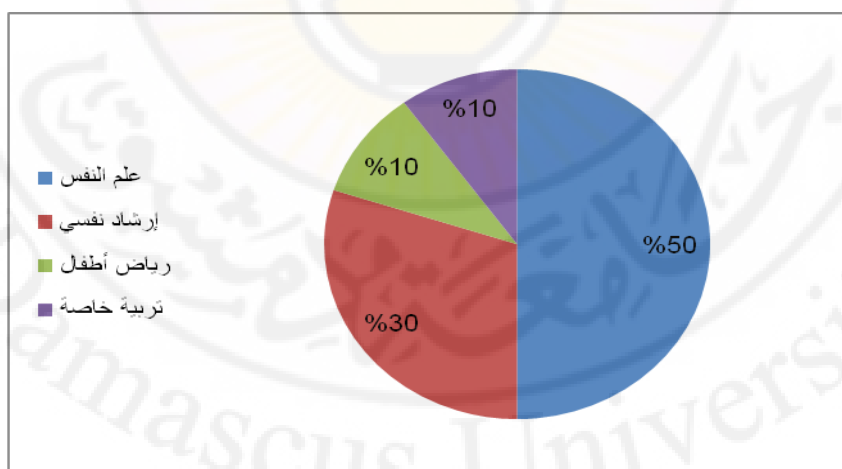
----- الفصل الثالث: أساليب تنظيم البيانات وعرضها

جدول (3-9)

الأعداد والنسب المئوية لاختصاصات الطلاب

الاختصاص	عدد الطلاب	النسبة المئوية
علم النفس	500	50%
إرشاد نفسي	300	30%
رياض أطفال	100	10%
تربية خاصة	100	10%
المجموع	1000	100%

يمكن تمثيل الجدول على شكل دائرة وذلك بحساب قطاعات الدائرة وفق المعلومات الواردة في الجدول السابق كما هو في الشكل التالي:



الشكل (3-2) دائرة بيانية

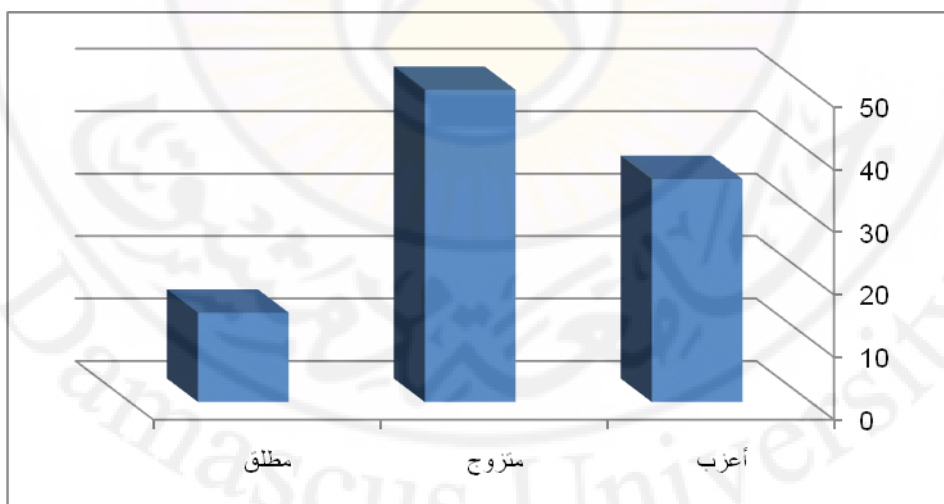
----- الفصل الثالث: أساليب تنظيم البيانات وعرضها

مثال:

لدينا التوزيع التكراري التالي الذي يمثل الحالة الاجتماعية لـ 28 فرداً، والجدول التكراري التالي يبين توزيع أفراد العينة.

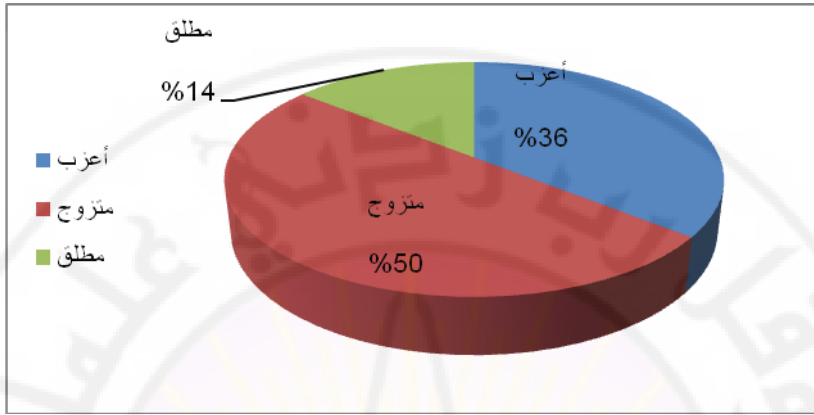
جدول (9-3) التوزيع التكراري للحالة الاجتماعية لأفراد العينة

الحالة الاجتماعية	العدد	النسبة المئوية %
أعزب	10	35.7
متزوج	14	50
مطلق	4	14.3
المجموع	28	100



الشكل (3-3) : أعمدة بيانية

----- الفصل الثالث: أساليب تنظيم البيانات وعرضها



الشكل (3-4): دائرة بيانية

ب- التمثيل البياني لمتغيرات كمية متصلة :

1- المدرج التكراري Frequency Histogram :

لرسم المدرج التكراري نرسم محورين متعامدين بحيث تمثل الحدود لكل فئة على المحور الأفقي، في حين تمثل التكرارات على المحور العمودي.

ويمكن إتباع الخطوات التالية في رسم المدرج:

1- نضع حدود الفئات على المحور الأفقي بحيث تكون الفئات متالصة ولا مسافات بينها والمسافات متساوية داخل الفئات .

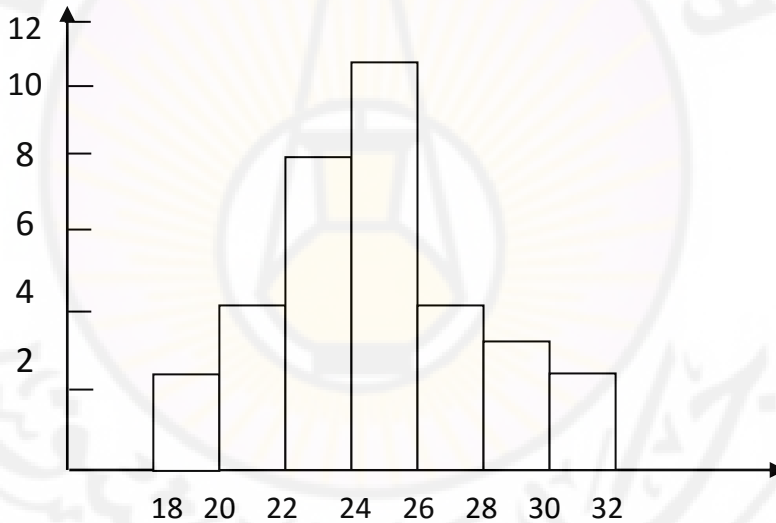
2- نقسم المحور العمودي (محور التكرارات) إلى أجزاء مناسبة باستخدام مقياس رسم مناسب، فإذا كان أعلى تكرار هو (20) فإنه يمكن أن نقسم المحور إلى 50، 10، 15، 20 أو 1، 2، 3 إلى 20 .

----- الفصل الثالث: أساليب تنظيم البيانات وعرضها

3- نرسم المستقيم الخاص بالفئة التي لها أكبر تكرار للمفردات بحيث يكون عرض المستطيل هو عرض الفئة وطوله، وارتفاعه متناسب مع تكرار الفئة، ثم نرسم باقي المستطيلات للفئات الأخرى .

مثال :

لو أخذنا البيانات السابقة من الجدول (3-6) والتي تمثل فئات الأعمار لمجموعة من الطلاب، فإننا نبدأ برسم المحورين الأفقي والعمودي، ونضع على المحورين الأفقي حدود الفئات، ولا حاجة لتكرار الحد الأعلى والحد الأدنى الحقيقي بل يكتب مرة واحدة.



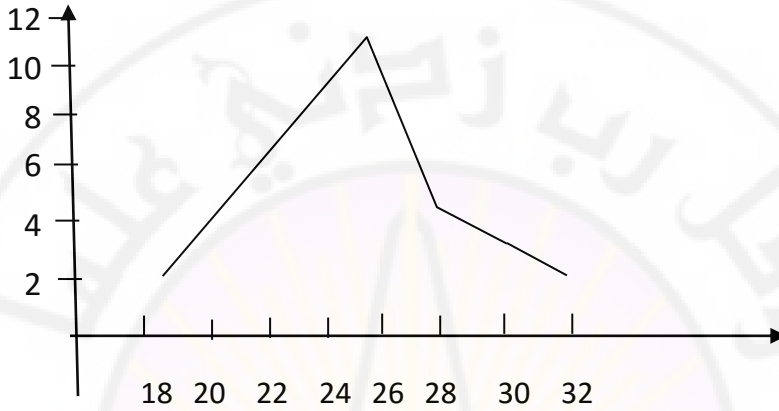
الشكل (1-5)

المدرج التكراري لأعمار 30 طالباً

2- المضلع التكراري Frequency Polygon: لرسم المضلع التكراري نستخدم مراكز الفئات بدلاً من حدود الفئات، ويمكن أن نحصل على المضلع التكراري بـتنصيف الأضلاع العلوية للمستطيلات في المدرج التكراري ثم وصل هذه النقاط بعضها مع بعض بقطع مستقيمة،

----- الفصل الثالث: أساليب تنظيم البيانات وعرضها

ولإغلاق الفئات نفترض وجود فئتين على الأطراف واحدة إلى أقصى اليسار وأخرى إلى أقصى اليمين بحيث يكون تكرار كل فئة مساوياً إلى الصفر .

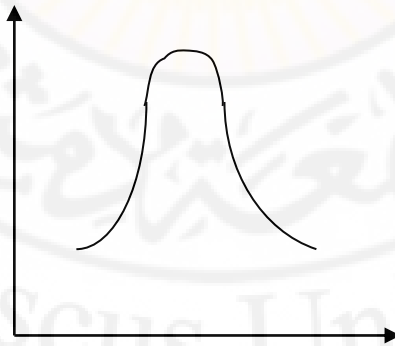


الشكل (3-6)

المضلع التكراري لأعمار 30 طالباً

3- المنحنى التكراري Frequency Curve:

يمكن التعبير عن البيانات بصورة منحنى تكراري، ونحصل على المنحنى التكراري بتسوية خطوط المضلع التكراري، ويفيد المنحنى التكراري في الدراسات الاستدلالية أكثر من الوصفية، وذلك للاستدلال على شكل التوزيع في المجتمع الذي أخذت منه عينة الدراسة



الشكل (3-7) : المنحنى التكراري لأعمار 30 طالباً

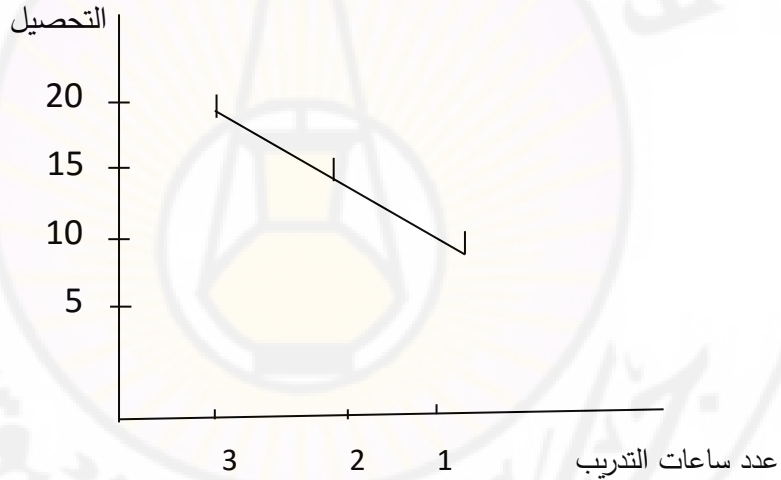
----- الفصل الثالث: أساليب تنظيم البيانات وعرضها

4- طريقة العرض بالخطوط البيانية The graphs lines:

إن طريقة العرض باستخدام الخطوط البيانية لا تختلف في جوهرها عن طريقة العرض باستخدام المستطيلات أو الأعمدة، وذلك برسم محورين أفقي لمسميات المتغير وعمودي للتكرار، وذلك باعتماد مقياس رسم مناسب .

والفرق يكمن أن الخطوط البيانية هي توصيل كل نقطتين متجاورتين بخط مستقيم بدلاً من استخدام الأعمدة أو المستطيلات .

ويمكن توضيح طريقة العرض هذه باعتماد مثال حول تزايد درجات التحصيل بازدياد عدد ساعات التدريب لدى عينة من التلاميذ.



الشكل (8-3)

الخط البياني لدرجات تحصيل التلاميذ وفق عدد ساعات التدريب

----- الفصل الثالث :أساليب تنظيم البيانات وعرضها

ج- التمثيل البياني لمتغيرات في مستوى الرتبة :

يتم بناء الجداول التكرارية للمتغيرات من مستوى الرتبة بنفس طريقة بناء جداول المتغيرات الأسمية. ويوضح الجدول (10-1) التوزيع التكراري لمتغير من مستوى الرتبة، إذ يبين التوزيع التكراري لأحد أسئلة مقياس من مقاييس الاتجاهات .

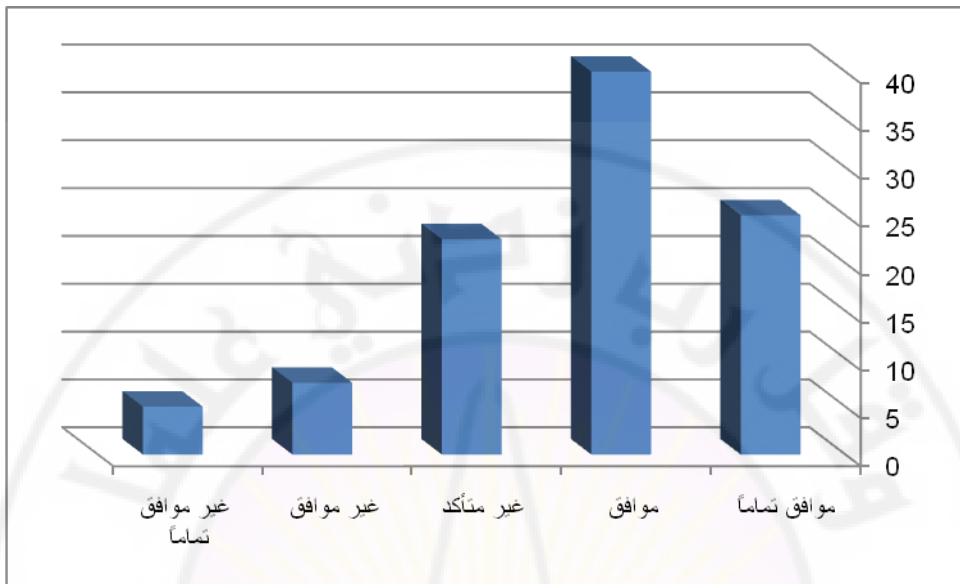
جدول (10-1)

التوزيع التكراري لأحد أسئلة الاتجاهات

الاستجابة	التكرار	النسبة المئوية%
(5) موافق تماماً	10	25
(4) موافق	16	40
(3) غير متأكد	9	22.5
(2) غير موافق	3	7.5
(1) غير موافق تماماً	2	5
المجموع	40	100

ويلاحظ أن وجود عمود للنسبة المئوية في جداول التوزيع التكراري للمتغيرات الأسمية والمتغيرات من مستوى الرتبة يساعد على توضيح البيانات وخاصة عندما تكون أعداد أفراد العينة كبيرة . ويتضح من الجدول السابق أن 65% من أفراد العينة وافقوا على مضمون السؤال المتعلق بالاتجاه والشكل البياني التالي يوضح ذلك.

الفصل الثالث: أساليب تنظيم البيانات وعرضها



الشكل (9-3): أعمدة بيانية

----- الفصل الثالث: أساليب تنظيم البيانات وعرضها

أسئلة وتمارين :

تمرين (1) :

فيما يلي بيانات لعينة مكونة من 100 فرد التي تمثل درجاتهم في اختبار نفسي والمطلوب:

1- تبويب المعطيات في جدول تكراري عدد فئاته 7.

2- إيجاد مراكز الفئات .

3- إيجاد التكرار المتجمع الصاعد والهابط.

4- استخراج التكرار المتجمع الصاعد النسبي والتكرار المتجمع الهابط النسبي.

5- عرض التكرارات بمضلع تكراري، مستخدماً قيم مراكز الفئات.

جدول (1-17)

درجات الأفراد

41	35	40	40	41	45	41
38	35	39	39	36	40	40
43	43	43	51	36	37	38
53	42	42	42	42	42	42
46	45	41	41	42	50	52
43	44	44	45	45	45	45
47	44	44	44	44	44	44
47	48	48	48	48	48	47
51	46	46	47	47	47	47
49	52	52	51	43	47	47
50	90	49	50	50	50	49
52	52	55	49	49	49	49
49	51	51	48	46	54	50
		53	53	53	55	48

----- الفصل الثالث :أساليب تنظيم البيانات وعرضها

تمرين (2) : في الجدول الأدنى التوزيع التكراري لعينة من الأفراد موزعة حسب فئات درجاتهم في مقياس التكيف النفسي، والمطلوب إيجاد مراكز الفئات وعرضها بيانياً مستخدماً المضلع التكراري :

جدول (3-11)

فئات الدرجات وتكراراتها

التكرار	فئات الدرجات
3	85-75
4	95-85
8	105-95
10	115-105
15	125-115
20	135-125
15	145-135
10	155-145
8	165-155
5	175-165
2	185-175

تمرين (3) : استخدم جدول معطيات التمرين (2) لإيجاد الدرجات التي تقل عن 105 درجة وتلك التي تزيد على 155 ليرة وذلك باستخدام التكرارات المتجمعة لهذا الغرض.

تمرين (4): الجدول التالي، يعطي عدد الطلاب المسجلين في الجامعة للفترة 2006-2000 مصنفة حسب الاختصاص، والمطلوب :

----- الفصل الثالث: أساليب تنظيم البيانات وعرضها

- 1- رسم المعطيات على شكل أعمدة بيانية .
- 2- استخدام الدائرة البيانية لعرض عدد الطلاب المسجلين في الجامعة للفترة 2000-2006

جدول (3-12)

السنة	حقوق	آداب	تربية	المجموع
2000	5575	6494	1848	13817
2001	5764	6757	2056	14477
2002	7338	12446	2216	22000
2003	9600	17073	2312	28985
2004	10311	16466	2508	29285
2005	9191	16899	2596	28686
2006	11472	17357	3159	31988

----- الفصل الثالث: أساليب تنظيم البيانات وعرضها

تمرين (5) :

لدينا جدول التوزيع التكراري التالي:

جدول (3-13)

جدول الفئات لعدد الطلاب

الفئات	f	x_i	$f \uparrow$	$f \downarrow$	$\%f \uparrow$	$\%f \downarrow$
51 55	1	53	1	25	4	100
55 59	1	57	2	24	8	96
59 63	2	61	4	23	16	92
63 67	2	65	6	21	24	84
67 71	4	69	10	19	40	76
71 75	3	77	13	15	52	60
75 79	3	81	16	12	64	48
79 83	4	85	20	9	80	36
83 87	1	89	21	5	84	20
87 91	2	93	23	4	92	16
91 95	1	97	24	2	96	8
95 103	1	101	25	1	100	4
	$\sum f=25$					

المصدر افتراضي

المطلوب:

1- ارسم مدرج تكراري

2- ارسم المضلع التكراري

3- احسب النسبة المئوية التي تقل عن الدرجة 75



الفصل الرابع

مقاييس النزعة المركزية

Measures of central tendency

مقدمة :

يساعد وضع البيانات في جدول تكراري في إعطاء فكرة أولية عن سير البيانات واتجاه الظاهرة مدار البحث، إلا أن أهم الخصائص التي يركز عليها الباحثون هي الوصول إلى قيمة تحدد الموقع العام في التوزيع التكراري للبيانات، الذي يحدد عدد الدرجات التي تقع أعلى أو أدنى من هذه الدرجة، فلو أجري اختبار لمجموعة من الطلاب فإن كل طالب سيقارن درجته بمعظم درجات الطلاب (وهي قيمة المنوال) ويقارن درجته بمتوسط معدل درجات زملائه، أو يحدد موقع درجته بين الدرجات وفي كل هذه الحالات يشير الباحث إلى ما يسمى النزعة المركزية "Central tendency" ذلك أن الدرجة الخام التي يحصل عليها الفرد ليس لها معنى إلا إذا كان هناك معيار محدد للمقارنة، مثال ذلك لو كانت درجة الطالب في امتحان ما تساوي (80) فحتى نستطيع تقييم درجة الطالب نحدد موقعه مقارنة بالآخرين فنقول إنه من المتفوقين إذ إن درجته أعلى من 85% من زملائه، أو أن درجته أعلى من متوسط درجات زملاؤه . وباختصار لابد من أن نبين شكل التوزيع حتى نصف توزيع معين من الدرجات، ومتوسطه وتشتته .

هناك عدة مقاييس للنزعة المركزية "Measures of central tendency" منها: المتوسط الحسابي، الوسيط، المنوال ويجب أن تتوفر في هذه المقاييس بعضاً من الصفات الآتية :

- أن يعتمد المقياس في حسابه على كل المشاهدات
- أن يكون المقياس سهل الحساب والفهم
- أن يتوفر فيه القابلية للتعامل الجبري
- ألا يتأثر بالقيم الشاذة أو المتطرفة

----- الفصل الخامس :مقاييس التشتت

ولا تتوافر كل الصفات السابق ذكرها في مقياس واحد ولكن كل مقياس من هذه المقاييس يفضل استخدامه في حالات معينة ولا يفضل استخدامه في حالات أخرى. (الصياد وربيع وسمره، 1429، ص 17) والجدول التالي يبين توفر الصفات السابقة في كل مقياس من مقاييس النزعة المركزية

جدول (1-4)

الصفات المتوفرة في كل مقياس من مقاييس النزعة المركزية

المقياس	المتوسط	الوسيط	المنوال	الخاصية
يعتمد في حسابه على كل المشاهدات	✓	×	×	
المقياس سهل الفهم والحساب	✓	✓	✓	
تتوفر فيه القابلية للتعامل الجبري	✓	✓	✓	
لا يتأثر بالقيم الشاذة أو المتطرفة	×	✓	✓	

وفيما يلي وصف وتوضيح لكل مقياس من هذه المقاييس :

1- المتوسط الحسابي Mean:

هو أكثر مقاييس النزعة المركزية استخداماً وهو يساوي مجموع القيم على عددها ويرمز له بـ $\bar{X}$ للعينة و M للمجتمع ويحسب بالمعادلة التالية:

$$\bar{X} = \frac{\sum x}{n}$$

حساب المتوسط الحسابي:

أولاً : حساب المتوسط الحسابي من الدرجات مباشرة للبيانات غير المبوبة :

يحسب المتوسط الحسابي من الدرجات مباشرة وذلك بقسمة مجموع هذه الدرجات على عددها. ولا يشترط أن يكون المتوسط الحسابي عدداً صحيحاً.

مثال : احسب المتوسط الحسابي لأعمار خمسة من الموظفين في إحدى دوائر الدولة
أعمار الموظفين :

40	28	25	35	30
----	----	----	----	----

$$\bar{X} = \frac{\sum x}{n}$$

$$\bar{X} = \frac{158}{5} = 31.6$$

ثانياً: حساب المتوسط من تكرار الدرجات غير المبوبة في فئات :

مثال : احسب المتوسط الحسابي لدرجات مجموعة من الطلبة الحاصلين على الشهادة الثانوية
كما هي مبينة في الجدول التالي:

جدول (2-4)

درجات مجموعة من الطلبة الحاصلين على الشهادة الثانوية

الدرجات X	التكرار f	الدرجة × التكرار fX
105	1	105
115	2	230
122	4	488
130	3	390
135	2	270
138	2	276
	$\Sigma f = 14$	1759

----- الفصل الخامس :مقاييس التشتت

حيث:

Σf : عدد الأفراد

Σfx : مجموع حاصل ضرب الدرجة بالتركرار .

لحساب المتوسط في هذه الحالة لابد من حساب عدد الأفراد Σf ، كذلك لابد من ضرب كل درجة بتركرارها، ثم جمع حاصل ضرب الدرجة في التكرار . وبهذا يحسب المتوسط وفق الصيغة التالية:

$$\bar{X} = \frac{\Sigma fx}{\Sigma f}$$

$$\bar{X} = \frac{\Sigma fx}{\Sigma f} = \frac{1759}{14} = 125.64$$

مثال :

فيما يلي توزيع درجات (15) طالباً في امتحان الإحصاء العلامة من (10) والمطلوب حساب المتوسط الحسابي.

الجدول (4-3)

توزيع درجات (15) تلميذاً في امتحان الرياضيات

X	f	$fX \uparrow$
3	3	9
5	7	35
8	4	32
4	1	4
	15	80

$$\bar{X} = \frac{\Sigma fX}{\Sigma f} = \frac{80}{15} = 5.33$$

----- الفصل الخامس :مقاييس التشتت

ثالثاً: حساب المتوسط من فئات القيم :

هناك ثلاث طرائق لحساب المتوسط للفئات هي:

1- : الأسلوب المباشر :

تقوم هذه الطريقة في جوهرها على فكرة النزعة المركزية والاعتماد على مركز الفئة باعتباره يعبر عن "متوسط" القيم التي تقع في نطاقها ويمكن استخدامه للتعبير عن مجموعة القيم في هذه الفئة، كما نستخدم المتوسط الحسابي للتعبير عن أي مجموعة من البيانات .

وهذا الأسلوب يقوم على تقدير متوسط للقيم من متوسط الفئات، وتزداد دقة هذا التقدير كلما كان مدى الفئات صغيراً، وتقل دقة هذا التقدير مع ازدياد مدى الفئات .

لحساب المتوسط الحسابي بهذه الطريقة يتم إتباع الخطوات التالية :

- 1- إيجاد مراكز الفئات .
- 2- ضرب مركز كل فئة بالتكرار المناظر .
- 3- إيجاد مجموع حاصل ضرب مراكز الفئات بالتكرارات المناظرة .
- 4- التعويض بالقانون :

$$\bar{X} = \frac{\sum(xif)}{\sum f}$$

حيث $\bar{X}$: المتوسط الحسابي

$\sum(xif)$: مجموع حاصل ضرب مركز الفئة × تكرارها

$\sum f$: مجموع التكرارات (حجم العينة)

مثال : طبق أحد المدرسين اختباراً تحصيلياً على 50 طالباً فحصل على البيانات التالية علماً أن القصوى هي (50) درجة:

----- الفصل الخامس :مقاييس التشتت

جدول (4-4)

التوزيع التكراري لفئات درجات الطلاب في الاختبار التحصيلي

الفئات	التكرار f	مركز الفئة X_i	مركز الفئة × التكرار fX_i
10-14	5	12	60
14-18	6	16	96
18-22	5	20	100
22-26	7	24	168
26-30	12	28	336
30-34	6	32	192
34-38	4	36	144
38-42	3	40	120
42-46	1	44	44
46-50	1	48	48
المجموع	$\Sigma f = 50$		$\Sigma X_i = 1308f$

$$\bar{X} = \frac{\Sigma(xif)}{\Sigma f} = \frac{1308}{50} = 26.16$$

2-حساب المتوسط بأسلوب الانحرافات :

خطوات الحل :

1- إيجاد مركز الفئات .

2- نختار مركز الفئة التي يقابلها أكبر تكرار والمتوسط الفرضي $\bar{X}_0$.

3- نحسب الانحراف:

$$d = X_i - \bar{X}_0$$

4- ضرب كل انحراف بالتكرار المقابل df . ثم نحصل على Σdf

----- الفصل الخامس: مقاييس التشتت

5- التعويض في القانون : $\bar{X} = \bar{X}_0 + \frac{\sum df}{\sum f}$

مثال: احسب المتوسط الحسابي بطريقة الانحرافات للبيانات السابقة.

الجدول (4-5)

التوزيع التكراري لفئات درجات الطلاب في الاختبار التحصيلي

الفئات X	f	X_i	d	Df
10-14	5	12	-16	-80
14-18	6	16	-12	-72
18-22	5	20	-8	-40
22-26	7	24	-4	-28
26-30	12	28	0	0
30-34	6	32	+4	+24
34-38	4	36	+8	+32
38-42	3	40	+12	+36
42-46	1	44	+16	+16
46-50	1	48	+20	+20
	$=50\sum f$			-220
				<u>+128</u>
				-92

$$\bar{X} = \bar{X}_0 + \frac{\sum df}{\sum f}$$

$$\bar{X} = 28 + \frac{-92}{50}$$

$$\bar{X} = 28 - 1.84 = 26.16$$

----- الفصل الخامس :مقاييس التشتت

3- حساب المتوسط الحسابي للفئات بأسلوب الانحرافات المختصرة :

خطوات الحل :

- 1- إيجاد مراكز الفئات .
 - 2- نختار مركز الفئة التي يقابلها أكبر تكرار ويسمى X_0 .
 - 3- نحسب الانحراف المختصر بتقسيم الانحراف على طول الفئة أي $d' = \frac{d}{c}$
 - 4- ضرب كل انحراف مختصر بالتكرار المقابل أي Fd' ، ثم نحصل على مجموع Fd'
 - 5- التعويض في القانون $X = X_0 + \frac{\sum d'f}{\sum f} \times C$
- مثال: احسب المتوسط الحسابي بطريقة الانحرافات المختصرة للبيانات السابقة.

الجدول (4-6)

التوزيع التكراري لفئات درجات الطلاب في الاختبار التحصيلي

الفئات X	f	X_i	d	d'	Fd'
10-14	5	12	-16	-4	-20
14-18	6	16	-12	-3	-18
18-22	5	20	-8	-2	-10
22-26	7	24	-4	-1	-7
26-30	12	28	0	0	0
30-34	6	32	+4	+1	+6
34-38	4	36	+8	+2	+8
38-42	3	40	+12	+3	+9
42-46	1	44	+16	+4	+4
46-50	1	48	+20	+5	+5
	$=50\sum f$				-55
					<u>+32</u>
					-23

$$\begin{aligned}\bar{X} &= \bar{X}_0 + \frac{\Sigma(Fd)}{\Sigma f} \times C \\ &= 28 + \frac{-23}{50} \times 4 \\ X &= 28 - 1.84 \\ &= 26.16\end{aligned}$$

يلاحظ أن قيمة المتوسط هي نفسها في الأساليب الثلاث.

ولكل من الأساليب الثلاثة السابقة مزاياها واستخدامها :فالأسلوب المباشر يفضل استخدامه عندما تكون قيم التكرارات وقيم النزعة المركزية صغيراً ويمتاز هذا الأسلوب بسهولة إلا أنه يتأثر بالتقريب الذي ينشأ من تلخيص جميع درجات كل فئة في منتصفها.

وفيد أسلوب الانحرافات في حال وجود بيانات كبيرة القيم، فاستخدام الأسلوب المباشر سوف يكون صعباً و عرضة للخطأ، فعليه يمكن استخدام أسلوب الانحرافات بغرض تصغير القيم التي لدينا حتى يسهل التعامل معها.

أما أسلوب الانحرافات المختصرة فيهدف إلى تبسيط العمليات الحسابية الطويلة التي ظهرت في أسلوب الانحرافات ، إلا أنه لا يستخدم إلا إذا كانت أطوال الفئات متساوية، في حين يمكن استخدام أسلوب الانحرافات سواء كانت الأطوال متساوية أم غير متساوية.

----- الفصل الخامس :مقاييس التشتت

مثال:

احسب المتوسط الحسابي لدرجات لطلاب في اختبار للذكاء بالطرائق الثلاث.

الجدول (4-7)

التوزيع التكراري لفئات درجات الطلاب في اختبار للذكاء

الفئات	f	xi	xif	d	df	d'	$d'f$
30-34	4	32	128	-8	-32	-2	-8
34-38	5	36	180	-4	-20	-1	-5
38-42	10	40	400	0	0	0	0
42-46	7	44	308	+4	+28	+1	+7
46-50	6	48	288	+8	+48	+2	+12
50-54	5	52	260	+12	+60	+3	+15
	$\Sigma f =$ 37		$= \Sigma(xif)$ 1564		-52 <u>+136</u> =+84	-3 <u>+6</u> =+3	-13 <u>+34</u> +21

1- الطريقة المباشرة:

$$\bar{X} = \frac{\Sigma(xif)}{\Sigma f}$$

$$\bar{X} = \frac{1564}{37}$$

$$=42.27$$

2- طريقة الانحرافات:

$$\bar{X} = \bar{X}_0 + \frac{\Sigma df}{\Sigma f}$$

----- الفصل الخامس: مقاييس التشتت

$$=40 + \frac{84}{37}$$

$$=40+2.27=42.27$$

3- طريقة الانحرافات المختصرة:

$$\bar{X} = \bar{X}_0 + \frac{\Sigma(Fd')}{\Sigma f} \times C$$

$$\bar{X} = 40 + \frac{(21)}{37} \times 4$$

$$=40 + 0.567 \times 4$$

$$=40+2.027=42.27$$

رابعاً: حساب المتوسط في حالة القيم المنفصلة:

لا تختلف طريقة حساب المتوسط الحسابي في هذه الحالة عنها في حالة القيم المتصلة، إلا في عدم وجود الفئات، وعلى ذلك نتخذ القيمة المعطاة لنا بدلاً من مركز الفئة.

مثال :

الجدول التالي يوضح لنا توزيع عدد الأبناء في 100 عائلة.

والمطلوب حساب متوسط عدد الأبناء .

نفرض أن المتوسط الفرضي = 4

الجدول (4-8)

جدول بعدد أبناء عدد من العائلات

عدد الأبناء	عدد العائلات f	d	fd
0	3	-4	-12
1	7	-3	-21
2	11	-2	-22
3	14	-1	-14
$X_0 = 4$	20	صفر	صفر
5	16	+1	+16
6	12	+2	+24
7	7	+3	+21
8	5	+4	+20
9	3	+5	+15
10	2	+6	+12
	$\Sigma = 100$		+108
			<u>-69</u>
			39

$$\bar{x} = x_0 + \frac{\Sigma df}{\Sigma f}$$

$$\bar{x} = 4 + \frac{39}{100} = 4.39$$

المتوسط الموزون *Weighted Mean*:

إذا أخذنا عدة عينات من مجتمع معين، فليس بالضرورة أن تكون متوسطات تلك العينات متطابقة، ولاحتمساب المتوسط لتلك العينات فإنه لابد من مراعاة المتوسط الحسابي لكل عينة وحجمها، وهو ما يسمى بالمتوسط المرجح أو الموزون، فإذا كانت $\bar{x}_1$ ، $\bar{x}_2$ ، ، $\bar{x}_m$ هي المتوسطات الحسابية لـ n من العينات حجم كل منها على الترتيب n_1 ، n_2 ، ، n_m فإن المتوسط الحسابي الموزون يحسب من خلال المعادلة :

$$\frac{\bar{x}_1.n_1 + \bar{x}_2.n_2 + \dots + \bar{x}_m.n_m}{n_1 + n_2 + \dots + n_m}$$

حقيقة ما نقوم به هو إيجاد المجموع الكلي لقيم المشاهدات في العينات جميعها ، وتصبح كأنها عينة واحدة حجمها يساوي مجموع أحجام العينات.

مثال : احسب المتوسط الموزون للجدول التالي

الجدول (9-4)

رقم العينة	متوسط العينة $\bar{x}$	حجم العينة n	$\sum_{m=1}^n x_m \times n_m$
1	30	15	450
2	40	30	1200
3	20	5	100
المجموع	-	50	1750

$$\text{المتوسط الموزون} = \frac{1750}{50} = \frac{5 \times 20 + 30 \times 40 + 15 \times 30}{5 + 30 + 15} = 35$$

لاحظ أننا لو حسبنا المتوسط الحسابي للأوساط لوجدناه =30، وهو يختلف عن المتوسط الموزون، وهذا يؤكد أن الخطأ يكون كبيراً ، إذا تجاهلنا أسلوب المتوسط الموزون. وهذا ما نحتاج إليه لحساب المتوسط الحسابي لعلامات طالب خريج جامعة يدرس بنظام الساعات المعتمدة فلكل مقرر عدد من الساعات المعتمدة، ولحساب المتوسط الحسابي لعلامات الطالب في فصل أو فصلين أو خلال فترة الدراسة يتم ضرب عدد ساعات كل مقرر في علامته. ويقسم

----- الفصل الخامس :مقاييس التشتت

مجموع الناتج على مجموع الساعات المعتمدة في المقررات جميعها، فنحصل على المتوسط الموزون الذي يسمى (المعدل التراكمي) للطالب.

خواص المتوسط الحسابي:

1- يستخدم المتوسط للبيانات الكمية (فئوية أو نسبية) ولا يستخدم للبيانات الوصفية (اسمية، ترتيبية)

2- قابل للعمليات الجبرية، ويدخل في حساب مقاييس أخرى كثيرة .

3- يعطى قيمة وحيدة لمجموعة القيم.

4- مجموع انحرافات القيم عن متوسطها الحسابي يساوي صفراً .

وللتحقق من هذه الخاصية نأخذ المثال التالي :

مثال : ليكن لدينا البيانات التالية :

3	4	6	8	9
---	---	---	---	---

خطوات الحل:

1- نحسب المتوسط كما يلي:

$$\bar{X} = \frac{30}{5} = 6$$

2- نحسب انحراف كل درجة عن متوسطها ثم نحسب مجموع الانحراف كما يلي:

الجدول (4-10)

X	d
3	$3-6=-3$
4	$4-6=-2$
6	$6-6=0$
8	$8-6=2$
9	$9-6=3$
	-5
	<u>+5</u>
	=0

----- الفصل الخامس :مقاييس التشتت

يلاحظ أن : مجموع الانحرافات يساوي صفراً.

5- يتأثر المتوسط بالقيم المتطرفة أو الشاذة، ويقال بأن مفردة ما متطرفة إذا كانت كبيرة جداً أو صغيرة جداً مقارنة ببقية المفردات .

مثال : ليكن لدينا القيم التالية والتي تمثل درجات خمس طلاب في مادة التاريخ

6	7	8	10	12
---	---	---	----	----

$$\bar{X} = \frac{43}{5} = 8.6$$

مثال : وليكن لدينا القيم التالية التي تمثل درجات خمس طلاب في مادة الجغرافية

6	7	8	10	25
---	---	---	----	----

$$\bar{X} = \frac{56}{5} = 11.2$$

يلاحظ أن المتوسط في مادة الجغرافية قد ازداد زيادة واضحة عن متوسط درجات الطلاب في مادة التاريخ نظراً لوجود قيمة متطرفة في الجغرافية وهي الدرجة 25، في حين كانت باقي الدرجات متطابقة، لهذا في حال وجود قيم متطرفة يفضل استخدام الوسيط .

6- لا يمكن حساب المتوسط في حال الجداول ذات الفئات الحدية المفتوحة، والفئة المفتوحة هي الفئة التي لا يعرف أحد حديها (الأعلى أو الأدنى).

الوسيط Median

المقياس الثاني من مقاييس النزعة المركزية هو الوسيط.و يعرف الوسيط بأنه: تلك القيمة التي تتوسط توزيع مجموعة قيم مرتبة تصاعدياً بحيث يسبقها نصف القيم ويلها النصف الآخر، أو القيمة التي يقع تحتها أو يقل عنها (50%) من القيم، وذلك بعد أن يتم ترتيبها تصاعدياً أو تنازلياً . يعتمد حساب الوسيط على عدد الدرجات من حيث نوعها سواء كان فردياً أم زوجياً .

ويحسب الوسيط للبيانات الرتبية لذلك يناسب المستوى الترتبي من مستويات القياس.

حساب الوسيط للبيانات غير المبوبة :

أ- حساب الوسيط عندما يكون عدد القيم فردياً :

مثال : لدينا القيم التالية :

10	8	12	11	14	13	7
----	---	----	----	----	----	---

خطوات الحل:

1- ترتب القيم ترتيباً تصاعدياً كما يلي :

7	8	10	11	12	13	14
---	---	----	----	----	----	----

2- نبحث بعد ذلك عن النقطة التي تتصف هذه الدرجات فنرى أنها تقع تماماً عند

الدرجة 11 لأن عدد الدرجات التي تسبقها 3 وهي 7،8،10 وعدد الدرجات التي تليها 3 وهي

12،13،14

3- يمكن أن نحدد ترتيب الوسيط من خلال حساب رتبة الوسيط وذلك بقسمة عدد

الدرجات على 2

$$\text{أي رتبة الوسيط} = \frac{\text{عدد الدرجات}}{2}$$

ويرمز لعدد الدرجات بـ n

----- الفصل الخامس :مقاييس التشتت

$$\frac{n}{2} = \text{أي رتبة الوسيط}$$

$$\frac{7}{2} = 3.5 \text{ وفي المثال السابق نقسم}$$

ثم نقرب الناتج إلى أقرب رقم صحيح وهو (4) ثم نبحث عن الدرجة التي تقابل هذا الترتيب. وبما أننا في هذه الحالة اضطررنا إلى تقريب العدد إلى أقرب رقم صحيح فإنه بالإمكان الاستغناء عن هذا التقريب بإضافة واحد صحيح إلى عدد الدرجات حتى يصبح زوجياً ويصبح الناتج عدداً صحيحاً .

$$\frac{\text{عدد الدرجات} + 1}{2} = \text{أي أن ترتيب الوسيط}$$

$$\frac{n+1}{2} = \frac{7+1}{2} = 4$$

والدرجة الرابعة هي 11.

ونلخص الخطوات السابقة كما يلي :

- نرتب الدرجات ترتيباً تصاعدياً .

- عدد الدرجات 7

$$4 = \frac{7+1}{2} = \text{ترتيب الوسيط}$$

- إذن الدرجة الوسطى هي 11

ب- حساب الوسيط عندما يكون عدد القيم زوجياً

احسب الوسيط للدرجات التالية :

20	13	10	8	5	4	3	7
----	----	----	---	---	---	---	---

خطوات الحل:

1- نرتب الدرجات تصاعدياً :

20	13	10	8	7	5	4	3
----	----	----	---	---	---	---	---

----- الفصل الخامس :مقاييس التشتت

2- نحسب رتبة الوسيط أي $\frac{n}{2}$

$$\frac{8}{2} = 4$$

الدرجة التي ترتيبها الرابع من الطرف الأول لتدريج الدرجات هي 7.

الدرجة التي ترتيبها الرابع من الطرف الثاني لتدريج الدرجات هي 8.

يكون الوسيط هو متوسط 7، 8

$$med = \frac{7+8}{2} = 7.5 \text{ أي}$$

2- حساب الوسيط من فئات القيم :

أ- حساب الوسيط من الفئات المغلقة:

مثال: ليكن لدينا بيانات لدرجات مجموعة مؤلفة من 28 طفلاً في اختبار ريفن للذكاء

الجدول (4-11)

درجات الأطفال في اختبار ريفن

فئات القيم	f	$f \uparrow$
17-20	2	2
20-23	3	5
23-26	5	10
26-29	7	17
29-32	6	23
32-35	5	28
	28	

الحل:

لحساب الوسيط من الفئات نتبع الخطوات التالية:

1- نحسب التكرار المتجمع الصاعد.

----- الفصل الخامس :مقاييس التشتت

2- نحسب ترتيب الوسيط وهو يساوي $\frac{\text{مجموع التكرارات}}{2}$ أي $\frac{\Sigma F}{2}$ وفي مثالنا يساوي أي

$$\frac{28}{2} = 14$$

3- نبحث في حقل التكرار التجميحي الصاعد عن قيمة ترتيب الوسيط فإذا لم نجدها نأخذ الأكبر منها مباشرة وهي هنا 17 فنجد أن فئة الوسيط هي الفئة (26-29) المقابلة للتكرار التجميحي 17 لأن التكرار التجميحي للفئة التي تسبقه 10 وهو أقل من 14.

4- نحسب الوسيط وفق المعادلة التالية

$$\text{الوسيط} = \text{الحد الأدنى لفئة الوسيط} + \left(\frac{\text{رتبة الوسيط} - \text{التكرار التجميحي لفئة ما قبل الوسيط}}{\text{تكرار فئة الوسيط}} \right) \times \text{طول فئة الوسيط}$$

$$\text{أي أن } Med = Lme + \left(\frac{\frac{\Sigma F}{2} - F}{fme} \right) \times Cme$$

حيث Lme الحد الأدنى لفئة الوسيط

Σf مجموع التكرارات (عدد الأفراد)

f التكرار التجميحي الصاعد لفئة ما قبل الوسيط

mef : تكرار فئة الوسيط

C : طول فئة الوسيط

وبتطبيق هذه المعادلة نحصل على

$$Med = 26 + \frac{14-10}{7} \times 3$$

$$Med = 26 + \frac{4}{7} \times 3$$

$$Med = 26 + \frac{12}{7}$$

$$Med = 26 + 1.71$$

$$Med = 27.71$$

نظراً لأن الوسيط لا يتأثر بالفئات المفتوحة فإنه يمكن أن نحسب الوسيط من الفئات المفتوحة.

----- الفصل الخامس :مقاييس التشتت

ب: حساب الوسيط من فئات مفتوحة:

مثال: لدينا مجموعة من القيم التي تمثل أسعار بعض السلع.

الجدول (4-12)

أسعار بعض السلع

فئات القيم	f	$f \uparrow$
20	3	3
20-23	5	8
23-26	4	12
26-29	6	18
29-32	4	22
32-35	3	25
المجموع	25	

الحل:

رتبة الوسيط :

$$\frac{\Sigma F}{2} = \frac{25}{2} = 12.5$$

قيمة الوسيط:

$$\text{Med} = \text{Lme} + \left(\frac{\frac{\Sigma F}{2} - F_{\uparrow}}{F_{me}} \right) \times C_{me}$$

----- الفصل الخامس: مقاييس التشتت

$$\begin{aligned} &= 26 + \frac{12.5 - 12}{6} \times 3 \\ &= 26 + \frac{0.5 \times 3}{6} \\ &= 26 + 0.25 \\ &= 26.25 \end{aligned}$$

خواص الوسيط:

- 1- يفضل استخدام الوسيط في حالة المتغيرات الترتيبية .
- 2- لا يتأثر الوسيط بالقيم المتطرفة أو الشاذة لذلك يفضل استخدامه في وصف المتغيرات الكمية في حال وجود مثل هذه القيم، كما هي الحال مع بيانات الدخل
- 3- يستخدم في حال وجود الفئات المفتوحة .
- يتحدد بعدد البيانات وليس بقيمتها .
- يستخدم في حال الفئات غير المتساوية في أطوالها . (فهمي، 2005، ص180)

المنوال Mode

يعرّف المنوال أنه: القيمة الأكثر تكراراً أو الأكثر شيوعاً بين القيم، أي القيمة التي

تتكرر أكثر من غيرها . ويرمز له بـ Mod

يستخدم المنوال في جميع مستويات القياس (الاسمي الترتيبي، الفئوي، النسبي)

1- حساب المنوال للبيانات غير المبوبة:

مثال : لدينا القيم التالية التي تمثل أعمار مجموعة من الموظفين

31 30 28 28 28 26 25 26

يلاحظ أن العمر الأكثر تكراراً هو 28، بمعنى أن غالبية الموظفين هم من عمر 28

3- حساب المنوال من فئات القيم :

هناك أساليب عديدة لحساب المنوال وسوف نقتصر هنا على أسلوب الفروق.

أ- حساب المنوال من فئات مغلقة:

مثال: ليكن لدينا فئات القيم التالية والتي تمثل درجات الأفراد في مقياس لتقدير الذات والمطلوب حساب منوال هذه القيم.

الجدول (4-13)

فئات الدرجات في مقياس تقدير الذات

فئات القيم	f
20-24	3
24-28	6
28-32	7
32-36	10
36-42	4
42-46	2
	32

----- الفصل الخامس :مقاييس التشتت

أولاً : نحدد فئة المنوال وهي في مثالنا الفئة (32-36) المقابلة لأعلى تكرار وهي 10

نحسب المنوال وفق المعادلة التالية:

$$Mo = Lmo + \frac{\Delta 1}{\Delta 1 + \Delta 2} \times C_{m0}$$

حيث Mo : المنوال

Lmo : الحد الأدنى لفئة المنوال

$\Delta 1$: تكرار فئة المنوال - تكرار الفئة التي قبلها أي $10 - 7 = 3$

$\Delta 2$: تكرار فئة المنوال - تكرار الفئة التي تليها أي $10 - 4 = 6$

نعوض بالقيم :

$$Mo = 32 + \frac{10 - 7}{(10 - 7) + (10 - 4)} \times 4$$

$$= 32 + \frac{3}{3 + 6} \times 4$$

$$= 32 + \frac{12}{9}$$

$$= 32 + 1.33$$

$$= 33.33$$

ب- حساب المنوال من فئات مفتوحة :

مثال: لدينا الفئات التالية :

الجدول (4-14)

الفئات	f
→18	8
18-22	12
22-30	20
30- 34	14
34- 40	10
40→	6
	$\Sigma f=70$

والمطلوب : حساب المنوال:

نحسب المنوال وفق المعادلة التالية:

$$Mo = L_{mo} + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times C_{mo}$$

$$= 22 + \frac{8}{8 + 6} \times 8 = 22 + \frac{64}{14} = 26.57$$

خواص المنوال :

1- يستخدم المنوال في جميع أنواع المتغيرات سواء كانت كمية أو نوعية، وهو المقياس

الوحيد المستخدم في وصف البيانات التي يمكن قياسها على المستوى الاسمي (فهومي،

2000، ص 180)

2- يفضل استخدام المنوال في حال وجود فئات مفتوحة .

3- لا يتأثر المنوال بالقيم المتطرفة .

- 4- يمتاز بسهولة حسابه .
- 5- قد يوجد أكثر من منوال في مجموعة القيم، إذا تكرر بعض الفئات بالدرجة نفسها. فيسمى عندها بالتوزيع متعدد المنوال .

فوائد مقاييس النزعة المركزية :

تتلخص أهم الفوائد التطبيقية لمقاييس النزعة المركزية في :

- 1- المعايير : تعتمد المعايير على المتوسط، ولهذا ينسب طول الفرد مع متوسط طول من هم في عمره .
- كذلك يقاس ذكاء الفرد بالنسبة لمتوسط ذكاء أقرانه وبمدى انحرافه عن هذا المعيار زيادة أو نقصاناً .
- 2- المقارنة : تستخدم المتوسطات لمقارنة مجموعة من الأفراد بمجموعة أخرى، من مثل مقارنة متوسط درجات الطلاب في مادة ما بمتوسط درجات نظرائهم في صف آخر .
- على أن تتم المقارنة للمجموعات المتجانسة من حيث بعض الصفات من مثل: المهنة، العمر، الصف الدراسي، حتى تصبح نتائجها صحيحة . (البهي السيد، 1978، ص65-66).
- هذا ويصلح الوسيط والمنوال للأغراض نفسها التي يصلح لها المتوسط ويجري استعمال الوسيط بدلاً عن المتوسط حال البيانات الكمية الترتيبية، أما المنوال فيستعمل بدلاً من المتوسط في البيانات الاسمية .
- وفيما يلي جدول يبين حالات استخدام كل من المتوسط والوسيط والمنوال:

الجدول (15-4)

حالات استخدام كل من المتوسط والوسيط والمنوال

المنوال	الوسيط	المتوسط
1- يستخدم إذا كان التوزيع متماثلاً تقريباً . 2- إذا كان الاهتمام منصّباً على القيمة العددية لجميع البيانات أي المجموع الكلي.	- يستخدم إذا كان التوزيع ملتوياً . - يستخدم في حال الرغبة في إيجاد قيمة ممثلة، أي قيمة نموذجية بدلاً من الاهتمام بالمجموع الكلي . - في حال فقدان بعض القيم وعرف ترتيبها، بحيث لا يمكن حساب المتوسط مباشرة.	- يستخدم إذا كان التوزيع متماثلاً تقريباً . 2- إذا كان الاهتمام منصّباً على القيمة العددية لجميع البيانات أي المجموع الكلي.

(أبو صالح، 57، ص2008)

- مقاييس النزعة المركزية ومستويات القياس:

إن تحديد مستويات القياس للبحث يعد أمراً هاماً للباحث الذي يقوم بتنفيذه ليتسنى له اختيار الأسلوب الإحصائي المناسب لوصف المتغير من جهة، وللإجابة عن أسئلة بحثه من جهة أخرى .

وبيّن الجدول التالي مدى ملائمة كل من مقاييس النزعة المركزية لكل مستوى من مستويات القياس المختلفة.

الجدول (4-16)

مقاييس النزعة المركزية ومستويات القياس

	الاسمي	الترتيبي	الفئوي	النسبي
المتوسط			*****	*****
الوسيط		*****	*****	*****
المنوال	*****	*****	*****	*****

- وتجدر الملاحظة أن المنوال هو المقياس الوحيد المستخدم في جميع مستويات القياس، في حين أن المستويين الفئوي والنسبي يمكن وصفهما بجميع مقاييس النزعة المركزية.

حساب مقاييس النزعة المركزية من علاقة بيرسون:

- تمكن الإحصائيون من إيجاد علاقة تقريبية بين المتوسطات الثلاثة: المتوسط والوسيط والمنوال. وتستخدم هذه العلاقة عندما يتعذر استخراج إحداها، كما يحدث عندما يراد إيجاد المتوسط الحسابي في جدول تكراري مفتوح. وهذه العلاقة تسمى علاقة بيرسون ويمكن توضيحها على النحو التالي:

المتوسط الحسابي - المنوال = 3 (المتوسط - الوسيط)

$$\bar{X} - \text{Mod} = 3(\bar{X} - \text{Med})$$

أي أن الفرق بين المتوسط الحسابي والمنوال يعادل ثلاثة أمثال الفرق بين المتوسط الحسابي والوسيط، ويمكن من هذه العلاقة الحصول على أي قيمة من هذه المتوسطات إذا عرف اثنان.

$$\text{المنوال} = (3 \times \text{الوسيط}) - (2 \times \text{المتوسط الحسابي})$$

$$\text{Mod} = 3\text{Med} - 2\bar{X}$$

----- الفصل الخامس :مقاييس التشتت

$$\frac{(3 \text{ الوسيط}) - (\text{المنوال})}{2} = \text{المتوسط الحسابي}$$

$$\bar{X} = \frac{3Med - Mod}{2}$$

$$\frac{(\text{المنوال}) + (2 \times \text{المتوسط})}{3} = \text{الوسيط}$$

$$Med = \frac{Mod + 2\bar{X}}{3}$$

مثال : احسب المتوسط الحسابي لتوزيع يمثل درجات مجموعة من الأفراد في اختبار التكيف

الاجتماعي علماً أن وسيط درجاتهم 35 والمنوال 34.5

الحل: المتوسط الحسابي وفق علاقة بيرسون يساوي :

$$\frac{(3 \text{ الوسيط}) - (\text{المنوال})}{2} = \text{المتوسط الحسابي}$$

$$\bar{X} = \frac{(3 \times 35) - (34.5)}{2} = 35.25$$

وللتأكد من صحة الناتج يمكننا حساب قيمة المنوال من قيمة المتوسط والوسيط

$$\text{المنوال} = (3 \times \text{الوسيط}) - (2 \times \text{المتوسط الحسابي})$$

$$M_0 = (3 \times 35) - (2 \times 35.25) = 34.5$$

وهي قيمة المنوال نفسها المعطاة في المثال.

العلاقة بين المتوسط الحسابي والوسيط والمنوال:

تظهر العلاقة بين مقاييس النزعة المركزية بأشكال عديدة :

1- أن يكون التوزيع التكراري متماثلاً وله قيمة واحدة وشكله يشبه شكل الجرس (التوزيع

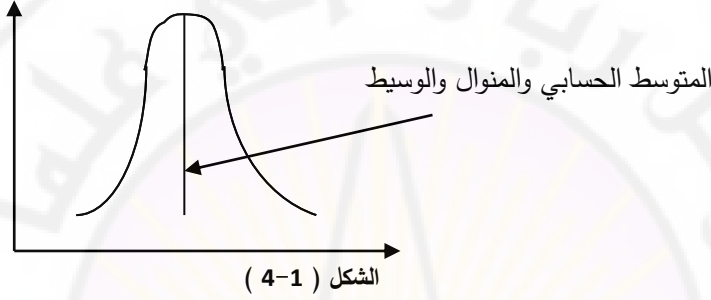
الطبيعي أو الاعتدالي). ويعد التوزيع متماثلاً إذا أمكننا إقامة عمود على المحور الأفقي

بحيث يقسم هذا العمود التوزيع إلى قسمين ينطبقان على بعضهما تمام الانطباق. وفي

----- الفصل الخامس :مقاييس التشتت

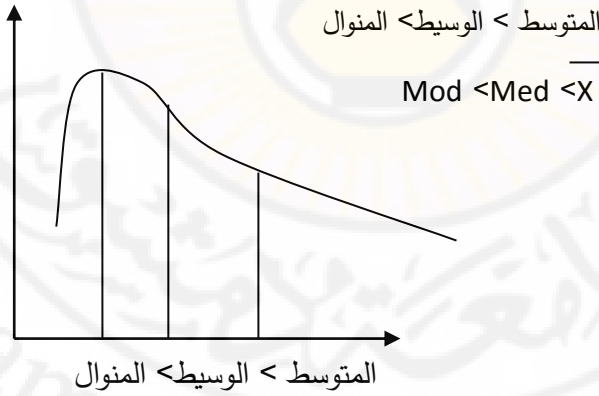
الحياة العملية يوجد عدد قليل من التوزيعات المتماثلة، ولكن يوجد كثير من التوزيعات تقريباً متماثلة. ومثالاً على توزيع متماثل درجات مجموعة من الأفراد في اختبار للذكاء . وفي هذه الحالة تكون العلاقة بين مقاييس النزعة المركزية على النحو التالي:

$$\overline{X} = \text{Med} = \text{Mod}$$



توزيع تكراري متماثل (المتوسط الحسابي = الوسيط = المنوال)

2- عندما يكون المتوسط أكبر من الوسيط والمنوال، فإن المنحنى التكراري سيكون ملتوياً ناحية اليمين ويسمى المنحنى منحنى ملتوي التواء موجباً كما يتضح من الشكل التالي .



الشكل (4-2)

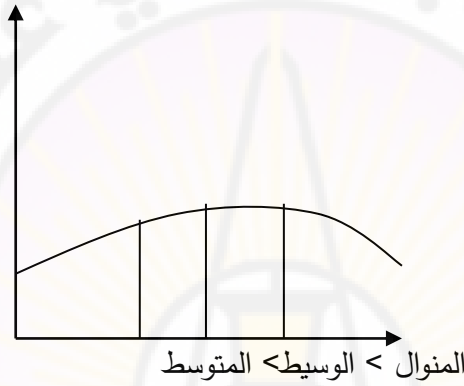
توزيع ملتوٍ ناحية اليمين

----- الفصل الخامس :مقاييس التشتت

3- عندما يكون المتوسط أصغر من الوسيط والمنوال، فإن المنحنى التكراري سيكون ملتوياً ناحية اليسار ويسمى المنحني ملتوي التواءً سالباً كما يتضح من الشكل التالي .

المنوال < الوسيط < المتوسط

$$\text{Mod} > \text{Med} > \bar{X}$$



الشكل (3-4) :توزيع ملتوٍ ناحية اليسار

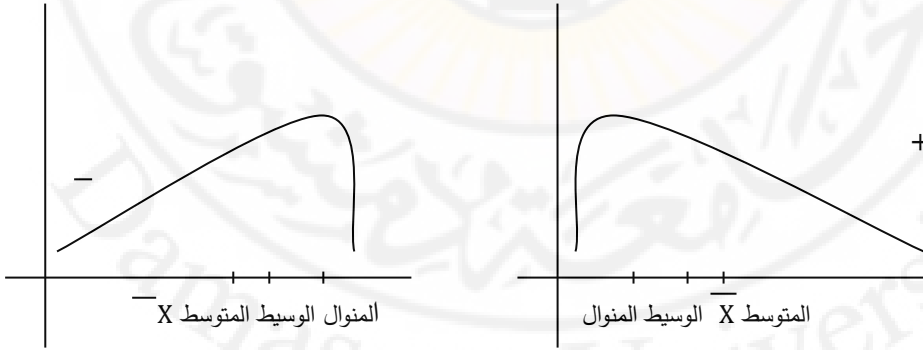
والالتواء يكون ناتجاً من صفة حقيقية في المجتمع الذي يجري عليه البحث، كما في حالة النتائج الدراسية للصفوف الضعيفة أو الفصول القوية، حيث يكون الالتواء موجباً في الحالة الأولى و سالباً في الحالة الثانية، أو راجعاً إلى سوء اختيار العينة، بحيث لو أحسن الباحث اختيار هذه العينة التي يجري عليها البحث لزال الالتواء أو خفت حدته، أو سوء الطريقة المستخدمة في القياس كما هي الحال في تطبيق اختبار أعلى أو أقل في مستواه عن مستوى العينة المختبرة، فالاختبار الصعب يعطي توزيعاً موجب الالتواء، بينما يعطي الاختبار السهل توزيعاً سالب الالتواء.(خيرى، 1997، ص34)

مقاييس الالتواء Coefficients of skewness:

تعني مقاييس الالتواء انحراف التوزيعات التكرارية عن التماثل، ولهذا تقسم التوزيعات التكرارية إلى عدة أشكال منها:

- 1- التوزيعات الملتوية نحو اليمين أو موجبة الالتواء إذا كان ممتداً أكثر نحو اليمين.
- 2- التوزيعات الملتوية نحو اليسار أو سالبة الالتواء إذا كان له طرف واحد ممتد نحو اليسار .
عندما يكون التوزيع ملتوياً نحو اليمين تؤثر على المتوسط الحسابي وتسحبه نحو اليمين وبذلك يكون المتوسط الحسابي أكبر من الوسيط أما إذا كان التوزيع ملتوياً نحو اليسار فإن المتوسط الحسابي يكون أصغر من الوسيط .
- 3- توزيعات معتدلة أو متماثلة. أي لا التواء فيها عندها يكون المتوسط الحسابي مساوياً للوسيط مساوياً للمنوال.

والأشكال التالية تمثل نماذج من هذه التوزيعات:



ويمكن قياس الالتواء لمجموعة من البيانات لتوزيع تكراري بالمعادلة التالية:

$$Y_I = \frac{3(\bar{x} - Med)}{S}$$

Y_I : معامل الالتواء

$\bar{x}$: المتوسط الحسابي

Med : الوسيط

S : الانحراف المعياري

ويستفاد من مقياس الالتواء في معرفة نوعية الالتواء للتوزيع التكراري، فإذا كان الالتواء موجباً فهذا يعني أن المتوسط الحسابي أكبر من الوسيط وأن للتوزيع طرفاً ممتداً إلى اليمين أكثر من الطرف الثاني الممتد إلى اليسار، أما إذا كان مقياس الالتواء سالباً فهذا يعني أن الالتواء نحو اليسار والطرف الأيسر هو الأكثر امتداداً .

وهناك فائدة ثانية وهي للمقارنة بين توزيعين تكراريين أو مجموعتين من البيانات، فالمجموعة التي تكون القيمة المطلقة لمقياس الالتواء لها أكبر يكون توزيعها ملتوياً أكثر من توزيع المجموعة الأخرى مع الأخذ بعين الاعتبار أن إشارة مقياس الالتواء تعين اتجاه الالتواء . فمثلاً إذا كان مقياس الالتواء لمجموعة من البيانات (أ) هو 0.57 ولمجموعة أخرى (ب) هو -0.63 فهذا يعني أن التواء المجموعة (ب) أكبر من التواء المجموعة (أ) لأن $|0.57| < |0.63|$ بينما التواء (أ) لليمين والتواء (ب) إلى اليسار .

مثال :

توزيع تكراري فيه المتوسط الحسابي $\bar{x}=50$ والوسيط $Med=48$ والتباين $S^2=144$ ما قيمة الالتواء فيه:

الحل :

$$Y_l = \frac{3(\bar{x} - Med)}{S}$$
$$= \frac{3(50 - 48)}{12} = 0.5$$

أي أنه ملتو نحو اليمين والتواءه يساوي 0.5.

أما القسمة على الانحراف المعياري في تعريف الالتواء فذلك لجعله غير معتمد على وحدة القياس المستعملة في البيانات.(أبو صالح ، 2000، ص 112-113 ؛ بدر، عباينة،2007 ص135-136)

ملاحظات هامة:

- المتوسط الحسابي أدق المقاييس الثلاثة لأنه يأخذ في الحسبان جميع فئات التوزيع، يليه الوسيط، وأخيراً المنوال لأن طرائق حسابه تقريبية.
- عادة ما تكون قيم مقاييس النزعة المركزية قريبة من بعضها البعض.(الصياد وربع وسمرة، 2008، ص24)

----- الفصل الخامس :مقاييس التشتت

أسئلة وتمارين خاصة بالفصل الرابع:

- 1- ما منوال التوزيع التالي : 6, 5,4,4,4,4,3
- 2- أكثر إحصائيات النزعة المركزية تأثراً بالقيم المتطرفة هو
- 3 - أفضل مقاييس النزعة المركزية لوصف المتغيرات الاسمية هو.....
- 4- أقل مقاييس النزعة المركزية تأثراً بتقلبات العينة هو.....
- 5- المقياس الذي يقسم التوزيع إلى قسمين متساويين مهما كان شكله هو.....
- 6- إذا كان المتوسط الحسابي لدرجات 50 طالباً في امتحان = 66 انسحب طالب درجته =15 فالمتوسط الحسابي لدرجات الباقيين يساوي.....
- 7- فيما يلي جدول توزيع تكراري للفئات والمطلوب :
اعتمد التوزيع التكراري التالي للإجابة عن الأسئلة التالية:

الجدول (4-17)

الفئات	f							
10 14	3							
14 18	5							
18 22	7							
22 26	10							
26 30	15							
30 38	9							
38 42	7							
42 46	5							

----- الفصل الخامس :مقاييس التشتت

أ- ما طول الفئة

ب- المنوال يساوي

ج- المتوسط الحسابي يساوي

د- الوسيط يساوي

10- أوجد كلا من المتوسط والوسيط والمنوال للتوزيع التالي:

الجدول (4-18)

الفئة	30 34	34 38	38 42	42 46	46 50
التكرار	1	2	5	4	3

11- ما الحالات التي يمتاز فيه استخدام كل من المتوسط الحسابي والوسيط والمنوال كمقاييس للنزعة المركزية؟

12- عند حساب مقاييس النزعة المركزية لتوزيع وحيد المنوال وجد أن المتوسط = 35 والوسيط = 32 فاحسب قيمة تقريبية للمنوال معتمداً على العلاقة بين مقاييس النزعة المركزية؟ حدد فيما إذا كان التوزيع ملتوياً وبأي اتجاه ؟

الفصل الخامس

مقاييس التشتت

Measures of Dispersion



----- الفصل الخامس :مقاييس التشتت

يشير مفهوم مقاييس النزعة المركزية إلى ميل البيانات للمركز حول قيمة ممثلة أو نموذجية في التوزيع، وتستخدم مقاييس النزعة المركزية لغايات المقارنة بين مجموعتين من البيانات ولوصف توزيع البيانات، وتساعد هذه المقاييس في فهم وتفسير سلوك الظواهر.

أما مقاييس التشتت مدى تقارب أو تباعد أو تتأثر أو اختلاف القيم عن بعضها البعض أو عن نقطة معينة كالمتوسط مثلاً. فالتشتت يرجع إلى اختلاف قيم البيانات، لذلك فإن القيم المتشابهة لا يوجد لها تشتت، ولكن باختلاف قيم البيانات فلا بد من وجود تشتت بين هذه القيم، حيث يختلف مقدار التشتت من مجموعة إلى أخرى.

وتبرز الحاجة إلى مقاييس التشتت لتقديم وصف دقيق حول التوزيعات التكرارية وخصوصاً عند إجراء المقارنات بين المجموعات أو بين الأفراد، فدالات مقاييس النزعة المركزية كالمتوسط الحسابي قد تقود إلى اتخاذ قرارات غير صحيحة.

فمثلاً لدينا ثلاث مجموعات من القيم :

الجدول (1-5)

المجموعات	الأولى	الثانية	الثالثة
القيم (البيانات)	5 , 5 , 5 , 5	4, 5, 6,8, 9,2,1	1, 9
المتوسط	5	5	5

يلاحظ أن قيمة المتوسط الحسابي لكل مجموعة يساوي (5) فإذا اكتفينا بهذا المقياس فإننا نقرر أن المجموعات الثلاث متشابهة ولكن في الحقيقة أن قيم المجموعة الثالثة أكثر تباعداً من قيم المجموعة الثانية والتي هي أكثر تباعداً من قيم المجموعة الأولى.

----- الفصل الخامس :مقاييس التشتت

" وبهذا يعد مقياس التشتت مكملاً لمقياس النزعة المركزية في وصف البيانات، كما أنه يعد مؤشراً على مدى كفاءة مقياس النزعة المركزية في تمثيل البيانات، عندما تكون البيانات أقل تشتتاً أو اختلافاً فيما بينها. فمثلاً إذا كنا نهتم بمعرفة عمر المفحوص، وكان متوسط هذا العمر (25) سنة، فكلما كان الانحراف المعياري (مقياس التشتت حول المتوسط الحسابي صغيراً دل ذلك على ثقة أكبر في المتوسط ممثلاً للقيم). (فهيم، 2005، ص86)

وهناك عدة مقاييس للتشتت منها:

1- المدى Range

يعد المدى أسهل وأبسط طريقة في حساب تشتت البيانات الكمية، ويعرف المدى أنه الفرق بين أكبر قيمة وأصغر قيمة في البيانات.

أولاً: حساب المدى لبيانات غير مبوبة:

مثال: لدينا البيانات التالية:

3, 4, 7, 8, 9, 10

المطلوب: حساب المدى.

المدى = أكبر قيمة - أصغر قيمة

$10 - 3 = 7$ = المدى

ثانياً: حساب المدى لفئات القيم:

يحسب المدى لفئات القيم بطرح الحد الأعلى للفئة الأخيرة من الحد الأدنى للفئة الأولى

----- الفصل الخامس :مقاييس التشتت

مثال : لدينا البيانات التالية والمطلوب حساب المدى :

الجدول (5-2)

الفئات	40 50	50 60	60 70	70 80	80 90
التكرار	2	9	15	10	5

المدى = الحد الأعلى للفئة الأخيرة - الحد الأدنى للفئة الأولى

المدى = 40-90

المدى = 50

خواص المدى:

- يعد المدى أسهل وأبسط مقاييس التشتت، ولذلك فهو يستخدم في الأوساط العامة
- يعطي فكرة سريعة عن طبيعة البيانات ويستخدم كثيراً في مراقبة جودة الإنتاج، وكذلك في وصف الأحوال الجوية. (بري، هندي، عبد الله، 95، ص1994)
- يتأثر بالقيم المتطرفة أو الشاذة، لذلك لا يفضل استخدامه في حال وجود قيم متطرفة.
- لا يمكن حسابه من الجداول التكرارية المفتوحة.

2- الانحراف الربيعي Quartile Deviation

للتغلب على بعض عيوب المدى، والتي من أهمها تأثره بالقيم الشاذة وعدم إمكانية حسابه في حال وجود الجداول التكرارية المفتوحة وجد مقياس آخر للتشتت هو ما يسمى بالانحراف الربيعي أو (نصف المدى الربيعي). والذي يعتمد في حسابه على إيجاد قيم الربيعات.

مفهوم الربيعات:

تعرف الربيعات أنها: النقاط التي تقسم التوزيع التكراري إلى أربعة أقسام متساوية، بحيث تكون درجات التوزيع مرتبة ترتيباً تصاعدياً.

الربيع الأول: هو النقطة التي تسبقها ربع القيم وتليها ثلاثة أرباع القيم ويرمز له بـ Q1.

الربيع الثاني: هو النقطة التي تسبقها نصف البيانات وتليها نصف البيانات ويرمز له بـ Q2 وهو قيمة الوسيط نفسها.

الربيع الثالث: هو النقطة التي تسبقها ثلاثة أرباع البيانات وتليها ربع البيانات ويرمز له بـ Q3.

تحسب الربيعات بالطريقة نفسها التي يحسب فيها الوسيط مع اختلاف بسيط في كيفية حساب رتبة كل ربع.

ومقياس التشتت الذي يعتمد على الربيعات يسمى نصف المدى الربيعي أو الانحراف الربيعي.

1- حساب الانحراف الربيعي لبيانات غير مبوبة:

مثال: لدينا البيانات التالية التي تمثل درجات الطلاب في مادة من المواد العملية والمطلوب حساب نصف المدى الربيعي لهذه الدرجات بعد تحديد الدرجات التي تمثل الربيع الأول والثاني والثالث مع تفسير تلك الدرجات .

الجدول (3-5)

X	
3	
8	
→	Q1
10	
12	
→	Q2
20	
24	
→	Q3
30	
50	

طريقة حساب الانحراف الربيعي للبيانات غير المبوبة:

1- نرتب البيانات تصاعدياً

2- نحسب ترتيب الربع الأول:

$$\frac{n+1}{4} = \frac{8+1}{4}$$

$$= 2.25$$

2.25 تعني القيمة الثانية +ربع المسافة بين الثانية والثالثة

$$Q1 = 8 + \frac{10-8}{4} = 8 + \frac{2}{4} = 8.5$$

تعني قيمة الربع الأول (8.5) أنها الدرجة التي يقل عنها 25% من درجات الطلاب ويزيد عنها 75% من درجات الطلاب.

3- نحسب ترتيب الوسيط:

$$\frac{n+1}{2} = \frac{8+1}{2} = 4.5$$

4.5 تعني القيمة الرابعة +نصف المسافة بين الرابعة والخامسة.

$$\text{med أو } Q_2 = 12 + \frac{20-12}{2} = 12 + \frac{8}{2} = 16$$

تعني قيمة الربع الثاني (16) أنها الدرجة التي يقل عنها 50% من درجات الطلاب ويزيد عنها 50% من درجات الطلاب.

4- نحسب ترتيب الربع الثالث:

$$\frac{3(n+1)}{4} = \frac{3 \times 9}{4} = 6.75$$

6.75 تعني القيمة السادسة + ثلاثة أرباع المسافة بين السادسة والسابعة

$$Q_3 = 24 + \frac{3(30-24)}{4} = 24 + \frac{18}{4} = 28.5$$

تعني قيمة الربع الثالث (28.5) أنها الدرجة التي يزيد عنها 25% من درجات الطلاب ويقل عنها 75% من درجات الطلاب.

$$Q = \frac{Q_3 - Q_1}{2}$$

يحسب نصف المدى الربيعي من خلال العلاقة التالية:

حيث Q هو نصف المدى الربيعي

Q1 الربع الأول

----- الفصل الخامس:مقاييس التشتت

Q3 الربع الثالث

$$Q = \frac{(Q3 - Q1)}{2}$$

$$= \frac{(28.5 - 8.5)}{2} = 10 \text{ Q}$$

الانحراف الربيعي النسبي: Q%

يستخدم الانحراف الربيعي النسبي للمقارنة بين تشتت عينتين، وتكون العينة الأكثر تشتتاً هي العينة صاحبة الانحراف الربيعي النسبي الأكبر. علماً أن الثقة تكون أكبر في العينة الأقل تشتتاً. مثال:

قارن بين تشتت العينتين، وحدد أيهما أكثر تشتتاً، وبأيهما ثقنا أعلى؟ ولماذا؟

الجدول (5-4)

العينة الأولى	العينة الثانية
Med=16	Med=36
نصف المدى الربيعي	المدى الربيعي
Q=10	Q3-Q1=14
	نحسب الانحراف الربيعي (نصف المدى الربيعي):
$Q\% = \frac{10}{16} \times 100 = 62.5\%$	$Q = \frac{Q3 - Q1}{2} = \frac{14}{2} = 7$
	$Q\% = \frac{7}{36} \times 100 = 19.44\%$

العينة الثانية أقل تشتتاً لهذا ثقنا فيها أكبر.

----- الفصل الخامس: مقاييس التشتت

2- حساب الانحراف الربيعي (نصف المدى الربيعي) لبيانات مبوبة:

مثال:

الجدول (5-5)

الفئات	f	$f \uparrow$
-12	6	6
12-18	14	20
18-20	24	44
20-24	16	60
24-30	10	70
30-36	6	76
36-	4	80
	$\sum f = 80$	

الحل:

لحساب الانحراف الربيعي نقوم بما يلي:

1- نحسب التكرار المتجمع الصاعد

2- نحدد ترتيب الربع الأول:

$$Q_1 \text{ رتبة} = \frac{\sum f}{4} = \frac{80}{4} = 20$$

3- نحسب الربع الأول:

$$Q_1 = L_0 + \frac{\frac{\sum f}{4} - f \uparrow - 1}{f_{Q_1}} \cdot C_{Q_1}$$

----- الفصل الخامس: مقاييس التشتت

$$Q_1 = 12 + \frac{14}{14} \times 6$$

$$= 12 + 6$$

$$Q_1 = 12 + \frac{14}{14} \times 6$$

$$Q_1 = 12 + 6$$

$$Q_1 = 18$$

4- نحسب ترتيب الربيع الثالث:

$$= \frac{3(80)}{4} = 60 \frac{3 \sum f}{4}$$

5- نحسب الربيع الثالث

$$= LQ_3 + Q_3 \times \frac{\frac{3 \sum f}{4} - f}{f Q_1}$$

$$\times 4 = 20 = Q_3 \frac{60 - 44}{16}$$

$$20 + 4 = 24$$

6- نحسب الانحراف الربيعي

$$Q = \frac{Q_3 - Q_1}{2} = \frac{24 - 18}{2}$$

$$Q = 3$$

----- الفصل الخامس: مقاييس التشتت

مثال : ليكن لدينا فئات القيم التالية والمطلوب حساب الانحراف الربيعي

الجدول (5-6)

الفئات	F	f
0-4	4	4
4-8	3	7
8-12	5	12
12-16	6	18
16-20	3	21
20-24	2	23
	$\Sigma f=23$	

الحل:

1- نحسب ترتيب الربيع الأول:

$$Q1 = \frac{\Sigma f}{4} = \frac{23}{4} = 5.75$$

2- نحسب الربيع الأول:

$$Q1 = 4 + 2.33 = 6.33 = 4 + \frac{5.75 - 4}{3} \times 4$$

3- نحسب ترتيب الربيع الثالث:

$$= \frac{3(23)}{4} = 17.25 = \frac{3\Sigma f}{4}$$

4- نحسب الربيع الثالث:

$$Q3 = 12 + \frac{17.25 - 12}{6} \times 4$$

$$= 12 + 3.5 = 15.5$$

5- نحسب الانحراف الربيعي

$$Q = \frac{Q3 - Q1}{2} = \frac{15.5 - 6.33}{2} = \frac{9.17}{2} = 4.585$$

الجدول (5-7)

الفئات	F	f ↑
12-	6	6
18-12	14	20
18-20	24	44
20-24	16	60
24-30	10	70
30-36	6	76
36-	4	80
	$\Sigma f=80$	

لحسابه نقوم بما يلي :

1- نحسب التكرار التجميعي الصاعد .

2- نحدد ترتيب الربع الأول :

$$\frac{\Sigma f}{4} = \frac{80}{4} = 20$$

ملاحظة : لا يوجد هنا فئة عددها 19.25 فنأخذ الأكبر منها مباشرة وهي هنا 20.

3- نحسب الربع الأول :

$$Q_1 = L_0 + \frac{\frac{\Sigma f}{4} - f_{\uparrow-1}}{f_{Q_1}} \cdot C_{Q_1}$$

$$Q_1 = 12 + \frac{20 - 6}{14} \cdot 6$$

$$= 12 + 6 = 18$$

4- نحسب ترتيب الربع الثالث :

$$Q_3 = \frac{3 \Sigma f}{4} = \frac{3 \times 80}{4} = 60$$

5- نحسب الربع الثالث :

$$\begin{aligned} &= 20 + \frac{60-44}{16} \times 4 \\ &= 20 + \frac{16}{16} \times 4 \\ &= 20+4 = 24 \end{aligned}$$

6- نحسب الانحراف الربيعي :

$$Q = \frac{Q_3 - Q_1}{2} = 3$$

بعض خواص الانحراف الربيعي:

- يفضل استخدامه مقياساً للتشتت حينما يستخدم الوسيط كمقياس للنزعة المركزية، وأيضاً حين يكون الجدول التكراري مفتوحاً أو شديد الالتواء أو عندما يكون هناك قيماً متطرفة.
- يتحدد بعدد البيانات وليس بقيمتها .

الفوائد التطبيقية للربيعات:

- قياس التشتت: عن طريق حساب نصف المدى الربيعي وهو يحسب بسرعة وسهولة.
- المعايير والمستويات: للربيعات أهمية قصوى في معرفة نقط التوزيع التكراري التي تحدد المستويات العليا والوسطى والدنيا للدرجات. وبهذا المعنى تحدد المستويات المختلفة للضعيف والمتوسط والممتاز، وهي تصلح لتقنيين الاختبارات والمقاييس المختلفة وللكشف عن معاييرها ومستوياتها وتحديدها تحديداً دقيقاً.

العشریات Deciles

مفهوم العشریات Deciles :

العشریات: هي النقط التي تقسم التوزيع التكراري إلى أجزاء عشرية. كما قسمته الإرباعیات إلى أربعة أقسام : كل قسم يحدد ربع التوزيع التكراري.

طرائق حساب والعشریات:

لا تختلف طريقة حساب العشریات عن طريقة حساب الإرباعیات أو إلا في الخطوة الأولى التي تقرر ترتيب الإرباعي وترتيب العشري، كما اختلفت الإرباعیات عن الوسيط في الخطوة نفسها. فعند حساب ترتيب الوسيط يقسم عدد الدرجات على 2 أي ترتيب الوسيط يساوي $\frac{n}{2}$ لأنه يقسم التوزيع التكراري إلى نصفين، وهو بذلك يقع في منتصف التوزيع. وعند حساب ترتيب الإرباعیات نقسم عدد الدرجات على أربعة. وبذلك يصبح ترتيب الإرباعي الأول مساوياً لـ $\frac{n}{4}$ وترتيب الإرباعي الثاني مساوياً لـ $\frac{2n}{4}$ أي $\frac{n}{2}$ ، وترتيب الإرباعي الثالث مساوياً لـ $\frac{3n}{2}$.

وهكذا يصبح ترتيب العشري الأول مساوياً لـ $\frac{10n}{100}$ أي $\frac{n}{10}$ وترتيب العشري الثاني مساوياً لـ $\frac{20n}{100}$ أي $\frac{2n}{10}$ وهكذا بالنسبة لبقية العشریات.

وبالطريقة نفسها يمكن أن تقسم التوزيع التكراري إلى تساعیات أو سباعیات أو غير ذلك من الأقسام المختلفة تبعاً لرغبة الباحث وهدف البحث. ويعتمد كل تقسيم من هذه التقسيمات على تحديد ترتيب القسم.

حساب العشریات :

حساب العشري الأول نتبع الخطوات التالية :

1- ترتيب العشري الأول

$$\frac{350}{10} \times 1 = 35$$

2- قيمة العشري الأول

$$\begin{aligned} &= 15 + \frac{35-12}{29} \times 5 \\ &= 15 + \frac{22}{29} \times 5 \\ &= 3.793 + 15 \\ &= 18.793 \end{aligned}$$

الفوائد العلمية والتطبيقية العشريات:

بما أن العشريات تقسم التوزيع التكراري إلى ما هو أكبر من، وما هو أقل من حد فاصل معين إذن فهي بذلك تحدد مستويات متدرجة للبيانات الرقمية التي يشتمل عليها التوزيع .
العشري الأول مثلاً يبين بوضوح جميع قيم الدرجات التي تقل عن مستواه، وبدراسة مثالنا السابق المبين بالجدول (43) نرى أن أي درجة تقل عن 18.3 تقل عن العشري الأول، أي أن مستوى جميع الأفراد الذين حصلوا على درجات تمتد من صفر إلى 18 هو أضعف المستويات بالنسبة لتدريجنا القياسي لمستويات الدرجات وأن أي درجة تقل عن 30 تقل بذلك عن العشري الخامس أي إن النقطة العشرية التي تقع عند 30 تحدد تماماً هذا المستوى المتوسط في التدرج.

----- الفصل الخامس :مقاييس التشتت

وهكذا تصلح هذه الطريقة إلى حد كبير في تحديد مستويات ومعايير الأفراد في أي اختبار وتبدو أهمية هذه المعايير في فهمنا للدرجات الخام التي يحصل عليها الفرد وذلك لأن هذه الدرجات تكتسب معنى واضحاً عندما تنسب إلى مستويات الجماعة التي أجرى عليها الاختبار. وعندما تكون هذه الجماعة كبيرة وممثلة تماماً لجميع الأفراد الذين يحتمل انتمائهم إليها وعندما يهذب التوزيع التكراري الدرجات بحيث يقترب من التوزيع الاعتدالي فإن هذه العشريّات تصبح مقاييس ومعايير صالحة للمقارنة بين درجات أي فرد في ذلك الاختبار والمستويات التي حددتها درجات تلك الجماعة .

فإذا أجرى اختبار للذكاء على آلاف الأفراد الذين تمتد أعمارهم مثلاً من 6 سنوات إلى 7 سنوات ثم حسبت العشريّات لدرجات هؤلاء الأفراد أمكن اتخاذ هذه النقط معايير لتحديد مستويات ذكاء أي فرد يمتد عمره الزمني من 6 سنوات إلى 7 سنوات .

ونستطيع أن نمّدت بتلك المعايير إلى جميع الأعمار بحيث نحدد لكل عمر زمني نقطه العشرية المتدرجة .وبما أن هذه النقط العشرية تحدد منتصف درجات كل اختبارات عند العشري الخامس، إذن فهي بذلك تنسب جميع التوزيعات التكرارية إلى منتصف واحد ثابت وهكذا نستطيع أن نقارن نتائج الاختبارات المختلفة بمقارنة نقطها العشرية، أو نقارن نتائج الجماعات المختلفة بالنسبة لاختبار واحد وذلك بمقارنة نقطها العشرية أيضاً كما قارنا نتائج الفرد بالنسبة للمعايير التي تحددها نتائج الجماعة .

الانحراف المعياري

Standard deviation

الانحراف المعياري أهم مقاييس التشتت وهو يقوم في جوهره على حساب انحرافات الدرجات عن متوسطها كما تدل تسميته عليه . وهو الجذر التربيعي لمجموع مربعات انحرافات القيم عن متوسطها مقسوماً على عددها.

أولاً: حساب الانحراف المعياري للبيانات غير المبوبة :

هناك طرائق عديدة لحساب الانحراف المعياري منها:

1- الأسلوب المباشر :

مثال: ليكن لدينا مجموعة من القيم والمطلوب حساب الانحراف المعياري.

الجدول (5-10)

الدرجات X	$X-X$	$(X-X)^2$	X^2	$d=(X-X_0)$	d^2
7	3+	9	49	4+	16
4	0	0	16	1+	1
2	2-	4	4	1-	1
1	3-	9	1	2-	4
7	3+	9	49	4+	16
4	0	0	16	1	1
4	0	0	16	1	1
3	1-	1	9	0	0
$\Sigma x= 32$ $\bar{X} = 4$	0	32	$\Sigma x^2=160$	$11-3=8+$	40

الحل:

1- الأسلوب المباشر :

خطوات حساب الانحراف المعياري بالأسلوب المباشر :

- حساب المتوسط الحسابي للقيم.
- حساب انحراف القيم عن المتوسط وذلك بطرح كل قيمة من متوسط القيم.
- حساب مربع انحراف القيم.
- حساب مجموع مربع انحراف القيم.
- حساب الانحراف المعياري وفق الصيغة التالية:

$$S = \sqrt{\frac{\sum (x - \bar{x})^2}{n}}$$

- بالتعويض بالقيم:

$$S = \sqrt{\frac{32}{8}} = \sqrt{4} = 2$$

2- أسلوب المربعات :أو حساب الانحراف المعياري من قيم المتغير ذاتها

خطوات حساب الانحراف المعياري بأسلوب المربعات :

- حساب مربع القيم.
- حساب مجموع مربع القيم.
- حساب الانحراف المعياري وفق الصيغة التالية:

$$S = \sqrt{\frac{\sum x^2}{n} - \left(\frac{\sum x}{n}\right)^2}$$

----- الفصل الخامس :مقاييس التشتت

- بالتعويض بالقيم:

$$S = \sqrt{\frac{160}{8} - \left(\frac{32}{8}\right)^2}$$

$$S = \sqrt{20 - (4)^2} = \sqrt{4} = 2$$

3- الأسلوب غير المباشر (الانحرافات) :

خطوات حساب الانحراف المعياري بأسلوب الانحرافات:

- نفترض متوسطاً فرضياً وليكن في مثالنا السابق (3)
- نحسب انحراف القيم عن المتوسط الفرضي، وذلك بطرح كل قيمة من المتوسط الفرضي.

- نحسب مربع الانحرافات.

- نحسب مجموع مربع الانحرافات.

- نحسب الانحراف المعياري وفق الصيغة التالية:

$$S = \sqrt{\frac{\sum d^2}{n} - \left(\frac{\sum d}{n}\right)^2}$$

- بالتعويض بالقيم:

$$S = \sqrt{\frac{40}{8} - \left(\frac{8}{8}\right)^2}$$

$$S = \sqrt{5 - (1)^2} = \sqrt{4} = 2$$

نلاحظ أن قيمة الانحراف المعياري هي واحدة في الأساليب.

----- الفصل الخامس :مقاييس التشتت

مثال آخر للبيانات غير المبوبة:

الجدول (5-11)

X الدرجات	$X - \bar{X}$	$(X - \bar{X})^2$	d	d^2	x^2
5	6.8-	46.24	7-	49	25
8	3.8-	14.14	4-	16	64
12	0.2	0.04	0	0	14
14	2.2	4.84	2	4	196
20	8.2	67.24	8	64	400
$\Sigma x = 59$	0	132.8	$\Sigma d = -1$	$\Sigma d^2 = 133$	829

1- الأسلوب المباشر:

$$S = \sqrt{\frac{\sum (x - \bar{x})^2}{n}}$$

$$\bar{x} = \frac{\Sigma x}{n} = \frac{59}{5} = 11.8 \quad \text{نحسب المتوسط:}$$

نعوض فيكون :

$$S = \sqrt{\frac{132.8}{5}} = 5.15$$

2- الأسلوب غير المباشر (الانحرافات) : نفترض متوسطاً فرضياً = 12

$$S = \sqrt{\frac{\sum d^2}{n} - \left(\frac{\sum d}{n}\right)^2}$$

----- الفصل الخامس: مقاييس التشتت

$$S = \sqrt{\frac{133}{5} - \left(\frac{-1}{5}\right)^2}$$

$$S = 5.15$$

3- أسلوب المربعات :

$$S = \sqrt{\frac{\sum x^2}{n} - \bar{x}^2} = \sqrt{\frac{829}{5} - (11.8)^2}$$

$$S = 5.15$$

مثال :

الجدول (5-12)

X	$X - \bar{X}$	$(X - \bar{X})^2$
5	6.4-	40.96
8	3.4-	11.56
10	1.4-	1.96
14	2.6	6.76
20	8.6	73.96
$\sum x = 57$	0	$\sum (X - \bar{X})^2$ 135.2

$$\bar{X} = \frac{57}{5} = 11.4$$

$$11.4 - 5 = -6.4$$

$$S = \sqrt{\frac{135.2}{5}} = 5.2$$

----- الفصل الخامس :مقاييس التشتت

ثانياً:حساب الانحراف المعياري من فئات الدرجات (بيانات مبوبة):

مثال :

ليكن لدينا الجدول التالي الذي يمثل أجور العمال بألوف الليرات السورية :

الجدول (5-13)

الفئات	f	X_i	$X - \bar{X}$	$(X - \bar{X})^2$	$f(X - \bar{X})^2$
4-6	6	5	4.74-	22.46	134.8
6-8	12	7	2.74-	7.5	90.09
8-10	20	9	0.74-	0.55	10.95
10-12	14	11	1.26	1.58	22.22
12-14	10	13	3.26	10.62	106.2
14-16	5	15	5.26	27.66	138.34
	$\Sigma f=67$				502.69

يحسب الانحراف المعياري للفئات بطرائق أربع:

1- الأسلوب المباشر :

$$S = \sqrt{\frac{\sum f(X - \bar{X})^2}{\sum f}}$$

2- أسلوب الانحرافات :

$$S = \sqrt{\frac{\sum f d^2}{\sum f} - \left(\frac{\sum f d}{\sum f}\right)^2}$$

3- أسلوب الانحرافات المختصرة :

$$S = \sqrt{\frac{\sum f d^{\circ 2}}{\sum f} - \left(\frac{\sum f d^{\circ}}{\sum f}\right)^2} . C$$

4- أسلوب المربعات :

$$S = \sqrt{\frac{\sum f x^2}{\sum f} - \bar{x}^2}$$

نقل التربيع الأول إلى خارج القوس

1- الأسلوب المباشر:

- لحساب الانحراف المعياري وفق الأسلوب المباشر يجب حساب الوسط الحسابي :

$$\bar{X} = \bar{X}_0 + \frac{\sum f d_0}{\sum f} \cdot C = 9 + \frac{25}{67} \cdot 2 = 9.74$$

- نطرح مركز كل فئة من المتوسط $(X - \bar{X})$.
- نربع انحراف مركز كل فئة عن المتوسط الفرضي $(X - \bar{X}_0)^2$.
- نضرب مربع الانحراف بتكرار كل فئة $f(X - \bar{X}_0)^2$ ثم نجمع العمود .
- نحسب الانحراف المعياري وفق الصيغة التالية:

$$S = \sqrt{\frac{\sum f (X - \bar{X})^2}{\sum f}}$$

$$= \sqrt{\frac{502.69}{67}} = 2.74$$

2- أسلوب الانحرافات :

الجدول (14-5)

الفئات	f	X_i	d	fd	d^2	fd^2
4-6	6	5	4-	24-	16	96
6-8	12	7	2-	24-	4	48
8-10	20	9	0	0	0	0
10-12	14	11	2	28	4	56
12-14	10	13	4	40	16	160
14-16	5	15	6	30	36	180
المجموع	$\Sigma f=76$			$\Sigma fd=50$		$\Sigma fd^2=540$

خطوات الحل:

- اختيار متوسط فرضي للقيم هو عبارة عن مركز الفئة الواقعة في الوسط، ويفضل أن تكون الفئة صاحبة أعلى تكرار، وإن كان من الممكن استخدام فئة وسطى ليست صاحبة أعلى تكرار.
- حساب الانحراف (d) وذلك بطرح مركز كل فئة من مركز الفئة الفرضية
- ضرب تكرار كل فئة بالانحراف، ثم القيام بالجمع الجبري لعمود حاصل الضرب (fd)
- تربيع قيم الانحراف (d^2)
- ضرب تكرار كل فئة بمربع الانحراف (fd^2)، ثم جمع حاصل الضرب.
- يحسب الانحراف وفق الصيغة التالية:

----- الفصل الخامس: مقاييس التشتت

$$S = \sqrt{\frac{\sum f d^2}{\sum f} - \left(\frac{\sum f d}{\sum f}\right)^2}$$

$$S = \sqrt{\frac{540}{67} - \left(\frac{50}{67}\right)^2}$$

$$S = \sqrt{8.059 - (0.746)^2}$$

$$= \sqrt{8.059 - 0.5567}$$

$$= \sqrt{7.50} = 2.739 \approx 2.74$$

3- أسلوب الانحرافات المختصرة :

الجدول (5-15)

الفئات	f	X_i	d°	d^{2°	Fd°	Fd^{2°
4-6	6	5	2-	4	12-	24
6-8	12	7	1-	1	12-	12
8-10	20	9	0	0	0	0
10-12	14	11	1	1	14	14
12-14	10	13	2	4	20	40
14-16	5	15	3	9	15	45
المجموع	67				25+	135

----- الفصل الخامس :مقاييس التشتت

خطوات الحل:

- اختيار متوسط فرضي للقيم هو عبارة عن مركز الفئة.
- حساب الانحراف المختصر وذلك بقسمة الانحراف المحسوب بالأسلوب السابقة على طول الفئة (d).
- ضرب تكرار كل فئة بالانحراف المختصر (fd)، ثم القيام بالجمع الجبري للعمود.
- تربيع الانحراف المختصر (d^2).
- ضرب الانحراف المختصر بتكرار كل فئة (fd^2)، ثم القيام بجمع العمود.
- حساب الانحراف المعياري وفق الصيغة التالية:

$$S = \sqrt{\frac{\sum fd^2}{\sum f} - \left(\frac{\sum fd}{\sum f}\right)^2} . C$$

$$S = \sqrt{\frac{135}{67} - \left(\frac{25}{67}\right)^2} . 2$$

$$S = \sqrt{2.0149 - (0.373)^2} \times 2$$

$$S = 2\sqrt{2.014 - 0.139}$$

$$= 2\sqrt{1.875} =$$

$$= 2 \times 1.369$$

$$= 2.738$$

$$\approx 2.74$$

4- أسلوب المربعات :

الجدول (5-16)

الفئات	f	X_i	X_i^2	fX_i^2
4-6	6	5	25	150
6-8	12	7	49	588
8-10	20	9	81	1620
10-12	14	11	121	1694
12-14	10	13	169	1690
14-16	5	15	225	1125
المجموع	67			

خطوات الحل:

- حساب مركز كل فئة، ثم تربيع مراكز الفئات
- ضرب تكرار كل فئة بمربع مركزها (fX^2)، ثم جمع العمود .
- حساب الانحراف المعياري وفق الصيغة:

$$S = \sqrt{\frac{\sum f x^2}{\sum f} - \bar{x}^2}$$

$$S = \sqrt{\frac{6867}{67} - (9.74)^2} = 2.74$$

----- الفصل الخامس :مقاييس التشتت

مثال: ليكن لدينا فئات القيم التي تمثل تقديرات المعلمين لامتحانات مجموعة من (350) طالباً في الامتحانات الشفهية، والمطلوب حساب الانحراف المعياري لهذه التقديرات:

الجدول (5-17)

الفئات	f	X_i	d	fd	d^2	fd^2
0-5	2	2.5	25-	50-	625	1250
5-10	3	7.5	20-	60-	400	1200
10-15	8	12.5	15-	120-	225	1800
15-20	29	17.5	10-	290-	100	2900
20-25	51	22.5	5-	255-	25	1275
25-30	72	27.5	0	0	0	0
30-35	97	32.5	5+	485	25	2425
35-40	48	37.5	10+	480	100	4800
40-45	24	42.5	15+	360	225	5400
45-50	15	47.5	20+	300	400	6000
50-55	1	52.5	25+	25	625	625
				775- 1650+ =875		27675

1-أسلوب الانحرافات :

خطوات الحل:

- اختيار متوسط فرضي للقيم هو عبارة عن مركز الفئة الواقعة في الوسط، ويفضل أن تكون الفئة صاحبة أعلى تكرار، وإن كان من الممكن استخدام فئة وسطى ليست صاحبة أعلى تكرار.
- حساب الانحراف (d) وذلك بطرح مركز كل فئة من مركز الفئة الفرضية
- ضرب تكرار كل فئة بالانحراف، ثم القيام بالجمع الجبري لعمود حاصل الضرب (fd)
- تربيع قيم الانحراف (d^2)
- ضرب تكرار كل فئة بمربع الانحراف (fd^2)، ثم جمع حاصل الضرب.
- يحسب الانحراف وفق الصيغة التالية:

$$S = \sqrt{\frac{\sum fd^2}{\sum f} - \left(\frac{\sum fd}{\sum f}\right)^2}$$

$$S = \sqrt{\frac{27675}{350} - \left(\frac{875}{350}\right)^2}$$

$$S = \sqrt{79.714 - (2.5)^2}$$

$$S = \sqrt{72.82} \approx 8.5$$

----- الفصل الخامس: مقاييس التشتت

الجدول (5-18)

الفئات	f	X_i	$d_{\bar{o}}$	$f d_{\bar{o}}$	$d_{\bar{o}}^2$	$f d_{\bar{o}}^2$
0-5	2	2.5	5-	10-	25	50
5-10	3	7.5	4-	12-	16	48
10-15	8	12.5	3-	24-	9	72
15-20	29	17.5	2-	58-	4	116
20-25	51	22.5	1-	51-	1	51
25-30	72	27.5	0	0	0	0
30-35	97	32.5	1+	97	1	97
35-40	48	37.5	2+	96	4	192
40-45	24	42.5	3+	72	9	216
45-50	15	47.5	4+	60	16	240
50-55	1	52.5	5+	5	25	25
	350			175		1107

2- أسلوب الانحرافات المختصرة :

خطوات الحل:

- اختيار متوسط فرضي للقيم هو عبارة عن مركز الفئة.
- حساب الانحراف المختصر، وذلك بقسمة الانحراف المحسوب بالأسلوب السابقة على طول الفئة ($d_{\bar{o}}$).
- ضرب تكرار كل فئة بالانحراف المختصر ($f d_{\bar{o}}$)، ثم القيام بالجمع الجبري للعمود.

----- الفصل الخامس :مقاييس التشتت

- تربيع الانحراف المختصر (d^2).
- ضرب الانحراف المختصر بتكرار كل فئة ($f d^2$)، ثم القيام بجمع العمود.
- حساب الانحراف المعياري وفق الصيغة التالية:

$$S = \sqrt{\frac{\sum f d^2}{\sum f} - \left(\frac{\sum f d}{\sum f}\right)^2} \times C$$

$$S = \sqrt{\frac{1107}{350} - \left(\frac{175}{350}\right)^2} \times 5$$

$$S = \sqrt{3.1629 - (0.5)^2} \times 5$$

$$S = \sqrt{3.1629 - 0.25} \times 5$$

$$S = \sqrt{3.1629 - 0.25} \times 5$$

$$S = \sqrt{2.912} \times 5 \approx 8.5$$

معامل الاختلاف النسبي Coefficient of Variation:

إن مقاييس التشتت السابقة تعتمد جميعها على الوحدات المستخدمة في القياس، وبالتالي لا يمكن استخدامها في المقارنة بين مجموعتين أو أكثر مقاسة بوحدات قياس مختلفة، مثل الأطوال والأوزان والدرجات ... الخ . ومن ناحية أخرى لو أردنا المقارنة من حيث التشتت حتماً سيتم استخدام أفضل تلك المقاييس، وهو الانحراف المعياري، وحيث أن الانحراف المعياري يقيس مقدار انحراف القيم عن متوسطها، وعليه فإن عملية المقارنة تكون غير واقعية إذا اختلفت المتوسطات الحسابية للمجموعات المراد مقارنتها، حتى ولو كانت مقاسة بنفس الوحدات. لذلك

----- الفصل الخامس :مقاييس التشتت

وجدت مقاييس أخرى لا تعتمد على الوحدات المستخدمة في القياس حيث تقيس الاختلاف النسبي دون وحدة تمييز ، أهمها وأكثرها استخداماً ما يسمى بمعامل الاختلاف النسبي .

وباختصار فإن معامل الاختلاف النسبي يستخدم عند مقارنة تشتت مجموعتين (أو أكثر) من القيم إذا كانت تختلف في متوسطها أو في وحدات قياسها، وهو عبارة عن النسبة المئوية للانحراف المعياري منسوبة إلى المتوسط الحسابي وبالطبع كلما كانت قيمة هذا المعامل كبيرة دل ذلك على وجود تشتت كبير بين مفردات التوزيع والعكس صحيح. ويرمز لها المعامل بالرمز C.V ويعرف كما يلي :

$$\text{معامل الاختلاف النسبي} = \frac{\text{الانحراف المعياري}}{100 \times \text{المتوسط الحسابي}}$$

$$C.V = \frac{S}{X} \times 100$$

وهناك صورة أخرى لمعامل الاختلاف إذا تعذر حساب المتوسط الحسابي و الانحراف المعياري، كما في حالة الفئات المفتوحة، أو وجود قيم شاذة أو متطرفة، في هذه الحالات يمكن استخدام الصورة التالية لمعامل الاختلاف والتي تستخدم الانحراف الربيعي بدلاً من الانحراف المعياري، والوسيط بدلاً من الوسط الحسابي، ويأخذ معامل الاختلاف في هذه الحالة الشكل التالي :

$$\text{معامل الاختلاف النسبي} = \frac{\text{الانحراف الربيعي}}{100 \times \text{الوسيط}}$$

$$C.V = \frac{Q}{Med} \times 100$$

----- الفصل الخامس :مقاييس التشتت

ملاحظة : عند المقارنة بين عدة مجموعات يفضل استخدام الصيغة نفسها لمعامل الاختلاف، فلا ينبغي استخدام الصيغة الأولى لمجموعة، واستخدام الصيغة الثانية لمجموعة أخرى، بل ينبغي استخدام نفس الصيغة للمجموعتين .

مثال :عند مقارنة تشتت درجات الطلاب في اختبارين للذكاء (ريفن ووكسلر)، وجد ما يلي :

الانحراف المعياري	المتوسط الحسابي	
10	40	اختبار ريفن
14	140	اختبار وكسلر

فهل نستطيع القول استناداً إلى قيمة الانحراف المعياري فقط إن التشتت في اختبار وكسلر أكبر من التشتت في اختبار ريفن؟ الإجابة لا، لسببين أولهما اختلاف وحدات القياس وثانيهما أنه حتى وإن كانت وحدات القياس متشابهة فهناك سبب آخر وهو اختلاف المتوسطات الحسابية في المجموعتين . لذلك سيصبح من الضروري استخدام معامل الاختلاف النسبي لغرض المقارنة، كمايلي :

$$\text{معامل الاختلاف النسبي (لاختبار ريفن)} = 100 \times \frac{10}{40}$$

$$\text{معامل الاختلاف النسبي (لاختبار ريفن)} = 25\%$$

$$\text{معامل الاختلاف النسبي (لاختبار وكسلر)} = 100 \times \frac{14}{140}$$

$$\text{معامل الاختلاف النسبي (لاختبار وكسلر)} = 10\%$$

----- الفصل الخامس :مقاييس التشتت

وعلى ذلك وبالاتماد على معامل الاختلاف النسبي نستطيع القول بأن تشتت الدرجات في اختبار رافن أكبر من تشتت في اختبار وكسلر، في حين أنه كان من الممكن أن نصل إلى استنتاج خاطئ لو اعتمدنا على قيم الانحراف المعياري فقط (فهمي، 2005، ص192-193).

$$S = \sqrt{\frac{\sum f(X - \bar{x})^2}{\sum f}}$$

مثال: ليكن لدينا البيانات التالية التي تمثل أجور مجموعة من عمال الصناعات الغذائية:

الجدول (5-19)

فئات الأجور	عدد العمال f	X_i	fX_i	$x - \bar{x}$	$(x - \bar{x})^2$	$f(x - \bar{x})^2$
14-20	5	17	85	14.35-	205.92	1029.61
20-24	10	22	220	9.35-	87.51	875.1
24-30	12	27	324	4.35-	18.96	227.59
30-36	20	33	660	1.65	2.72	54.45
36-40	15	38	570	6.65	44.22	663.34
40-44	8	42	336	10.65	113.42	907.38
	$\sum f = 70$					3757.47

كما سحبت عينة ثانية من عمال الصناعات النسيجية فكان متوسط أجورهم الشهرية 32.500 ل.س وتباين $(S)^2$ الأجور 3600 ل.س، فأأي الأجور ذات تشتت أكبر الصناعات الغذائية أم النسيجية ولماذا ؟
متوسط العينة الأولى :

$$\bar{X} = \frac{\sum fX_i}{\sum f} = \frac{2195}{70} = 31.35$$

----- الفصل الخامس :مقاييس التشتت

انحراف العينة الأولى:

$$S = \sqrt{\frac{\sum f(x-\bar{x})^2}{\sum f}}$$

$$S = \sqrt{\frac{3757.47}{70}} = 7.326$$

معامل الاختلاف للعينة الأولى :

$$C.V = \left(\frac{S}{\bar{x}}\right) \times 100$$

$$C.V = \frac{7.326}{31.35} \times 100$$

$$= 23.36\%$$

العينة الثانية :

$$= \bar{x} 32500$$

$$S^2 = 3600 \implies S = 60$$

$$C.V = \frac{60}{32500} \times 100 = 18\%$$

نلاحظ أن أجور الصناعات الغذائية ذات تشتت أكبر من أجور الصناعات النسيجية لأن معامل اختلافها أكبر .

ملاحظات على الانحراف المعياري:

- الانحراف المعياري حساس لبعد أو قرب العلامات من المتوسط ولذلك كلما صغرت قيمته دل ذلك على طبيعة البيانات مقارنة ومتراكمة حول المتوسط، وبالتالي التشتت قليل والعكس هو الصحيح.

----- **الفصل الخامس :مقاييس التشتت**

- يأخذ في الحسبان جميع القيم في حسابه وبالتالي يستخدم عند المقارنة بين المجموعات والعينات الإحصائية والاستنتاجات الإحصائية المتعلقة بالفرضيات.
- يتأثر الانحراف المعياري بالقيم الشاذة أو المتطرفة.
- يصعب إيجاد الانحراف المعياري في حالة الفئات المفتوحة (المنزل وغريبة ، 2005 ، ص70) .

الدرجة المعيارية

الدرجات المعيارية Z Scores

تمر بنا مواقف يومية نحتاج فيها للمقارنة بين القيم المختلفة، كأن نقارن بين دخل موظفين في بلد ما، فالمقارنة هنا بين قيمتين من نفس التوزيع يمكن إجراؤها بصورة مباشرة، أما إذا كانت المقارنة بين توزيعين مختلفين كالمقارنة بين راتب موظف في سورية حيث راتبه 20000 ليرة، وموظف آخر في مصر راتبه 300 جنيه، فكيف تتم المقارنة؟

هنا المقارنة المباشرة غير عادلة ولا يمكن إجراؤها من خلال القيم الظاهرة، ولابد من إجراء عملية تحويلية لهذه القيم، أي تحديد هذه القيم من خلال توزيع معياري ودرجات معيارية يمكن أن نقيم المقارنة من خلالها. وقد اتفق إحصائياً على تحديد التوزيع الطبيعي المعياري والدرجات المعيارية (Z Scores) كمرجعية موحدة لكل التوزيعات.

الدرجة المعيارية Z :

الدرجة المعيارية Z هي عبارة عن عدد الانحرافات المعيارية (للدرجات الحقيقية في التوزيع الطبيعي) التي تقع فوق المتوسط (للدرجات الموجبة) والتي تقع تحت المتوسط (للدرجات السالبة). وهي تحسب بنسبة انحراف درجة الفرد عن المتوسط إلى انحراف درجات أفراد المجموعة كلها عن هذا المتوسط.

فالانحراف المعياري هو وحدة المقياس المطلوبة. واعتمد التوزيع الطبيعي المعياري على أنه التوزيع الطبيعي الذي يكون متوسطه صفراً وانحرافه المعياري (1)

تحويل الدرجات الخام إلى درجات معيارية Z :

نستطيع تحويل درجات خام من توزيعات مختلفة إلى درجات معيارية حيث نحتاج إلى معرفة العلامة الخام x ، والمتوسط الحسابي $\bar{x}$ ، والانحراف المعياري S ومن خلال المعادلة التالية نستطيع تحويل العلامة الخام إلى علامة معيارية:

$$Z = \frac{x - \bar{x}}{S}$$

مثال :

تقدم 20 طالباً لامتحان فكان المتوسط الحسابي للعلامات = 70، والانحراف المعياري للعلامات = 8 وحصل طالب على علامة 75 في هذا الامتحان، وحصل الطالب نفسه على علامة 75 في اختبار آخر كان متوسط علامات هذا الاختبار 65، والانحراف المعياري للعلامات 15، ففي أي من الاختبارين كان تحصيله أفضل؟

الحل : نحول العلامتين إلى درجات معيارية

$$Z_1 = \frac{75 - 70}{8} = \frac{5}{8} = 0.625$$

$$Z_2 = \frac{75 - 65}{15} = \frac{10}{15} = 0.666$$

بما أن $Z_2 > Z_1$ تحصيله في الامتحان الثاني أفضل.

مثال (2) : إذا كان $\bar{X} = 20$ ، $S = 3$ أوجد :

1- القيمة المعيارية للقيمة الخام 29 .

2- القيمة الخام التي قيمتها المعيارية (-2)

الحل :

$$1-Z = \frac{29-20}{3} = \frac{9}{3} = 3$$

$$-2 = \frac{x-20}{3}$$

$$x-20 = -6$$

$$x = 14$$

العلامات التائية T Scores T :

نظراً إلى احتواء العلامات المعيارية Z على كسور عشرية في الغالب، ولأنها تكون سالبة أحياناً فإنها لا تستعمل كثيراً في الأمور العملية إلا أنها تستعمل كثيراً في القضايا العلمية لما ذكر فيها من خصائص، وللتخلص من ذلك استحدثت علامات جديدة ناتجة من العلامات المعيارية Z وذلك بضربها بعشرة وإضافة الرقم (50) إليها مما يؤدي إلى التخلص من القيم السالبة والكسور العشرية على النحو التالي :

$$T = Z10 + 50$$

ومن الجدير بالذكر أن متوسط توزيع العلامة المعيارية T يساوي 50 وانحرافها المعياري يساوي 10، كذلك لا تؤثر هذه العملية التحويلية على شكل التوزيع لأنه تحويل خطي.

مثال :

إذا كانت العلامة المعيارية لطالب في امتحان تحصيلي تساوي (-2.4)، فما هي العلامة المعيارية T ؟

الحل :

$$\begin{aligned} T &= Z10 + 50 \\ &= 10 \times (-2.4) + 50 \\ &= -24 + 50 = 26 \end{aligned}$$

لاحظ كيف أن الإشارة السالبة قد اختفت، وأن العلامة (T) قيمة غير كسرية .

ونظراً إلى ما تتصف به العلامات المعيارية من ميزات فقد أخذت مقاييس الذكاء مثل مقياس ستانفورد وبينيه، ومقياس كسلر، وغيرها في استخدامها في معايرها بعد إدخال بعض التعديلات المناسبة عليها.

في الإمكان تحويل أي توزيع من العلامات الخام إلى توزيع آخر ذي متوسط حسابي وانحراف معياري جديدين، فمثلاً قد نحول توزيعاً لعلامات خام إلى توزيع بوسط حسابي = 40 وانحراف معياري = 12، وذلك بتحويل العلامات الخام إلى علامات معيارية Z ، ومن ثم نحول هذه إلى العلامات المطلوبة حسب العلاقة : $40 + Z12$.

التوزيع الطبيعي المعياري

Standard Normal Distribution

يعد التوزيع الطبيعي (أو ما يعرف بالتوزيع المعتدل) من التوزيعات الاحتمالية المهمة في الإحصاء وفي الدراسات التربوية والاجتماعية والإنسانية، إذ يسود اعتقاد عام مفاده أن معظم السمات والخصائص الإنسانية (درجات الذكاء، أطوال الأشخاص، التحصيل الدراسي ... الخ) تتوزع طبيعياً أو تقترب من ذلك عندما يكون عدد المشاهدات كثيرة، هذا على الرغم من نسبة هذا الطرح وعدم صحته في كثير من الأحيان.

أما ما يقود إلى هذا الاعتقاد هو أن افتراض التوزيع الطبيعي يخلص الدارسين من متاعب كثيرة، إضافة إلى أن معظم الأساليب الإحصائية المستخدمة في الإجابة عن أسئلة العديد من الدراسات أو تحليل بياناتها تتطلب الاعتدالية *Normality* كافتراض رئيس. والمعروف كذلك أن الأساليب الإحصائية التي تستوجب توافر بعض الافتراضات حول التوزيع الاحتمالي لتوزيع البيانات تدعى الأساليب البارامترية (أو المعلمية) *Parametric*، أما تلك التي لا تتطلب توافر ذلك الافتراض فتعرف بالأساليب اللابارامترية (أو اللامعلمية) *Non-Parametric*.

ومنحنى التوزيع الطبيعي يشبه الجرس حيث يتمثل حول المستقيم العمودي على الخط الأفقي المار بنقطة المنتصف، ويتقارب من الصفر عندما يقترب من المالا نهاية من جهته، فالمتوسط الحسابي يقع على خط تماثل المنحنى، فإذا تحرك المتوسط الحسابي يميناً ويساراً فإن شكل

----- الفصل الخامس :مقاييس التشتت

المنحنى لا يتغير ولكن يتغير مركزه، أما إذا تغير الانحراف المعياري فإنه يؤدي إلى تشتت وتباعده في المنحنى وتباعده في المنحنى حول المركز .

ويرجع اكتشاف المنحنى الطبيعي إلى عالم الرياضيات الألماني كارل فريدريك جاويس *Karl F.Gauss* ولذلك يشير كثير من الإحصائيين إلى المنحنى الطبيعي بالمنحنى الجاوسي . وقد خلصت الدراسات النفسية والاجتماعية بعد ذلك إلى أن معظم الصفات الشخصية والنفسية تتوزع توزيعاً طبيعياً، معتمدين على افتراض مفاده أن معظم الظواهر الشخصية التي تتغير تغيراً متصلاً لها توزيعات طبيعية أو قريبة من الطبيعية، وعلى الرغم من شيوع مثل الآراء والافتراضات عن علاقة التوزيع الطبيعي بالظواهر النفسية والاجتماعية التي ندرسها، إلا أن حقيقة الأمر تخالفها وتدحضها في كثير من الأحيان.

وعلى الرغم من كل ما يحيط بهذا الاصطلاح من غموض وإبهام، وبغض النظر عن استعمالاته الخاطئة أو غير الدقيقة، فإنه شائع الاستعمال في الدراسات والبحوث المختلفة، ولهذا السبب فإنه من المستحسن دراسته بشكل وافٍ للتعرف على حقيقة أمره، وعلى الاعتبار التي يقوم على أساسها، وعلى الحالات أو المجالات التي يمكن استعماله فيها.

خصائص التوزيع الطبيعي:

المنحنى الطبيعي هو منحنى توزيع تكراري له قمة واحدة، ويمثل المحور الأفقي قيم المتغير X بينما المحور الرأسي يمثل قيم دالة الاحتمال أو التكرارات النسبية $F(X)$ لهذه القيم، و للمنحنى الطبيعي ست خصائص تميزه عن التوزيعات التكرارية الأخرى وهي :

- 1- تتجمع معظم القيم (الدرجات) في المنحنى الطبيعي حول متوسط التوزيع، حيث تقع قمة التوزيع، ومع زيادة المسافة عن المتوسط(من الجهتين) تقل تكرارات الدرجات

----- الفصل الخامس :مقاييس التشتت

وينحدر المنحنى ليقترّب من المحور الأفقي عند طرفيه. بمعنى أنه وحيد المنوال، أي لا يوجد له سوى منوال واحد.

1- تتساوى في المنحنى الطبيعي قيم النزعة المركزية الثلاث (المتوسط والوسيط والمنوال) حيث تكون في نفس النقطة وهي مركز أو منتصف التوزيع. كما أن توزيع القيم توزيع متصل .

2- المنحنى الطبيعي متماثل *Symmetric* ويقصد بذلك أنه إذا أسقطنا عموداً من قمته إلى المحور الأفقي، فإنه يقسم المنحنى إلى نصفين متطابقين تماماً وتكون مساحة كل قسم مساوية 50% من المساحة الكلية تحت المنحنى.

4- نمطية المساحة : أن المساحة الواقعة تحت المنحنى تساوي (واحداً) أو نسبة (100%) ويحصر المنحنى الطبيعي ما يقارب 68% من المساحة بين انحراف معياري واحد على يمين محور التماثل (المتوسط) وانحراف معياري على يسار ذلك المحور، وعليه تكون المساحة تحت المنحنى بين المتوسط $\pm$ واحد انحراف معياري هي (68%) تقريباً. كما يحصر 95% تقريباً من المساحة بين انحرافين معياريين على يمين المحور وانحرافين على يساره، وعليه تكون المساحة تحت المنحنى بين المتوسط ± 2 انحراف معياري هي 95% تقريباً، وكذلك يحصر 99% تقريباً من المساحة المحصورة بين ثلاثة انحرافات معيارية على يمين المحور وثلاثة انحرافات على يساره وعليه تكون المساحة تحت المنحنى بين المتوسط ± 3 انحراف معياري هي 99% تقريباً.

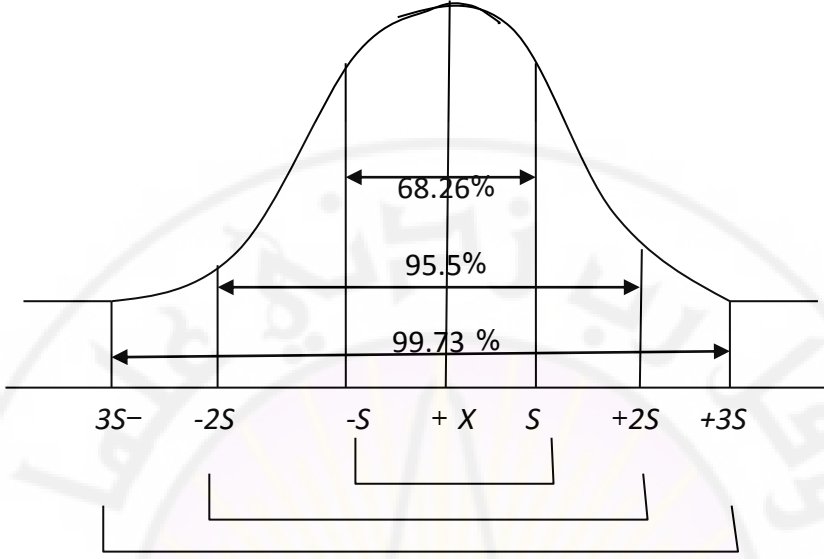
3- طرفا المنحنى الطبيعي متقاربان *Asymptotic* مع المحور الأفقي، بمعنى أنهما لا يمسان المحور الأفقي مهما كان امتداده، أي أن طرفيه غير موازيين ولكن لا يلتقيان مع المحور.

4- من أهم خواص المنحنى الطبيعي أن نقطتي الانقلاب للمنحنى وهما : النقطتان اللتان يتغير عندهما اتجاه المنحنى تقعان على بعد $\pm$ واحد انحراف معياري من المتوسط الحسابي. (فهيمى، 2005 ، ص 269-270 ; أبو زينة، الشايب، 2006، ص 203،202)

المساحات تحت التوزيع الطبيعي :

بما أن أي توزيع طبيعي يحدد بالمتوسط $\bar{X}$ ، والتباين S^2 ، فإن المساحة لأي فترة تحت التوزيع الطبيعي يعتمد على $\bar{X}$ ، S^2 وبالتالي لا يمكن وضع جداول لجميع قيم X ، S^2 ، ولذلك نعتد في حساب المساحات تحت التوزيع الطبيعي على تحويله إلى توزيع طبيعي معياري، ومن ثم نجد المساحة المطلوبة من جداول التوزيع الطبيعي المعياري ويلاحظ من الشكل الموصوف أعلاه ما يلي :

- 1- المساحة الواقعة ضمن انحراف معياري واحد عن المتوسط ($S \pm \bar{X}$) تشكل حوالي 68% من المساحة الكلية .
- 2- المساحة الواقعة ضمن انحرافين معياريين واحد عن المتوسط ($2S \pm \bar{X}$) تشكل حوالي 95% من المساحة الكلية .
- 3- المساحة الواقعة ضمن ثلاثة انحرافات معيارية واحد عن المتوسط ($3S \pm \bar{X}$) تشكل حوالي 99% من المساحة الكلية .



الشكل (5-1)

مثال (4):

لو كان لدينا درجة معيارية ($Z=1.83$) مثلاً، وأردنا أن نجد المساحة التي تقع على يسار هذه النقطة، فإن علينا الرجوع إلى جدول القيم المعيارية ثم :

- نحدد القيمة ($Z=1.8$) (دون الكسر المئوي 0.03) على حقل الدرجات المعيارية الواقع في يسار الجدول ونقف عندها عمودياً.
- ثم نتحرك من هذه النقطة (1.8) أفقياً حتى نصل إلى الحقل الذي يقع تحت الكسر المئوي (0.03) فنقف عند نقطة التقاء الحقتين (العمودي والأفقي).

----- الفصل الخامس :مقاييس التشتت

- نجد أن المساحة التي يشير لها هذا الالتقاء في الجدول المعيارية هي 0.4664 أو 46.64%.
- هذه المساحة هي المساحة المحصورة بين الدرجة المعيارية المطلوبة (1.83) وخط تنصيف المنحنى (عند نقطة الصفر).
- ولأننا نبحث عن كل المساحة التي تقع عن يسار الدرجة المعيارية (1.8) وهي موجبة أي تقع عن يمين خط التنصيف فإننا يجب أن نضيف إليها مساحة نصف المنحنى الواقع عن يسار خط التنصيف، وهو يساوي بالطبع - نصف المساحة الكلية أو (50%).
- وبذلك تصبح المساحة الكلية للدرجة المعيارية (1.83) والواقعة عن يسارها مساوية $46.64+50=96.64\%$
- أما إذا أردنا إيجاد المساحة الكلية الواقعة عن يمين الدرجة المعيارية (1.83) فإنه يمكن استخراجها مباشرة من طرح المساحة الواقعة عن يسارها من المساحة الكلية للمنحنى (1) : $1-96.64=3.36\%$
- أما إذا كانت القيمة المعيارية (سالبة) مثلاً، ولتكن (-1.83) فإننا نستخرج المساحة بذات الأسلوب السابقة وهي % 46.64 ولكننا في هذه الحالة نطرحها من 50 بدلاً من جمعها عندما كانت القيمة المعيارية موجبة .
- أي أن المساحة التي ستقع على يسار القيمة المعيارية $Z=-1.83$ ستكون مساوية إلى $50-46.64=3.36\%$

----- الفصل الخامس :مقاييس التشتت

تمارين وأسئلة :

احسب كلا من

أ-مدى التوزيع

ب- نصف المدى الربيعي للتوزيع التالي:

الجدول (20-5)

الفئة	30 34	34 38	38 42	42 46	46 50
التكرار	1	2	5	4	3

2- احسب كلا من

أ- التباين

ب- الانحراف المعياري

ج- معامل الاختلاف

3- جرى حساب عدد الناجحين في 200 مدرسة ولخصت النتائج كما يلي:

$$\sum X = 920 \quad \sum X^2 = 5032 \quad \text{حيث تشير } X \text{ إلى عدد الناجحين في المدرسة الواحدة}$$

1- احسب المتوسط الحسابي لعدد الناجحين في المدرسة الواحدة

2- حساب الانحراف المعياري.

الفصل السادس

الارتباط

Correlation



معنى الارتباط:

يجب التنبيه إلى أن الارتباط بين متغيرين معناه أن التغير في أحدهما يكون عادة مصحوباً بتغير في الآخر، وأنه توجد علاقة معينة بين اتجاهي التغير فيهما إيجاباً وسلباً .

كما يجب التنبيه أيضاً إلى أن الارتباط بين ظاهرتين متغيرتين ليس دليلاً على أن أحدهما نتيجة للأخرى، أو أن التغير في واحدة تابع للتغير في الأخرى ولا ينشأ إلا بسببه، بل هو يشير فقط إلى احتمال وجود هذه العلاقة . لأن هذه العلاقة ما هي إلا نوع خاص من أنواع العلاقات التي يدل الارتباط على وجودها. وهذه الأنواع المختلفة للعلاقات تتمثل فيما يأتي :

- 1- العلاقة السببية المباشرة أي أن يكون أحد المتغيرين نتيجة مباشرة للمتغير الآخر كالعلاقة بين نظام معين للتعزير وكفاءة التعلم.
- 2- العلاقة السببية غير المباشرة، كأن يكون أحد المتغيرين سبب غير مباشر للثاني يؤثر فيه بواسطة متغير ثالث أو أكثر كالعلاقة بين الطول والوزن في بحوث النمو، فهذه العلاقة تنشأ عن متغير ثالث هو العمر الزمني أو الصحة الجسمية.
- 3- أن يؤثر عامل واحد في المتغيرين معاً، وفي ذلك يكون كل من المتغيرين المرتبطين نتيجة عامل آخر ثالث مشترك بينهما يؤثر فيهما في وقت واحد فيكون التغير في أحدهما مصحوباً بالتغير في الآخر. مثال ذلك الارتباط بين أسعار سلعتين تمتلكها طبقة معينة من السكان فإن أسعارهما تكون مرهونة بالحالة الاقتصادية لهذه الطبقة، أو ارتباط أسعار السلعتين بأسعار النقل.
- 4- أن تكون بعض العوامل مشتركة بين المتغيرين ومن ذلك مثلاً لو اخترنا عدداً من التلاميذ في مادتين مثل جغرافيا العالم الإسلامي وتاريخ العالم الإسلامي فإننا نجد

----- الفصل السادس :الارتباط

الارتباط شديداً بين درجات هاتين المادتين والسبب في ذلك أن الأداء في الاختبارين يوجد فيه بعض العوامل المشتركة .

تعريف الارتباط:

الارتباط في معناه العلمي الدقيق هو التغير الاقتراني، أو بمعنى آخر هو النزعة إلى اقتران التغير في ظاهرة بالتغير في ظاهرة أخرى، ويقاس الارتباط بمعامل الارتباط.

يأخذ معامل الارتباط قيماً تتراوح بين $+1$ و -1 مروراً بالصفر . معامل الارتباط $+1$ يعني وجود ارتباط موجب، أي أن الشخص الذي حاز على المرتبة الأولى في الاختبار الأول حصل على المرتبة الأولى في الاختبار الثاني .

-1 معامل ارتباط سالب، أي أن الشخص الذي حصل على المرتبة الأولى في الاختبار الأول حصل على المرتبة الأخيرة في الاختبار الثاني، والشخص الذي حصل على الثانية في الاختبار الأول كان ترتيبه الثاني قبل الأخير في الاختبار الثاني وهكذا ...

أما الصفر فيعني عدم وجود علاقة بين المتغيرين.

يعرف الارتباط على أنه علاقة بين متغيرين أو أكثر . وبالرغم من أن الحد الأقصى لعدد المتغيرات ترك مفتوحاً في هذا التعريف، إلا أن هذا العدد يفرض نفسه من الناحية العملية، ولكن ليس من السهل توضيح هذه الفكرة هنا، إلا أن بعض الملامح ستتضح في فصل الانحدار والتنبيؤ . وبناء على عدد المتغيرات التي تدخل في حساب الارتباط فإنه يمكن حصر أنواع الارتباط بما يلي :

أنواع الارتباط:

1- الارتباط الثنائي حيث يندرج تحته أيضاً نوعان هما :

----- الفصل السادس :الارتباط

أ- الارتباط البسيط Simple corrolation: ويبحث في العلاقة بين متغيرين بصرف النظر عن الأثر الناتج عن أي متغير آخر له علاقة بأي من المتغيرين على أي منهما أو كليهما. ويدرس العلاقة بين متغير واحد تابع ومتغير واحد مستقل مع افتراض عدم وجود متغيرات أخرى وتمثل المعادلة على شكل خط مستقيم أي معادلة من الدرجة الأولى .

ب- الارتباط الجزئي Partial corrolation. ويبحث في العلاقة بين متغيرين مع الأخذ بالاعتبار الأثر الناتج عن ارتباط متغير أو متغيرات أخرى على أي من المتغيرين أو كليهما. فإذا أزيل هذا الأثر من أحد المتغيرين فقط، فإن الارتباط بينهما يسمى ارتباطاً جزئياً part أو شبه جزئي Semi-partial أما إذا أزيل هذا الأثر من كل منهما، فإن الارتباط بينهما هو ارتباط جزئي partial ولعل أبسط هذا النوع من الارتباط (أي الارتباط الجزئي من نوع Partial) هو الارتباط الجزئي من الدرجة الأولى First order part or partial corr وهو الارتباط بين متغيرين عند إزالة أثر أي منهما أو كليهما .

2- الارتباط المتعدد multiple Corrolation:

1-" ارتباط متغير بعدة متغيرات:

وهو الارتباط الذي يبحث في العلاقة بين متغير من جهة ومجموعة متغيرات (اثان أو أكثر) من جهة أخرى . وأبسطها هو الارتباط بين متغير ما من جهة ومتغيرين معاً من جهة أخرى. أو ما يشار إليه بالارتباط المتعدد من الدرجة الأولى.

----- الفصل السادس :الارتباط

وهو يدرس العلاقة بين متغير تابع وعدد من المتغيرات المستقلة وتدرس جملة واحدة ويمثل بمعادلة من الدرجة الثانية فأكثر ، أي خط منحنى أو قطع كامل أو قطع ناقص.(عودة والخليلي، 1988، ص135)

2- الارتباط القانوني :

ينصب هذا النوع من الارتباط على قياس العلاقة بين عدة متغيرات من جهة (تحصيل الطالب في العلوم العامة) وعدة متغيرات من جهة أخرى (تحصيل الطالب في المواد الاجتماعية) (الجضي،2005،ص190)

مؤشرات الارتباط :

1- معامل الارتباط:

يجري الكشف عن درجة الارتباط بين متغيرين بواسطة مؤشرين هما شكل الانتشار، ومعامل الارتباط وتتراوح قيمة معامل الارتباط من +1 إلى -1 حيث تدل القيمة المطلقة على قوة الارتباط، وتدل الإشارة على اتجاه العلاقة. أي يمكن أن يكون الارتباط سالباً (عكسياً) أو موجباً (طردياً). يقصد بالارتباط السالب نقصان قيم أحد المتغيرين مع زيادة قيم المتغير الآخر (عودة، الخليلي، 1988، ص136) .

2- شكل الانتشار scattar plot :

يمكن التعبير عن العلاقة بين المتغيرين باستخدام ما يسمى شكل الانتشار scattar plot ويمكن الحصول عليه مباشرة من جدول المتغيرين لنفس المفحوصين كما هو الحال في الجدول التالي الذي يمثل درجات 10 أطفال في كل من العمر والذكاء

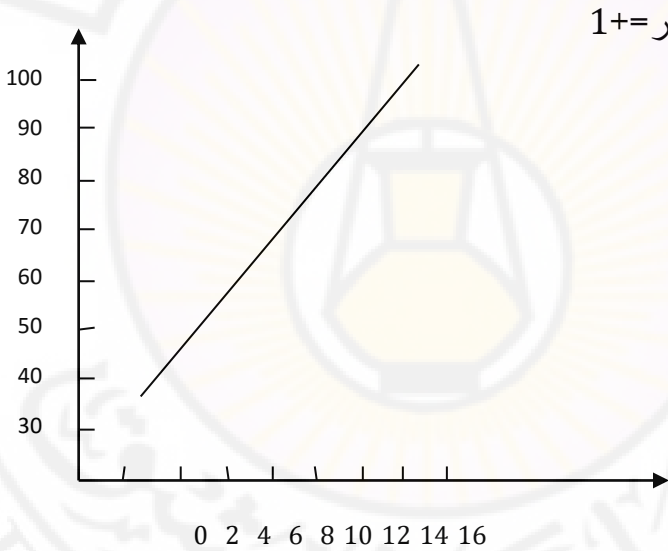
----- الفصل السادس :الارتباط

الجدول (1-6)

درجات 10 أطفال في كل من العمر والذكاء

الأطفال	1	2	3	4	5	6	7	8	9	10	التعبير عن المتغير بالرمز
العمر	2	4	5	6	7	8	9	10	12	13	x
الذكاء	40	50	55	60	65	70	75	80	90	95	y

ويمكن تحويل هذا الجدول إلى شكل انتشار كما هو الحال في الشكل التالي:



الشكل (1-6) علاقة موجبة كاملة بين متغيرين

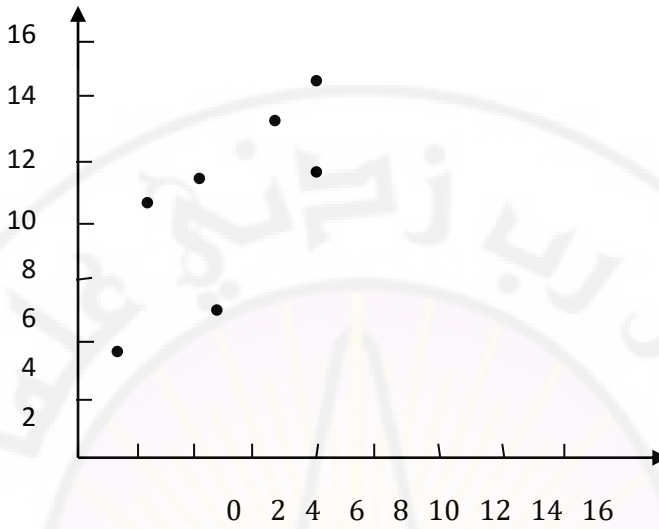
يلاحظ من بيانات الجدول والرسم البياني السابقين أنه بالنسبة لكل من زوج من الدرجات التي حصل عليها كل طفل (في كل من العمر والذكاء) أن كل درجة حصل عليها الطفل في المتغير (y) أي الذكاء تزيد عن درجته في المتغير (x) أي العمر بنسبة ثابتة. فإذا

----- الفصل السادس :الارتباط

زاد عمر الطفل سنة واحدة يناظر ذلك زيادة في الذكاء مقدارها 5 درجات وكذلك النقص . ولهذا حصلنا على شكل خط مستقيم كامل، وهذا الخط يدل على علاقة موجبة كاملة تدل على أن الزيادة في المتغير x تناظرها زيادة بنفس النسبة في المتغير (y) والنقص في المتغير (x) يناظره نقص بنفس النسبة في المتغير (y) وهذه العلاقة يعبر عنها بمعامل الارتباط $+1$ وهو أقصى معامل يمكن الحصول عليه، فلا يمكن لمعامل الارتباط أن يتجاوز الواحد الصحيح سلباً أو إيجاباً، ويعبر عن علاقة حتمية، أي أن المفحوص الذي يكون له عمر معين فإن من المحتم أن يكون ذكاؤه محدداً . وكذلك يمكن استنتاج العمر من الذكاء . إلا أنه تجدر الإشارة أن هذا المثال هو مثال افتراضي ويندر وجوده في العلوم الإنسانية والاجتماعية، فمعاملات الارتباط دائماً فيها أقل من الواحد الصحيح.

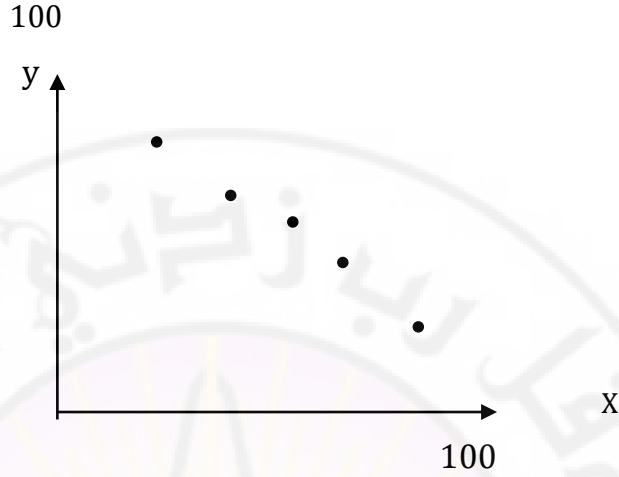
والآن تأمل الشكل (2-6) الذي يمثل شكل الانتشار لمتغيرين أحدهما درجات 10 تلاميذ في اختبار في الحساب (X) والثاني درجات اختبار في اللغة العربية (y). فإذا نستنتج من هذا الرسم أن الشكل بالطبع ليس خطاً مستقيماً كما هو الحال في الشكل السابق. ومع ذلك نلاحظ بصفة عامة أن الشخص الذي يحصل على درجة عالية في المتغير x أي الحساب يحتمل أن يحصل على درجة عالية في المتغير y أي اللغة العربية، كما أن الشخص الذي يحصل على درجة منخفضة في x يحتمل أن يحصل على درجة منخفضة في y أيضاً . وهذا النوع من العلاقات يسمى العلاقة الجزئية الموجبة، وحين تحسب كمعامل ارتباط يكون مقدارها كسراً عشرياً أعلى من الصفر وأقل من الواحد الصحيح وتكون إشارته الجبرية موجبة .

----- الفصل السادس :الارتباط



الشكل (2-6): علاقة موجبة جزئية بين متغيرين

وبالتأكيد يمكن للعلاقات أن تكون سالبة، فمعامل الارتباط السالب التام أو الكامل -1 يعبر عن علاقة حتمية بين المتغيرين في اتجاه عكس العلاقة الموجبة التامة. وهو نادر الحدوث في العلم عادة، إن لم يكن مستحيلاً في العلوم الإنسانية والاجتماعية. ولعل أقرب الأمثلة التي توضحه العلاقة بين حجم الغاز x وضغطه y ، فمن المعروف أن الزيادة في حجم الغاز x تؤدي إلى قلة الضغط y ، والنقص في الحجم يؤدي إلى زيادة الضغط. ويمكن التعبير عن هذه العلاقة في الشكل التالي :

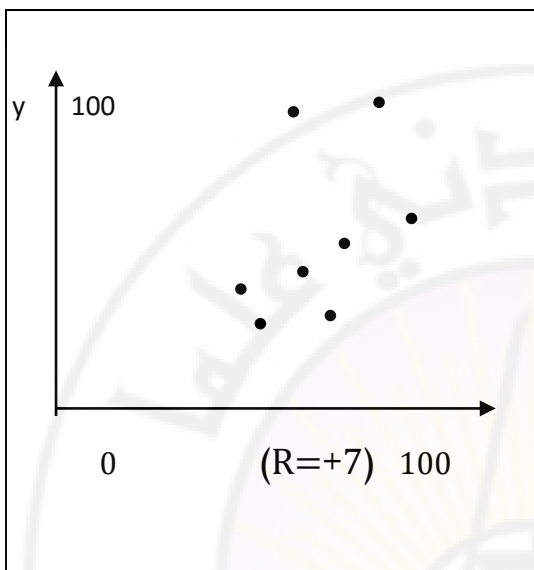


الشكل (3-6): علاقة سالبة كاملة

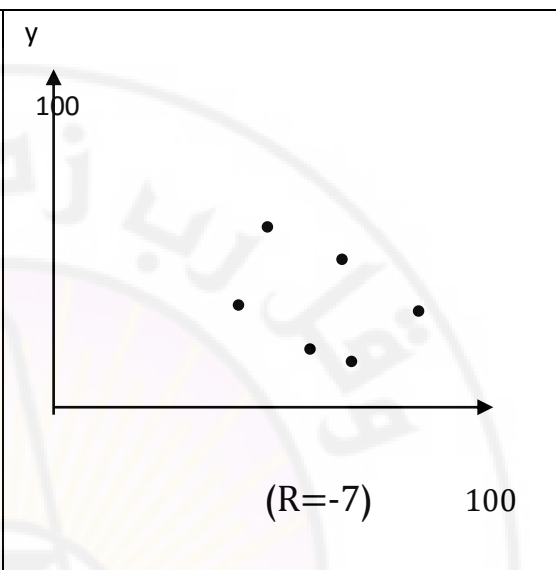
أما العلاقات الجزئية فقد تكون موجبة أو سالبة، وقد تكون كبيرة أو صغيرة، ويتحدد اتجاه العلاقة بالإشارة الجبرية لمعامل الارتباط عند حسابه، أما حجم العلاقة فيتحدد في ضوء مقدار الكسر العشري المحسوب، وبالطبع فإنه كلما ابتعدت العلاقة عن الصفر واقتربت من الواحد الصحيح تكون أكبر. ويوضح الشكل التالي علاقات جزئية موجبة وسالبة، كبيرة وصغيرة

----- الفصل السادس: الارتباط

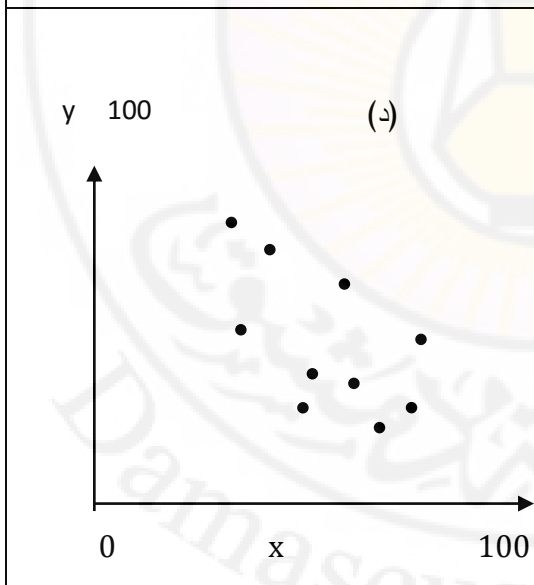
(أ)



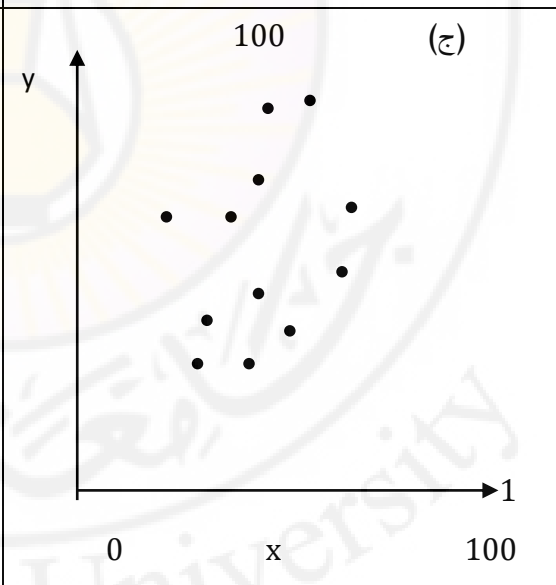
(ب)



(د)



(ج)



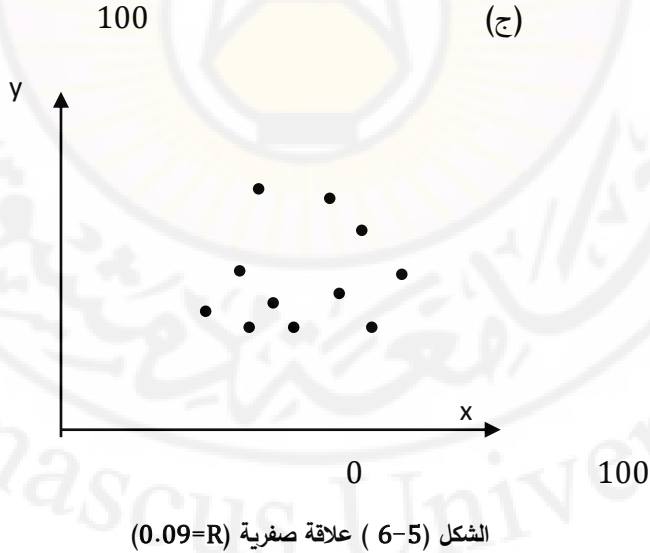
الشكل (4-6): علاقات جزئية

----- الفصل السادس :الارتباط

الشكل (أ) علاقات جزئية مختلفة (أ) علاقة جزئية موجبة كبيرة ($R=+7$) (ب) علاقة سالبة كبيرة ($R=-7$)، (ج) علاقة جزئية موجبة صغيرة ($R=+2$)، (د) علاقة جزئية سالبة صغيرة ($R=-2$)

وهذه العلاقات الجزئية هي الأكثر شيوعاً في البحوث العلمية عامة، والبحوث النفسية والتربوية والاجتماعية خاصة.

وقد تكون العلاقة بين المتغيرين صفيرية، والتي تعني عدم وجود علاقة بين المتغيرين ومن أمثلة ذلك العلاقة بين طول القامة والذكاء . والصفر هنا لا يقصد به الصفر بمعناه الرياضي المعتاد، وإنما يدل على أن معامل الارتباط ليست له دلالة إحصائية مهما كان مقداره. وحسبنا أن ننتبه هنا إلى خطأ فادح شائع في كثير من البحوث الحديثة حين يفسر بعض الباحثين معامل الارتباط بأنه موجب أو سالب بينما هو غير دال إحصائياً . إن معامل الارتباط في هذه الحالة ليس إلا صفراً، ولا يحمل معنى العلاقة الموجبة أو السالبة الحال (أبو حطب وصادق، 1991، ص245-251).



خلاصة لأنواع العلاقات :

في ضوء ما سبق يمكن تصنيف العلاقات في ضوء الشكل الآتي :

1+ علاقة موجبة كاملة (محيط الدائرة وقطرها)

العلاقة بين الذكاء والتحصيل	{	R0.9
		R0.8
		R0.7
		R0.6
		R0.5
		R0.4
		R0.3
		R0.2
		R0.1
		صفر لا علاقة (طول الجسم والذكاء)
		R0.1-
		R0.2-
		R0.3-
		R0.4-
		R0.5-
		R0.6-
		R0.7-
العلاقة بين القلق والتحصيل	{	R0.8-
		R0.9-
1- علاقة سالبة كاملة (حجم الغاز وضغطه)		

الشكل (6-6) أنواع العلاقات المختلفة بين المتغيرات

شروط حساب معامل الارتباط :

هناك شروط عديدة يجب أن تتوفر بين الظاهرتين المدروستين حتى نستطيع أن ندرس علاقتهما المشتركة وهذه الشروط هي :

- أن تكون بين الظاهرتين المدروستين علاقة جدلية واضحة.
- أن تكون كل من الظاهرتين قابلتين للقياس بواسطة وحدة قياس معينة لكل منهما (مستقلة أو مشتركة).
- أن تكون القياسات المأخوذة متقابلة من حيث الزمان أو المكان أو كلاهما معاً .

(العلي، 1998 ،ص7)

طرائق حساب معامل الارتباط:

ويمكن استخدام معامل الارتباط بين متغيرين بطرائق عديدة نذكر منها:

1. **معامل بيرسون (Pearson):** يستخدم إذا كان كلا المتغيرين مقاسا بمقياس كمي مثل إيجاد معامل الارتباط بين الدخل والإنفاق.
2. **معامل سبيرمان (Spearman):** يستخدم إذا كان كلا من المتغيرين مقاسا بمقياس ترتيبي مثل إيجاد علاقة المستوى التحصيلي (مرتفع - متوسط - منخفض) والمستوى التعليمي للألم (ابتدائي - إعدادي - ثانوي- جامعة - دراسات عليا) كما يمكن استخدام مقياس سبيرمان في حالة المتغيرات الكمية أيضا.

معامل ارتباط بيرسون:

يستخدم معامل ارتباط بيرسون لدراسة العلاقة بين متغيرين كميين. ولحساب هذا المعامل لا بد من توفر قيمة لكل متغير عند نفس الفرد، بمعنى أن هذه القيم عبارة عن قيم مزدوجة وهذا يعني أيضاً تساوي عدد هذه القيم في المتغيرين.

نعرض فيما يلي للطرائق المختلفة لحساب معامل الارتباط، وهي جميعاً مشتقة من

المعادلة الأساسية :

----- الفصل السادس :الارتباط

أولاً : حساب معامل الارتباط باستخدام المعادلة الأساسية :

المعادلة الأساسية لمعامل الارتباط تعتمد على متوسط حاصل ضرب الدرجات المعيارية المتناظرة، وتصاغ على النحو الآتي :

$$R = \frac{\sum Z_x \cdot Z_y}{n}$$

ويوضح المثال التالي هذه الطريقة، وفيه درجات 10 تلاميذ في اختبارين أحدهما يقيس فهم القراءة x والثاني يقيس الذكاء اللغوي y وفيه تم تحويل الدرجات الخام في الاختبارين إلى درجات معيارية حيث أن $x = 7.5$ ، $\bar{S}_x = 3.528$ ، $y = 8$ ، $\bar{S}_y = 3.795$.

الجدول (2-6) حساب معامل الارتباط بالمعادلة الأساسية

(متوسط حاصل ضرب الدرجات المعيارية المتناظرة)

x	y	Z_x	Z_y	$Z_x \cdot Z_y$
2	1	1.56-	1.84-	2.87
3	6	1.27-	0.52-	0.66
5	7	0.71-	0.26-	0.18
6	3	0.42-	1.32-	0.55
6	11	0.42-	0.79+	-0.33
8	9	0.14+	0.26+	0.04
10	7	0.71+	0.26-	-0.18
10	11	0.71+	0.79+	0.56
12	14	1.27+	1.58+	2.01
13	11	1.56+	0.79+	1.23
$\sum x = 75$	$\sum y = 80$			$\sum Z_x \cdot Z_y = 7.58$

وبتطبيق المعادلة السابقة يصبح معامل الارتباط

$$R = \frac{7.58}{10} = 0.758$$

الفصل السادس: الارتباط

ثانياً: حساب معامل الارتباط باستخدام انحرافات المتغيرين عن متوسطيهما :
يمكن الوصول إلى مزيد من التبسيط في إجراءات حساب الارتباط وذلك بالتخلص نهائياً من استخدام الانحراف المعياري في المعادلة والاعتماد كلياً على الانحرافات عن المتوسط ومربعات هذه الانحرافات وفيما يلي المعادلة :

$$R = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sqrt{\sum (x - \bar{x})^2 \sum (y - \bar{y})^2}}$$

مثال: احسب معامل الارتباط بين عدد ساعات الدراسة والتحصيل الدراسي
الجدول (3-6)

x	y	(x - $\bar{x}$)	(y - $\bar{y}$)	(x - $\bar{x}$)(y - $\bar{y}$)	(x - $\bar{x}$) ²	(y - $\bar{y}$) ²
2	3	3.6-	2.4-	8.64	12.96	5.76
4	5	1.6-	0.4-	0.64	2.56	0.16
5	6	0.6-	0.6	0.36-	0.36	0.36
7	5	1.4	0.4-	0.56-	1.96	0.16
10	8	4.4	2.6	11.44	19.36	6.76
$\Sigma=28$	27	0	0	21.64	37.2	13.2

$$\bar{X} = \frac{\sum x}{n} = \frac{28}{5} = 5.6$$

$$\bar{y} = \frac{\sum y}{n} = \frac{27}{5} = 5.4$$

$$R = \frac{21.64}{22.159} = 0.976$$

التفسير : الارتباط بين ساعات الدراسة والتحصيل شديد وطردي .

شديد لأنه أكبر من 0.5 وطردي لأنه موجب .

رابعاً: الطريقة العامة:

يمكن حساب معامل الارتباط من القيم مباشرة وذلك بتربيع قيمة المتغيرين وحساب حاصل جمع ضرب المتغيرين ببعضهما . ويحسب معامل الارتباط وفق الصيغة التالية:

----- الفصل السادس : الارتباط

$$R = \frac{n\sum xy - \sum x \sum y}{\sqrt{[n\sum x^2 - (\sum x)^2][n\sum y^2 - (\sum y)^2]}}$$

مثال :

احسب معامل الارتباط بين علامات الطلبة في اختبار القدرة العددية واختبار القدرة اللغوية

الجدول (4-6)

x	y	x ²	y ²	xy
20	12	400	144	240
18	16	324	256	286
16	10	256	100	160
15	14	225	196	210
14	12	196	144	168
12	10	144	100	120
12	9	144	81	108
10	8	100	64	80
8	7	64	49	56
5	2	25	4	10
$\sum x = 130$	$\sum y = 100$	$\sum x^2 = 1878$	$\sum y^2 = 1138$	$\sum xy = 1440$

$$\bar{X} = \frac{130}{10} = 13$$

$$\bar{y} = \frac{100}{10} = 10$$

$$R = \frac{n\sum xy - \sum x \sum y}{\sqrt{[n\sum x^2 - (\sum x)^2][n\sum y^2 - (\sum y)^2]}}$$

----- الفصل السادس : الارتباط

$$R = \frac{10 \times 1440 - 100 \times 130}{\sqrt{[10 \times 1878 - (130)^2][10 \times 1138 - (100)^2]}}$$

$$R = \frac{14400 - 13000}{\sqrt{(1880)(1380)}}$$

$$R = \frac{1400}{1610.7}$$

$$R = 0.87$$

مثال :

الجدول (5-6)

x	y	x ²	y ²	xy
10	7	100	49	70
5	8	25	64	40
3	4	9	16	12
4	3	16	9	12
7	6	49	36	42
8	9	64	81	72
37	37	263	255	248

$$R = \frac{n \sum xy - \sum x \sum y}{\sqrt{[n \sum x^2 - (\sum x)^2][n \sum y^2 - (\sum y)^2]}}$$

$$R = \frac{6 \times 248 - 37 \times 37}{\sqrt{[6 \times 263 - (37)^2][6 \times 255 - (37)^2]}}$$

----- الفصل السادس :الارتباط

$$R = \frac{1488 - 1369}{\sqrt{(209)(161)}}$$

$$R = \frac{119}{183.43} = 0.648 \approx 0.65$$

ارتباط متوسط الشدة

معامل ارتباط سبيرمان (لترتب):

يهدف معامل ارتباط الرتب لسبيرمان إلى قياس التغير الاقتراني القائم بين ترتيب الأفراد أو الأشياء إلى صفة وترتيبهم لصفة أخرى. وبمقارنة كفاءة هذا المعامل مع كفاءة معامل بيرسون فإنه يقترب من 91%، معنى ذلك أنه عند استخدام معامل سبيرمان بدلاً عن معامل بيرسون عند تحقق شروط استخدام بيرسون وهي اعتدالية التوزيع وكون المتغير هو من المستوى الفئوي، فإن معامل ارتباط سبيرمان يكون 91% من معامل بيرسون، فإذا وجد ارتباط بين X, Y في ذلك المجتمع مع 100 حالة، فإن معامل ارتباط سبيرمان سيوضح ذلك الارتباط عند مستوى الدلالة التي يحققها معامل ارتباط بيرسون مع 91% حالة (حبيب، 2000، ص 635) .

هذا ويحسب معامل ارتباط سبيرمان وفق الصيغة التالية:

$$R_s = 1 - \frac{6\sum D^2}{n(n^2 - 1)}$$

مثال:

الجدول (6-6)

x	y	رتبة x	رتبة y	فرق الرتب D	D^2
5	3	6	7	1-	1
4	6	7	4	3	9
7	7	3.5	2	1.5	2.25
10	6	1	4	3-	9
7	5	3.5	6	2.5-	6.25
6	9	5	1	4	16
8	6	2	4	2-	4
					47.5

نرتب القيم المفردة الأكبر تأخذ المرتبة الأولى وهكذا. أما إذا تساوت قيمتان في الرتبة فإنهما تأخذان مرتبتين متتاليتين وذلك بجمع المرتبتين وتقسيمهما على 2 فنحصل على رتبة كل منهما:

$$R_s = 1 - \frac{6 \sum D^2}{n(n^2 - 1)}$$

$$R_s = 1 - \frac{6 \times 47.5}{7(49 - 1)} = 1 - \frac{285}{336}$$

$$= 1 - 0.84 = 0.16$$

التفسير : الارتباط بين رتب x ورتب y ضعيف وطردي وموجب .

مثال :

الجدول (6-7)

x	y	رتبة x	رتبة y	D	D ²
10	7	1	3	2-	4
5	8	4	2	2	4
3	4	6	5	1	1
4	3	5	6	1-	1
7	6	3	4	1-	1
8	9	2	1	1	1
					12

$$R_s = 1 - \frac{6 \sum D^2}{n(n^2 - 1)}$$

$$R_s = 1 - \frac{6 \times 12}{6(36 - 1)}$$

$$= 1 - 0.34 = 0.657 \approx 0.66$$

ارتباط متوسط الشدة

مثال :

قام أحد المحكمين بتقييم عينة تتكون من 11 طالباً في مادتي الجغرافية والتاريخ، وكانت نتائج التقييم كما يلي، والمطلوب إيجاد العلاقة بين أداء الطالب في المادتين باستخدام معامل ارتباط

الرتب R_s

الجدول (8-6)

التسلسل	تقييم مادة الجغرافية X	تقييم التاريخ Y
1	جيد	ضعيف جداً
2	ضعيف	ممتاز
3	مقبول	ممتاز
4	جيد	جيد
5	ممتاز	مقبول
6	مقبول	جيد جداً
7	ضعيف جداً	مقبول
8	جيد جداً	ضعيف
9	ممتاز	جيد
10	ضعيف	ضعيف جداً
11	جيد جداً	مقبول

الحل :

أ- نرتب قيم كل من المتغيرين Y, X للعينة من الأدنى إلى الأعلى، فتكون رتبة كل طالب حسب مستوى تقييمه في مادة الجغرافية X تصاعدياً، فتعطي الطالب الذي حاز على تقييم ضعيف جداً الرتبة 1، والطالب الذي تسلسله 2 والذي تسلسله 10 اللذان حازا على تقييم ضعيف الرتب 2.5 على اعتبار $2.5 = \frac{2+3}{2}$. والرتبة التي تلي ذلك هي للطلبة الذين حازوا على تقييم مقبول وهما اللذان تسلسلها 3 و 6 فتكون الرتبة 4.5 ثم الطالب الذي يحمل تسلسل 4 وتقييمه جيد ورتبته 6، وهكذا ونفس الإجراءات يتم تطبيقها مع تقييم مادة

----- الفصل السادس :الارتباط

التاريخ Y ، ومن ثم نحسب الفرق D بين X و Y لكل مشاهدة، وإيجاد $\sum D$ وكذلك $\sum D^2$ فيكون لدينا :

الجدول (9-6)

D^2	D	y رتبة	X رتبة
25	5	1.5	6.5
64	-8	10.5	2.5
36	-6	10.5	4.5
1	-1	7.5	6.5
30.25	5.5	5	10.5
20.25	-4.5	9	4.5
16	-4	5	1
30.25	5.5	3	8.5
9	3	7.5	10.5
1	1	1.5	2.5
12.25	3.5	5	8.5
$\sum D^2 = 245$	$\sum D = 0$		

ب- نطبق الصيغة فنحصل على :

$$R_s = 1 - \frac{6 \sum D_i^2}{n(n^2 - 1)}$$

$$= 1 - \frac{6(245)}{11(120)} = -0.114$$

أي إن العلاقة بين نتيجتي التقييم سالبة وضعيفة .

مثال :حالة القيم الكمية : استخدم المعطيات التالية التي تمثل معدل استهلاك الفرد من أحد أنواع الأطعمة (كغ/سنة) y ، ومعدل دخله السنوي بالليرة X لعينة عددها 10 أشخاص لاحتساب معامل ارتباط الرتب R_s :

الجدول (6-11)

معدل دخل الفرد y	معدل استهلاك الفرد x	التسلسل
541	1.7	1
541	13.4	2
530	11.2	3
595	11.7	4
826	11.6	5
996	12.5	6
569	12.7	7
463	11.1	8
411	12.6	9
398	13.5	10

الحل :

نتبع نفس الخطوات في المثال السابق في ترتيب المعطيات فيكون لدينا :

الجدول (6-10)

x	y	D	D^2
5.5	4.5	1	1
5.5	9	-3.5	12.25
4	2	2	4
8	4.5	3.5	12.25
9	3	6	36
10	6	4	16
7	8	-1	1
3	1	2	4
2	7	-5	25
1	10	-9	81
		$\sum D = 0$	$\sum D^2 = 192.5$

نطبق الصيغة فنحصل على :

$$R_s = 1 - \frac{6 \sum D_i^2}{n(n^2 - 1)}$$

$$= 1 - \frac{6(192.5)}{10(99)} = -0.167$$

أي إن العلاقة بين معدل الاستهلاك ومعدل دخل الفرد سالبة وضعيفة .

الأهمية الإحصائية لمعامل الارتباط:

للتأكد من الأهمية الإحصائية لمعامل الارتباط نحسب الخطأ المعياري لمعامل الارتباط S_r ونتأكد من تحقق العلاقة :

$$r > 3 S_r$$

ليكون له أهمية إحصائية أو مدلول إحصائي .

$$S_r = \sqrt{\frac{1-r^2}{n-2}}$$

$$S_r = \sqrt{\frac{1-(0.893)^2}{5-2}}$$

$$S_r = 0.265$$

$$r = 0.976 > 3(0.260)$$

$$0.893 > 0.78$$

تحققت العلاقة، إذن لمعامل الارتباط أهمية إحصائية .

أي أن معامل ارتباط المجتمع الذي سحبت منه العينة لا يساوي الصفر والعينة تمثل مجتمعها الإحصائي، وإذا لم تتحقق العلاقة يمكن أن يكون معامل ارتباط المجتمع مساوياً للصفر والعينة

----- الفصل السادس :الارتباط

لا تمثل مجتمعها الإحصائي وقد يكون الارتباط زائفاً . والارتباط الزائف هو الارتباط الذي تنتقي فيه العلاقة السببية بين المتغيرين .

معامل التحديد $0.950 = (R^2) = (0.976)$ أي أن 0.95 من التباين في المتغير التابع يحدده التباين في المتغير المستقل .

----- الفصل السادس: الارتباط

أسئلة وتمارين محلولة:

1- معامل الارتباط بين متغيرين بإزالة أثر متغير ثالث من إحداهما يسمى ارتباط

.....

2- معامل الارتباط سبيرمان رو هو معامل ارتباط بين متغيرين كل منهما بمستوى

القياس.....

3- أقل قيمة يمكن أن يأخذها معامل الارتباط تساوي

4- إذا وقعت جميع النقاط في شكل الانتشار على خط مستقيم واحد يسمى ارتباط

.....

5- إذا كان معامل الارتباط بين متغيرين $= 0.70$ فإن معامل التحديد يساوي

الأجوبة

(1) شبه الجزئي

(2) الرتبى

(3) 1-

(4) تام

(5) 0.49

----- الفصل السادس :الارتباط

6 - قام أحد الباحثين بدراسة العلاقة بين درجات ثمانية أفراد في الإنتاجية ودرجاتهم في الرضا المهني فكانت النتائج :

الجدول (6-11)

	1	2	3	4	5	6	7	8
X الإنتاجية	10	20	30	40	50	60	70	80
Y الرضا المهني	21	40	62	78	100	122	135	164

والمطلوب دراسة العلاقة الارتباطية بين الإنتاجية والرضا المهني :

الجدول (6-12)

حساب معامل الارتباط

x	y	x^2	y^2	$x.y$
10	21	100	441	210
20	40	400	1600	800
30	62	900	3844	1860
40	78	1600	6084	3120
50	100	2500	10000	5000
60	122	3600	14884	7320
70	135	4900	18225	9450
		6400	26896	13120

----- الفصل السادس :الارتباط

$$R = \frac{n \sum x_i y_i - (\sum x_i) (\sum y_i)}{\sqrt{[n \sum x_i^2 - (\sum x_i)^2][n \sum y_i^2 - (\sum y_i)^2]}}$$

$$R = \frac{67120}{\sqrt{33600 \times 134508}} = 0.998$$

أي أن العلاقة بين X و Y قوية جداً

9 - احسب معامل ارتباط بيرسون بين المتغيرين y, x المبينة في الجدول :

الجدول (6-13)

x	y	x.y	x ²	y ²
1	0	0	1	0
2	4	8	4	16
3	8	24	9	64
4	6	24	16	36
5	2	10	25	4
$\sum = 15$	$\sum = 20$	$\sum = 66$	$\sum = 55$	$\sum = 120$

$$\bar{x}=3$$

$$S_x=1.414$$

$$\bar{y}=3$$

$$S_y=2.828$$

----- الفصل السادس :الارتباط

$$R = \frac{5 \times 66 - 15 \times 20}{\sqrt{5 \times 120 - (20)^2} \sqrt{5 \times 55 - (15)^2}} = 0.30$$

8-احسب معامل ارتباط بيرسون للبيانات في الجدول التالي :

الجدول (6-14)

x	y	x.y	x ²	y ²
5	8	40	25	64
4	6	24	16	36
3	2	6	9	4
2	1	2	4	1
1	3	3	1	9
$\sum x = 15$	$\sum y = 20$	$\sum x^2 = 55$	$\sum y^2 = 114$	$\sum xy = 75$

$$R = \frac{5 \times 75 - 15 \times 20}{\sqrt{(5 \times 55 - 225)(5 \times 114 - 400)}}$$

----- الفصل السادس :الارتباط

$$R = \frac{75}{\sqrt{(50)(170)}}$$

$$R = \frac{75}{10\sqrt{85}} = 0.81$$

9-احسب معامل الارتباط بين المتغيرين (x و y)

الجدول (6-15)

x	y	x.y	x ²	y ²
1	1	1	1	1
1	2	2	1	4
1	3	3	1	9
1	5	5	1	25
2	3	6	4	9
2	1	2	4	1
3	5	15	9	25
3	0	0	9	0
4	6	24	16	36
4	3	12	16	9
5	7	35	25	49
5	4	20	25	16
$\sum x = 15$	$\sum y = 20$	$\sum xy = 125$	$\sum x^2 = 184$	$\sum y^2 = 112$

$$R_{xy} = \frac{12 \times 32 - 32 \times 40}{\sqrt{(12 \times 113 - (32)^2)(12 \times 184 - (40)^2)}}$$

$$R_{xy} = \frac{1500 - 1280}{\sqrt{(1344 - 1024)(2208 - 1600)}}$$

$$R_{xy} = \frac{220}{\sqrt{194560}} = 0.50$$



الفصل السابع

الانحدار

Regression



مقدمة :

يهدف الانحدار إلى الإفادة من الارتباط في التنبؤ، فإذا علمنا معامل ارتباط درجات اختبار الحساب بدرجات اختبار الجبر، وعلمنا درجة أي طالب في اختبار الحساب فإننا نستطيع أن نتنبأ بدرجته في الجبر، وإذا علمنا درجة طالب آخر في اختبار الجبر فإننا نستطيع أن نتنبأ بدرجته في الحساب، ولهذا التنبؤ أهميته النفسية في الإفادة من اختبارات الاستعدادات العقلية المختلفة التي تهدف إلى التنبؤ بمستويات الأفراد في نواحي النشاط الجديدة التي لم يمارسوها من قبل .

مفاهيم إحصائية متعلقة بتحليل الانحدار :

1. **خط الانحدار Regression Line**: عند رسم شكل الانتشار لعلاقة خطية بين متغيرين، يلاحظ أن النقاط منتشرة في اتجاه معين، فإذا حاولنا الوصل بين هذه النقاط، فإننا نحصل في أغلب الحالات على خط منكسر لا تحكمه أي معادلة إحصائية. ولا تتحقق الفائدة العلمية إلا برسم خط مستقيم تحكمه معادلة إحصائية ويمثل النقاط أفضل تمثيل، ويسمى هذا الخط بخط الملاءمة الأفضل، أو خط الانحدار. ويمكن التوصل إلى خط الانحدار الأفضل إما بأسلوب الرسم أو بطريقة المربعات الصغرى.

2. **الخطأ المعياري للتقدير Standard Error of estimation** : يعرف خطأ التقدير إحصائياً بأنه الفرق بين الدرجة الحقيقية على المتغير التابع و الدرجة المتنبأ بها. الانحراف المعياري لأخطاء التقدير هو الخطأ المعياري للتقدير،(عودة، الخليلي، 2000، ص469) ويحسب الخطأ المعياري لتقدير الانتشار للبيانات المعطاة حول خط الانحدار. كلما كبرت قيمة الخطأ المعياري للتقدير، قلت الثقة بالقيم المقدرة و قلت الثقة بمعادلة مستقيم الانحدار .

3. **معامل التحديد Coefficient of determination**: إن معامل الارتباط يعكس مقدار الترابط الخطي بين المتغيرين اللذين يعبر معامل الارتباط عن العلاقة بينهما، حيث يمكن من خلال معامل الارتباط والعلاقة الخطية المفترضة بين المتغيرين اللذين تربطهما علاقة خطية إيجاد مقدار التباين في أحدهما الذي يمكن أن يفسر عن طريق التباين في الثاني، ويتم إيجاد

----- الفصل السابع :الانحدار

عادة عن طريق إيجاد مربع معامل الارتباط والذي يسمى معامل التحديد *Determination* ويساوي R^2 . فإذا كانت قيمة معامل الارتباط بين المتغير X والمتغير Y تساوي 0.50 يمكن القول أن 0.25 من التباين في المتغير X يفسر عن طريق التباين في المتغير Y ، في حين أن 0.75 من التباين في المتغير X يفسر عن طريق التباين لدى الأفراد في المتغيرات الأخرى . يعرف معامل التحديد بأنه: مقدار التغير في المتغير المستقل الذي تفسره معادلة خط الانحدار، (بري وآخرون، 1998، ص454). وهو مقدار التخفيف في تشتت القيم الفعلية الناتج من استخدام خط الانحدار ، ويحسب بتربيع الفرق بين القيم الفعلية ومتوسط القيم .

التباين المفسر والتباين غير المفسر:

هناك ثلاثة أنواع من مجموع المربعات وهي:

- أ- تباين القيم حول متوسط العينة وهو يستخدم في تحديد التباين والانحراف المعياري للعينة .
- ب- تباين القيم حول خط الانحدار أو حول القيم المتنبأ بها ويسمى التباين غير المتنبأ به أو التباين الذي لا نستطيع تفسيره.

ويمكن أن يتضح سبب هذه التسمية إذا رجعنا إلى تفسير معامل الارتباط بين متغيرين، حيث أنه إذا كان معامل الارتباط بين متغيرين يساوي (± 1) أي معامل ارتباط تام ، فإن جميع القيم تقع على خط الانحدار . وهذا يعني أننا نكون قد فسرنا التباين الكلي للمتغير Y بمعلومية المتغير X ، أي أننا نستطيع القول أنه في حالة الارتباط التام يمكننا تفسير التباين الكلي. ولكن لكي يكون هذا الاستنتاج صحيحاً يجب أن نفترض أن قيمة معامل الارتباط هي القيمة الفعلية التي لا ترجع إلى الصدفة. وهذا يعني عدم اختلاف قيمة معامل الارتباط اختلافاً ملحوظاً باختلاف العينات المستمدة من المجتمع الأصلي. أما إذا لم يكن معامل الارتباط تاماً فسوف نجد أن كثير من القيم لا تقع على خط الانحدار، وانحرافات هذه القيم عن خط الانحدار تمثل التباين الذي لا نستطيع تفسيره بمعلومية الارتباط بين المتغيرين . ولذلك استخدمنا عبارة التباين الذي لا نستطيع تفسيره أو التباين غير المتنبأ به.

----- الفصل السابع :الانحدار

ج- تباين القيم المتنبا بها حول متوسط التوزيع، ويسمى التباين المتنبا أو التباين الذي يمكن تفسيره. وكلما زادت قيمة معامل الارتباط زاد مقدار التباين الذي يمكن تفسيره أو التنبؤ به. وعندما يكون مقدار هذا التباين أكبر ما يمكن يكون معامل الارتباط تاماً، وتكون نسبة التباين الذي يمكن تفسيره (100%)

الافتراضات حول تحليل الانحدار :

- 1- التوزيع الطبيعي : تعني أن قيم المتغير التابع والمستقل تنتوزع على نحو طبيعي.
- 2- تجانس التباين: ويعني أن التباين حول خط الانحدار يكون متناسقاً لكل قيم X (المتغير المستقل)، أي أن قيم (X) تتباين بالمقدار نفسه. و افتراض التباين هام من أجل استخدام طريقة المربعات الصغرى ولتحديد معامل الانحدار . و في حال وجود انحراف جوهري في تجانس التباين تحول البيانات إلى نمط آخر .
- 3- استقلال الأخطاء : يجب أن يكون اختيار العينة بطريقة عشوائية، و لا تعتمد قيم أي مفردة من مفردات العينة على قيم أي فرد آخر. و يترتب على هذا أن تكون الأخطاء مستقلة (James & Terry, 2000, p 325; 1998 Levin, R. & Rubin pp.639-640; Zar, 1996, P. 325).

استخدامات تحليل الانحدار :

"يخدم تحليل الانحدار الأغراض الرئيسة :

- 1- الوصف .
 - 2- السيطرة
 - 3- التنبؤ .
- وتتداخل الأغراض المتعددة لتحليل الانحدار في الواقع العملي" (كنجو وآخرون، 2000، ص34)

ويقصد بالوصف وصف طبيعة العلاقة بين المتغير التابع والمتغير (أو المتغيرات) المستقل (المستقلة) مثل وصف طبيعة العلاقة بين الذكاء و بعض المتغيرات البيئية (تعليم الأب

----- الفصل السابع :الانحدار

ومهنته، التحصيل الدراسي، تعليم الأم) للكشف عما إذا كانت علاقة خطية أم منحنية، وعما إذا كانت إيجابية أم سلبية، قوية أم ضعيفة. أما السيطرة فيقصد بها التحكم بقيم المتغير التابع عند تغيير قيم المتغيرات المستقلة، (حبيب، 1996، ص 498). فمثلاً إذا قمنا بدراسة لأثر عدد من المتغيرات على أداء مجموعة من الأطفال في اختبارات الذكاء التي تقوم على حل المشكلات، وجرى لهذا الغرض استخدام تحليل الانحدار المتعدد، وأثبتت الدراسة أن التدريب له أثر في تحسين أدائهم على تلك الاختبارات، هنا يمكننا أن نقوم بتدريب الأطفال على حل المشكلات وتنمية قدرتهم على التعامل مع العلاقات المجردة، وفي هذه الحالة نكون قد قمنا بعملية السيطرة على أحد المتغيرات .

ويستخدم تحليل الانحدار أيضاً في التقدير، أي تقدير القيمة الوسطية للمتغير التابع من أجل قيمة معينة من قيم العينة، (Mendenhall ,1995,p.572). أما التنبؤ فهو الغاية الأكثر أهمية من غايات تحليل الانحدار فيما يختص بالقضايا التربوية والنفسية . والآن يتبادر إلى الذهن سؤال : لماذا نهتم بالتنبؤ بقيم المتغير التابع عندما يكون لدينا قيم حقيقية منه، و على نحو خاص عندما تكون القيم المتنبأ بها على درجة ما من الخطأ، فإذا كنا مهتمين بخصائص عينة محددة فإن تحليل الانحدار الخطي ليس له أهمية تذكر. وتتمثل الفائدة من تحليل الانحدار في وصف العلاقة بين X و Y في حدود كمية . وعلى كل حال، نهتم غالباً بالذهاب إلى ما هو أبعد من البيانات الحقيقية المتوافرة لدينا، ونظرياً يمكننا أن نتنبأ بأي قيمة من (Y) لأجل أي قيمة من (X) ، إلا أن صدق القيمة المتنبأ بها والمحسوبة بفترة الثقة، تقل عندما نتحرف عن المدى الحقيقي لقيم (X) . أي أن النتائج لا يوثق بها إلا إذا كانت ضمن المدى المحسوب، واستخدام قيمة متطرفة لا يعطي نتيجة دقيقة . وهناك إمكانية أخرى لاستخدام تحليل الانحدار في التنبؤ، في حالة عدم وجود أي معلومات على الإطلاق عن المتغير التابع (Y) . كأن نتنبأ بتحصيل الطلاب في الجامعة من خلال معرفة تحصيلهم في الثانوية،(Raymond,1999,P.255) وعملياً تتداخل هذه الأهداف مع بعضها فيحقق تحليل الانحدار الغايات جميعها .

نماذج الانحدار The Regression Models :

ذكرنا سابقاً أن الهدف من قياس معاملات الارتباط هو معرفة درجة العلاقة أو مقدار الترابط بين المتغيرات، أو درجة اقتران متغير بمتغير آخر . وقلنا إن هذا الاقتران ليس معناه أن أحد المتغيرين يسبب المتغير الآخر. إلا أنه إذا وجدت علاقة قوية بين متغيرين فإننا ربما نحتاج إلى تقدير أحد المتغيرين بدلالة المتغير الآخر (أو التنبؤ به) ويسمى المتغير الذي يراد دراسة سلوكه (أو التنبؤ به). ومعرفة مدى تأثيره بالمتغيرات الأخرى، بالمتغير التابع Dependent Variable ويطلق على كل متغير يؤثر في سلوك المتغير التابع المتغير المستقل (المنبأ) Independent Variable.

ويستخدم تحليل الانحدار كأسلوب إحصائي في تقدير العلاقة بين متغيرين أو أكثر على شكل علاقة دالية، يمكن عن طريقها معرفة التغير في أحد المتغيرات على أساس تأثيره بالمتغيرات الأخرى، أو بمعنى آخر يتم عن طريقها التنبؤ بقيمة أحد المتغيرات (المتغير التابع) عن طريق معرفة قيم المتغيرات الأخرى (المتغيرات المستقلة) فمثلاً قد نرغب في تقدير أو التنبؤ بدرجة كفاءة أداء العامل (متغير تابع) لدرجة الرضا عن العمل (متغير مستقل). وبمدى تطبيق المنظمة للجودة الشاملة (متغير مستقل آخر)، أو بمعنى آخر معرفة أكثر المتغيرات المستقلة (الرضا عن العمل، مدى تطبيق الجودة) تأثيراً في المتغير التابع (أداء العامل).

ويمكن القول إن تحليل الانحدار يعدّ من أهم الفروع الإحصائية التي تستخدم على نطاق واسع في جميع مجالات العلم والمعرفة . فهناك كثير من المشكلات الإدارية تتضمن التنبؤات. فالتخطيط للمستقبل يعدّ جزءاً متكاملاً للإدارة، ويحتاج بالضرورة إلى التنبؤ فمثلاً تحديد ميزانية العام القادم يحتاج إلى تنبؤ بمستوى العمليات، ورأس المال المستخدم في الآلات

----- الفصل السابع :الانحدار

والأجهزة، والأفراد اللازمين لتنفيذ العمليات، وعوامل أخرى كثيرة وهناك أمثلة لحالات كثيرة نحتاج فيها إلى التنبؤ في الإدارة والاقتصاد والصناعة والزراعة..... الخ

وتعد مشكلة تحديد النموذج المستخدم في عملية التنبؤ أو التقدير، باستخدام نماذج الانحدار، من أهم المشكلات التي تواجه الباحث، حيث تتداخل العديد من العوامل في تحديد اختيار النموذج الملائم للتحليل، منها نوعية البيانات، وطبيعة المتغيرات المستخدمة في التحليل، ومستويات قياسها، وطبيعة العلاقة بين المتغيرات محل الدراسة، ومن بين نماذج الانحدار نذكر:

1- نماذج الانحدار البسيط **linear Regression**: يستخدم الانحدار الخطي البسيط

في وصف العلاقة الخطية بين متغيرين، أحدهما يؤثر في الآخر، ويسمى بالمتغير المستقل، ويرمز له بالرمز X ، والآخر يتأثر بالأول، ويسمى بالمتغير التابع، ويرمز له بالرمز (Y) ويجب أن يكون المتغير التابع متغيراً متصلاً ومستوى قياسه لا يقل عن المستويين الفئوي أو النسبي، بينما المتغير المستقل قد يكون مستوى قياسه ترتيبياً أو فئوياً أو نسبياً .

2- نماذج الانحدار المتعدد **Multiple Regression** :

تستخدم في حالة وجود متغير تابع، وأكثر من متغير مستقل. (فهيمى،2005: 608-610)

3- نماذج الانحدار متعدد التباين **Multivariate Regression**: تستخدم عندما يكون

هناك متغيرات تابعة عديدة ومتغيرات مستقلة عديدة. (Rencher,2002,p. 322)

-واستنادا إلى شكل العلاقة بين المتغيرات يمكن أن نميز بين نوعين من أنواع نماذج الانحدار:

----- الفصل السابع : الانحدار

1- نماذج الانحدار الخطي Linear Regression

2- نماذج الانحدار المنحني (غير الخطي) Nonlinear Regression

ج- مستوى القياس للمتغيرات: هنا يوجد تقسيم شائع أيضاً :

1- نماذج الانحدار التقليدية : تستخدم في حالة ما إذا كان المتغير التابع متغيراً كمياً (فئوياً أو نسبياً).

2- نماذج الانحدار اللوجستي : تستخدم في حالة ما إذا كان المتغير التابع متغيراً كيفياً (ترتيباً أو اسمياً).

تفسير معامل الانحدار :

يفسر معامل الانحدار في معادلة الانحدار البسيط على أنه الزيادة في Y لكل وحدة زيادة في X ، أما في حالة الانحدار المتعدد فهناك عدد من المتنبئات يسهم كل واحد منها بجزء خاص في هذه الزيادة و قيمة هذه المساهمة تعتمد على قيمة معامل الانحدار، و لذلك يمكن تسميته معامل الانحدار الجزئي . وهو يعطي التغير في Y عندما تتغير X بوحدة واحدة و تبقى آثار المتغيرات الأخرى ثابتة، (عودة والخليلي، 2000، ص 488).

اعتبارات في تطبيق تحليل الانحدار :

1- كثيراً ما يستخدم تحليل الانحدار للقيام بالتنبؤ بالمستقبل، وتأتي مشروعية تحليل الانحدار من أن الشروط الأساسية في المستقبل تماثل الشروط الحالية التي أجري فيها تحليل الانحدار .

2- عند التنبؤ بقيم جديدة للمتغير التابع، غالباً ما نضطر إلى التنبؤ بالمتغير المستقل نفسه. مثل التنبؤ بعدد طلبة السنة الأولى في الجامعة للعام القادم من خلال عدد

----- الفصل السابع :الانحدار

الطلاب الذين سوف يسجلون في الثانوية العامة في العام القادم بناء على عدد الناجحين في الصف الحادي عشر في هذا العام . فهذا التنبؤ مشروط تعتمد صحته صحة عدد المسجلين وعدد الناجحين منهم، ومن السهل إغفال الطبيعة الشرطية في هذا النوع من التنبؤ .

3- وهناك محذور آخر عندما نتنبأ بمستويات المتغير المستقل الواقعة خارج مدى البيانات، ومثل هذا الأمر كثيراً ما يقع . فإذا كانت تلك القيمة ليست بعيدة عن ذلك المدى فقد تتوافر فيه ثقة مقبولة. أما إذا كانت بعيدة كثيراً عن مدى البيانات السابقة، فيجب ممارسة الحذر الشديد لأنه لا يمكن التأكد من دالة الانحدار التي تصلح للبيانات السابقة من أنها ستبقى صالحة فوق مدى أوسع للمتغير المستقل .

4- الاختبار الإحصائي الذي جعلنا نستنتج أن قيمة معامل الانحدار لا تساوي صفراً لا يبين علاقة سبب ونتيجة بين المتغيرين التابع والمستقل، فمثلاً في البيانات غير التجريبية قد يكون المتغيران التابع والمستقل متأثرين معاً بمتغيرات أخرى لم يشملها نموذج الانحدار. فقد يكون هناك علاقة خطية واضحة بين المفردات اللغوية وسرعة الكتابة عند أطفال مدرسة ابتدائية، إلا أن هذا قد يكون نتيجة عمر الطفل، وعدد سنوات الدراسة وعوامل مشابهة تؤثر فيهما، ومن جهة أخرى وجود علاقة انحدار في تجارب تحت السيطرة تعد دليلاً جيداً على وجود علاقة سبب ونتيجة. (كنجو وآخرون، 2000، ص 106، 105)

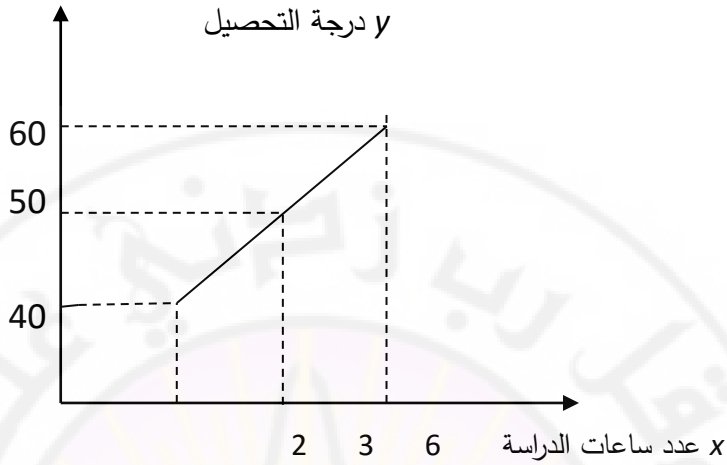
شكل الانتشار : Scatter plot

الأسلوب المفيد لبدء تحليل الانحدار هو تمثيل البيانات بيانياً وهو ما يعرف بشكل الانتشار، نسمي المستقيم الذي يصل بين النقاط باسم الوسط الحسابي، أما النقاط التي تقع على جانبه فإنها تسمى نقاط انتشار، وهي تمثل في الوقت نفسه الانحراف المعياري عن المتوسط الحسابي.

نستفيد من شكل الانتشار فيما يلي:

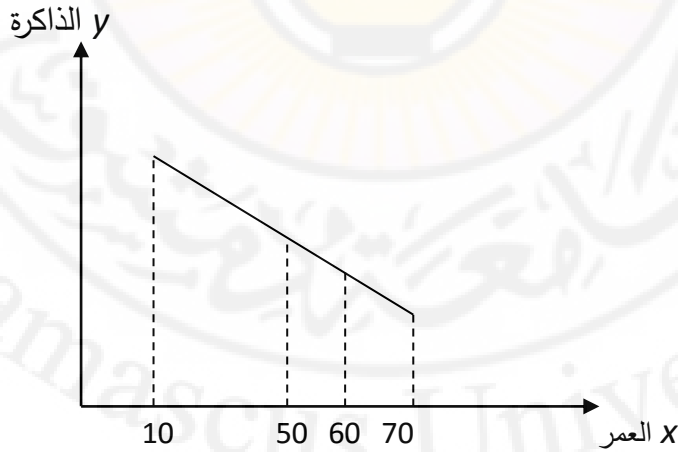
- 1- يوضح فيما إذا كان هناك علاقة ظاهرة بين المتغيرين أم لا .
 - 2- يوضح شكل العلاقة هل هي خطية أم لا.
 - 3- يوضح فيما إذا كانت العلاقة إيجابية أم سلبية. (عبد المنعم، 2005، ص9)
- كيفية التمييز بين المتغير المستقل والمتغير التابع :

- 1- زمنياً : المتغير الأسبق هو المستقل .
 - 2- التأثير : المتغير المؤثر هو المستقل .
 - 3- العلاقة المنطقية . أي يكون من المنطقي أن أحد المتغيرين لا بد أن يكون هو المؤثر وهو حينئذ المتغير المستقل، والمتغير الآخر هو المتغير الذي يقع عليه التأثير وهو المتغير التابع. مثل العلاقة بين جنس الفرد وسرعة تعلم قيادة السيارة، إذ من المنطقي أن الجنس لا بد أن يكون هو المتغير المستقل وسرعة تعلم القيادة المتغير التابع، إذ من غير المنطقي أن سرعة تعلم القيادة تؤثر في الجنس بحيث يتغير تبعاً لسرعة تعلم القيادة. كذلك من الممكن دراسة العلاقة بين شدة الضوضاء أثناء الامتحان ودرجات الطلاب في ذلك الامتحان، فالمتغير المستقل هنا شدة الضوضاء، ودرجات الامتحان هي المتغير التابع.
- وفيما يلي رسم بياني للعلاقة بين عدد ساعات الدراسة (y) وبين التحصيل الدراسي (الدرجات)



شكل رقم (7-1)

يلاحظ من الشكل السابق أن العلاقة هي علاقة خطية إيجابية تامة بمعنى أن كل زيادة في X المتغير المستقل تؤدي إلى زيادة في المتغير التابع (y) أما العلاقة الخطية السببية التامة فيمكن تمثيلها من خلال العلاقة بين العمر والذاكرة على النحو التالي :



شكل رقم (7-2)

----- الفصل السابع : الانحدار

أما العلاقة بين العمر والذاكرة فهي علاقة سلبية (عكسية) إذ أن الذاكرة تتناقص مع تقدم العمر

نموذج الانحدار البسيط :

الفكرة الأساسية في تحليل الانحدار البسيط أن متوسط قيم المتغير التابع Y تتوضع على خط مستقيم عندما يرسم بمقابل المتغير المستقل X والتباين في المتغير Y للقيمة المعطاة لـ X تفترض أن تكون X مستقلة. (Brion & s.Everitt, 2001, p.161) وبصورة أكثر تحديداً معادلة مستقيم الانحدار على النحو التالي

$$\hat{y} = a + bx$$

$\hat{y}$: المتغير التابع

x : المتغير المستقل

$a:b$ ثوابت

أما قيمة a فهي قيمة y عندما x تساوي الصفر

$$\hat{y} = a \quad \longleftarrow \quad x = 0 \text{ أي}$$

b : تدعى معامل الانحدار ويقصد به مقدار التغير في قيمة المتغير التابع y عندما تتغير قيمة المتغير المستقل (x) وحدة واحدة.

معادلة مستقيم الانحدار : تعرف بأنها المعادلة التي تمثل العلاقة الوسطية بين المتغير المستقل x والمتغير التابع y .

ولهذه المعادلة شرطان هما :

1- الشرط الأول : أن تكون العلاقة بين متغيرين فقط، حيث يكون أحد المتغيرين مستقل بينما يكون المتغير الآخر تابع .

----- الفصل السابع :الانحدار

2- الشرط الثاني : يمكن تمثيل العلاقة بخط مستقيم .

لحساب الثوابت a و b نطبق الحل المشترك للمعادلتين الطبيعييتين التاليتين :

$$(1) \Sigma y = n a + b \Sigma x$$

$$(2) \Sigma x y = a \Sigma x + b \Sigma x^2$$

أو بالطريقة المختصرة

$$b = \frac{\Sigma xy - \bar{x} \Sigma y}{\Sigma x^2 - \bar{x} \Sigma x}$$

أو بصيغة أخرى

$$b = \frac{\Sigma xy - \frac{\Sigma x \Sigma y}{n}}{\Sigma x^2 - \frac{(\Sigma x)^2}{n}}$$

أو بصيغة ثالثة

$$b = \frac{\Sigma xy - n \bar{x} \bar{y}}{\Sigma x^2 - n \bar{x}^2}$$

$$a = \bar{y} - b \bar{x}$$

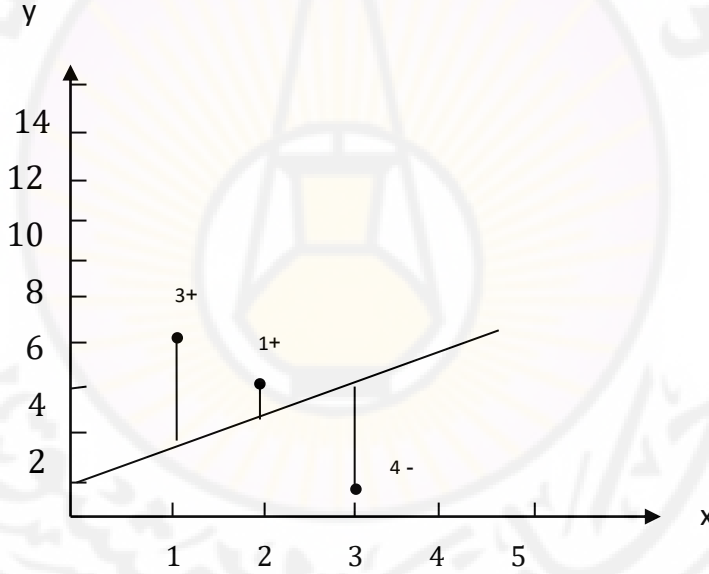
وهناك عدة طرائق للحصول على خد الانحدار أهمها:

طريقة المربعات الصغرى The method of least squares :

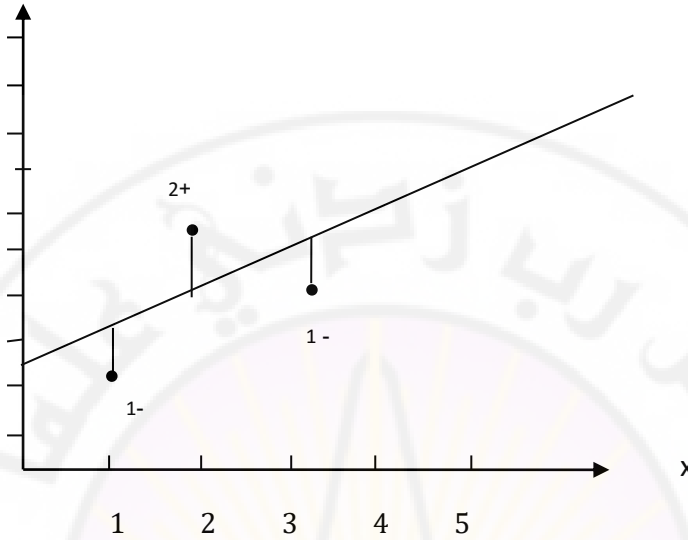
بالرغم من وجود العديد من الطرائق للحصول على تقديرات للمعلمتين a و b , إلا أن أفضل هذه الطرائق هي طريقة المربعات الصغرى. ترجع هذه الطريقة إلى عالم الرياضيات الألماني كارل فريدريكس جاوس Carl Friendrich Gauss . وبما أن الخط المطلوب يكون لأغراض التنبؤ لذلك من المناسب أن يكون الخط من الدقة بحيث تكون أخطاء التقدير صغيرة. والمقصود هنا

----- الفصل السابع :الانحدار

بأخطاء التقدير الفروق بين القيم المشاهدة y والقيم المناظرة $\hat{y}$ (البواقي) على الخط المستقيم. أي أن أخطاء التقدير هي $1, 2, \dots, n$ و $(y - \hat{y})$ أخطاء التقدير موضحة في شكل (7-3) وشكل (7-4) بأجزاء الخطوط الرأسية التي تصل بين النقاط والخط المستقيم . النقطة الواقعة فوق الخط تعطي خطأً (باقي) موجب والنقطة الواقعة تحت الخط تعطي خطأً سالب، واحد من الطرائق لتقليل الأخطاء هو جعل $\sum_{i=1}^n (y - \hat{y})$ أقل ما يمكن، ولكن جعل $\sum_{i=1}^n (y - \hat{y})$ أقل ما يمكن لا يعني الحصول على توفيق جيد . ففي الشكل (7-4) ثلاثة أخطاء واحد موجب والآخرين سالبين حيث $\sum (y - \hat{y}) = 0$ في هذه الحالة بتقليل الخطأ فإننا حصلنا على توفيق يبدو جيداً.

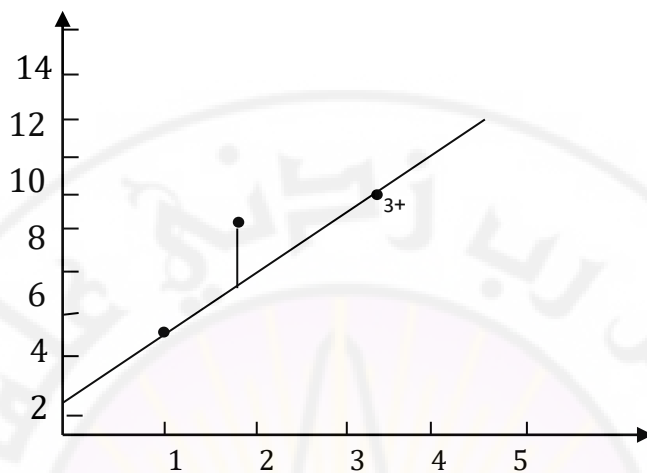


الشكل : (7-3)

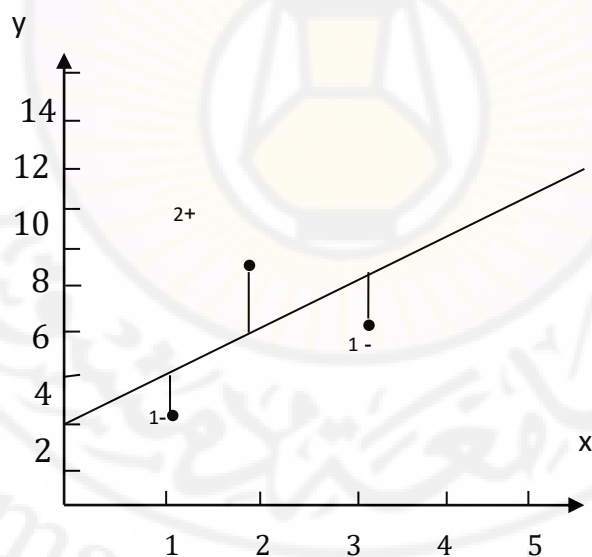


الشكل (7-4):

الآن بالنظر إلى الشكل (7-4) فإن خط الانحدار أدى إلى جعل $\Sigma = (y - \hat{y}) = 0$ وبالرغم من ذلك يتضح أن التوفيق رديء . الآن ماذا يحدث عند إهمال الإشارة وإيجاد الخط المقدر الذي يجعل $\Sigma = (y - \hat{y}) = 0$ أقل ما يمكن ؟ مرة أخرى لم نضمن أن الخط يمثل أفضل توفيق. في الشكل أن الخط في (7-5) أفضل من الخط في (7-6) بالرغم من أن الخط في الشكل (7-6) جعل $\Sigma = (y - \hat{y}) = 0$ أقل من الشكل (7-5)



الشكل : (7-5)



الشكل : (7-6)

----- الفصل السابع :الانحدار

وعلى ذلك نجد أن استخدام القيم المطلقة ليس مناسباً في المعالجة الرياضية ولذلك فإن هذه الصعوبة يمكن تلافيها بأن نطلب أن يكون مجموع مربعات الأخطاء صغيراً بقدر الإمكان.

مما تجدر الإشارة إليه أن طريقة المربعات الصغرى تقلل قيم المعالم هذه إلى أقصى حد، وبالتالي تحدد ما يعرف بأفضل خط مستقيم (أو خط الملائمة الأفضل) الذي يوفق النقاط المشاهدة بطريقة المربعات الصغرى. ومما يجدر الإشارة إليه أيضاً أن طريقة المربعات الصغرى لتوفيق خط مستقيم لمجموعة من النقاط يمكن تطبيقها سواء كانت قيم X حددت مسبقاً أو تمثل قيم لمتغير عشوائي، أي إذا كان المتغير المستقل والمتغير التابع يمثلان متغيرات عشوائية. وفي هذه الحالة تطبق طريقة المربعات الصغرى إذا تحقق الشرطان التاليان :

1- التوزيعات الشرطية للمتغيرات التابعة $\hat{y}$ تمثل توزيعات طبيعية مستقلة .

2- المتغيرات X هي متغيرات عشوائية مستقلة .(عبد المنعم، 2005، ص13-15)

الأخطاء المعيارية للتقدير:

على الرغم من أن طريقة المربعات الصغرى ينتج عنها الخط الملائم للبيانات مع الحد الأدنى من التباين، فإن معادلة الانحدار ليست متنبأ تام، إلا إذا كانت جميع نقاط البيانات واقعة على خط الانحدار. بما أننا لا نستطيع توقع أن تحدد جميع قيم البيانات بدقة من متوسطها الحسابي، وبالطريقة نفسها لا نستطيع أن نتوقع أن جميع النقاط واقعة على خط الانحدار. فإن خط الانحدار يخدم بوصفه متنبأ تقريباً لقيم Y لكل قيمة من X لذلك تحتاج لمقياس إحصائي يقيس التباين بين القيم الحقيقية لـ $\hat{y}$ والقيم المتنبأة لـ $\hat{y}$. والمقياس الذي يقيس هذا التباين حول خط الانحدار يسمى الخطأ المعياري للقياس.

هذا ويحسب الخطأ المعياري وفق المعادلة التالية:

$$S_e = \sqrt{\frac{\sum (y - \hat{y})^2}{n-2}} \quad (\text{Bearson, \&Mlevine,2002,p.591})$$

----- الفصل السابع : الانحدار

مثال :

نفرض أنه تم دراسة العلاقة بين مصاريف الإعلان لسلعة ما x والمبيعات للسلعة Y والبيانات موضحة في الجدول

الجدول (3-7)

x	y	x^2	xy
100	9	10000	900
105	8	11025	840
90	5	8100	450
80	2	6400	160
80	4	6400	320
85	6	7225	510
87	4	7569	348
92	7	8464	644
90	6	8100	540
95	7	9025	665
93	5	8649	465
85	5	7225	425
85	4	7225	340
70	3	4900	210
85	3	7225	255
1322	78	117532	7072

ومن الجدول (73) يمكن حساب قيمة α, b كالتالي :

$$(1) 78 = 15a + 1322b$$

----- الفصل السابع : الانحدار

$$(2) 7072 = 1322a + 117532b$$

بضرب المعادلة الأولى في 1322 (معامل a في المعادلة الثانية) نحصل على

$$(3) 103116 = 19830a + 1747684b$$

الآن بضرب المعادلة الثانية في 15 (معامل a في المعادلة الأولى) نحصل على :

$$(4) 106080 = 19830a + 1762980b$$

الآن سوف يكون لدينا زوج من المعادلات الآتية حيث معامل a واحد في الاثنين. بطرح المعادلتين نحصل على معادلة لا تحتوي على الحد a :

$$(4) 106080 = 19830a + 1762980b$$

$$(3) 103116 = 19830a + 1747684b$$

$$2964 = 15296b$$

$$b = 0.193776$$

الآن بعد الحصول على قيمة b نعوض عنها في المعادلة الأولى وذلك لإيجاد قيمة a :

$$78 = 15a + 0.193776 \times 1322$$

$$78 = 15a + 256.1721$$

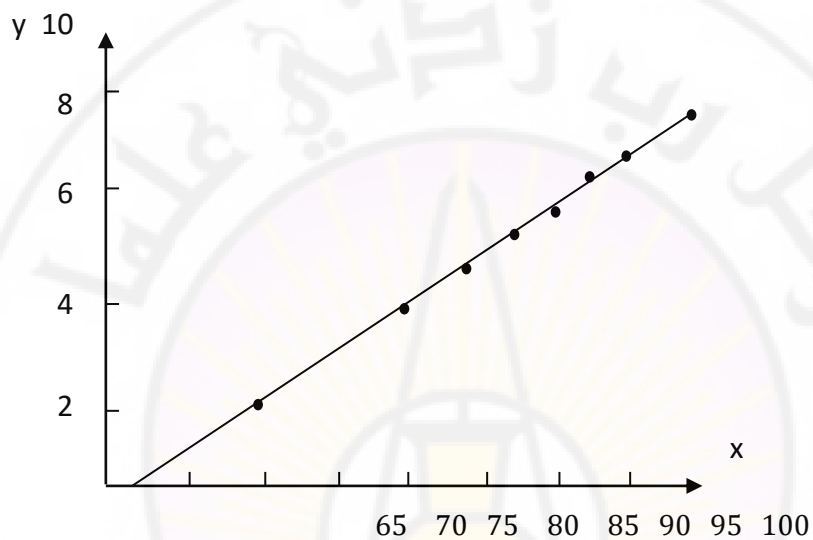
$$a = -11.8781$$

وعلى ذلك معادلة الانحدار المقدرة سوف تكون :

$$\hat{y} = -11.8781 + 0.19378x$$

أو بصورة بسيطة :

$$\hat{y} = -11.9 + 0.19x$$



الشكل (7-7)

مستقيم الانحدار

مثال :

تعد كمية الرطوبة في منتج ما لها تأثير على كثافة المنتج النهائي، تم مراقبة المنتج وقياس كثافته والبيانات المسجلة معطاة في الجدول (74) في شكل شفرة

الجدول (4-7)

x	y	x ²	xy
4.7	3	14.1	22.09
5	3	15	25
5.2	4	20.8	27.04
5.2	10	52	27.04
5.9	2	11.8	34.81
4.7	9	42.3	22.09
5.9	3	17.7	34.81
5.2	7	36.4	27.04
5.9	6	35.4	34.81
5.6	6	33.6	31.36
5	4	20	25
58.3	57	299.1	311.09

قدر معال نموذج الانحدار الخطي البسيط 299.1

الحل :

$$b = \frac{\sum xy - \frac{\sum x \sum y}{n}}{\sum x^2 - \frac{(\sum x)^2}{n}}$$

$$= \frac{311.09 - \frac{(58.3)(57)}{11}}{299.1 - \frac{(58.3)^2}{11}}$$

----- الفصل السابع : الانحدار

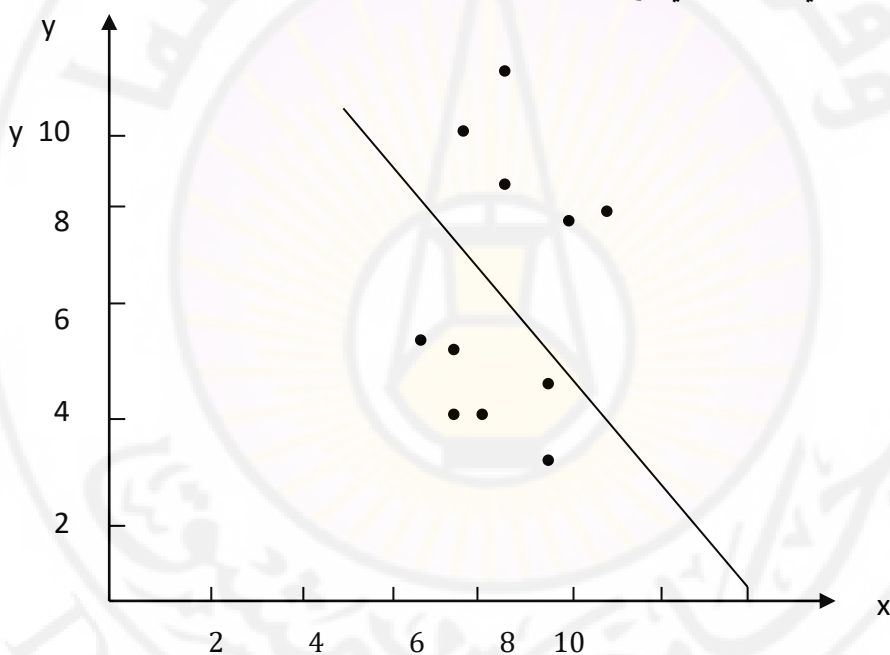
$$= \frac{-3}{2.1} = -1.42857$$

$$a = y - b_1x = 5.18182 - 1.42857(5.3) \\ = 12.7532$$

معادلة الانحدار المقدرة سوف تكون على الشكل

$$\hat{y} = 12.7532 - 1.42857x$$

والممثلة بيانياً في الشكل التالي مع شكل الانتشار



الشكل (7-8)

----- الفصل السابع : الانحدار

مثال :

لتكن لدينا البيانات التالية التي تمثل العلاقة بين عدد ساعات الدراسة والتحصيل بالعشرات:

10	7	5	4	2	ساعات الدراسة
8	5	6	5	3	التحصيل بالعشرات

والمطلوب : حساب معادلة مستقيم الانحدار التي تمثل العلاقة بين ساعات الدراسة والتحصيل وتفسير ثوابتها .

الجدول (5-7)

ساعات الدراسة x	التحصيل y	xy	x^2
2	3	6	4
4	5	20	16
5	6	30	25
7	5	35	49
10	8	80	100
$\Sigma=28$	27	171	194

نعوض بالمعادلتين الطبيعيين :

$$\Sigma y = n a + b \Sigma x$$

$$\Sigma x y = a \Sigma x + b \Sigma x^2$$

نقسم أمثال a في (2) على أمثالها في (1) فنحصل على $5.6 = 5 a + 28 b$ (1)

----- الفصل السابع : الانحدار

نضرب 5.6 بـ المعادلة (1) فنحصل على المعادلة (3) $171 = 28a + 194b$ (2)

نضرب المعادلة (2) بـ (1) فنحصل على المعادلة (4) $151.2 = 28a + 156.8b$ (3)

$$(4) 28a + 194b$$

نطرح (4) - (3)

$$19.8 = 0 + 37.2b$$

$$b = \frac{19.8}{37.2} = 0.53$$

نعوض في (1) :

$$27 = 5a + 28(0.53)$$

$$27 = 5a + 14.84$$

$$a = \frac{27 - 14.84}{5}$$

$$a = 2.43$$

وتكون المعادلة : $\hat{y} = 2.43 + 0.53x$

التفسير :

a : عندما تكون ساعات الدراسة معدومة فإن التحصيل 23.4 درجة .

b : كلما ازدادت ساعات الدراسة ساعة واحدة ينخفض التحصيل 0.53 درجة .

نستخدم معادلة مستقيم الانحدار للتقدير أو التنبؤ :

مثلاً : قدر متوسط تحصيل طالب يدرس 8 ساعات وآخر يدرس ساعة واحدة

$$x=8 \rightarrow \hat{y} = 2.43 + 0.53(8) = 6.67$$

----- الفصل السابع :الانحدار

إن الطالب الذي يدرس 8 ساعات يقدر أن يكون تحصيله 66.7 أي 67 درجة .

$$X=1$$

لا نستطيع التقدير لطالب يدرس ساعة واحدة لوقوع هذه القيمة خارج مجال قيم x

والسبب أننا لسنا متأكدين من استمرار العلاقة خطية خارج مجال هذه القيم .

مثال : لدينا البيانات التالية التي تمثل درجات الطلاب في مادة الحساب (X) والجبر (y):

الجدول (6-7)

x	y	xy	x^2	y^2
0	2	0	0	4
1	1	1	1	1
2	3	6	4	9
3	2	6	9	4
4	4	16	16	16
5	3	15	25	9
6	5	30	36	25
21	20	74	91	68
Σx	Σy	Σxy	Σx^2	Σy^2

$$\bar{y} = 2.86 \quad \bar{x} = 3$$

وبتعويض القيم التي حصلنا عليها في المعادلتين الطبيعيتين، فإننا نجد :

$$(1) 20 = 7a + 21b$$

$$(2) 74 = 21a + 91b$$

فإذا ضربنا المعادلة (1) بـ (3) والمعادلة (2) بـ (1) فإننا نحصل على ما يلي :

----- الفصل السابع : الانحدار

$$(3) 60 = 21 a + 63b$$

$$(4) 74 = 21a + 91 b$$

وبطرح المعادلة (3) من المعادلة (4) فإن الناتج هو :

$$14 = 0 + 28b$$

$$b = \frac{14}{28} = \frac{1}{2} = 0.5$$

وللحصول على قيمة a ، فإننا نعوض (b) بقيمتها في المعادلة (1) فنحصل على ما يلي :

$$20 = 7 a + 21(0.5)$$

$$20 - 10.5 = 7a$$

ومنه :

$$a = \frac{9.5}{7} = 1.36$$

وهكذا فإن معادلة مستقيم الانحدار لبيانات الجدول تصبح على النحو التالي :

$$\hat{y} = 1.36 + 0.5 x$$

تفسير الثوابت :

$$a \text{ عندما } x = 0 \quad 1.36 = y$$

a : عندما تكون درجة الحساب = 0 تكون درجة الجبر 1.36 درجة

b : عندما تزداد درجة الحساب درجة واحدة فإن درجة الجبر تزداد بمقدار 0.5 درجة .

$$\text{عندما } x = 2$$

$$\hat{y} = 1.36 + 0.5 (2)$$

$$= 1.36 + 1 = 2.36$$

$$\hat{y} = 1.36 + 0.5 (3)$$

----- الفصل السابع : الانحدار

$$= 1.36 + 1.5 = 2.86$$

مثال :

تبين القيم التالية عدد السجائر اليومية لعدد من المدخنين وعدد نوبات السعال كما يلي :

الجدول (7-7)

14	12	14	10	6	8	4	عدد السجائر
11	9	11	7	3	5	1	عدد النوبات

والمطلوب حساب معادلة مستقيم الانحدار وتفسير ثوابتها

الجدول (7-8)

x	y	xy	x ²	y ²
4	1	4	16	1
8	5	40	64	25
6	3	18	36	9
10	7	70	100	49
14	11	154	196	121
12	9	108	144	81
14	11	154	196	121
68	47	548	752	407
Σx	Σy	Σxy	Σx^2	Σy^2

----- الفصل السابع :الانحدار

$$\bar{x} = \frac{68}{7} = 9.71$$

$$\bar{y} = \frac{47}{7} = 6.71$$

$$b = \frac{\sum xy - \bar{x} \sum y}{\sum x^2 - \bar{x} \sum x}$$

$$b = \frac{548 - 9.71 \times 47}{752 - 9.71 \times 68}$$

$$b = \frac{91.63}{91.72} = 0.999 = 1$$

$$a = y - b x$$

$$a = 6.71 - 1 \times 9.71 = -3$$

$$\hat{y} = -3 + 1 x$$

التفسير : a : عندما يكون عدد السجائر معدوماً فإن عدد نوبات السعال -3 مقبول إحصائياً مرفوض واقعياً .

b : كلما ازداد عدد السجائر المدخنة وحدة واحدة تزداد عدد النوبات نوبة واحدة

مثال: لدينا البيانات التالية:

$$\sum x = 25 \quad \sum x^2 = 151 \quad \sum y = 40 \quad \sum y^2 = 354 \quad \sum xy = 227$$

$$\bar{x} = 5 \quad \bar{y} = 8$$

$$(1) \sum y = na + b \sum x$$

$$(2) \sum xy = a \sum x + b \sum x^2$$

نضرب (1) بـ 5 و 2 بـ 1

$$(1) 40 = 5a + 25b$$

$$(2) 227 = 25a + 151b$$

نطرح :

$$200 = 25a + 125b$$

----- الفصل السابع : الانحدار

$$227 = 25a + 151b$$

$$27 = 0 + 26b$$

ومنه :

$$b = \frac{27}{26} = 1.038$$

وللحصول على قيمة a فإننا نعوض b بقيمتها في المعادلة (1) فنحصل على ما يلي :

$$40 = 5a + 25 \times (1.038)$$

$$40 = 5a + 25.961$$

$$5a = 25.961 - 40$$

$$a = \frac{14.038}{5} = 2.807 \approx 2.81$$

$$y = 2.81 + 1.038x$$

أو $b =$

$$b = \frac{\sum xy - \bar{x} \sum y}{\sum x^2 - \bar{x} \sum x}$$

$$b = \frac{227 - 5 \times 40}{151 - 5 \times 25}$$

$$b = \frac{27}{26} = 1.038$$

$$a = \bar{y} - b \bar{x}$$

$$a = 8 - 1.038 \times 5 = 2.81$$

$$\hat{y} = 2.81 + 1.038x$$

أسئلة وتمارين:

1- درست عينة من 6 أفراد لمعرفة العلاقة بين العمر العقلي والعمر الزمني فأعطت النتائج التالية :

$$\Sigma y = 27. \Sigma xy = 169. \Sigma x = 36. \Sigma x^2 = 230$$

$$\Sigma (y - \bar{y})^2 = 3.5 \Sigma D^2 = 27.5$$

$$\Sigma (x - \bar{x})^2 = 4.5$$

$$\Sigma (x - \bar{x}) (y - \bar{y}) = 3.2$$

فإذا علمت أن العمر الزمني يتراوح بين 4-8 سنوات
والمطلوب :

- 1- حساب معادلة مستقيم الانحدار وتفسير ثوابتها.
- 2- حساب معامل الارتباط وتفسيره وتبيان أهميته الإحصائية .
- 3- حساب معامل ارتباط الرتب وتفسيره .
- 4- تقدير العمر العقلي لطفل عمره الزمني 7 سنوات وآخر عمره 3 سنوات .
- 5- رسم مستقيم الانحدار بيانياً .

الحل :

1- حساب ثوابت المعادلة :

$$\bar{x} = \frac{\Sigma x}{n} = \frac{36}{6} = 6$$

$$\bar{y} = \frac{\Sigma y}{n} = \frac{27}{6} = 4.5$$

----- الفصل السابع : الانحدار

$$b = \frac{\sum xy - \bar{x} \sum y}{\sum x^2 - \bar{x} \sum x}$$

$$b = \frac{169 - 6(27)}{230 - 6(36)} = \frac{7}{14} = 0.5$$

$$a = \bar{y} - b \bar{x} = 4.5 - 0.5 (6) = 1.5$$

معادلة مستقيم الانحدار : $\hat{y} = 1.5 + 0.5 x$

التفسير :

a : عندما يكون العمر الزمني معدوماً فإن العمر العقلي 1.5 سنة (مقبول إحصائياً مرفوض واقعياً . لعدم إمكانية وجود عمر عقلي دون عمر زمني) .

b : كلما ازداد العمر الزمني سنة واحدة يزداد العمر العقلي وسطياً 0.5 سنة .

2- حساب معامل الارتباط وتفسيره وتبيان أهميته الإحصائية:

$$R = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sqrt{\sum (x - \bar{x})^2 \sum (y - \bar{y})^2}} = \frac{3.2}{\sqrt{4.5(3.5)}} = 0.8$$

التفسير : الارتباط بين العمر العقلي والعمر الزمني شديد وطردى

لمعرفة الأهمية الإحصائية لمعامل الارتباط نحسب الخطأ المعياري لعامل الارتباط:

$$S_r = \sqrt{\frac{1 - r^2}{n - 2}} =$$

$$S_r = \sqrt{\frac{1 - (0.8)^2}{6 - 2}} = 0.3$$

ويجب تحقق العلاقة $r > 3 S_r$ لكي يكون له أهمية إحصائية

$$R = 0.8 \not> 3(0.3)$$

$$0.8 \not> 0.9$$

----- الفصل السابع : الانحدار

لم تحقق العلاقة ليس لمعامل الارتباط أهمية إحصائية ومعامل ارتباط المجتمع يمكن أن يساوي الصفر . والعينة لا تمثل مجتمعها الإحصائي .

3- حساب معامل ارتباط الرتب وتفسيره:

$$R_s = 1 - \frac{6 \sum D^2}{n(n^2 - 1)}$$

$$= 1 - \frac{27.5}{6(36 - 1)} = 1 - \frac{27.5}{210}$$

$$= 0.87$$

التفسير : الارتباط بين رتب العمر الزمني والعمر العقلي شديد وطردي

4- تقدير العمر العقلي لطفل عمره الزمني 7 سنوات وآخر عمره 3 سنوات:

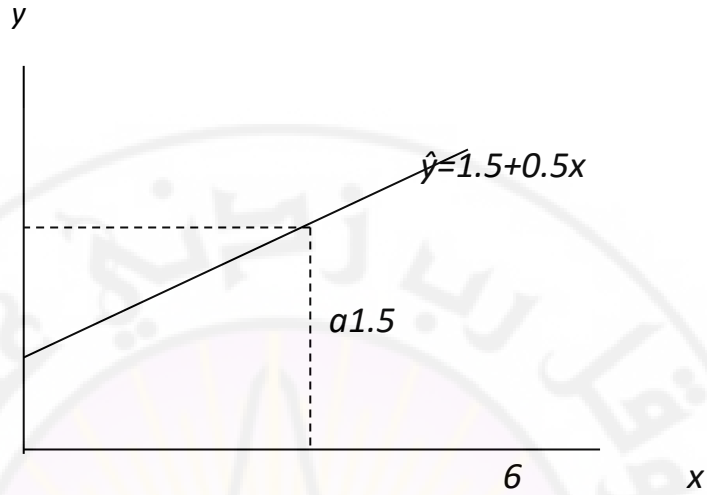
$$x=7 \quad \hat{y} = 1.5 + 0.5(7) = 5 \quad \longrightarrow$$

أي عمره العقلي 5 سنوات

لا نستطيع التقدير لأن هذه القيمة $x=3$ خارج مجال قيم x وبالتالي لسنا متأكدين من استمرار العلاقة خطية خارج مجال هذه القيم .

6- رسم مستقيم الانحدار : مستقيم الانحدار يمر من نقطة تقاطع الوسطين الحسابيين

ويمر من النقطة $(0, a)$



الشكل (7-9)

----- الفصل السابع :الانحدار

مثال :

يمثل الجدول التالي العلامات النهائية لثمانية طلاب في مقرري الإحصاء X، والرياضيات Y.

الجدول (7-9)

95	85	65	80	45	60	65	85	X
87	82	57	72	52	62	67	77	Y

احسب معامل الارتباط بيرسون الخطي .

الحل : نرتب خطوات الحل كما في الجدول التالي :

الجدول (7-10)

x	y	x^2	$x.y$	y^2
85	77	7225	6545	5929
65	67	4225	4355	4489
60	62	3600	3720	3844
45	52	2025	2340	2704
80	72	6400	5760	5184
65	57	4225	3705	3249
85	82	7225	6970	6724
95	87	9025	8265	7569
580	556	43950	41660	39692

لاحظ أنك تحتاج في حساب R إلى المقادير التالية :

----- الفصل السابع :الانحدار

وقد حسبت في الجدول السابق حيث :

$$\bar{y} = \frac{580}{8} = 72.5, \bar{x} = \frac{556}{8} = 69.5$$

وبقية المقادير محسوبة في معادلة R نجد :

$$R = \frac{39692 - 8 \times 72.5 \times 69.5}{\sqrt{41660 - 8 \times (72.5)^2} \sqrt{43950 - 8 \times (69.5)^2}}$$

$$R = \frac{1350}{\sqrt{1900} \sqrt{1014}} = 0.97$$

مثال :

يمثل الجدول التالي الطول X والوزن Y لخمسة طلاب في إحدى الكليات :

الجدول (7-11)

X	172	170	165	169	164
Y	70	71	74	67	68

احسب معامل ارتباط بيرسون الخطي

الحل : نرتب خطوات الحل حيث نحتاج إلى

$$\sum xy, \sum y^2, \sum x^2, y, x, n$$

----- الفصل السابع :الانحدار

الجدول (7-12)

x	y	x.y	x ²	y ²
172	70	12040	29584	4900
170	71	12070	28900	5041
165	74	12210	27225	5476
169	67	11323	28561	4489
164	68	11152	26896	4624
840	350	58795	141166	24530

$$\bar{y} = \frac{340}{5} = 68, \bar{x} = \frac{840}{5} = 168$$

وبتعويض المقادير المحسوبة في المعادلة R نجد:

$$R = \frac{58795 - 5 \times 840 \times 350}{\sqrt{141166 - 5 \times (168)^2} \sqrt{23270 - 5 \times (68)^2}}$$

$$R = \frac{35}{\sqrt{46}\sqrt{150}} = 0.42$$

مثال : أوجد معادلة انحدار علامة الرياضيات (Y) على علامة الإحصاء (X) ثم أوجد تقديراً لعلامة الرياضيات لطالب حصل في الإحصاء على 80 ثم أوجد الخطأ في التقدير .

الحل : افرض معادلة انحدار X على Y بطريقة المربعات الصغرى هي :

$$\hat{y} = \alpha + bx$$

----- الفصل السابع : الانحدار

احسب قيم α ، b من المعادلتين، وذلك بتعويض المقادير المحسوبة في الجدول

$$b = \frac{41660 - 8 \times 72.5 \times 69.5}{43950 - 8(72.5)^2}$$

$$b = \frac{1350}{1900} = 0.71$$

$$\alpha = \bar{y} - b\bar{x} = 69.5 - 0.71 \times 72.5 = 18.03$$

إذن تكون معادلة الانحدار Y على X هي :

$$y = 18.03 + 0.71x$$

ويكون تقدير علامة الرياضيات إذا كانت علامة الإحصاء 80 بتعويض $x=80$ في المعادلة وإيجاد قيمة y وهي :

$$y = 18.03 + 0.71 \times 80$$

$$= 74.8 \approx 75$$

وهو تتبؤ علامة الرياضيات .

أما الخطأ في التقدير فهو :

$$e = y - \hat{y} = 72 - 75 = -3$$

حيث 72 هي القيمة المشاهدة المقابلة للعلامة 80، والعلامة 75 هي القيمة المتنبأ.

----- الفصل السابع : الانحدار

مثال :

أوجد معادلة انحدار الوزن Y على الطول X

حسب المعطيات التالية

$$\sum y = 68 \quad \bar{x} = 168 \quad n = 5 \quad \sum xy = 57155 \quad \sum x^2 = 141166$$

الحل : نعوض بالمعادلة

$$b = \frac{\sum xy - n \bar{x} \bar{y}}{\sum x^2 - n \bar{x}^2}$$

$$b = \frac{57155 - 5 \times 168 \times 68}{141166 - 5 \times (168)^2}$$

$$b = \frac{35}{46} = 0.76$$

$$a = \bar{y} - b\bar{x} = 68 - 0.76 \times 168 = -59.8$$

إذا معادلة انحدار الوزن Y على الطول X هي :

$$\hat{y} = -59.8 + 0.76x$$

إذا كان طول طالب 180 سم فإننا نقدر وزنه :

$$\hat{y} = -59.8 + 0.76 \times 180$$

$$= 77$$

----- الفصل السابع : الانحدار

مثال : أوجد معادلة انحدار Y على X للبيانات في الجدول التالي:

الجدول (7-13)

X	4	10	9	12	8	5
y	2	6	8	11	5	4

2- بماذا نقدر قيمة y المقابلة للقيمة $x=9$

3- ما الخطأ في تقدير y عندما تكون $x=9$

الحل :

الجدول (7-14)

X	Y	XY	X ²
4	2	8	16
10	6	60	100
9	8	72	81
12	11	132	144
8	5	40	64
4	4	20	25
المجموع = 48	36	332	430

معادلة انحدار Y على X هي :

$$y = \alpha + bx$$

----- الفصل السابع : الانحدار

حيث :

$$\bar{x} = \frac{48}{6} = 8, \bar{y} = \frac{36}{6} = 6$$

$$b = \frac{332 - 6 \times 8 \times 6}{430 - 6(8)^2}$$

$$b = \frac{332 - 8 \times 36}{430 - 6 \times 64} = 0.96$$

$$\alpha = 6 - 0.96 \times 8 = -1.68$$

إذا معادلة انحدار Y على X هي :

$$\hat{y} = 1.68 + 0.96x$$

ب- القيمة التقديرية (المتنبأة) للمتغير Y عندما $x=9$ هي :

$$\hat{y} = 1.68 + 0.96 \times 9 = 6.96$$

وذلك بتعويض $x=9$ في معادلة الانحدار .

ج- من الجدول المعطى في المثال والذي يبين المشاهدات X والمشاهدات المقابلة لها Y نجد

أن المشاهدة التي تقابل $x=9$ هي $y=8$ ، وكما وجدنا في (ب) نقدر قيمة Y بالمقدار

$$\hat{y} = 6.96$$

إذا الخطأ في التقدير هو

$$e = y - \hat{y} = 8 - 6.96 = 1.04$$

----- الفصل السابع : الانحدار

مثال : يعطي الجدول التالي علامات 12 طالباً في الامتحان الأول X والامتحان الثاني Y.

الجدول (7-15)

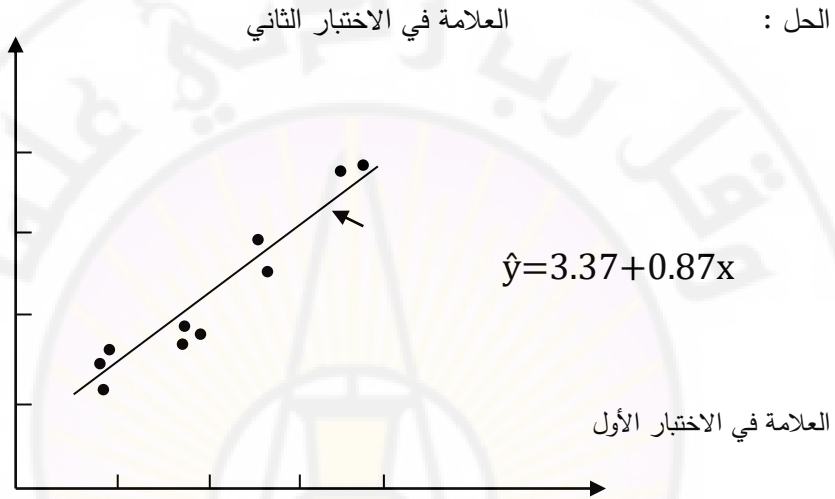
Y	X
20	18
11	14
14	10
16	15
10	7
10	12
17	13
11	8
12	9
20	17
18	15
12	12

- 1- ارسم لوحة الانتشار .
- 2- أوجد معادلة خط الانحدار Y على X
- 3- أخذ طالب في الاختبار الأول العلامة 16 وغاب عن الاختبار الثاني، ما هي العلامة التقديرية التي يحصل عليها الطالب في الامتحان الثاني؟

----- الفصل السابع : الانحدار

4- ما هو الخطأ في تقديرك للعلامة في الاختبار الثاني إذا حصل الطالب على 10 في الاختبار الأول موضحاً ذلك بالرسم.

الحل :



الشكل (10-7)

نعوض بالمعادلة

$$b = \frac{\sum xy - n \bar{x} \bar{y}}{\sum x^2 - n \bar{x}^2}$$

$$b = \frac{2255 - 12 \times 12.5 \times 14.25}{2010 - 12(12.5)^2}$$

$$b = \frac{117.5}{135} = 0.87$$

وبالتعويض في المعادلة :

----- الفصل السابع : الانحدار

$$\alpha = \bar{y} - b\bar{x} = \frac{171}{12} - 0.87 \times 12.5 = 14.25 - 10.875 = 3.37$$

إذا معادلة انحدار الوزن Y على الطول X هي :

$$\hat{y} = 3.37 + 0.87x$$

3- نعوض قيمة $x=16$ في معادلة خط الانحدار فنجد العلامة التقديرية في الامتحان الثاني :

$$\hat{y} = 3.37 + 0.87 \times 16 = 17.29 = 17$$

4- نعوض $x=10$ في معادلة الانحدار فنجد :

$$\hat{y} = 3.37 + 0.87 \times 10 = 12.07 = 12$$

ويكون الخطأ في التقدير

$$e = y - \hat{y} = 14 - 12 = 2$$

----- الفصل السابع : الانحدار

مثال :

أوجد معادلة خط الانحدار Y على X ثم أوجد الخطأ المعياري للتقدير للبيانات :

الجدول (7-16)

X	1	4	3	7	5
y	3	5	4	10	8

الحل :

معادلة انحدار Y على X هي :

$$\hat{y} = a + bx$$

حيث :

$$b = \frac{\sum xy - n \bar{x} \bar{y}}{\sum x^2 - n \bar{x}^2}$$

$$a = \bar{y} - b\bar{x}$$

ولحساب هذه المقادير نرتب خطوات الحل كما في الجدول :

----- الفصل السابع: الانحدار

الجدول (18-7)

x	y	xy	x ²	Ŷ	(y- ŷ)	(y- ŷ) ²	y ²
1	3	3	1	2.26	0.751	0.56	9
4	5	20	16	6.00	1.0016-	1.00	25
3	4	12	9	4.75	0.759-	0.58	16
7	10	70	49	9.75	0.2549	0.06	100
5	8	40	25	7.25	0.7525	0.57	64
20	30	145	100			2.77	214

$$\bar{y} = \frac{30}{5} = 6, \bar{x} = \frac{20}{5} = 4 \text{ إذن}$$

$$b = \frac{145 - 5 \times 4 \times 6}{100 - 5(4)^2} = \frac{145 - 120}{100 - 80} = 1.25$$

إن معادلة الانحدار هي:

$$\hat{y} = 1 + 1.25x$$

ولحساب الخطأ المعياري للتقدير تحتاج إلى حساب y المقابلة لكل قيمة x في معادلة الانحدار

ف نجد y المقابلة لها، كما يلي :

ŷ المقابلة للقيمة x=1 هي :

$$\hat{y} = 1 + 1.25 \times 1 = 2.25$$

----- الفصل السابع : الانحدار

$\hat{y}$ المقابلة للقيمة $x=4$ هي :

$$\hat{y} = 1 + 1.25 \times 4 = 6$$

وهكذا نجد القيم $\hat{y}$ المقابلة للقيم x في الجدول، ونضع هذه القيم في العمود (5) في الجدول.
ثم نجد الانحرافات $(y - \hat{y})$ ونضعها في العمود (6) ونربع الانحرافات ونضعها في العمود (7)،
نجمع مفردات العمود (7) ونقسم على $n-2$ أي $5-2=3$ ونجد الجذر التربيعي للنتائج فنحصل
على الخطأ المعياري للتقدير :

$$S_e = \sqrt{\frac{\sum (y - \hat{y})^2}{n - 2}} = \sqrt{\frac{2.75}{5 - 2}} = 0.96$$

لاحظنا أننا حسبنا S_e من التعريف ولذلك كان العمل مضمناً حيث اضطررنا لحساب $\hat{y}$ لجميع $I = 1, \dots, n$ ثم وجدنا $\sum (y - \hat{y})^2$ ، ولكن إذا استعملنا القاعدة السابقة فإننا نحتاج إلى حساب y_i^2 ونضعها في العمود (8) في الجدول ونحسب:

$$S_{yy} = 214 - 5(6)^2 = 34$$

$$S_{xx} = 100 - 5(4)^2 = 20$$

$$S_{xy} = 145 - 5(4)(6) = 25$$

$$SSE = 34 - \frac{(25)^2}{20} = 2.75$$

$$S_e = \sqrt{\frac{SSE}{n - 2}}$$

ومنه :

$$S_e = \sqrt{\frac{2.75}{3}} = 0.96$$

ويمكن حساب S_e من العلاقة التالية :

$$S_e = \sqrt{\frac{\sum y^2 - a \sum y - b \sum xy}{n - 2}}$$

لاحظ في تطبيق هذه النظرية أنك تستعمل المقادير المحسوبة أصلاً لحساب معادلة الانحدار، وهذه المقادير هي $a, b, \sum y, \sum xy$ وتحتاج إلى حساب $\sum y^2$ إضافة إلى ما تكون قد قمت بحسابه، ففي مثالنا السابق يكون الحل كالتالي :

الجدول (7-19)

x	y	xy	x ²	y ²
1	3	3	1	9
4	5	20	16	25
3	4	12	9	16
7	10	70	49	100
5	8	40	25	64
20	30	145	100	214

لقد تم حساب a, b سابقاً وكانت $a=1, b=1.25$

من العلاقة :

$$S_e = \sqrt{\frac{214 - 1 \times 30 - 1.25 \times 145}{5 - 2}}$$

$$= \sqrt{\frac{2.75}{3}} = 0.96$$

لاحظ سهولة الحل حسب قاعدة (1) أو حسب العلاقة (3) وذلك لعدم حاجتنا لحساب $\hat{y}_i$.

بعد أن وجدنا S_e ينشأ لدينا الحاجة إلى تقدير الانحراف المعياري للخطأ في التنبؤ :

إذا علمت $x = x^*$ وكان التنبؤ بالمعدل .

إذا علمت قيمة محددة للمتغير المستقل ولكن x^* فإنك تتنبأ بقيمة المتغير التابع لتكون في المعدل مساوية إلى $(a + bx^*)$.

ولهذا التنبؤ خطأ معياري يعتمد على الانحراف المعياري للمتغير Y . وما تحتاجه هو تقدير للانحراف المعياري للخطأ في التنبؤ وهو يعطى بالمقدار :

$$S_e \sqrt{\frac{1}{n} + \frac{(x^* - \bar{x})^2}{S_{xx}}}$$

حيث S_e هي تقدير الخطأ المعياري، و $\bar{x}$ هو المتوسط الحسابي للقيم، x^* هي القيمة المعطاة والمراد إيجاد التنبؤ y المقابل لها، n هو عدد أزواج المشاهدات المعطاة

$$S_{xx} = \sum (x - \bar{x})^2$$

إذا علمت x وكان تنبؤ y بقيمة مفردة أي إذا علمت أن $x = x^*$ قيمة محددة للمتغير المستقل وتنبأت المتغير التابع بقيمة واحدة فإن تقدير الانحراف المعياري للخطأ في التنبؤ هو :

$$S_e \sqrt{1 + \frac{1}{n} + \frac{(x^* - \bar{x})^2}{S_{xx}}}$$

مثال :

أوجد تقديراً للخطأ المعياري لمعدل التنبؤ إذا كان $x=6$ ثم أوجد تقديراً للخطأ المعياري عندما تتنبأ مشاهدة واحدة y والتي تقابل $x=6$.

الحل : عندما $x=6$ فإن المعدل يكون $A+B \times 6$ وهذا يقدر بالقيمة $y=a+b \times 6$ وتقدير الخطأ المعياري يكون

$$S_e \sqrt{\frac{1}{n} + \frac{(x^* - \bar{x})^2}{S_{xx}}}$$

$$= 0.96 \sqrt{\frac{1}{5} + \frac{(6 - 4)^2}{20}}$$

$$= 0.96 \sqrt{0.4} = 0.61$$

أما تقدير الخطأ المعياري عند تنبؤ قيمة مفردة فهو

$$S_e \sqrt{1 + \frac{1}{n} + \frac{(x^* - \bar{x})^2}{S_{xx}}}$$

$$= 0.96 \sqrt{1 + \frac{1}{5} + \frac{(6 - 4)^2}{20}}$$

$$= 0.96 \sqrt{1.4} = 1.61$$

أسئلة وتمارين عامة

السؤال الأول: أجب باختصار عن الأسئلة التالية :

- 1- ما الفرق بين الارباعي الثاني والوسيط ؟
- 2- ما مقياس النزعة المركزية والتشتت اللذان تختارهما عند وجود فئات مفتوحة؟
- 3- ماذا نستفيد من شكل الانتشار ؟
- 4- كيف تفسر الارتباط إذا كان الخطأ المعياري للتقدير يساوي الصفر ؟
- 5- ماذا يعني وجود منوالين للبيانات وما هو السبب ؟
- 6- ما المدرج التكراري ؟
- 7- لماذا يترك الباحث إحدى الفئات مفتوحة ؟
- 8- ما العينة ومتى تكون عشوائية ؟
- 9- ما معنى أن المجتمع الاحصائي مفهوم مرن ؟
- 10- ما أنواع الارتباط ؟
- 11- ما الفرق بين الثابت الاحصائي والتابع الاحصائي؟
- 12- ما الفرق بين المتوسط الحسابي والوسيط ؟
- 13- ما معنى البيانات المبوبة؟
- 14- ما الهدف من العرض الاحصائي بالرسم البياني ؟
- 15- متى يستخدم الانحدار ومتى يستخدم معامل الارتباط ؟
- 16- هل يقاس التشتت حول المنوال ؟ ولماذا؟
- 17- ما سبب وجود منوالين في توزيع ما

الحل :

- 1- الارباعي الثاني = الوسيط.
- 2- نختار الوسيط للنزعة المركزية والانحراف الربيعي للتشتت.
- 3- نستفيد من شكل الانتشار في تحديد طبيعة العلاقة بين المتغيرين .
- 4- يكون الارتباط تاماً عندما ينعدم الخطأ المعياري للتقدير .
- 5- يعني وجود منوالين أن التوزيع ثنائي المنوال بسبب عدم تجانس البيانات .
- 6- المدرج التكراري هو أحد منحنيات التكرار وهو مجموعة من المستطيلات المتلاصقة متساوية في القاعدة ومختلفة في الارتفاع .
- 7- يترك الباحث أحد الفئات مفتوحة للمحافظة على سرية البيانات وإمكانية إضافة قيم أخرى .
- 8- العينة هي جزء صغير من المجتمع وتكون عشوائية عند إعطاء كل مفردة في المجتمع نفس الفرصة للظهور في العينة .
- 9- معنى أن المجتمع الإحصائي مفهوم مرن لإمكانية التحكم بإطاره الزماني والمكاني
- 10- أنواع الارتباط : متعدد - جزئي - بسيط
- 11- الثابت الإحصائي هو مقياس مجتمع أما التابع الإحصائي فهو مقياس عينة .
- 12- المتوسط الحسابي مقياس قيم أما الوسيط فمقياس ترتيب.
- 13- البيانات المبوبة: هي بيانات منظمة في جدول تكراري.
- 14- الهدف من العرض الإحصائي هو إعطاء فكرة أوضح عن الظاهرة ثم إظهار خصائص لا تستطيع الجداول التكرارية إظهارها .
- 15- الانحدار يشترط وجود علاقة سببية بين المتغيرين أما الارتباط فلا يشترط ذلك.

أسئلة وتمارين عامة

- 16- لا يقاس التشتت حول المنوال لأنه في قمة التوزيع وهو دوماً قيمة ممثلة للنزعة المركزية، كما انه يمكن أن يوجد أكثر من منوال للبيانات .
 - 17- التقدير بنقطة يتمثل بالحصول على قيمة واحدة اما التقدير بفاصلة فهو الحصول على حد أدنى وحد أعلى .
 - 18- يكون للجدول التكراري أكثر من منوال لعدم تجانس البيانات .
- السؤال الثاني :** ضع إشارة (✓) أما العبارة الصحيحة و (×) أمام العبارة الخاطئة :
- 1- نستفيد من شكل الانتشار في تحديد طبيعة العلاقة بين المتغيرين (×)
 - 2- يمكن أن يقاس التشتت حوال المنوال (×)
 - 3- العرض الاحصائي يظهر ميزات لا تستطيع الجداول التكرارية إظهارها (✓)
 - 4- البيانات الأولية أكثر أهمية من البيانات الثانوية (✓)
 - 5- المتغير التابع هو الأسبق زمنياً (×)
 - 6- المضلع التكراري هو منحنى أزيلت زواياه (×)
 - 7- التباين هو جذر الانحراف المعياري (×)
 - 8- المجتمع الاحصائي مفهوم مرن لامكانية التحكم بإطاره الزمني والمكاني (✓)
 - 9- يقيس الخطأ المعياري للتقدير انتشار القيم حول المتوسط الحسابي (✓)
 - 10- النوعيات وحدات يعبر عنها بصورة رقمية (×)
 - 11- الجداول التكرارية تعقد العمل وتضيع الوقت (×)
 - 12- يمكن أن يقاس التشتت حول المنوال (×)
 - 13- يمكن أن يكون للبيانات أكثر من منوال (✓)
 - 14- يمكن أن يكون للبيانات أكثر من وسيط (×)
 - 15- معامل الاختلاف هو الانحراف الربيعي النسبي (×)

- 16- التباين هو مربع الانحراف المعياري (×)
- 17- علاقة بيرسون تعطي قيمة صحيحة دوماً (×)
- 18- الارتباط يشترط وجود علاقة سببية بين المتغيرين (×)
- 19- المتغيرات وحدات يعبر عنها بنسب مئوية (×)
- 20- العينة أصغر جزء داخل المجتمع (×)
- 21- عينات الحصص هي عينات عشوائية (×)
- 22- يمكن أن يكون للبيانات أكثر من متوسط حسابي (×)
- 23- التباين هو مربع الانحراف الربيعي (×)
- 24- ينعدم الارتباط إذا كان التباين المفسر معدوماً ($\sqrt{1}$)
- 25- ينعدم الارتباط إذا كان التباين غير المفسر معدوماً (×)
- 26- ينعدم الارتباط إذا كان الخطأ المعياري للتقدير معدوماً (×)
- 27- ينعدم الارتباط إذا كان معامل التحديد معدوماً ($\sqrt{1}$)
- 28- عينات الحصص يسحبها الباحث بإرادته ($\sqrt{1}$)

السؤال (3): أكمل الفراغات التالية بالكلمات المناسبة :

يقيس الانحراف تشتت البيانات حول عادة، كما يقيس الانحراف الربيعي
التشتت حول ويعتبر أفضل مقاييس التشتت إذا كانت الفئات عند وجود
منوالين للبيانات يسمى التوزيع

مسألة (1) :

بالاطلاع على سجلات الامتحانات وجد أن مجموع درجات 700 طالب في مقرر علم النفس قد
بلغ 36750 درجة، والانحراف المعياري 10.5 درجة . أما في مقرر الاحصاء فكان متوسط
درجاتهم 50 درجة والتباين 49 درجة والمطلوب :

----- أسئلة وتمارين عامة

- أي العينتين ذات تشتت أكبر، درجات الطلاب في الإحصاء أم درجاتهم في علم النفس.

الحل :

$$n_1 = 700 \quad \Sigma x = 36750 \quad S_1 = 10.5$$

$$n_2 = 700 \quad \bar{x} = 50 \quad S_2^2 = 49 \quad S_2 = 7 \quad \text{لأنها جذر}$$

$$\bar{x} = \frac{\Sigma x}{n} = \frac{36750}{700} = 52.5$$

$$C.V = \frac{s}{\bar{x}} \times 100$$

$$C.V \text{ علم النفس} = \frac{10.5}{52.5} \times 100 = 20\%$$

$$C.V \text{ إحصاء} = \frac{7}{50} \times 100 = 14\%$$

درجات علم النفس ذات تشتت أكبر لأن معامل اختلافها أكبر

مسألة (2) : تبين القيم التالية عدد ساعات الدراسة لعينة من 8 طلاب وترتيبهم في الصف كما يلي:

10	9	8	7	6	4	3	2	عدد الساعات
9	8	10	6	3	5	4	1	ترتيب في الصف

----- أسئلة وتمارين عامة

إذا علمت أن $\sum xy = 336$ ، $\sum y^2 = 332$ ، $\sum x^2 = 359$ ، معامل الانحدار = 0.92.

1- احسب معامل ارتباط الرتب وفسره وبين أهميته الإحصائية.

الحل :

عدد ساعات الدراسة الترتيب في الصف :

الجدول (1)

x	y	رتبة x	رتبة y	D	D ²
2	1	8	8	0	0
3	4	7	6	1	1
4	5	6	5	1	1
6	3	5	7	2-	4
7	6	4	4	0	0
8	10	3	1	2	4
9	8	2	3	1-	1
10	9	1	2	1-	1
$\Sigma x = 49$	$\Sigma y = 46$				$\Sigma D^2 = 12$

$$R_s = 1 - \frac{6 \sum D_i^2}{n(n^2 - 1)}$$

$$R_s = 1 - \frac{6 \times 8}{8(64 - 1)}$$

$$R_s = 1 - \frac{72}{504} = 1 - 0.142$$

أسئلة وتمارين عامة

$$R_s = 0.857$$

الأهمية الإحصائية

$$S_r = \sqrt{\frac{1-r^2}{n-2}} = \sqrt{\frac{1-(0.857)^2}{8-2}}$$

$$S_r = \sqrt{\frac{1-0.734}{6}}$$

$$S_r = \sqrt{\frac{0.266}{6}}$$

$$S_r = \sqrt{0.044} = 0.2105$$

$$0.857 > 0.2105 \times 3$$

$$0.857 > 0.6316$$

تحققت العلاقة، إذن لمعامل الارتباط أهمية إحصائية

التفسير : معامل الارتباط بين عدد ساعات الدراسة وترتيب الطالب في الصف ارتباط قوي

----- أسئلة وتمارين عامة

المسألة (3) : لدينا البيانات التالية :

الجدول (2)

الفئات	f	Xi	d _o	fd _o	² d _o	f. d _o ²	%f ↑
أقل من 20	14	10	2-	28-	4	56	14
20-40	16	30	1-	16-	1	16	30
40-60	40	50	0	0	0	0	70
60-80	16	70	1	16	1	16	86
80 فأكثر	14	90	2	28	4	56	100
						$\Sigma = 144$	$\Sigma = 27344$

والمطلوب:

- 1- ما طبيعة التوزيع الذي تخضع له البيانات .
- 2- حساب معامل الاختلاف .
- 3- ما الدرجة التي يقل عنها 75% من البيانات.
- 4- ما عدد الطلاب الذين تزيد درجاتهم عن 74

الحل :

لتحديد اتجاه الالتواء لابد من حساب المتوسط والمنوال .

----- أسئلة وتمارين عامة

$$Mod = 40 + \left(\frac{24}{24+24}\right) \times 20$$

$$Mod = 40 + 10 = 50$$

المتوسط :

$$\bar{X} = 50 + \frac{0}{100} \times 10$$

$$\bar{X} = 50$$

بما أن المتوسط = المنوال فالتوزيع متماثل ولا يوجد فيه التواء

2- معامل الاختلاف :

$$S = \sqrt{\frac{144}{100} - \frac{(0)^2}{100}} \times 20 = 24$$

$$S = \frac{24}{50} \times 100 = 48\%$$

3) الدرجة التي حصل عليها أقل من 75% من الطلاب هي الربع الثالث

$$Q_3 = 3 \times \frac{100}{4} = 75$$

$$Q_3 = 60 + \left(\frac{75-70}{16}\right) \times 20 = 66.25$$

4) عدد الطلاب الذين تزيد درجاتهم عن 74 درجة هو :

$$= \frac{74-50}{24} = +1$$

المساحة 34.13%

----- أسئلة وتمارين عامة

المساحة المطلوبة 90% + 34.13% = 84.13%

$$\text{العدد المطلوب} = 100 \times \frac{84.13}{100} = 84$$

المسألة (4) :

لدينا البيانات التالية والتي تمثل درجات الطلاب في مادتي اللغة العربية والإنكليزية

الجدول (3)

x	y	x.y	x ²	y ²
4	1	4	16	1
8	5	40	64	25
6	3	18	36	9
10	7	70	100	49
14	11	154	196	121
12	9	108	144	81
14	11	154	196	121
Σx=68	Σy=47	=548 Σx.y	Σx ² = 752	Σy ² = 407

المطلوب :

- 1- حساب معامل الارتباط وبيان أهميته .
- 2- حساب معادلة مستقيم الانحدار وتفسير ثوابتها.

$$\bar{x} = \frac{68}{7} = 9.71$$

----- أسئلة وتمارين عامة

$$\bar{y} = \frac{47}{7} = 6.71$$

(1) حساب معادلة مستقيم الانحدار وتفسير ثوابتها:

$$R = \frac{n \sum xy - \sum x \cdot \sum y}{\sqrt{[n \sum x^2][n \sum y^2 - (\sum y)^2]}}$$

$$R = \frac{7 \times 548 - 68 \times 47}{\sqrt{[7 \times 752 - (68)^2][7 \times 407 - (47)^2]}}$$

$$R = \frac{3836 - 3196}{\sqrt{[5264 - 4624][2849 - 2209]}}$$

$$R = \frac{640}{\sqrt{640 \times 640}} = \frac{640}{640} = 1$$

الارتباط تام

$$Sr = \sqrt{\frac{1 - (1)^2}{7 - 2}} = 0$$

$$Sr = \sqrt{\frac{0}{5}} = 0$$

$$0 \times 3 < 1$$

تحققت العلاقة لمعامل الارتباط أهمية.

(2) حساب معادلة مستقيم الانحدار:

$$\hat{y} = a + bx$$

$$b = \frac{548 - 971 \times 47}{752 - 971 \times 68}$$

$$= \frac{91.63}{91.72} = 0.999 \approx 1$$

$$a = \bar{y} - b\bar{x}$$

----- أسئلة وتمارين عامة

$$\alpha = 6.71 - (1 \times 9.71) = -3$$

تفسير α : عندما تكون درجة الطالب في اللغة العربية صفر، فإن درجته في اللغة الإنكليزية -
3 مقبول إحصائياً مرفوض منطقياً

تفسير b : عندما تزداد درجة الطالب في اللغة العربية درجة واحدة تزداد درجته درجة واحدة في اللغة الإنكليزية.

(3) حساب معامل الارتباط وبيان أهميته الإحصائية :

$$Sr = \sqrt{\frac{1 - (1)^2}{7 - 2}} = 0.99$$

$$Sr = 0$$

$$1 > 3 \times 0$$

تحققت العلاقة لمعامل الارتباط أهمية إحصائية .

أسئلة وتمارين عامة -----

(4) حساب معامل ارتباط الرتب:

الجدول (4)

x	y	رتبة x	رتبة y	D	D ²
4	1	7	7	0	0
8	5	5	5	0	0
6	3	6	6	0	0
10	7	4	4	0	0
14	11	1.5	1.5	0	0
12	9	3	3	0	0
14	11	1.5	1.5	0	0
					Σ

$$R_s = 1 - \frac{6 \times 0}{7(49-1)} = 1 - 0 = 1$$

الارتباط تام وطردي

----- أسئلة وتمارين عامة -----

المسألة (5) : يمثل الجدول التالي درجات عينة من الطلاب في مقرر الاحصاء

فئات الدرجة	أقل من 25	25 وأقل من 35	35 وأقل من 45	45 وأقل من 55	55 وأقل من 65
عدد الطلاب	10	20	30	20	10

المطلوب :

- 1- حدد اتجاه الالتواء إن وجد.
- 2- احسب الوسيط عن طريق علاقة بيرسون .
- 3- وجد أن عينة مماثلة معامل اختلافها 20% فأَي العينتين تظهر تشتتاً أكبر ولماذا.
- 4- احسب الربيع الأول وفسره

الحل :

الجدول (5)

الفئات	f	X	d _o	fd _o	d _o ²	f. d _o ²	%f↑
أقل من 25	10	20	2-	20-	4	40	10
35-25	20	30	1-	20-	1	20	30
45-35	30	40	0	0	0	0	60
55-45	20	50	1	20	1	20	80
65-55	10	60	2	20	4	40	90
	90				$\sum = 0$	$\sum = 120$	

(1) حدد اتجاه الالتواء إن وجد:

$$Mod = 35 + \left(\frac{10}{10+10} \right) \times 10 = 40$$

أسئلة وتمارين عامة

$$\bar{X} = 40 + \left(\frac{0}{90}\right) \times 10 = 40$$

لا يوجد التواء والتوزيع متماثل

(2) احسب الوسيط عن طريق علاقة بيرسون :

$$Mod = 3Med - 2\bar{x}$$

$$40 = 3Med - 2 \times 40$$

$$3Med = 80 + 40$$

$$Med = \frac{120}{3} = 40$$

(3) وجد أن عينة مماثلة معامل اختلافها 20% فأبي العينتين تظهر تشتتاً أكبر ولماذا.

$$S = \sqrt{\frac{120}{90} - \frac{(0)^2}{90}} \times 10 = 11.54$$

$$= \frac{11.54}{40} \times 100 = 28.85\%$$

القيمة الأولى ذات تشتت أكبر لأن معامل اختلافها أكبر.

(4) احسب الربيع الأول وفسره

$$= \frac{90}{4} = 22.5$$

$$= 25 + \left(\frac{22.5 - 10}{20}\right) \times 10 = 31.25$$

التفسير : إن 25% من الطلاب درجاتهم أقل من 31.25 و 75% أكثر من 31.25 درجة .

----- أسئلة وتمارين عامة

المسألة (6) : درست العلاقة بين أطوال وأوزان عينة من 6 أفراد فكانت النتائج كما يلي :

$27 = \sum y$	$27.5 = \sum D^2$	$0.91 = S_y^2$
$36 = \sum x$	$169 = \sum xy$	$127 = \sum y^2$
		$230 = \sum x^2$

المطلوب :

- 1- حساب معادلة الانحدار وتفسير ثوابتها مبيناً شرطاً المعادلة .
- 2- حساب معامل الارتباط (برسون) وتفسيره وتبيان أهميته الاحصائية.
- 3- حساب معامل ارتباط الرتب (سبيرمان) وتفسيره

الحل :

1) حساب معادلة الانحدار وتفسير ثوابتها مبيناً شرطاً المعادلة .

$$\bar{x} = \frac{36}{6} = 6$$

$$\bar{y} = \frac{27}{6} = 4.5$$

$$b = \frac{169 - 6 \times 27}{230 - 6 \times 36} = \frac{7}{14} = 0.5$$

$$\alpha = 4.5 - 0.5 \times 6 = 1.5$$

$$y = 1.5 + 0.5b$$

----- أسئلة وتمارين عامة

التفسير :

α : عندما يكون الطول معدوماً فإن الوزن 1.5 مقبول إحصائياً مرفوض واقعياً.

b : كلما ازداد الطول مرة واحدة يزداد الوزن وسطياً 0.5 وحدة

(2) حساب معامل الارتباط (بيرسون) وتفسيره وتبيان أهميته الإحصائية:

$$R = \frac{n \sum xy - \sum x \cdot \sum y}{\sqrt{[n \sum x^2 - (\sum x)^2][n \sum y^2 - (\sum y)^2]}}$$

$$R = \frac{6 \times 169 - 36 \times 27}{\sqrt{[6 \times 230 - (36)^2][6 \times 127 - (27)^2]}}$$

$$R = \frac{1014 - 972}{\sqrt{[1380 - 1296][762 - 729]}}$$

$$R = \frac{42}{\sqrt{84 \times 33}} = \frac{42}{52.649} = 0.79$$

$0.79 < 0.3 \times 3 < 0.9$ لم تتحقق العلاقة ليس لمعامل الارتباط أهمية إحصائية.

(3) حساب معامل ارتباط الرتب (سبيرمان) وتفسيره:

$$R_s = 1 - \frac{6 \times 27.5}{6(36-1)} = 1 - \frac{165}{210} = 1 - 0.78 = 0.22$$

الارتباط ضعيف وطردي

----- أسئلة وتمارين عامة

المسألة (7) : في اختبار القياس والتقويم عملي لطلاب دبلوم التأهيل التربوي كانت نتائج عينة

عشوائية من 100 طالب كما يلي :

المطلوب :

فئات الدرجات	10 وأقل من 14	14-18	18-22	22-26	26 فما فوق
عدد الطلاب	8	10	35	30	17

- حساب الدرجة التي يقل عنها 75% من الدرجات
- حساب الدرجة الثانية للدرجة 26.5
- حساب عدد الطلاب الذين تزيد درجاتهم عن 18 درجة
- رسم مدرج تكراري للجدول السابق .

الحل :

الجدول (6)

الفئات	f	X	d _o	fd _o	d _o ²	f. d _o ²	f ↑	f ↓
10-14	8	12	2-	16-	4	32	8	100
14-18	10	16	1-	10-	1	10	18	92
18-22	35	20	0	0	0	0	53	82
22-26	30	24	1	30	1	30	83	47
26 فأكثر	17	28	2	34	4	68	100	17
	100			$\Sigma = 38$		$\Sigma = 140$		

----- أسئلة وتمارين عامة

(1) حساب الدرجة التي يقل عنها 75% من الدرجات :

هو الربيع الثالث :

$$Q_3 = \frac{3 \times 100}{4} = 75$$

$$Q_3 = 22 + \left(\frac{75 - 53}{30} \right) \times 4 = 24.93$$

(2) حساب الدرجة الثانية للدرجة 26.5

$$= 20 + \left(\frac{38}{100} \right) \times 4 = 21.52$$

$$= \sqrt{\frac{140}{100} - \frac{(38)^2}{100}} \times 4 = 4.48$$

$$= 50 + 10 \times \left(\frac{26.5 - 21.525}{4.48} \right) = 61.11 = 61 \text{ درجة}$$

(3) حساب عدد الطلاب الذين تزيد درجاتهم عن 17 درجة :

$$= \frac{17 - 21.52}{4.48} = -1$$

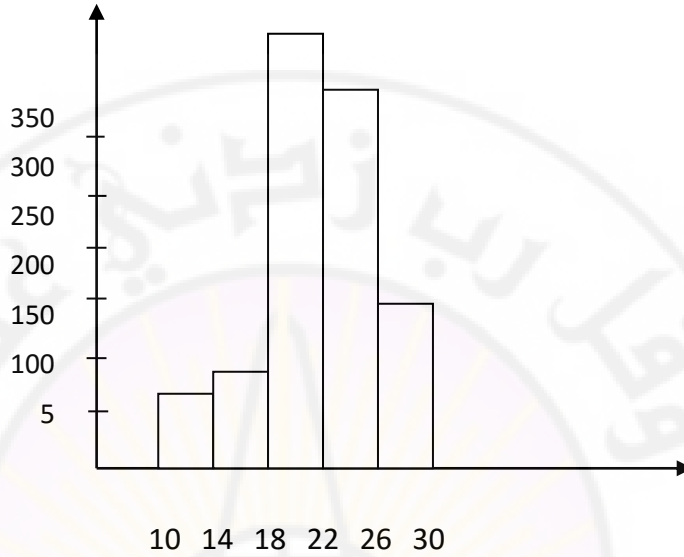
$$\%84.13 = \%50 + \frac{78.26\%}{2} = \text{المساحة}$$

$$\text{العدد المطلوب} = 100 \times \frac{84.13}{100} = 84 \text{ طالباً}$$

(4) حساب عدد الطلاب الذين تزيد درجاتهم عن 18 درجة :

العدد هو 82 طالب

(6) رسم مدرج تكراري للجدول السابق :



المسألة (8) :

درست العلاقة بين أسعار الكتب وعدد القراء لعينة عشوائية من 6 مكتبات فكانت النتائج التالية:

معامل التحديد = 0.52

معامل الانحدار = -0.58

فإذا علمت ان الوحدات مقدرة بالعشرات، وأن الأسعار تتراوح بين 50 و 100 ليرة

والمطلوب :

- بين التباين المفسر بالرسم.
- بين الأهمية الإحصائية لمعامل الارتباط.

أسئلة وتمارين عامة

- احسب معامل ارتباط الرتب وفسره .
- قدر بالتالي عدد القراء لكتاب سعره 95 ليرة، وآخر سعره 150 ليرة

الحل :

(1) بين التباين المفسر بالرسم.

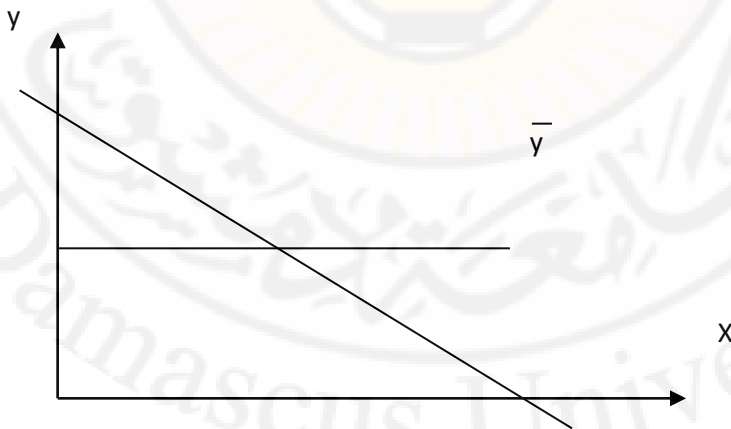
$$\bar{x} = \frac{44}{6} = 7.33$$

$$\bar{y} = \frac{31}{6} = 5.16$$

$$a = \bar{y} - b \bar{x}$$

$$a = 5.16 - (-0.58) \times 7.33 = 9.41$$

$$\hat{y} = 9.41 - 0.58x$$



----- أسئلة وتمارين عامة

(2) بين الأهمية الإحصائية لمعامل الارتباط:

$$Sr = \sqrt{\frac{1-0.52}{6-2}} = 0.34$$

$$R = \sqrt{(0.52)^2} = 0.72$$

$0.72 < 0.34 \times 3$ أي $0.72 < 1.02$ لذلك لم تتحقق العلاقة إذن ليس لمعامل الارتباط أهمية إحصائية.

(3) احسب معامل ارتباط الرتب وفسره :

$$R = 1 - \frac{6 \times 10.5}{6(36-1)} = 1 - \frac{63}{210} = 1 - 0.3 = 0.7$$

الارتباط شديد وطردى.

(4) قدر بالتالي عدد القراء لكتاب سعره 95 ليرة، وآخر سعره 150 ليرة

$$x = 95 \rightarrow = 9.41 - 0.58 \times 9.5 = 3.9$$

$$x = 150$$

لا تستطيع التقدير لأنه خارج مجال قيم x

أسئلة وتمارين عامة

المسألة (9) : درست عينة من 130 طالب وقيس حاصل الذكاء حسب مقياس فكسلر فكان توزيعهم التكراري كما يلي :

فئات حاصل الذكاء	أقل من 70	70-80	80-90	90-110	110-120	120-130	130 فأكثر
عدد الطلاب	5	20	25	30	25	20	5

والمطلوب :

- 1- ما عدد الطلاب الذين يقل حاصل الذكاء عندهم عن 80 درجة .
- 2- ما حاصل الذكاء الذي يقل عنه 75% من الطلاب .
- 3- وُجِدَ أن عينة ثانية ربعها الثاني 80 ومداها الربيعي 40 فأَي العينتين ذات تجانس أكبر ولماذا .
- 4- ارسم مدرج تكراري للجدول السابق .
- 5- حدد اتجاه الالتواء دون حساب .

الحل :

الجدول (7)

الفئات	f	$f \uparrow$
أقل من 70	5	5
70-80	20	25
80-90	25	50
90-110	30	80
110-120	25	105
120-130	20	125
130 فأكثر	5	130
المجموع	130	

----- أسئلة وتمارين عامة

(1) عدد الطلاب الذين يقل حاصل الذكاء عندهم عن 80 درجة هو 25 طالب من f

(2) ما حاصل الذكاء الذي يقل عنه 75% من الطلاب :

$$Q_3 = \frac{3 \times 130}{4} = 97.5$$

$$Q_3 = 110 + \left(\frac{97.5 - 80}{25} \right) \times 10 = 117$$

(3) نحسب الانحراف الربيعي النسبي لكل من العينين ونقارنهما

نحسب الربيع الأول :

$$Q_1 = 80 + \left(\frac{32.5 - 25}{25} \right) \times 10 = 83$$

نحسب الوسيط:

$$Med = 90 + \left(\frac{65 - 50}{30} \right) \times 20 = 100$$

$$Q\% = \frac{117 - 83}{2 \times 100} \times 100 = 34\%$$

$$Q\% = \frac{40}{2 \times 80} \times 100 = 50\%$$

تشنت العينة الثانية أكبر

(5) اتجاه الالتواء لا يوجد التواء لأن التكرارات تصعد وتهبط بانتظام

----- أسئلة وتمارين عامة

المسألة (10) : درست العلاقة بين العمر الزمني والعمر العقلي لعينة من 8 أفراد فأعطت

النتائج التالية :

$\sum x = 80$ ، $\sum y = 128$ ، $\sum (y - \bar{y})^2 = 38$ ، $6 =$ معامل الانحدار $= 0.65$
--

والمطلوب :

1- من المتغير المستقل ومن هو المتغير التابع ولماذا ؟

الحل :

المتغير المستقل : العمر الزمني

المتغير التابع : العمر العقلي

3) معامل الانحدار $= 0.65$

$$\bar{x} = \frac{80}{8} = 10$$

$$\bar{y} = \frac{128}{8} = 16$$

$$\alpha = 16 - (0.65 \times 8) = 10.8$$

$$\hat{y} = 10.8 + 0.65x$$

----- أسئلة وتمارين عامة

المسألة (11) : درست عينة من 66 طالب من الملتحقين بمعاهد التربية الفكرية فكان التوزيع التكراري لدرجات ذكائهم كما يلي :

فئات الذكاء	50-44	56-50	62-56	68-62	74-68	74 فأكثر
عدد الطلاب	3	7	13	25	12	6

والمطلوب :

- 1- ما مقياس النزعة المركزية والتشتت المناسبان لحساب هذه البيانات ولماذا؟
- 2- حدد اتجاه الالتواء إن وجد
- 3- وُجِدَ أن عينة مماثلة من طلاب المدارس العادية متوسطها الحسابي 82 ومعامل اختلافها 9.2% فأَي العينتين أكثر تشتتاً ولماذا ؟
- 4- احسب المتوسط الحسابي للعينتين معاً
- 5- من العينة الأولى احسب الدرجة التائية لحاصل الذكاء 71.35

الحل :

الجدول (8)

الفئات	f	X	d _o	fd _o	d _o ²	f. d _o ²
44-50	3	47	3-	9-	9	27
50-56	7	53	2-	14-	4	28
62-56	13	59	1-	13-	1	13
62-68	25	65	0	0	0	0
62-74	12	71	1	12	1	12
74 فأكثر	6	77	2	12	4	24
	$\sum f = 60$			$\sum = -12$		$\sum = 104$

----- أسئلة وتمارين عامة

(1) نختار الوسيط للنزعة المركزية والانحراف الربيعي للتشتت لوجود فئة مفتوحة.

(2) تحديد اتجاه الالتواء :

$$Mod = 62 + \left(\frac{12}{12+13}\right) \times 6 = 64.88$$

$$\bar{X} = 65 + \left(\frac{-12}{66}\right) \times 6 = 63.9$$

← $Mod > \bar{X}$ الالتواء سالب نحو اليسار .

(3)

$$S = \sqrt{\frac{104}{66} - \frac{(-12)^2}{66}} \times 6 = 7.45$$

$$C.V = \frac{7.45}{63.9} \times 100 = 11.65\%$$

$$= 9.2\%$$

القيمة الأولى ذات تشتت أكبر لأن معامل اختلافها أكبر .

(4) المتوسط الحسابي المرجح أو الموزون

$$= \frac{63.9+82}{2} = 72.95 \quad \text{لأن } n_1 = 2$$

$$= 50 + 10 \times \left(\frac{71.3-63.9}{7.45}\right) = 60 \quad \text{درجة}$$

----- أسئلة وتمارين عامة

المسألة (12) : درست العلاقة بين ذكاء الآباء والأبناء لعينة من 6 أزواج فأعطت النتائج التالية :

$\sum xy = 458$	$S^2y = 1.18$	$\sum y = 50$	$\sum X = 53$
$\sum D^2 = 13$	$= \sum y^2 = 438$	$\sum x^2 = 493$	

المطلوب :

- 1- احسب المعادلة وفسرها
- 2- احسب معامل الارتباط وبين أهميته الإحصائية
- 3- احسب معامل ارتباط للرتب وفسره

الحل :

$$\bar{x} = \frac{53}{6} = 8.83$$

$$\bar{y} = \frac{50}{6} = 8.33$$

$$1)b = \frac{458 - 8.83 \times 50}{493 - 8.83 \times 53} = 0.65$$

$$\alpha = 8.33 - 0.65 \times 8.83 = 2.59$$

$$\hat{y} = 2.59 + 0.65x$$

وهي معادلة مستقيم الانحدار

التفسير :

- ج: عندما يكون ذكاء الآباء معدوماً فإن ذكاء الأبناء 2.59
- ب: عندما ازداد ذكاء الآباء وحدة واحدة يزداد ذكاء الأبناء 0.65 وحدة .

أسئلة وتمارين عامة

$$R = \frac{n \sum xy - \sum x \cdot \sum y}{\sqrt{[n \sum x^2][n \sum y^2 - (\sum y)^2]}}$$

$$R = \frac{6 \times 458 - 53 \times 50}{\sqrt{[6 \times 493 - (53)^2][6 \times 438 - (50)^2]}}$$

$$R = \frac{2748 - 2650}{\sqrt{[2958 - 2809][2628 - 2500]}}$$

$$R = \frac{98}{\sqrt{[149 \times 128]}} = \frac{98}{138.10} = 0.7$$

$$Sr = \sqrt{\frac{1 - (0.7)^2}{6 - 2}} = 0.35$$

$$0.35 \times 3 < p.7$$

لم تتحقق العلاقة ولذلك ليس لمعامل الارتباط أهمية إحصائية

-3

$$R = 1 - \frac{6 \times 13}{6(36 - 1)} = 1 - \frac{78}{210} = 1 - 0.37 = 0.63$$

التفسير : الارتباط يبين رتب نكاح الآباء ونكاح الأبناء شديد وطردي

المسألة (13):

تبين القيم التالية درجة الإحصاء والرياضيات لعينة من الطلاب مقاسة بالعشرات

$\Sigma(x-\bar{x})(y-\bar{y})=170$	$\Sigma(x-\bar{x})^2=269$	$\Sigma(y-\bar{y})^2=269$	$\Sigma D^2=40$
$\Sigma x = 70 \quad \Sigma y = 61$	$\Sigma xy=444$	$\Sigma x^2= 504$	$\Sigma y^2=399 \quad n=10$

مجال الدرجات بين 90-40 درجة .

المطلوب :

- إذا علمت أن معامل الانحدار 1.2 احسب معادلة مستقيم الانحدار وفسر ثوابتها .
- بين الفائدة من شكل الانتشار وما أنواع الارتباط ؟
- حساب معامل ارتباط سبيرمان وتفسيره .
- حساب معامل التحديد وتفسيره بضوء البيانات، وتبيان الأهمية الإحصائية لمعامل الارتباط مبيناً معنى النتيجة التي توصلت إليها .
- قدر درجة طالب في الإحصاء إذا كانت درجة الرياضيات 75 وآخر 92 درجة.

الحل :

1) معامل الانحدار هو $b = 1.21$ نحسب a :

$$a = \bar{y} - b \bar{x}$$

$$\bar{x} = \frac{\Sigma x}{n} = \frac{70}{10} = 7$$

$$\bar{y} = \frac{\Sigma y}{n} = \frac{61}{10} = 6.1$$

----- أسئلة وتمارين عامة

$$a = 6.1 - 1.21 (7) = 6.1 - 8.47 = -2.37$$

وتكون المعادلة :

$$Y = a + b x = -2.37 + 1.21 x$$

ويمكن حسابها عن طريق المعادلتين الطبيعيين:

$$\Sigma y = na + b \Sigma x$$

$$\Sigma xy = a \Sigma x + b \Sigma x^2$$

نعوض فيكون :

$$61 = 10a + 70 b$$

$$444 = 70a + 504b$$

نحسب a من المعادلة الأولى ونعوضها في المعادلة الثانية :

$$10a = 61 - 70b \rightarrow a = \frac{61 - 70b}{10}$$

بالتعويض بالثانية

$$444 = 70\left(\frac{61 - 70b}{10}\right) + 504 b$$

$$444 = 427 - 490b + 504b$$

ننقل المعاليم إلى طرف والمجاهيل إلى الطرف الآخر :

$$444 - 427 = 504b - 490b$$

أسئلة وتمارين عامة

$$17=14b \rightarrow b = \frac{17}{14} = 1.21$$

نعوض b بقيمتها في a فيكون

$$a = \frac{61-70(1.21)}{10} = -2.37$$

وتكون المعادلة :

$$\hat{y} = -2.37 + 1.21x$$

التفسير :

a : عندما تكون درجات الرياضيات معدومة فإن درجات الإحصاء -2.37 درجة (مقبول إحصائياً مرفوض واقعياً) لأن الدرجات لا يمكن أن تكون سالبة .

b : كلما ازدادت درجات الرياضيات عشر درجات تزداد درجات الإحصاء وسطياً 12.1 درجة .

(2) بين الفائدة من شكل الانتشار وما أنواع الارتباط ؟

1- الفائدة من شكل الانتشار هو معرفة طبيعة العلاقة بين المتغيرين .

2- أنواع الارتباط : بسيط - جزئي - متعدد .

(3) حساب معامل ارتباط سبيرمان وتفسيره :

$$N=10, \Sigma D^2= 40$$

$$R_s = 1 - \frac{6 \Sigma D^2}{n(n^2-1)}$$

أسئلة وتمارين عامة

$$R_s = 1 - \frac{6(40)}{10(100-1)}$$

$$R_s = 1 - 0.24$$

$$R_s = 0.76$$

التفسير : إن معامل الارتباط بين رتب درجات الإحصاء والرياضيات شديد وطردى .

(4) حساب معامل التحديد وتفسيره بضوء البيانات . وتبيان الأهمية الإحصائية لمعامل الارتباط مبيناً من النتيجة التي توصلت إليها .

معامل التحديد R^2 هو مربع معامل الارتباط R :

$$r = \frac{\sum(x-\bar{x})(y-\bar{y})}{\sqrt{\sum(x-\bar{x})^2 \sum(y-\bar{y})^2}} = \frac{170}{\sqrt{140(269)}}$$

$$r = \frac{170}{190.06} = 0.89$$

ويكون معامل التحديد :

$$R^2 = (0.89)^2 = 0.79$$

التفسير : إن 79% من التباين في درجات الإحصاء يفسره التباين في درجات الرياضيات .

(5) الأهمية الإحصائية لمعامل الارتباط :

يجب حساب الخطأ المعياري لمعامل الارتباط :

$$S_r = \sqrt{\frac{1-r^2}{n-2}} = \sqrt{\frac{1-0.79}{10-2}} = 0.16$$

ويجب تحقق العلاقة : $R > 3S_r$

----- أسئلة وتمارين عامة

لكي يكون لمعامل الارتباط أهمية إحصائية:

$$R = 0.89 > 3(0.16)$$

$$0.89 > 0.48$$

تحققت العلاقة إذن لمعامل الارتباط أهمية إحصائية .

ومعنى النتيجة : أن معامل ارتباط المجتمع لا يساوي الصفر والعينة تمثل مجتمعها الإحصائي.

6) قدر درجة طالب في الإحصاء إذا كانت درجة الرياضيات 76 وآخر 92 درجة.

1- الدرجة $x=76$ نقسمها على 10 لأن الدرجات بالعشرات :

$$\frac{76}{10} = 7.6$$

نعوض بالمعادلة

$$y = -2.37 + (1.21 (7.6)) = 6.826$$

نضرب الناتج بـ 10 ← $y=68$

يتوقع أن تكون درجة الإحصاء 68 درجة لطالب درجته بالرياضيات 76 درجة.

2- الدرجة 92 لا نستطيع التقدير لهذه القيمة لأنها خارج مجال قيم x لأن مجال الدرجات كما

هو بنص المسألة (40-90) درجة .

----- أسئلة وتمارين عامة

المسألة (14) :

لدينا البيانات التالية التي تبين درجات خمس أفراد في التحصيل الدراسي في اختبار القلق

القلق	3	4	7	5	8
التحصيل	2	4	5	7	6
$D=$	3.5-	1.5-	1	2	3.5

والمطلوب :

- 1- بين الفائدة من شكل الانتشار ؟
- 2- حساب معامل ارتباط سبيرمان وتفسيره للبيانات الموجودة في الصف D .
- 3- حساب معامل التحديد وتفسيره بضوء البيانات. وتبيان الأهمية الإحصائية لمعامل الارتباط مبيناً معنى النتيجة التي توصلت إليها .

الحل :

- (1) الفائدة من شكل الانتشار هي تحديد طبيعة العلاقة بين المتغيرين .
- (2) حساب معامل ارتباط سبيرمان وتفسيره :

$$R_s = 1 - \frac{6 \sum D^2}{n(n^2 - 1)}$$

نحسب $\sum D^2$ لدينا قيم D بنص المسألة نربعها ثم نجمع :

$$D = -3.5 \quad -1.5 \quad 1 \quad 2 \quad 3.5$$

$$\sum D^2 = 12.25 + 2.25 + 1 + 4 + 12.25 = 31.75$$

$$r_s = 1 - \frac{6(31.7)}{5(25-1)} = 1 - \frac{190.5}{120} = 1 - 1.58 = -0.58$$

التفسير : معامل الارتباط متوسط وعكسي .

----- أسئلة وتمارين عامة

3) حساب معامل التحديد وتفسيره بضوء البيانات وتبيان الأهمية الإحصائية لمعامل الارتباط

مبيناً معنى النتيجة التي توصلت إليها :

القلق هو المتغير المستقل x .

التحصيل هو المتغير التابع y .

نحسب متطلبات تطبيق الدستور :

الجدول (9)

x	y	$x - \bar{x}$	$y - \bar{y}$	$x - \bar{x} \times y - \bar{y}$	$(x - \bar{x})^2$	$(y - \bar{y})^2$
2	3	2.8-	4.2-	6.72	7.84	5.75
4	4	0.8-	1.4-	1.12	0.64	1.96
5	7	0.2	1.6	0.32	0.04	2.56
7	5	2.2	0.4-	0.88-	4.84	0.16
6	8	1.2	2.6	3.12	1.44	6.76
24	27	0	0	7.28	14.8	17.2

$$\bar{x} = \frac{\sum x}{n} = \frac{24}{5} = 4.8$$

$$\bar{y} = \frac{\sum y}{n} = \frac{27}{5} = 5.4$$

$$r = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sqrt{\sum (x - \bar{x})^2 \sum (y - \bar{y})^2}}$$

أسئلة وتمارين عامة

$$= \frac{7.28}{\sqrt{14.8(17.2)}} = \frac{7.28}{15.95} = 0.45$$

التفسير : الارتباط بين القلق والتحصيل ضعيف وطردى .

الأهمية الإحصائية لمعامل الارتباط :

نحسب الخطأ المعياري لمعامل الارتباط :

$$Sr = \sqrt{\frac{1-r^2}{n-2}} = \sqrt{\frac{1-(0.45)^2}{5-2}} = 0.51$$

يجب تحقق العلاقة : $Sr > 3r$ لكي يكون لمعامل الارتباط أهمية إحصائية :

$$0.45 \nless 1.53$$

لم تتحقق العلاقة، ليس لمعامل الارتباط أهمية إحصائية .

هذا يعني :

1- معامل ارتباط المجتمع يساوي الصفر .

2- العينة لا تمثل مجتمعها .

3- العينة مسحوبة من مجتمع آخر .

----- أسئلة وتمارين عامة

المسألة (15) :

يمثل الجدول التالي عدد المدارس والشعب في إحدى المناطق:

عدد المدارس	3	5	6	4	2	5
عدد الشعب	7	9	10	10	5	10

احسب معامل ارتباط بيرسون وبين أهميته الإحصائية، مبيناً معنى وجود أهمية أو عدم وجودها.

احسب معامل ارتباط سيرمان وفسره، مبيناً الفرق بين أنواع الارتباط .

احسب الربيع الثالث لعدد الشعب .

الحل :

(1) حساب معامل الارتباط :

الجدول (10)

x	y	$x - \bar{x}$	$y - \bar{y}$	$x - \bar{x} \times y - \bar{y}$	$(x - \bar{x})^2$	$(y - \bar{y})^2$
7	3	1.5-	1.16-	1.74	2.25	1.35
9	5	0.5	0.84	2.1	0.25	0.71
10	6	1.5	1.86	2.79	2.25	3.46
10	4	1.5	0.16	0.24-	2.25	0.02
5	2	3.5-	2.16-	7.56	12.25	4.66
1	25	0	0	15.21	21.5	10.28

أُسئلة وتمارين عامة

$$\bar{X} = \frac{\sum x}{n} = \frac{51}{6} = 8.5$$

$$\bar{y} = \frac{\sum y}{n} = \frac{25}{6} = 4.16$$

$$r = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sqrt{\sum (x - \bar{x})^2 \sum (y - \bar{y})^2}}$$

$$= \frac{15.21}{\sqrt{21.5(10.9)}} = \frac{15.21}{15.31} = 0.99$$

التفسير :

الارتباط بين عدد الشعب وعدد المدارس شديد جداً وطردى .

(2) معامل ارتباط سبيرمان

$$r_s = 1 - \frac{6 \sum D^2}{n(n^2 - 1)}$$

نحسب رتب x ورتب y ثم نحسب D و D^2 :

x	y	رتبة x	رتبة y	D	D^2
7	3	5	5	0	0
9	5	4	2.5	1.5	2.25
10	6	2	1	1	1
10	4	2	4	2-	4
5	2	6	6	0	0
10	5	2	2.5	0.5-	0.25
					$\sum D^2 = 6.25$

نطبق الدستور :

$$r_s = 1 - \frac{6(6.25)}{6(36-1)} = 1 - 0.18 = 0.82$$

أسئلة وتمارين عامة

التفسير :

الارتباط بين رتب المدارس والشعب شديد وطردي

- الفرق بين أنواع الارتباط :

أ- ارتباط بسيط : يدرس العلاقة بين متغيرين فقط .

ب- ارتباط جزئي : يدرس العلاقة بين متغير تابع مستقل مع افتراض ثبات المتغيرات

المستقلة الأخرى وهكذا .

ت- ارتباط متعدد : يدرس العلاقة بين متغير واحد تابع وعدد من المتغيرات المستقلة

3) احسب الربع الثالث لعدد الشعب :

أ- نرتب البيانات تصاعدياً :

$$\frac{3n}{4} + \frac{1}{2} = \frac{3(6)}{4} + \frac{1}{2} = 4.5 + 0.5 = 5$$

ب- نحدد ترتيب الربع الثالث

وتكون القيمة 5 هي قيمة الربع الثالث

----- أسئلة وتمارين عامة

المسألة (16): احسب معامل الارتباط بين المتغيرين Y ، X حيث تكون قيمة كل من Y ، X كما في الجدول التالي :

4	7	7	4	6	1	8	X
2	4	5	4	3	4	6	Y

الحل : نرتب خطوات الحل كما في الجدول التالي :

الجدول (10)

x	y	x^2	y^2	$x.y$
8	6	64	48	36
1	4	1	4	16
6	3	36	18	9
4	4	16	16	16
7	5	49	35	25
7	4	49	28	16
2	2	4	4	4
35	28	219	122	153

لاحظ أنك تحتاج في حساب r إلى المقادير التالية :

$$\sum xy, \sum y^2, \sum x^2, y, x, n$$

----- أسئلة وتمارين عامة

وقد حسبت في الجدول السابق حيث :

$$\bar{y} = \frac{28}{7} = 4, \bar{x} = \frac{35}{7} = 5$$

وبقية المقادير محسوبة في معادلة r نجد :

$$r = \frac{153 - 7 \times 5 \times 4}{\sqrt{219 - 7 \times (5)^2} \sqrt{122 - 7 \times 4^2}}$$

$$r = \frac{13}{\sqrt{44} \sqrt{10}} = 0.62$$

المصطلحات المستخدمة

statistics	الإحصاء
Inferential statistics	الإحصاء الاستدلالي
Descriptive statistics	الإحصاء الوصفي
<i>Correlation</i>	الارتباط
Simple corrolation	الارتباط البسيط
Partial corrolation	الارتباط الجزئي
First order part or partial corr	الارتباط الجزئي من الدرجة الأولى
multiple <i>Correlation</i>	الارتباط المتعدد
<i>Parametric</i>	الأساليب البارامترية (أو المعلمية)
Graphical Presentation	أساليب العرض البياني للمتغيرات
<i>Non- Parametric</i>	الأساليب اللابارامترية (أو اللامعلمية)
Methods of data organization and Presentation	أساليب تنظيم البيانات وعرضها
Deciles	العشرية
<i>Regression</i>	الانحدار

المصطلحات

Quartile Deviation	الانحراف الربيعي
Coefficient of Quartile Q% Variation	الانحراف الربيعي النسبي
Standard deviation	الانحراف المعياري
Scientific Research	البحث العلمي
Independent Variable	التابع المتغير المستقل (المنبأ)
Predicted Variance	التباين المفسر (المتنبأ به)
Un Predicted Variance	التباين غير المفسر (غير المتنبأ به)
Frequency	التكرار
Standard Normal Distribution	التوزيع الطبيعي المعياري
Constant	الثابت
Regression Constant	ثابت الانحدار
Line-Regression	خط الانحدار
Standard Error of estimation	الخطأ المعياري للتقدير
experimental studies	الدراسات التجريبية
Correlative studies	الدراسات الارتباطية

المصطلحات

Survey studies	الدراسات المسحية
Z Scores	الدرجات المعيارية
scattar plot	شكل الانتشار
Bar Chart	طريقة الأعمدة
The graphs lines	طريقة العرض بالخطوط البيانية
The method of least squares	طريقة المربعات الصغرى
Probabilistic samples	العينات العشوائية أو الاحتمالية
Non probability sample	العينات غير الاحتمالية اللا عشوائية
sample	العينة
Quota Sample	عينة الحصص
Accidental sample	العينة العرضية
Simple Random sample	العينة العشوائية البسيطة
Satisfied random sample	العينة العشوائية الطبقية
Cluster Random Sample	العينة العشوائية العنقودية
Multi-stage sample	العينة العشوائية المتعددة المراحل
Systematic Random Sample	العينة العشوائية المنتظمة

المصطلحات

Purposeful	العينة المقصودة
Percentiles	المئينات
Artificial Variable	المتغير الاسمي الصناعي
Dependent Variable	المتغير التابع
Variables	المتغيرات
Dependent Variable	المتغيرات التابعة
Independent Variable	المتغيرات المستقلة
Categorized Variable	متغيرات تصنيفية
Continuous Variables	متغيرات كمية متصلة
Discrete Variables	متغيرات كمية منفصلة
Mean	المتوسط الحسابي
Weighted Mean	المتوسط الموزون
Statistical Population	المجتمع الإحصائي
Frequency Histogram	المدرج التكراري
Range	المدى
Nominal Scales	المستوى الاسمي أو التصنيفي

المصطلحات

ordinal scales	المستوى الترتيبي
Interval Scales	المستوى الفئوي
Ratio Scales	مستوى النسبة
Prisoner sources	المصادر الأولية لجمع البيانات
Historical sources	المصادر التاريخية
secondary sources	المصادر الثانوية لجمع المعلومات
Field sources	المصادر الميدانية
Frequency Polygon	المضلع التكراري
Regression Slope	معامل الانحدار (ميل الانحدار)
<i>Determination</i>	معامل التحديد
Coefficient of determination	معامل التحديد
Pearson	معامل بيرسون
Spearman	معامل سبيرمان
Phi	معامل فاي
Kandell,s tau	معامل كاندل تاو
Cramers	معامل كريمر

المصطلحات

Coefficient of skewness	مقاييس الالتواء
Measures of Dispersion	مقاييس التشتت
Measures of central tendency	مقاييس النزعة المركزية
Frequency Curve	المنحنى التكراري
Mode	المنوال
Percentage	النسبة
central limit Theorem	نظرية النزعة المركزية
The Regression Models	نماذج الانحدار
linear Regression	نماذج الانحدار البسيط
Linear Regression	نماذج الانحدار الخطي
Multiple Regression	نماذج الانحدار المتعدد
Nonlinear Regression	نماذج الانحدار المنحني (غير الخطي)
Multivariate Regression	نماذج الانحدار متعدد المتباين
Median	الوسيط

المراجع

- أبو النيل، محمود (1987). الإحصاء النفسي والاجتماعي والتربوي، دار النهضة العربية، بيروت .
- أبو حطب، فؤاد، صادق، آمال (1991). مناهج البحث وطرق التحليل الإحصائي في العلوم النفسية والتربوية والاجتماعية، مكتبة الأنجلو المصر، القاهرة، مصر .
- أبو زينة، فريد، وآخرون (2006). مناهج البحث العلمي الإحصاء في البحث العلمي . دار المسيرة. عمان.
- أبو صالح، محمد صبحي (2000). الطرق الإحصائية، دار البارودي العلمية للنشر والتوزيع، عمان، الأردن.
- أبو صالح، محمد صبحي، عوض، عدنان (2008). مقدمة في الإحصاء مبادئ وتحليل باستخدام SPSS، دار المسيرة للنشر والتوزيع، عمان، الأردن.
- أبو علام، رجاء . (2004). مناهج البحث في العلوم النفسية والتربوية، الطبعة الرابعة، دار النشر للجامعات - القاهرة - مصر .
- أبو علام، رجاء (2004). مناهج البحث في العلوم النفسية والتربوية. الطبعة الرابعة. دار النشر للجامعات القاهرة.
- بدر، سالم، عباينة، عماد (2007). مبادئ الإحصاء الوصفي والاستدلالي، دار المسيرة للنشر والتوزيع والطباعة، عمان، الأردن.
- بركات، نافذ (2007). التحليل الإحصائي باستخدام برنامج SPSS. الجامعة الإسلامية.

المراجع

- بري، عدنان بن ماجد الرحمن، وآخرون (1998): أساسيات طرق التحليل الإحصائي، النشر العلمي والمطابع، السعودية.
- البطش، محمد وليد، أبو زينة، فريد . (2006). مناهج البحث العلمي - تصميم البحث والتحليل الإحصائي، دار المسيرة للنشر والطباعة، عمان، الأردن .
- البلداوي، عبد الحميد . (2004). الأساليب الإحصائية التطبيقية، دار الشروق للنشر والتوزيع، عمان - الأردن.
- الجادري، عدنان . (2007). الإحصاء الوصفي في العلوم التربوية، دار المسيرة للنشر والتوزيع والطباعة، عمان، الأردن.
- الجضعي، خالد (2005). تقنيات صنع القرار تطبيقات حاسوبية . الجزء الأول. دار الأصحاب. الرياض.
- حبيب، مجدي عبد الكريم (1996). التقويم والقياس في التربية وعلم النفس. الطبعة الأولى. المجلد الأول. مكتبة النهضة المصرية. القاهرة.
- حبيب، مجدي عبد الكريم (2000). التقويم والقياس في التربية وعلم النفس. الطبعة الأولى. المجلد الثاني. مكتبة النهضة المصرية. القاهرة.
- خيرى، السيد محمد . (1997). الإحصاء النفسي، دار الفكر العربي، القاهرة .
- السيد، فؤاد البهي . (1978). علم النفس الإحصائي وقياس العقل البشري، جامعة عين شمس، القاهرة .
- الشربيني، زكريا. (1995). الإحصاء وتصميم التجارب في البحوث النفسية والتربوية والاجتماعية. مكتبة الأنجلو المصرية. القاهرة.

المراجع

- الصياد، جلال، ربيع، عبد الحميد، سمرة، عادل (2008). مقدمة في الإحصاء. دار حافظ للنشر والتوزيع. الرياض.
- عبد المنعم، ثروت (2005). الإنحدار. مكتبة الأنجلو المصرية. القاهرة.
- علام، صلاح الدين (2005). الأساليب الإحصائية الاستدلالية في تحليل بيانات التربية والاجتماعية، الطبعة الأولى، دار الفكر العربي، القاهرة.
- علوان، حسين . (1993). طرق المعاينة، دار الفرقان - عمان - الأردن.
- العلي، ابراهيم.(1998).مدخل في نظرية الارتباط.مديرية الكتب والمطبوعات الجامعية.حلب.
- عودة، أحمد، الخليلي، خليل. (1988). الإحصاء للباحث في التربية والعلوم الإنسانية، دار الفكر للنشر والتوزيع، عمان، الأردن.
- عودة، أحمد الخليلي، خليل (2000). الإحصاء للباحث في التربية والعلوم الإنسانية، الطبعة الثانية دار الفكرة للنشر والتوزيع، عمان.
- عودة، أحمد وملكوي فتحي حسن (1992). أساسيات البحث العلمي، الطبعة الثانية مركز ببيضون للكمبيوتر، أريد.
- فهمي، محمد شامل (2005). الإحصاء بلا معاناة المفاهيم مع التطبيقات باستخدام برنامج SPSS، الجزء الأول، مكتبة فهد الوطنية أثناء النشر، الرياض.
- كروكر، ليندا، الجينا، جميس(2009). نظرية القياس التقليدية والمعاصرة. دار الفكر.عمان

-----المراجع

- كنجو، أنيس ترجمة (2000) : نماذج إحصائية خطية تطبيقية، انداد تحليل تباين و تصاميم تجريبية، الجزء الأول، النشر العلمي و المطابع السعودية .
- الكيلاني، عبد الله زيد، الشريفين، نضال(2007).مدخل إلى البحث في العلوم التربوية والاجتماعية.دار المسيرة،عمان .
- المنيزل، عبد الله، غرايبة، عايش (2006). الإحصاء التربوي، دار المسيرة، عمان.
- المنيزل، عبد الله فلاح . (2000). الإحصاء الاستدلالي وتطبيقاته في الحاسوب باستخدام الرزم الإحصائية (SPSS)، دار وائل للطباعة.عمان.
- النبهان،موسى(2004).أساسيات القياس في العلوم السلوكية.دار الشروق للنشر والتوزيع.عمان.
- هندي، محمود، عبد الله، أنور (1994). مبادئ الإحصاء والاحتمالات، الطبعة الثانية، مطابع الملك سعود، الرياض.

المراجع الأجنبية

- James, S. & Terry, M. (2000). *Statistics*, Prentice Hall. U.S.
- Lomax, R, L. (2001). *an introduction to statistical concept for education and behavioral sciences*. Lawrence erblum associates Publishers Mahwah, U.S.A.
- Raymond, J (1999). *Statistical Analysis in The behavioral sciences*, Mc Craw- Hill. U.S.A.
- MendeHall. S. (1995). *statistics for Engineering and the sciences forth edition*, prentice Hill international, INC, U.S.A.
- Zar, J (1996). *Biostatistical analysis*. Third Edition, prentice Hall. International, INC, U.S.A.
- Levin, R. & Rubin, D. (1998). *statistics for management*, Prentice-Hall international, INC. U.S.A.
- Bearson, M. M. & Levine, D. (2002). *Basic Business Statistics*. prentice Hall. International, INC, U.S.A.
- Leech, N. Barrett, K. Morgan, G. (2005). *Spss for intermediate Statistics use and interpretation*. second Edition. Lawrence Erlbaum associates. Publishers . London
- Evertt, B. Er (2001). *statistics for psychologists An Intermediate Course*.
- Stonghton, H. (). *rescher Methods and Statistics in Psychology*. second Edition.
- Rencher. A. (2002). *Methods of Multivariate Analysis* second Edition. Wiley interscience .Canada .
- Lawrence Erlbaum Associates, Publishers mahwah, New Jersey,

المراجع -----

- Norman.G&Streiner.D.(2002).Statistics,ThirdEdition.BD Decker.Inc.Hamilton.London.
- Rumsey,D.(2007).Intermediate Statistics For Dummies .Wiley Publishing.Inc.Indiana
- Elliott,R.(2002).Research Methods in Clinical Psychology 2nd Edittion. Wiley Sonns.Ltd.England.



ملحق (1)

تدريبات على استخدام برنامج SPSS

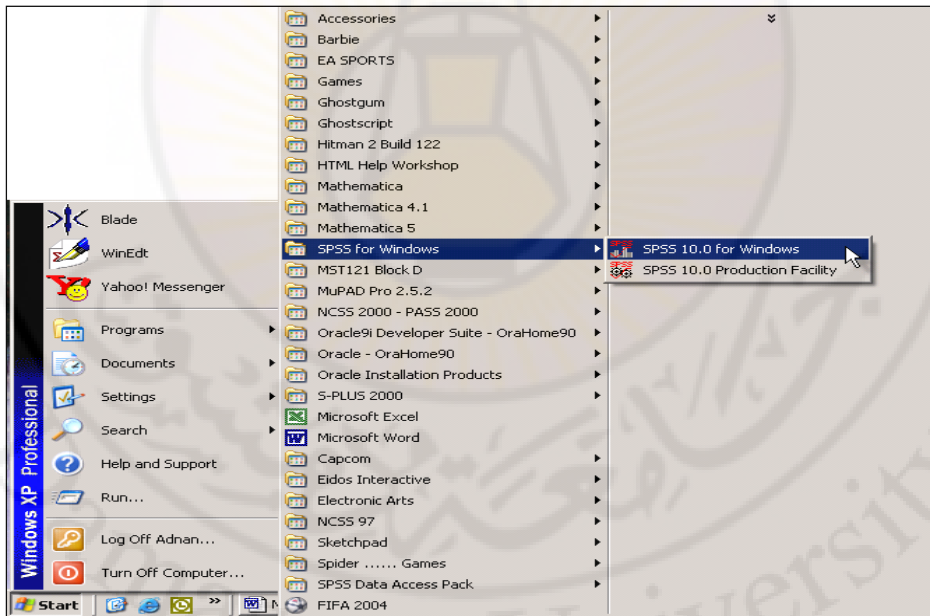
SPSS مختصر (Statistical Package For Social Sciences)

دليل استخدام SPSS :

1. أولاً: إنشاء ملف بيانات

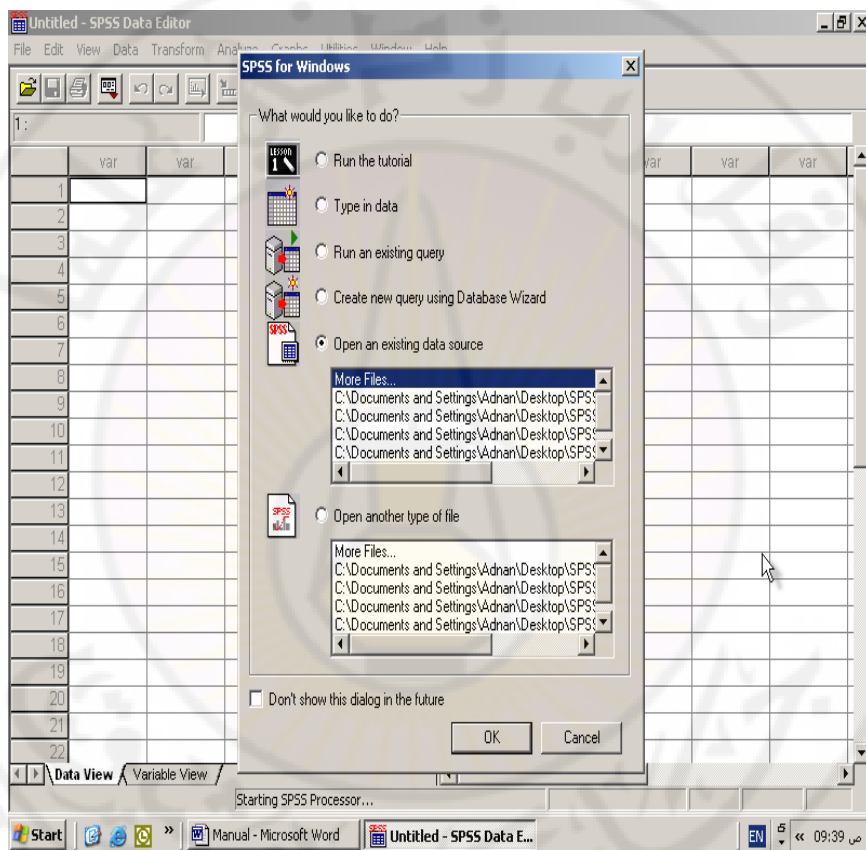
فيما يلي أبرز الخطوات التي تحتاجها لتشغيل برمجية SPSS ومن ثم فتح ملف بيانات أو إنشاء ملف جديد.

1- لتشغيل SPSS اشر على (Start) ثم (Program) ثم (SPSS) كما هو موضح في الشكل التالي:

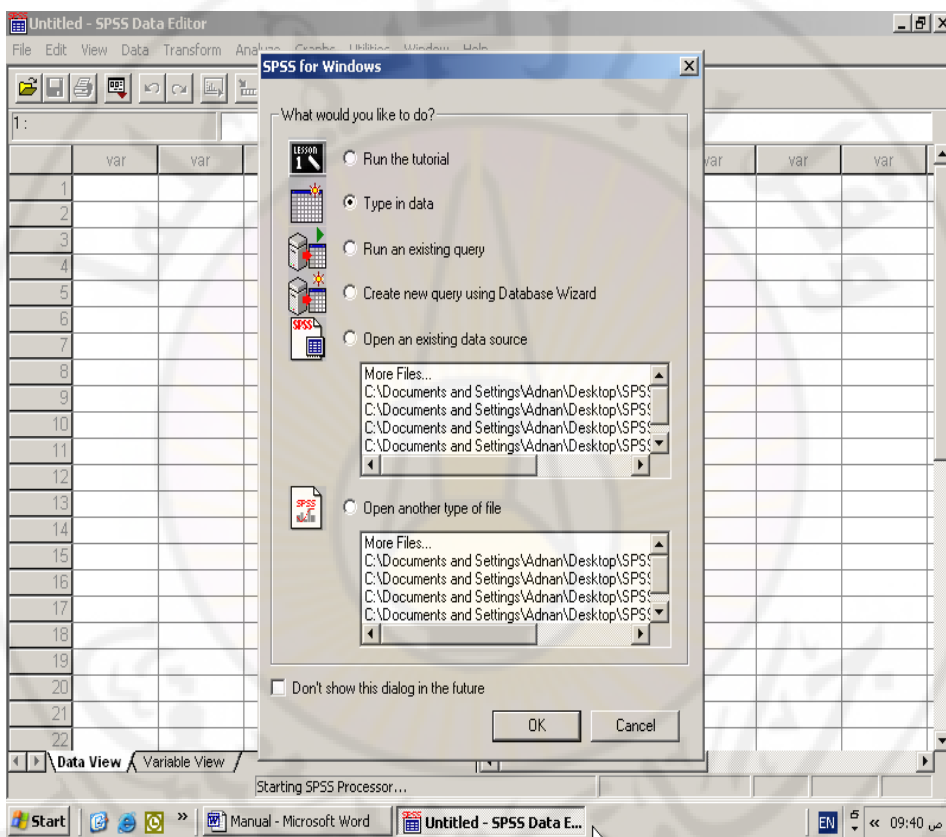


الملاحق

وبالنقر على (SPSS) تحصل على صندوق الحوار التالي. فإذا أردت فتح ملف موجود أشر على (Open an existing data source) كما هو موضح في الشكل التالي.



أما إذا أردت أن تدخل بيانات جديدة فأشر على (Type in data) كما هو موضح في الشكل التالي.



وبالنقر على (ok) تحصل على جدول فارغ كما هو ظاهر في الشكل التالي.

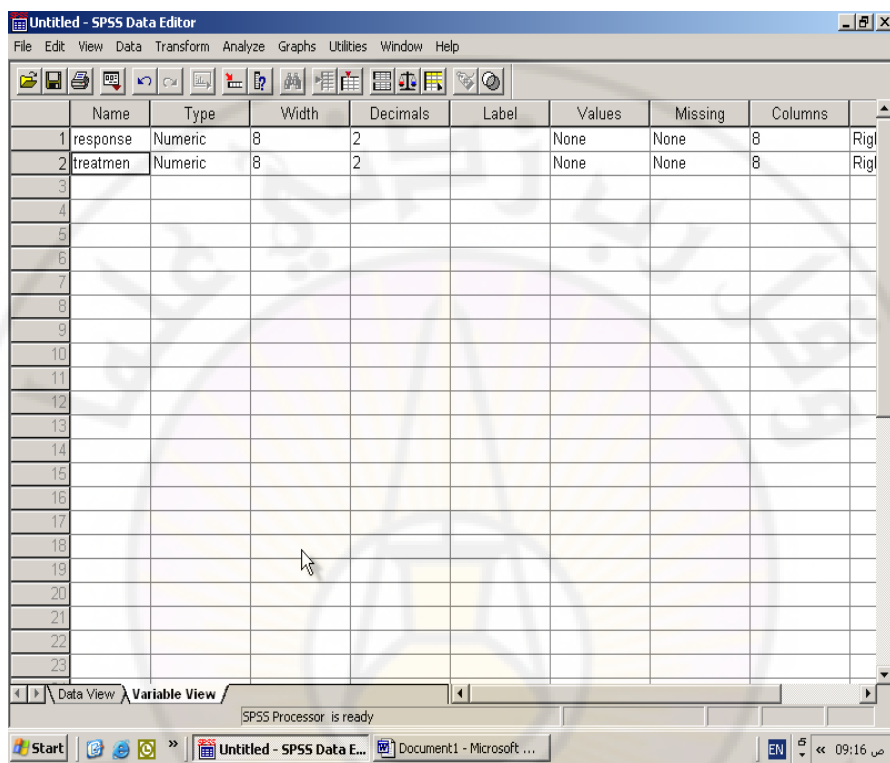
الملاحق

أدخل البيانات في أعمدة ولاحظ ضرورة أن تكون الأعمدة متساوية الطول وحيث كل عمود يعطي قيم متغير وكل سطر يعطي مشاهدات حالة.

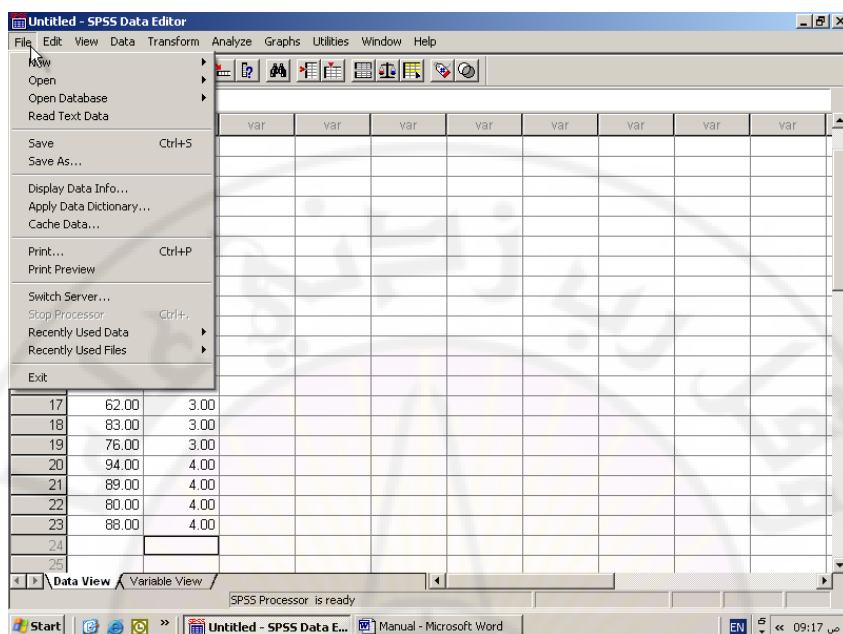
	response	treatmen	var	var	var	var	var	var	var	var
4	79.00	1.00								
5	81.00	1.00								
6	69.00	1.00								
7	75.00	2.00								
8	69.00	2.00								
9	83.00	2.00								
10	81.00	2.00								
11	72.00	2.00								
12	79.00	2.00								
13	90.00	2.00								
14	59.00	3.00								
15	78.00	3.00								
16	67.00	3.00								
17	62.00	3.00								
18	83.00	3.00								
19	76.00	3.00								
20	94.00	4.00								
21	89.00	4.00								
22	80.00	4.00								
23	88.00	4.00								
24										
25										

ولإعطاء أسماء للمتغيرات انقر (Variable view) افتتح صفحة كما يلي ثم ادخل في عمود (Name) أسماء المتغيرات كما هو ظاهر في الشكل التالي.

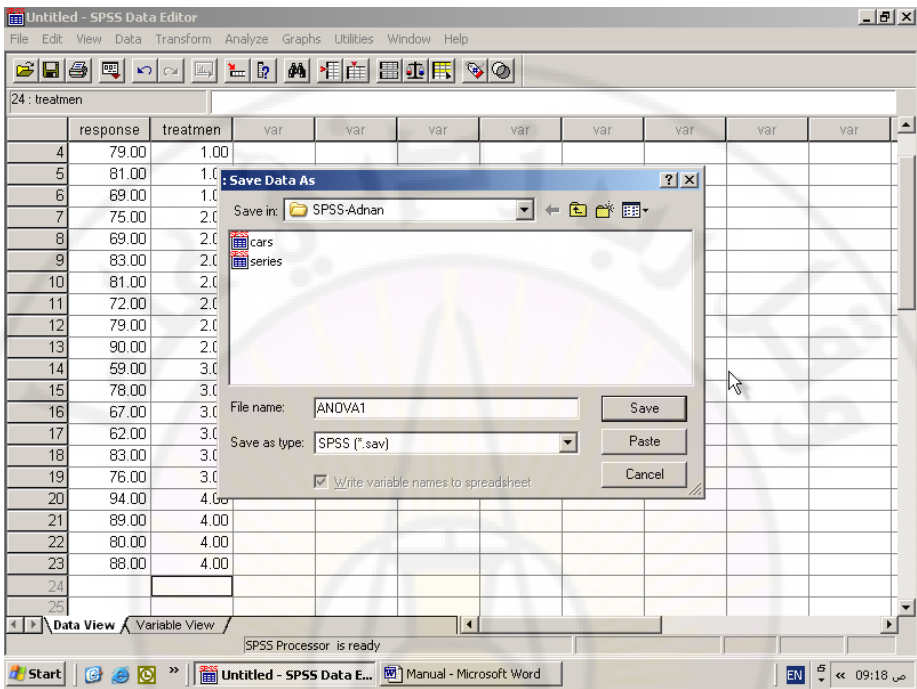
الملاحق -----

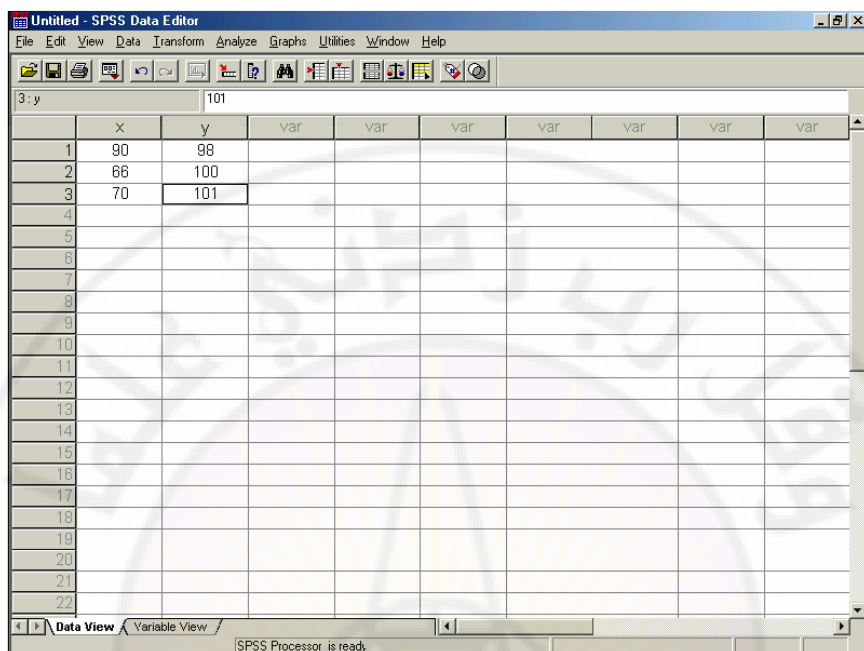


ولتخزين البيانات اختر (File \ Save As).



وبعدها تحصل على صندوق الحوار التالي لتحديد مكان التخزين واسم الملف المراد تخزينه كما هو موضح في الشكل التالي وأخيراً أنقر (Save) ليتم التخزين.





نلاحظ أن الاسم الافتراضي للملف هو (Untitled) إلى أن يتم حفظ هذا الملف باسم آخر.

يوجد أسفل شريط Spss Data Editor شريط القوائم ويتضمن :

File – edit – view – Data – Transform – analyze ...etc .

و أسفل منه يوجد شريط الأدوات (الاختصارات) للقيام بالمهام على نحو أسرع .

----- الملاحق

ورقة العمل في برنامج spss تشبه ورقة العمل في برنامج Excel ، حيث تتكون من عدد من الأعمدة و الصفوف (الحالات) ، وكل عمود من هذه الأعمدة يمثل متغير و كل صف من الصفوف يمثل حالة .

تنقسم صفحة DATA EDITOR إلى قسمين :

القسم الأول : صفحة Data view و يتم فيها إدخال البيانات وعرضها .

القسم الثاني : صفحة variable view و يتم فيها كتابة اسم المتغير (Name) و نوعه (Type) و عرض المتغير (columns) ... الخ .

كيفية إدخال البيانات :

أولاً : نذهب إلى صفحة **variable view** و نقوم بإدخال أسماء المتغيرات في

خانة Name

شروط كتابة اسم المتغير :

- ❖ لا يزيد طول اسم المتغير عن ثمانية رموز .
- ❖ يجب أن يبدأ اسم المتغير بحرف أما بقية الرموز فقد تكون أحرفاً أو أرقاماً .
- ❖ لا يجب أن ينتهي اسم المتغير بـ (Dot) (.)
- ❖ لا يتضمن اسم المتغير فراغات و بعض الرموز الخاصة مثل ! ، ? ، * ، %
^ : -
- ❖ لا يجوز تسمية المتغير مرتين في نفس ملف البيانات ، بمعنى أنه لا يجوز تكرار اسم المتغير في ملف البيانات نفسها.
- ❖ هناك بعض الكلمات المحجوزة للبرنامج و الممنوع استخدامها مثل:
LT, LE ,GT,GE ,EQ , NE , AND , OR , NOT , BY , WITH

ثانياً : نقوم باختيار نوع المتغيرات و هي كالتالي :

- ❖ **Numeric**(رقمي) و هي خاصة بالمتغيرات الرقمية و الكمية مثل الدخل و الاستهلاك و الدرجات..الخ .

❖ String (وصفي) وهي خاصة بالبيانات الوصفية مثل الاسم و الجنس ..الخ .

❖ Date (التاريخ) يتم استخدامها في حالة ذكر التاريخ ك اسم للمتغير .

❖ Dollar (العملة) يتم استخدامها في حالة العملة ك اسم للمتغير حيث يتم كتابة القيمة ويظهر إلى جانبها علامة \$.

ثالثاً : Width : وفيها يتم تحديد الأرقام الصحيحة.

رابعاً : Decimals : وفيها يتم تحديد الأرقام العشرية .

خامساً : Label : و فيه يتم كتابة عنوان للمتغير ، فعلى سبيل المثال لو كان اسم المتغير الدخل الشهري فإننا لا نستطيع كتابة ذلك في خانة Name و ذلك لوجود فراغات في اسم المتغير كما أن اسم المتغير الدخل الشهري يزيد عن ثمانية أحرف و لحل هذه المشكلة نقوم بكتابة الدخل مثل اسم للمتغير و نكتب في خانة Label و التي تتسع لـ 120 حرف الدخل الشهري حيث يظهر مصطلح الدخل الشهري في المخرجات و ليس في المدخلات .

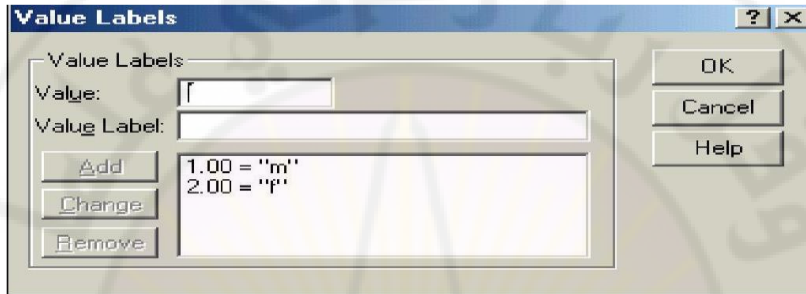
سادساً : Values : حيث يتم إعطاء قيم للمتغير و لتوضيح ذلك دعنا نأخذ المثال التالي:

الجنس ينقسم إلى نوعين ذكور و إناث ، ويمكن لنا أن نعوض عن الذكور برقم 1 و عن الإناث برقم 2

الحل العملي :

نذهب إلى صفحة variable view ونضغط على Values فيظهر لنا مربع

الحوار التالي :



نقوم بكتابة الرقم 1 في خانة values ثم ننقل إلى خانة values label ونكتب "m" كمختصر لـ "ذكر" ثم نضغط Add ونكرر العملية مرة أخرى بكتابة الرقم 2 في خانة Values وكتابة "F" بوصفها مختصراً لـ "أنثى" ثم نضغط Add ثم نضغط ok و بالتالي يصبح الرقم 1 يشير إلى الذكور و الرقم 2 يشير إلى الإناث.

سابعاً : القيم المفقودة (missing values) : ففي حالة الضغط على missing يظهر مربع الحوار التالي و يكون لدينا ثلاثة خيارات :

Missing Values

☐ No missing values

☐ Discrete missing values

☒ Range plus one optional discrete missing value

Low: 7.5 High: 10.3

Discrete value: 15

OK Cancel Help

❖ **No missing values** : و هذا يعني عدم وجود قيم مفقودة و هي الحالة الافتراضية .

❖ **Discrete missing values** : و يستخدم في حالة وجود ثلاث قيم منفصلة كحد أقصى .

❖ **Range of missing values** : و يستخدم فقط في حالة المتغيرات الرقمية ، و فيها يمكن كتابة مدى القيم المفقودة و ذلك بتحديد الحد الأدنى و الحد الأعلى في الخانتين low , High على الترتيب و كذلك قيمة مفقودة إضافية اختيارية.

ملاحظة : لا يمكن تعريف قيم مفقودة للمتغيرات الوصفية .

ثامناً : عرض العمود (columns) : و يستخدم لتحديد عرض العمود لمتغير معين و ذلك بتحديد المتغير المراد تحديد عرض عموده ، كما و يمكن زيادة أو تقليل عرض العمود بالأسهم إلى الأعلى أو إلى الأسفل (أو كتابة عرض العمود مباشرة) .

تاسعاً : محاذاة النص **Alignment** : و ذلك لضبط النص داخل الخلايا حيث يتم اختيار واحد من الثلاث خيارات :

❖ Left : محاذاة النص إلى اليسار .

❖ Right : محاذاة النص إلى اليمين .

❖ Center : محاذاة النص إلى الوسط .

عاشراً : **Measurement (القياس)** : و هي على ثلاث خيارات

❖ Scale : يستعمل للبيانات العددية مثل الطول والوزن و الدخل .

❖ Ordinal : يستعمل للمتغيرات الترتيبية حيث أن هذا المتغير ذو عدد

محدد من الفئات يمكن ترتيبها تصاعدياً أو تنازلياً مثل تقدير طالب في

امتحان (ممتاز ، جيد جداً ، جيد ، مقبول ، راسب) .

❖ Nominal : يستعمل للمتغيرات الاسمية التي لا تحمل معنى الترتيب ،

مثل الجنس (ذكر ، أنثى) فقد يأخذ الذكر رقم 1 و الأنثى تأخذ رقم 2 و

العكس و كذلك المدن مثل (دمشق ، حمص ، حماة ، حلب) .

قوائم البرنامج :

File: تتضمن قائمة file العديد من الأوامر منها :

- ❖ جديد (New) : بمعنى فتح ملف جديد ، وهي على خيارين أما ملف بيانات data أو ملف مخرجات output .
- ❖ فتح (open) : بمعنى فتح ملف مخزن مسبقاً و هو علي خياران أما :
File—Open—Data فتح ملف بيانات مخزن مسبقاً
File—Open—Output فتح ملف مخرجات مخزن مسبقاً
- ❖ حفظ (Save As) : بمعنى حفظ الملف باسم لأول مرة .
- ❖ حفظ التغيرات (Save) : بمعنى حفظ التغيرات التي أجريت على الملف المخزن مسبقاً.

Edit:

- ❖ Undo : تراجع عن الكتابة .
- ❖ Redo : إعادة الكتابة .
- ❖ Copy : نسخ .
- ❖ Paste : لصق .
- ❖ Find : بحث .
- ❖ Cut : قطع .
- ❖ Clear : مسح .

View: تتضمن قائمة view :

❖ Fonts : حيث يتم من خلاله التحكم في حجم و نوع و نمط الخط و من خلاله يتم أيضاً تكبير الخلايا .

❖ Value Label : يحدد يتم إظهار عناوين القيم .

❖ Variables: يتم من خلاله إظهار المتغيرات للتعرف على كل منها ففي حالة فتح ملف جديد و نريد التعرف على المتغيرات و طبيعتها فإنه يتم اختيار هذا الأمر .

Data : ومن الأوامر التي تتضمنها هذه القائمة ما يلي :

❖ Insert variable : إضافة متغير إلى قائمة المتغيرات الموجودة في الملف .

❖ Insert case : إضافة صف (حالة) إلى الصفوف الموجودة في الملف .

❖ Go to case : الذهاب إلى حالة أو صف ما .

❖ Sort case : ترتيب القيم سواء أكانت رقمية أو اسمية ترتيباً تصاعدياً أو تنازلياً

❖ Merge file : دمج الملفات مع بعضها البعض ، وعملية الدمج على نوعين :

1- Add variables : يتيح هذا الأمر إمكانية دمج الملف العامل (الحالي) مع الملف الخارجي و اللذان يحتويان نفس الحالات و لكن المتغيرات مختلفة .

2- Add cases : يتيح هذا الأمر إمكانية دمج الملف العامل (الحالي) مع الملف الخارجي و اللذان يحتويان نفس المتغيرات و لكن الحالات مختلفة .

مثال على add cases :

----- الملاحق

قم بإنشاء ملف باسم الملف الأول والذي يتكون من الآتي :

الاسم	تربية	معلم صف	علم نفس	إرشاد نفسي
أحمد	85	75	90	66
رامز	88	60	85	70

و كذلك إنشاء ملف باسم الملف الثاني و الذي يتكون من الآتي :

الاسم	تربية	معلم صف	علم نفس	إرشاد نفسي
محمد	83	66	87	80
فيصل	70	77	60	73

المطلوب : دمج الملفين المذكورين في ملف واحد ليصبح بالشكل التالي :

الاسم	تربية	معلم صف	علم نفس	إرشاد نفسي
أحمد	85	75	90	66
رامز	88	60	85	70
محمد	83	66	87	80
فيصل	70	77	60	73

الحل العملي :

نلاحظ أن الملفين السابقين يشمان المتغيرات نفسها و هي (الاسم- تربية -

إحصاء- علم نفس - إرشاد نفسي)

و لدمج الملفين نقوم بفتح الملف الأول ثم نذهب إلى قائمة data ثم merge

file ثم نختار Add cases ثم نقوم بفتح الملف الثاني ثم نضغط ok وبالتالي

يصبح الملفين مدمجين .

----- الملاحق

ملاحظة / السبب في اختيار Add cases أن المتغيرات في كلا الملفين متشابهة.

مثال على add variables :

قم بإنشاء ملف باسم كلية التربية والذي يتكون من الآتي :

الاسم	تربية	معلم صف	علم نفس	إرشاد نفسي
أحمد	85	75	90	66
رامز	88	60	85	70

و كذلك إنشاء ملف باسم كلية العلوم و الذي يتكون من الآتي :

الاسم	فيزياء	كيمياء	أحياء
حسام	83	66	87
سامر	70	77	60

المطلوب : دمج الملفين المذكورين في ملف واحد ليصبح بالشكل التالي :

الاسم	تربية	معلم صف	علم نفس	إرشاد نفسي	فيزياء	كيمياء	أحياء
أحمد	85	75	90	66	83	66	87
رامز	88	60	85	70	70	77	60

الحل العملي :

نلاحظ أن ملف كلية التربية يشتمل على متغيرات تختلف عن متغيرات ملف كلية العلوم و لكن هناك تشابه في الحالات (أحمد ، رامز) .

----- الملاحق

قم بفتح ملف كلية التجارة ، ثم أذهب إلى قائمة data ثم نختار merge file ثم نضغط على add variables ، ثم نقوم بفتح ملف كلية العلوم ثم نضغط ok ، نلاحظ أنه تم استبعاد الأسماء في ملف كلية العلوم و ذلك لوجود متغير الاسم في ملف كلية التجارة و من المعروف لدينا أنه من أحد شروط كتابة اسم المتغير عدم تكرار اسم المتغير في نفس ملف البيانات و لحل هذه المشكلة نقوم بترتيب متغير الاسم في كلا الملفين ترتيبا تصاعديا أو تنازليا ثم بعد ذلك نقوم بإدخال متغير الاسم في كلية العلوم و التي تم استبعاده إلى خانة key variables مع تنشيط البند match cases on key variables in sorted files ثم نضغط ok و من ثم نلاحظ أنه تم تضمين الأسماء في كلية العلوم مع الأسماء في كلية التجارة.

❖ Aggregate : يستخدم هذا الأمر لتجميع البيانات

الملف salary يحتوي أسماء مجموعة من الموظفين و درجاتهم الوظيفية degree و الأجر الشهري salary و كما يظهر في شاشة data editor :

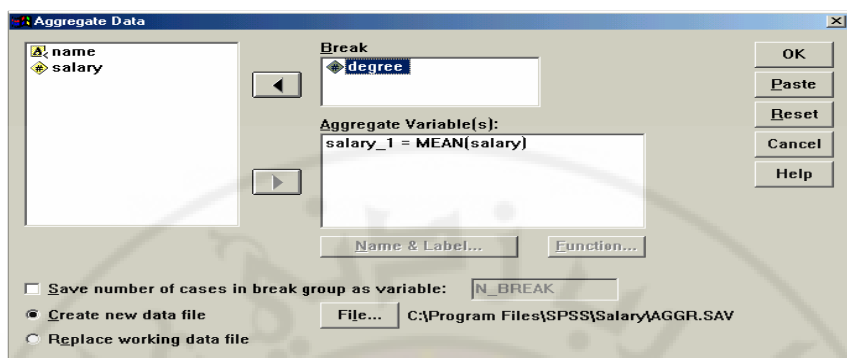
الملف salary

الاسم	الدرجة الوظيفية	الراتب الشهري
أحمد	3	40
سمير	3	35
لؤي	3	50
محمود	1	80
إياد	1	70
ياسين	2	66
سامح	1	85
رزق	1	77
كمال	2	59
عباس	3	45
مهدي	1	90
سالم	2	62
صباح	2	57
فلاح	2	55
عماد	1	82

المطلوب : تجميع متغير الراتب (استخراج المتوسط الحسابي) حسب الدرجة الوظيفية

الحل العملي : من قائمة data اختر aggregate فيظهر الشكل التالي :

الملاحق -----



أنقل متغير الدرجة الوظيفية degree إلى خانة break و متغير الراتب إلى خانة aggregate variables ثم اضغط ok .

❖ Spilt file (تجزئة الملف) : يستعمل هذا الأمر لتجزئة (فصل) ملف

البيانات.

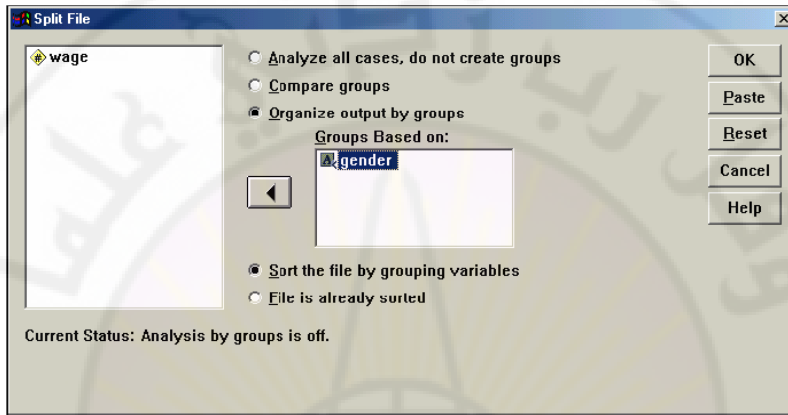
مثال: الجدول التالي يمثل رواتب مجموعة من الأشخاص حسب الجنس

الراتب	الجنس
60	ذكر
30	أنثى
70	ذكر
35	أنثى
65	ذكر
40	أنثى

المطلوب : تجزئة الملف إلى جزئين ، الأول يمثل رواتب الذكور و الثاني يمثل رواتب الإناث.

الحل العملي:

أدخل البيانات إلى شاشة data editor ثم اختر من قائمة data الأمر spilt
file: فيظهر صندوق الحوار التالي :



نقوم بإدخال متغير الراتب إلى خانة groups based on و نقوم باختيار
organize output by groups ثم نضغط ok .

ملاحظة / اختيار البند organize output by groups يعطي نفس نتائج البند
compare groups و لكن الاختلاف يظهر في شكل المخرجات .

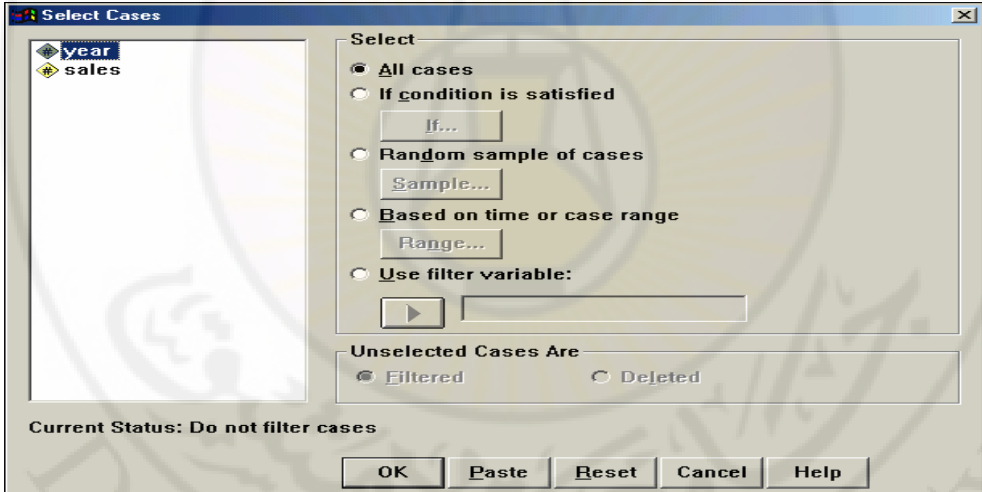
Select cases: يستخدم هذا الأمر لاختيار حالات معينة و فق شروط

معينة .

مثال :من ملف employee data :

1. اختر الحالات التي تزيد رواتبهم عن \$50000 .
2. اختر الموظفين من الذكور و الذين تزيد سنوات تعليمهم عن 17 سنة .

3. اختر الموظفين من المدراء و الذي تزيد خبرتهم (job time) عن 91 شهر .
 4. اختيار الموظف الذي يعمل في وظيفة كاتب أو مدير .
 5. اختيار الموظفين الذين تعلموا بين 18 سنة و 20 سنة .
 6. اختيار الموظفين من غير الكتبة .
 7. اختيار الموظفين من الحراس و سنوات تعليمهم أكثر من 10 سنة و راتبهم أكثر من \$30000 .
 8. اختيار الموظفين من الذكور و سنوات تعليمهم 17 سنة فأكثر و مدراء .
- من قائمة data اختر select cases : فيظهر صندوق الحوار التالي :



ملاحظة / المتغيرات الظاهرة في يسار صندوق الحوار (year , sales) هي متغيرات افتراضية.

في صندوق الحوار الأبيض يظهر أسماء المتغيرات بالشكل التالي:

رقم الموظف	Employee code(id)
الجنس	gender
تاريخ الميلاد	Date of birth(bdate)
سنوات التعليم	Education level(years)
نوع الموظف	Employee category(job cat)
الراتب الحالي	Current salary(salary)
الراتب في بداية العمل	Beginning salary(salbegin)
الخبرة بالشهور	Months since hire(job time)
الخبرة السابقة	Previous experience(perv exp)
الأقلية	Minority classification

معلومات عن ملف employee data

يتكون من تسعة متغيرات ذكرت في الأعلى ،رقم الموظف ، و تاريخ الميلاد ، و سنوات التعليم ، و نوع الموظف والموظفين مقسمين إلى ثلاثة أنواع (مدير يأخذ الرقم 3 ، و كاتب يأخذ الرقم 1 ، و حارس يأخذ الرقم 2) ، و الجنس مقسم إلى قسمين (ذكر ويأخذ الرمز f ، و أنثى تأخذ الرمز m) .

في الجانب الأيمن لصندوق الحوار و أسفل select يوجد خمسة خيارات :

❖ All cases : اختيار كل الخلايا و هي الحالة الافتراضية .

❖ If condition is satisfied : اختيار خلايا معينة وفق شروط معينة .

❖ Random sample of cases : اختيار عينة عشوائية بحجم معين و

لتنفيذ ذلك نشط هذا الاختيار ثم أضغط sample فيظهر مربع حوار

بعنوان select cases :random sample : بحيث يوجد في مربع

الحوار هذا خياران :

Approximately : يستخدم لاختيار نسبة مئوية تقريبية من الحالات ، فمثلا

يمكن اختيار 20% تقريباً من كل الخلايا .

Exactly : يستخدم لاختيار عينة عشوائية ذات حجم معين من أول عدد مناسب

من الخلايا .

❖ Based on time or case range : يستخدم لاختيار عينة ذات مدى

معين .

❖ Use filter variable : يستخدم هذا الاختيار في حالة استخدام متغير

رقمي كمتغير لتصفية الخلايا المطلوبة ، و في هذه الحالة فان الخلايا

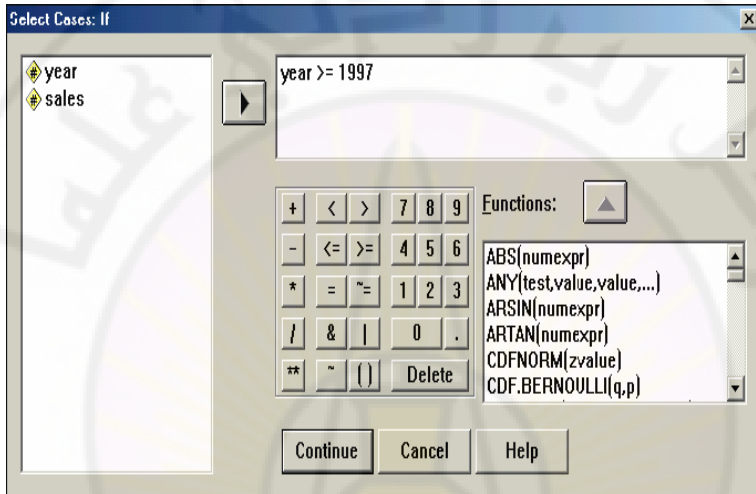
التي قيمها صفراً أو ليست قيم مفقودة لمتغير التصفية سوف يتم اختيارها.

حل المثال :

1- لاختيار الحالات التي تزيد رواتبهم عن \$50000 نقوم بالتالي :

----- الملاحق

نذهب إلى قائمة data ثم نختار select cases فيظهر لنا مربع الحوار السابق ،
نقوم بتنشيط البند If condition is satisfied ثم نضغط If : فيظهر لنا مربع
الحوار التالي :



ملاحظة / المتغيرات الظاهرة في يسار صندوق الحوار (year , sales) هي
متغيرات افتراضية .

نقوم بإدخال المتغير salary إلى المستطيل الأبيض ثم نختار إشارة > من شاشة
المفاتيح الموجودة في صندوق الحوار ثم نكتب الرقم 50000 ثم نضغط
continue ثم نختار البند Deleted لإزالة الخلايا الغير مطلوبة.

salary>50000

2- لاختيار الموظفين من الذكور و الذين تزيد سنوات تعليمهم عن 17 سنة
نذهب إلى قائمة data ثم نختار select cases فيظهر لنا مربع حوار ، نقوم

بتنشيط البند If condition is satisfied ثم نضغط If فيظهر لنا مربع الحوار السابق ، ثم نقوم بإدخال متغير gender ثم إشارة = ثم حرف "m" ثم & ثم نقوم بإدخال متغير التعليم educ ثم إشارة > ثم نكتب الرقم 17 ، ثم نضغط continue ثم نختار البند deleted ثم نضغط ok .

gender="m"&educ>17

3- لاختيار الموظفين من المدراء و الذي تزيد خبرتهم (job time) عن 91 شهر نتبع نفس الخطوات السابقة .

cat=3&jobtime>91

4- اختيار الموظف الذي يعمل في وظيفة كاتب أو مدير نتبع نفس الخطوات السابقة .

Job cat=3& jobcat=1

5- اختيار الموظفين الذين تعلموا بين 18 سنة و 20 سنة نذهب إلى قائمة data ثم نختار select cases فيظهر لنا مربع حوار ثم نقوم بكتابة الصيغة التالية :

educ>=18&educ<=20

6- لاختيار الموظفين من غير الكتبة نتبع نفس الخطوات السابقة بكتابة الصيغة التالية :

Job cat =~1

7- اختيار الموظفين من الحراس و سنوات تعليمهم أكثر من 10 سنة و راتبهم أكثر من \$30000 نكتب الصيغة التالية :

Jobcat=2&educ>10&salary>3000

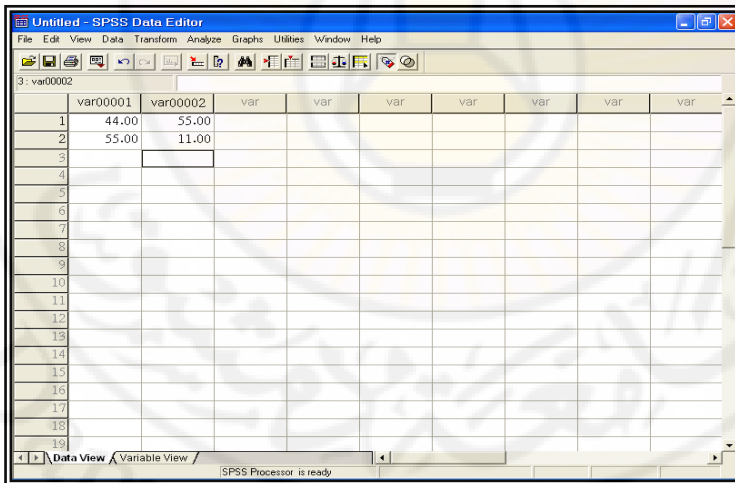
8- اختيار الموظفين من الذكور و سنوات تعليمهم 17 سنة فأكثر و مدراء نكتب الصيغة التالية :

Gender="m"&educ>=17&jobcat=3

عملية إدخال البيانات في SPSS

نحن نفترض هنا أن برنامج SPSS موجود على جهازك ولتشغيله انقر فوق زر البدء " ابدأ " أو "Start" من شاشة تشغيل النوافذ اختر " برامج Programs " انقر فوق أيقونة " SPSS for windows " ثم تنتج قائمة فرعية اختر " SPSS " فيتم فتح الشاشة التالية والتي تسمى نافذة محرر البيانات (Data Editor) :

لاحظ أن محرر البيانات هو عبارة عن شبكة من الصفوف والأعمدة تستخدم لإنشاء وتحرير ملفات البيانات. وفي محرر البيانات فإن كل صف يمثل حالة (Case) أي إن الصف الأول يفرغ فيه إجابات الاستبيان الأول والصف الثاني يفرغ فيه إجابات الاستبيان الثانية وهكذا....



الملاحق

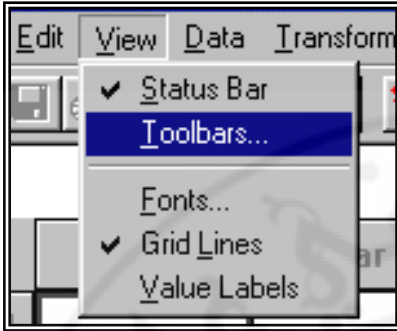
أما الأعمدة فتتمثل المتغيرات أي أن كل سؤال في الاستبيان يمثل بمتغير (Variable) أي بعمود. وتسمى نقاط التقاطع بين الصف والعمود بالخلية (Cell).

كما يوجد في أعلى شاشة محرر البيانات شريط العنوان وشريط القوائم وشريط محرر البيانات وفي أسفل شاشة محرر البيانات يوجد عرض البيانات (Data View) لعرض البيانات وكذلك يوجد عرض المتغيرات (Variable View) لعرض خصائص المتغيرات (اسم المتغير ونوعه و...) وكذلك نشاهد أشرطة التمرير الرأسية والأفقية على الجانب الأيمن والجهة السفلي لشاشة محرر البيانات.

وقبل البدء في كيفية إدخال البيانات سنشير إلى وظائف الأيقونات التي يحتويها شريط الأدوات (شريط محرر البيانات Data Editor) و الموضح بالشكل التالي:

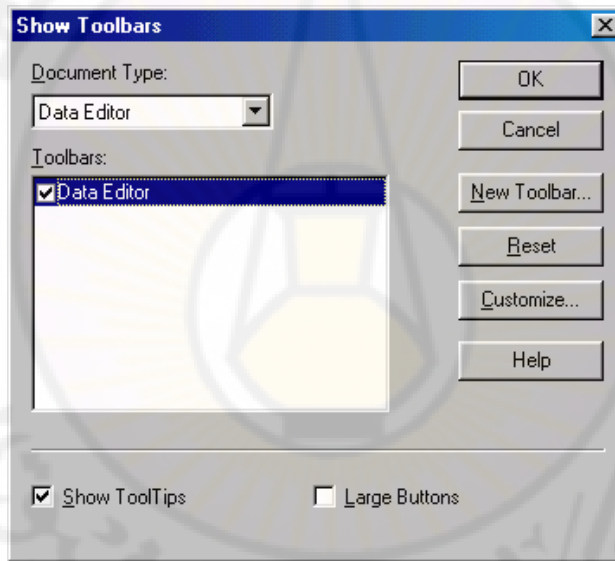


الوظيفة	العنوان	الأيقونة
فتح ملف مخزن	open	
تخزين ملف	Save	
طباعة ملف	Print	
إظهار آخر مجموعة من الإجراءات التي تم استخدامها	Dialog Recall	
تراجع عن آخر عملية قمت بها	Undo	
الرجوع عن آخر عملية تراجع عنها	Redo	
الانتقال إلى تخطيط	Goto Chart	
الانتقال إلى حالة (صف)	Goto Case	
إعطاء معلومات عن المتغير	Variable	
بحث عن	Find	
إدراج حالة جديدة إلى الملف	Insert Case	
إدراج متغير جديد إلى الملف	Insert Variable	
شطر الملف إلى جزأين	Split File	
إعطاء أوزان للحالات	Weight Cases	
اختيار مجموعة حالات	Select Cases	
إظهار (أو إخفاء) عناوين (دلالات) القيم	Value Labels	
استخدام مجموعات من المتغيرات	Use Sets	

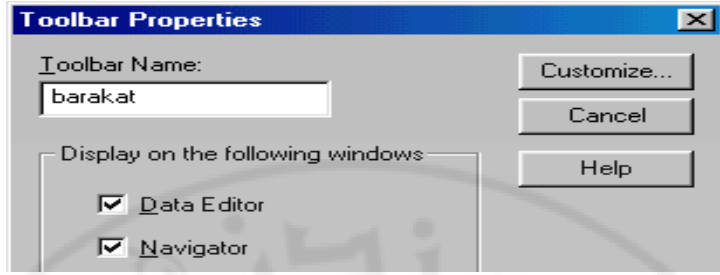


أيقونات spss

لإيجاد الشريط الموجود تحت شريط القوائم
نضغط من شريط القوائم على View ثم
نختار Toolbars فيظهر مربع الحوار التالي
نضغط في المربع المقابل لـ Data Editor



فتظهر علامة الصح، وإذا أردنا تكبير أزار الشريط نضغط أمام Large Buttons . أما إذا أردنا إيجاد شرائط جديدة نحن في حاجة لها فإننا نضغط على زر New Toolbar فيظهر مربع الحوار التالي :



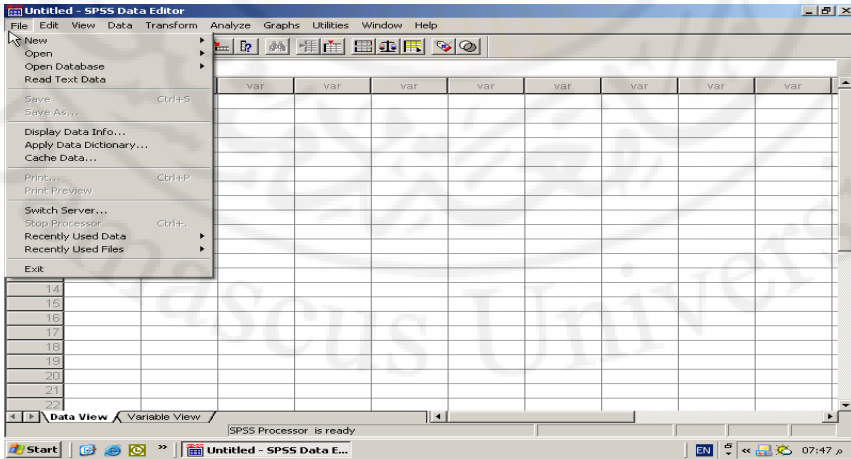
نكتب اسم الشريط الجديد على سبيل المثال barakat ثم نضغط على customize فيظهر الشكل التالي:

نختار من القائمة Categories ما نراه مناسباً ومن المستطيل المقابل نختار الـ Items المناسب بالضغط على الزر الأيسر للفارة مرتين متتاليتين فينتقل الزر إلى المستطيل الأفقي Customizing Toolbar المسمى barakat تم نضغط أخيراً على موافق فيظهر شريط جديد باسم barakat

استخدام شريط الأوامر SPSS :

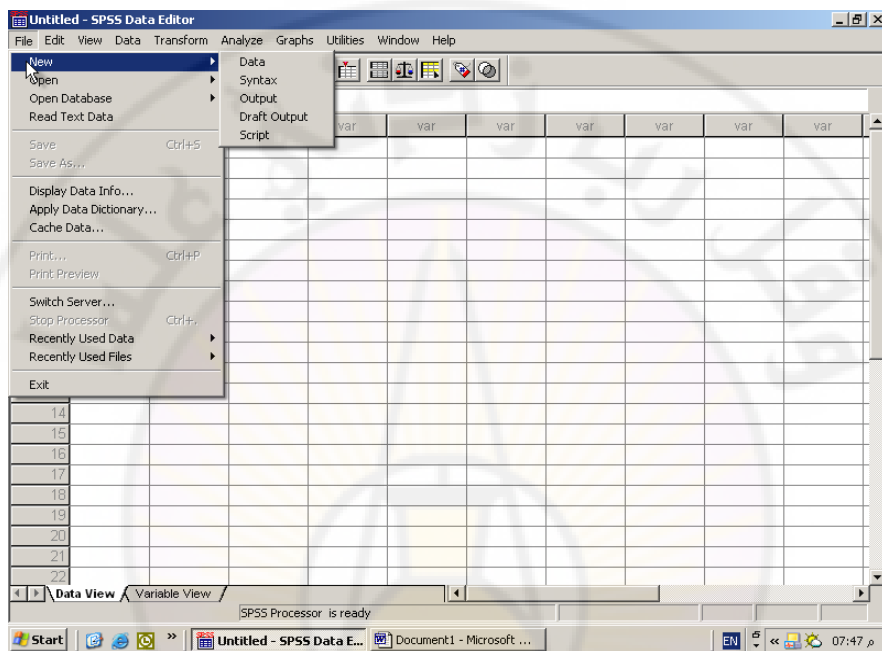
فيما يلي أبرز مكونات شريط الأوامر في SPSS والتي ستتعرف على استخداماتها في فصول الكتاب المرفق.

لاحظ أن (الفأرة) تشير إلى الأمر (File) ملف وبالتأشير عليها تنسدل قائمة تحتوي الأوامر الفرعية المدرجة.



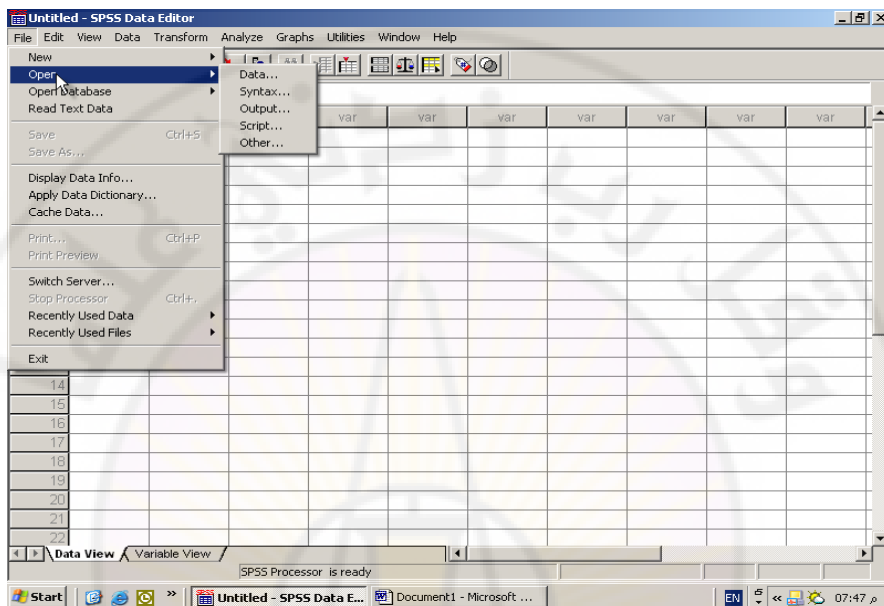
الملاحق -----

وبالتأشير على (New) جديد تظهر قائمة فرعية يهملك منها (Data) لإنشاء ملف بيانات جديد وكذلك (Output) التي تتعامل مع إنشاء ملف مخرجات جديد

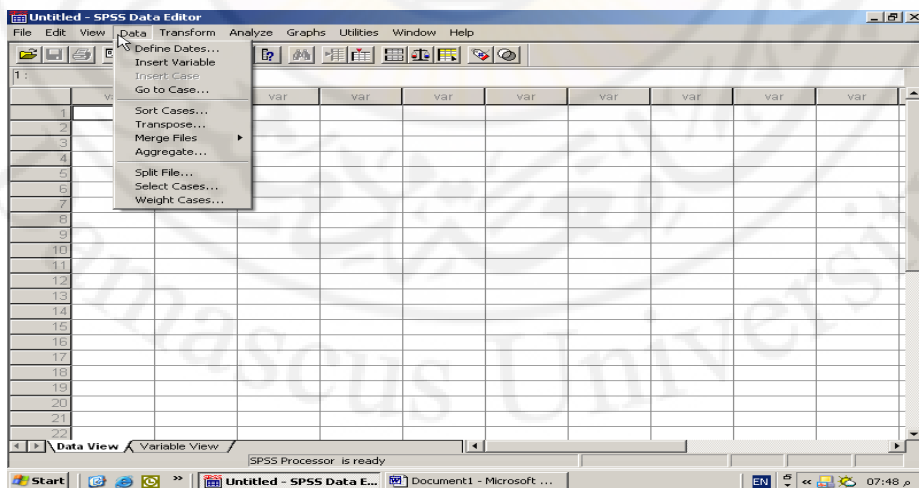


الملاحق -----

وبالمثل يعطي المؤشر على (Open) عدة أوامر فرعية أهمها (Data) و (Output) اللذان تستخدمان لفتح ملف بيانات مخزن أو ملف مخرجات مخزن.

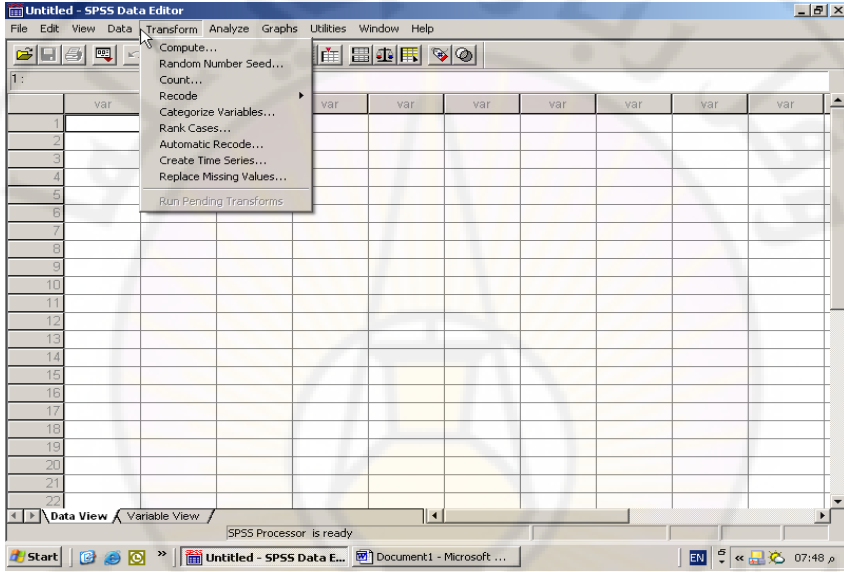


الفأرة تشير إلى (Data) للتعامل مع ملف البيانات عن طريق (Define) للتعريف (insert Variable) لإدخال عمود بيانات، (Sort Cases) لترتيب حالات البيانات وهكذا.

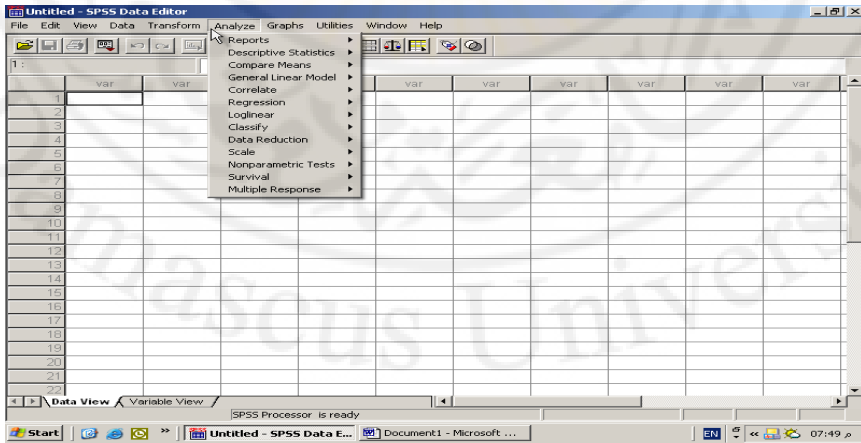


الملاحق

الفأرة تشير إلى (Transform) التي يمكن أن تستخدم لعدة أهداف منها (Compute) لإجراء الحسابات، (Recode) لإعادة ترميز البيانات بتحويلات معطاة الفأرة تشير إلى (transform) والتي يمكن أن تستخدم لعدة أهداف منها (compute) لإجراء الحسابات ، (Recode) لإعادة ترميز البيانات بتحويلات معطاه .

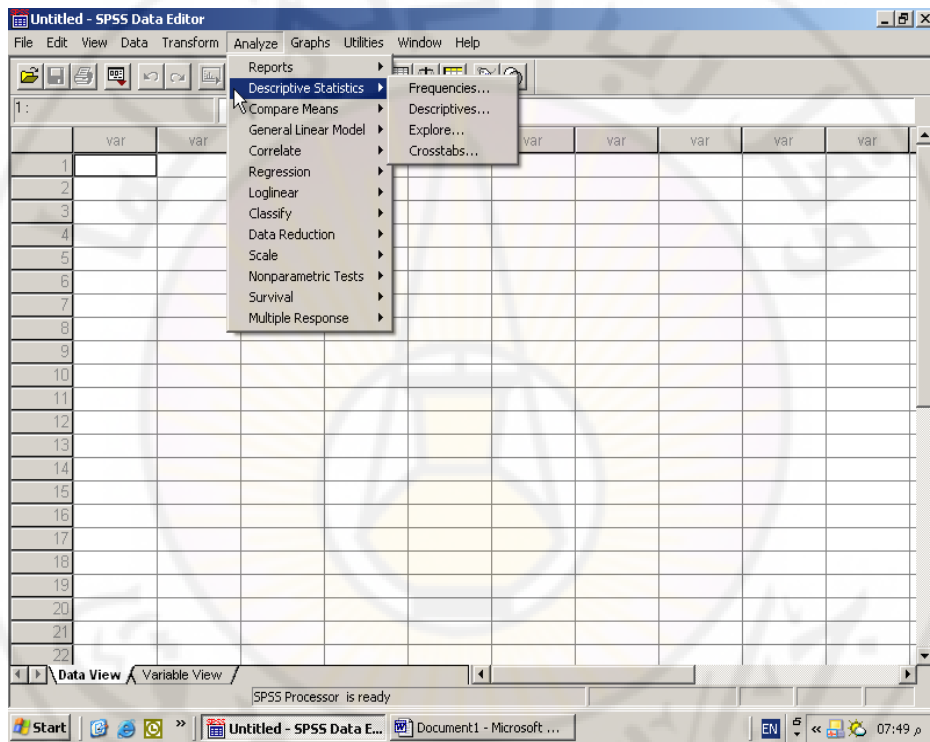


الفأرة تشير إلى (Analyze) وينسدل عنها أوامر العمليات الإحصائية.



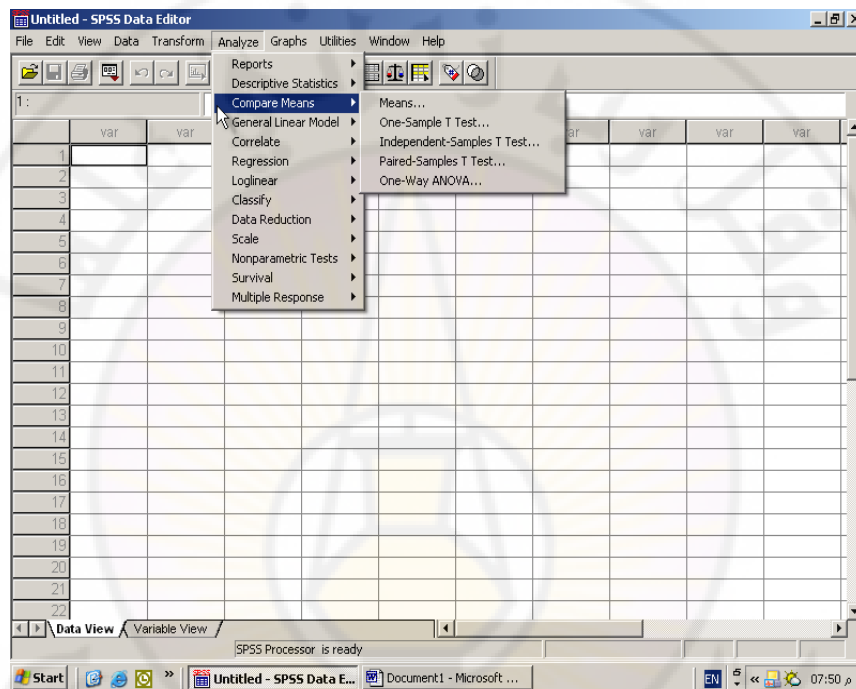
الملاحق

الفأرة تشير إلى (Descriptive Statistics) أي الإحصاءات الوصفية ويتفرع عنها (Frequencies) التكرارات، (Descriptive) المقاييس الوصفية كمقاييس النزعة المركزية ومقاييس التشتت، (Cross tabs) تعطي جداول التوافق وما يتعلق بها من إحصاءات.



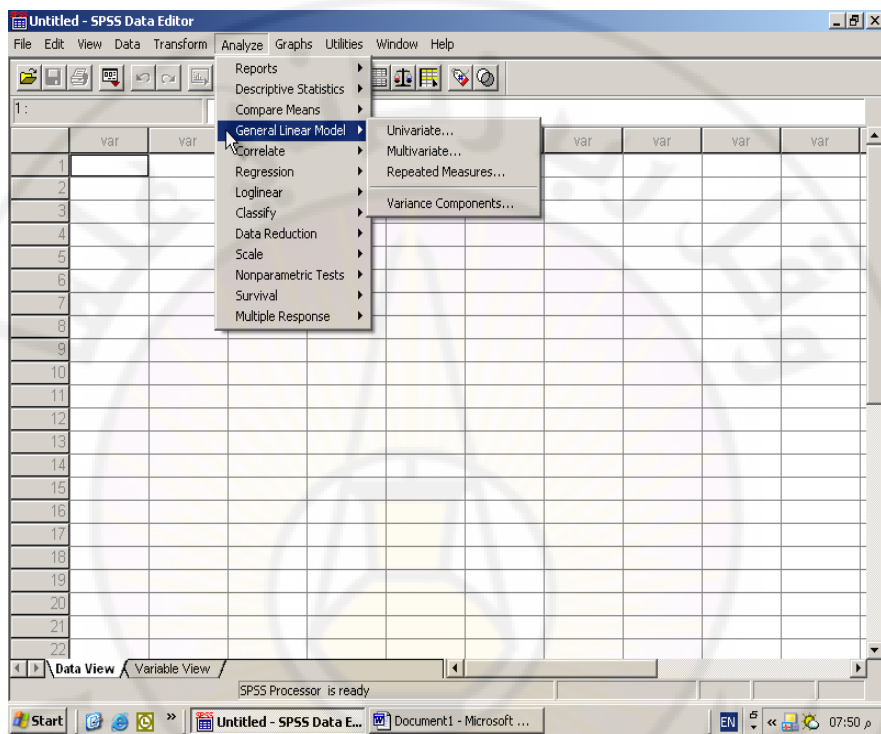
الملاحق

تشير الفأرة (Compare Means) لمقارنة الأوساط ويتفرع عنها اختبار t واختبار t لمقارنة وسطي مجتمعين مستقلين واختبار t لمقارنة وسطي مجتمعين مرتبطين.



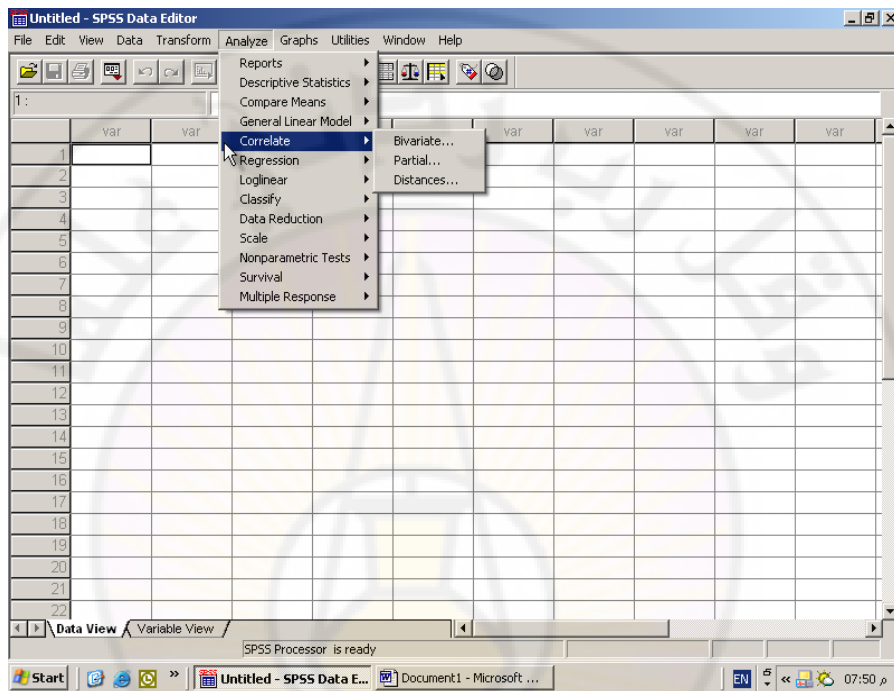
الملاحق -----

تشير الفأرة إلى (General Liner Model) والذي يهتم منه هو الأمر الفرعي (Univariate) الذي يستخدم لتحليل التباين.

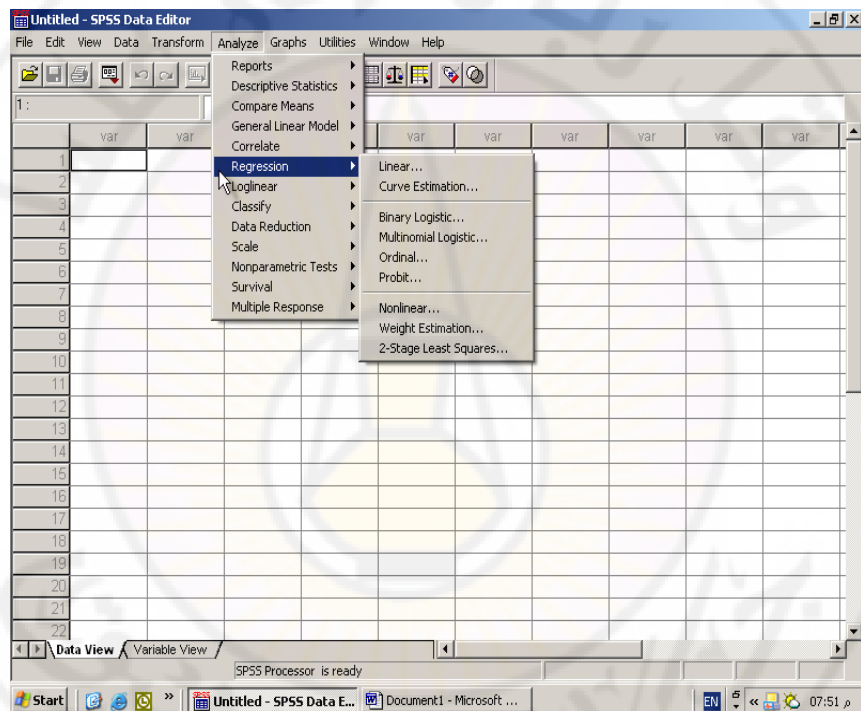


الملاحق -----

تشير الفأرة إلى (Correlate) وما يهكم مما يتفرع عنها (Bivariate) الذي يستخدم لحساب معاملات الارتباط.

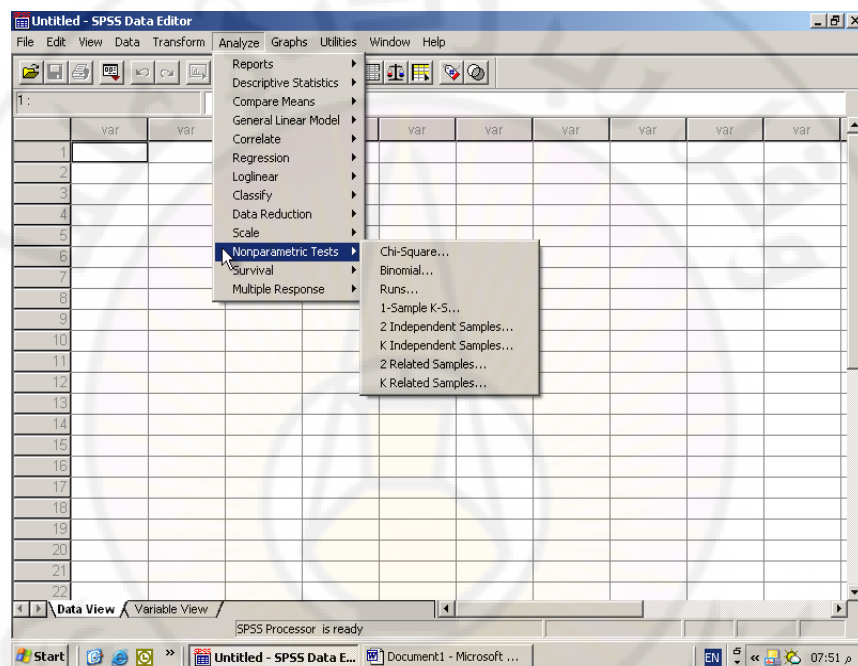


تشير الفأرة إلى (Regression) ومايهمك منه الامر الفرعي (Linear) الذي يستخدم لإجراء تحليل خط الانحدار البسيط أو المتعدد.



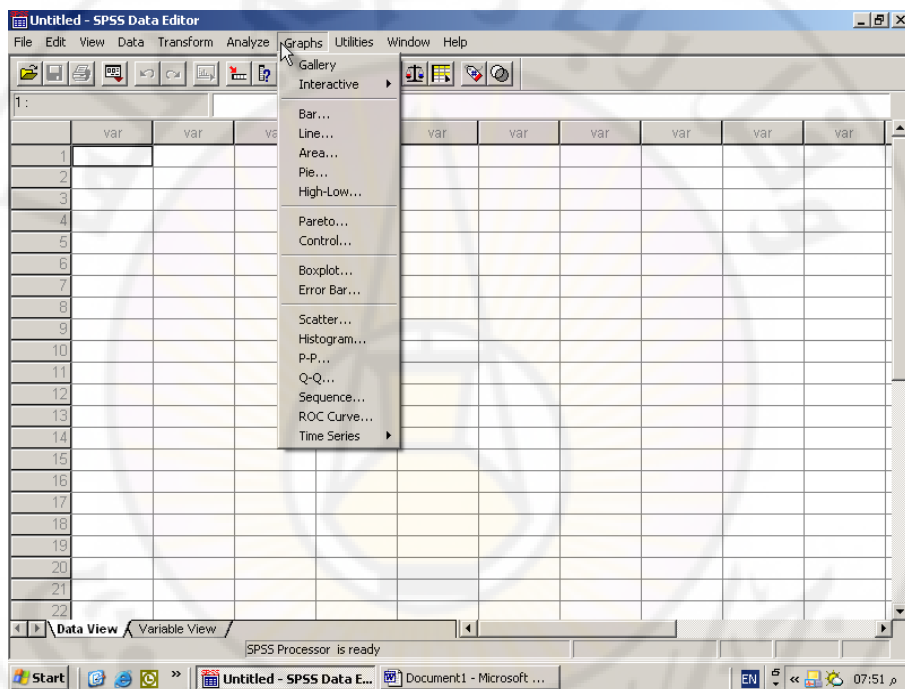
الملاحق -----

تشير الفأرة إلى (Nonparametric Tests) الاختبارات اللامعلمية وينسدل عنها عدة أوامر تعالج اختبار كاي تربيع، واختبار كروسكال واليس واختبارات مقارنة مجتمعين مستقلين ومجتمعين غير مستقلين ومقارنة عدة مجتمعات مستقلة وعدة مجتمعات مترابطة.



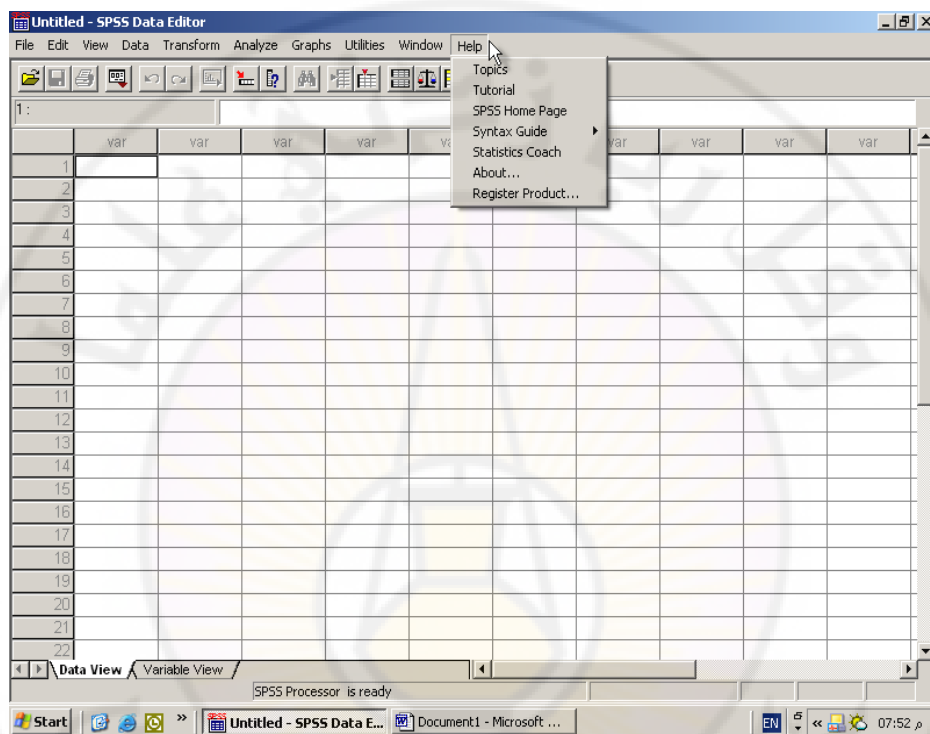
الملاحق -----

تشير الفأرة إلى (Graphs) الرسومات وينسدل عنها عدة أوامر لتمثيل البيانات بالرسم مثل (Bar) الأعمدة، (Scatter) شكل الانتشار، (Histogram) المدرج التكراري وهكذا.



الملاحق -----

تشير الفأرة إلى (Help) المساعدة ويمكن الاستعانة بها للبحث عن مساعدة في كيفية استخدام أمر ما.



----- الملاحق

والآن نوضح كيفية إدخال البيانات التالية والتي تهدف إلى معرفة اتجاهات المعلمين نحو الوسائل التعليمية:

استبيان

المؤهل العلمي: ☐ دبلوم ☐ بكالوريوس فما فوق ☐

الخبرة: ☐ أقل من 5 سنوات ☐ من 5-10 سنوات ☐ أكثر من 10 سنوات ☐

الرقم	الفقرة	موافق بشدة	موافق	محايد	معارض	معارض بشدة
1	أشعر بارتياح لاستخدام الوسيلة التعليمية					
2	أفضل عرض الوسيلة التعليمية في وقتها المناسب					
3	أرى أن استخدام الوسيلة التعليمية تحسن نوعية التعليم					

- نقوم بعملية الترميز للمتغيرات:

أولاً : متغير المؤهل العلمي:

المؤهل العلمي	دبلوم	بكالوريوس فما فوق
التصنيف	1	2

ثانياً: الخبرة:

الخبرة	أقل من 5 سنوات	من 5-10 سنوات	أكثر من 10 سنوات
التصنيف	1	2	3

ثالثاً: يتم تفرغ البيانات وفقاً للتصنيف التالي:

التصنيف	موافق بشدة	موافق	محايد	معارض	معارض بشدة
الدرجة	5	4	3	2	1

نعطي أسماء لمتغيرات أسئلة الدراسة كالتالي: المؤهل، الخبرة، q1, q2, q3

الملاحق

* نضغط على Variable View تظهر الشاشة التالية والتي تستخدم في تعريف متغيرات الدراسة .

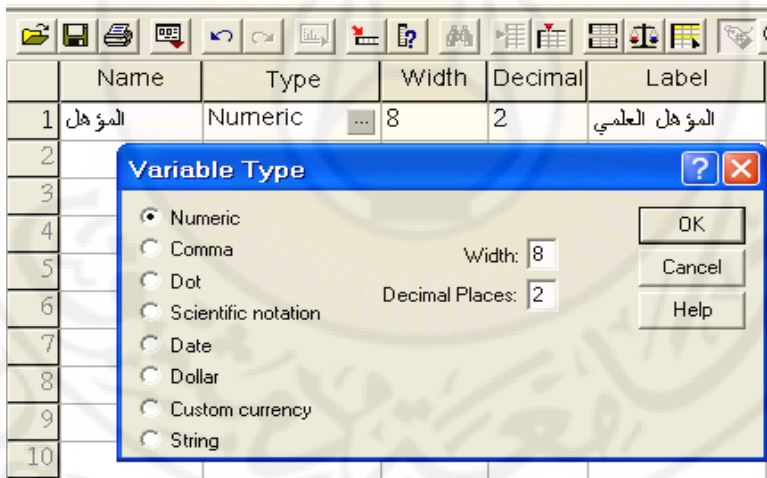
نلاحظ من الشاشة أن للمتغير عدة خواص هي الاسم Name والنوع Type ووصف المتغير Label وغيرها كما تشاهد في الشكل أعلاه وسوف نأتي بالتفصيل لكيفية إدخال متغير المؤهل العلمي، وسوف يكون إدخال بقية المتغيرات مشابه تاماً:

المرحلة الأولى: كتابة اسم المتغير

نضغط في الخلية أسفل Name في السطر الأول لنكتب اسم المتغير " المؤهل "

المرحلة الثانية: تعيين نوع المتغير

نضغط في الخلية أسفل Type فتظهر أيقونة عليها ثلاث نقاط نضغط عليها فيظهر لنا الشكل التالي:



□ **Numeric** من الشكل نلاحظ أن SPSS يعتبر أن جميع المتغيرات رقمية وعرضها 8 Width أي 8 أرقام وكذلك عدد الأرقام العشرية Decimal Places 2 ويمكن تغيير عدد أرقام العدد وكذلك عدد الأرقام العشرية بالضبط داخل المربع المعني أو في الخلية أسفل العمود Width أو أسفل

الملاحق

العمود **Decimal** في شاشة محرر البيانات ونقوم بتغيير عدد أرقام العدد وكذلك عدد الأرقام العشرية كما هو مبين بالشكل:-

☐ **Comma** لتعريف متغير رقمي يراد عرض قيمه بحيث تشتمل على فاصلة

Type	Width	Decimal
Numeric	6	3

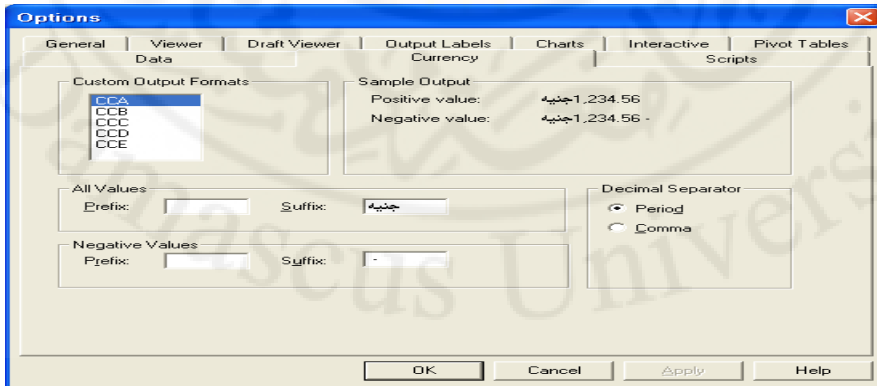
☐ كل ثلاثة أرقام (للأرقام الأكبر من 1000) مع نقطة لفصل الخانات العشرية. وكمثال على ذلك 545,445,555.000 .

☐ **Dot** - لتعريف متغير رقمي يراد عرض قيمه بحيث تشتمل على نقطة كل ثلاثة أرقام (للأرقام الأكبر من 1000) مع فاصلة لفصل الخانات العشرية وكمثال على ذلك العدد 545.445.555,000 .

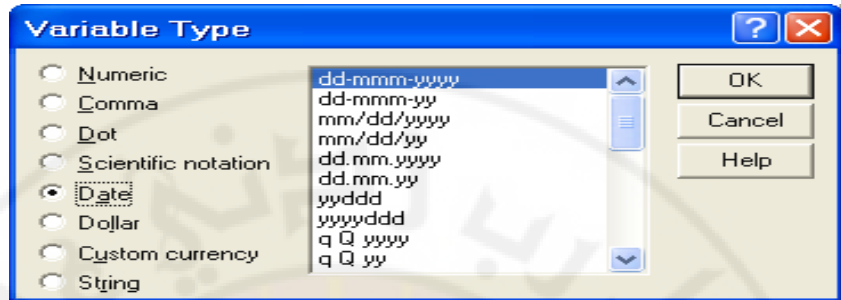
☐ **Scientific Notation** لتعريف متغير رقمي يراد عرض قيمة بشكل تعبير أسي وفي هذا النوع يستخدم الحرف (E) ليسد مسد الأساس (10) فالرقم 4.51×10^2 يعبر عنه حسب هذا النوع كما يلي 4.51E2

☐ **Date** لتعريف متغير رقمي يراد عرض قيمة بشكل تاريخ أو تاريخ مع الوقت وصندوق الحوار التالي يبين أشكال خاصة من هذا النوع

☐ وكمثال يمكن اختيار الشكل mm/ dd/ yy وهو التاريخ على الطريقة الأمريكية ورمز mm يعني الشهر و dd تعني اليوم و yy تعني السنة.

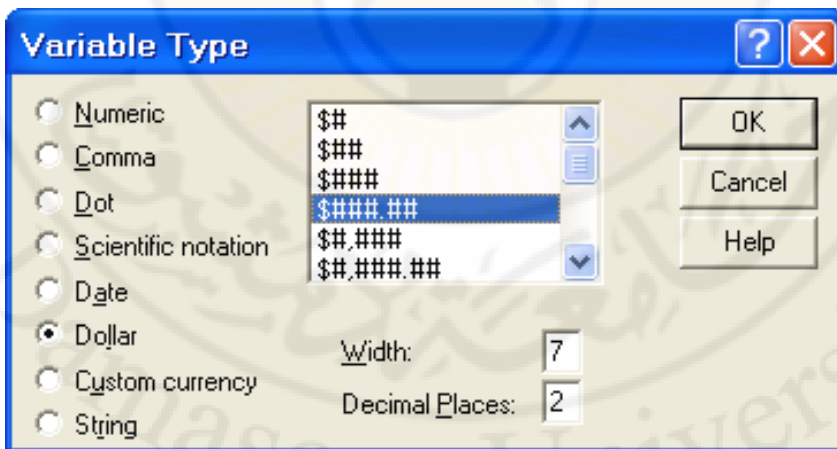


وكمثال 05/06/99 .



☐ **Dollar** لتعريف متغير رقمي يراد عرض قيمة بحيث تشمل على إشارة الدولار \$ مع فاصلة كل ثلاثة أرقام (العدد اكبر من 1000) مع نقطة لفصل الخانات العشرية.

والشكل التالي يبين هذا النوع :وكمثال على قيم متغير من هذا النوع



\$,505,487.14

الملاحق

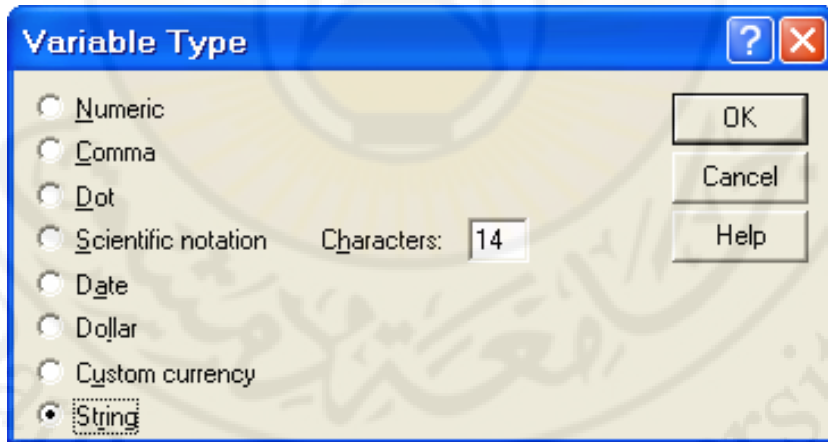
□ **Custom Currency** : لتعريف متغير رقمي يراد عرض قيمة بحيث تشمل على عملة دولة معينة تم تعريف مواصفاتها حسب الطلب، لذلك قبل اختيار هذا النوع فانه يجب أولاً إنشاء العملة المطلوبة كما يلي:

- اختار القائمة Edit ثم اختيار الأمر Options فيظهر مربع الحوار التالي،
اختار النافذة Currency ثم في مربع All Values اكتب في المربع المقابل لـ
Suffix " جنيه " وفي مربع Negative Values اكتب إشارة السالب " - "
في المربع المقابل لـ Suffix ثم موافق.

وكمثال على هذا النوع: - 454.000 جنيه .

□ **String** : لتعريف متغير حرفي قيمه تحتوي على أحرف أو أرقام أو أي رموز أخرى، والشكل التالي يبين هذا النوع:

□ في مربع Characters ادخل أقصى عدد ممكن للرموز، ويجب معرفة انه يوجد فرق بين الحروف الصغيرة والكبيرة أي أن الحرف a يختلف عن الحرف A .



المرحلة الثالثة: تعيين الأوصاف للمتغير

□ لتعيين وصفا للمتغير (variable Label) وتعيين رموزا (Values) تستخدم كأوصاف لقيم المتغير (Value Labels) اضغط داخل الخلية أسفل Label في شاشة Variable View لكتابة نص السؤال وهو " المؤهل العلمي".

□ في الخلية أسفل Values اضغط على المربع المنقط يظهر مربع الحوار التالي:

□ اكتب 1 أمام Value و دبلوم أمام Value Label ثم اضغط على زر Add ، ثم اكتب 2 في المستطيل المقابل لـ Value ثم اكتب " بكالوريوس فما فوق" في المستطيل المقابل لـ Value Label ثم اضغط على Add .



- لتغيير وصف قيمة المتغير: ظلل الوصف المطلوب بنقره بالفارة ثم ادخل القيمة الجديدة في مستطيل Value أو الوصف في مستطيل Value Label) ثم انقر الزر Change ، فيظهر الوصف الجديد.

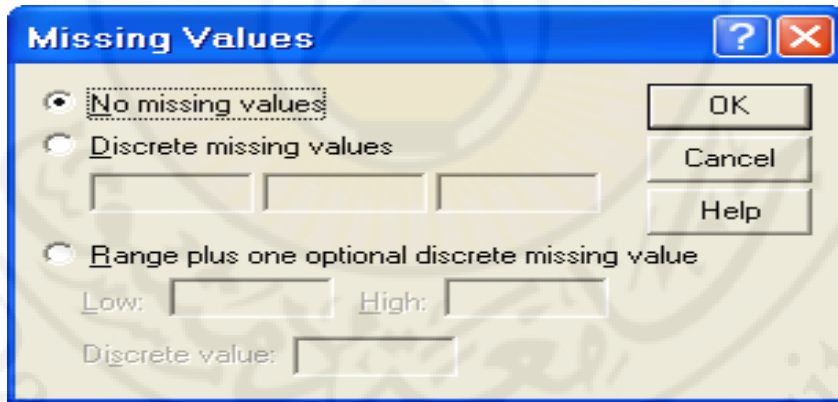
لحذف وصف قيمة في المتغير: ظلل الوصف المطلوب من القائمة بنقره بالفارة ثم انقر زر Remove ، فيتم حذف الوصف من القائمة.

المرحلة الرابعة: تحديد القيم المفقودة

أحيانا قد يقوم بعض الأشخاص بعدم الإجابة على سؤال ما تبقى إجابة ذلك السؤال مفقود وتسمى بالقيمة المفقودة، ويجب إبلاغ الجنرال SPSS بذلك، وهناك عدة طرق لتعيين القيم المفقودة، نذكر منها:

□ عندما يكون هناك سؤال ليس له إجابة فما عليك إلا أن تفقر عنه، ليقوم محرر البيانات بعرض تلك الخلية المفقودة بنقطة، وتسمى تلك القيم المفقودة " قيم نظام مفقودة (System Missing Values) " وجدير بالذكر انه بالنسبة للمتغيرات الرقمية فان الخلايا تحول إلى قيم نظام مفقودة ، أما بالنسبة للمتغيرات النصية فان الخلايا الفارغة تعامل كقيمة صحيحة، بمعنى آخر لا يوجد قيم مفقودة في المتغيرات النصية.

□ يمكنك أن تضع رمزا بدل القيم المفقودة لتصبح تلك القيم " قيم المستخدم المفقودة User Missing Values " ولتحديد قيم مستخدم مفقودة نضغط في الخلية الموجودة اسفل Missing في شاشة " محرر البيانات " ثم الضغط على المربع المنقط بثلاث نقط ليظهر الشكل التالي:



و يظهر من مربع الحوار عدة خيارات لتعيين القيم المفقودة كالتالي:

No missing values

يتم اختياره عند عدم وجود قيم مستخدم مفقودة وعادة يكون هذا الخيار محددا.

Discrete missing values

يمكنك إدخال حتى ثلاث قيم مختلفة لمتغير واحد تعامل كقيم مستخدم مفقودة وهذا الخيار يصلح للمتغيرات الرقمية والنصية.

Range of missing values

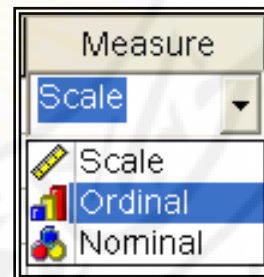
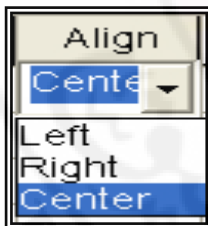
يمكنك هذا الخيار من تحديد مدى معين من قيم المستخدم المفقودة بحيث تعامل اقل قيمة واكبر قيمة وما بينهما من القيم كقيم مفقودة. ويصلح هذا الخيار فقط للقيم الرقمية ولا يصلح للمتغيرات النصية.

Range plus one discrete missing value

يمكنك هذا الخيار من تحديد مدى معين من قيم مستخدم مفقودة إضافة إلى قيمة خارج المدى، ويصلح هذا الخيار للمتغيرات الرقمية ولا يصلح للمتغيرات النصية. وجدير بالذكر أن قيم المستخدم المفقودة لا تدخل في الحسابات.

المرحلة الخامسة : تحديد شكل العمود

يقصد بشكل العمود عرض العمود (Column width) وموقع البيانات داخل العمود (Text Format) بحيث يمكن توجيهها بحيث تكون في يسار العمود أو في وسطه أو في يمينه. ولتغيير ذلك نضغط في الخلية اسفل Column واسفل Align ونختار المناسب.



المرحلة السادسة : تحدي مقياس المتغير

لتحديد مقياس المتغير نضغط داخل الخلية أسفل Measure ثم نضغط على السهم الموجود داخل الخلية فتظهر الخيارات التالية كما بالشكل أعلاه ، نختار منها Nominal .

تمرين :إليك الاستبيان التي عرضت في بداية هذا الفصل والمطلوب توزيعها على عينة عدد مفرداتها 10 وتفرغها في SPSS .

----- الملاحق

المؤهل العلمي: ☐ دبلوم ☐ استبانة ☐ بكالوريوس فما فوق ☐

الخبرة: ☐ اقل من 5 سنوات ☐ من 5-10 سنوات ☐ اكثر من 10 سنوات ☐

الرقم	الفقرة	موافق بشدة	موافق	محايد	معارض ض	معارض بشدة
1	أشعر بالارتياح لاستخدام الوسيلة التعليمية					
2	أفضل عرض الوسيلة التعليمية في وقتها المناسب					
3	أرى أن استخدام الوسيلة التعليمية تحسن نوعية التعليم					

الحل:

- نقوم بعملية الترميز للمتغيرات:

أولا : متغير المؤهل العلمي:

المؤهل العلمي	دبلوم	بكالوريوس فما فوق
التصنيف	1	2

ثانيا: الخبرة:

الخبرة	اقل من 5 سنوات	من 5-10 سنوات	اكثر من 10 سنوات
التصنيف	1	2	3

ثالثا: يتم تفريغ البيانات وفقا للتصنيف التالي:

التصنيف	موافق بشدة	موافق	محايد	معارض	معارض بشدة
الدرجة	5	4	3	2	1

نعطي أسماء لمتغيرات أسئلة الدراسة كالتالي: المؤهل، الخبرة، q1، q2، q3

بعد تفريغ البيانات تظهر شاشة محرر المتغيرات كالتالي:

وكالة - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Window Help

	Name	Type	Width	Decimals	Label	Values	Missing	Columns
1	المؤهل	Numeric	8	0	المؤهل العلمي	{1... [ليقوم]	None	8
2	الخبرة	Numeric	8	0	الخبرة	{1... [من 5 سنوات]	None	8
3	q1	Numeric	8	0	شعر بالارتياح عن	{1... [معارض بشدة]	None	8
4	q2	Numeric	8	0	شعر بالارتياح عن	{1... [معارض بشدة]	None	8
5	q3	Numeric	8	0	أرى أن في استخدام	{1... [معارض بشدة]	None	8
6	q4	Dollar	4	0	الراتب الشهري	{1... [معارض بشدة]	None	8

والبيانات بعد التفريغ تظهر على شاشة محرر البيانات كالتالي:

The screenshot shows the SPSS Data Editor window. The title bar reads 'وكالة - SPSS Data Editor'. The menu bar includes File, Edit, View, Data, Transform, Analyze, Graphs, Utilities, Window, and Help. The toolbar contains various icons for file operations, editing, and data management. Below the toolbar, a label 'الخبرة: 13' is visible. The main data grid has 6 columns: an empty column, 'المؤهل' (Degree), 'الخبرة' (Experience), 'q1', 'q2', and 'q3'. The data is as follows:

	المؤهل	الخبرة	q1	q2	q3
1	دبلوم	أقل من 5	موافق	موافق بشدة	محايد
2	دبلوم	من 5-10	محايد	موافق	معارض
3	دبلوم	أقل من 5	موافق	موافق بشدة	موافق بشدة
4	دبلوم	أكثر من 10	موافق بشدة	موافق	موافق
5	بكالوريوس	من 5-10	موافق بشدة	موافق	موافق
6	بكالوريوس	أكثر من 10	موافق	موافق	موافق بشدة
7	بكالوريوس	من 5-10	محايد	محايد	محايد
8	بكالوريوس	أكثر من 10	موافق بشدة	موافق بشدة	موافق بشدة
9	بكالوريوس	من 5-10	معارض	معارض	موافق
10	بكالوريوس	أكثر من 10	موافق بشدة	موافق بشدة	موافق بشدة

الفصل الثاني

العمليات الحسابية واختيار الحالات

عند تحليل الإستبانة يلزم في بعض الأحيان إيجاد بعض العمليات الحسابية على بعض المتغيرات وهنا سنركز على بعض الدوال الهامة التي لها اتصال مباشر بتحليل الإستبانة.

□ حساب مجموع عدة متغيرات

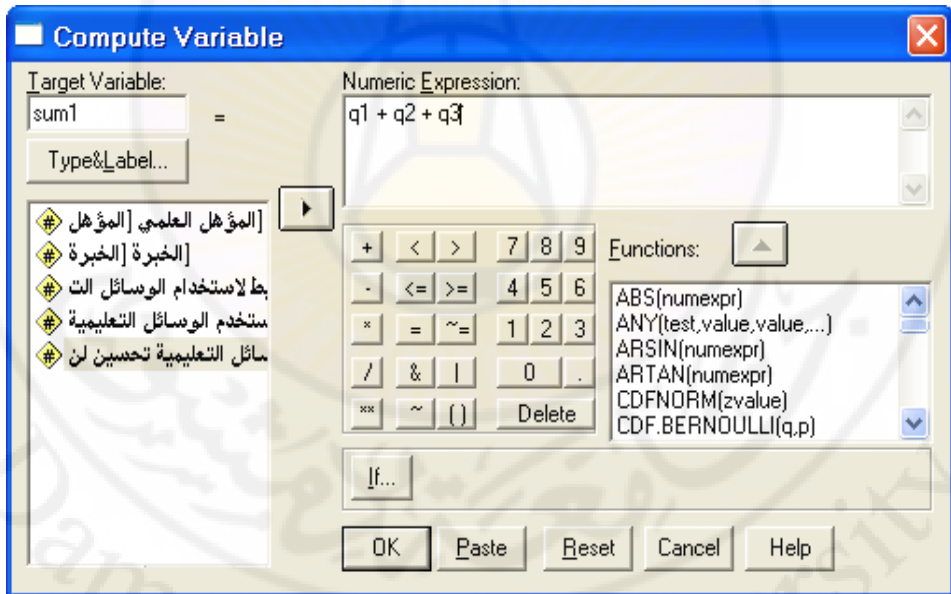
❖ عملية الجمع

مثال: احسب مجموع المتغيرات q_1, q_2, q_3 الواردة في الإستبانة السابقة

الحل: لحساب مجموع المتغيرات الثلاثة

- نختار Compute من شريط القوائم Transform فيظهر مربع الحوار التالي:

في المستطيل Target Variable ادخل اسم المتغير الجديد المطلوب



وليكن sum1 ويجب أن يكون الاسم مخالف لأسماء المتغيرات في الإستبانة.

في المستطيل Numeric Expression اكتب $q_1+q_2+q_3$ ويمكنك

كتابة ذلك باستخدام لوحة المفاتيح أو باستخدام أزرار الآلة الحاسبة الموجودة في

----- الملاحق

مربع الحوار أو بالنقر على اسم المتغير مرتين من قائمة المتغيرات أو بنقر المتغير مرة واحدة ثم الضغط على السهم ليدخل داخل صندوق Numeric Expression

إذا أردت أن تكتب وصف للمتغير اضغط على الزر Type&Label فيظهر مربع الحوار التالي:

Compute Variable: Type and Label

Label

☒ Label: q1.q2.q3 مجموع المتغيرات

☐ Use expression as label

Type

☒ Numeric

☐ String Width: 8

Continue

Cancel

Help

اكتب في المستطيل المقابل لـ Label ثم اضغط على Continue فينتقل إلى مربع الحوار السابق ، اضغط على Ok فتظهر النتائج التالية:

	المؤهل	الخبرة	q1	q2	q3	sum1
1	1	1	4	5	3	12.00
2	2	2	5	4	4	13.00
3	2	3	4	4	5	13.00
4	2	2	3	3	3	9.00
5	1	2	3	4	2	9.00
6	2	3	5	5	5	15.00
7	1	1	4	5	5	14.00
8	2	2	2	2	4	8.00
9	1	3	5	4	4	13.00
10	2	3	5	5	5	15.00

ملاحظة هامة 1: عند استخدام طريقة الجمع السابقة إذا كانت إحدى قيم المتغيرات مفقودة فإن نتيجة الجمع للمتغيرات ستكون مفقودة، ولذلك يفضل استخدام دالة SUM من فئة الدوال Functions وكتابة الصيغة التالية داخل مستطيل **Numeric Expression** ، $\text{sum}(q1, q2, q3)$ أو $\text{sum}(q1 \text{ to } q3)$ فإنه يتم جمع قيم المتغيرات الغير مفقودة حاول أن تجرب هذه الملاحظة مع اختيار اسم جديد للمتغير الناتج.

ملاحظة هامة 2: من الممكن أن نحدد الحد الأدنى للمتغيرات غير المفقودة في المتغيرات المراد جمعها، وهذا يمكن أن يتم بإلحاقه نقطة مرفقة بالحد الأدنى لعدد المتغيرات التي لا تحتوي على قيم مفقودة في اسم الدالة كالتالي:

$\text{sum.2}(q1 \text{ to } q3)$

هذا يعني أن عملية الجمع تتم إذا وجد على الأقل متغيرين يحملان قيم أو بيانات وإلا فالنتيجة ستكون مفقودة.

ملاحظة هامة 3: يحتوي مربع الحوار Compute Variable على آلة حاسبة تحتوي على أرقام ورموز حسابية ورموز علائقية ورموز منطقية. ويمكن استخدام هذه الحاسبة مثل أية حاسبة يدوية وذلك بنقر الزر باستخدام الفأرة. ويبين الجدول التالي الرموز المستخدمة في الآلة الحاسبة:

الرموز المنطقية		الرموز العلائقية		الرموز الحسابية	
العملية	الرمز	العملية	الرمز	العملية	الرمز
يجب أن تكون جميع العلاقات صحيحة	& أو and	أقل من	<	الجمع	+
		أكبر من	>	الطرح	-
واحدة من العلاقات يجب أن تكون صحيحة	أو or	أقل أو يساوي	<=	الضرب	*
		أكبر أو يساوي	>=	القسمة	/
تقيد النفي	~ أو not	يساوي	=	الأس	**
		لا يساوي	~=	ترتيب العمليات	()

❖ عملية إيجاد المعدل

ليكن أننا نريد إيجاد معدل المتغيرات الثلاثة في كل حالة:

مثال: أوجد معدل المتغيرات الثلاثة لكل حالة من الحالات

من القائمة Transform اختر Compute فيظهر مربع الحوار المسمى (Compute Variable) .. انظر المثال السابق.. اختر الدالة mean من قائمة الإقتران Functions لتتقلها في داخل المستطيل Numeric Expression وتكتب الصيغة التالية mean(q1 to q3) ، ثم اختر اسما جديد للمتغير الجديد واكتبه داخل مستطيل Target Variable وليكن المعدل ثم اضغط على Ok ، فيظهر عمود جديد في شاشة البيانات باسم " المعدل "

ملاحظة هامة: إذا أردت إيجاد معدل المتغيرات الخاصة للمعلمين الذي خبرتهم أقل من 5 سنوات فقط اضغط على الزر "If" فيظهر مربع الحوار التالي:

اضغط على Include if case satisfied condition:

الملاحق

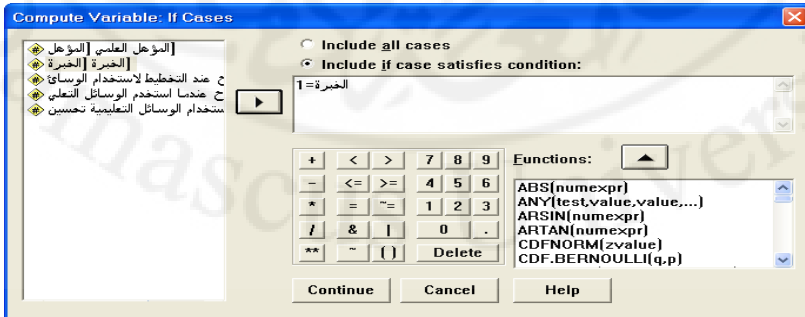
قم بإدخال الشرط المطلوب وهو الخبرة = 1 لأن الخبرة أقل من 5 سنوات رمزنا لها بالرمز 1 أتذكر ذلك ؟

انقر الزر Continue فيظهر مربع الحوار Compute Variable وتظهر عبارة الشرط بجانب الزر If ... انقر الزر Ok تلاحظ ظهور متغير جديد باسم "المعدل" في نهاية ملف البيانات يحمل قيم جديدة لمعدل المتغيرات الثلاثة

المعدل	q3	q2	q1	الخبرة	المؤهل
4.00	محايد	موافق بشدة	موافق	أقل من 5	دبلوم
.	موافق	موافق	موافق بشدة	من 5-10	بكالوريوس
.	موافق بشدة	موافق	موافق	أكثر من 10	بكالوريوس
.	محايد	محايد	محايد	من 5-10	بكالوريوس
.	معارض	موافق	محايد	من 5-10	دبلوم
.	موافق بشدة	موافق بشدة	موافق بشدة	أكثر من 10	بكالوريوس
4.67	موافق بشدة	موافق بشدة	موافق	أقل من 5	دبلوم
.	موافق	معارض	معارض	من 5-10	بكالوريوس
.	موافق	موافق	موافق بشدة	أكثر من 10	دبلوم
.	موافق بشدة	موافق بشدة	موافق بشدة	أكثر من 10	بكالوريوس

لكل حالة في حالة أن تكون الخبرة أقل من 5 سنوات بناء على الشرط .

- من الممكن أن يكون الشرط مركب ، فإذا أردنا إيجاد معدل المتغيرات الثلاثة للمعلمين الذي خبرتهم من 5 إلى 10 من حملة البكالوريوس فإننا نكتب في مستطيل الشرط الموضح في مربع الحوار Compute Variable: if



Cases الصيغة التالية:

الخبرة=1 & المؤهل=2 أو الخبرة=1 and المؤهل=2

لاحظ وجود فراغ من اليمين ومن اليسار حول كلمة and

* إذا أردنا إيجاد معدل المتغيرات الثلاثة بشرط أن للمعلمين الذي خبرتهم من 5 إلى 10 أو حاصلون على درجة البكالوريوس فننا نكتب في مستطيل الشرط العبارة التالية:

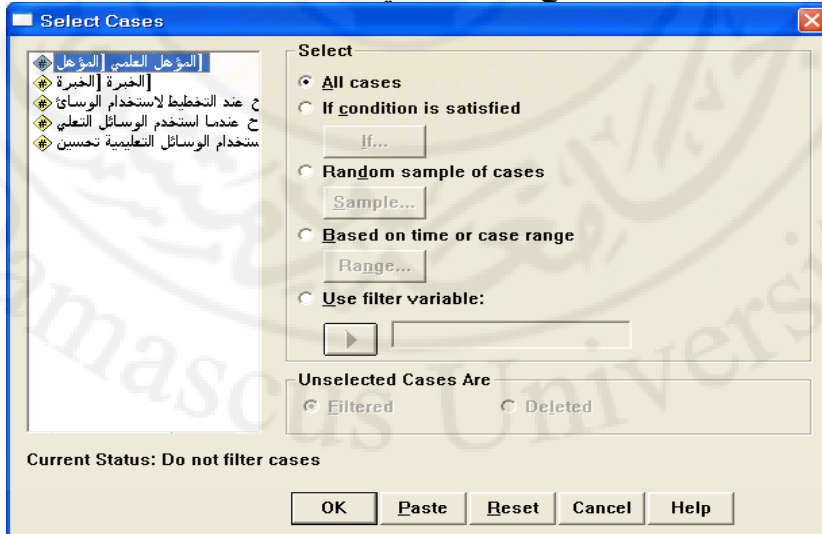
الخبرة=1 | المؤهل=2 أو الخبرة=1 or المؤهل=2

تمرين : أوجد المعدل للمتغيرات الثلاثة للمعلمين ذوى الخبرة اكبر من 5

تمرين : أوجد المعدل للمتغيرات الثلاثة للمعلمين ذوى الخبرة اكبر من 5 من حملة الدبلوم.

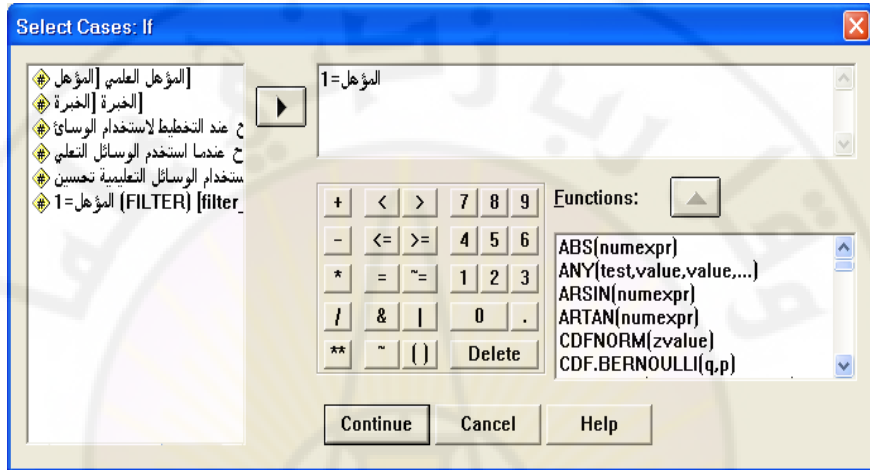
طرائق اختيار عدة حالات ☐

بإمكاننا اختيار عدة حالات يمكن للباحث إجراء التحليل عليها ولاختيار عدة حالات أو مجموعة جزئية من الحالات نختار من القائمة Data الأمر Select Cases فيظهر مربع الحوار التالي:



الملاحق

مثال: افترض أننا نريد تحديد الحالات للمعلمين الذين مؤهلهم العلمي دبلوم فقط، من مستطيل Select نختار الخيار If condition is satisfied ، ثم ننقر على الزر If فيظهر مربع الحوار التالي:



ادخل الشرط المؤهل=1 ثم اضغط Continue

في أسفل مربع الحوار Select Cases يوجد مستطيل يسمى Unselected Cases Are يوجد خياران يحددان الطريقة التي سوف نستثني الحالات المستثناه وهما

Filtered : هذا الخيار يؤدي إلى إضافة متغير في نهاية ملف البيانات يسمى filter_\$ يأخذ قيمتين، القيمة (1 او Selected) للحالات المختارة والرقم (0 او Not Selected) للحالات غير المختارة، كما أن هذا الخيار يؤدي إلى وضع إشارة " / " للحالات غير المختارة. وإذا أردت إيقاف هذا الخيار والرجوع لجميع البيانات اختر All Cases من المستطيل Select .

Deleted: هذا الخيار يؤدي إلى حذف الحالات غير المختارة ولا يمكن الرجوع إلى البيانات الأصلية إلا إذا قمنا بإغلاق البرنامج مع عدم التخزين وفتح الملف من جديد.

على كل حال سنختار Filtered ثم نضغط على Ok فتظهر النتائج التالية:

الملاحق

SPSS Data Editor - وكالة

File Edit View Data Transform Analyze Graphs Utilities Window Help

1 : filter_\$ 1

	المؤهل	الخبرة	q1	q2	q3	filter_\$
1	دبلوم	أقل من 5	موافق	موافق بشدة	محايد	Selected
2	بكالوريوس	من 5-10	موافق بشدة	موافق	موافق	Not Select
3	بكالوريوس	أكثر من 10	موافق	موافق	موافق بشدة	Not Select
4	بكالوريوس	من 5-10	محايد	محايد	محايد	Not Select
5	دبلوم	من 5-10	محايد	موافق	معارض	Selected
6	بكالوريوس	أكثر من 10	موافق بشدة	موافق بشدة	موافق بشدة	Not Select
7	دبلوم	أقل من 5	موافق	موافق بشدة	موافق بشدة	Selected
8	بكالوريوس	من 5-10	معارض	معارض	موافق	Not Select
9	دبلوم	أكثر من 10	موافق بشدة	موافق	موافق	Selected
10	بكالوريوس	أكثر من 10	موافق بشدة	موافق بشدة	موافق بشدة	Not Select
11						

إذا اخترنا Deleted من المستطيل Unselected Cases فإن النتائج تكون كالتالي:

SPSS Data Editor - وكالة

File Edit View Data Transform Analyze Graphs Utilities Window Help

3 :

	المؤهل	الخبرة	q1	q2	q3	filter_\$
1	دبلوم	أقل من 5	موافق	موافق بشدة	محايد	Selected
2	دبلوم	من 5-10	محايد	موافق	معارض	Selected
3	دبلوم	أقل من 5	موافق	موافق بشدة	موافق بشدة	Selected
4	دبلوم	أكثر من 10	موافق بشدة	موافق	موافق	Selected

❖ لاختيار عينة عشوائية من البيانات نتبع الخطوات التالية:

- من القائمة Data اختر Select Cases فيظهر مربع الحوار Select Cases كما في المثال السابق، نضغط على Random sample of cases من مستطيل Select ، ثم نضغط على Sample فيظهر مربع الحوار التالي:

Select Cases: Random Sample

Sample Size

☐ Approximately 6 % of all cases

☒ Exactly 6 cases from the first 10 cases

Continue Cancel Help

الملاحق

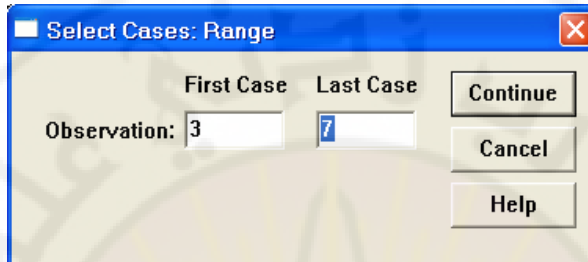
يشتمل هذا الحوار على خيارين هما Approximately وهو يحدد نسبة الحالات المئوية وذلك بادخال رقم في مستطيل هذا الخيار وليكن 60 وعلية سيتم اختيار 60% من الحالات عشوائيا. اما الخيار Exactly فيحدد عدد الخيارات من عينة حجمها n من الحالات، فاذا ادخلنا الرقم 6 امام في المستطيل الايسر المقابل لـ Exactly والرقم 10 في المستطيل الايسر، فهذا يعني اختيار 6 حالات من اول 10 حالات.

في مثالنا سنختار حالة Exactly. ونضغط على الزر Continue ثم على Ok فتظهر النتائج التالية:

SPSS Data Editor - وكالة						
File Edit View Data Transform Analyze Graphs Utilities Window Help						
المؤهل : 11						
	المؤهل	الخبرة	q1	q2	q3	filter_\$
1	دبلوم	اقل من 5	موافق	موافق بشدة	محايد	1
2	بكالوريوس	من 5-10	موافق بشدة	موافق	موافق	1
3	بكالوريوس	اكثر من 10	موافق	موافق	موافق بشدة	0
4	بكالوريوس	من 5-10	محايد	محايد	محايد	1
5	دبلوم	من 5-10	محايد	موافق	معارض	1
6	بكالوريوس	اكثر من 10	موافق بشدة	موافق بشدة	موافق بشدة	0
7	دبلوم	اقل من 5	موافق	موافق بشدة	موافق بشدة	1
8	بكالوريوس	من 5-10	معارض	معارض	موافق	0
9	دبلوم	اكثر من 10	موافق بشدة	موافق	موافق	1
10	بكالوريوس	اكثر من 10	موافق بشدة	موافق بشدة	موافق بشدة	0

❖ اختيار مدى معين من الحالات

لتحديد مدى معين من الحالات نختار من مربع الحوار Select Cases الخيار Based on time or case range فيظهر مربع الحوار التالي:



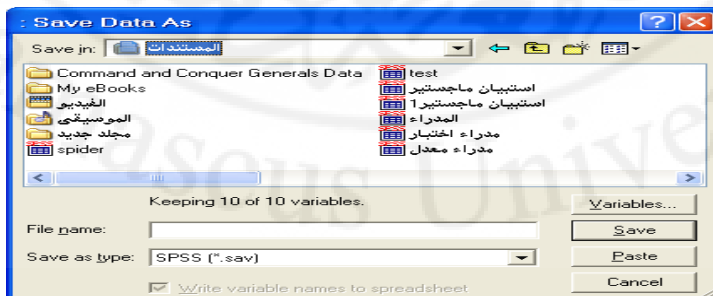
أكتب الرقم 3 أسفل First Cases والرقم 7 أسفل Last Cases وبذلك يتم إختيار الحالات من الحالة الثالثة إلى الحالة السابعة.

❖ تصفية حالات معينة

نستطيع من خلال هذا الأمر اختيار الحالات التي لا تساوي قيمتها في هذا المتغير صفراً وتحذف الحالات التي تساوي قيمتها الصفر وذلك بالنقر على Use Filter Variable ثم إدخال المتغير الذي يحتوي على بيانات تساوي الصفر وبيانات لا تساوي الصفر ، ثم نضغط Ok فنحصل على الحالات التي لا تساوي الصفر.

✓ حفظ أو تخزين البيانات Saving Data

1. لحفظ البيانات لأول مرة اختر الأمر Save As من القائمة File فيظهر مربع الحوار التالي:



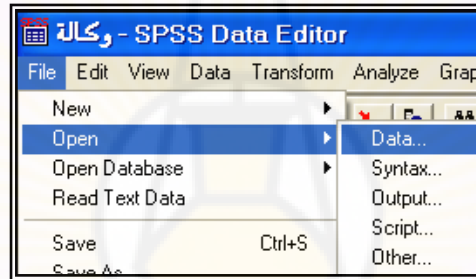
الملاحق

2. حدد الدليل الذي تريد من مربع Save In نحن اخترنا "المستندات"، ثم ادخل اسم الملف "وكالة" في مربع File Name ، لاحظ أن امتداد ملفات البيانات SPSS*.sav المقابل لمربع Save as type كما هو موضح بالشكل. ثم اضغط على الزر Save.

3. للخروج من نظام SPSS بالنقر على Exit SPSS من قائمة File .

✓ فتح ملف بيانات مخزن

1. من القائمة File اختر Open ومن القائمة الفرعية اختر Data كما هو بالشكل.



2- يظهر مربع الحوار التالي: اختر الملف المطلوب ثم اضغط الزر Open.



الملاحق

✓ لحذف متغير ، نحدد المتغير ثم نضغط Delete ، ولنسخ متغير أو عدة متغيرات حدها ثم اختر من القائمة Edit الأمر Copy ، وللصق المتغيرات بعد نسخها نختار من القائمة Edit الأمر Past .

✓ إدراج متغير (عمود) Insert Variable

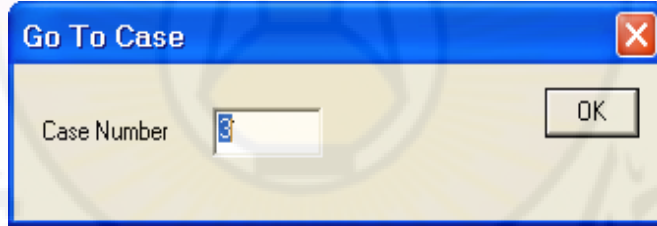
1. ضع مؤشر الفارة على العمود الذي تريد إضافة عمود جديد إلى يساره.
2. من قائمة بيانات اختر الأمر Insert Variable (أو بالنقر على الزر في شريط الأزرار الخاص بإدراج متغير) فيظهر عمود جديد باسم افتراضي Var00001

✓ إدراج حالات (صفوف) Insert Cases

1. ضع مؤشر الفارة على الصف الذي تريد إضافة صف جديد فوقه.
2. من قائمة بيانات اختر الأمر Insert Case (أو بالنقر على الزر في شريط الأزرار الخاص بإدراج صف) فيظهر صف جديد باسم افتراضي.

✓ الإنتقال إلى Go To Case

1. انقر فوق الأمر Go To Case من قائمة Data فيظهر مربع الحوار Go To Case كما هو بالشكل ثم اكتب رقم الحالة التي تريد الانتقال إليها



✓ البحث عن القيم Finding Values

1. إذا رغبت في البحث عن قيم لمتغيرات معينة (مثلا المتغير q3) انقر فوق أي خلية في المتغير q1.
2. من القائمة Edit اختر Find فيظهر مربع الحوار التالي:



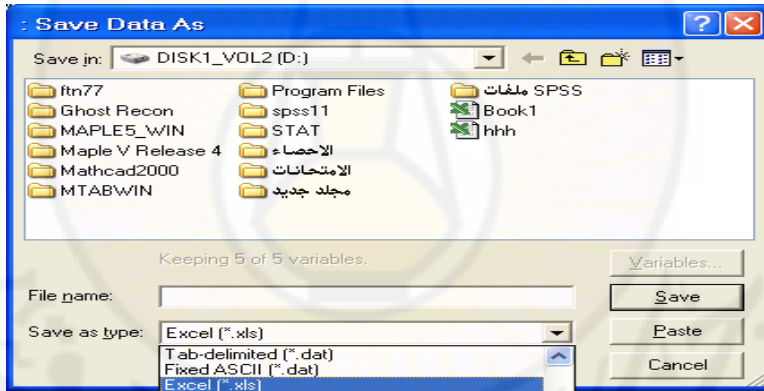
3. اكتب الرقم المراد البحث عنه وليكن 4 في المستطيل أمام Find what

❑ إستيراد وتصدير البيانات Exporting and Importing

تعتبر عملية الحصول على البيانات من الأولويات التي تشغل بال الباحثين، ولكن ليس بالضرورة أن تكون هذه البيانات مخزنة في ملفات SPSS إذ قد تكون ضمن برنامج Excel أو Access وغيرها (تسمى هذه العملية استيراد البيانات). كذلك فانك قد تحتاج في بعض الأحيان تخزين بياناتك التي قمت بمعالجتها في تطبيقات أخرى مثل Excel أو Access (تسمى هذه العملية تصدير البيانات).

✓ تصدير البيانات Exporting Data

1. إذا أردت تخزين ملف SPSS في برنامج Excel نختار من Save As القائمة File ليظهر مربع الحوار التالي:



2. من المربع Save as type نحدد نوع الملف Excel *.xls الذي يستطيع تطبيق Excel التعرف عليه، ثم اكتب اسم الملف "المخزون" في المستطيل أمام File name. ثم اضغط على زر Save.

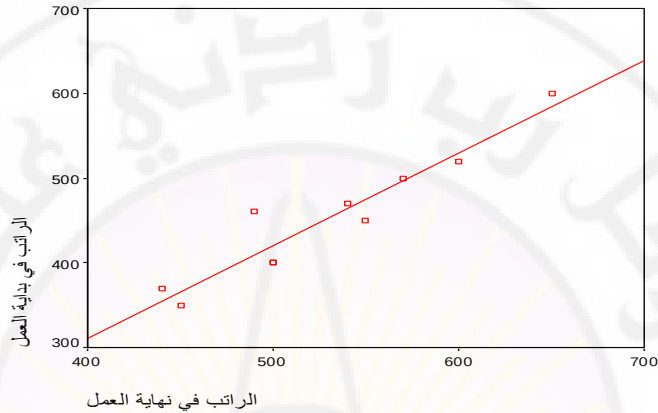
3. افتح تطبيق Excel ثم اضغط على Open من شريط القوائم وافتح الملف "المخزون".

✓ إستيراد البيانات Importing Data

نستطيع استيراد البيانات من تطبيق آخر مثل Excel وتحويله إلى تطبيق SPSS باتباع الخطوات التالية:

الملاحق

1. افتح برنامج جديد في SPSS ثم اختر Open من القائمة File ، ثم اضغط على السهم يمين القائمة File of Type ستظهر قائمة بأنواع الملفات التي يمكن لبرنامج SPSS التعامل معها، حدد على سبيل المثال Excel*.xls



2- حدد الملف الذي تريد فتحه بالنقر عليه، ثم اضغط Ok.

□ إيجاد المقاييس الإحصائية الرقمية للمتغيرات

2. المقاييس الإحصائية المراد إيجادها هي

1. مقاييس النزعة المركزية (Central Tendency)

- ✓ الوسط الحسابي mean مجموع القيم على عددها.
- ✓ الوسيط Median القيمة التي يقل عنها 50% من مفردات العينة.
- ✓ المنوال Mode القيمة الأكثر تكرارا.

2. مقاييس التشتت Dispersion

- ✓ الانحراف المعياري Slandered Deviation مقدار تشتت القيم عن وسطها الحسابي مقاسا بوحدات المتغير نفسها.
- ✓ التباين Variance مربع الانحراف المعياري

- ✓ المدى Range الفرق بين أكبر قيمة وأصغر قيمة.
- ✓ أقل قيمة Minimum
- ✓ أكبر قيمة Maximum
- ✓ الخطأ المعياري S.E.mean مقدار الخطأ الموجود في الوسط الحسابي وهو دلالة على دقة الوسط الحسابي كتقدير لوسط المجتمع.
- 3. شكل التوزيع Distribution

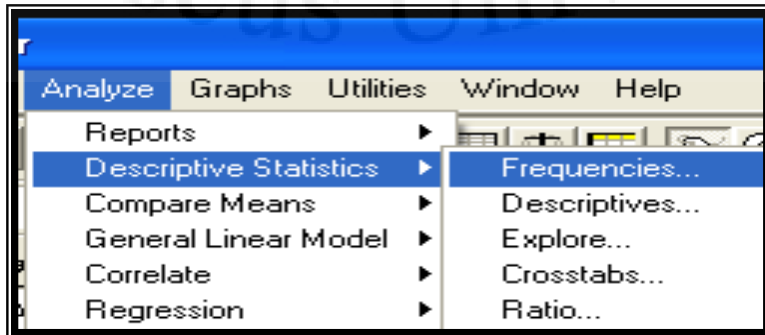
✓ الالتواء Skew ness : يعطى مقياس الالتواء فكرة عن تمركز قيم المتغير ، فإذا ما كانت قيم هذا المتغير تتمركز باتجاه القيم الصغيرة أكثر من تمركزها باتجاه القيم الكبيرة فإن توزيع هذا المتغير ملئ نحو اليمين ويسمى موجب الالتواء وتكون قيمة الالتواء موجبة. أما إذا كان العكس فإن هذا الالتواء يمين سالباً أو ملئ نحو اليسار وتمون قيمة الالتواء سالبة. أما إذا كانت قيمة معامل الالتواء صفراً فإن التوزيع يكون طبيعياً.

● التفلطح او التفرطح Kurtosis : يمثل تكرارات القيم على طرفي هذا المتغير و هو يمثل أيضاً درجة علو قمة التوزيع بالنسبة للتوزيع الطبيعي. فإذا كانت قيمة التفرطح كبيرة كانت للتوزيع قمة منخفضة، ويسمى التوزيع كبير التفلطح، إما إذا كانت قيمة التفلطح صغيرة فإن للتوزيع قمة عالية ويسمى التوزيع مدبباً أو قليل التفلطح.

- ✓ الربعيات Quartiles تقسيم البيانات إلى أربعة أرباع
- ✓ المئينات Percentile(s) تقسيم البيانات أجزاء من مائة

❖ لإيجاد المقاييس الإحصائية السابقة بالإضافة إلى بعض الرسوم البيانية التي تساعد على التوضيح تتبع الخطوات التالية:

- استخدام الخيار Frequencies
- 1. من شريط القوائم Analyze اختر Descriptive Statistics ومن

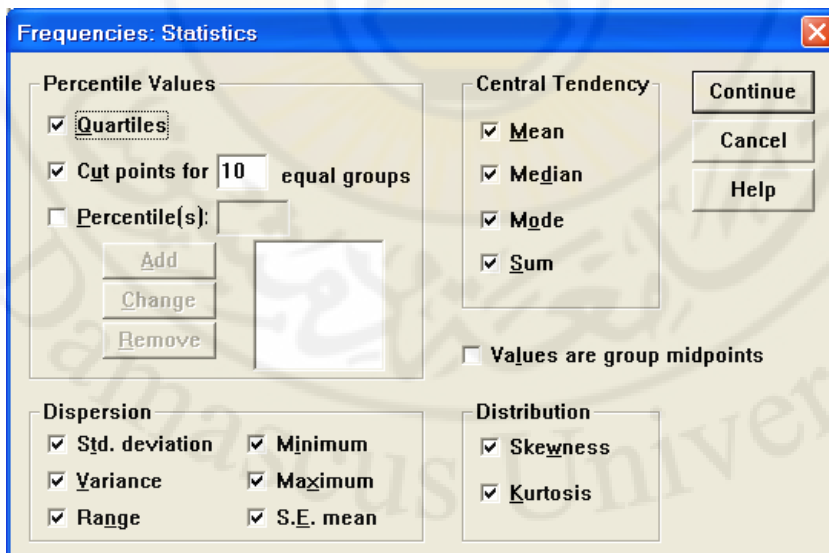


الملاحق

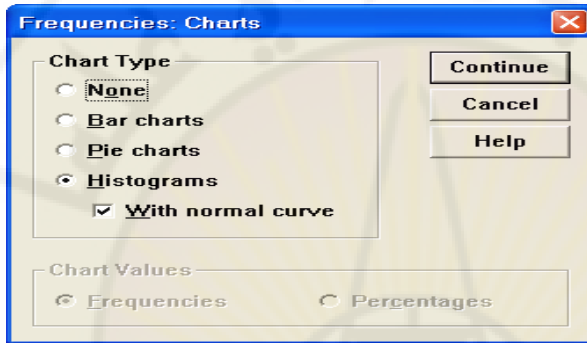
القائمة الفرعية اختر Frequencies كما هو موضح بالشكل ينتج مربع الحوار التالي:



2. اضغط على الزر Statistics يظهر مربع الحوار التالي:



3. اضغط على جميع الإحصاءات المطلوبة ، ثم اضغط على Continue
فانرجع إلى مربع الحوار السابق: اضغط على الزر Charts يظهر مربع
الحوار التالي:



4. اضغط على Histograms و داخل المربع With normal carve ثم
Continue نرجع لمربع الحوار Frequency اضغط على Ok تظهر
النتائج التالية:

Frequencies

Statistics

		الراتب في بداية العمل	الراتب في نهاية العمل
N	Valid	10	10
	Missing	0	0
Mean		\$452.00	\$529.00
Std. Error of Mean		\$23.981	\$20.894
Median		\$455.00	\$520.00
Mode		\$400	\$500
Std. Deviation		\$75.836	\$66.072
Variance		\$5,751.111	\$4,365.556
Skewness		.567	.435
Std. Error of Skewness		.687	.687
Kurtosis		.113	-.351
Std. Error of Kurtosis		1.334	1.334
Range		\$250	\$210
Minimum		\$350	\$440
Maximum		\$600	\$650
Sum		\$4,520	\$5,290
Percentiles	10	\$352.00	\$441.00
	20	\$376.00	\$458.00
	25	\$392.50	\$480.00
	30	\$400.00	\$493.00
	40	\$420.00	\$500.00
	50	\$455.00	\$520.00
	60	\$466.00	\$546.00
	70	\$491.00	\$564.00
	75	\$505.00	\$577.50
	80	\$516.00	\$594.00
	90	\$592.00	\$645.00

Frequency Table

الملاحق -----

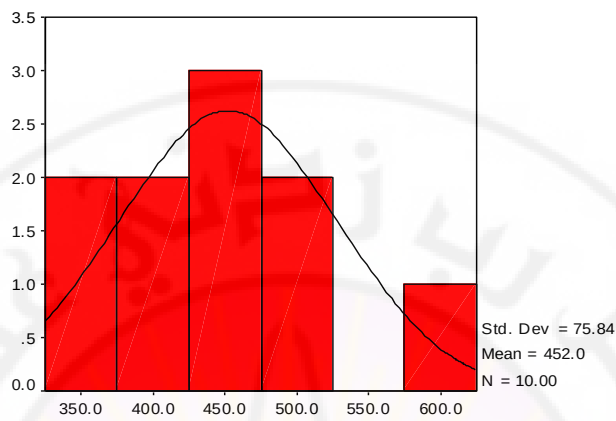
الراتب في بداية العمل

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	\$350	1	10.0	10.0	10.0
	\$370	1	10.0	10.0	20.0
	\$400	2	20.0	20.0	40.0
	\$450	1	10.0	10.0	50.0
	\$460	1	10.0	10.0	60.0
	\$470	1	10.0	10.0	70.0
	\$500	1	10.0	10.0	80.0
	\$520	1	10.0	10.0	90.0
	\$600	1	10.0	10.0	100.0
	Total	10	100.0	100.0	

الراتب في نهاية العمل

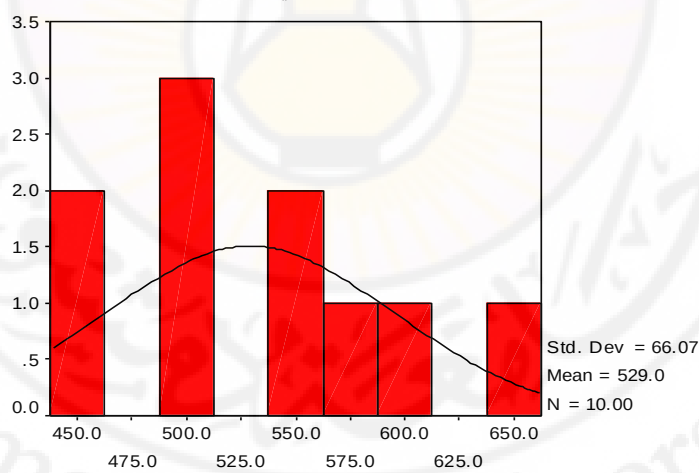
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	\$440	1	10.0	10.0	10.0
	\$450	1	10.0	10.0	20.0
	\$490	1	10.0	10.0	30.0
	\$500	2	20.0	20.0	50.0
	\$540	1	10.0	10.0	60.0
	\$550	1	10.0	10.0	70.0
	\$570	1	10.0	10.0	80.0
	\$600	1	10.0	10.0	90.0
	\$650	1	10.0	10.0	100.0
	Total	10	100.0	100.0	

الراتب في بداية العمل



الراتب في بداية العمل

الراتب في نهاية العمل

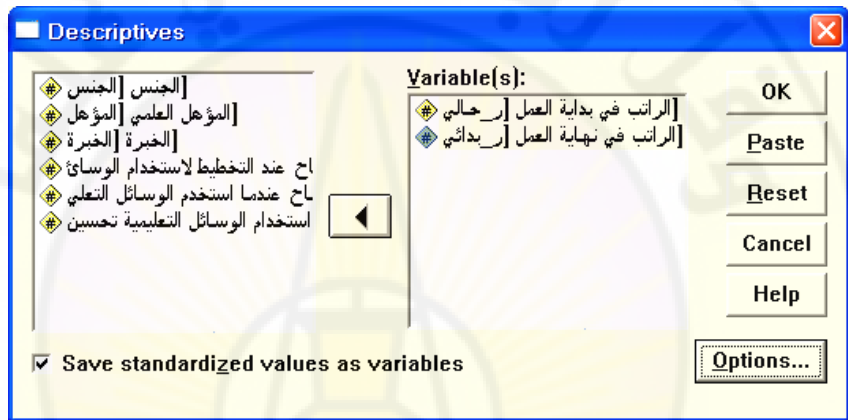


الراتب في نهاية العمل

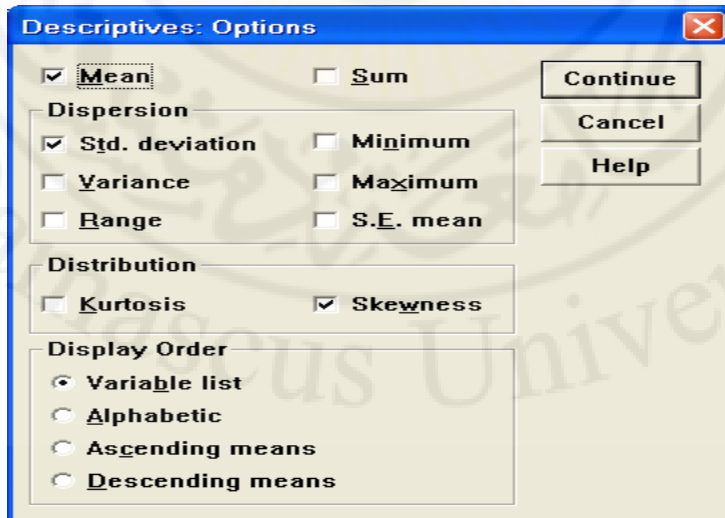
● استخدام الأمر Descriptive

يستخدم هذا الخيار لإيجاد بعض المقاييس الإحصائية أيضا ولعمل ذلك

1. نختار من القائمة Analyze الخيار Descriptive Statistics ومن القائمة الفرعية الخيار Descriptives يظهر مربع الحوار التالي:



2. ندخل المتغيرات " ر_حالي " و " ر_بدائي " داخل المستطيل أسفل



. Variable(s)

3. اضغط على الزر Option ليظهر مربع الحوار التالي:

5. اختر المقاييس المطلوبة ، ثم اضغط على Continue لنعود لمربع الحوار Descriptives .

6. اضغط داخل المربع بجانب Save standardized values as variables (ليحول البيانات إلى قيم معيارية) ثم اضغط Ok تظهر النتائج التالية:

Descriptive Statistics

	N	Mean	Std.	Skewness	
	Statistic	Statistic	Statistic	Statistic	Std. Error
الراتب في بداية العمل	10	\$452.00	\$75.836	.567	.687
الراتب في نهاية العمل	10	\$529.00	\$66.072	.435	.687
Valid N (listwise)	10				

7. يتم حساب القيم المعيارية وفق العلاقة $z = \frac{x - \mu}{\sigma}$ حيث x هي القيمة

المدخلة ، μ هي المتوسط الحسابي للمتغير x ، σ هو الانحراف المعياري للمتغير x والقيم المعيارية تظهر عادة في نهاية ملف البيانات وهذا جزء من ملف البيانات كما يلي:

	q1	q2	q3	ر_حالي	ر_بدائي	ز_حالي	ز_بدائي
1	موافق	موافق بشدة	محايد	\$400	\$500	-0.68569	-0.43891
2	محايد	موافق	معارض	\$350	\$450	-1.34501	-1.19566
3	موافق	موافق بشدة	موافق بشدة	\$370	\$440	-1.08128	-1.34701
4	موافق بشدة	موافق	موافق	\$400	\$500	-0.68569	-0.43891
5	موافق بشدة	موافق	موافق	\$500	\$570	0.63294	0.62053
6	موافق	موافق	موافق بشدة	\$450	\$550	-0.02637	0.31783
7	محايد	محايد	محايد	\$460	\$490	0.10549	-0.59026
8	موافق بشدة	موافق بشدة	موافق بشدة	\$470	\$540	0.23735	0.16648
9	معارض	معارض	موافق	\$520	\$600	0.89667	1.07458
10	موافق بشدة	موافق بشدة	موافق بشدة	\$600	\$650	1.95158	1.83133

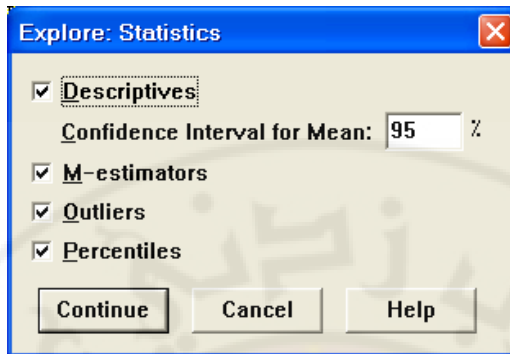
● استخدام الأمر Explore (مستكشف البيانات)

يستخدم هذا الخيار لإيجاد بعض المقاييس الإحصائية لمتغير أو أكثر وفقا لتصنيف متغير آخر أو أكثر ، وكذلك نحصل منه على بعض الرسوم البيانية وعملية تلخيص البيانات وغيرها وللتعرف عليه نتبع ما يلي:

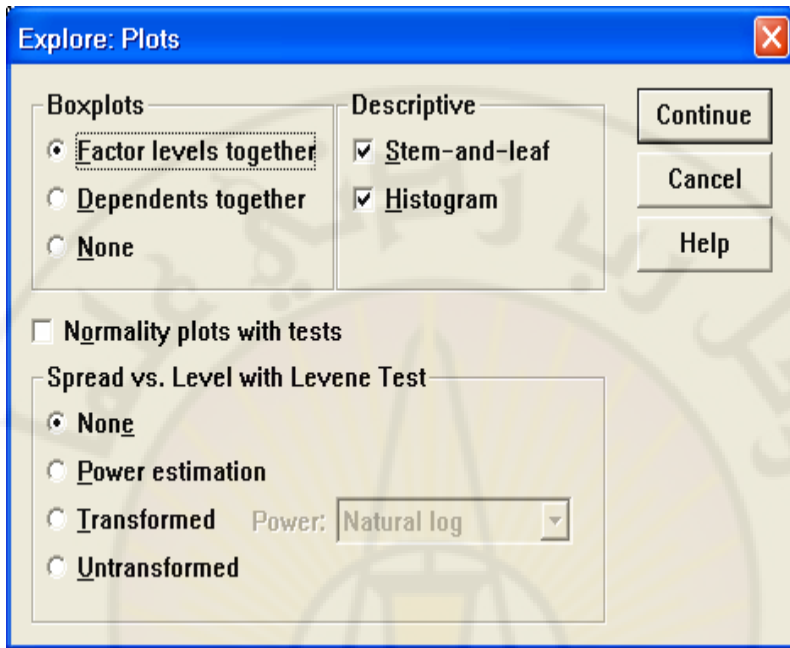
1. من القائمة Analyze نختار Descriptive Statistics ومن القائمة الفرعية نختار Explore يظهر مربع الحوار التالي:



2. ندخل المتغير " ر_حالي " في المستطيل أسفل Dependent List والمتغير "الجنس" في المستطيل أسفل Factor List (لاحظ وجود عدة خيارات داخل المستطيل Display وهي Both و Statistics و Plots وهي تعني اختيار الإحصاءات أو الرسم البياني أو كليهما ، سوف نختار كليهما Both) ثم اضغط على Statistics ليظهر مربع الحوار التالي:



3. اختر Descriptive (الإحصاءات الوصفية) و M-Estimators (تقدير لمقاييس النزعة المركزية التي لا تتأثر بالقيم المتطرفة أو الشاذة) و Outliers (تحديد ما إذا كانت هناك قيم شاذة واستخراج أكبر خمس قيم وأقل خمس قيم شاذة، وذلك تمهيدا لحذفها من البيانات حتى لا تؤثر على الاختبارات الأخرى و اختر كذلك الخيار Percentiles (وتعني المئينات) ثم اضغط على Continue لتعود لمربع الحوار Explore.
4. اضغط على الزر Plots ليظهر مربع الحوار التالي:



5. اضغط على Factor level together و من المستطيل
Descriptive اختر Stem-and-leaf و Histogram ثم اضغط
على Continue لنعود مرة ثانية لمربع الحوار Explore ، اضغط
Ok لتظهر النتائج التالية:

تفسير النتائج:

الجدول التالي: يظهر عدد ونسبة القيم المدخلة والمفقودة لكلا الجنسين وذلك
لمتغير الجنس.

----- الملاحق

الجنس

Case Processing Summary

الجنس	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
ذكر الراتب في بداية العمل	6	100.0%	0	.0%	6	100.0%
انثى	4	100.0%	0	.0%	4	100.0%

الجدول التالي: يظهر بعض المقاييس الإحصائية الجديدة مثل.

95% Confidence interval for mean وهي تعني

فترة الثقة للوسط الحسابي بنسبة دقة 95% ولها حد أدنى وحد أعلى وذلك لكل من الذكور والإناث كل على حدة.

5% Trimmed Mean

وهو الوسط الحسابي الذي يتم حسابه بعد استبعاد أكبر 5% وأصغر 5% حتى يتم استبعاد القيم الشاذة.

Interquartile Range

تمثل المدى الربيعي وهو الفرق بين قيمتي الربع الثالث والربع الأول.

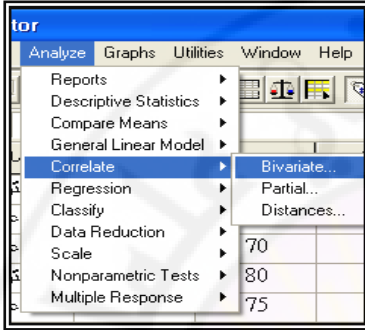
لاحظ أن باقي الإحصاءات قد تم شرحها سابقا.

Descriptives

الجنس			Statistic	Std. Error
ذكر الراتب في بداية العمل	Mean		\$488.33	\$28.097
	95% Confidence Interval for Mean	Lower Bound	\$416.11	
		Upper Bound	\$560.56	
	5% Trimmed Mean		\$487.04	
	Median		\$480.00	
	Variance		4736.667	
	Std. Deviation		\$68.823	
	Minimum		\$400	
	Maximum		\$600	
	Range		\$200	
	Interquartile Range		\$102.50	
	Skewness		.605	.845
	Kurtosis		.620	1.741
انثى	Mean		\$397.50	\$26.260
	95% Confidence Interval for Mean	Lower Bound	\$313.93	
		Upper Bound	\$481.07	
	5% Trimmed Mean		\$396.11	
	Median		\$385.00	
	Variance		2758.333	
	Std. Deviation		\$52.520	
	Minimum		\$350	
	Maximum		\$470	
	Range		\$120	
	Interquartile Range		\$97.50	
	Skewness		1.165	1.014
	Kurtosis		1.085	2.619

الإرتباط والإنحدار

لإيجاد معامل الارتباط بين كل درجة الطالب في الرياضيات والإحصاء أو بمعنى



آخر اختبر الفرضية التي تقول

بأنه لا يوجد ارتباط بين علامة

الرياضيات وعلامة الإحصاء "

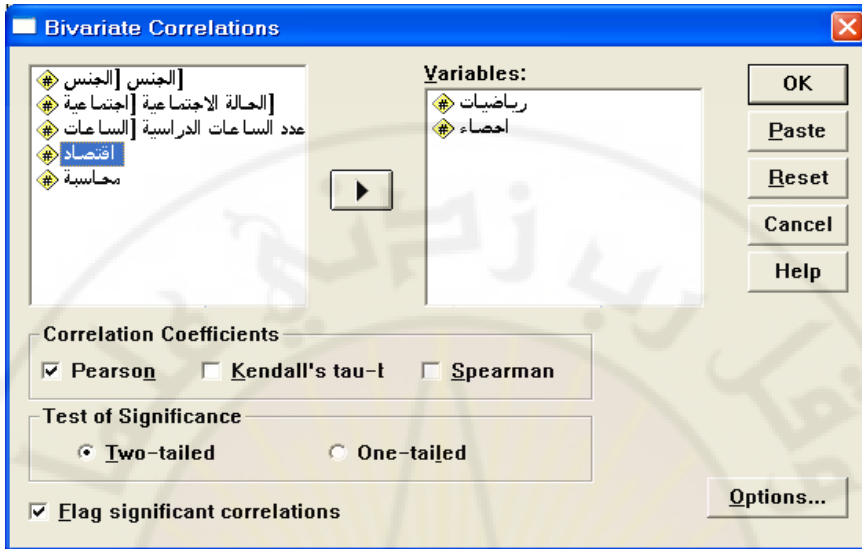
تسمى هذه الفرضية الصفرية

اتبع الخطوات التالية:

1. من القائمة Analyze اختر Correlate ومن القائمة الفرعية

اختر Bivariate كما تلاحظ بالشكل المقابل:

يظهر مربع الحوار التالي:



2. ادخل المتغيرين " رياضيات " و " إحصاء " داخل المستطيل Variables

3. لاحظ أن اختيار معامل ارتباط بيرسون هو المختار في الأصل وإذا أردت اختيار مقياس آخر لمعامل الارتباط عليك أن تضغط في المربع الذي بجانبه، كذلك لاحظ أن المربع بجانب Flag significant correlations مفعّل أي موجود بداخله إشارة "صح" وفائدته وضع نجمة أو نجمتين على المتغيرات الذي لها معامل ارتباط مقبول أي عرض مستوى الدلالة .

1. اضغط Ok نحصل على النتائج التالية:

Correlations

		رياضيات	احصاء
رياضيات	Pearson Correlation	1	.959**
	Sig. (2-tailed)	.	.000
	N	10	10
احصاء	Pearson Correlation	.959**	1
	Sig. (2-tailed)	.000	.
	N	10	10

** . Correlation is significant at the 0.01 level

2. نلاحظ من النتائج الواردة في مصفوفة المعاملات أن
 $2\text{-tailed Significance} = 0.000$ وهو اقل من $\alpha = 0.05$ وهذا يدل على أن
 هناك ارتباط قوي بين علامات الرياضيات والفيزياء ويساوي $r = 0.959$ أي
 علينا رفض الفرضية الصفرية.

✓ إيجاد مصفوفة معاملات الارتباط

مصفوفة معاملات الارتباط هي مصفوفة يتم فيها عرض معاملات الارتباط بين كل
 زوجين من المتغيرات ولإيجاد ذلك، ادخل جميع المتغيرات داخل مستطيل



الملاحق -----

Variables فى مربع الحوار Bivariate Correlations كما فى الشكل التالى:

اضغط على Ok تظهر النتائج التالية:

Correlations

		رياضيات	احصاء	اقتصاد	محاسبة
رياضيات	Pearson Correlation	1	.959**	.780**	.833**
	Sig. (2-tailed)	.	.000	.008	.003
	N	10	10	10	10
احصاء	Pearson Correlation	.959**	1	.746*	.811**
	Sig. (2-tailed)	.000	.	.013	.004
	N	10	10	10	10
اقتصاد	Pearson Correlation	.780**	.746*	1	.890**
	Sig. (2-tailed)	.008	.013	.	.001
	N	10	10	10	10
محاسبة	Pearson Correlation	.833**	.811**	.890**	1
	Sig. (2-tailed)	.003	.004	.001	.
	N	10	10	10	10

** . Correlation is significant at the 0.01 level (2-tailed).

* . Correlation is significant at the 0.05 level (2-tailed).

من مصفوفة معاملات الارتباط نجد انه توجد علاقة ارتباط قوي بين كل متغيرين بعضها

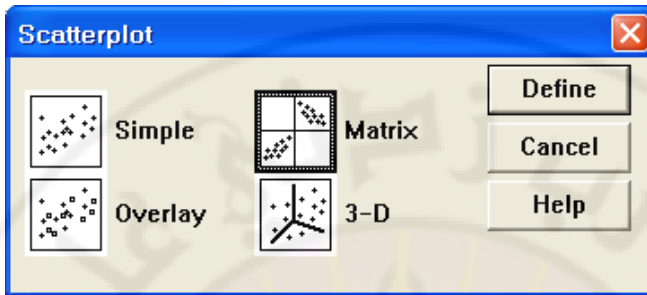
عند مستوى دلالة $\alpha=0.01$ وبعضها الآخر عند مستوى معنوية $\alpha=0.05$

ولتمثيل النتائج باستخدام لوحة الانتشار Scatter Plot لتمثيل شكل وقوة العلاقة بين

متغيرين كميين بيانيا نتبع الخطوات التالية:

الملاحق -----

1. من قائمة Graphs نختار Scatter سيظهر لنا مربع الحوار Scatterplot

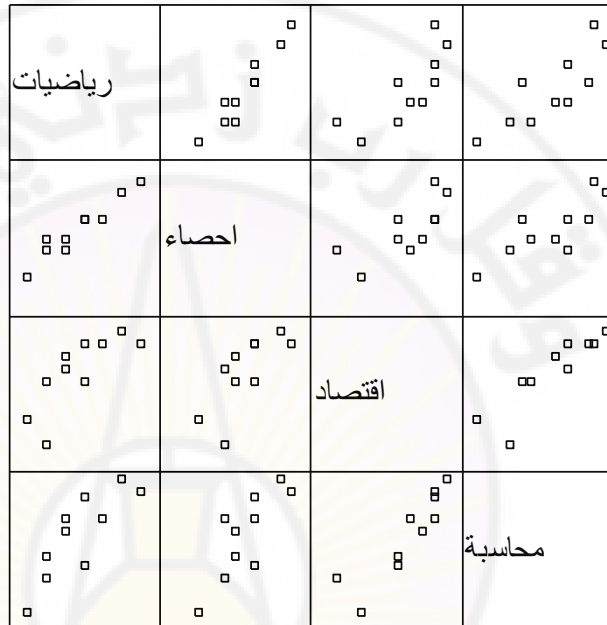


المبين بالشكل التالي:

2. اضغط على Matrix ثم على Define سيظهر مربع الحوار Matrix كما يلي:



3. ادخل المتغيرات في المستطيل Matrix Variables ثم اضغط Ok ستظهر النتائج التالية:



تحليل الانحدار الخطي

يستخدم تحليل الانحدار الخطي للتنبؤ بقيمة متغير، يسمى المتغير التابع، من خلال مجموعة من المتغيرات المستقلة، وذلك من خلال تمثيل العلاقة بين المتغير التابع والمتغيرات المستقلة :

$$Y = a x + b z + c w + \dots + d + \text{خطأ}$$

حيث المتغير Y هو المتغير التابع، والمتغيرات $X, Z, W, \dots$ متغيرات مستقلة، و $a, b, c, \dots, d$ ثوابت .

الملاحق -----

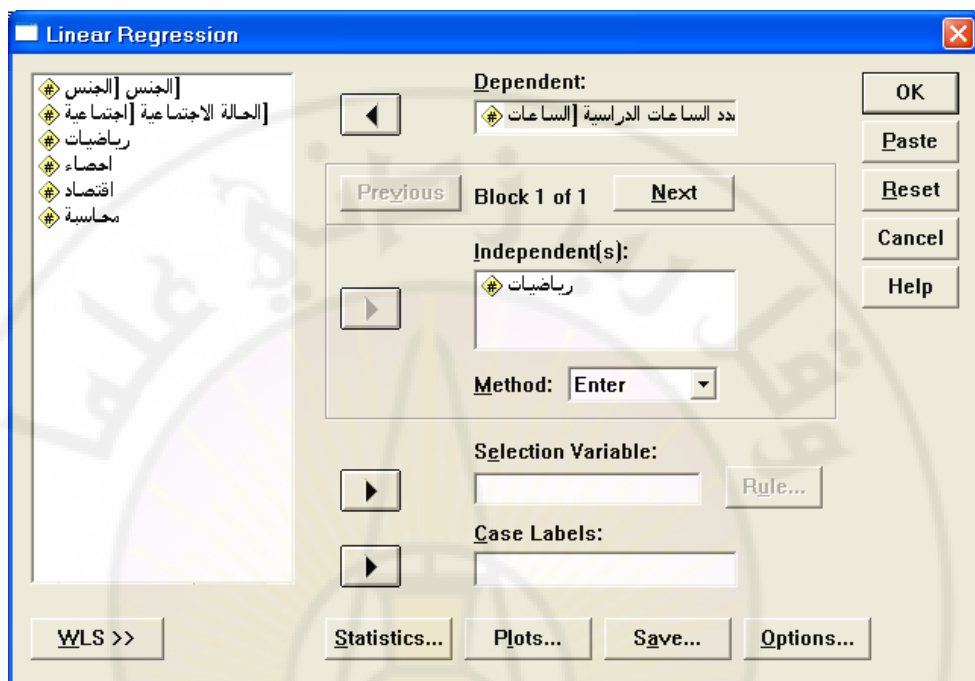
وتحليل الانحدار يسمى ثنائيا إذا كان هناك متغيرين فقط الأول متغير مستقل والآخر متغير تابع، أما إذا كان هناك عدة متغيرات مستقلة ومتغير تابع واحد سمي تحليل الانحدار بتحليل الانحدار المتعدد.

أولاً: تحليل الانحدار الثنائي:

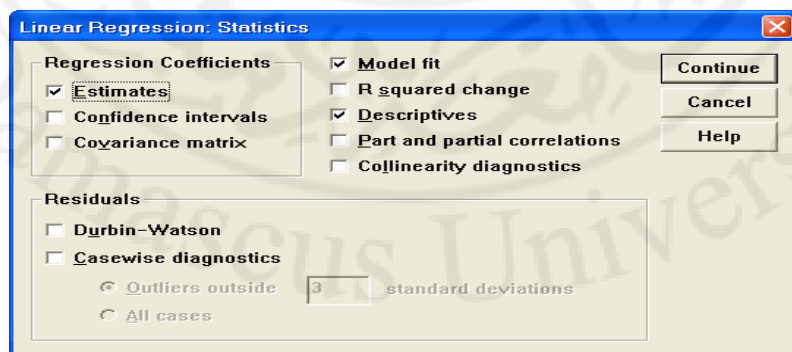
مثال: ما أثر ساعات الدراسة على التحصيل الدراسي للطالب في مادة الرياضيات؟

للجواب عن هذا السؤال نجري تحليل الانحدار الثنائي التالي:

1. من القائمة Analyze نختار Regression ثم من القائمة الفرعية نختار Linear سيظهر مربع الحوار التالي:



2. انقل المتغير " رياضيات " إلى المستطيل (s) Independent والمتغير " الساعات " داخل المستطيل Dependent . ثم اضغط على Statistics يظهر مربع الحوار التالي:

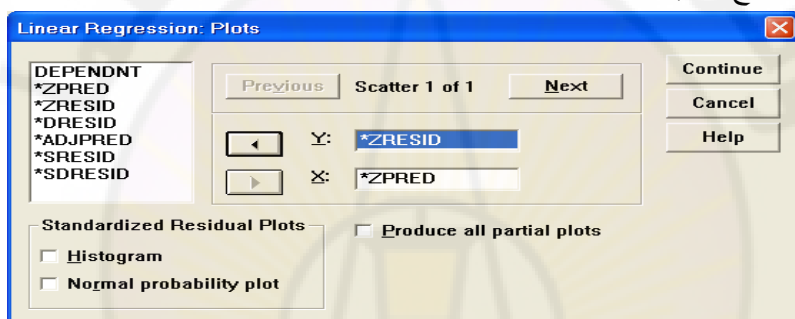


----- الملاحق

4. اضغط على Estimates, Model fit, Descriptives ، ثم اضغط على Continue نعود إلى مربع الحوار الأصلي Linear Regression

5. اضغط على Plots لعمل لوحة انتشار Scatter plot لأخطاء التقدير Residuals والقيم المتنبأ بها Predicted values سيظهر مربع الحوار Linear Regression: Plots التالي:

6. انقل المتغير ZRESID إلى مستطيل Y والمتغير ZPRED إلى المستطيل X ثم اضغط Continue سنعود إلى مربع الحوار الأصلي، اضغط Ok تظهر النتائج التالية:



تحليل النتائج: الجدول التالي يبين المتوسطات للمتغيرات المدخلة وكذلك الانحراف المعياري وعدد المفردات في كل متغير.

Regression

Descriptive Statistics

	Mean	Std. Deviation	N
عدد الساعات الدراسية	4.50	2.014	10
رياضيات	78.50	9.443	10

الجدول التالي هو مصفوفة معامل الارتباط بين المتغير المستقل والمتغير التابع وتساوي 0.949 وهو ارتباط قوي جدا أي كلما زاد عدد ساعات الدراسة زادت تحصيل الطالب.

Enter الجدول التالي يبين المتغيرات المدخلة والنموذج المستخدم وهو نموذج وسيأتي شرحه لاحقا

Variables Entered/Removed^a

Model	Variables Entered	Variables Removed	Method
1	رياضيات	.	Enter

a. All requested variables entered.

b. Dependent Variable: تقديرنا تا علل ا ددع

Correlations

		عدد الساعات الدراسية	رياضيات
Pearson Correlation	عدد الساعات الدراسية	1.000	.949
	رياضيات	.949	1.000
Sig. (1-tailed)	عدد الساعات الدراسية	.	.000
	رياضيات	.000	.
N	عدد الساعات الدراسية	10	10
	رياضيات	10	10

----- الملاحق

الجدول التالي يبين معامل الارتباط R ومعامل التحديد R^2 وتساوي 0.902 وهي مرتفعة وهذا يدل على أن معادلة الانحدار أو التنبؤ جيدة.

Model Summary^a

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.949 ^a	.902	.889	.670

a. Predictors: (Constant), تخطيط

b. Dependent Variable: تغييرات معدل ددع

الجدول التالي هو جدول تحليل التباين ويوضح المتغير المستقل هو الرياضيات والمتغير التابع هو عدد الساعات وقد كانت قيمة $\text{Sig.} = 0.000$ وهي أقل من 0.05 وهذا يعني قبول معادلة الانحدار

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	32.905	1	32.905	73.224	.000 ^a
	Residual	3.595	8	.449		
	Total	36.500	9			

a. Predictors: (Constant), تخطيط

b. Dependent Variable: تغييرات معدل ددع

الجدول التالي يسمى جدول المعاملات ويحتوي على:

----- الملاحق

-معاملات المتغيرات التي دخلت المعادلة الموجودة في العمود B .

-الخطأ المعياري لكل عمود في عمود Std.Error.

-معاملات المتغيرات المستقلة التي دخلت المعادلة بعد تحويلها إلى علامات معيارية Standardization والموجودة في عمود Beta المقابلة لكل متغير، وفي العمودين الأخيرين من هذا الجدول تظهر قيمة الإحصائي t ومستوى الدلالة الخاصتين باختبار دلالة قيمة Beta فإذا كانت قيمة Sig. المقابلة لأي من قيم Beta أقل من 0.05 فهذا يعني أن المتغير المقابل لهذه القيم له اثر كبير ذو دلالة إحصائية.ومن خلال هذا الجدول يمكن كتابة معادلة الانحدار أو التنبؤ التالية:

$$\text{عدد ساعات الدراسة} = 0.202 \times \text{علامة الرياضيات} - 11.396$$

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-11.396	1.870		-6.095	.000
	رياضيات	.202	.024	.949	8.557	.000

a. Dependent Variable: تحصيلات اعلیٰ ددع

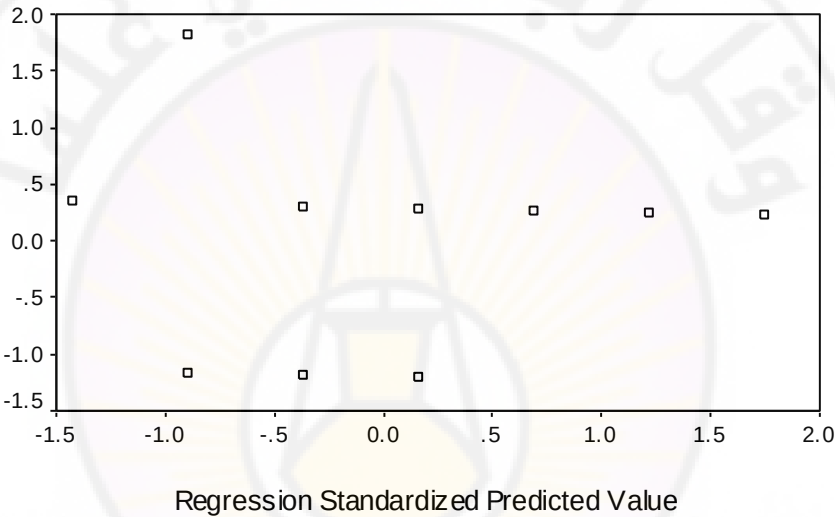
- في المخطط التالي (لوحة الانتشار)بين القيم المتنبأ بها Predicted values وأخطاء التقدير Residual values ، نلاحظ شكل الانتشار عشوائيا وهذا يدل على أن العلاقة بين المتغيرين خطية وأن شروط تحليل الانحدار متوفرة، ولكن إذا ظهرت في امثل أخرى أن نمط شكل الانتشار يشبه شكل الدالة التربيعية أو التكعيبية أو غيرها فهذا دليل أن على أن العلاقة بين المتغيرين غير خطية.

Charts

Scatterplot

Dependent Variable: تقدير لارتفاع الدرع

لوحة انتشار القيم المعيارية للقيم المتنبأ بها مع القيم المعيارية للخطأ



رياضيات			
احصاء			
اقتصاد			
محاسبة			

