مبادئ الإحصاء

علم الإحصاء:

هو العلم أو مجموعة القواعد والطرق والنظريات التي تهتم بجمع البيانات وتبويبها وعرضها بيانياً ثم تحليلها وتفسيرها وإجراء المقارنات واستنتاج العلاقات بهدف استخدامها في اتخاذ القرارات المناسبة.

من هذا التعريف نستنتج الملاحظات التالية:

1/أن المراحل الأساسية للعملية الإحصائية هي 4 مراحل: جمع البيانات / تبويبها / العرض البياني لها / تحليلها.

2/الهدف الأساسي من العملية الإحصائية هو تحليل البيانات وتفسيرها.

3/ يمكن تطبيق عملية الإحصاء في مختلف المجالات.

4/أن البيانات هي المجال الرئيسي لمراحل علم الإحصاء.

أنواع البيانات:

البيانات الوصفية		البيانات الكمية	
أ/بيانات وصفيه اسمية: مثل تصنيف المواليد (ذكور إناث) وتصنيف أماكن تواجد الإدارات(الشرقية/الوسطى) نصفها بالاسم..	ب/ بيانات وصفيه ترتيبية: مثل التقديرات (ممتاز، جيد جدا، جيد، مقبول)، بمعنى إنها ترتب الممتاز أولا ثم الجيد جدا ثم الجيد وهكذا.	أ/ بيانات كميه متصلة: وهي التي تأخذ جميع القيم داخل نطاق معين مثلاً (أعمار درجات طلاب من0 إلى20)	ب/ بيانات كمية متقطعة: وهي التي تأخذ قيماً منقطعة عن بعض مثلا (عدد أفراد الأسرة أسرة أفرادها 4 وأسرة أفرادها 8 وهكذا) وممكن لا يكون هناك أسرة عدد أفرادها ما بين ال4 وال8 ، إذا البيانات هنا منقطعة ، وممثله عدد الغرف في المباني

والبيانات الكمية يعبر عنها بأرقام، ويمكن ترتيبها تصاعدياً وتنازلياً، وكذلك يمكن إجراء العمليات الحسابية عليها...

جمع البيانات

1. هي المرحلة الأولى من مراحل علم الإحصاء وهي المرحلة الأساسية والمهمة.
2. تعتبر أكثر تكلفة وأكثر جهد.
3. يُنشأ لها أجهزة ومؤسسات متخصصة.

مصادر البيانات:

1/ المصادر التاريخية

وتشمل الإحصاءات والنشرات الإحصائية التي تصدرها المؤسسات الحكومية والخاصة لتبين أوجه التغير والتطور في المجال الذي تختص به هذه المؤسسات. وهذه البيانات هي من مسؤولية هذه الشركة أو المؤسسة.

ونلاحظ على هذا المصدر:

1. عدم توفر جميع البيانات. (لا تغطي جميع جوانب البحث)
2. قد تكون قديمة.
3. قد تكون غير دقيقة.
4. قد تكون البيانات المنشورة لا تغطي جميع جوانب البحث.

2 / المصادر الميدانية للبيانات:

يتم تجميعها عن طريق الاستبيانات الخاصة ويحصل الباحث على البيانات مباشرة من أفراد العينة.

أساليب البحث الميداني

أ/ الحصر الشامل.

ب/ العينة.

أ / أسلوب المسح الشامل

وهو جمع البيانات من جميع أفراد المجتمع محل الدراسة.

يلاحظ عليه:

1. الشمول.
2. تنوع المجتمعات الإحصائية (بشري، نباتي، حيواني).
3. المجتمع الإحصائي قد يكون محدود أو غير محدود.
4. تشابه خصائص أفراد المجتمع محل الدراسة.

مزايا الحصر الشامل:

1. يوفر المعلومات الدقيقة عن جميع أفراده.
2. نتائجه نهائية، لا تحتاج إلى تعديل.

عيوبه:

1. طول الوقت.
2. جهود كبيرة.
3. تكاليف كبيرة.
4. في بعض الحالات قد يؤدي إلى تلف جميع مكونات المجتمع.

ب /أسلوب العينة:

يتم جمع البيانات من مجموعة من أفراد المجتمع وليس المجتمع ككل.

مزايا أسلوب العينة (توفير الوقت والجهد والمال)

عيوبه/

1. قد تكون النتائج غير دقيقة.
2. قد لا تكون نهائية.
3. العينات لا تصلح في بعض الحالات.

العرض الجدولي للبيانات الوصفية:

ألوان سيارات **20** طالبا (ابيض. اسود. ازرق. أبيض. احمر. ذهبي. اسود. ابيض. ازرق. فضي. اسود. أبيض. ذهبي. احمر . ابيض. ذهبي. اسود. أبيض. اسود. ذهبي).

توزيع ألوان سيارات 20 طالباً

اللون	العلامات	عدد السيارات أو التكرار ورمزه (ك) أو F
أبيض	~~\\\\~~ \	**6**
اسود	\\\\	**4**
ازرق	\\\	**3**
احمر	\\	**2**
ذهبي	\\\\	**4**
فضي	\	**1**
المجموع		**20**

اللون	التكرار (ك) (f)	التكرار المئوي = $f \div \sum f$	التكرار النسبي =$f \div \sum f$ ×100
ابيض	**6**	0,3	**30%**
اسود	**4**	0,2	**20%**
ازرق	**3**	0,15	**15%**
احمر	**2**	0,1	**10%**
ذهبي	**4**	0,2	**20%**
فضي	**1**	0,05	**5%**
المجموع	**20**	$\sum$ **f** (المجموع الكلي للتكرارات)	**100**

العرض الجدولي للبيانات الكمية

العرض للبيانات الكمية المتقطعة

مثال:

دراسة عدد أفراد (20) أسرة حصل الباحث على الإعداد الآتية

(7,7,3,6,4,6,3,4,66,5,2,5,3,5,4,34,3,2):

التكرار عدد الأسرة	العلامات	حجم الأسرة
2	\\	2
5	\\\\\	3
4	\\\\	4
3	\\\	5
4	\\\\	6
2	\\	7
20		

عرض البيانات الكمية المتصلة

دراسة لأعمار (30) مراجعا لأحد المراكز الصحية حصل الباحث على الأرقام التالية:

(4/9/17/20/12/17/10/26/6/23/18/5/16/20/14/16/19/23/11/13/18/9/27/23/8/10/15/21/19/9)

هنا لابد من تكوين الفئات أو المجموعات التي تتدرج تحتها الأرقام السابقة وتصميم الفئات يرجع للباحث ويمكن أن يسترشد **بالمدى وهو الفرق بين أكبر قيمه وأصغر قيمة** المدى = **27 – 4 = 23**

طول الفئة = المدى ÷ عدد الفئات = 23 ÷ 6 + 3.8 تقريباً 4.

ملاحظات

- لا بد أن يكون أقل رقم ضمن الفئة الأولى، وأعلى رقم ضمن الفئة الأخيرة.
- لا يلزم أن نبدأ بأقل رقم بل ممكن أن نبدأ بأقل منه مثل من 0 إلى 4 أو من 2إلى 6، وكذلك لا يلزم أن يكون أعلى رقم هو حد الفئة الأخيرة.
- يوجد الفئة المفتوحة من الأسفل: مثل الفئة الأولى 12 فأقل، أو من الأعلى مثل: 24 فأكثر

الفئات	العلامات	التكرار	التكرار المئوي	التكرار النسبي
4إلى أقل من ثمانية 8	\\\	3	0،1	15%
8 – (الشرطة تعني أقل من)	卌 \\	7	0،23	32%
12 –	\\\\	4	0.13	13%
16 –	卌 \\\	8	0،23	27%
20 –	卌 \	6	0،2	20%
24– 28	\\	2	0،067	6.7%
المجموع		30		

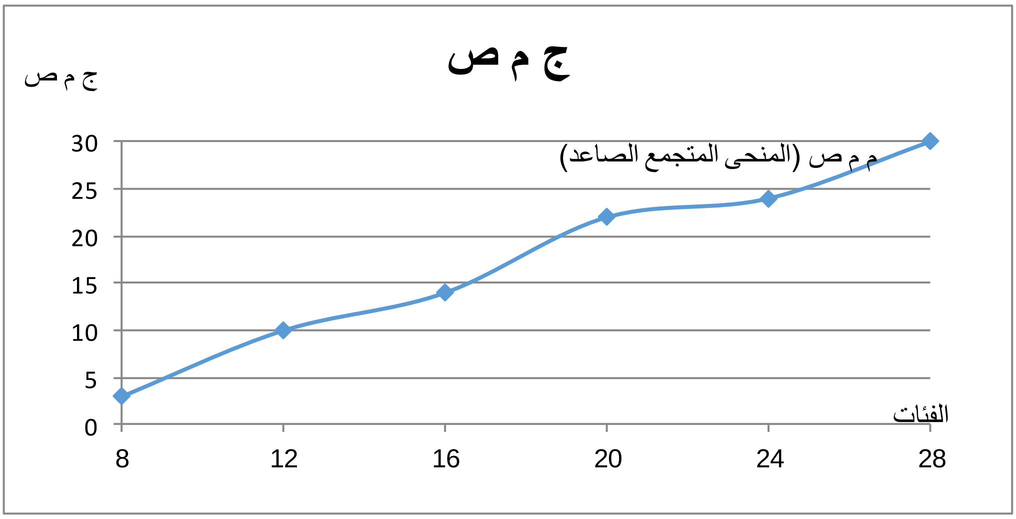
مركز الفئة = الحد الأدنى + الحد الأعلى ÷ 2

الفئات	التكرار	مركز الفئة رمزه (X)	التكرار النسبي
4إلى أقل من ثمانية 8	3	4+8=12÷2= 6	15%
8 – (الشرطة تعني أقل من)	7	10	32%
12 –	4	14	13%
16 –	8	18	27%
20 –	6	22	20%
24– 28	2	26	6.7%
المجموع	30		

الجدول المتجمع الصاعد

وهو عبارة عن جمع متتالي للتكرارات في كل فئة والتي قبلها، وتكون آخر فئة هي عبارة عن مجموع التكرارات

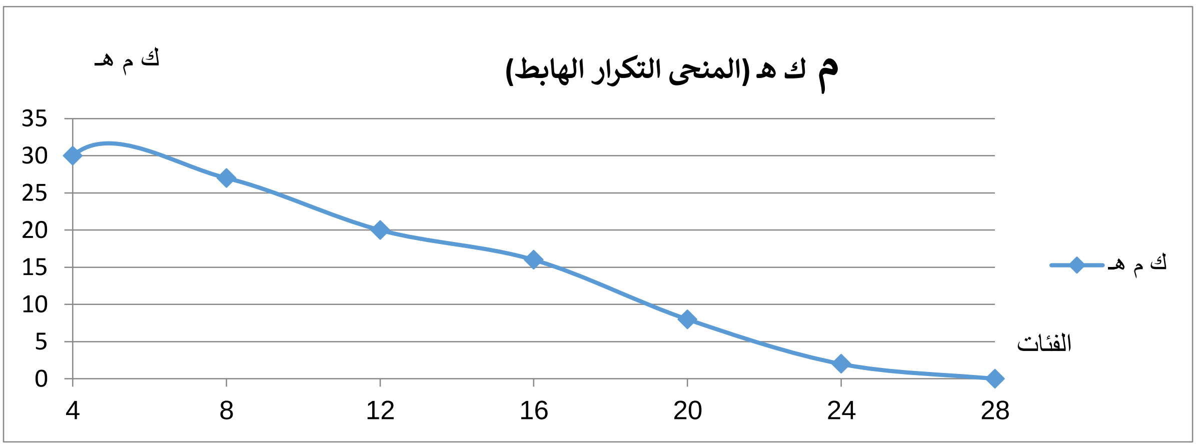
الحدود العليا للفئات	ج م ص
أقل **8** (الفئة الأولى فقط)	3
اقل من **12** (الفئة الثانية والتي قبلها)	10
اقل من **16** (الفئة الثالثة والتي قبلها)	14
اقل من **20** (الفئة الرابعة والتي قبلها)	22
اقل من **24** (الفئة الخامسة والتي قبلها)	28
اقل من **28** (نفس المجموع الكلي)	30



التكرار المتجمع الهابط (ك م هـ)

في الفئة الأولى يكون المجموع الكلي، وفي باقي الفئات نطرح عدد التكرار في الفئة التي قبلها، كما أن الفئة الأخيرة يكون العدد صفر.

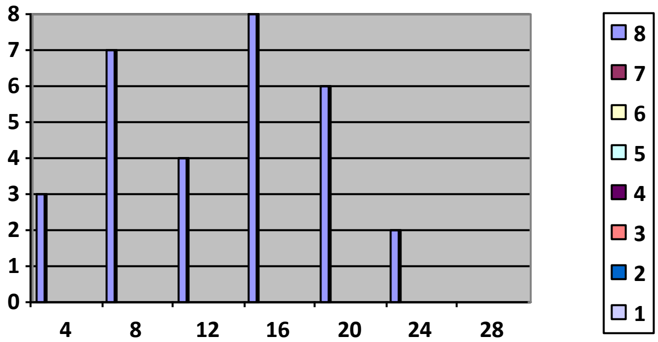
أكبر من	ك م هـ
2	
4	**30** (نفس المجموع الكلي)
8	**27** (جميع الفئات إلا الأولى)
12	**20** (جميع الفئات إلا الأولى والثانية)
16	**16**
20	**8**
24	**2**
28	**0**



العرض البياني للبيانات

1/ الأعمدة البيانية

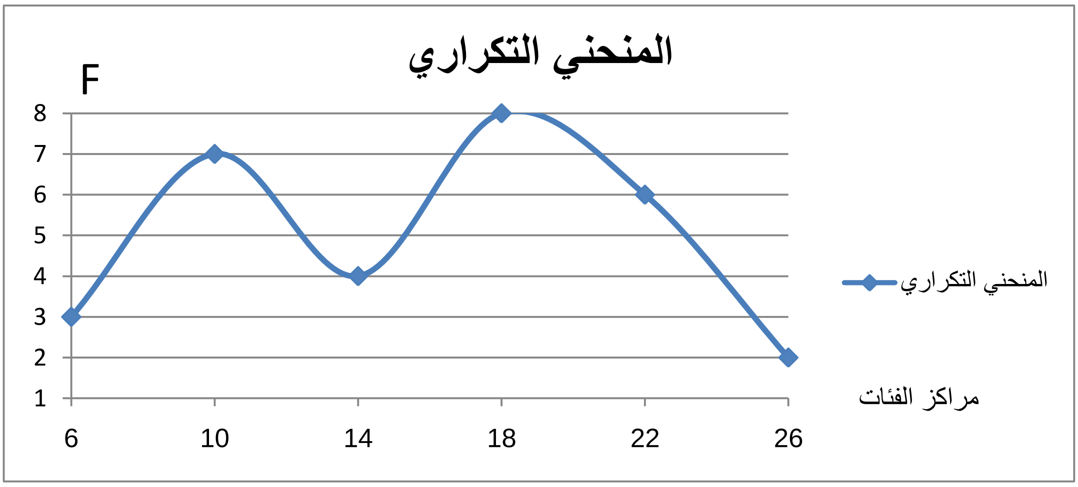
تكون قاعدة الأعمدة (الأفقي) الفئات، ويكون الارتفاع (الرأسي) تكرار الفئة



2/ المنحني التكراري

يمثل العلاقة بين مراكز الفئات، المحور الأفقي يمثل مراكز الفئات، والمحور الرأسي التكرارات.

المضلع تكون زوايا ، والمنحنى لا يكون زوايا . كما يمكن رسم المضلع التكراري بنفس جدول الأعمدة البيانية .



مقاييس النزعة المركزية

الوسط الحسابي $\overline{X}$

هو: القيمة التي لو حلت محل جميع القيم لم يتغير مجموعها.

أ/ البيانات غير المبوبة مثل (7،4،2..)

ب/ البيانات المبوبة: تكون على شكل فئات وتكرارات مثل الجدول المقابل.

الفئات	(F)
5	2
10	3
15	8

قاعدة حساب الوسط الحسابي للبيانات غير المبوبة هي: **مجموع القيم ÷ عددها** $\overline{X} = \frac{\sum x}{n}$

بحيث إن $\sum$**x** هي مجموع القيم، و **n**هي عدد القيم، مثال الوسط الحسابي ل (6،4،2) هو: 12÷3=4

قاعدة حساب الوسط الحسابي للبيانات المبوبة بدون فئات (على شكل قيم وتكرارات). $\overline{X} = \frac{\sum f x}{\sum f}$

مثال: الوسط الحسابي لأعمار الطلاب:

العمر **X**	عدد **F**	العمر ×عدد الطلاب **FX**
13	**7**	91 = 13×3
14	**11**	154 = 11×14
15	**7**	105
	25	$\sum$**fx** 350

فيكون الوسط الحسابي لأعمار الطلاب هو: $\overline{X} = \frac{350}{25} = 14$

مثال حساب المعدل التراكمي:

أولاً: يجب تحديد كل من **f** و**x**، ثم نحدد **fx**

المادة	النقاط **X**	عدد الساعات **F**	**FX**
الجزئي	3.5	**3**	10.5
الإحصاء	4.5	**2**	**9**
النحو	**4**	**1**	**4**
المحاسبة	**4**	**3**	**12**
المجموع		**9**	35.5

ثم نحسب الوسط الحسابي = $\overline{X} = \frac{\sum f x}{\sum f}$ $\overline{X} = \frac{35,5}{9} = 3,94$

قاعدة الوسط الحسابي للبيانات المبوبة بفئات: (مجموع القيم ÷ عددها) بعد استخدام مركز الفئة

X هنا هي مركز الفئة

$$\overline{X} = \frac{\sum f x}{\sum f}$$

الوسط الحسابي = 240 ÷ 30 = 8

الفئات	التكرار F	X	FX
−2	**4**	3	13
−4	**6**	5	30
−6	**3**	7	21
−8	**9**	9	81
−10	**4**	11	44
12−14	**4**	13	52
المجموع	30		240

خصائص الوسط الحسابي:

$$\sum (x-\bar{x}) = 0$$

أ/ مجموع الانحرافات عن الوسط الحسابي يساوي صفر

فمثلاً لو قلنا إن الوسط الحسابي للمثال أدناه هو 6، فيكون كالتالي:

$x-\bar{x}$	x
2− =6−4	**4**
0 = 6−6	**6**
2 =6−8	**8**
المجموع صفر	

ب/ مجموع مربع انحرافات القيم عن وسطها الحسابي يكون اقل ما يمكن $\sum (x-\overline{x})^2$

$(x-\bar{x})^2$	$x-\bar{x}$	x
4	**−2**	**4**
0	**0**	**6**
4	**2**	**2**
8	**0**	المجموع

ج/ لو أجرينا العمليات الحسابية على جميع الأرقام فإن الوسط الحسابي الجديد هو الوسط الحسابي للأرقام الأصلية مع إجراء نفس عملية الحسابية. (أي نفس العملية التي نجريها على الأرقام نجريها على الوسط الحسابي).

X	X+2	2−X	2÷X	2÷X
4	**6**	**2**	**8**	2
6	**8**	**4**	**12**	3
8	**10**	**6**	**16**	4
6 الوسط الحسابي	**8**	**4**	**12**	3

ملاحظه حول الوسط الحسابي:

1. يحسب للبيانات الكمية فقط.
2. تعريفه دقيق ويعطي قيمة واحده فقط.
3. يسهل التعامل معه جبرياً.
4. جميع الأرقام تدخل في حساب الوسط الحسابي.
5. يتأثر الوسط الحسابي بالقيم الشاذة. (مثل لو حسبنا معدل درجات الطلاب وكان هناك مجموعة قليلة حصلوا على صفر فإن هذه الأرقام تؤثر على تدني الوسط الحسابي).
6. لا يمكن حسابه في ظل وجود خانات مفتوحة.

الوسط الهندسي

تعريف الوسط الهندسي: هو الجذر النوني لحاصل ضرب الأعداد ببعضها، ورمزه **(G)**.

مثال الوسط الهندسي للأرقام: 7/5/3/2 هو:

$$G = \sqrt[4]{2 \times 3 \times 5 \times 7} = \sqrt[4]{210} = 3.8$$

قانون الوسط الهندسي : $G = \sqrt[n]{x_1.x_2.x_3.x_n}$

حيث إن **n** هي عدد الأرقام

مثال آخر: احسب الوسط الهندسي للقيم التالية: **4،6،3،8،9**

الحل: حيث أن عدد القيم **n=5** فإن الوسط الهندسي لها هو الجذر الخامس لحاصل ضرب هذه القيم

أي أن الوسط الهندسي هو:

$$G = \sqrt[5]{9 \times 8 \times 3 \times 6 \times 4}$$

$$G = \sqrt[5]{5184}$$

G=5.53

ملاحظات على الوسط الهندسي:

1. لحساب الوسط الهندسي لمجموعة من القيم لا يصح أن تكون إحدى القيم (أو بعضها) مساوياً للصفر.
2. لحساب الوسط الهندسي لمجموعة من القيم يجب أن تكون جميعها موجبه.
3. يجب أن يكون مجموع التكرارات مرة واحده فقط لكل فئة بحيث تصبح مجموع التكرارات = **N**

الوسط التوافقي

تعريف الوسط التوافقي: مقلوب الوسط الحسابي لمقلوبات هذه القيم. يرمز له بالرمز **(H)**.

يفضل استخدامه في بعض الحالات مثل حساب متوسط السرعة لمجموعة من السيارات والقطارات.

قانون الوسط التوافقي هو:

$$\frac{1}{H} = \frac{1}{n} \sum \frac{1}{Xi}$$

الوسيط

تعريف الوسيط

هو القيمة التي تقع في منتصف المجموعة بعد ترتيب القيم تصاعدياً أو تنازلياً. أي هو القيمة التي يكون نصف عدد القيم أصغر منها أو يساويها والنصف الآخر أكبر منها أو يساويها وسنرمز له بالرمز (**Q2**).

من هذا التعريف للوسيط نجد أنه يعالج العيوب الثلاثة التي يعاني منها الوسط الحسابي، فالوسيط لا يتأثر بالقيم الشاذة أو المتطرفة. كما أنه يمكن حسابه في حالة الفئات المفتوحة، ويمكن إيجاده بيانياً.

أولاً: الوسيط للبيانات غير المبوبة:

مثال (8،6،8،2،5،2) لحساب الوسيط نقوم أولاً بترتيبها تصاعدياً أو تنازليا :

الرقم المتسلسل	6	5	4	3	2	1
ترتيب الأرقام تصاعديا	8	8	6	5	2	2

ثم نضع لها رقماً متسلسلاً (1_6) في هذا المثال.

رتبة الوسيط: هي الرقم التسلسلي الذي تم وضعه أعلى الأرقام أعلاه (1_6) في هذا المثال

إذا كان عدد الأرقام (المسلسل) زوجي فتكون رتبة الوسيط رتبتين كالتالي:

الرتبة الأولى: $\mathbf{n}\div2 = 6\div2 = 3$، والرتبة الثانية: $\mathbf{n} \div 2+1 = 3+1 = 4$

فلدينا رتبتين للوسيط هي الرقمين 3 و4، ويكون حساب الوسيط بجمع العددين ÷ 2 كالتالي:

$\mathbf{Q2} = 5+6\div2 = 5{,}5$ (وهذا هو الوسيط) ، مثال: إذا كان عدد الأرقام فردي (التسلسل)

الرقم المتسلسل	7	6	5	4	3	2	1
ترتيب الأرقام تصاعديا	8	8	7	6	5	2	2

يكون حساب رتبة الوسيط $\mathbf{n} +1\div2 = 7+1\div2= 4$، ويكون الوسيط في المثال هو الرقم (6).

N هي عدد الأرقام (التسلسل).

ثانياً: الوسيط للفئات المبوبة على شكل فئات:

أ/ لا بد من حساب التكرار المتجمع الصاعد (ك م ص).

ب/ نوجد رتبة الوسيط وهي تساوي مجموع التكرارات $\sum \mathbf{f} \div 2$ ، وهي في المثال أدناه = 30÷2=15

الفئات	التكرار	ك م ص		أقل من	
2−	4	4		4	
4−	6	10		6	
6−	3	13	**f1**	8	**A**
8−	9	22	**f2**	10	
10−	4	26		12	
12−14	4	30		14	
المجموع	30				

ج/ نحدد موقع رتبة الوسيط في الجدول وهي في المثال تقع بين 13و22 (لو كانت رتبة الوسيط هي نفس أحد أرقام الجدول، فإن الوسيط هو الحد الأعلى المقابل لهذا الرقم.

د/ قانون الوسيط : $Q2 = A + \dfrac{\dfrac{\sum F}{2} - F_1}{F_2 - F_1}.L$

مثال: إذا كانت بيانات الجدول التالي احسب الوسيط لأعمار عمال أحد المصانع.

فئات العمر	عدد العمال
15−	**5**
25−	**8**
35−	**7**
45−	**6**
60−55	**4**
المجموع	**30**

الحل

تكوين الجدول المتجمع الصاعد.

الحدود العليا للفئات	التكرار المتجمع الصاعد
أقل من 25	**5**
أقل من 35	**13** F_1
أقل من 45	**20** F_2
أقل من 55	**26**
أقل من 60	**30**

الفئة الوسيطية

ترتيب الوسيط = $\dfrac{\sum F}{2}$ = $\dfrac{30}{2} = 15$

نجد ترتيب الوسيط يقع بين التكرارين المتجمعين الصاعدين 20،13

وبالتالي تكون الفئة الوسيطية هي من 35إلى أقل من 45 وبالتالي نجد الوسيط حسب القانون التالي:

$$Q2 = A + \frac{\frac{\sum F}{2} - F_1}{F_2 - F_1} \times L$$

$$Q2 = 35 + \frac{\frac{30}{2} - 13}{20 - 13} \times 10 = 35 + \frac{2}{7} \times 10 = 35 + \frac{20}{7} = 35 + 2.85 = 37.85$$

إذا الوسيط هو = 37.85 لأعمار عمال أحد المصانع.

المنوال

تعريف المنوال: هو القيمة التي تتكرر أكثر من غيرها ويرمز له بالرمز **M**.

الحالة الأولى: لا يوجد منوال: إذا كانت جميع الأرقام تكررت بنفس العدد مثل (2،3،5،7،8).

الحالة الثانية: يوجد منوال واحد: إذا كان هناك رقم واحد تكرر أكثر من (2،3،3،5،7،8) فالمنوال = 3.

الحالة الثالثة: يوجد أكثر من منوال: إذا كان هناك أكثر من رقم تكرر بنفس العدد (2،3،3،5،5،7،8)

M = 3و5.

أما المنوال في حالة البيانات المبوبة بقيم وتكرارات فيكون حسابه كالتالي:

أولاً: **تحديد الفئة الموالية:** وهي الفئة التي بها أكثر التكرارات، في المثال أدناه (8_) بها 9 تكرارات.

الفئات		التكرار	
2-		4	
4-		6	
6-		3	**f1**
8-	**A**	9	
10-		4	**f2**
14-12		4	
المجموع		30	

ملاحظات على المنوال

1. إذا كانت الفئة المنوالية هي الفئة الأولى (أو الأخيرة) في الجدول فالمنوال يساوي **مركز الفئة المنوالية.**
2. إذا كان التكرار السابق يساوي التكرار اللاحق فان كل الطرق تؤدي إلى النتيجة نفسها ويكون المنوال **مساوياً لمركز الفئة المنوالية.**
3. إذا كان الجدول التكراري أكثر من فئة منوالية يكون **هناك أكثر من منوال.**
4. المنوال يعتبر أبسط المتوسطات.
5. المنوال هو المتوسط الوحيد الذي يمكن إيجاده للبيانات الوصفية سواء الاسمية أو الترتيبية.

مقاييس التشتت

مقاييس التشتت (المدى – الانحراف المتوسط – الانحراف المعياري والتباين –– معامل الاختلاف –التغير المعياري)

أولاً: المدى

المدى هو: الفرق بين أكبر قيمة وأصغر قيمة ويرمز له بالرمز **D**

حيث المدى = أكبر قيمة – اصغر قيمة

المدى يعتبر ابسط مقاييس التشتت لسهوله حسابه ووضوح معناه.

الحالة الأولى: إذا كانت على شكل قيم وتكرارات فيطبق التعريف مباشرة هو (الفرق بين أكبر قيمة وأصغر قيمة).

الحالة الثانية: في حالة الفئات فان المدى يكون الفرق بين الحد الأعلى للفئة والحد الأدنى للفئة.

ثانياً: الانحراف المتوسط.

هو الوسط الحسابي للقيم المطلقة لانحراف القيم عن وسطها الحسابي.

خطوات حساب الانحراف المتوسط هي:

1. حساب الوسط الحسابي.
2. حساب الانحرافات المطلقة لجميع القيم عن وسطها الحسابي.
3. حساب متوسط هذه الانحرافات المطلقة (بجمعها والقسمة على عددها).

مثال: احسب الانحراف المتوسط للقيم التالية: **3،9،8،4،6**

الحل: 1. حساب الوسط الحسابي:

$$\overline{X} = \frac{6+4+8+9+3}{5} = \frac{30}{5} = 6$$

2. الانحرافات المطلقة للقيم عن الوسط الحسابي

$$|6-6| , |6-4| , |6-8| , |6-9| , |6-3|$$

0 2 2 3 3

الانحراف المتوسط = الوسط الحسابي للانحرافات المطلقة:

$$MD = \frac{\sum x - x}{n}$$

$$= \frac{0+2+2+3+3}{5} = \frac{10}{5} = 2$$

إيجاد الوسط الحسابي وبطريقة الفئات المبوبة:

حسب القانون التالي: حاصل مجموع ضرب انحرافات القيم المطلقة عن وسطها الحسابي في تكرار كل فئة تقسيم مجموع التكرارات $MD = \frac{\sum F\left|X - \overline{X}\right|}{\sum F}$

أ/ إيجاد الوسط الحسابي $\overline{X}$ حسب القانون $\overline{X} = \frac{\Sigma f x}{\Sigma f}$ ويساوي 240÷30= 8

ب/ الانحراف عن الوسط الحسابي $X - \overline{X}$

ج/ نضرب كل انحراف في تكرار كل فئة (الضرب للقيم المطلقة يعني إهمال الإشارات للعدد).

د/ نجمع حاصل الضرب في الخطوة السابقة (144) حسب المثال.

هـ/ نقسم الحاصل على مجموع الفئات

الفئات	التكرار f	X مركز الفئة	fx	$X - \overline{X}$	القيم المطلقة $\left|X - \overline{X}\right|$	$f\left|X - \overline{X}\right|$
−2	4	3	13	5−	5	15
−4	6	5	30	3−	3	15
−6	3	7	21	1−	1	7
−8	9	9	82	1	1	9
−10	4	11	44	3	3	44
14−12	4	13	52	5	5	65
المجموع	30		240			144

إذن بناء على القانون $MD = \frac{\sum F\left|X - \overline{X}\right|}{\sum F}$ يصبح الانحراف المتوسط 144÷30=4.8

ثالثاً: التباين والانحراف المعياري.

التباين : هو الوسط الحسابي لمربعات انحرافات القيم عن وسطها الحسابي ، ويرمز له بالرمز σ

تعريف الانحراف المعياري : هو الجذر التربيعي للتباين ورمزه **S**

حساب الانحراف المعياري للقيم المطلقة:

مثال: (6،4،8،9،3)، وسطها الحسابي = المجموع ÷ العدد أي 30 ÷ 5 = 6

x	الانحراف عن الوسط الحسابي $X-\overline{X}$	$(X-\overline{X})2$	f I $X-\overline{X}$ I
6	0	0	
4	2−	4	
8	2	4	
9	3	9	
3	3−	9	
		26	

قانون التباين هو : $\sigma^2 = \frac{\sum(X-X)^2}{n}$ = 5÷26 = 5,2

الانحراف المعياري : هو جذر التباين = $\sigma = \sqrt{5,2} = 2,28$

حساب التباين والانحراف المعياري للبيانات المبوبة:

حساب التباين حسب القانون التالي : $\sigma^2 = \frac{\sum F(X-\overline{X})^2}{\sum F}$

ومن القانون تتضح خطوات حساب التباين والانحراف المعياري كالتالي:

إيجاد الوسط الحسابي $\overline{X}$ يساوي 240÷30= 8

ب/ حساب انحرافات القيم عن وسطها الحسابي.

ج/ حساب مربعات انحرافات القيم عن وسطها الحسابي.

د/ نضرب تكرار كل فئة في التربيع السابق، ونجمع حاصل الضرب

ه/ نقسم المجموع الناتج على مجموع التكرارات لنحصل على التباين.

و/ نحسب الجذر التربيعي للتباين فنحصل على الانحراف المعياري **S**.

الفئات	التكرار f	مركزX الفئة	fx	$X-\overline{X}$	$(X-\overline{X})^2$	$f(X-\overline{X})^2$
−2	4	3	13	5−	25	**100**
−4	6	5	30	3−	9	**54**
−6	3	7	21	1−	1	**3**
−8	9	9	82	1	1	**9**
−10	4	11	44	3	9	**36**
14−12	4	13	52	5	25	**100**
المجموع	30		240			**302**

وحسب القانون أعلاه يصبح التباين = 30÷302 = 10.1

والانحراف المعياري يساوي جذر التباين : $\sqrt{10{,}1}$ = 3،18

ملاحظات مهمة على التباين والانحراف المعياري:

1. يقاس الانحراف المعياري بالوحدات نفسها التي يقاس بها المتغير بينما يقاس التباين بوحدات مربعة.
2. لا يتأثر التباين والانحراف المعياري بالجمع والطرح.
3. يتأثر الانحراف المعياري بالضرب والقسمة.

رابعاً: التشتت النسبي (معامل الاختلاف)

يستخدم عند مقارنة تشتت مجموعتين أو أكثر من القيم إذا كانت تختلف في وسطهما الحسابي أو تختلف في وحدات القياس، ويرمز له بالرمز **C.V**. وقاعدة حسابه حس القاعدة التالية : (الانحراف المعياري ÷ الوسط الحسابي×100)

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

خامساً: المتغير المعياري

يستخدم بين التوزيعات المختلفة، يستخدم لدراسة ومقارنة التشتت بين مجموعتين أو أكثر ، يقيس الانحرافات عن المتوسط الحسابي.

وقانونه هو = $Z=\dfrac{X-\overline{X}}{\sigma}$ مثال:

المادة	الدرجة X	الوسط الحسابي	الانحراف المعياري
الرياضيات	**85**	**80**	**3**
الإحصاء	**80**	**70**	**4**

المتغير المعياري لمادة الرياضيات = $Z=\dfrac{85-80}{3}=\dfrac{5}{3}=1.67$

المتغير المعياري لمادة الإحصاء = $Z=\dfrac{80-70}{4}=\dfrac{10}{4}=2.5$

المتغير المعياري للإحصاء أكبر من المتغير المعياري للرياضيات ومنه نستنتج أن الطالب أكثر استيعابا في مادة الإحصاء.

الارتباط

الارتباط: هو قياس قوة العلاقة بين متغيرين واتجاه هذه العلاقة. ويرمز له بالرمز **(r)**.

العامل الذي يؤثر يسمي العامل المستقل **X**، والعامل الذي يتأثر يسمي العامل التابع **Y**.

قانون الارتباط هو =

$$r = \frac{n\sum XY - (\sum X)(\sum Y)}{\sqrt{n\sum X^2 - (\sum X)^2}\sqrt{n\sum Y^2 - (\sum Y)^2}}$$

مثال: احسب معامل الارتباط الخطي البسيط للبيانات التالية:

عدد العمال X	الإنتاج Y
1	**6**
2	**11**
3	**15**
4	**19**
5	**21**
15	72

الحل: يجب تطبيق القانون التالي

$$r = \frac{n\sum XY - (\sum X)(\sum Y)}{\sqrt{n\sum X^2 - (\sum X)^2}\sqrt{n\sum Y^2 - (\sum Y)^2}}$$

ولتطبيق القانون يجب إيجاد قيم كل من

قيمة **N**= هي عدد القيم وهي (5)

قيمة = **X** وقيمة **Y** وقيمة X^2 وقيمة Y^2 وقيمة **XY** وطريق الحل حسب الجدول التالي

عدد العمال X	الإنتاج Y	X^2	Y^2	XY
1	**6**	**1**	**36**	**6**
2	**11**	**4**	**121**	**22**
3	**15**	**9**	**225**	**45**
4	**19**	**16**	**361**	**76**
5	**21**	**25**	**441**	**105**
15	**72**	**55**	**1184**	**254**

أيضاً نوجد $(\sum Y)^2$ وهو = 72×72 = 5184

أيضاً نوجد $(\sum x)^2$ وهو = 15×15 = 225

بعد إيجاد القيم نطبق القانون دون تغيير

$$r = \frac{n\sum XY - (\sum X)(\sum Y)}{\sqrt{n\sum X^2 - (\sum X)^2}\sqrt{n\sum Y^2 - (\sum Y)^2}}$$

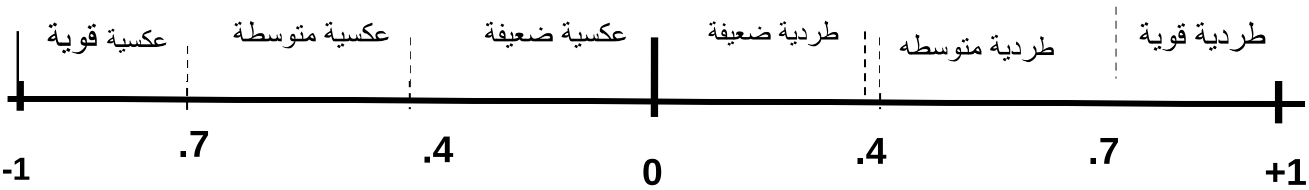
$$r = \frac{5(254) - (15 \times 72)}{\sqrt{5(55) - (15)^2}\sqrt{5(1184) - (72)^2}}$$

$$r = \frac{1270 - 1080}{\sqrt{275 - 225\sqrt{5920 - 5184}}}$$

$$r = \frac{190}{\sqrt{50\sqrt{736}}} = \frac{190}{\sqrt{36800}} = \frac{190}{191.83} = .99$$

وبهذه النتيجة أصبح حساب معامل الارتباط الخطي هو (**.99**) ويعتبر ارتباط طردي قوي (لأن الإشارة موجبة) ، وإذا كانت الإشارة سالبة فإن النتيجة هي ارتباط عكسي .

ولكن إذا كانت النتيجة بإشارة موجب تصبح ارتباط طردي وإذا كانت النتيجة بإشارة السالب تكون ارتباط عكسي



.

ملاحظة : يجب أن تكون إجابة معامل الارتباط أن لا تتجاوز (**+1**) صحيح ، فإذا وجدنا قيمة اكبر من **(1)** فنقول الناتج خاطئ. ، فيجب أن تنحصر القيمة بين $-1 \leq \& \leq +1$.

إذا كانت العلاقة (0) فنقول عنها معدومة ، وإذا كانت (1) فنقول علاقة تامة .

بعض خصائص معامل الارتباط.

1. معامل الارتباط بين **Y،X** هو نفسه بين **X،Y** .
2. تنحصر قيمة معامل الارتباط بين **1−،1+** وإذا كانت قيمة معامل الارتباط بالموجب فان الارتباط يكون طردياً وإذا كانت قيمة معامل الارتباط بالسالب فن الارتباط يكون عكسياً.
3. قيمة معامل الارتباط بين العامل المتغير ونفسه هي (1) .
4. لا يتأثر معامل الارتباط بأي من العمليات الحسابية أو الجبرية التي تجري على **Y،X** أي لا يتأثر بالطرح والجمع أو الضرب والقسمة.
5. إذا كان المتغيران **Y،X** مستقلين فإن قيمة معامل الارتباط صفراً ولكن العكس غير صحيح.

الانحدار

الانحدار: يهتم بصياغة العلاقة بين المتغيرين **Y،X** على شكل معادلة رياضية بحيث يمكن التنبؤ بقيمة أحد المتغيرين بمعرفة قيمة المتغير الآخر.

فإذا كانت **X** هي المتغير المستقل، و**Y** هي المتغير التابع، فإن الانحدار الخطي البسيط = **Y= A+B X**

إذا كانت **X** = فر فإن **Y=A**).

ولإيجاد **.b**

نتبع القانون التالي : $b = \dfrac{\sum XY - \sum X \times \sum Y}{n\sum X^2 - (\sum X)^2}$

وعندما نستخدم القانون السابق نجد قيمة **b**

وبعد ذلك نطبق القانون التالي لنجد قيمة **a**

$$a = \frac{\sum Y}{n} - b\frac{\sum X}{n}$$

مثال: نطبق التمرين السابق الخاص بالارتباط لنجد قيمة الانحدار.

عدد العمال **X**	الإنتاج **Y**	X^2	Y^2	**XY**
1	**6**	**1**	**36**	**6**
2	**11**	**4**	**121**	**22**
3	**15**	**9**	**225**	**45**
4	**19**	**16**	**361**	**76**
5	**21**	**25**	**441**	**105**
15	**72**	**55**	**1184**	**254**

الحل: أولا نجد قيمة **b**

نستخدم القانون التالي

$$b = \frac{n\sum XY - \sum X \times \sum Y}{n\sum X^2 - (\sum X)^2}$$

$$b = \frac{5 \times 254 - (15 \times 72)}{5(55) - (225)} = \frac{1270 - 1080}{275 - 225} = \frac{190}{50} = 3.8$$

وبذلك أصبحت قيمة **b= 3.8**

بعد ذلك نجد قيمة **a**

$$a = \frac{\sum Y}{n} - b\frac{\sum X}{n}$$

$$a = \frac{72}{5} - 3.8\frac{15}{5} = 14.4 - 11.4 = 3$$

وإذا وجدنا قيمة **a & b** نستخدم القانون التالي لإيجاد الانحدار حسب القانون التالي: **Y= A+BX** $\left(Y = 3 + (3.8)(X)\right)$

وإذا ورد في السؤال **10**عمال كم الإنتاج المتوقع : نعوض في الناتج السابق $Y = 3 + (3.8)(10)$ = $Y = 3 + 38 = 41$

السلاسل الزمنية

السلاسل الزمنية : هي مجموعة القيم التي يأخذها هذا المتغير في فترات زمنية متتالية .

مكونات السلسلة الزمنية: 1- التغيرات الاتجاهية أو الاتجاه العام 2-التغيرات الموسمية 3- التغيرات الدورية 4- التغيرات العشوائية.

التغيرات الاتجاهية او الاتجاه العام نرمز لها بالرمز **T**

مثال يوضح معادلة خط الاتجاه العام.

السنوات	قيمة الانتاج	X
1423	3	1
1424	5	2
1425	4	3
1426	7	4
1427	5	5
1428	8	6

كم الانتاج المتوقع للعام 1432هـ

حسب المعادلة التالية

Y= a+bx

السنوات	قيمة الانتاجy	X (نفترضها)
1423	3	1
1424	5	2
1425	4	3
1426	7	4
1427	5	5
1428	8	6
1429		7
1430		8
1431		9
1432		10

قيمة **x** باللون الأصفر هي قيمة افتراضية

حسب المعادلة: y=2.5+.8X وقيمة Xفي 1432 هي 10

2.5+.8(10) =

= 2.5+8 = 10.5

10.5 الإنتاج المتوقع للعام 1432.

الأرقام القياسية

الرقم القياسي هو : رقم نسبي يقيس التغير الذي يطرأ على ظاهرة أو مجموعه من الظواهر خلال فترتين زمنيتين أو منطقتين جغرافيتين .

مثال على ذلك :

الأساس 1420	المقارنة 1428
30	**40**

نوجد الرقم القياسي بتطبيق القانون التالي = $I = \frac{40}{30} \times 100 = 133\%$

<u>الرقم القياسي التجميعي البسيط:</u>

وقانونه كالتالي : $I = \frac{\sum P_1}{\sum P_0} \times 100$

مثال:

أسعار **1420**	أسعار **1428**
رمز سنة الأساس p_0	رمز سنة الأساس p_1
5	**15**
25	**35**
35	**40**
65	**90**

نطبق القانون التالي = $I = \frac{\sum P_1}{\sum P_0} \times 100$ $\quad I = \frac{90}{65} \times 100 = 138\%$

<u>الرقم القياسي النسبي البسيط:</u>

قانون الرقم النسبي البسيط = $I = \frac{1}{n} \sum (\frac{p_1}{p_0}) \times 100$

ولو طبقناه على المثال السابق = $I = \frac{1}{3}(\frac{15}{5} + \frac{35}{25} + \frac{40}{35}) \times 100$ $\quad I = \frac{1}{3}(3 + 1.4 + 1.14) \times 100$

$$I = \frac{1}{3}(5.45) \times 100 = 1.84 \times 100 = 184$$

<u>الأرقام المرجحة:</u>

<u>الرقم القياسي التجميعي المرجح (لاسبير):</u>

وقانونه كالتالي : $I = \frac{\sum P_1 Q_0}{\sum P_0 Q_0} \times 100$

مثال:

أسعار 1420	أسعار 1428	كميات 1420	كميات 1428	الوزن المرجح للأساس	الوزن المرجح للمقارنة	$P_1 \times Q_0$	$P_0 \times Q_0$
p_0	p_1			Q0	Q1		
5	15	20	60	0.17=120÷20	0.3=200÷60	2.55	.85
25	35	40	90	0.33=120÷40	0.45=200÷90	11.55	8.25
35	40	60	50	0.5=120÷60	0.23=200÷50	20	17.5
65	90	120	200			34.1	26.6

نطبق القانون التالي $I = \frac{\sum P_1 Q_0}{\sum P_0 Q_0} \times 100$ = $I = \frac{34.1}{26.6} \times 100 = 1.28 \times 100 = 128$

<u>**الرقم القياسي النسبي المرجح (لاسبير) : قانونه كالتالي :**</u> $\sum (\frac{P_1}{P_0}) Q_0 \times 100 =$

أسعار 1420	أسعار 1428	كميات 1420	كميات 1428	الوزن المرجح للأساس	الوزن المرجح للمقارنة	$P_1 \times Q_0$	$P_0 \times Q_0$	$\frac{P_1}{P_0}$	$(\frac{P_1}{P_0}) Q_0$
p_0	p_1			Q0	Q1				
5	15	20	60	.17	.3	2.55	.85	3	.51
25	35	40	90	.33	.45	11.55	8.25	1.4	.46
35	40	60	50	.5	.25	20	17.5	1.14	.57
65	90	120	200			34.1	26.6	5.54	1.54

ثم نقوم بتطبيق القانون التالي : $\sum (\frac{P_1}{P_0}) Q_0 \times 100 =$

1.54 X100=154

الرقم القياسي التجميعي المرجح (باش) : قانونه هو : $I = \frac{\sum P_1 Q_1}{\sum P_0 Q_1} \times 100$

مثال :

$\frac{P_1}{P_0}$	$P_0 \times Q_1$	$P_1 \times Q_1$	الوزن المرجح للمقارنة	الوزن المرجح للأساس	كميات 1428	كميات 1420	أسعار 1428	أسعار 1420
			Q1	Q0			p_1	p_0
3	1.5	4.5	.3	.17	60	20	15	5
1.4	11.25	15.75	.45	.33	90	40	35	25
1.14	8.75	10	.25	.5	50	60	40	35
5.54	21.5	30.25			200	120	90	65

نعوض في القانون (باش)

$$I = \frac{\sum P_1 Q_1}{\sum P_0 Q_1} \times 100 \qquad I = \frac{30.25}{21.5} \times 100 = 1.82 \times 100 = 182$$

الرقم القياسي النسبي المرجح (باش) : قانونه هو : $\sum (\frac{P_1}{P_0}) Q_1 \times 100 =$

$(\frac{P_1}{P_0})Q_1$	$\frac{P_1}{P_0}$	$P_0 \times Q_1$	$P_1 \times Q_1$	الوزن المرجح للمقارنة	الوزن المرجح للأساس	كميات 1428	كميات 1420	أسعار 1428	أسعار 1420
				Q1	Q0			p_1	p_0
.9	3	1.5	4.5	.3	.17	60	20	15	5
.63	1.4	11.25	15.75	.45	.33	90	40	35	25
.28	1.14	8.75	10	.25	.5	50	60	40	35
1.82	5.54	21.5	30.25			200	120	90	65

نعوض في القانون : $\sum (\frac{P_1}{P_0}) Q_1 \times 100 =$ = $1.82 \times 100 = 182$