المشرف على مقرر

الإحصاء الاجتماعي

د. سعيد سيف الدين

جامعة الملك فيصل
عمادة التعلم الإلكتروني والتعليم عن بعد

كلية الآداب

المحاضرة المباشرة الأولي

مراجعة على المحاضرات الخمس الأولي مع حل التمارين الخاصة بها في امتحانات سابقة

المحاضرة الأولى : مقدمة في علم الإحصاء
المحاضرة الثانية : تبويب وعرض البيانات الإحصائية
المحاضرة الثالثة : مقاييس النزعة المركزية (الوسط الحسابي – الوسيط – المنوال)
المحاضرة الرابعة : مقاييس التشتت (المدى – الانحراف المتوسط – الانحراف المعياري)
المحاضرة الخامسة : الفروض الإحصائية

(1) البيانات

هي مجموعة من " المشاهدات أو القياسات" التي تخص الظاهرة تحت الدراسة ، والكمية التي نقوم بمشاهدتها أو قياسها تُسمى بالمتغير

أمثلة :

مثال	العملية الإحصائية هي دراسة :	البيانات (القياسات أو المشاهدات)	المتغير
(1)	لون العين لبعض الأطفال حديثي الولادة	أخضر – أزرق – بني –	لون العين
(2)	عدد الطلاب في فصول مدرسة	15 – 18 – 20 – 25 – 17 –	عدد الطلاب
(3)	أطوال مجموعة من الطلاب في فصل ما (بالمتر)	1.5 – 1.52 – 1.71 – 1.83 –	طول الطالب
(4)	أوزان بعض العاملات بمصنع معين (بالكيلوجرام)	55.2 – 60.1 – 63.35 –.....	وزن العاملة
(5)	تقديرات عدد من الطلاب في مقرر الإحصاء	A – B – C – D – F –	تقدير الطالب

والمتغير إما أن يكون متغير نوعياً (أي لايمكن التعبير عنه بعدد) كما في حالة الأمثلة (1) ، (5) أو أن يكون كمياً (أي يوصف بعدد) كما في الأمثلة (2) ، (3) ، (4) . وإذا أخذ المتغير الكمي قيماً محددة دون أن يأخذ القيم الأخرى بينها سمي متغير كمي متقطع (أو منفصل) [كما في حالة مثال (2)] ، أما إذا كان المتغير الكمي يمكن أن يأخذ كل القيم بين قيمتين معينتين فإنه يُسمى متغير كمي متصل [كما في حالة الأمثلة (3) ، (4)] .

وتُسمى البيانات باسم نوع المتغير ، بمعنى أنها تكون بيانات نوعية عندما تكون المتغيرات نوعية وتكون كمية متقطعة (أو متصلة) عندما تكون المتغيرات كمية متقطعة (أو متصلة) . وإذا كانت البيانات نوعية أو كمية متقطعة فإنها نُعرف بأنها بيانات غير مبوبة ، أما إذا كانت بيانات كمية متصلة فإنها تُعرف بأنها بيانات مبوبة .

(2) المتغير المستقل والمتغير التابع

عند دراسة العلاقة بين متغيرين ، فالمتغير الذي يخضع للتحكم والسيطرة يُسمى بالمتغير المستقل والذي بتغير قيمه تتغير قيم المتغير الآخر الذي يُسمى بالمتغير التابع .

أسئلة اختبارات

1. تقديرات الطلاب (A+، A، B+ ، B) تمثل:
   (أ) بيانات كمية
   (ب) بيانات كمية متصلة
   (ج) بيانات كمية منفصلة
   ✓ (د) بيانات نوعية

2. جنسية الطلاب (مصري، سعودي، بحريني) تمثل:
   (أ) بيانات كمية
   ✓ (ب) بيانات نوعية
   (ج) بيانات كمية متصلة
   (د) بيانات كمية منفصلة

3. أراد باحث أن يدرس أثر التدريس في الفصول الافتراضية على تحصيل الطلاب في الاحصاء الاجتماعي. فإن المتغير المستقل هو:
   (أ) الطلاب
   ✓ (ب) التدريس في الفصول الافتراضية
   (ج) التحصيل
   (د) الاحصاء الاجتماعي

4. أراد باحث أن يدرس أثر عدد ساعات الدراسة على تحصيل الطلاب في الاحصاء . فإن المتغير التابع هو:
   (أ) الباحث
   (ب) الطلاب
   (ج) عدد ساعات الدراسة
   ✓ (د) تحصيل الطلاب في الإحصاء

5. عدد حبات البطيخ التي يبيعها أحد المحلات في أسبوع هو :

(أ) متغير نوعي
(ب) متغير كمي متقطع ✓
(ج) متغير كمي متصل
(د) خلاف ذلك

6. وزن البطاطس الذي تنتجه إحدى المزارع في شهر هو :

(أ) متغير نوعي
(ب) متغير كمي متقطع
(ج) متغير كمي متصل ✓
(د) خلاف ذلك

7. البيانات المجمعة عن درجات الحرارة ساعة الظهيرة (مقربة لأقرب عدد صحيح) في عدد من مدن المملكة هي :

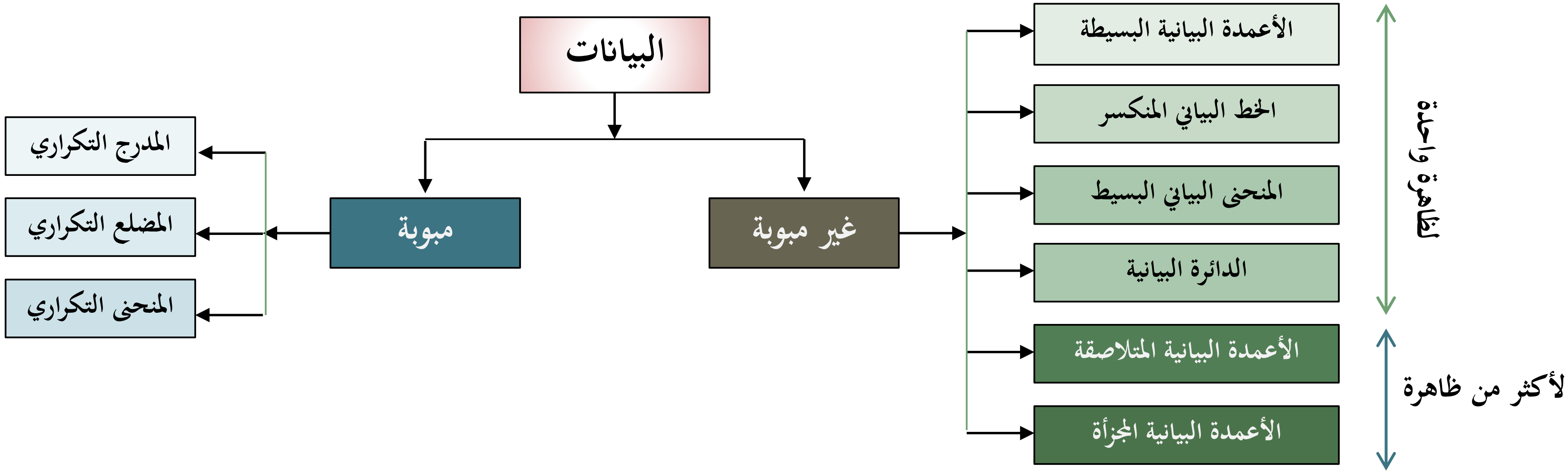
(أ) بيانات نوعية
(ب) بيانات كمية متقطعة ✓
(ج) بيانات كمية متصلة
(د) خلاف ذلك

8. البيانات المجمعة عن النسبة المئوية للمجموع الكلي لطلبة الثانوية العامة في المملكة هي :

(أ) بيانات نوعية
(ب) بيانات كمية متقطعة
(ج) بيانات كمية متصلة ✓
(د) خلاف ذلك

9. البيانات المجمعة عن الحالة الاجتماعية لسكان منطقة معينة هي

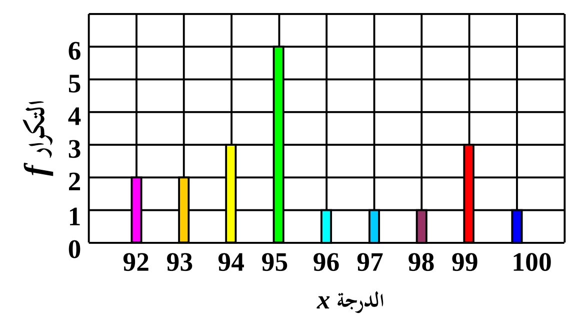
(أ) بيانات نوعية ✓
(ب) بيانات كمية متقطعة
(ج) بيانات كمية متصلة
(د) خلاف ذلك



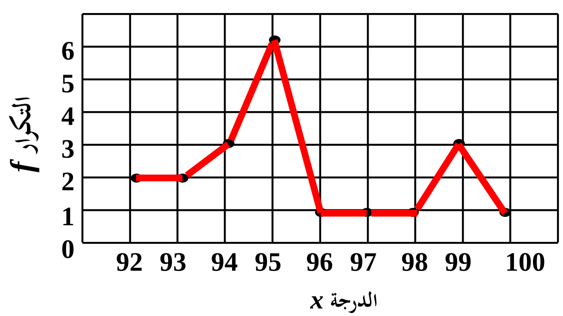
(1) عرض البيانات غير المبوبة (أي الكمية المنفصلة أو النوعية) لظاهرة واحدة

مثال : الجدول المقابل يبين التوزيع التكراري لدرجات 20 طالبة في مقررالإحصاء الاجتماعي . هذه البيانات يمكن تمثيلها بيانياً بطرق عديدة منها :

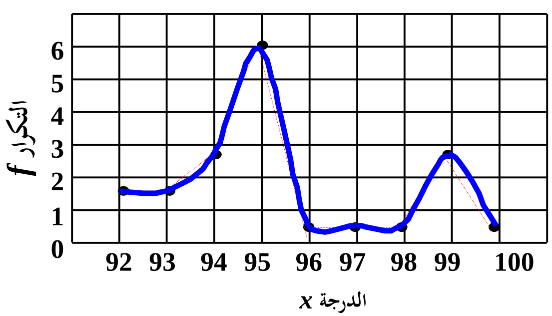
الدرجة x	92	93	94	95	96	97	98	99	100
التكرار f	2	2	3	6	1	1	1	3	1



طريقة الأعمدة البيانية البسيطة

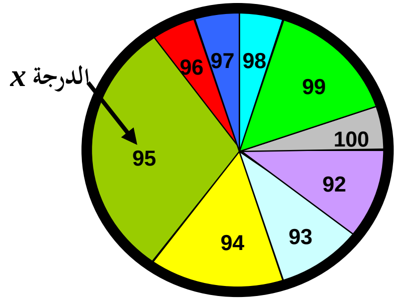


طريقة الخط البياني المنكسر



طريقة المنحنى البياني البسيط

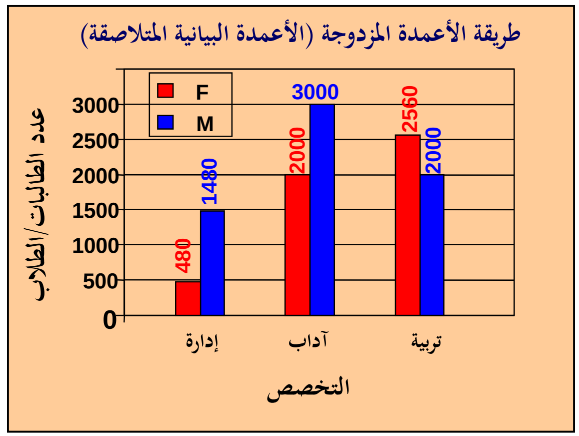
طريقة الدائرة البيانية



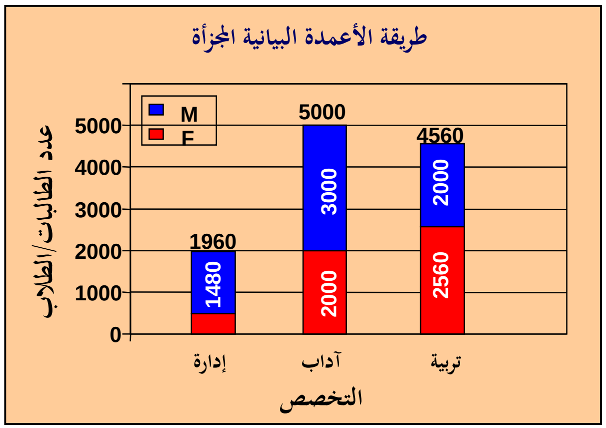
(2) عرض البيانات غير المبوبة لظاهرتين

مثال : الجدول المقابل يبين عدد الطالبات F وعدد الطلاب M في تخصصات إدارة أعمال وآداب وتربية خاصة المنضمين لبرنامج الانتساب المطور بجامعة الملك فيصل في أحد الفصول الدراسية . هذه البيانات يمكن تمثيلها بطرق عديدة منها :

	طالبات F	طلاب M
إدارة أعمال	480	1480
آداب	2000	3000
تربية خاصة	2560	2000



أي أن كل تخصص يُمثل بعمود مزدوج مكون من عمودين بسيطين متلاصقين

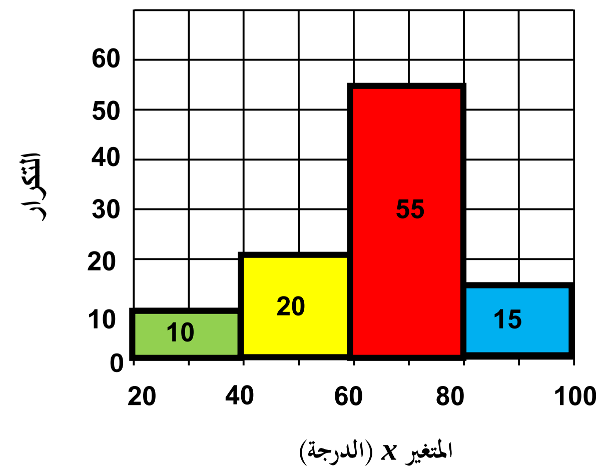


أي أن كل تخصص يُمثل بعمود طوله يُعبر عن مجموع عدد طالباته وطلابه معاً ثم يتم تجزئته إلى عمودين كل منهما يمثل فئة من الفئات

(3) عرض البيانات المبوبة (أي البيانات الكمية المتصلة) لظاهرة

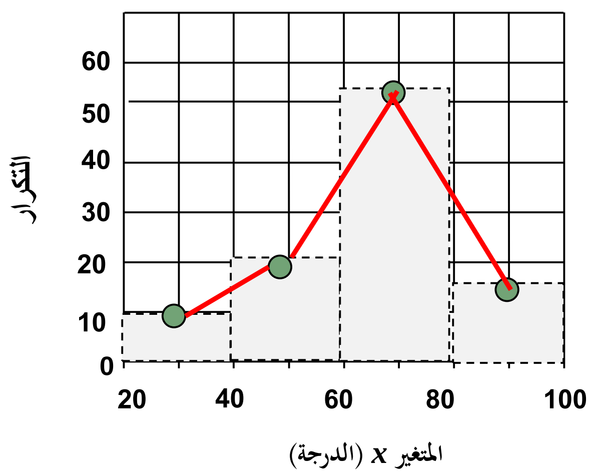
مثال : الجدول المقابل يبين التوزيع التكراري لدرجات 100 طالب في اللغة الإنجليزية بأحد فصول مدرسة ثانوية (الدرجة العظمى 100) . هذه البيانات يمكن تمثيلها بطرق عديدة منها :

الجدول التكراري	
المتغير x	التكرار
$20 \le x < 40$	10
$40 \le x < 60$	20
$60 \le x < 80$	55
$80 \le x < 100$	15
	$\sum f = 100$

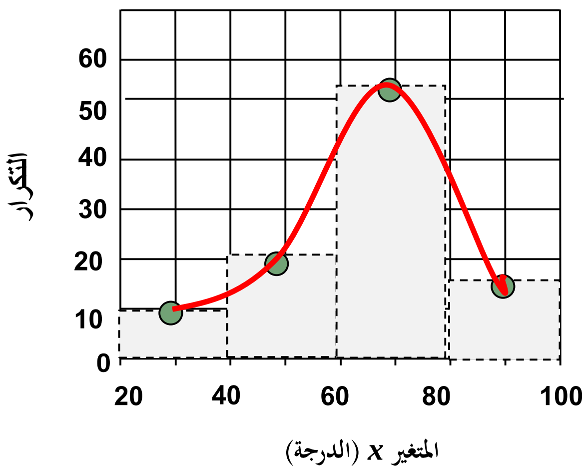


المدرج التكراري

لا فراغات موجودة بين المستطيلات [حيث أن البيانات هنا بيانات متصلة] بخلاف طريقة الأعمدة في حالة البيانات غير المبوبة حيث يجب ألا تكون الأعمدة متلاصقة .



المضلع التكراري



المنحنى التكراري

أسئلة اختبارات

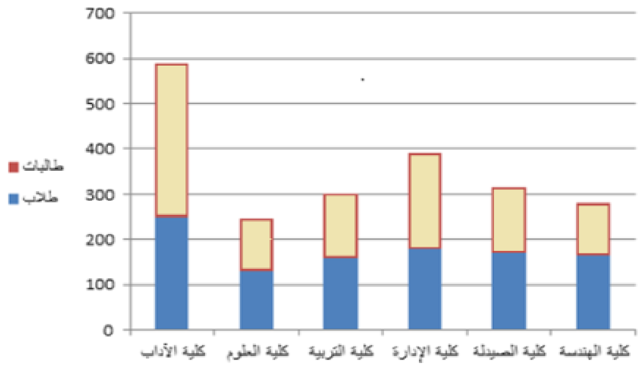
1. الشكل المقابل يبين طريقة .... لتمثيل البيانات بيانياً:

(أ) الأعمدة البيانية البسيطة

(ب) الأعمدة البيانية المتلاصقة

(ج) الأعمدة البيانية المجزأة ✓

(د) المضلع التكراري



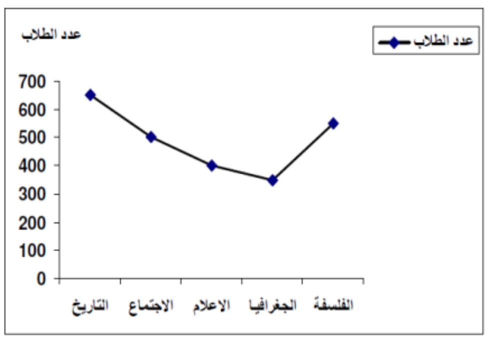
2. تمثيل البيانات في الشكل المقابل تسمى:

(أ) الأعمدة البيانية المجزأة

(ب) المنحنى البياني البسيط

(ج) المضلع التكراري

(د) الخط البياني المنكسر ✓



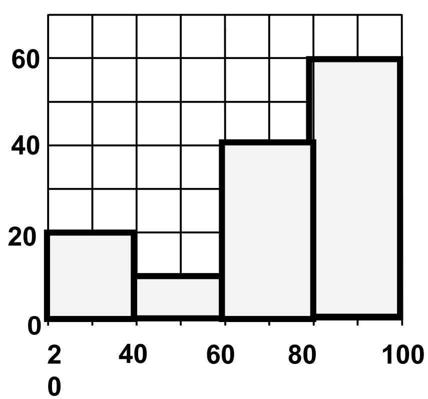
3. تمثيل البيانات في الشكل المقابل تسمى:

(أ) الخط البياني المنكسر

(ب) المدرج التكراري ✓

(ج) المنحنى التكراري

(د) المضلع التكراري



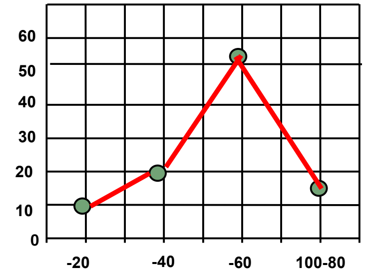
4. تمثيل البيانات في الشكل المقابل تسمى:

(أ) الخط البياني المنكسر

(ب) المدرج التكراري

(ج) المنحنى التكراري

(د) المضلع التكراري ✓



5. الشكل التالي يستخدم مع:

(أ) البيانات المبوبة

(ب) البيانات غير المبوبة ✓

(ج) البيانات النوعية فقط

(د) كل ما سبق

6. تمثيل البيانات في الشكل المقابل تسمى:

(أ) الخط البياني المنكسر

(ب) المدرج التكراري

(ج) المنحنى التكراري ✓

(د) المضلع التكراري

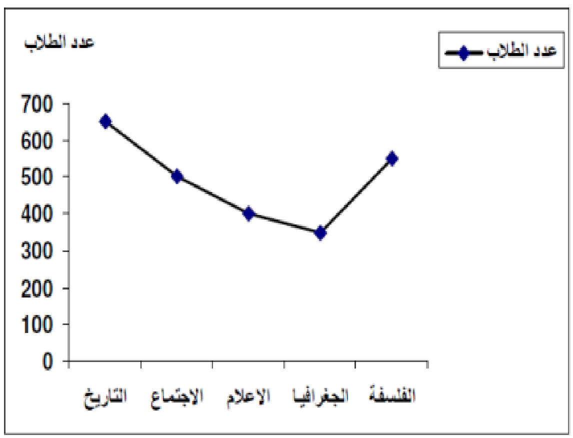
7. تمثيل البيانات في الشكل التالي تسمى:

(أ) الأعمدة البيانية المجزأة

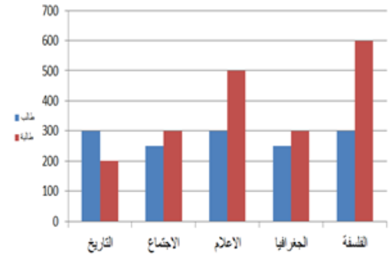
(ب) المنحنى البياني البسيط

(ج) المضلع التكراري

(د) الخط البياني المنكسر ✓

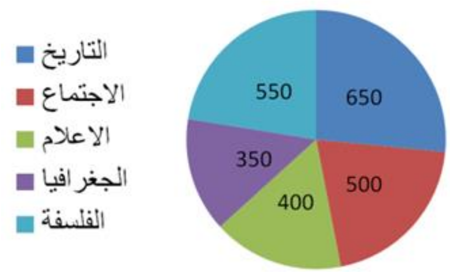


8. تمثيل البيانات في الشكل المقابل تسمى:



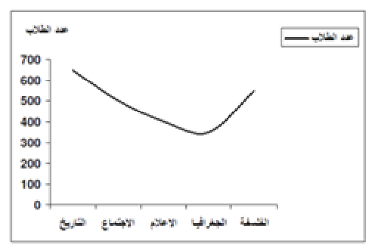
(أ) المضلع التكراري
(ب) الأعمدة البيانية المجزأة
(ج) الأعمدة البيانية المتلاصقة ✓
(د) الأعمدة البيانية البسيطة

9. الشكل المقابل يستخدم مع:



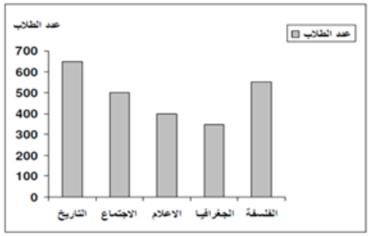
(أ) البيانات المبوبة
(ب) البيانات غير المبوبة ✓
(ج) البيانات النوعية فقط
(د) البيانات الاسمية

10. تمثيل البيانات في الشكل المقابل تسمى:



(أ) الخط البياني المنكسر
(ب) المضلع التكراري
(ج) المنحنى البياني البسيط ✓
(د) الأعمدة البيانية المجزأة

11. الشكل المقابل يستخدم مع:



(أ) البيانات المبوبة
(ب) البيانات غير المبوبة ✓
(ج) البيانات النوعية فقط
(د) البيانات الاسمية

12. كل ما يلي من أنواع العرض البياني للبيانات غير المبوبة <u>فيما عدا واحدة</u> هي:

(أ) الخط البياني المنكسر
(ب) الدائرة البيانية
(ج) الأعمدة البيانية المجزأة
(د) المنحنى التكراري ✓

المحاضرة الثالثة : مقاييس النزعة المركزية

مقاييس النزعة المركزية هي قيم نموذجية يمكن أن تمثل مجموعة من البيانات بحيث تعطي دلالة معينة لتلك البيانات . وحيث أن مثل هذه القيم تميل إلى الوقوع في المركز داخل مجموعة البيانات (عند ترتيبها حسب قيمها) ، فإن هذه القيم سميت بهذا الاسم . وهناك صور عديدة من هذه المقاييس وإن كان الأكثر شيوعاً :

الوسط الحسابي (أو باختصار الوسط أو المتوسط)　　　الوسيط　　　المنوال (أو الشائع)

(1) الوسط الحسابي

- لبيانات غير مبوبة :

مثال : أوجد الوسط الحسابي للقيم **9 , 2 , 7 , 12 , 10**

$$\text{الوسط الحسابي} = \frac{\text{مجموع قيم البيانات}}{\text{عددها}} = \frac{\sum x}{n} = \bar{x}$$

$$\bar{x} = \frac{\sum x}{n} = \frac{9+2+7+12+10}{5} = \frac{40}{5} = 8$$

مثال : أوجد الوسط الحسابي لقيم x المبينة بالجدول التكراري المعطى.

المتغير x	التكرار f
5	6
3	2
6	2
4	5
2	2
8	3

المتغير x	التكرار f	fx
5	6	30
3	2	6
6	2	12
4	5	20
2	2	4
8	3	24
	20	96
	$\sum f = 20$	$\sum fx = 96$

x

الجدول التكراري

$$\bar{x} = \frac{\sum fx}{\sum f}$$

حيث $\sum f$ هو مجموع التكرارات
$\sum fx$ هو مجموع حاصل ضرب كل قيمة في تكرارها

$$\therefore \qquad \bar{x} = \frac{\sum fx}{\sum f} = \frac{96}{20} = 4.8$$

- لبيانات مبوبة :

مثال : أوجد الوسط الحسابي لقيم x المبينة بالجدول التكراري المعطى.

الجدول التكراري

الفئة	المتغير x	التكرار f
الأولى	$0 \le x < 20$	4
الثانية	$20 \le x < 30$	16
الثالثة	$30 \le x < 35$	12
الرابعة	$35 \le x < 40$	10
الخامسة	$40 \le x < 50$	6
السادسة	$50 \le x < 60$	2

الجدول التكراري

الفئة	المتغير x	التكرار f	مركز الفئة x_0	fx_0
1	$0 \le x < 20$	4	10	40
2	$20 \le x < 30$	16	25	400
3	$30 \le x < 35$	12	32.5	390
4	$35 \le x < 40$	10	37.5	375
5	$40 \le x < 50$	6	45	270
6	$50 \le x < 60$	2	55	110
		$\sum f = 50$		$\sum f\, x_0 = 1585$

$$\boxed{\bar{x} = \frac{\sum f\, x_0}{\sum f}} = \frac{1585}{50} = 31.7$$

حيث $\sum f$ هو مجموع التكرارات ، $\sum f\, x_0$ هو مجموع حاصل ضرب مركز كل فئة في تكرار الفئة ، مركز أي فئة = $\dfrac{\text{حد الفئة الأدنى + حدها الأعلى}}{2}$

خواص هامة للوسط الحسابي :

1. يمكن تحديد قيمة الوسط الحسابي بالضبط لأي بيانات كمية (ميزة) لكن لا يمكن تحديده للبيانات النوعية (عيب) ولا للفئات المفتوحة (عيب).
2. يأخذ في الاعتبار جميع البيانات (ميزة) ولا يتأثر بترتيب القيم (ميزة) .
3. يتأثر بالقيم المتطرفة في البيانات (عيب) .
4. إذا أضفنا (أو طرحنا) عدد ثابت c لكل قيمة من قيم البيانات ، فإن الوسط الحسابي للقيم الجديدة = الوسط الحسابي للقيم القديمة + (أو −) هذا الثابت c
5. إذا ضربنا (أو قسمنا) كل قيمة من قيم البيانات في (على) عدد ثابت c ، فإن الوسط الحسابي للقيم الجديدة = الوسط الحسابي للقيم القديمة مضروباً في (مقسوماً على) هذا الثابت c

(2) الوسيط

(ببساطة) يُعرف الوسيط [وسنرمز له بالرمز M] لمجموعة من القيم (المرتبة تصاعدياً أو تنازلياً حسب قيمها) على أنه القيمة التي تقسم مجموعة القيم إلى مجموعتين متساويتين في العدد ، أو بتعبير آخر هي القيمة التي في المنتصف

- **لبيانات غير مبوبة :**

لتحديد قيمة الوسيط لعدد n من القيم :

- قم أولاً بترتيب البيانات تصاعدياً أو تنازلياً .
- حدد ما إذا كانت هناك قيمة واحدة بالمنتصف أم قيمتين ، وهذا يتوقف على قيمة n (عدد القيم) ، فإذا كانت n فردية ستكون هناك قيمة واحدة بالمنتصف وتكون هذه القيمة هي الوسيط ، أما إذا كانت n زوجية ستكون هناك قيمتان في المنتصف وتكون قيمة الوسيط هي الوسط الحسابي لتلك القيمتين .

مثال : مجموعة الأرقام (بعد ترتيبها تصاعدياً) : 9 7 6 6 $\boxed{5}$ 4 3 3 2 وسيطها هو **5** [لاحظ أن عدد القيم $n = 9$ (فردي)]

في حين مجموعة الأرقام (بعد ترتيبها تصاعدياً) : 24 20 13 $\boxed{11}$ $\boxed{9}$ 7 4 3 وسيطها هو $\frac{9+11}{2} = 10$ [لاحظ أن عدد القيم $n = 8$ (زوجي)]

- **لبيانات مبوبة :**

لتحديد قيمة الوسيط لتوزيع تكراري كما هو مبين نتبع الخطوات التالية :

الخطوة الأولى : نحدد الفئة الوسيطية (أي الفئة التي يقع داخلها الوسيط)

- نحسب أولاً نصف مجموع التكرارات $\frac{1}{2}\sum f$
- إبدأ بالصفر في ذهنك وزود تكرارات الفئات على التوالي تكرار تلو الآخر ومع كل زيادة لتكرار نقارن الناتج بنصف مجموع التكرارات السابق حتى نصل إلى نصف مجموع التكرارات أو يزيد عنه فتكون آخر فئة زودنا تكرارها هي الفئة الوسيطية .

الجدول التكراري

الفئة	المتغير x	التكرار f
1	$1 \le x < 3$	14
2	$3 \le x < 5$	29
3	$5 \le x < 7$	18
4	$7 \le x < 10$	9
		$\sum f = 70$

الجدول التكراري		
الفئة	المتغير x	التكرار f
1	$1 \le x < 3$	14
2	$3 \le x < 5$	29
3	$5 \le x < 7$	18
4	$7 \le x < 10$	9
		$\sum f = 70$

كما يلي :

- احسب $\frac{1}{2}\sum f$: $\longleftarrow$ $\frac{1}{2}\sum f = \frac{70}{2} = 35$
- نبدأ بالصفر [في ذهننا] ، نزود على الصفر السابق تكرار الفئة الأولى [14] ينتج 14

14 أقل من 35 ، يبقى الفئة الأولى ليست الفئة الوسيطة

- نزود على الـ 14 الأخيرة تكرار الفئة الثانية [29] ينتج 43

43 أكبر من 35 ، يبقى الفئة الثانية هي الفئة الوسيطة

الخطوة الثانية : نحدد الفئة الوسيطية (أي الفئة التي يقع داخلها الوسيط)

- للفئة الوسيطية نحدد حدها الأدنى وطولها وتكرارها ونحسب ما يُسمى بالتكرار المتجمع السابق وهو مجموع تكرارات الفئات السابقة للفئة الوسيطة

الفئة الوسيطية هي الفئة الثانية : إذن حدها الأدنى = 3 وطولها يساوي 2 [حدها الأعلى 5 – حدها الأدنى 3] ، وتكرارها 29
التكرار المتجمع السابق = تكرار الفئة الأولى فقط = 14

- نحسب الوسيط من العلاقة :

الوسيط M = الحد الأدنى للفئة الوسيطية + $\left[\frac{\text{(نصف مجموع التكرارات – التكرار المتجمع السابق)}}{\text{تكرار الفئة الوسيطية}} \times \text{طول الفئة الوسيطية} \right]$

$$M = 3 + \left[\frac{35-14}{29} \times 2 \right] = 4.4$$

(3) المنوال (الشائع)

يُعرف المنوال لمجموعة من القيم على أنه القيمة التي تتكرر أكثر من غيرها أو القيمة الأكثر شيوعاً [لذا يُسمى في بعض الأحيان بال "الشائع"] . وأحياناً يُرمز للمنوال بالرمز $\hat{X}$

فمثلاً مجموعة القيم : 18 12 11 10 10 9 9 9 7 5 2 2 لها منوال **9**

ومجموعة القيم : 18 15 12 10 8 5 3 9 ليس لها منوال (أو عديمة المنوال)

ومجموعة القيم : 9 7 7 7 5 5 4 4 4 3 2 لها منوالان (7 ، 4)

أي أن مجموعة القيم قد تكون وحيدة المنوال [لها منوال واحد] ، وقد تكون عديدة المنوال [منوالان أو أكثر] وقد تكون عديمة المنوال [لا يوجد لها منوال]

أما مجموعة القيم : 7 7 6 6 5 5 4 4 فقد تتسرع وتقول أنها رباعية المنوال ومناولها : **7 , 6 , 5 , 4**

لكن [حيث أن جميع القيم لها نفس التكرار] هذه المجموعة الأخيرة عديمة المنوال

والمنوال [مقارنةً بالوسط الحسابي والوسيط] به العديد من العيوب منها :

- أنه لا يأخذ في الاعتبار جميع البيانات ولكنه يهتم فقط بالقيم الأكثر تكراراً .
- أنه قد لا يتواجد أو قد يكون هناك أكثر من منوال للبيانات .

إلا أنه أيضاً يتميز ببعض المزايا منها :

- أنه أسرع في تحديده من الوسط والوسيط
- من الممكن تحديده للتوزيعات التكرارية للبيانات غير المبوبة سواء كانت تلك البيانات كمية متقطعة أو نوعية [والبيانات الأخيرة (النوعية) لا يمكن حساب الوسط الحسابي لها أو الوسيط]

سيارات في أحد المواقف	
عدد السيارات	لون السيارة
10	أحمر R
23	أزرق B
12	أبيض W
5	أصفر Y

بيانات نوعية

لها منوال وهو "اللون الأزرق"

علاقة اعتبارية (تقريبية) بين المتوسطات الثلاثة : الوسط والوسيط والمنوال

المنحنيات التكرارية التي تظهر في الناحية العملية تأخذ أشكالاً مميزة منها :

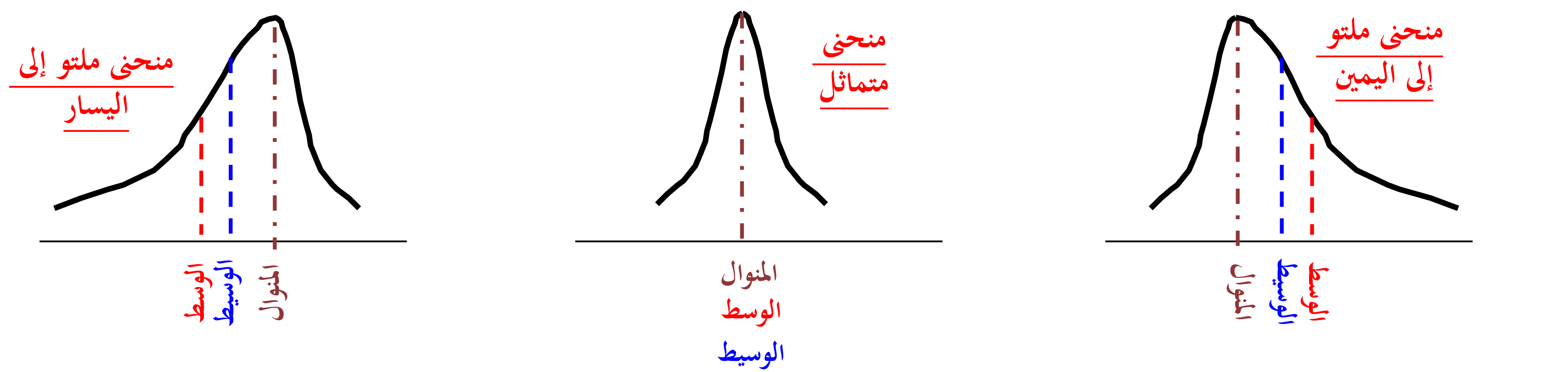
الوسط > الوسيط > المنوال

الوسط أكبر من الوسيط أكبر من المنوال

الوسط = الوسيط = المنوال

المنوال > الوسيط > الوسط

المنوال أكبر من الوسيط أكبر من الوسط



تذكر دائماً أن المنوال هو القيمة المناظرة لأعلى نقطة في المنحنى والوسيط يقع دائماً بين الوسط و المنوال

والمنحنيات التكرارية وحيدة المنوال والبسيطة الالتواء تحقق العلاقة الاعتبارية التالية :

الوسط – المنوال = 3 × (الوسط – الوسيط)

أسئلة اختبارات

1. **البيانات التالية، تمثل أعمار (7) طلاب (18 ، 23 ، 20 ، 19 ، 25 ، 21 ، 17)، فإن العمر الوسيط هو:**

(أ) 18 (ب) 19 ✓ (ج) 20 (د) 32

2. **في مقرر الاحصاء الاجتماعي كانت درجات خمسة طلاب كالتالي (9 ، 10 ، 12 ، 7 ، 2)، أراد المعلم تحسين الدرجات بإضافة ثلاث درجات لكل طالب. وفقاً لذلك فإن الوسط الحسابي الجديد للبيانات يساوي:**

(أ) 8 (ب) 9 (ج) 10 ✓ (د) 11

3. **أحد مقاييس النزعة المركزية الأكثر استخداماً في البحوث الاجتماعية هو:**

(أ) الانحراف المعياري (ب) المدى ✓ (ج) الوسط الحسابي (د) الانحراف المتوسط

4. **القيمة الأكثر شيوعاً أو التي تتكرر أكثر من غيرها في المجموعة أو التوزيع تسمى:**

(أ) المدى ✓ (ب) المنوال (ج) الانحراف المتوسط (د) الوسط الحسابي

5. **المنوال لمجموعة القيم (9 ، 11 ، 10 ، 12 ، 18 ، 9، 10، 18 ، 12 ، 9) هو:**

✓ (أ) 9 (ب) 10 (ج) 12 (د) 18

6. **قيمة الوسيط للبيانات (0 ، 1 ، 2 ، 15 ، 11 ، 10 ، 8 ، 17) هي:**

✓ (أ) 9 (ب) 11 (ج) 15 (د) 17

7. **يصنف المنوال ضمن مقاييس:**

(أ) الاحتمالات (ب) التشتت ✓ (ج) النزعة المركزية (د) العينات

8. **المقياس الذي تعتمد قيمته على قيم البيانات جميعها هو:**

✓ (أ) الوسط الحسابي (ب) الوسيط (ج) المنوال (د) المدى

9. **يوضح الجدول المقابل درجات الطلاب في أعمال السنة في مقرر الاحصاء الاجتماعي وتكرار الطلاب الحاصلين على كل درجة. لوسط الحسابي لدرجات الطلاب هو:**

الدرجة	4	5	6	7
التكرار	20	40	30	10

(أ) 0,53 ✓ (ب) 5,3
(ج) 3,5 (د) 0,35

التكرار	فئات الدرجات
14	$1 \le x < 3$
29	$3 \le x < 5$
18	$5 \le x < 7$
9	$7 \le x < 10$

10. يوضح الجدول التالي درجات (70) طالب وتكراراتها في اختبار مناهج البحث. الفئة الوسيطية للدرجات هي:

(أ) $1 \le x < 3$

(ب) $3 \le x < 5$

(ج) $5 \le x < 7$

(د) لا يمكن تحديدها

11. إذا كان الوسط الحسابي لمجموعة من القيم 70، والوسيط 60، والمنوال 40، فإن المنحنى التكراري للبيانات:

(أ) متطابق (ب) متماثل (ج) ملتو لليسار (د) ملتو لليمين

12. في التوزيع الملتوي إلى اليسار يكون الوضع النسبي للمتوسطات:

(أ) الوسيط > المنوال > الوسط الحسابي

(ب) الوسط الحسابي > المنوال > الوسيط

(ج) المنوال > الوسيط > الوسط الحسابي

(د) الوسط الحسابي > الوسيط > المنوال

13. القيمة التي تقسم مجموعة القيم إلى مجموعتين متساويتين في العدد تسمى:

(أ) الوسط الحسابي (ب) الوسيط (ج) المنوال (د) المدى

14. من عيوب الوسط الحسابي:

(أ) يتأثر بترتيب البيانات

(ب) لا يأخذ في الاعتبار جميع البيانات

(ج) لا يتأثر بالقيم المتطرفة للبيانات

(د) يتأثربالقيم المتطرفة للبيانات

المحاضرة الرابعة : مقاييس التشتت

تعريف التشتت

الدرجة التي تتجه بها البيانات الكمية للانتشار حول قيمة متوسطة (أحد مقاييس النزعة المركزية) تُسمى تشتت أو تغير البيانات ، ومن أكثر المقاييس شيوعا لقياس هذا التشتت :

المدى – الانحراف المتوسط – الانحراف المعياري

أولاً : المدى R : مدى مجموعة من البيانات الكمية هو الفرق بين أكبر قيمة في البيانات وأقل قيمة فيها

فمثلاً لمجموعة القيم : 15 3 7 6 12 18 5 3 13 15 ← المدى : $R = 18 - 3 = 15$

وللتوزيع التكراري المبين يكون

المتغير x	التكرار f
15	6
3	2
20	2
4	5

$R = 20 - 3 = 17$

أما للتوزيع التكراري المبين فإن

الجدول التكراري		
الفئة	المتغير x	التكرار f
1	$1 \le x < 3$	14
2	$3 \le x < 5$	29
3	$5 \le x < 7$	18
4	$7 \le x < 10$	9

$R = 10 - 1 = 9$

وبالرغم من بساطة تحديده إلا أن له بعض العيوب مثل تأثره بالقيم المتطرفة ، كما أنه لا يمكن تحديده في حالة التوزيعات التكرارية المفتوحة .

الفئة	العمر x
1	$x < 6$
2	$6 \le x < 12$
3	$12 \le x < 15$
4	$15 \le x < 18$

مفتوح من أسفل

الفئة	العمر x
1	$6 \le x < 12$
2	$12 \le x < 15$
3	$15 \le x < 18$
4	$x \ge 18$

مفتوح من أعلى

الفئة	العمر x
1	$x < 6$
2	$6 \le x < 12$
3	$12 \le x < 15$
4	$x \ge 15$

مفتوح من الطرفين

لا يمكن تحديد مدى البيانات

ثانياً : الانحراف المتوسط [أو متوسط الانحرافات] $M.D$

يُعرف الانحراف المتوسط (أو متوسط الانحرافات) [وسنرمز له بالرمز $M.D$] على أنه متوسط القيم المطلقة للانحرافات عن قيمة متوسطة للبيانات [عادةً تكون الوسط الحسابي . أي أن :

$$M.D = \frac{\sum |d|}{n}$$

حيث $d = x - \bar{x}$ هي انحراف القيمة x عن الوسط الحسابي ، $|d|$ هي القيمة المطلقة للانحراف d .

فمثلاً لمجموعة القيم : 10 12 7 2 9 يمكننا حساب الانحراف المتوسط لها كالتالي :

x
9
2
7
12
10
40

عدد القيم $n = 5$

$\sum x = 40$ مجموع القيم

$\therefore \quad \bar{x} = \frac{\sum x}{n} = \frac{40}{5} = \underline{\underline{8}}$ الوسط الحسابي

بعد ذلك نحسب الانحرافات عن الوسط الحسابي والقيم المطلقة لهذه الانحرافات

x	d	$\lvert d \rvert$
9	1	1
2	-6	6
7	-1	1
12	4	4
10	2	2
		14

$\sum |d| = 14$

وبالتالي يكون الانحراف المتوسط مساوياً لـ

$$M.D = \frac{\sum |d|}{n} = \frac{14}{5} = \underline{\underline{2.8}}$$

أما في حالة جدول تكراري كمل هو مبين ، فإن الانحراف المتوسط **M.D** يتحدد من :

$$M.D = \frac{\sum f\,|d\,|}{\sum f}$$

أي نضرب القيمة المطلقة لانحراف كل قيمة [عن الوسط] في تكرارها ، ثم نقسم الناتج على مجموع التكرارات

فمثلاً : إذا كان المطلوب حساب الانحراف المتوسط للبيانات المبينة بالجدول التكراري التالي :

المتغير x	التكرار f
4	20
5	40
6	30
7	10

الجدول التكراري

المتغير x	التكرار f	fx	$d = x - \bar{x}$	$\|d\|$	$f\|d\|$
4	20	80	4 – 5.3 = -1.3	1.3	20 × 1.3 = 26
5	40	200	5 – 5.3 = -0.3	0.3	40 × 0.3 = 12
6	30	180	6 – 5.3 = 0.7	0.7	30 × 0.7 = 21
7	10	70	7 – 5.3 = 1.7	1.7	10 × 1.7 = 17
	100	530			76
	$\sum f = 100$	$\sum fx = 530$			$\sum f\,|d\,|$

دول عشان نحسب الوسط الحسابي

$$\bar{x} = \frac{\sum fx}{\sum f} = \frac{530}{100} = 5.3$$

وده الانحرافات عن الوسط والقيم المطلقة لها

وده حواصل ضرب التكرارات في القيم المطلقة ومن ثم المجموع

$$\therefore \quad M.D = \frac{\sum f\,|d\,|}{\sum f} = \frac{76}{100} = 0.76$$

الانحراف المتوسط

مزايا الانحراف المتوسط : من السهل حسابه – يأخذ في الاعتبار جميع البيانات – لا يحتاج لترتيب معين للبيانات

عيوب الانحراف المتوسط : يتأثر بشدة بالقيم المتطرفة – لا يمكن حسابه للتوزيعات التكرارية المفتوحة

ثالثاً : التباين s^2 والانحراف المعياري s

يُعرف متوسط مربعات الانحرافات عن الوسط الحسابي على أنه تباين مجموعة البيانات [ويُرمز له بالرمز s^2] ، ويُعرف الجذر التربيعي للتباين على أنه الانحراف المعياري للبيانات [ويُرمز له بالرمز s] ، أي أن :

التباين $s^2 = \frac{\sum d^2}{n}$ ومنه يكون ← الانحراف المعياري $= s = \sqrt{s^2} = \sqrt{\frac{\sum d^2}{n}}$

فمثلاً لمجموعة القيم : 10 12 7 2 9 التي سبق وحسبنا لها الانحراف المتوسط ، يمكن حساب التباين والانحراف المعياري كالتالي :

x	d	d^2
9	1	1
2	-6	36
7	-1	1
12	4	16
10	2	4
40		58

عدد القيم $n = 5$

$\sum x = 40$ $\quad \sum d^2 = 58$

الوسط الحسابي $\bar{x} = \frac{\sum x}{n} = \frac{40}{5} = 8$

التباين $s^2 = \frac{\sum d^2}{n} = \frac{58}{5} = 11.6$

الانحراف المعياري $s = \sqrt{11.6} \cong 3.4$

أما في حالة جدول تكراري كمل هو مبين ، فإن التباين s^2 يتحدد من : $s^2 = \frac{\sum fd^2}{\sum f}$ أي نضرب (مربع انحراف كل قيمة عن وسطها) في (تكرارها) ، ثم نقسم الناتج على مجموع التكرارات

فمثلاً : إذا كان المطلوب حساب الانحراف المعياري للبيانات المبينة بالجدول التكراري التالي :

المتغير x	التكرار f
4	20
5	40
6	30
7	10

الجدول التكراري

المتغير x	التكرار f	fx	$d = x - \bar{x}$	d^2	fd^2
4	20	80	4 – 5.3 = -1.3	1.69	20 × 1.69= 33.8
5	40	200	5 – 5.3 = -0.3	0.09	40 × 0.09= 3.6
6	30	180	6 – 5.3 = 0.7	0.49	30 × 0.49= 14.7
7	10	70	7 – 5.3 = 1.7	2.89	10 × 2.89= 28.9
	100	530			81
	$\sum f = 100$	$\sum fx = 530$			$\sum fd^2$

دول عشان نحسب الوسط الحسابي

وده الانحرافات عن الوسط ومربعاتها

وده حواصل ضرب التكرارات في مربعات الانحرافات ومن ثم المجموع

$$\bar{x} = \frac{\sum fx}{\sum f} = \frac{530}{100} = 5.3$$

$$\therefore \quad s^2 = \frac{\sum fd^2}{\sum f} = \frac{81}{100} = 0.81$$ التباين

الانحراف المعياري $s = \sqrt{s^2} = \sqrt{0.81} = \underline{\underline{0.9}}$

ومزايا وعيوب الانحراف المعياري هي نفسها مزايا وعيوب الانحراف المتوسط ؛ أي أن

مزايا الانحراف المعياري : من السهل حسابه – يأخذ في الاعتبار جميع البيانات – لا يحتاج لترتيب معين للبيانات

عيوب الانحراف المعياري : يتأثر بشدة بالقيم المتطرفة – لا يمكن حسابه للتوزيعات التكرارية المفتوحة

أسئلة اختبارات

1. **يصنف المدى ضمن مقاييس:**

✓ (أ) مقاييس التشتت (ب) الصدق (ج) مقاييس النزعة المركزية (د) الثبات

2. **يصنف الانحراف المتوسط ضمن مقاييس:**

(أ) النزعة المركزية ✓ (ب) التشتت (ج) الصدق (د) العينات

3. **من عيوب الانحراف المعياري:**

✓ (أ) لا يمكن حسابه للتوزيعات التكرارية المفتوحة (ب) لا يأخذ في الاعتبار جميع البيانات
(ج) لا يحتاج لترتيب بعض البيانات (د) لا يتأثر بالقيم الطرفية

4. **من عيوب الانحراف المتوسط :**

(أ) لا يأخذ في الاعتبار جميع البيانات (ب) لا يحتاج لترتيب بعض البيانات
(ج) لا يتأثر بالقيم الطرفية ✓ (د) لا يمكن حسابه للتوزيعات التكرارية المفتوحة

5. **قيمة المدى للدرجات التالية (6 ، 5 ، 4 ، 5 ، 7 ، 8 ، 7 ، 3 ، 4 ، 6) هو:**

(أ) 3 (ب) 4 ✓ (ج) 5 (د) 6

6. **في الجدول المبين ، مدى الدرجات يساوي :**

فئات الدرجات	6 -	12 -	18 -	24 -	30 -	36 – 42
عدد الطلاب	6	22	36	12	10	3

(أ) 6 (ب) 30 (ج) 33 ✓ (د) 36

7. **مقياس التشتت الذي يعتمد على أعلى وأقل قيمة هو:**

✓ (أ) المدى (ب) الوسط الحسابي (ج) المنوال (د) الوسيط

8. **يوضح الجدول التالي العمر الزمني لمجموعة من الطلاب.**
المدى للعمر الزمني يساوي:

العمر	الفئة
$6 \le x < 12$	الأولى
$12 \le x < 15$	الثانية
$15 \le x < 18$	الثالثة
$x \ge 18$	الرابعة

(أ) 6 (ب) 9
(ج) 12 ✓ (د) لا يمكن تحديده

خاص بالأسئلة من (9) إلى (15) :

للدرجات : 14 7 12 17 15

9. الوسط الحسابي يساوي

✓ (أ) 13 (ب) 12 (ج) 10 (د) غير موجود

10. المنوال يساوي

(أ) 13 (ب) 12 (ج) 10 ✓ (د) غير موجود

11. المدى يساوي

(أ) 13 (ب) 12 ✓ (ج) 10 (د) غير موجود

12. الوسيط يساوي

(أ) 13 ✓ (ب) 14 (ج) 10 (د) غير موجود

13. التباين يساوي

(أ) 2.8 (ب) 3.4 (ج) 10 ✓ (د) 11.6

14. الانحراف المتوسط يساوي

✓ (أ) 2.8 (ب) 3.4 (ج) 10 (د) 11.6

15. الانحراف المعياري يساوي

(أ) 2.8 ✓ (ب) 3.4 (ج) 10 (د) 11.6

المحاضرة الخامسة : الفروض الإحصائية

يعرف **الفرض** بأنه إجابة متوقعة لسؤال من الأسئلة التي تراود ذهن الباحث أو المهتم ، وهذه الإجابة لا تكون نهائية وإنما خاضعة للدراسة والتحقق من مدى صحتها فإما أن تكون الإجابة صحيحة وإما أن تكون الإجابة خاطئة.

الفرضية الصفرية (فرضية العدم) H_0:

هي الفرضية حول معلمة المجتمع التي نجري اختبار عليها باستخدام بيانات من عينة والتي تشير أن الفرق بين معلمة المجتمع والإحصائي من العينة ناتج عن الصدفة ولا فرق حقيقي بينهما. وهي الفرضية التي ننطلق منها ونرفضها عندما تتوفر دلائل على عدم صحتها، وخلاف ذلك نقبلها .

الفرضية البديلة H_a :

وهي الفرضية التي يضعها الباحث كبديل عن فرضية العدم ونقبلها عندما نرفض فرضية العدم باعتبارها ليست صحيحة بناء على المعلومات المستقاة من العينة.

وفي اختبارات الفروض يمكن أن نرتكب نوعين من الخطأ :

الخطأ من النوع الأول : وهو أن **نرفض** الفرض العدمي بينما هو **صحيح** .

الخطأ من النوع الثاني : وهو أن **نقبل** الفرض العدمي بينما هو **خاطئ** .

		الفرض العدمي Ho	
		✓	×
قرار	قبول	صواب	خطأ من النوع الثاني
	رفض	خطأ من النوع الأول	صواب

أي أن :

- فرضية صحيحة نتائج العينة تؤيد صحتها. (قبول صواب)
- فرضية صحيحة نتائج العينة غير مؤيدة لصحتها. (رفض صواب) وهذا يعطينا خطأ من النوع الأول
- فرضية خاطئة نتائج تؤيد صحتها (قبول خطأ) وهذا يعطينا خطأ من النوع الثاني
- فرضية خاطئة نتائج غيرمؤيدة صحتها (رفض خطأ)

الفروض البحثية

هي الفروض التي يصيغها الباحث في بحثه بناءً على خلفيته النظرية ونتائج دراسات سابقة

	الفروض العلاقية	الفروض الفارقة	الفروض التنبؤية	الفروض السببية
الفرض البديل غير الموجه	توجد علاقة ذات دلالة إحصائية بين الاتجاه نحو الدراسة والبيئة الدراسية	توجد فروق ذات دلالة إحصائية بين متوسطي درجات الذكور والإناث في الذكاء	يمكن التنبؤ من المتغيرات المستقلة (مثل القلق وحب الاستطلاع) بالمتغير التابع (ليكن التحصيل الدراسي لدى طلاب جامعة ما	يمكن التوصل إلى نمذج سببي يفسر العلاقة بين المتغيرات المستقلة (معاملة الوالدين ، الضغط النفسي) والمتغير التابع (الطموح) لدي طلاب جامعة ما
الفرض البديل الموجه	توجد علاقة ذات دالة إحصائية إيجابية بين الاتجاه نحو الدراسة والبيئة الدراسية	توجد فروق ذات دلالة إحصائية بين متوسطي درجات الذكور والإناث في الذكاء لصالح الإناث	يمكن التنبؤ من المتغيرات المستقلة (القلق) كمنبئ سالب وحب الاستطلاع كمنبئ موجب) بالمتغير التابع (التحصيل الدراسي)	يمكن التوصل إلى نمذج سببي يفسر العلاقة بين المتغيرات المستقلة (معاملة الوالدين كتأثير موجب والضغط النفسي كتأثير سالب) والمتغير التابع (الطموح)
الفرض الصفري	لا توجد علاقة ذات دالة إحصائية بين الاتجاه نحو الدراسة والبئة الدراسية	لا توجد فروق ذات دلالة إحصائية بين متوسطي درجات الذكور والإناث في الذكاء	لا يمكن التنبؤ من المتغيرات المستقلة بالمتغير التابع	لا يمكن التوصل إلى نموذج سببي يفسر العلاقة بين المتغيرات المستقلة والمتغير التابع
الأسلوب الإحصائي المناسب	معاملات الارتباط	إختبارات «ت»	تحليل الانحدار	تحليل المسار

أسئلة اختبارات

1. **الخطأ من النوع الثاني في الفروض يسمى:**
(أ) قبول صواب ✓ (ب) قبول خطأ (ج) رفض صواب (د) رفض خطأ

2. **عندما يكون لدينا فرضية خاطئة نتائج البحث غير مؤيدة صحتها، فهذا يسمى:**
(أ) قبول خطأ ✓ (ب) رفض خطأ (ج) قبول صواب (د) قبول صواب

3. **الخطأ من النوع الأول في الفروض يسمى:**
(أ) قبول خطأ (ب) قبول صواب (ج) رفض خطأ ✓ (د) رفض صواب

4. **عندما يكون لدينا فرضية صحيحة نتائج البحث غير مؤيدة صحتها، فهذا يسمى:**
(أ) قبول صواب (ب) قبول خطأ ✓ (ج) رفض صواب (د) رفض خطأ

5. **" توجد فروق ذات دلالة احصائية بين متوسطي درجات طلاب قسم التاريخ وطلاب قسم الاجتماع في الاحصاء الاجتماعي لصالح طلاب علم الاجتماع". نوع الفرض هو:**
(أ) فرض صفري (ب) فرض صفري موجه ✓ (ج) فرض بديل موجه (د) فرض بديل غير موجه

6. **" توجد فروق ذات دلالة احصائية بين متوسطي درجات طلاب كلية الدراسات التطبيقية وطلاب كلية الأداب في مشروع التخرج". نوع الفرض هو:**
(أ) فرض بديل موجه ✓ (ب) فرض بديل غير موجه (ج) فرض صفري (د) فرض صفري موجه

7. **"لا توجد فروق ذات دلالة احصائية بين متوسطي درجات طلاب كلية الآداب وطلاب كلية التربية في الاحصاء الاجتماعي". نوع الفرض هو:**
(أ) فرض بديل غير موجه (ب) فرض بديل موجه ✓ (ج) فرض صفري (د) فرض صفري موجه

8. **الأسلوب الاحصائي المناسب للتحقق من الفرض "يمكن التنبؤ بأداء طلاب التعليم الالكتروني في مقرر الاحصاء بمعلومية الذكاء العملي لديهم" هو:**
(أ) معامل الارتباط (ب) تحليل التباين ✓ (ج) تحليل الانحدار (د) اختبار مربع كاي

تمت
بحمد الله