الجامعة المستنصرية
كلية التربية الأساسية
قسم الرياضيات/م٢

محاضرات في الإحصاء التربوي

إعداد

ا.م. بيداء محمد احمد

2021 - 2022

المفردات

الفصل الأول : الاحصاء

تعريف الاحصاء ، القياس ، المجتمع ، العينة، الرموز الاحصائية .

أنواع الاحصاء : الاحصاء الوصفي ، الاحصاء الاستدلالي .

طرق عرض البيانات :

- **أنواع لبيانات : البيانات النوعية ، البيانات الكمية**

أ- العرض الجدولي (جدول التوزيع التكراري).

ب- التمثيل البياني : المدرج التكراري ، المضلع التكراري ، المنحني التكراري .

الفصل الثاني: المقاييس الاحصائية (مقاييس النزعة المركزية)

الوسط الحسابي (المتوسط) ، الوسيط ، المنوال .

الفصل الثالث : مقاييس التشتت

- **أنواع مقاييس التشتت**

الانحراف الوسطي ، الانحراف المعياري ، التباين .

الفصل الرابع : مقاييس الارتباط

أنواع معاملات الارتباط :

أ- معامل ارتباط بيرسون ب- معامل ارتباط سبيرمان

الفصل الاول

الاحصاء Statistics

الاحصاء Statistics: يقصد بالإحصاء العد أو التعداد أو عدد الأشياء أو جمع بيانات عنها .

و**علم الإحصاء** "هو ذلك العلم الذي يعمل على استخدام الاسلوب العلمي في طرق جمع البيانات وتبويبها وتلخيصها وعرضها وتحليلها بهدف الوصول الى استنتاجات وقرارات مناسبة ".

أهمية الاحصاء في التربية :

يلعب الاحصاء دوراً كبيراً في مجال التربية وعلم النفس ، إذ أصبح من الضروري لكل من له علاقة بهذين المجالين أن يكون ملماً بالطرق الاحصائية سواء كان معلماً أو ادارياً أو مخططاً ...الخ،

لذا يمكن ايجاز أهمية دراسة الاحصاء بالنقاط الآتية :

١) تساعد الطرق الإحصائية المختلفة على وصف الظواهر النفسية والتربوية وصفاً دقيقاً.
٢) تمكن الباحث من أن يكون دقيقاً ومحدداً في خطوات تفكيره لحل المشكلات.
٣) تساعد على تلخيص نتائج البحوث بطريقة سهلة ومفيدة.
٤) يمكن الوصول الى نتائج يمكن الإستفادة منها وتعميمها .
٥) تساعد على التنبؤ بالظواهر المختلفة وعلى معرفة امكانية حدوث مثل هذه الظواهر ومقدار وشروط حدوثها وكيفية تعديل مواعيد حدوثها.

أما أهميته بالنسبة للمعلم فتتمثل في:

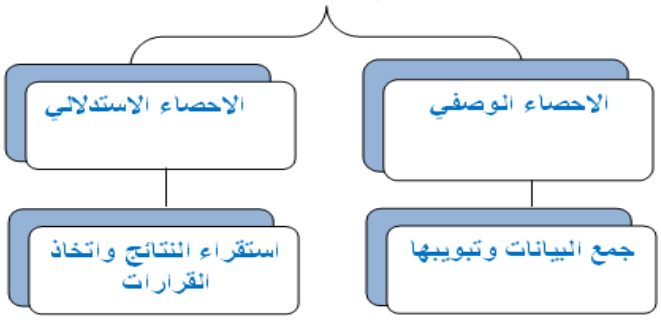
١. يساعده في عرض نتائج تلاميذه بشكل بيانات مبوبة .
٢. يساعده في التعرف على نسبة النجاح و الرسوب .
٣. يساعده في ترتيب تلاميذه بحسب درجاتهم .
٤. يساعده في المقارنة بين الصفوف والشعب .
٥. يساعده في فهم وتفسير النتائج للبحوث التربوية والنفسية .
٦. يساعده في اجراء البحوث التربوية والنفسية .

الخطوات المنهجية للتحليل الاحصائي في البحث العلمي :

١) **جمع البيانات :** هو عملية الحصول على المعلومات أو قيم المشاهدات أو القياسات للتجارب التي يجريها الإحصائي .
٢) **تنظيم البيانات وعرضها :** هي عملية وضع المعلومات في جداول هندسية وعرضها بطريقة مناسبة كالأشكال الهندسية والرسوم والبيانية وغيرها.
٣) **تحليل البيانات :** هي عملية ايجاد القيم لمقاييس و اقترانات معينة تحدد قيمها من البيانات موضع الدراسة .

٤) **استقراء النتائج واتخاذ القرارات :** وهي الاستنتاجات التي يصل إليها الباحث وتكون علل شكل تقديرات أو تنبؤات أو تعميمات أو قرار برفض أو قبول الفرضيات الإحصائية .

اقسام علم الإحصاء



يقسم علم الاحصاء إلى قسمين:

١. الاحصاء الوصفي Descriptive :

ويتضمن هذا النوع الطرق والأساليب المستخدمة في جمع البيانات والمعلومات عن ظاهرة معينة أو مجموعة ظواهر وكيفية تنظيم وتصنيف وتبويب هذه البيانات مع امكانية عرضها في جداول ورسومات بيانية ، وحساب بعض المؤشرات الاحصائية منها .

- ملخص جيد لمجموعة كبيرة من المعلومات و البيانات .
- أهم صور التصنيف جداول التوزيع التكراري و الرسوم البيانية التي تعبر عن هذا التوزيع .
- أما التلخيص فيتخذ ثلاثة صور هي :
  * النزعة المركزية " المتوسط ، الوسط المحرج ، الوسيط ، المنوال "
  * التشتت " الانحراف المعياري ، الانحراف المتوسط / التباين ، معامل الالتواء ".
  * العلاقة أو الارتباط و الانحدار .

٢. الاحصاء الاستدلالي Inferential :

وهو الشطر الآخر من علم الاحصاء الذي يشمل الطرق الإحصائية التي تهدف الى عمل استنتاجات او استدلالات حول المصدر الذي جمعت منه البيانات .

- يعد مكملاً للإحصاء الوصفي ، إذ يقدم الاحصاء الوصفي خصائص الظاهرة ومواصفاتها ، في حين يؤكد الإحصاء الاستدلالي مدى صحة حدوث الظاهرة .
- يهتم الإحصاء الاستدلالي بالتحليل والتفسير والاستنتاج واختيار الفروض واتخاذ القرارات بالاعتماد على نتائج العينات التي تمثل جزءاً مسحوباً من المجتمع .

وتضم فرعين هما :

أ- **التقدير :** ويهتم بإيجاد قيم تقديرية للاستدلال منها على القيم الحقيقية لمصدر جمع البيانات .

ب- **اختبار الفرضيات :** ويتضمن اختبار الفرضيات التي توضع كتفسير أولي للظاهرة المراد دراستها للوصول منها الى قرار بقبولها أو رفضها .

المفاهيم الأساسية في دراسة الاحصاء

طبيعة البيانات الاحصائية : عند جمع بيانات حول ظاهرة ما فإننا نرمز للظاهرة بالرمز **(X)** ونرمز لكل مفردة أو مشاهدة منها بالرمز (X_i) .

المتغير Variable: هو أي ظاهرة تظهر اختلافات بين مفرداتها ويرمز له بالرمز **X** أو أي رمز آخر مثل Y , Z .... الخ .

وتقسم المتغيرات الى نوعين :

١- متغيرات وصفية أو نوعية Variables Qualitative :

وهي تلك الظواهر أو الصفات التي لا يمكن قياسها مباشرة بالأرقام العددية ، ولا يمكن اجراء العمليات الحسابية عليها . مثل : صفة لون العين (أزرق ، أسود ، بني ، ...) ، والحالة الاجتماعية (غني ، متوسط الحال ، فقير ،...) ، وغيرها من الظواهر .

٢- المتغيرات الكمية Quantitative Variables :

وهي تلك الظواهر أو الصفات التي يمكن قياسها مباشرة بالأرقام العددية ، ويمكن اجراء العمليات الحسابية عليها . مثل : صفة الطول ، الوزن ، كمية المحاصيل ، .... الخ .

والمتغيرات الكمية تنقسم الى قسمين:

أ- **المتغيرات الكمية المتصلة (مستمرة) Continuous Variables :**
وهي المتغيرات التي تأخذ المشاهدة أو الظاهرة فيها أية قيمة رقمية في مدى معين سواء كانت أعداد صحيحة أو كسرية .

مثلاً فرضنا أن أطوال طلبة يتراوح بين 150 - 190 سم فهذا يعني :

$$150 \leq x \leq 190$$

ومن أمثلته أيضاً : الوزن ، كمية المحاصيل الزراعية ، درجات الحرارة ، الزمن ، ... الخ.

ب- **المتغيرات الكمية المنفصلة (غير مستمرة ، متقطعة) Discrete Variables :**
وهي المتغيرات التي تأخذ المشاهدة أو الظاهرة فيها قيماً متباعدة أو متقطعة غير مستمرة . أي تكون منفصلة عن بعضها ، مثل غرف المنزل .

مثلاً إذا كان عدد طلبة في ٤ صفوف هو ٢٥ ، ٢٧ ، ٣٠ ، ٣٢ فإن قيم المتغير X هي

$$X = 25 , 27 , 30 , 32$$

وأيضاً عند رمي حجر الزهر نرد نجد إن النتيجة تكون أحد الوجوه الآتية ١، ٢ ، ٣ ، ٤ ، ٥ ، ٦ أي أن : $X = 1 , 2 , 3 , 4 , 5 , 6$

ومن أمثلته أيضاً : عدد الطلبة ، عدد الثمار في الاشجار ، عدد أفراد العائلة .... الخ .

أي أنها تكون أعداد صحيحة فقط بدون كسور .

مثال : حدد نوع المتغيرات الآتية

نوع المتغير	وصفي	كمي متصل	كمي منفصل
درجة الذكاء			
نوع الكلية			
رقم الهاتف			
الدخل الشهري للأسرة			
فصيلة الدم			
عدد أطفال الأسرة			
لون البشرة			
الجنس			

مثال : إذا كانت أعمار (٥) طلاب هي : ٢٠ ، ١٩ ، ٢١ ، ٢٢ ، ٢٤ سوف تكتب

$$X = 20 , 19 , 21 , 22 , 24$$

أي أن

$X_1 = 20$ القيمة الأولى للمتغير X

$X_2 = 19$ القيمة الثانية للمتغير X

$X_5 = 24$ القيمة الخامسة للمتغير X

القياس Measurment : هو عملية تحديد درجة امتلاك الفرد سمة معينة . مثلاً نقول أن الطالب حصل على ٧٢% في مادة الاحصاء .

المجتمع Population : هو جميع القيم أو المفردات التي يمكن أن يأخذها المتغير .

مثل : درجات طلبة في كلية ما فإن المجتمع هو درجات جميع الطلبة في تلك الكلية .

والمجتمع أما أن يكون :

أ- **مجتمع غير محدد أو غير منتهي Finite Population :**
أي يمكن حصر عدد أفراده ، مثل : عدد أفراد عائلة ، عدد الطلبة الناجحين ، .... الخ .

ب- **مجتمع غير محدد أو غير منتهي Finite Population :**
وهو المجتمع الذي من الصعب أو المستحيل حصر عدد أفراده ، مثل : عدد الأسماك في النهر ، عدد الطيور ، عدد الميكروبات في الجو .... الخ .

العينة Sample هي مجموعة من المشاهدات أو المفردات اختيرت بطريقة ما من المجتمع ، وهي مجموعة جزئية منه .

وإن سبب اختيار العينة من المجتمع هو صعوبة دراسة المجتمع ككل إذ إنه قد يتطلب وقت وجهد كبيرين . ويجب أن تكون العينة ممثلة للمجتمع أي أنها تكون ممثلة لصفات المجتمع الأصلي ويمكن استنتاج خواصه منها .

أنواع العينات :

يوجد نوعان من العينات في ضوء طريقة اختيارها وهي :

أ- **العينة القصدية :**
ويتم سحبها بطريقة ليست عشوائية (قصدية) وبحسب غرض الباحث ، وتستخدم في الحالات التي منها الحصول على تقديرات تقريبية لتكوين فكرة سريعة عن مشكلة معينة
أو لأختيار الاستمارة الاحصائية للتأكد من صلاحيتها .

ب- **العينة العشوائية :**
وتعني الاختيار العشوائي وإتاحة الفرصة أمام جميع أفراد المجتمع ليظهروا في العينة.

المؤشر:

وهو قيمة تصف المجتمع مثل المتوسط الحسابي والانحراف المعياري للمجتمع ... الخ .

الرموز الاحصائية Statistical Notations

وهي عبارة عن أحرف أو أشكال تستخدم بدل الكلمات للسهولة . ومنها :

الكلمة	الرمز
المجموع	Σ
التكرار	Fi
المدى	R
قيمة المشاهدة	x_i
الوسط الحسابي للعينة	$\bar{X}$
الوسط الحسابي للمجتمع	M
الوسيط	M_E
المنوال	Mo
التباين للمجتمع	6^2
التباين للعينة	S^2
الانحراف المعياري للمجتمع	6
الانحراف المعياري للعينة	S
معامل ارتباط بيرسون	R
معامل ارتباط سبيرمان	p_r

المجموع (Σ) (Summtion) :

إن الرمز Σ يشير إلى عملية الجمع ، وهو حرف أغريقي يلفظ (Sigma) .

وإن (i , n) هما حدا المجموعة .

(i) يمثل دليل لتسلسل الأعداد عند عملية الجمع ، فإذا كان **i = 1** فإن ذلك يعني القيمة الأولى ، وإذا كان **i = 5** فإن ذلك يعني القيمة الخامسة ، وهكذا .

والعملية $\sum_{i=1}^{50} X_i$ تقرأ :

(مجموع مشاهدات (قيم المتغير X) ابتداءً من القيمة الأولى (i=1) أي x_1 وانتهاءً بالقيمة الأخيرة (i=50) أي x_{50} أو بشكل أكثر اختصاراً (مجموع x_i : i من 1 الى 50).

وبشكل عام إذا كان هناك عدد من المشاهدات (قيم المتغير x) لكميات عددها n فإن المجموع الكلي يتم التعبير عنه بالشكل :

$$\sum_{I=1}^{n} Xi \qquad i = (1 - n)$$

القواعد الخاصة بالمجموع :

١. يرمز لمجموع عدد المشاهدات (قيم المتغير X) ابتداءً من المشاهدة الأولى وحتى الأخيرة بالرمز

$$\sum_{I=1}^{n} X_i = x_1 + x_2 + x_3 + \cdots + x_n$$

EX: If X = 1 , 2 , 3 , 4 , 5 find

$$\sum_{I=1}^{n} X_i = x_1 + x_2 + x_3 + x_4 + x_5$$

$$= 1 + 2 + 3 + 4 + 5 = 15$$

$$\sum_{I=1}^{3} X_i = x_1 + x_2 + x_3$$

$$= 1 + 2 + 3 = 6$$

٢- وهناك مجموع جزئي مثل $\sum_{I=3}^{5} X_i$ (أي مجموع المشاهدات الثالثة والرابعة والخامسة).

٣- ويرمز لمجموع مربعات عدد من المشاهدات (قيم المتغير X) بالرمز :

$$\sum_{i=1}^{n} x_i^2 = x_1^2 + x_{i2}^2 + x_3^2 + \cdots + x_n^2$$

ففي المثال السابق :

$$\sum_{I=1}^{n} x_i{}^2 = (1)^2 + (2)^2 + (3)^2 + (4)^2 + (5)^2$$

$$= 1 + 4 + 9 + 16 + 25 \quad = 55$$

٤- يرمز لمربع مجموع عدد من المشاهدات (قيم المتغيرX) بالرمز :

$$\left(\sum_{i=1}^{n} x_i\right)^2 = (x_1 + x_2 + x_3 + \cdots + x_n)^2$$

ففي المثال السابق فإن :

$$\left(\sum_{I=1}^{5} x_i\right)^2 = (1 + 2 + 3 + 4 + 5)^2$$

$$= (15)^2 = 125$$

ملاحظة : نستنتج من ٣ ، ٤ بأن :

$$\sum_{i=1}^{n} x_i^2 \neq \left(\sum_{i=1}^{n} x_i\right)^2$$

٥- ويرمز لحاصل ضرب مجموعي قيم المتغيرين x , y بالرمز :

$$\sum_{I=1}^{n} x_i \sum_{I=1}^{n} y_i = (x_1 + x_2 + x_3 + \cdots + x_n)(y_1 + y_2 + y_3 + \cdots + y_n)$$

مثال: إذا كان x= 2, 3,5,1 و y= 4,2,3,5 أوجد ما يأتي :

1) $\sum_{I=1}^{3} x_i = x_1 + x_2 + x_3$

$$= 2 + 3 + 5 = 10$$

2) $\sum_{I=1}^{n} x_i \sum_{I=1}^{n} y_i = (x_1 + x_2 + x_3 + \cdots + x_n)(y_1 + y_2 + y_3 + \cdots + y_n)$

$$= (2 + 3 + 5 + 1)(4 + 2 + 3 + 5)$$

$$= (11)(14) = 154$$

٦- يرمز لمجموع حاصل ضرب متغيرين y , x بالرمز :

$$\sum_{I=1}^{n} X_i Y_i = X_1 Y_1 + X_2 Y_2 + X_3 Y_3 + \cdots + X_n$$

مثال : إذا كانت قيم كل من x ، y هي كالآتي $y= 4,2,3,5$ $x= 2,3,5,1$ أوجد كل مما يأتي :

1) $$\sum_{I=1}^{3} x_i = x_1 + x_2 + x_3$$

$$= 2 + 3 + 5 = 10$$

2) $$\sum_{I=1}^{n} X_i\ Y_i = (X_1\ Y_1 + X_2\ Y_2 + X_3\ Y_3 + \cdots + X_n\ Y_n)$$

$$= 2(4) + 3(2) + 5(3) + 1(5)$$

$$= 8 + 6 + 15 + 5 = 34$$

ملاحظة : نستنتج من ٦ ، ٥ أن

$$\sum_{I=1}^{n} X_i \quad \sum_{I=1}^{n} Y_i \neq \sum_{I=1}^{n} X_i Y_i$$

٧- يرمز لمجموع مقلوب قيم المتغير X بالرمز :

$$\sum_{i=1}^{n} \frac{1}{X_i} = \frac{1}{X_1} + \frac{1}{X_2} + \cdots + \frac{1}{X_n}$$

مثال : إذا كانت $X = 2, 3, 1$ أوجد :

1) $$\sum_{I=1}^{3} \frac{1}{X_i} = \frac{1}{X_1} + \frac{1}{X_2} + \frac{1}{X_3}$$

$$= \frac{1}{2} + \frac{1}{3} + \frac{1}{1} = \frac{11}{6}$$

٨- يرمز لمقلوب مجموع قيم المتغير X بالرمز :

$$\frac{1}{\sum_{i=1}^{n} x_i} = \frac{1}{x_1 + x_2 + \cdots + x_n}$$

مثال : إذا كانت قيم كل من x و y هي كالآتي : X = 2 , 3 , 1 , 4 اوجد كل مما يأتي :

$$\frac{1}{\sum_{I=1}^{n} x_i} = \frac{1}{x_1 + x_2 + x_3 + x_4}$$

$$= \frac{1}{2+3+1+4} = \frac{1}{10}$$

بعض القواعد الأساسية في عملية الجمع :

١- إذا كان (c) أي عدد ثابت فإن :

$$\sum_{I=1}^{n} c = nc$$

البرهان :

$$\sum_{I=1}^{n} c = c_1 + c_2 + c_3 + \cdots + c_n = nc$$

مثال : أوجد كل مما يأتي :

$$\sum_{I=1}^{4} 7 = 7 + 7 + 7 + 7$$

$$= 7\,(4) = 28$$

٢- اذا كان (c) اي عدد ثابت فان:

$$\sum_{I=1}^{n} cx_i = c \sum_{I=1}^{n} x_i$$

البرهان :

$$\sum_{I=1}^{n} cx_i = cx_1 + cx_2 + cx_3 + \ldots + cx_n$$

$$= c\,(x_1 + x_2 + x_3 + \cdots + x_n)$$

$$= c \sum_{I=1}^{n} x_i$$

مثال : إذا كانت $x = 1, 2, 3, 4, 5$ أوجد ناتج ما يأتي :

$$1)\ \sum_{I=1}^{n} 4X_i = 4(1+2+3+4+5)$$

$$= 4\,(\,15\,) = 60$$

$$2)\ \sum_{I=1}^{n} 2\,X_i^2 = 2(\,x_1^2 + x_2^2 + \cdots + x_5^2)$$

$$= 2\,(\,1+4+9+16+25\,)$$

$$= 2\,(\,55\,) = (\,110\,)$$

٣- جمع متغيرين أو اكثر هو مجموع جمعهم ، أي ان :

$$\sum_{I=1}^{n}(x_i y_i) = \sum_{I=1}^{n} X_I + \sum_{I=1}^{n} y_i$$

البرهان :

$$\sum_{I=1}^{n}(x_i + y_i) = (x_1 + y_1) + (x_2 + y_2) + \cdots + (\,x_n + y_n\,)$$

$$= (\,x_1 + x_2 + \cdots + x_n) + (y_1 + y_2 + \cdots + y_n)$$

$$= \sum x_i + \sum y_i$$

مثال / إذا كان $\sum_{I=1}^{4} x_i = 5\,,\ \sum_{I=1}^{4} y_i = 8$ جد :

$$\sum_{I=1}^{4}(2X_i - 3y_i + 5) = 2\sum_{I=1}^{4} x_i + 3\sum_{I=1}^{4} y_i + \sum_{I=1}^{4} 5$$

$$= 2 \times 5 - 3 \times 8 + 4 \times 5$$

$$= 10 - 24 + 20$$

$$= -14 + 20 \quad = 6$$

ملاحظة :

* يرمز لمجموع حاصل قسمة المتغيرين x , y بالرمز :

$$\sum_{I=1}^{n} \frac{x_i}{y_i} = \frac{x_1}{y_1} + \frac{x_2}{y_2} + \cdots + \frac{x_n}{y_n}$$

* ويرمز لحاصل قسمة مجموعي المتغيرين x , y بالرمز :

$$\frac{\sum_{I=1}^{n} x_i}{\sum_{I=1}^{n} y_i} = \frac{x_1 + x_2 + \cdots + x_n}{y_1 + y_2 + \cdots + y_n}$$

H . W :

If X = 3 , 6 , 9 , 12 , 15 Y= 2 , 4 , 6 , 8 , 10 Find the following :

1) $\sum_{I=1}^{3} X_i$

2) $\sum_{I=1}^{n} \sqrt{1}^2$

3) $\sum_{I=1}^{3} X_i\ y_i \quad 2$

4) $\left(\sum_{I=1}^{3} y_i \right)^2$

5) $\sum_{I=1}^{n} \frac{1}{X_i y_i}$

6) $\sum_{I=1}^{n} 2\ x_i\ y_i^2$

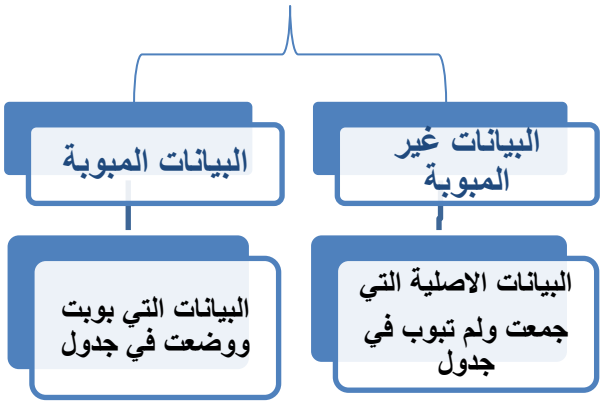
طرق عرض البيانات

البيانات DATA مجموعة القيم التي يتم جمعها من مفردات المجتمع أو العينة لخاصية معينة (متغير).

وإن البيانات الخام لا معنى لها ولا يمكن قياسها إلا إذا خضعت للمعالجات الاحصائية،

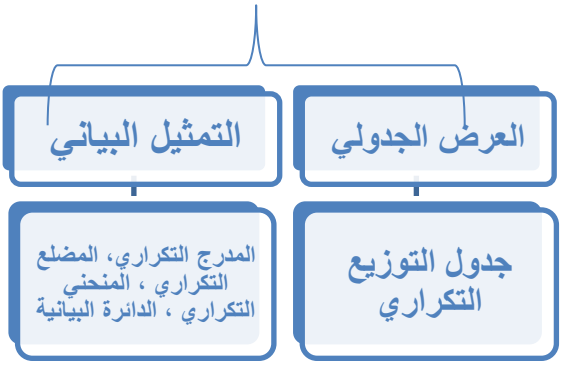
فالبيانات قبل عرضها تسمى **بيانات غير مبوبة (Ungrouped)** وبعد عرضها تسمى **بيانات مبوبة (Grouped)** .

أنواع البيانات



١. **البيانات غير المبوبة (Ungrouped Date):** وهي البيانات الأصلية التي جمعت ولم تبوب في جدول.
٢. **البيانات المبوبة (Grouped Date):** وهي البيانات التي بوبت ووضعت في جدول توزيع تكراري.

طرق عرض البيانات



تقسم طرق عرض البيانات الى قسمين رئيسين هما:

أولاً: العرض الجدولي (Tabular Presentation):

ويطلق عليه أيضاً **الجدول الاحصائي** : وهو عبارة عن جدول منظم ، يسمى (جدولاً بسيطاً) إذا تألف من بعد واحد ، و (جدولاً مركباً) إذا تألف من بعدين .

الجدول البسيط: وهو الجدول الذي توزع فيه البيانات بحسب صفة واحدة ، ويتألف عادة من عمودين ، **الأول** يتضمن تقسيمات الصفة الى فئات أو مجموعات ، و**الثاني** يبين عدد المفردات التابعة لكل فئة أو مجموعة .

مثال: إعرض البيانات الآتية في جدول تكراري

1 , 2, 2, 2 , 3 , 3 , 4 , 5 , 5 , 5 , 5 , 5 , 6 , 7 , 8 , 8 , 8 , 9 , 9 , 9

الحل :

X	1	2	3	4	5	6	7	8	9	ΣX
f_i	1	3	2	1	4	1	1	2	3	18

الجدول المركب: وهو الجدول الذي توزع فيه البيانات حسب صفتين أو ظاهرتين أو أكثر في نفس الوقت.

مثال: الجدول أدناه يبين توزيع مجموعة من الطلبة بحسب الوزن والطول :

الوزن (كغم) / الطول	51 - 60	61 - 70	71 - 80	المجموع
121- 140	20	6	4	30
141 – 160	2	40	10	52
161 – 180	2	6	10	181
المجموع	24	52	24	100

ويتمثل العرض الجدولي بجدول التوزيع التكراري .

جدول التوزيع التكراري (Frequency Distribution Table):

وهو جدول بسيط يتكون من عمودين **الأول يسمى عمود الفئات (Classes)** ، وتقسم فيه قيم المتغيرالى أقسام ومجموعات ، ويرمز له بالرمز (C) ، **والثاني يسمى عمود التكرارات (Frequency)** ويتضمن عدد المرات التي تكررت فيها الفئة ، ويرمز له بالرمز (f_i) .

بعض المفاهيم ذات الصلة بجدول التوزيع التكراري :

١. **التوزيع التكراري :** هو عبارة عن تلخيص وترتيب لبيانات المتغير االعشوائي التي سبق و أن جمعت وصنفت مقسمة الى عدد من المجاميع يسمى كل منها بـ (الفئة Class) . وهذه الفئات قد تكون مرتبة تصاعدياً أو تنازلياً حسب طبيعة البيانات . ويسمى توزيع عدد قيم X حسب الفئات بـ (التوزيع التكراري).

٢. **المدى الكلي (Total Range):** يعرف المدى الكلي بأنه الفرق ما بين أكبر قيمة و أصغر قيمة في المجموعة . ويرمز له بالرمز (T.R)

$$T.R = X_L - X_S$$

٣- **الفئات Class :** هي عدد المجاميع التي يتألف منها التوزيع التكراري ، وكل مجموعة تتحدد بحدين ، الأول يسمى **الحد الأدنى للفئة (Lower Class Limit)** ويرمز له بالرمز **L.L.** ، والثاني يسمى **الحد الأعلى للفئة (Upper Class Limit)** ويرمز له بالرمز **U.L.** .

مثل: الفئة (16 - 10) حدها الأدنى 10 وحدها الأعلى 16

٤- الحدود الحقيقية للفئات :

لكل فئة حدان حقيقيان حد أدنى حقيقي وحد أعلى حقيقي

الحد الأدنى الحقيقي للفئة = الحد الأدنى - 0.5

مثل: إذا كان الحد الأدنى لفئة هو (10) فإن الحد الأدنى الحقيقي لها هو ؟

الحد الأدنى الحقيقي = الحد الأدنى – 0.5 = 9.5

الحد الأعلى الحقيقي للفئة = الحد الأعلى + 0.5

مثل: إذا كان الحد الأعلى لفئة هو (16) فإن الحد الأعلى الحقيقي لها هو ؟

الحد الأعلى الحقيقي = الحد الأعلى + 0.5 = 16.5

٥- **عدد فئات التوزيع m .** ويمثل عدد المجاميع التي يتألف منها التوزيع التكراري . فإذا رمزنا لعدد الفئات بالرمز m فإن :

$$m = 2.5 \sqrt[4]{n}$$

حيث n تمثل عدد المفردات (قيم المتغير x) .

ويجب أن لا يقل عدد الفئات عن خمسة فئات ولا يزيد عن خمسة عشر فئة .

٦- تكرار الفئة (Classes Frequency) : هو عدد القيم التي تقع في مدى تلك الفئة ، ويرمز له بالرمز (f_i).

إذ إن f_i يمثل تكرار الفئة الأولى ، f_2 يمثل تكرار الفئة الثانية ، .. وهكذا .

وإن مجموع التكرارات يجب أن يكون دائماً مساوياً للعدد الكلي لقيم المفردات .

مثلاً: تكرار الفئة الثالثة = 5 ، فهذا يعني أن هناك 5 قيم من قيم المتغير واقعة في المدى الفئة ..

٧- مركز الفئة (Center of Classes): ويمثل منتصف المدى بين حدي للفئة ، ويرمز له بالرمز **x** .

أي إن :

$$\text{مركز الفئة} = \frac{\text{الحدالأدنى} + \text{الحدالأعلى}}{2}$$

$$X = \frac{L.L + U.L}{2}$$

مثال: إذا كانت لدينا الفئة (16 – 10) أوجد مركز الفئة :

الحل:

$$X = \frac{L.L+U.L}{2}$$

$$= \frac{10+16}{2} = 13$$

٨- طول الفئة (Length of Classes): ويسمى أحياناً **بالمدى الفئوي** ، ويمثل مقدار المدى بين حدي للفئة ، ويرمز له بالرمز **(L)** .

وأن طول الفئة يتناسب عكسياً مع عدد الفئات ، فكلما زاد طول الفئة قل عدد الفئات والعكس صحيح.

ويجب أن تكون أطوال الفئات متساوية لتسهيل العمليات الحسابية ، ويجب أن تكون أعداد صحيحة موجبة دائماً .

طول الفئة = الحد الأعلى - الحد الأدنى + ١

$$L = U.L - L.L + 1$$

مثال: إذا كان لدينا الفئة (16 - 10)، أوجد طول الفئة :

الحل: طول الفئة = الحد الأعلى - الحد الأدنى + ١

$$L = 16 - 10 + 1 = 7$$

كما يمكن استخراجه من :

أ) إذا أمكن استخراج المدى الكلي وعدد الفئات فإن طول الفئة يستخرج من القانون :

$$L = \frac{T.R}{m}$$

حيث T.R : هو المدى الكلي

m : عدد الفئات

ب) أو من خلال : الفرق بين الحدين الأعلى أو (الأدنى) لفئتين متتاليتين إذا كانت هناك فئتين متتاليتين معلومتين :

$$L = (L.L)_2 - (L.L)_1$$

$$L = (u.L)_2 - (u.L)_1$$

ج) أو من (الفرق بين مركزي فئتين متتاليتين) :

حيث : X_2 يمثل مركز الفئة الثانية و X_1 يمثل مركز الفئة الأولى.

$$L = X_2 - X_1$$

خطوات إنشاء التوزيع التكراري (تبويب البيانات) :

لنأخذ المثال الافتراضي الآتي :

لتكن $X_1 , X_2 , X_3 , \dots , X_n$ بيانات المتغير العشوائي X من عينة عشوائية ما عدد مفرداتها n مفردة . ونرغب في تلخيص هذه البيانات في توزيع تكراري عدد فئاته هو m .

ولنفرض أن X_S هي أصغر قيمة ، X_L هي أكبر قيمة في مجموعة بياناتت المتغير X ، عندئذ نتبع الخطوات الآتية في إنشاء جدول التوزيع التكراري :

١- تحديد أصغر قيمة وأعلى قيمة من قيم المتغير العشوائي X .

٢- استخراج المدى الكلي R . T .

٣- استخراج عدد فئات التوزيع m .

٤- استخراج طول الفئة L .

٥- تحديد الحد الأدنى والحد الأعلى للفئات ، وكتابة حدود الفئات .

٦- استخراج عدد التكرارات لكل فئة f_i .

ملاحظة مهمه : يفضل أن يمتاز التوزيع التكراري بما يأتي :

١- أن تكون الفئات متساوية في الطول .
٢- أن يكون التوزيع التكراري توزيع مغلق .
٣- أن تبدأ فئات التوزيع وتنتهي (حدود الفئات) بأعداد صحيحة ، وذلك لتسهيل اجراء العمليات الحسابية خصوصاً إذا أجريت بدون استخدام الحاسبات .
٤- أن لايقل عدد فئات التوزيع التكراري عن 5 فئات ولا يزيد عن 15 فئة .
٥- أن تكون حدود الفئات محددة بشكل واضح بحيث أن كل قيمة من قيم المتغير تبوب ضمن فئة واحدة فقط من فئات التوزيع .

مثال : البيانات الآتية تمثل درجات 36 طالباً في مادة الإحصاء . المطلوب انشاء جدول تكراري لهذه النباتات أو (عرض البيانات الآتية في جدول توزيع تكراري) .

80	84	80	87	97	81	56	92	70	71	78	82
83	93	65	51	75	68	93	68	86	43	74	73
61	80	91	75	67	72	80	95	97	70	74	81

الحل :

1) $X_L = 97 \qquad X_s = 43$

2) $T.R = X_L - X_s$

$= 97 - 43 = 54$

3) $m = 2.5 \sqrt[4]{n}$

$= 2.5 \sqrt[4]{36}$

$= 5.12 = 7$

4) $L = \frac{T.R.}{m} = \frac{54}{7} = 7.71 \cong 8$

ت	C	f_i
1	43 - 50	1
2	51 - 58	2
3	59 - 66	2
4	67 - 74	10
5	75 - 82	10
6	83 - 90	4
7	91 - 98	7
		36

- **ملاحظة :** اذا كان هناك فئة موجودة في الجدول نستخرج طول الفئة مباشرة من القانون **طول الفئة = الحد الأعلى – الحد الأدنى+ ١** ثم أكمال الجدول دون الحاجة الى استخراج المدى الكلي ، ونتوفق بعدد الفئات عندما تكون أعلى قيمة واقعة ضمن آخر فئة.

مثال :نظم البيانات الاتية : (15 , 21 , 15 , 19 , 11 , 10 , 16 , 15 , 5 , 24) في جدول تكرار فئته الأولى (9 – 5)

الحل :

١) نستخرج طول الفئة (L)

طول الفئة = الحد الأعلى – الحد الأدنى + ١

= ٩ - ٥ + ١

= ٥

٢) بتطبيق القانون طول الفئة يساوي الفرق ما بين طولي فئتين متباينتين نكمل كتابة الفئات .

طول الفئة = الحد الأدنى للفئة الثانية ــ الحد الأدنى للفئة الأولى

٥ = الحد الأدنى للفئة الثانية ــ ٥

الحد الأنى للفئة الثانية = ١٠

طول الفئة = الحد الأعلى للفئة الثانية ــ الحد الأعلى للفئة الأولى

٥ = الحد الأعلى للفئة الثانية ــ ٩

الحد الأعلى للفئة الثانية = ١٤

وهكذا لبقية الفئات .

٣) ثم نحسب التكرارات

	C	F_i
1	5 - 9	1
2	10 - 14	2
3	15 - 19	5
4	20 - 24	2
5	25 - 29	0

مثال : إذا كان عدد مفردات ظاهرة ما هو (50) مفردة وإن أقل قيمة فيها هي (10) وأعلى قيمة هي (90) . أكمل الجدول الآتي :

	C	X	الحدود الحقيقية للفئات
1			
2			
3			
4			
5			
6			
7			
8			
9			

الحل :

1) $X_L = 90 \qquad X_s = 10$

2) $T.R. = X_L - X_s$

$= 90 - 10 + = 80 = 80$

3) $m = 2.5\ 4\&n$

$= 2.5\ \sqrt[4]{150}$

$= 6.64 \cong 7$

4) $L = \frac{T.R.}{m} = \frac{80}{7} = 11.48 \cong 12$

5) $L = u.L. - L.L. + 1$

$12 = u.L. - 10 + 1$

$u.L. = 21 \ldots.$

6) $X_1 = \frac{u.L. + L.L.}{2} = \frac{21 + 10}{2} = 15.5$

$L = X_2 - X_1 = 15.5 + 12 = 27.5$

..... وهكذا .

الحد الأدنى الحقيقي للفئة الأولى = الحد الأدنى للفئة – 0.5 10 – 0.5 = 9.5

الحد الأعلى الحقيقي للفئة الأولى = الحد الأعلى للفئة + 0.5 21 + 0.5 = 21.5

وهكذا ....

	C	X	الحدود الحقيقية للفئات
1	10 - 21	15.5	9.5 - 21.5
2	22 - 33	27.5	21.5 - 33.5
3			
4			
5			
6			
7			

التوزيع التكراري المتجمع (Cumulativ frequency distribution) :

وهو التوزيع الذي يبين كمية التكرار المتجمع عن قيمة معينة من قيم المتغير العشوائي .

وهو على نوعين :

أ – التوزيع التكراري المتجمع الصاعد : ($F_i \uparrow$)

وهو التوزيع الذي يبين تراكم التكرارات ابتداءً من الفئة الأولى في التوزيع وانتهاءً بالفئة الأخيرة منه . ويتم حساب التكرارات المتجمعة على أساس الحدود العليا للفئات .

ب – التوزيع التكراري المتجمع النازل : ($F_i \downarrow$)

وهو التوزيع الذي يبين تناقص التكرارات ابتداءً من الفئة الأولى في التوزيع وانتهاءً بالفئة الأخيرة منه . ويتم حساب التكرارات المتجمعة على أساس الحدود الدنيا للفئات .

مثال : الآتي جدول توزيع تكراري يمثل توزيع 60 عائلة فلاحية حسب ملكيتها من عدد أشجار البرتقال . المطلوب تكوين التوزيع التكراري المتجمع الصاعد والنازل :

	C	f_i	$F_i \uparrow$	$F_i \downarrow$
1	60 - 74	4	4	60
2	75 - 89	5	9	56
3	90 - 104	10	19	51
5	120 - 134	28	47	41
6	135 - 149	7	54	13
7	150 - 164	6	60	6
		60		

تمرين 1: أكمل الجدول الآتي :

	C	X	الحدود الحقيقية للفئات
1	10 -	12	
2			
3			
4			
5			
6			

تمرين 2 : أكمل الجدول الآتي :

	C	X	الحدود الحقيقية للفئات
1			
2			
3			
4			
5			
6	90 - 99		

تمرين 3 : البيانات الآتية تمثل درجات 30 طالب في مادة الإحصاء :

١- أنشئ جدول التوزيع التكراري لدرجات الطلبة . ٢- استخرج الحدود الحقيقية للفئات ، مراكز الفئات . ٣- التكرار المتجمع الصاعد ولتكرار المتجمع النازل .

13 , 15 , 25 , 67 , 88 , 33 , 90 , 98 , 78 , 75 , 44 , 54 , 31 , 40 , 89 , 55 , 76
38 , 51 , 34 , 72 , 69 , 37 , 33 , 42 , 88 , 78 , 79 , 29 , 43 ,

تمرين ٤ : نظم البيانات الاتية (15 , 21 , 15 , 19 , 11 , 10 , 16 , 15 , 5 , 24) في جدول تكرارات فئته الثانية (9 – 13) ، ثم استخرج التكرار المتجمع الصاعد والتكرار المتجمع النازل .

تمرين ٥ : البيانات الآتية تمثل درجات 25 طالباً في مادة الإحصاء . المطلوب انشاء جدول تكراري لهذه البيانات أو (اعرض البيانات الآتية في جدول توزيع تكراري) ثم جد مراكز الفئات والتكرارات المتجمعة الصاعدة والنازلة .

84	80	87	97	81
92	70	71	78	82
93	65	51	75	68
68	86	43	74	73
80	91	75	67	72

ثانياً : التمثيل البياني لجدول التوزيع التكراري

١- المدرج التكراري :

هو عبارة عن مستطيلات رأسية تمتد قواعدها على المحور الأفقي لتمثل الحدود الحقيقية للفئات بينما يمثل ارتفاعها تكرار الفئات .

خطوات رسم المدرج التكراري :

١- رسم المحورين الأفقي و العمودي .
٢- ٢- تدريج المحور الأفقي الى أقسام متساوية بحيث يشمل جميع الحدود الحقيقية للفئات ، ويقسم المحور العمودي الى اقسام متساوية بحيث يشمل على اكثر التكرارات .
٣- ٣- يرسم على كل فئة مستطيلاً رأسياً تمثل قاعدته طول تلك الفئة وارتفاعه يمثل تكرارات تلك الفئة.

مثال : ارسم المدرج التكراري لجدول التوزيع التكراري الآتي :

	C	F_i
1	5 - 9	3
2	10 - 14	6
3	15 - 19	5
4	20 - 24	8
5	25 - 29	12
6	30 - 39	7
$\sum f_i$		41

الحل :

نستخرج الحدود الحقيقية للفئات :

الحد الأدنى الحقيقي للفئة الأولى = الحد الأدنى للفئة الأولى – 0.5

= 5 - 0.5 = 4.5

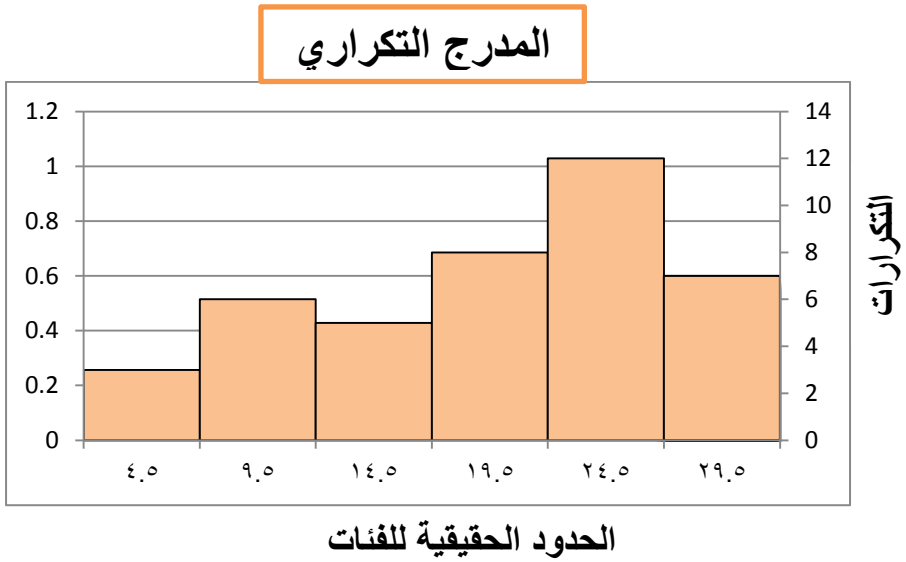
الحد الأعلى الحقيقي للفئة الأولى = الحد الأعلى للفئة الأولى + 0.5

= 9 + 0.5 = 9.5

وبنفس الطريقة نكمل بقية الحدود الحقيقية للفئات

كما في الجدول أدناه

	C	F_i	الحدود الحقيقية للفئات
1	5 - 9	3	4.4 - 9.5
2	10 - 14	6	9.5 - 14.5
3	15 - 19	5	14.5 - 19.5
4	20 - 24	8	19.5 - 24.5
5	25 - 29	12	24.5 - 29.5
6	30 - 34	7	29.5 - 34.5
$\sum f_i$		41	



٢- المضلع التكراري :

هو عبارة عن خطوط مستقيمة متكسرة تصل بين نقاط كل منها واقعة فوق مركز فئة على ارتفاع يمثل تكرار تلك الفئة .

ملاحظة: عادة يغلق المضلع بأن نصل بداية المضلع بالمحور الأفقي بمركز فئة (خيالية) واقعة الى يسار أول فئة تكرارها صفراً ، ونصل نهاية المضلع بالمحور الأفقي بمركز فئة (خيالية) واقعة الى يمين اخر فئة تكرارها صفراً أيضاً .

خطوات رسم المضلع التكراري :

١- رسم المحورين الأفقي والعمودي .
٢- ٢- تدريج المحور الأفقي الى أقسام متساوية بحيث يشمل جميع مراكز الفئات ، ويقسم المحور العمودي الى اقسام متساوية بحيث يشمل على اكثر التكرارات .
٣- ٣- وضع نقطة أمام مركز كل فئة ارتفاعها يعادل تكرار تلك الفئة .
٤- ٤- توصيل تلك النقاط بخطوط مستقيمة .

مثال : ارسم المضلع التكراري للمثال السابق :

الحل :

أولاً نستخرج مراكز الفئات

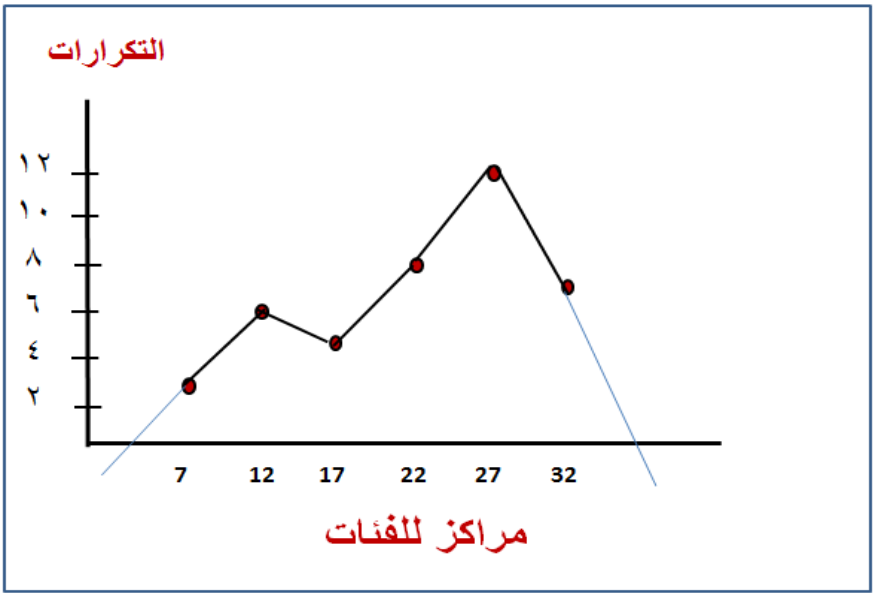
$$\text{مركز الفئة} = \frac{\text{الحد الأدنى للفئة} + \text{الحد الأعلى للفئة}}{2}$$

$$X_1 = \frac{5+9}{2} = \frac{14}{2} = 7$$

وهكذا نكمل مراكز بقية الفئات ، كما في الجدول أدناه

	C	F_i	X
1	5 -9	3	7
2	10 - 14	6	12
3	15 - 19	5	17
4	20 - 24	8	22
5	25 - 29	12	27
6	30 - 34	7	32
$\sum f_i$		41	

ثانياً نبدأ بالرسم



٣- المنحني التكراري :

وهو عبارة عن منحني يمر بمعظم النقاط الواقعة على مراكز الفئات والتي ارتفاعها يمثل تكرارات تلك الفئات .

ملاحظة : يفضل غلق المنحني التكراري بأن نصل بدايته بالحد الأدنى للفئة الأولى ونهايته بالحد الأعلى للفئة الأخيرة . وتكون مساحة المنحني (مكافئة) وليست مساوية للمضلع التكراري .

مثال : ارسم المضلع التكراري للمثال السابق :

الحل :

أولاً نستخرج مراكز الفئات

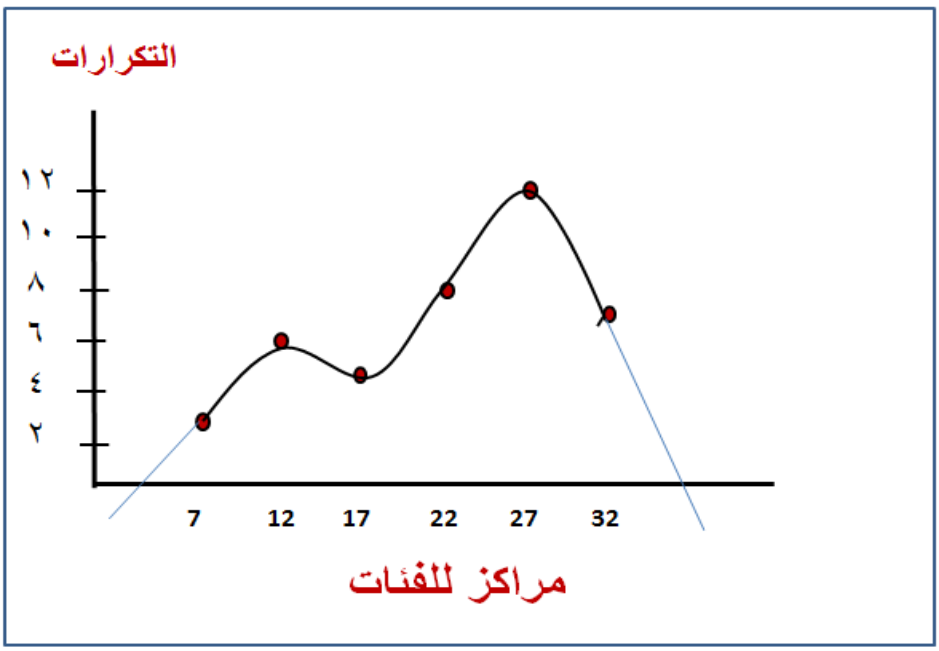
$$\text{مركز الفئة} = \frac{\text{الحد الأدنى للفئة} + \text{الحد الأعلى للفئة}}{2}$$

$$X_1 = \frac{5+9}{2} = \frac{14}{2} = 7$$

وهكذا نكمل بقية مراكز الفئات ، كما في الجدول أدناه

	C	F_i	X
1	5 -9	3	7
2	10 - 14	6	12
3	15 - 19	5	17
4	20 - 24	8	22
5	25 - 29	12	27
6	30 - 34	7	32
$\sum f_i$		41	

ثانياً نبدأ بالرسم



٤- الدائرة البيانية :

تعد هذه الطريقة أفضل الطرق لتمثيل االبيانات ذات الصفة المشتركة ونستطيع بواسطتها أن نقارن الأجزاء بعضها ببعض ثم الجزء (القطاع الدائري) بالكل (الدائرة) .

خطوات رسم الدائرة البيانية :

١- استخراج زاوية القطع $= \left(360 \times \frac{\text{الجزء}}{\text{الكل}}\right)$

٢- نرسم دائرة معينة ونرسم عليها نصف القطر .

٣- نرسم الزاوية المركزية التي ضلعها الابتدائي نصف القطر والممثلة بالقطاع .

مثال/ مجموعة من الفاكهة وزعت على طلاب القسم الداخلي وكانت كالآتي :

نوع الفاكهة	تفاح	موز	برتقال	رمان	المجموع
العدد	180	540	90	270	1080

المطلوب تمثيل هذه البيانات بالقطاع الدائري .

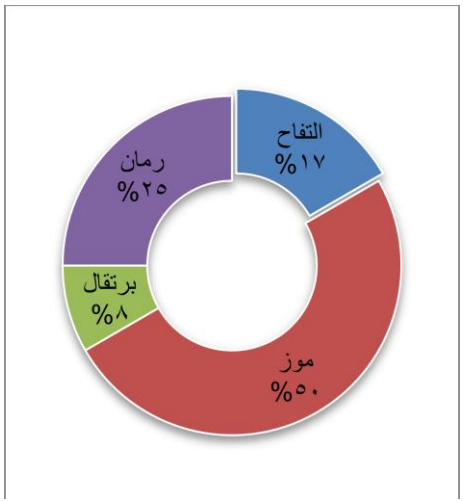
الحل :

زاوية قطاع التفاح $= 360 \times \frac{180}{1080} = 60^\circ$

زاوية قطاع الموز $= 360 \times \frac{540}{1080} = 180^\circ$

زاوية قطاع البرتقال $= 360 \times \frac{90}{1080} = 30^\circ$

زاوية قطاع الرمان $= 360 \times \frac{270}{1080} = 90^\circ$



تمارين :

١- الآتي توزيع تكراري لدرجات (40) طالباً في مادة الأحصاء الترب

93	82	78	71	70	92	56	81
74	73	72	68	75	51	65	93
35	90	83	73	74	43	86	68
92	76	71	90	72	67	75	91
80	61	72	97	91	88	81	74

اعرض البيانات أعلاه في جدول توزيع تكراري ، ثم ارسم المدرج والمضلع التكراري .

٢- أكمل جدول التوزيع التكراري الآتي ، ثم اعرض البيانات في الجدول بإستخدام التمثيل البياني (المدرج والمضلع والمنحني) التكراري :

	C	f_i	X
1	1 -	2	4
2		4	
3		7	
4		12	
5		10	
6		15	
7		8	

الفصل الثاني

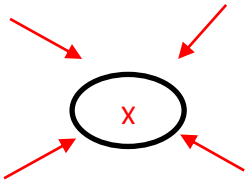
المقاييس الإحصائية

مقاييس النزعة المركزية Measures of Central Tendency

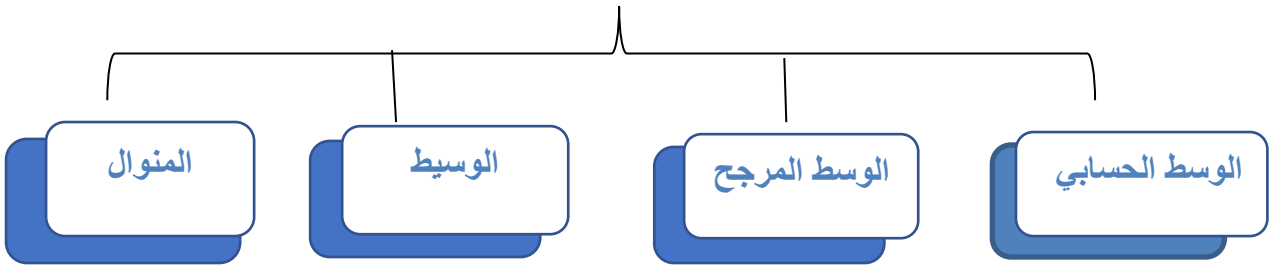
إن مصطلح النزعة المركزية يعني التمركز والتكثف نحو رقم معين .

تعرف مقاييس النزعة المركزية بأنها تلك المقاييس التي تحث عن قيمة تتمركز حولها أغلبية البيانات ،

وإن هذه القيمة المتوسطة هي رقم واحد يعبر عن جميع بيانات تل المجموعة .



مقاييس النزعة المركزية



١- الوسط الحسابي (المتوسط) (The Arithmetic Mean):

وهو القيمة الناتجة عن قسمة مجموع القيم على عددها ، ويرمز له بالرمز $\overline{X}$.

طرق إحتساب الوسط الحسابي :

١- في حالة البيانات غير المبوبة :

إذا كان لدينا $X_1 , X_2 , X_3 , ... , X_n$ بيانات المتغير العشوائي X عددها n من المفردات ، فإن الوسط الحسابي لها :

$$\overline{X} = \frac{\sum x_i}{n}$$

حيث : X : القيم للمتغير

n : عدد القيم

مثال : البيانات الآتية تمثل درجات 15 طالب في مادة الإحصاء (43 , 32 , 55 , 81 , 80 , 36 , 49 , 58 , 69 , 75 , 65 , 72 , 9060 , 95 ,) أوجد الوسط الحسابي لدرجات الطلاب .

الحل :

$$\overline{X} = \frac{\sum x_i}{n}$$

$$= \frac{60+72+65+75+69+58+49+36+80+81+55+32+43+95+90}{15}$$

$$= \frac{987}{15} = \mathbf{65.8}$$

أي أن معدل درجات الطلبة هو 65.8 .

٢- في حالة البيانات المبوبة :

إذا كانت $x_1 , x_2 , x_3 , \dots , x_n$ تمثل مراكز فئات في توزيع تكراري عدد فئاته m فئة ، و $f_1 , f_2 , \dots , f_n$ هي تكرارات تلك الفئات فإن الوسط الحسابي يستخرج من الصيغة الآتية :

$$\overline{X} = \frac{\sum (X)(f_i)}{\sum f_i}$$

حيث : X تمثل مراكز الفئات

f_i تمثل التكرارات

مثال : أوجد الوسط الحسابي للبيانات الآتية :

	C	f_i	X	Xf_i
1	2 - 6	2	4	8
2	7 - 11	8	9	72
3	12 - 16	6	14	84
4	17 - 21	1	19	19
5	22 - 26	5	24	110
Σ		22		203

الحل :

$$\overline{X} = \frac{\sum (Xf_i)}{\sum f_i} = \frac{203}{22} = \mathbf{9.22}$$

خصائصه الوسط الحسابي :

١- مجموع انحرافات قيم المتغير العشوائي X عن وسطها الحسابي الذي أحتسب منها يساوي صفر .

للبيانات غير المبوبة : $\sum(x_i - \bar{X}) = 0$

البرهان :

$$\sum(x_i - \bar{X}) = \sum x_i - \sum \bar{X}$$

$$= \sum x_i - n.\ \bar{X}$$

$$= \sum x_i - \frac{n\sum x_i}{n}$$

$$= \sum x_i - \sum x_i = 0$$

للبيانات المبوبة : $\sum f_i(x_i - \bar{x}) = 0$

البرهان :

$$\sum f_i\,(x_i - \bar{X}) = \sum f_i x_i - \bar{X}\sum f_i$$

$$= \sum f_i x_i - \frac{\sum f_i x_i - \sum f_i}{\sum f_i}$$

$$= \sum f_i x_i - \sum f_i x_i = 0$$

٢- عند إضافة عدد ثابت k إلى كل قيمة من قيم المفردات (المشاهدات) فإن قيمة الوسط الحسابي تزداد بمقدار العدد الثابت . أي إن :

$$X_i = y_i + k \longrightarrow \bar{X} = \bar{y} + k$$

البرهان :

$$X_i = y_i + k$$

$$\sum x_i = \sum y_i + n.k$$

نقسم طرفي المعادلة على n

$$\frac{\sum x_i}{n} = \frac{\sum y_i}{n} + \frac{n.\ k}{n}$$

$$\bar{X} = \bar{y} + k$$

٣- عند ضرب عدد ثابت k إلى كل قيمة من قيم المفردات (المشاهدات) فإن قيمة الوسط الحسابي تكون مساوية لحاصل ضرب العدد الثابت في الوسط الحسابي .

$$X_i = y_i . k \longrightarrow \bar{X} = \bar{y} . k$$

البرهان :

$$X_i = y_i . k$$

$$\sum x_i = \sum y_i . n . k$$

نقسم طرفي المعادلة على n

$$\frac{\sum X_i}{n} = \frac{\sum y_i}{n} . \frac{n . k}{n}$$

$$\bar{X} = \bar{y} . k$$

٤- اذا ضربت كل مشاهدة من المشاهدات بمقدار ثابت a واضيف لها مقدار ثابت b فإن الوسط الحسابي الجديد = الوسط الحسابي الأصلي مضروب في a ومضاف له b .

$$x_i = ay_i + b \longrightarrow x = -_{ay} + -_b$$

البرهان :

$$\sum x_i = \sum (ay_i + b)$$

$$\sum X_i = a \sum y_i + \sum b$$

$$\sum x_i = a \sum y_i + n . b$$

نقسم طرفي المعادلة على n

$$\frac{\sum X_i}{n}^- = \frac{a \sum y_i}{n}^- + \frac{n.k}{n}$$

$$\bar{X} = ay^- + b$$

٥- الوسط الحسابي لمجموع قيم متغيرين = مجموع الوسطين الحسابيين للمتغيرين .

$$z_i = x_i + y_i \longrightarrow z = x^- + y^-$$

البرهان :

$$z_i = x_i + y_i$$

$$\sum z_i = \sum (x_i + y_i)$$

$$\sum z_i = \sum x_i + \sum y_i$$

نقسم طرفي المعادلة على n

$$\frac{\sum z_i^-}{n} = \frac{\sum x_i^-}{n} + \frac{\sum y_i^-}{n}$$

$$z^- = x^- + y^-$$

٦- لا يمكن ايجاده في التوزيعات التكرارية و الجداول المفتوحة (غير المنتهية) .

٧- الوسط الحسابي يتأثر بالقيم المتطرفة .

٨- مجموع مربعات انحرافات قيم المتغير العشوائي X عن وسطها الحسابي الذي أحتسب منها يكون أقل ما يمكن .

٩- يتأثر الوسط الحسابي بالقيم المتطرفة (صغيرة أو كبيرة) التي تقع على جانبي التوزيع .

H . W :

١) : إذا علمت أن x = 1 , 2 ,3 , 4 و y = 2 , 4 , 6 , 8 أوجد ما يأتي :

١- الوسط الحسابي بعد إضافة ٣ الى كل قيمة من قيم المتغير x.

٢- أثبت أن $\sum(x_i - \bar{X}) = 0$ بالأعتماد على قيم x في السؤال .

٣- قيمة **y-** بعد ضرب كل قيمة من قيم y في ٢ .

٤- قيمة $\bar{x}$ بعد ضرب كل قيمة من قيم x في ٤ وجمعها في ٢ .

٥- قيمة $\bar{z}$ إذا علمت أن $z = x + y$

2) الآتي توزيع تكراري لدرجات 92 طالب في مادة الإحصاء التربوي ، المطلوب إيجاد متوسط درجات الطلبة في مادة الإحصاء :

	C	F_i
1	21 - 30	4
2	31 - 40	7
3	41 - 50	10
4	51 - 60	15
5	61 - 70	18
6	71 - 80	12
7	81 - 90	6
8	91 - 100	18

٣) الآتي توزيع تكراري لدرجات (25) طالب في مادة الإحصاء التربوي ، استخرج معدل درجات الطلبة مستخدماً :

أ) حالة البيانات غير المبوبة **ب) البيانات المبوبة**

66 , 90 , 44 , 31 , 24 , 35 , 88 , 65 , 67 , 55 , 20 , 71 , 65 , 77 , 48 , 84 , 80 , 50 , 56 , 44 , 70 , 74 , 79 , 78 , 45 ,

٤) أ) أكمل جدول التوزيع التكراري أدناه

ب) ارسم المدرج والمضلع التكراري

ج) استخرج الوسط الحسابي

	C	F_i	X
1	10 -	4	14.5
2		7	
3		10	
4		25	
5		18	
6		12	
7		6	
8		18	

٥) جد قيمة z التي تجعل قيمة الوسط الحسابي في المجموعة 20 , 25 , 28 , 31 , 37, z مساوٍ إلى 30 .

٦) إذا علمت أن الوسط الحسابي للتوزيع التكراري الآتي هو 39 ، جد قيمة f .

	C	f_i
1	10 - 19	2
2		4
3		6
4		8
5	50 - 59	F

٧) صنف صف إل عدد من الطلاب والطالبات ، بحيث كان متوسط عدد الذكور فيه 20 سنة ، ومتوسط عمر الاناث 19.5 سنة ، أوجد متوسط عمر الذكور والاناث معاً .

٢- الوسط الحسابي المرجح (s الموزون) : (Weighted Mean)

يستخدم الوسط الحسابي المرجح $(\overline{X}_w)$ عندما تكون بعض المفردات أكثر أهمية من الأخرى مما يستوجب أخذ هذه الأهمية بنظر الاعتبار عند استخراج الوسط الحسابي .

ومن الأمثلة عليه (استخراج معدل الطالب الجامعي في فصل دراسي) ، إذ أن الأمر يتطلب احتساب معدل درجات الطالب موزونة بعدد الساعات الأسبوعية لكل مادة من المواد التي تدخل في حساب المعدل .

فإذا كانت $x_1 , x_2 , x_3 , \dots , x_n$ بيانات المتغير العشوائي x قوامها (عددها) n من المفردات ، وكانت $w_1 , w_2 , \dots , w_n$ تمثل أوزان هذه البيانات إن الوسط الحسابي المرجح يمكن استخراجه كما يأتي :

$$\overline{X}_w = \frac{\sum w_i \ \sum x_i}{\sum w_i}$$

حيث : $\overline{X}_w$ هو الوسط الحسابي المرجح

$\sum w_i x_i$ هو مجموع حاصل ضرب كل مفردة بوزنها

$\sum w_i$ هو مجموع أوزان المفردات

مثال : طالب جامعي في السنة الثالثة كانت درجاته في نهاية الفصل كالاتي :

اسم المادة	قياس وتقويم	علم النفس	اللغة العربية	E	ارشاد تربوي	صحة
الدرجة	72	80	90	65	65	70
عدد الساعات	3	3	2	2	2	2

فما هو معدله في الفصل الدراسي

الحل :

$$\overline{X}_w = \frac{\sum w_i \ x_i}{\sum w_i}$$

$$= \frac{70.2+65.2+65.2+90.2+80.3+72.3}{14}$$

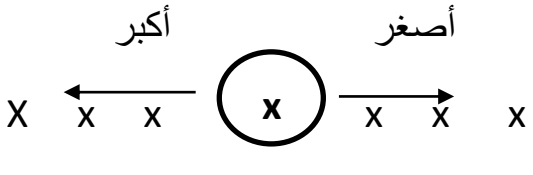
$$= \frac{1036}{14}$$

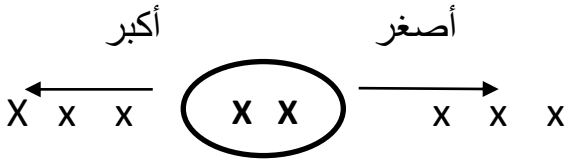
$$= 74$$

ملاحظة : إن الوسط الحسابي الاعتيادي هو حالة خاصة من الوسط الحسابي المرجح .

٣- الوسيط : (The Median)

هي القيمة التي تفصل القيم الى مجموعتين متساويتين أحدهما أكبر من القيمة والأخرى أصغر من القيمة . ويرمز له بالرمز (Me).





طرق حسابه :

أ _ في حالة البيانات غير المبوبة :

إذا كان لدينا $X_1 , X_2 , X_3 , ... , X_n$ بيانات المتغير العشوائي X عددها n من المفردات ، ورتبت تصاعدياً أو تنازلياً فإن الوسيط يحتسب من :

١- إذا كانت n عدداً فردياً فإن الوسيط هو القيمة التي يكون ترتيبها $\frac{n+1}{2}$

٢- إذا كانت n عدداً زوجياً فغن الوسيط هو متوسط للقيمتين $\frac{n}{2}$, $\frac{n}{2}+1$

خطوات استخراجه فى حالة البياناتغير المبوبة :

١- ترتب البيانات تصاعدياً أو تنازلياً .
٢- تحديد قيمة n هل هي عدد فردي أم زوجي .
٣- تحديد ترتيب الوسيط .
٤- القيمة التي تقابل ترتيب الوسيط تكون هي الوسيط إذا كان n فردي ، ومتوسط القيمتين إذا كان n عدد زوجي .

مثال ١: أوجد الوسيط لدرجات خمسة طلاب في امتحان مادة الإحصاء إذا علمت ان الدرجات هي : 80 , 82 , 76 , 87 , 84

الحل :

١- ترتيب البيانات ترتيباً تصاعدياً : 76 , 80 , 82 , 84 , 87

٢- تحديد قيمة n : n = 5 (عدد فردي)

٣- تحديد ترتيب الوسيط : الوسيط هو القيمة التي ترتيبها

$$\frac{n+1}{2} = \frac{5+1}{2} = 3$$

٤- استخراج قيمة الوسيط : $Me = X_3 = 82$

مثال ٢ : أوجد الوسيط للقيم الآتية : 84 , 87 , 76 , 82 , 80 , 88

الحل : ترتب البيانات ترتيباً تصاعدياً 76 , 80 , 82 , 84 , 87 , 88

تحديد قيمة n : n = 6 (عدد زوجي)

الوسيط هو القيمة التي ترتيبها $\frac{n}{2}, \frac{n}{2}+1 \quad \frac{6}{2}, \frac{6}{2}+1 =$

$= X_3 , X_4 = 82 , 84$

$Me = \frac{82+84}{2} = 83$

ب – في حالة البيانات المبوبة :

إذا كانت $x_1 , x_2 , x_3 , ... , x_n$ تمثل مراكز فئات في توزيع تكراري عدد فئاته m فئة ، وكانت $f_1 , f_2 , f_3 , ... , f_n$ تمثل التكرارات المقابلة لتلك الفئات ، وافرض أن $F_1 , F_2 , ... , F_n$ تمثل التكرارات المتجمعة الصاعدة المقابلة للحدود العليا لفئات التوزيع ، ولتكن $\frac{\sum f_i}{2}$ تمثل ترتيب الوسيط في هذا التوزيع ، فإذا كان $F_1 < \frac{\sum f_i}{2} < F_2$

عندئذٍ يقال أن قيمة الوسيط هي الفئة التي تسلسلها هو K ، وسوف يتم حساب قيمة الوسيط وفق الصيغة الآتية :

$$M_e = L_{.}L_{.} + \frac{\frac{\sum f_i}{2} - F_1\uparrow}{f_2} (L_{.})$$

حيث :

$L_{.}L_{.}$: الحد الأدنى الحقيقي للفئة الوسيطية .

$\frac{\sum f_i}{2}$: مجموع التكرارات مقسوم على 2 (ترتيب الوسيط)

$F_1\uparrow$: التكرار المتجمع الصاعد للفئة قبل الوسيطية .

f_2 : تكرار الفئة الوسيطية .

$L_{.}$: طول الفئة .

خطوات إيجاد الوسيط :

١- عمل جدول توزيع تكراري متجمع صاعد .

٢- نستخرج ترتيب الوسيط $= \frac{\sum f_i}{2}$

٣- تحديد الفئة الوسيطية : وهي الفئة التي يكون تكرارها المتجمع الصاعد يساوي ترتيب الوسيط أو أكبر منه .

٤- تطبيق القانون .

مثال : أوجد الوسيط لجدول التوزيع التكراري الآتي جبرياً وبيانياً :

	C	F_i
1	20 – 29	4
2	30 – 39	6
3	40 – 49	2
4	50 – 59	8
5	60 – 69	12
6	70 – 79	8
7	80 – 89	3

الحل :

١) إيجاد التكرار المتجمع الصاعد :

	C	f_i	$F_i\uparrow$
1	20 – 29	4	4
2	30 – 39	6	10
3	40 – 49	2	12
4	50 – 59	8	20
5	60 - 69	12	32
6	70 – 79	8	40
7	80 – 89	3	43
Σ		43	

٢) نستخرج ترتيب الوسيط $\frac{\sum f_i}{2} = \frac{43}{2} = 21.5$

٣) نحدد الفئة الوسطية وهي (60- 69)

٤) تطبيق القانون

$$Me = L.L. + \frac{\frac{\sum f_i}{2} - F_1\uparrow}{f_i} \times (L.)$$

$$= 59.5 + \frac{21.5-20}{12} \times (10)$$

$$= 59.5 + 1.25 = 60.75$$

خصائص الوسيط :

١) لا يتأثر اطلاقاً بالقيم المتطرفة .
٢) يمكن تعيينه هندسياً .
٣) يمكن حسابه في حالة التوزيعات التكرارية المفتوحة س طرف واحد أو طرفين .

H . W .

١) استخرج الوسط الحسابي المرجح للتوزيعات الآتية :

X_i	10	18	15	11	19	22	20
W_i	1	4	3	2	5	7	6

X_i	1	0	2	2	3	1	4
W_i	2	4	3	5	3	2	2

٢) في جدول التوزيع التكراري الآتي :
أ) رسم المدرج والمضلع التكراري لهذا التوزيع .
ب) أوجد الوسط الحسابي .
ج) أوجد الوسيط جبرياً وبيانياً .

	C	f_i
1	11 – 15	3
2	16 – 20	8
3	21 – 25	15
4	26 – 30	25
5	31 – 35	20
6	36 – 40	17
7	41 – 45	6
8	46 – 50	10

٣) الآتي توزيع تكراري لكميات الكهرباء المستهلكة (كيلو واط) من قبل 40 دار سكنية خلال شهر واحد

٤)

الاستهلاك (كيلو واط)	800 -	900 -	1000 -	1100 -	1200 -	فأكثر 1300
عدد الدور	7	8	15	5	3	2

جد الوسط والوسيط للكمية المستهلكة من الكهرباء ثم تحقق من ذلك بالرسم .

٥) أكمل جدول التوزيع التكراري أدناه ، وارسم المدرج والمضلع التكراري ، واستخرج الوسيط جبرياً وبيانياً .

	C	F_i	X
1	10 -	4	14.5
2		7	
3		10	
4		25	
5		18	
6		12	

٦) جد قيمة z التي تجعل قيمة الوسط الحسابي في المجموعة 20 , 25 , 28 , 31 , 37 , z مساوٍ الى 30 ، ثم استخرج الوسيط .

٤- المنوال (The Mode):

يعرف المنوال بأنه القيمة الأكثر تكراراً وشيوعاً بين مجموعة القيم ، ويرمز له بالرمز (Mo).

طرق حساب المنوال :

١- في حالة البيانات غير المبوبة :

لتكن x_n , ... , x_3 , x_2 , x_1 بيانات المتغير العشوائي x في توزيع تكراري قوامها(عددها)n من المفردات ، ولنفرض أن x_j قيمة من قيم هذه المجموعة لحظ أنها تكررت أكثر من غيرها، عندئذ وبحسب تعريف المنوال فإن x_j تمثل المنوال لهذه المجموعة.

أ- إذا لم تتكرر أية قيمة فلا يوجد منوال :

مثال : أوجد المنوال للقيم الآتية : 19 , 18 , 17 , 22 , 16 , 12 , 15

الحل : لا يوجد منوال لهذه القيم لعدم وجود قيمة مكررة فيها .

ب- إذا تكرر أحد القيم أكثر من غيرها فإن هناك منوال واحد .

مثال : أوجد المنوال للقيم الآتية : 19 , 18 , 15 , 22 , 16 , 12 , 15

الحل : المنوال (MO) = 15 لأنها القيمة الأكثر تكراراً .

ج- إذا كان لقيمتين من قيم التوزيع نفس العدد من التكرارات فإن هناك منوالين .

مثال : أوجد المنوال للقيم الآتية : 19 , 19 , 18 , 15 , 22 , 16 , 12 , 15

الحل : يوجد منوالان هما 19 , 15 لأن لهما نفس التكرارات .

٢-في حالة البيانات المبوبة :

إذا كانت $x_1 , x_2 , \dots , x_n$ تمثل مراكز فئات في توزيع تكراري عدد فئاته m فئة ، و $f_1 , f_2 , \dots , f_n$ هي تكرارات تلك الفئات **فإن المنوال يستخرج بإتباع الخطوات الآتية :**

١- تحديد الفئة المنوالية : وهي الفئة التي تقابل أعلى تكراراً بين الفئات .

٢- تحديد الفرق ما بين تكرار الفئة المنوالية والفئة السابقة لها ، وليكن f_1 .

٣- تحديد الفرق ما بين تكرار الفئة المنوالية والفئة اللاحقة لها ، وليكن f_2 .

٤-تطبيق العلاقة الآتية لإيجاد المنوال :

$$Mo = L_{.} L_{.} + \frac{f_1}{f_1 + f_2} (L_{.})$$

حيث $L_{.}L_{.}$ الحد الأدنى الحقيقي للفئة المنوالية ، $L_{.}$ طول الفئة

مثال : أوجد المنوال جبرياً وبيانياً لجدول التوزيع التكراري الآتي :

	C	f_i
1	20 – 29	2
2	30 – 39	4
3	40 – 49	7
4	50 – 59	10
5	60 – 69	8
6	70 – 79	5
7	80 – 89	3

الحل :

1) 50 – 59 الفئة المنوالية

2) $f_1 = 10 - 7 = 3$

3) $f_2 = 10 - 8 = 2$

4) $Mo = L.L. + \frac{f_1}{f_1 + f_2} (L.)$

$= 49.5 + \frac{3}{3+2} (10)$

$= 55.5$

خصائص المنوال :

١- لا يتأثر بالقيم المتطرفة (صغيرة أو كبيرة) التي تقع على جانبي التوزيع .

٢- يمكن إيجاده بسهولة لأنه القيمة الأكثر تكراراً بين القيم .

٣- يمكن ايجاده في التوزيعات التكرارية المفتوحة (غير المنتهية) والجداول المفتوحة .

العلاقة بين مقاييس النزعة المركزية :

تحدد العلاقة بالقانون الآتي :

$$\bar{X} - Mo = 3 (\bar{X} - Me)$$

مثال : في احد التوزيعات التكرارية إذا كانت قيمة Mo=17 و $\bar{X}$ = 12 ، جد Me

الحل :

$$\bar{X} - Mo = 3 (\bar{X} - Me)$$

$12 - 17 = 3 (12 - 3Me)$

$-5 = 36 - 3Me$

$3Me = 36 + 5$

$3Me = 41$

$Me = \frac{41}{3}$

تمارين : H . W .

١) إذا كانت البيانات الآتية تمثل درجات 30 طالب في امتحان مادة الإحصاء

80 , 84 , 48 , 77 , 65 , 71 , 25 , 55 , 67 , 65 , 88 , 35 , 27 , 31 , 44

88 , 66 , 45 , 78 , 79 , 74 , 70 , 44 , 56 , 50 , 72 , 58 , 67 , 55 , 78

المطلوب :

أ) اعرض البيانات في جدول توزيع تكراري .

ب) ارسم المدرج التكراري والمضلع التكراري على نفس الرسم .

ج) استخراج الوسط الحسابي والوسيط والمنوال لهذه البيانات .

٧) أوجد المنوال جبرياً وبيانياً لجدول التوزيع التكراري الآتي :

	C	f_i
1	11 - 15	3
2	16 - 20	8
3	21 - 25	15
4	26 - 30	25
5	31 - 35	20
6	36 - 40	17

٣) اكمل جدول التوزيع التكراري الآتي ، ثم احسب : والوسيط والمنوال جبرياً وبيانياً :

	C	f_i	X
1		3	
2	28 -	8	30
3		15	
4		25	
5		20	

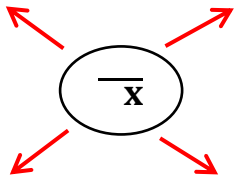
٤) إذا كانت Mw = 10 ، Mo = 44 فما قيمة

الفصل الثالث

مقاييس التشتت

وردت كلمة تشتت في الأدبيات الأحصائية بمعنى مبعثر (Scattered) أي انتشار قيم مجموعة من البيانات .

ويعرف التشتت بأنه الدرجة التي تتجه بها البيانات الكمية للتباعد والانتشار عن بعضها البعض ، أو حول قيمة متوسطة وغالباً ما تكون أحد مقاييس النزعة المركزية (الوسط و الوسيط) .



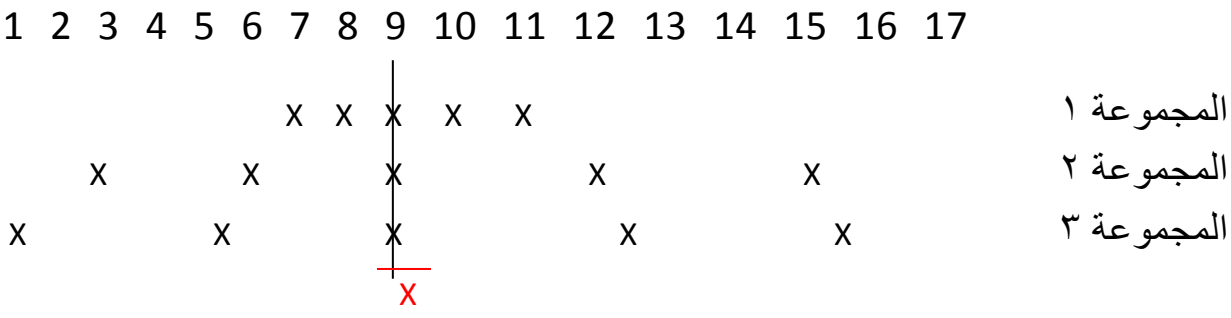
وإن الهدف من دراسة التشتت هو تكوين فكرة عن مدى تجانس قيم مجموعة من المفردات ، وهذا يعني أن دراسة التشتت أمر مفيد في اجراء المقارنة بين قيم مجموعة أو أكثرمن البيانات عن ظاهرة معينة .

فعلى سبيل المثال : إن الوسط الحسابي للمجاميع الآتية هو (٩)

المجموعة ١ : ٧ ، ٨ ، ٩ ، ١٠ ، ١١ المجموعة ٢ : ٣ ، ٦ ، ٩ ، ١٢ ، ١٥

المجموعة ٣ : ١ ، ٥ ، ٩ ، ١٣ ، ١٧

نلاحظ أن المجموعة الأولى أكثر تجانساً (أقل انتشاراً) من والمجموعتين الثانية والثالثة كذلك فإن المجموعة الثانية أكثر تجانساً من المجموعة الثالثة ، كما موضح في الشكل :



نلاحظ من المثال إنه يمكن اجراء المقارنة بين المجموعات الثلاث بعد أن كانت المقارنة بينهما على أساس الوسط الحسابي عديمة الفائدة .

مثال : كانت درجات الحرارة لثلاثة من المحافظات الشمالية هي كالآتي :

المحافظة الأولى : -٢ ، -١ ، ٠ ، ١ ، ٢ ، ٣ ، ٤

المحافظة الثانية : -٣ ، -١ ، ٠ ، ١ ، ٢ ، ٤ ، ٧

المحافظة الثالثة : -١ ، ٠ ، ٠ ، ١ ، ٢ ، ٢ ، ٣

في أية محافظة كان الجو أكثر استقراراً ؟

مقاييس التشتت

تعرف مقاييس التشتت بأنها المقاييس التي تبحث في مقدار الاختلافات بين البيانات وهي على الأنواع الآتية :

١- المدى:

هو الفرق بين اعلى قيمة واقل قيمة في تلك المجموعة ويرمز له بالرمز R .

مثال/جد المدى للبيانات الاتية:- 8,3,5,2,4

R=8-2=6

٢-الانحراف المتوسط :- ويرمز له بالرمز M.D

طريقة حسابهد

أ- الانحراف المتوسط للبيانات غير مبوبة:-

نستخدم القانون الاتي:-

$$M.D=\frac{\sum |xi-\bar{x}|}{n}$$

حيث نتبع الخطوات الاتية لايجاد الانحراف المتوسط:-

١-نجد الوسط الحسابي لبيانات.

٢-نجد الفرق بين كل قيمة والوسط الحسابي.

٣-نجد القيمة المطلقة للفروق ثم نجد المجموع.

جد الانحراف المتوسط للبيانات الاتية:-

Xi	$Xi-x-\bar{x}$	|
3	-11	1
5		1
6	2	2
2	-22	2

$\bar{x}=\frac{16}{4}=4$

$M.D=\frac{6}{4}$

=1.5

ب- الانحراف المتوسط للبيانات المبوبة:-

نستخدم القانون الأتي:-

$$M.D=\frac{\sum f i\,|\bar{x}i-x|}{\sum fi}$$

مثال/ اوجد الانحراف المتوسط لجدول الاتي:-

C	fi	Xi	fixi	$\|Xi-\bar{x}\|$	$fi\|xi-\bar{x}\|$
10- 14	1	12	12	-9	9
15- 19	7	17	119	-4	28
20- 24	8	22	176	1	8
25- 29	3	27	81	6	18
30-34	1	32	32	11	11

$$\bar{X}=\frac{\sum fixi}{\sum fi}$$

$$= \frac{420}{20}$$

=21

$$M.D= \frac{74}{20}= 3.7$$

٣- الانحراف المعياري (القياسي) :

ويسمى في بعض الأحيان بالانحراف القياسي ، ويعد أفضل مقاييس التشتت على الاطلاق .

ويعرف بأنه الجذر التربيعي الموجب لمتوسط مجموع مربعات انحرافات قيم المتغير العشوائي عن وسطها الحسابي ، ويرمز له بالرمز s في حالة العينة وبالرمز ٦ في حالة المجتمع .

٤- التباين :

ويعرف بأنه متوسط مجموع مربعات انحرافات قيم المتغير العشوائي عن وسطها الحسابي ، ويرمز له بالرمز S^2 في حالة العينة وبالرمز $٦^٢$ في حالة المجتمع .

أي إنه مربع الانحراف المعياري .

طرق حساب الانحراف المعياري والتباين :

١) الانحراف المعياري والتباين للبيانات غير المبوبة :

لحساب الانحراف المعياري والتباين للعينة نستخدم القوانين الآتية :

الطريقة المطولة

الانحراف المعياري $S=\sqrt{\frac{\sum(xi-\bar{x})^2}{n-1}}$

التباين $S^2=\frac{\sum(xi-\bar{x})^2}{n-1}$

الطريقة المختصرة

الانحراف المعياري $S=\sqrt{\frac{\sum xi^2-\frac{(\sum xi)^2}{n}}{n-1}}$

التباين $S^2=\frac{\sum xi^2-\frac{(\sum xi)^2}{n}}{n-1}$

اما لحساب الانحراف المعياري والتباين للمجتمع فنستخدم القوانين الآتية :

الطريقة المطولة

الانحراف المعياري $6=\sqrt{\frac{\sum(xi-\bar{x})^2}{n}}$

التباين $6^2=\frac{\sum(xi-x)^2}{n}$

الطريقة المختصرة

الانحراف المعياري $6=\sqrt{\frac{\sum xi^2-\frac{(\sum xi)^2}{n}}{n}}$

التباين $6^2=\frac{\sum xi^2-\frac{(\sum xi)^2}{n}}{n}$

ملاحظة : **يتم استخدام الانحراف المعياري للمجتمع في أغلب الاحيان .**

مثال : البيانات الآتية ٧ ، ٥ ، ٦ ، ٨ ، ٩ تبين كمية محصول القطن في خمس مزارع احسب الانحراف المعياري والتباين لها .

الحل بالطريقة الاولى

x_i	$(x_i - x)$	$(x_i - x)^2$
9	2	4
8	1	1
6	-1	1
5	-2	4
7	0	0
		10

$$\overline{X} = \frac{35}{5} = 7$$

$$S=\sqrt{\frac{\sum(xi-\bar{x})^2}{n-1}} = \sqrt{\frac{10}{5-1}}$$

$$=\sqrt{\frac{10}{4}} \quad =\sqrt{2.5} \quad = \quad 1.58$$

$$S^2 \quad = \quad 2.5$$

الحل بالطريقة الثانية

X_i	X_i^2
٩	٨١
٨	٦٤
٦	٣٦
٥	٢٥
٧	٤٩
	٢٥٥

$$S=\sqrt{\frac{\sum xi^2-\frac{(\sum xi)^2}{n}}{n-1}} = \sqrt{\frac{255-\frac{(35)^2}{5}}{5-1}} = \sqrt{\frac{255-245}{4}} = \sqrt{2.5}$$

$$S^2 \quad = \quad 2.5$$

نلاحظ أن الناتج هو نفسه بإستخدام الطريقتين

٢) الانحراف المعياري والتباين للبيانات المبوبة :

لحساب الانحراف المعياري والتباين للعينة نستخدم القوانين الآتية :

الطريقة المطولة

$$S=\sqrt{\frac{\sum fi\,(xi-\bar{x})^2}{\sum fi-1}}$$ الانحراف المعياري

$$S^2=\frac{\sum fi\,(xi-\bar{x})^2}{\sum fi-1}$$ التباين

الطريقة المختصرة

$$S=\sqrt{\frac{\sum fix^2 \; -\frac{(\sum fixi)^2}{\sum fi}}{\sum fi-1}}$$ الانحراف المعياري

$$S^2 = \frac{\sum fixi^{\,2}-\frac{(\sum fixi)^2}{\sum fi}}{\sum fi-1}$$ التباين

اما لحساب الانحراف المعياري والتباين للمجتمع فنستخدم القوانين الآتية :

الطريقة المطولة

الانحراف المعياري $6=\sqrt{\frac{\sum fi(xi-\bar{x})^2}{\sum fi}}$

التباين $6^2=\frac{\sum fi(xi-\bar{x})^2}{\sum fi-1}$

الطريقة المختصرة

الانحراف المعياري $6=\sqrt{\frac{\sum fixi^2-\frac{(\sum fixi)^2}{\sum fi}}{\sum fi}}$

التباين $6^2=\frac{\sum fixi^2-\frac{(\sum fixi)^2}{\sum fi}}{\sum fi}$

مثال : احسب الانحراف المعياري لجدول التوزيع التكراري الاتي :

C	fi	Xi	fi xi	$(xi-\bar{x})$	$(xi-\bar{x})^2$	Fi $(xi-\bar{x})^2$
50 – 60	5	52.5	262.5	-18.7	349.69	1748.45
60 – 65	14	62.5	875	-8.7	75.69	1059.66
70 – 75	22	72.5	1595	1.3	1.69	37.18
80 – 85	9	82.5	742.5	11.3	127.69	1149.51
90 – 95	4	92.5	370	21.3	453.69	1814.76
	54		3845			5809.56

$\bar{X}=\frac{\sum fixi}{\sum fi}=\frac{3845}{54}=71.2$

$6=\sqrt{\frac{\sum fi(xi-\bar{x})^2}{\sum fi}}=\sqrt{\frac{5809.56}{54}}$

$6^2=\frac{5809.56}{54}=107.58$

تمارين :

١) استخرج الانحراف المعياري والتباين للبيانات الآتية : 12, 4,33, 23, 21, 15, 14, 21,1113 ,22,31,15, باستخدام الطريقتين المختصرة و المطولة .

٢) اكمل جدول التوزيع التكراري الآتي ، ثم احسب الانحراف المعياري والتباين :

	C	f_i	X
1		3	
2	5-	8	6
3		15	
4		25	
5		20	

الفصل الرابع

مقاييس الارتباط

يعرف الارتباط بأنه درجة أو قيمة العلاقة بين التي تربط متغيرين أو أكثر مع بعضهما البعض ، وهي قيمة حقيقية خالية من وحدات قياس المتغيرات المرتبطة بالعلاقة .

وهناك عدة أنواع من الارتباط يختص كل منها بنوع معين من العلاقات ، وستقتصر دراستنا على معاملي ارتباط بيرسون للمتغيرات الكمية ، ومعامل ارتباط الرتب لسبيرمان الخاص بالمتغيرات الوصفية .

خواص معامل الارتباط :

١) تتراوح قيمة معامل الارتباط بين (-١،١) أي أن : $-1 \le r \le 1$

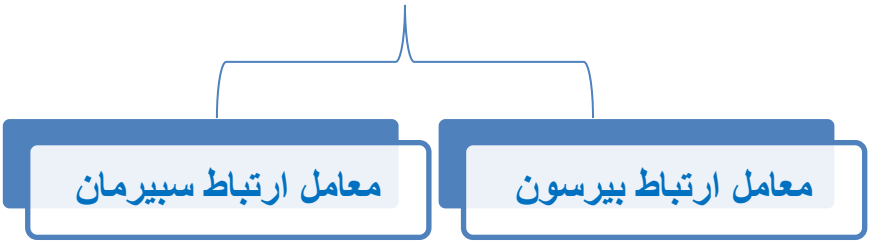
٢) إذا كان r=1 فهذا يعني وجود علاقة موجبة تامة طردية بين المتغيرين ، أي ان زيادة في قيم احد المتغيرين يصحبه زيادة في قيم المتغير الآخر . (ترابط طردي)

٣) إذا كان r= -1 فهذا يعني وجود علاقة سالبة تامة عكسية بين المتغيرين ، أي ان زيادة في قيم احد المتغيرين يصحبه نقصان في قيم المتغير الآخر . (ترابط عكسي)

٤) إذا كانت r = 0 فهذا يعني عدم وجود ترابط خطي بينهما ، أي لا توجد علاقة بينهما .

(لا يوجد ترابط)

أنواع معاملات الأرتباط



١) معامل الارتباط الخطي البسيط (معامل ارتباط بيرسون) :

ويعرف بأنه القيمة العددية للعلاقة بين متغيرين كميين ، ويعد العالم الإنجليزي كارل بيرسون أول من وضع صيغة لهذا المعامل . **ونرمز له بالرمز r) أو $r_{x,y}$ أو $r_{1,2}$).**

حساب قيمة معامل ارتباط بيرسون :

لتكن (x_i , y_i) ; I = 1 , 2 , 3 , ... , n تمثل أزواج القيم التي يتم الحصول عليها من عينة من البيانات عددها n من البيانات ، عندئذ نستخدم القانون اللآتي لاستخراجه :

$$r = \frac{\sum x_i y_i - \frac{\sum x_i \sum y_i}{n}}{\sqrt{\left(\sum x_i\, 2 - \frac{(\sum x_i)^2}{n}\right)\left(\sum y_i\, 2 - \frac{(\sum y_i)^2}{n}\right)}}$$

مثال : احسب معامل ارتباط بيرسون للبيانات الاتية والتي تمثل الدرجات الفصلية ٦ ، ٧ ، ١٠ ، ٨ ، ٥ والدرجات النهائية ٦ ، ٨ ، ٦ ، ٧ ، ٨ لخمسة طلاب .

X_i	y_i	Y_i^2	X_i^2	x_iy_i
7	8	49	64	56
6	7	36	49	42
5	6	25	36	30
8	8	64	64	64
10	6	100	36	60

$$r = \frac{\sum x_i y_i - \frac{\sum x_i \sum y_i}{n}}{\sqrt{\left(\sum x_i\, 2 - \frac{(\sum x_i)^2}{n}\right)\left(\sum y_i\, 2 - \frac{(\sum y_i)^2}{n}\right)}}$$

$$= \frac{252 - \frac{36 \times 35}{5}}{\sqrt{\left(274 - \frac{(36 \times 36)}{5}\right)\left(249 - \frac{(35 \times 35)}{5}\right)}}$$

$$= \frac{252 - 252}{\sqrt{(274 - 259.2)(249 - 245)}}$$

$$= \frac{0}{\sqrt{14.8 \times 4}} = 0$$ **لا يوجد ارتباط**

٢) معامل ارتباط سبيرمان للرتب :

ويعرف بأنه القيمة العددية للعلاقة بين متغيرين وصفيين . (ونرمز له بالرمز r_s).

ويقيس معامل ارتباط الرتب التغير الاقتراني القائم بين ترتيب الافراد بالنسبة لصفة معينة ، ففي بعض الأحيان يمكن وصف مركز الفرد في جماعته عن طريق ترتيبه بينهم في سمة معينة .

حساب قيمة معامل ارتباط سبيرمان للرتب :

لنفرض أن x , y متغيرين من النوع الوصفي ، ولنفرض أن البيانات التي جمعت عن المتغيرين x , y عددها n هي بهيئة صفات غير قابلة للقياس الكمي ، فلو فرضنا أن $(x_i , y_i , \ldots, n؛I=1,2,3)$ ممكنة الترتيب تصاعدياً أو تنازلياً وفق معيار معين يمتاز به كل متغير (مثلاً درجات مجموعة من الطلبة يمكن ترتيبها تصاعدياً على أساس معيار الأقل الى الأعلى درجة أي ضعيف ، مقبول ، متوسط ، جيد ، جيد جداً ، ممتاز وبالعكس أي الترتيب التنازلي . عندئذ و استناداً إلى هذا التعريف إذا تم تخصيص سلسلة الأعداد الطبيعية ١ ، ٢ ، ٣ ، ... ، للتعبير عن صفات الترتيب لكل متغير بحيث أن كل صفة يخصص لها أحد أعداد هذه السلسلة (مثلاً : ضعيف يخصص لها ١ ، ومتوسط يخصص لها ٣ ، وهكذا) في حالة عدم تكرار قيمة أي صفة من صفات المتغير ، عندها يستخدم القانون الآتي لإيجاد معامل الارتباط بين المتغيرين x , y :

$$R_S = 1 - \frac{6\sum(R_x - R_y)^2}{n(n^2 - 1)}$$

مثال ١ : كانت تقديرات درجات ستة طلبة في مادتي الاحصاء والرياضيات هي كالآتي :

الإحصاء : جيد ، متوسط ، ضعيف ، مقبول ، جيد جداً ، ممتاز

الرياضيات : متوسط ، جيد ، مقبول ، ضعيف ، ممتاز ، جيد جداً

أوجد معامل الارتباط بين تقديرات درجات الطلاب في امتحاني الاحصاء والرياضيات (هل توجد علاقة ارتباطية بين تقديرات درجات الطلاب في امتحاني الاحصاء والرياضيات ؟ وما نوعها ؟)

الحل :

ترتب التقديرات وفق ترتيب تصاعدي أو تنازلي ، وليكن الترتيب تصاعدي وتخصص لها سلسلة الأعداد الطبيعية .

التقديرات : ضعيف ، مقبول ، متوسط ، جيد ، جيد جداً ، ممتاز

١ ٢ ٣ ٤ ٥ ٦

ثم يعوض عن كل تقدير بما يقابله من اعداد طبيعية

x الاحصاء	y الرياضيات	R_x	R_y	$d_i = R_x - R_y$	$d_i^2 = (R_x - R_y)^2$
جيد	متوسط	4	3	1	1
متوسط	جيد	3	4	-1	1
ضعيف	مقبول	1	2	-1	1
مقبول	ضعيف	2	1	1	1
جيد جداً	ممتاز	5	6	-1	1
ممتاز	جيد جداً	6	5	1	1
					6

$$R_S = 1 - \frac{6\sum(R_x - R_y)^2}{n(n^2 - 1)}$$

$$= 1 - \frac{(6)6}{6(35)} = 0.8$$

علاقة موجبة طرية جيدة جداً بين درجات الطلاب

مثال ٢ : أوجد معامل ارتباط سبيرمان للرتب بين قيم المتغيرين في الجدول أدناه :

N	X	Y
١	٦	٩
٢	٨	١٠
٣	١٠	٧
٤	٧	٦
٥	٩	٨

الحل :

نرتب قيم المتغيرين تصاعدياً ونعبر عنها بسلسلة الأعداد الطبيعية ١ ، ٢ ، ٣ ، ...

القيم : ٦ ، ٧ ، ٨ ، ٩ ، ١٠

١ ، ٢ ، ٣ ، ٤ ، ٥

N	X	Y	R_x	R_y	$(R_x - R_y)$	$(R_X - R_y)^2$
١	٦	٩	١	٤	٣ −	٩
٢	٨	١٠	٣	٥	٢ −	٤
٣	١٠	٧	٥	٢	٣	٩
٤	٧	٦	٢	١	١	١
٥	٩	٨	٤	٣	١	١
						٢٤

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

$$= 1 - \frac{6 \times 24}{5(5^2 - 1)}$$

$$= 1 - \frac{144}{5 \times 24}$$

$$= 1 - \frac{144}{120}$$

= - 20 علاقة موجبة طردية ضعيفة

تمت المادة بعون الله