الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne Démocratique et Populaire
وزارة التعليم العالي والبحث العلمي
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
جامعة الشهيد حمه لخضر - الوادي
Université Echahid Hamma Lakhdar-El oued

كلية العلوم الاقتصادية و التجارية و علوم التسيير
المجلس العلمي للكلية

Faculté des sciences économiques, commerciales et sciences de gestion

جامعة الشهيد حمه لخضر - الوادي
Université Echahid Hamma Lakhdar - El-Oued

الرقم: ..... /م ع/ك ع ا ق ت وع ت/2022 الـوادي في: 2022/05/23

شهادة اعتماد دروس عبر الخط

يشهد رئيس المجلس العلمي لكلية العلوم الاقتصادية والتجارية وعلوم التسيير بجامعة الشهيد حمه لخضر الوادي على أن:

د/ أحمد بن أحمد............................ أستاذ محاضر أ

قدم للمجلس العلمي للكلية عن طريق اللجنة العلمية لقسم العلوم المالية والمحاسبية دروس عبر الخط في مقياس **الإحصاء 1 ملخصات ومسائل محلولة** مقدمة لطلبة سنة أولى جذع مشترك (مسار العلوم الاقتصادية والتجارية وعلوم التسيير)، للموسم الجامعي 2022/2021.

وبعد التقارير الايجابية المقدمة من طرف لجنة القراءة والتحكيم، فإنه يتم اعتماد الدروس عبر الخط وترفع في موقع الدروس عبر الخط https://elearning.univ-eloued.dz/course/view.php?id=5156

رئيس المجلس العلمي

رئيس المجلس العلمي لكلية العلوم الإقتصادية والتجارية وعلوم التسيير
أ.د/ إلياس شاهد

الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne démocratique et Populaire
وزارة التعليم العالي والبحث العلمي
Ministère de l'Enseignement Supérieure et de la Recherche Scientifique
جامعة الشهيد حمّه لخضر - الوادي
Université Echahid Hamma Lakhdar - El Oued

Faculté de science économique et commerce et science gestion
Département des Sciences financières et comptables
Comité scientifique du Département

كلية العلوم الاقتصادية والتجارية وعلوم التسيير
قسم العلوم المالية والمحاسبية
اللجنة العلمية للقسم

الرقم: **007**/ ل ع ق ع م م /ك.ع. ق.ت.ع.ت/2022

الوادي في:2022/05/18

شهـــادة إداريـــة
(اعتماد دروس على الخط)

يشهد السيد رئيس اللجنة العلمية لقسم العلوم المالية والمحاسبية بأن الدكتور: **بن أحمد أحمد** ، قد قدم أمام اللجنة العلمية لقسم العلوم المالية والمحاسبية دروس على الخط بعنوان:

الإحصاء 1 ملخصات ومسائل محلولة

مقدمة لطلبة السنة أولى جذع مشترك (مسار العلوم الاقتصادية والتجارية وعلوم التسيير) ، وبعد تعين لجنة الخبرة وبعد استلام تقارير الخبرة الإيجابية، واستيفاء المطبوعة الجوانب الشكلية والمنهجية والشروط المعتمدة في كلية العلوم الاقتصادية والتجارية وعلوم التسيير بجامعة الشهيد حمه لخضر، تم اعتماد دروس على الخط للتدريس في المستوى والتخصص المذكور أعلاه (عدد الصفحات 63).

رئيس اللجنة العلمية للقسم

رئيس اللجنة العلمية لقسم العلوم المالية والمحاسبية
أ.د محمد الهادي ضيف الله

وزارة التعليم العالي والبحث العلمي

جامعة الشهيد حمة لخضر – الوادي

كلية العلوم الاقتصادية والعلوم التجارية وعلوم التسيير

موجهة للسنة الأولى جذع مشترك

دروس على الخط في مقياس :

الاحصاء 1

ملخصات ومسائل محلولة

من إعداد:

د. بن أحمد أحمـد

أستاذ محاضر أ

السنة الجامعية: 2021 / 2022

1) تعريف علم الإحصاء:

يهتم علم الإحصاء بجمع البيانات (حول ظاهرة ما) وتلخيصها وتبويبها للاستفادة منها في وصف الظواهر وتحليلها وقياس واستقراء الوقائع للتنبؤ بسلوك الظاهرة حاضرا ومستقبلا لاتخاذ قرارات سليمة في ظروف عدم التأكد

ونستفيد من هذا التعريف أن:

- البيانات تكون متجانسة أي من نفس الصنف.
- توجد بين البيانات علاقات عددية مستقلة عن الصدفة.
- قياس الظواهر لاستقراء الواقع واستخراج العلاقات.
- تحليل البيانات للتنبؤ بالقيم المستقبلية للظاهرة المدروسة وفق شروط معينة.
- اتخاذ القرارات المناسبة من خلال تحليل سلوك الظواهر.

2) وظائف علم الإحصاء:

2-1) وصف البيانات:

تكون البيانات خاما ولا يمكن الاستفادة منها إلا إذا تم جمعها وعرضها إما في شكل جدولي أو بياني وحساب بعض المؤشرات الإحصائية (الوسط، المنوال، الانحراف، التباين.....الخ) التي تدلنا على طبيعة البيانات.

2-2) الاستدلال الإحصائي:

ويستند إلى فكرة اختيار جزء من المجتمع يسمى عينة بطريقة علمية مناسبة، واستخدام بيانات هذه العينة في التوصل إلى نتائج يمكن تعميمها على مجتمع الدراسة، ويمر الاستدلال الإحصائي بمرحلتين:

2-2-1) التقدير:

وهو معرفة معالم المجتمع (الخصائص التي تميزه كالمتوسط والتباين والتوزيع) عن طريق حساب مؤشرات تقديرية من بيانات العينة تسمى إحصائيات، نستخدم إحصائيات العينة لتقدير معالم المجتمع. أي تقدير متوسط المجتمع مثلا من خلال معرفة متوسط العينة، وهذا ما يسمى التقدير بنقطة أما إذا أردنا أن نقدر من خلال العينة المجال الذي يمكن أن تقع فيه معلمة المجتمع (متوسط المجتمع) فنسمي ذلك التقدير بفترة.

2-2-2) اختبار الفرضيات(الفروض): يتم تحديد فروض معينة حول خصائص ظاهرة مدروسة (معالم المجتمع) واستخدام بيانات العينة للوصول إلى قرار علمي سليم بخصوص الفروض المحددة حول معالم المجتمع.

(فرضية: مثلا تتأثر ظاهرة التسرب المدرسي بالدخل الشهري للعائلة)

3-2-2) التنبؤ: نستخدم نتائج الاستدلال الإحصائي التي تدلنا على سلوك الظاهرة في الماضي لمعرفة ما يمكن أن يحدث لها في الحاضر والمستقبل.

3) مفاهيم أساسية:

1-3) الظاهرة: وهي الخاصية أو المتغير المدروس في المجتمع الإحصائي (الطول، السن، علامة الامتحان، الادخار، الاستهلاك، الدخل....)، وقد تتكون الظاهرة من متغير واحد أو عدة متغيرات فظاهرة التسرب المدرسي يمكن أن نجد فيها عدة متغيرات: (سن الطفل، بعد المدرسة عن المنزل، دخل الأسرة، عدد الأفراد في الأسرة، معاملة المعلم للطفل).

3-2) المجتمع الإحصائي: وهو مجموعة من الوحدات الإحصائية تتعلق بظـــاهرة قابلة للقياس وقد يكون المجتمع متجانسا (الأسر، المؤسسات) أو غيـــر متجـــانس (كالديانات، والجنسيات).

3-3) الوحدة الإحصائية: وهي أصغر وحدة أساسية لتكوين المجتمع المدروس وتسمى أيضا المفردة أو المشاهدة، المجتمع الإحصائي (مجتمع الأسر، مجتمع المؤسسات)، والوحدات الإحصائية بالترتيب (الأسرة، المؤسسة).

3-4) العينة: هي جزء من المجتمع والذي يمثل المجتمع أحسن تمثيل .

4) أنواع البيانات:

4-1) بيانات وصفية: (نوعية أو كيفية) وهي بيانات غير رقمية مرتبة في شكل مستويات أو شكل فئات رقمية، ، ويتم قياسها بمعيارين هما:

4-1-1) معيار اسمي: تقاس به البيانات غير الرقمية التي تتكون من مجموعات (حالات) متنافية، وكل مجموعة لها خصائص تميزها عن المجموعات الأخرى حيث لا يمكن مقارنة المجموعات (الحالات) أو ترتيبها، وقد يتم ترميزها أي إعطاؤها أرقاما لتسهيل معالجتها بالحواسيب حيث لا تمثل الأرقام قيمتها ولكن تمثل نوع الحالة فقط.

أمثلة: متغير الجنس: وحالاته (ذكر 1، أنثى2).

متغير الحالة الاجتماعية: وحالاته (أعزب1، متزوج2، أرمل3، مطلق4).

متغير أصناف التمر: وحالاته (......).

متغير الجنسية: وحالاته (جزائري1، مصري2، فرنسي3....).

4-1-2) معيار ترتيبي: تقاس به البيانات غير الرقمية التي تتكون من مستويات أو فئات يمكن ترتيبها تنازليا أو تصاعديا: أمثلة: متغير علامات الطلبة في المنهجية الإنجليزية: وحالاته ($A\ A^+\ B\ B^+\ C\ C^+\ D\ D^+$).

متغير المستوى التعليمي: وحالاته (باحث، جامعي، ثانوي، أساسي، ابتدائي).

متغير الرأي: وحالاته (موافق تماما، موافق، غير موافق، غير موافق تماما).

4-2) بيانات كمية: هي بيانات يعبر عنها بأرقام عددية تمثل القيمة الفعلية للظاهرة وتنقسم لقسمين هما.

4-2-1) بيانات فترة (متقطعة أو منفصلة): هي بيانات رقمية تقاس بمقدار بعدها عن الصفر، ومتقطعة لأنها تأخذ قيما صحيحة ثابتة لا يمكن تجزئتها فعدد الأطفال يأخذ القيم1-2-3-4... ولكن لا يمكن أن يأخذ القيم 2,5 3,256....

4-2-1) بيانات نسبة (مستمرة أو متصلة): وتأخذ قيما غير متناهية في مجال صغير حيث يمكن للمتغير أن يأخذ قيما لا متناهية بين قيمتين مثل أطوال الطلبة التي قد تأخذ أي قيمة بين 175سم و176سم مثل 175,12 سم أو 175,47 سم...الخ.

قاعدة: كل متغير كمي يقاس قياسا هو متغير مستمر، وكل متغير كمي يعد ويحسب حسابا فهو متغير متقطع.

5) مصادر البيانات:

5-1) مصادر أولية: تمدنا بالبيانات بشكل مباشر من مصدرها حيث يقوم الباحث بجمعها من مفردات محل البحث مباشرة، إذا كان البحث حول الأسرة فجمع البيانات يتم من رب الأسرة مباشرة أو من ينوب عنه.

- تتميز بيانات هذه المصادر بالشمولية، التعميم، الدقة والثقة، لأن الباحث يجمعها بنفسه.
- ومن عيوبها أنها تحتاج للوقت والجهد وتكلفتها المادية الكبيرة نسبيا.

5-2) مصادر ثانوية: تمدنا بالبيانات بشكل غير مباشر (من أشخاص آخرين غير معنيين بالبحث)، كالأجهزة والهيئات الرسمية المتخصصة (نشرات الوزارات، وتقارير المعهد الوطني للإحصاء والتخطيط....). وتسميتها بالثانوية لا يعني التقليل من أهميتها ولكنها تعبر ثانوية بالنسبة للأولى.

- تتميز بيانات هذه المصادر بتوفير الجهد والوقت والمال.
- ومن عيوبها أنها تنقصها الثقة والدقة وليست بنفس الدرجة كالمصادر الأولية لأن هدف الباحث من جمعها يختلف عن هدف الهيئات التي نشرتها.

6) طرق جمع البيانات:

وتعتبر من أهم مراحل البحث الإحصائي حيث أن جمعها بأسلوب علمي يفيدنا بالوصول إلى نتائج دقيقة في التحليل، وتجد طريقتين لجمع البيانات وذلك حسب كبر المجتمع والهدف من البحث:

- **أسلوب الحصر الشامل:** أي حصر وعد كامل المجتمع ويتميز بالشمول وعدم تحيز البيانات ودقة النتائج، ومن عيوبه أنه يحتاج للوقت والمجهود والتكلفة العالية.

- **أسلوب المعاينة:** اختيار جزء من المجتمع محل الدراسة بطريقة علمية سليمة عن طريق مجموعة من العمليات بهدف تكوين عينة، ودراستها ثم تعميم النتائج على المجتمع، حيث يفضل هذا الأسلوب في حالة كبر حجم المجتمع ويصعب إجراء حصر شامل لمفرداته مثل دراسة المصابيح الكهربائية المنتجة في مصنع ما، تعداد أسماك البحر، معاينة نوع فصيلة الدم للسكان في بلد ما، ويتميز هذا الأسلوب بتقليل الجهد والوقت والتكلفة وبياناته أكثر تفصيلا، ومن عيوبه نقص الدقة والثقة (خاصة إذا كانت العينة لا تمثل المجتمع جيدا). ويتوقف نجاح استخدام هذا الأسلوب على:
- كيفية تحديد حجم العينة n،
- طريقة اختيار مفردات العينة $X_1, X_2, X_3 \dots X_n$،
- نوع العينة المختارة.

7) أنواع المعاينات:

- **المعاينة غير الاحتمالية:** يتم اختيار مفردات العينة بطريقة غير عشوائية أي مقصودة وينزع منها عنصر الصدفة، حيث تختار مفردات العينة بصورة تحقق الهدف من المعاينة، ومن أنواعها:

_ **المعاينة العرَضِية:** سحب عينة من مجتمع البحث حسبما يليق بالباحث (نختار مثلا من عمال مصنع ما فقط العمال الذين يترددون على مقهى معين).

_ **المعاينة العَمْدية (القصدية):** سحب عينة من مجتمع البحث بانتقاء عناصر مثالية (إذا كان البحث حول طبيعة الاهتمامات الاجتماعية للطلبة فسنختار طلبة العلوم الاجتماعية لأنهم أكثر اهتماما وباعتبارهم عناصر مثالية).

- **المعاينة الاحتمالية (العشوائية):** إذا كان لكل عنصر من مجتمع البحث نفس الحظ للظهور في العينة عن طريق السحب العشوائي دون تكرار أو نسيان ومن أنواعها:

_ **المعاينة البسيطة:** إعطاء السحب بالصدفة ميزة علمية لأخذ عينة من كل عناصر البحث.

- **المعاينة المنتظمة:** نفس طريقة المعاينة البسيطة ولكن بترتيب وترقيم مجتمع البحث، فإذا كان حجم مجتمع البحث هو 1000 فإننا نرقم مفرداته من 1 إلى 1000 ثم نقسم 1000 على حجم العينة ليفترض أنه 100 فنجد 10=100/1000 فينتج لنا 100 جزءا كل جزء فيه 10 مفردات مرتبة.

الجزء الأول: من 1 إلى 10

الجزء الثاني: من 11 إلى 20

الجزء الثالث: من 21 إلى 30

الجزء الرابع: من 31 إلى 41

الجزء التاسع والتسعون: من 981 إلى 990

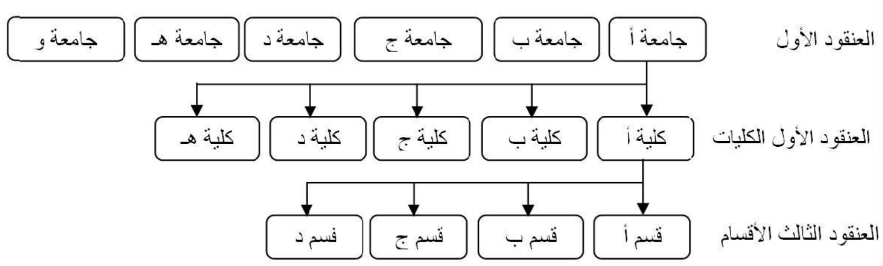
الجزء المئة: من 991 إلى 1000.

فنختار عينتنا التي تتكون من 100 مفردة، حيث نختار من الجزء الأول رقما عشوائيا أي بطريقة المعاينة البسيطة وليكن الرقم 7 ثم نختار بقية الأرقام بانتظام حيث من كل جزء نختار نفس رتبة الرقم الأول وهي 7 فنجد الأرقام من الأجزاء الأخرى هي 17–27–37–47–57––––––987–997.

أما إذا اخترنا عشوائيا من الجزء الأول الرقم 4 مثلا فإن الأرقام الأخرى ستكون 14–24–34–––––984–994.

- **المعاينة الطبقية:** حيث يتم تقسيم مجتمع البحث إلى طبقات ولكل طبقة خصائص مشتركة تختلف عن الأخرى، وتسمح بدرجة عالية من التمثيل وتقليل هامش الخطأ (تقسيم مجتمع بلد ما إلى طبقات حسب الديانات).

- **لمعاينة العنقودية:** السحب بالصدفة من خلال تقسيم المجتمع إلى عناقيد (عند القيام ببحث حول أساتذة جامعات الجزائر فإننا نقسم المجتمع حسب الجامعات وهو العنقود الأول ثم الكليات وهو العنقود الثاني ثم الأقسام وهو العنقود الثالث ولو أردانا أن نضيف عنقودا رابع فسيكون المقاييس) كما في الشكل التالي:



ثم اختيار عينة جزئية من كل قسم تمثل في مجموعها العينة الكلية

8) عرض البيانات:

8-1) جدوليا: في حالة دراسة ظاهرة تحوي متغير وصفي واحد فإننا نعرض البيانات في جدول تكراري حيث:

مثال الجنسيات

- العمود الأول نضع فيه (X_i) المجموعات المكونة للظاهرة أو الحالات.
- العمود الثاني نضع فيه f_i تكرار الظاهرة في كل مجموعة أو حالة (fréquences).
- العمود الثالث وفيه **التكرار النسبي** $f\% = f_i$ تكرار الحالة / مجموع التكرارات n **(حجم العينة)**. ($f\% = \frac{f_i}{n}$)

وفي حالة بيانات أو متغيرات كمية فإننا نضعها في فئات متساوية، ونكون جدول التوزيع التكراري حيث: وقبل تكوين التوزيع التكراري يتم تحديد المدى الذي تتراوح فيه البيانات، تحديد عدد الفئات، تحديد طول الفئة، تحديد حدود الفئات (الحد الأدنى والأعلى لكل فئة)، مراكز الفئات:

- حساب المدى Range الذي تتراوح فيه البيانات: وهو الفرق بين أكبر قيمة وأصغر قيمة للبيانات في العينة.
- تحديد عدد الفئات Classes : توجد عدة قوانين إلا أن أغلب الباحثين يرون أنها يجب أن لا تقل عن 5 فئات وأن لا تزيد عن 15 فئة وأكثر عدد الفئات استعمالا هو 8 فئات .
- تحديد طول الفئة L ويساوي حاصل قسمة المدى على عدد الفئات، range / L = nombre de classes.
- حدود الفئات:

مثال أعمار 70 طالب جامعي

- الحد الأدنى للفئة الأولى هو أقل قيمة للبيانات.
- الحد الأعلى للفئة الأولى هو الحد الأدنى لها زائدا طول الفئة.
- **(وهكذا الحد الأعلى لكل فئة = حدها الأدنى + طول الفئة)**
- الحد الأدنى للفئة الثانية هو الحد الأعلى للفئة الأولى.
- **(الحد الأدنى لكل فئة = الحد الأعلى للفئة التي قبلها)** إلا الفئة الأولى فإن حدها الأدنى هو أقل قيمة في البيانات.

- مراكز الفئات: وهي القيمة التي تقع في منتصف الفئة وتحسب كما يلي:

مركز الفئة = (الحد الأدنى للفئة + حدها الأعلى) / 2

تكوين جدول التوزيع التكراري:

- العمود الأول نضع فيه (X_i) المجموعات المكونة للظاهرة أو الحالات.
- العمود الثاني نضع فيه f تكرار الظاهرة في كل مجموعة.
- العمود الثالث ونضع فيه التكرار المتجمع الصاعد f↑.
- العمود الرابع ونضع فيه التكرار المتجمع النازل f↓.
- العمود الخامس ونضع فيه التكرار النسبي f% = تكرار كل مجموعة f_i / مجموع التكرارات n ($f\% = \frac{f_i}{n}$)

- العمود السادس ونضع فيه التكرار المتجمع الصاعد النسبي f↑% ($f\uparrow\% = \frac{f\uparrow}{n}$).
- العمود السابع ونضع فيه التكرار المتجمع النازل النسبي f↓% ($f\uparrow\% = \frac{f\downarrow}{n}$).

إذا كان التكرار البسيط والتكرار النسبي يعبران بالترتيب عن حجم الحالة ونسبة الحالة في العينة فإننا:

نستعمل التكرارات **المتجمعة <u>الصاعدة</u>** البسيطة أو النسبية لمعرفة عدد أو نسبة المشاهدات **<u>التي تقل</u>** عن قيمة معينة.

نستعمل التكرارات **المتجمعة <u>النازلة</u>** البسيطة أو النسبية لمعرفة عدد أو نسبة المشاهدات **<u>التي تزيد</u>** عن قيمة معينة.

f↓%	f↑%	f%	f↓	f↑	f	X_i

2-8) بيانيا: نحتاج لرسم كل من المدرج التكراري (الأعمدة البيانية التكرارية)، المضلع التكراري والمنحنى البياني للتوزيع التكراري البسيط؟، نحتاج لرسمها إلى:

- لرسم المدرج البياني أو الأعمدة البيانية يحدد طول الخط الأفقي أو العمودي من خلال أصغر وأكبر قيمة:

ـ على الخط الأفقي نكتب أرقام الفئات (1-2...7-8) أو نكتب قيم حدودها (18-22...46-50).

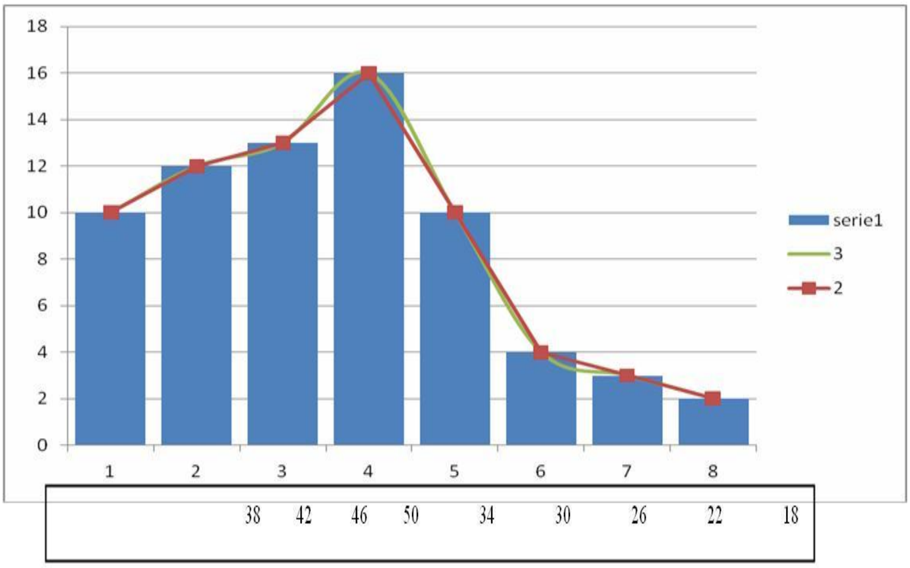
– وعلى الخط العمودي نبحث عن أعلى قيمة للتكرار وحسب تمرين مثال أعمار الطلبة فإن: أكبر قيمة للتكرار هي 16 وأدنى قيمة هي 2 وتقترب من 0، لذا نقسم الخط العمودي إلى أجزاء متساوية بين0 و16 حسبما يفيد الطالب في الرسم (ولا يشترط أن تكون أرقام هذه الأجزاء موجودة في التكرارات).

ونرسم الأعمدة البيانية حيث: عرض العمود = طول الفئة وطول المدرج = تكرار الفئة

- ولرسم المضلع أو المنحنى نحدد كل الثنائيات (مركز الفئة، تكرار الفئة) ونصل بالمضلع أو المنحنى بينها.

الشكل في حالة التوزيع التكراري البسيط المدرج والمضلع والمنحنى.

الشكل (1– 1): التوزيع التكراري البسيط المدرج والمضلع والمنحنى

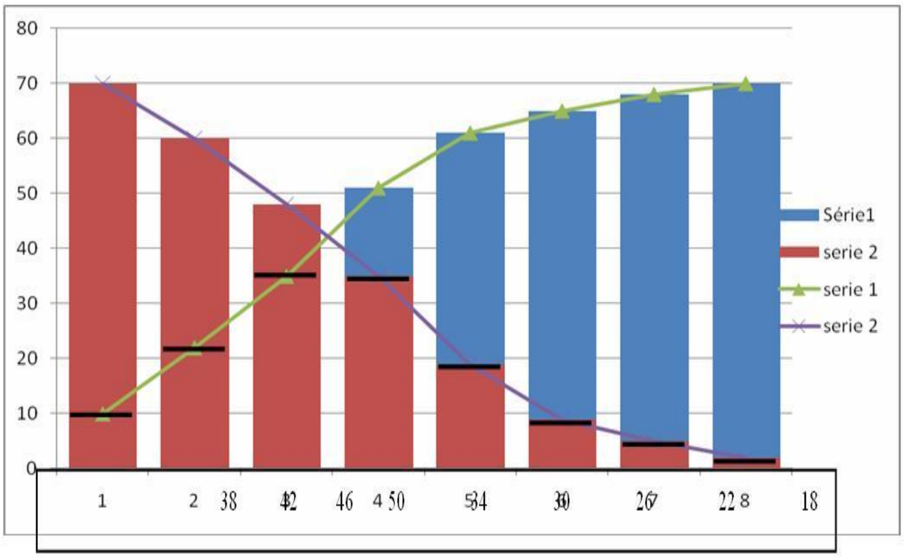


ثم يتم تحديد شكل المنحنى (ملتوي جهة اليمين أي موجب الالتواء ، متماثل، ملتوي جهة اليسار أي سالب الالتواء).

أما في حالة التوزيع التكراري المتجمع الصاعد أو النــازل فيكــون الشــكل كالتالي:

وبنفس الخطوات السابقة للخطين الأفقي والعمودي فقط تتبدل القيمة العليا للخط العمودي فتصبح 70 لأنه تكرار صاعد إلى 70 أو نازل من 70، فيقسم الخط حيث يستحسن أن يبدأ من 0 إذا كانت القيمة الدنيا تقترب من 0.

الشكل (1− 2): التوزيع التكراري للمتجمع الصاعد والنازل



أما لرسم التكرارات المتجمعة النسبية (البسيطة، الصاعدة والنازلة) في عرضها البياني (مدرج، مضلع ومنحنى)، نتبع نفس الخطوات من بداية هذه الصفحة، حيث تبقى قيم الخط الأفقي كما هي وتتبدل القيمة العليا للخط العمودي في كل التوزيعات النسبية لتصبح نفسها قيمة التكرار الأعلى في الجدول (حيث أن القيمة العليا للتكرار النسبي البسيط هي قسمة أكبر تكرار بسيط على مجموع التكرارات، أما التكرارات النسبية الصاعدة والنازلة فإن أعلى قيمة لها هي 1)، وكل قيمة لتكرار متجمع **صاعد** تقابله **قيمة الحد الأدنى** لفئته، وكل قيمة لتكرار متجمع **نازل** تقابله **قيمة الحد الأعلى** لفئته.

أولا – العرض الجدولي للبيانات

تلخص الجداول التكرارية البيانات الكمية الكثيرة في وضعها على صورة جدول منتظم يوضح كيفية توزيع القيم التي حصلنا عليها من الظاهرة المدروسة حيث يدل العمود (السطر) الأول على قيم الظاهرة، ويدل العمود (السطر) الثاني على التكرار المقابل لهذه القيم.

مثال1: إذا كانت لدينا درجات 30 طالب في أحد الاختبارات كما يلي:

4، 8، 7، 6، 4، 11، 12، 4، 6، 8، 4، 6، 9، 6، 7، 6، 8، 4، 11، 9، 8، 11 ، 9، 4، 8، 9، 11، 12، 4، 5.

وفي حالة ما إذا كنت الدرجات كثيرة ومنتشرة داخل مجال واسع (مثلا التنقيط كان من 100) فإنه توجد طريقة أخرى أكثر اختصارا من السابقة يمكن بواسطتها وضع البيانات في جدول يبين ويوضح الخصائص العامة لهذه البيانات، يسمى هذا الجدول بجدول التوزيع التكراري ولتكوين مثل هذا الجدول نقوم بالتالي:

أولا: نحدد المجال (المدى) الذي تنتشر فيه البيانات، وهو الفرق بين أكبر قيمة للبيانات وأصغر قيمة لها، أي أن:

المدى = أكبر قيمة – أصغر قيمة

ثانيا: نقسم المدى إلى فئات متساوية الطول بحيث يكون عددها مناسبا و هناك عدة طرق لحساب عدد الفئات نذكر منها:

1) معادلة ستيرجس Sturgers التي تنص على أن:

$$\text{عدد الفئات} = 1 + 3.322\log\left(\text{عدد البيانات}\right)$$

2) معادلة يول yule التي تنص على: $\text{عدد الفئات} = 2.5\sqrt[4]{\text{عدد البيانات}}$

ثالثا: نحسب طول الفئة وهو يساوي المدى مقسوما على عدد الفئات $\text{طول الفئة} = \frac{\text{المدى}}{\text{عدد الفئات}}$

مما سبق نستخلص الطريقة المرنة في تحديد عدد الفئات وأطوالها والتي لا تعتمد على المعدلات الرياضية بل أن هذه الطريقة مرنة بطبيعتها وهي: طول الفئة × عدد الفئات ≥ المدى

ملاحظـــات:

I. عند تفريغ البيانات فإنه يجب أن تنتمي كل مفردة إلى فئة واحدة فقط.

II. عند كتابة الفئات فإنه:

- يذكر الحد الأدنى والأعلى لكل فئة(فئات مغلقة)إذا كان المتغير متقطع.
- يذكر الحد الأدنى ويحدد الحد الأعلى ضمنيا أو العكس (فئات مغلقة من جهة و مفتوحة من الجهة الاخرى) إذا كان المتغير متصل.
- يفضل استخدام الفئات المتساوية الطول، إلا أنه في بعض الحالات يمكن أن يستخدم الفئات غير المتساوية، من هذه الحالات ما يلي:

– إذا كان الغرض من الدراسة هو الاهتمام ببعض الفئات والتركيز عليها وإهمال باقي الفئات، فيمكن عندها دمج الفئات التي لا تهم الباحث في فئة واحدة.

– إذا كان التكرار لبعض الفئات صغير جدا مقارنة بباقي الفئات، يمكن دمج هذه الفئات معا.

__مثال 2:__ البيانات التالية تمثل إنتاج 60 ورشة من الكراسي خلال يوم.

87	93	21	46	57	77	62	25	31	72
68	29	89	66	62	73	83	81	72	54
57	81	58	62	73	12	73	83	88	96
67	87	97	52	63	17	29	36	71	63
62	92	73	57	65	71	36	54	21	33
42	58	89	46	49	36	56	62	51	91

المطلوب:1) ما هو نوع المتغير؟ ولماذا؟.

2) أي نوع من الفئات تستخدم في مثل هذه الحالة؟

3) تحديد عدد الفئات باستخدام معادلة Sturgers؟

4) تحديد عدد الفئات باستخدام معادلة Yule؟

5) تكوين جدول تكراري من 10 فئات متساوية الطول؟.

مثال 3: إذا كانت بيانات المثال السابق تمثل إنتاج الحليب باللترات في يوما ما بـــ 60 مزرعة أعد الإجابة على الأسئلة المطروحة؟.

العرض الجدولي في حالة البيانات النوعية

يتكون الجدول من عمودين (سطرين) يحتوي العمود (السطر) الأول على رموز كتابية للخاصية المدروسة (صفات) أما الثاني فيحتوي على تكرارات كل رمز كتابي.

مثال4: أخذت عينة من طلبة جامعة الوادي متكونة من 100 طالب، حيـــث كانـــت الخاصـــية المدروسة هي الانتماء إلى كلية من كليات الجامعة فكانت النتائج كما هي مبينة في الجـــدول أدناه

جدول توزيع مجموعة من طلبة الوادي على كليات الجامعة

الكلية	الحقوق	الاقتصاد	الآداب	العلوم	المجموع
عدد الطلبة	35	30	25	10	100

التكـــرارات المتجمعـــة

في بعض الحالات نرغب في معرفة التكرارات أو البيانات التي تزيد قيمتها عن قيمة معينة أو تقل عن قيمة معينة، فمثلا عندما نرغب في معرفـــة عـــدد النـــاجحين (أي الطلبـــة المتحصلين على درجة تساوي أو تزيد 10) فإن هذه المعلومات غير واضـــحة فـــي جـــدول

التوزيع التكراري فنكون لهذا الغرض ما يسمى بالجدول التكـــراري المتجمـــع الصـــاعد أو النازل، وكتعريف فإن:

- التكرار المتجمع الصاعد لأي فئة هو تكرار هذه الفئة مضافا إليه مجموع تكرارات الفئات السابقة.

- التكرار المتجمع النازل لأي فئة هو عبارة عن مجموع التكرارات مطروحا منه تكرارات الفئات السابقة.

التكرار النسبي

هو وضع تكرار كل فئة كنسبة من التكرار الكلي، وهذه الطريقة لها عدة استخدامات وفوائد حيث توضح نسبة تكرار كل فئة إلى التكرار الكلي.

مثال 5: الجدول الآتي يبين توزيع دخول عينة من عمال مؤسسة صناعة الكوابل الكهربائية حسب دخولهم بآلاف الدينارات. المطلوب حساب التكرار المتجمع الصاعد والتكرار المتجمع النازل و التكرار النسبي لهذه البيانات.

فئة الدخل	10-15	15-20	20-25	-3025	30-35	35-40	**المجموع**
عدد العمال	4	6	12	8	6	4	**40**

ثانيا – العرض البياني للبيانات

بالإضافة إلى العرض الجدولي للبيانات هناك طريق أخرى تستخدم لتوضيح وتلخيص البيانات وهي طرق العرض البياني. وهذه الطرق أسهل وأبسط مـــن ســـابقتها، فـــإذا كـــان العرض الجدولي لا يفهمه إلا المختصين فإن طرق العرض البياني يمكـــن أن تجلـــب غيـــر المتخصصين، كما تمكن من القيام بتحليل سريع للظاهرة المدروسة. وتستخدم أنواع مختلفـــة للعرض البياني حسب نوعية المتغير المدروس.

1) المتغير الكمي المتقطع: وهنا يمكن استخدام الأعمـــدة البســـيطة، العـــرض البيـــاني للتكرارات المتجمعة الصاعدة والنازلة.

2) المتغير الكمي المتصل: هنا يمكن استخدام المدرج التكراري، المضلع التكراري، المنحنى البياني للتكرارات المتجمعة الصاعدة والنازلة.

3) المتغير النوعي: وهنا يمكن استخدام الدوائر، الأعمدة المجزأة، الأعمدة المستطيلة.

I – العرض البياني في حالة متغير كمي متقطع:

الأعمدة البسيطة (Diagrammes en bâtons): هي عبارة عن أعمدة بسيطة تتناسب أطوالها مع التكرار المقابل لقيمة العينة للمتغير المدروس

مثال 6: يبين الجدول التالي عدد الأطفال في العائلة لعينة تكون من 125 أسرة، المطلوب عرض هذه البيانات بطريقة الاعمدة البسيطة.

المجموع	8	7	6	5	4	3	2	1	0	عدد الأطفال في كل أسرة
125	10	15	25	20	26	14	10	9	6	عدد الأسر

II -العرض البياني في حالة متغير كمي متصل (مستمر):

إن العروض البيانية للمتغير الإحصائي الكمي المستمر هي أكثر العروض البيانية استعمالا ومن أهمها:

1 – المدرج التكراري Histogramme:

وهو عبارة عن مستطيلات متلاصقة طول كل منها يتناسب مع التكرار المقابل، وقاعدة كل منها تساوي طول الفئة المقابلة، ويمكن أن نميز بين حالتين عند رسم المدرج التكراري:

أ) الحالة الأولى: عندما تكون الفئات متساوية.

مثال7: يبين الجدول التالي توزيع الدرجات التي حصل عليها 50 طالب في مادة الإحصاء.
أرسم المدرج التكراري الذي يمثل توزيع الدرجات؟.

المجموع	16-20	12-16	8-12	4-8	0-4	الفئة
50	4	6	20	12	8	التكرار

ب) الحالة الثانية: عندما تكون الفئات غير متساوية.

إذا كانت هناك فئات التوزيع غير متساوية نقوم بتعديل التكرارات، حتى يكون هناك تناسب بين طول الفئة والتكرار المقابل لها، وقيم تعديل التكرارات باستخدام المعادلة الآتية.

$$\text{التكرار المعدل} = \frac{\text{تكرار الفئة}}{\text{طول الفئة}} \times \text{طول الفئة المختار}$$

مثال8: يبين الجدول التالي توزيع عينة من 100 عامل حسب الأجر اليومي المطلوب تمثيل هذه البيانات باستخدام المدرج التكراري؟.

فئة الأجر	200-250	250-350	350-400	400-550	550-750	750-800	المجموع
عدد العمال	5	15	20	25	30	5	100

2 – **المضلع التكراري** Polygone de fréquence:

هو مجموعة من القطع المستقيمة متصلة ومنكسرة تتحدد بنقاط أحداثياتها مركز الفئة والتكرارات المقابلة.

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

مثال9: ليكن التوزيع التكراري الآتي، أرسم المدرج التكراري والمضلع التكراري؟.

الفئة	2-4	4-6	6-8	8-10	10-12	12-14	المجموع
التكــرار	4	8	16	12	6	2	48

ملاحظة: المساحة التي تقع تحت المضلع التكراري تساوي المساحة التي تقع تحت المدرج التكراري، وحتى نحافظ على المساحة التي تقع أسفل المضلع التكراري، نفترض أن لهذا التوزيع فئات إحداهما في بدايته والأخرى في نهايته تكرار كل منهما يساوي صفر، بحيث ننطلق في رسم المضلع من مركز الفئة الافتراضية الأولى (الفئة ما قبل الأولى)، وننتهي عند مركز الفئة الافتراضية الأخيرة.

3 – منحنى التكرارات المتجمعة الصاعدة والنازلة:

يرسم منحنى التكرار المتجمع الصاعد عن طريق إيصال مجموعة النقاط ذات الإحداثيات التالية: الحدود العليا للفئات والتكرار المتجمع الصاعد المقابل لها، ويرسم منحنى التكرار المتجمع النازل بإيصال مجموعة النقاط التي إحداثياتها: الحدود الدنيا للفئات والتكرار المتجمع النازل مقابل لها.

مثال10: أرسم على نفس المعلم كل من منحنى التكرار المتجمع الصاعد ومنحنى التكرار المتجمع النازل لبيانات التكراري للمثال السابق؟.

III – العرض البياني في حالة متغير نوعي:

1) العرض الدائري: Diagramme circulaire:

ويتمثل في دائرة مقسمة إلى عدة أجزاء كل جزء يقابل زاوية مركزية تتناسب مع التكرارت المقابلة لكل خاصية من الخصائص المدروسة، ولتحقيق ذلك نضيف عمودا إلى جدول المعطيات يحتوي على الزوايا المركزية المقابلة لكل تكرار.

مثال11: يبين الجدول التالي عدد طلبة كلية العلوم الاقتصادية و التجارية و علوم التسيير سنة 2011 مقسمين على أقسام الكلية المختلفة.

القسم	تجارة	إقتصاد	تسيير	سنة اولي ل. م. د	**المجموع**
عدد الطلبة	120	130	250	520	**1020**

المطلوب: عرض البيانات باستخدام القطع الدائرية؟.

ملاحظة: تحسب الزوايا المركزية المقابلة لكل تكرار بالطريقة التالية:

$$\text{الزاوية المركزية} = \frac{\text{تكرار الخاصية}}{\text{مجموع التكرارات}} \times 360$$

2) العمود المجزأ Diagrammes en Barres: وهو عبارة عن مستطيل مقسم إلى عدة أجزاء كل جزء يقابل تكرار معين للخاصية المدروسة.

<u>مثال12</u>: أعرض بيانات المثال السابق باستخدام العمود المجزأ.

3) الأعمدة المستطيلة: وهو عبارة عن مجموعة من الأعمدة المتجاوزة ذات القواعد المتساوية إلا أن ارتفاعها تتناسب مع تكرار كل خاصية، كما أن هذه الأعمدة تكون متباعدة بمسافات متساوية.

<u>مثال13</u>: أعرض بيانات المثال السابق باستخدام المستطيلة

سلسلة تمارين حول: عرض البيانات

<u>التمرين الأول:</u> حدد المجتمع الإحصائي، الوحدة الإحصائية، المتغير الإحصائي ونوعه من واقع العبارات التالية:

- مدة حياة المصابيح الكهربائية المنتجة في مصنع.
- الأجور الشهرية لعمال مؤسسة ما.
- أوزران طلبة جامعة الوادي.
- الرياضة الممارسة من طرف طلبة كلية ع الاقتصادية والتجارية و ع التسيير بجامعة الوادي.
- عدد الغرف المملوكة من طرف كل عائلة في حي معين.
- تصنيف عمال مصنع حسب المؤهل.
- عدد السيارات في دولة حسب الصنف.
- تصنيف الأحزاب السياسية حسب عدد الأصوات المكتسبة في الانتخابات.

<u>التمرين 02</u>: في محاولة لدراسة حالة التعليم كانت نتائج البحث الميداني الذي خص 30 شخص كما يلي:

ماجيستير	**بكالوريا**	**دكتوراة**	**ماجيستير**	**دكتوراة**	**ليسانس**	**ليسانس**	**بكالوريا**	**بكالوريا**	**ليسانس**
دكتوراة	**بكالوريا**	**بكالوريا**	**بكالوريا**	**بكالوريا**	**ماجيستير**	**ليسانس**	**تقني**	**ماجيستير**	**ليسانس**
ليسانس	**تقني**	**تقني**	**تقني**	**ماجيستير**	**دكتوراة**	**ماجيستير**	**ليسانس**	**تقني**	**تقني**

1) حدد المتغير الإحصائي ونوعه؟
2) لخص هذه البيانات في جدول إحصائي ثم أوجد نسب التمثيل لكل صفة في العينة المدروسة؟
3) أنشئ التمثيل البياني المناسب للجدول؟

<u>التمرين 03</u>: البيانات التالية تبين عدد الغيابات التي سجلها عمال مؤسسة ما خلال الثلاثي الأول من السنة:

4	3	5	2	3	2	2	2	5	3
5	5	4	0	0	5	1	2	5	0
1	1	5	3	0	2	3	2	1	4

1) حدد المجتمع الإحصائي والمتغير الإحصائي ونوعه؟
2) لخص هذه البيانات في جدول إحصائي؟ ثم أنشئ التمثيل البياني المناسب للجدول؟
3) أحسب عدد ونسبة العمال الذين لديهم أقل من 3 غيابات؟

<u>التمرين 04</u>: البيانات التالية تمثل مبيعات كشك معين من جريدة الخبر خلال 50 يوم.

36	45	31	28	41	32	29	26	48	52
30	28	33	27	40	31	30	40	35	45
32	37	36	31	35	33	38	39	41	30
38	42	35	20	36	38	39	37	45	23
36	46	30	35	32	51	43	37	22	34

1) شكل جدول توزيع تكراري المناسب؟ ثم أنشئ التمثيل البياني لهذا الجدول؟

<u>التمرين 05</u>: لدينا البيانات التالية والتي تمثل الأجر اليومي لأربعين عاملا بالدينار:

370	250	140	250	490	300	490	170	180	370
450	560	150	600	510	140	220	270	190	450
230	460	640	370	450	260	340	170	320	150
600	160	270	300	700	100	490	190	160	690

1) لخص هذه البيانات في جدول إحصائي؟ ثم أنشئ التمثيل البياني المناسب للجدول؟
2) أحسب عدد و نسبة العمال الذين تساوي أو تزيد أجورهم عن 200 دينار؟

<u>التمرين 06</u>: الجدول التالي يمثل توزيع الأجور الأسبوعية لعمال مؤسسة ما:

الفئات	التكرارات	التكرارات النسبية
0−10	n_1	0.08

10—20	08	–
20—30	12	–
30—50	14	–
50—60	n_5	–
المجموع	50	1

1) أكمل الجدول ثم أحسب نسبة العمال الذين تساوي أو تزيد أجورهم عن 10 ولكنها تقل عن 50 ؟

<u>التمرين 07</u>: الجدول التالي يمثل توزيع عمال مؤسسة ما حسب التخصص:

الشهادة	التكرارات	التكرارات النسبية
تقني	55	0.44
تقني سامي	n_2	0.24
ليسانس	25	0.20
مهندس	n_4	f_4
المجموع	N	1

1) برهن على أن مجموع التكرارات النسبية يساوي 1.

2) استنتج كل من N ، n_4 ، n_2 و f_4 ؟

<u>التمرين 08</u>: فيما يلي بيانات عن المستوى التعليمي لعينة من 50 فرد.

متوسط	يقرأ ويكتب	ثانوي	متوسط	ثانوي	أعلى من جامعي	متوسط	ابتدائي
يقرا ويكتب	متوسط	ثانوي	ثانوي	متوسط	ثانوي	ابتدائي	متوسط
ابتدائي	ثانوي	يقرا ويكتب	جامعي	ثانوي	ابتدائي	يقرا ويكتب	ثانوي
متوسط	ابتدائي	متوسط	ثانوي	ابتدائي	متوسط	جامعي	متوسط
ثانوي	متوسط	ابتدائي	ثانوي	يقرا ويكتب	ابتدائي	ثانوي	ابتدائي
جامعي	ثانوي	جامعي	ابتدائي	جامعي	أعلى من جامعي	ثانوي	ثانوي
متوسط	يقرا ويكتب						

والمطلوب: **1**- اعرض البيانات في شكل جدول تكراري.

2- كون التوزيع التكراري النسبي، ثم علق على النتائج.

التمرين 09:

فيما يلي بيانات درجات 70 طالب في الاختبار النهائي لمقرر مادة الإحصاء التطبيقي.

56	65	70	65	55	60	66	70	75	56
60	70	61	67	61	71	67	62	71	66
68	72	57	68	72	69	57	71	69	75
72	62	67	73	58	63	66	73	63	65
58	73	74	76	74	80	81	60	74	58

72	94	78	91	85	77	83	77	82	76
62	78	88	64	87	55	79	57	64	79

والمطلوب:

1- كون التوزيع التكراري لدرجات الطلاب.

2- كون التوزيع التكراري النسبي.

3- ما هو نسبة الطلاب الحاصلين على درجة ما بين 70إلى أقل من 80؟

4- ما هو نسبة الطلاب الحاصلين على درجة أقل من 70 درجة؟

5- ما هو نسبة الطلاب الحاصلين على درجة 80 أو أكثر ؟

التمرين 10:

فيما يلي التوزيع التكراري لأوزان عينة من الدواجن بالجرام، حجمها 100 اختيرت من أحد المزارع بعد 45 يوم.

الوزن	600-	620-	640-	660-	680-	700-720	المجموع
عدد الدجاج	10	15	20	25	20	10	100

والمطلوب:

1- ما هو طول الفئة؟

2- ارسم المدرج التكراري.

3- ارسم المدرج التكراري النسبي، ثم علق على الرسم.

التمرين 11:

الجدول التكراري التالي يبين توزيع 40 بقرة في مزرعة حسب كمية الألبان التي تنتجها البقرة في اليوم باللتر.

كمية الألبان	18-	22-	26-	30-	34-38	المجموع
عدد الأبقار	4	9	15	8	4	40

والمطلوب:

1– كون جدول التوزيع التكراري المتجمع الصاعد.

2– كون جدول التوزيع التكراري المتجمع الصاعد النسبي.

3– ارسم المنحنى التكراري المتجمع الصاعد النسبي.

4– من المنحنى المتجمع أوجد الآتي:

- نسبة الأبقار التي يقل إنتاجها عن 28 لتر.
- كمية الإنتاج التي يقل عنها %25 من الأبقار.
- كمية الإنتاج التي يقل عنها %50 من الإنتاج.

التمرين 12:

الجدول التكراري التالي يبين توزيع عينة حجمها 500 أسرة حسب المنطقة التي تنتمي إليها.

المنطقة	الرياض	الشرقية	القصيم	الغربية	المجموع
عدد الأسر	150	130	50	170	500

مثل البيانات أعلاه في شكل دائرة بيانية.

عند التمعن في الظواهر التي حولنا والقيم التي تأخذها نلاحظ أن غالبية هذه القيم نقترب من بعضها البعض وتتجمع حول قيمة معينة غير منظورة، فذكاء أو طول أو وزن مجموعة من الأشخاص مثلا وتتجمع حول قيمة معينة متوسطة، والقليل من الأشخاص لهم ذكاء أو طول أو وزن يبتعد كثيرا عن هذه القيمة من ناحية الصغر أو الكبر. سميت هذه الظاهرة بظاهرة النزعة المركزية والتي لها عدد من المتوسطات للتعبير عنها تختلف باختلاف الغرض الذي تستخدم فيه، وطبيعة البيانات المحسوبة منها، أهم هذه المتوسطات: المتوسط الحسابي، المتوسط الهندسي، المتوسط التوافقي، الوسيط، المنوال والتي سنتطرق لها في الصفحات القادمة.

وميزة هذه المتوسطات كقيم عددية وحيدة توفر لنا فكرة عامة عن البيانات، وتصف الظاهرة المدروسة مثل الجداول الإحصائية، إلا أنها أكثر اختصارا وأكثر فائدة، حيث تمكننا من المقارنة بين مجموعة من القيم ومجموعة أخرى، أو بين ظاهرة وأخرى.

أولا – المتوسط الحسابي Moyenne arithmétique

يعتبر المتوسط الحسابي من أسهل وأكثر متوسطات النزعة المركزية استخداما في الإحصاء، هو عبارة عن مجموع القيم مقسوما على عددها.

فإذا كانت لدينا القيم $x_1,x_2,x_3,.........x_n$ فإن متوسطها الحسابي يساوي

$$\overline{X} = \frac{x_1 + x_2 + x_3 + ...+ x_n}{n} = \frac{\sum_{i=1}^{n} Xi}{n}$$

حيث: $\overline{X}$ = المتوسط الحسابي

X_i = تمثل قيم الظاهرة

n = تمثل عدد البيانات

<u>**مثال1:**</u>

إذا كانت الدرجات التي تحصل عليها الطالب في خمس مواد هي: 8، 10، 13، 14، 15.

أحسب متوسط درجات هذا الطالب؟.

الحل:

المتوسط = مجموع الدرجات

$$\overline{X} = \frac{\sum Xi}{n} = \frac{8+10+13+14+15}{5} = \frac{60}{5} = 12$$

المتوسط الحسابي في حالة بيانات مكررة

إذا كانت لدينا القيم $x_1, x_2, x_3, \ldots\ldots\ldots\ldots\ldots x_4$

ولها تكرارات $n_1, n_2, n_3, \ldots\ldots\ldots\ldots\ldots n_4$

فإن المتوسط الحساب لها يعطي بالعلاقة

$$\overline{X} = \frac{n_1 x_1 + n_2 x_2 + n_3 x_3 + \ldots\ldots\ldots\ldots n_n x_n}{n_1 + n_2 + n_3 + n_4 + \ldots\ldots\ldots\ldots n_n} = \frac{\sum_{i=1}^{n} n_i x_i}{\sum n_i}$$

أي أن المتوسط الحسابي لبيانات متكررة يساوي مجموع حاصل ضرب كل قيمة في تكرارها على مجموع التكرارات

وحيث: X_i = تمثل قيم الظاهرة.

n_i = تكرار مل قيمة.

Σn_i = مجموع التكرارات.

مثال2:

في امتحان فجائي في مادة الإحصاء الوصفي تحصل طلبة فوج معين على الدرجات المبينة في الجدول التالي:

الدرجة	3	4	5	6	8	9
عدد الطلبة	4	3	6	5	4	2

المطلوب: حساب متوسط الدرجات التي تحصل عليها طلبة هذا الفوج؟.

الحل:

لحساب هذا المتوسط فإننا نقوم أولا بضرب كل قيمة في تكرارها ثم نطبق العلاقة التي تحسب المتوسط الحسابي لبيانات متكررة، وسنقوم بذلك من خلال الجدول التالي:

الدرجة x_i	عدد الطلبة n_i	$n_i\,x_i$
3	4	12
4	3	12
5	6	30
6	5	30
8	4	32
9	2	18
المجموع	**24**	**134**

$$\overline{X} = \frac{\sum n_i x_i}{\sum n_i} = \frac{134}{24} = 5{,}58$$

أي أن متوسط درجات طلبة هذا الفوج الحسابي في الامتحان الفجائي يساوي 5.58.

المتوسط الحسابي في حالة توزيع تكراري:

تعتمد طريقة حساب المتوسط الحسابي لبيانات مبوبة على مراكز الفئات التي يفترق أنها تمثل الفئات التي أخذت منها، ويكون المتوسط الحسابي في هذه الحالة يساوي إلى مجموع حاصل ضرب مراكز الفئات في تكرارها على مجموع التكرارات

$$\overline{X} = \frac{\sum_{i=1}^{n} n_i x_i}{\sum_{i=1}^{n}}$$

حيث: x_i = تمثل مراكز الفئات.

n_i = تكرار الفئات.

Σn_i = مجموع التكرارات.

مثال3:

في دراسة إحصائية حول مادة الحليب بالمزارع الموجودة على مستوى ولاية ســوق أهراس توصلنا إلى إعداد الجدول التالي:

الإنتاج باللترات	240−200	280−240	320−280	360−320
عدد المزارع	5	6	8	4

المطلوب: إيجاد متوسط إنتاج الحليب بهذه المزارع

الحل:

الإنتاج باللترات	**عدد المزارع n_i**	**مراكز الفئات x_i**	**$n_i\ x_i$**
240−200	5	220	1100
280−240	6	260	1560
320−280	8	300	2400
360−320	4	240	1360
400−360	2	380	760
المجموع	**25**	/	**7180**

$$\overline{X} = \frac{\sum n_i x_i}{\sum n_i} = \frac{7180}{25} = 287.2$$

أي أن متوسط إنتاج الحليب بهذه المزارع هو أكثر بقليل من 287 لتر للمزرعة.

حساب المتوسط الحسابي بالطريقة المختصرة:

عندما تكون البيانات الموجودة في الجداول المعطاة كبيرة القيم فإن استخدام الطريقــة السابقة يصبح صعب، كما يزداد احتمال الوقوع في الأخطاء بذلك فإنه في مثل هذه الحــالات يفضل استخدام طريقة مختصرة الهدف منها تبسيط العمليات الحسابية الطويلة حتــى يســهل التعامل معها.

فإذا قمنا مثلا بطريقة قيمة ثابتة (a) من جميع القيم (جميع مراكز الفئات) فإن المتوسط الحساب يصبح

$$\boxed{\overline{X} = a + \frac{\sum_{i=1}^{n}(x_1 - a)}{n}}$$

إذا كانت البيانات مفردة أو إذا كانت البيانات مبوبة في جداول توزيع التكراري

$$\boxed{\overline{X} = a + \frac{\sum_{i=1}^{n} n_1(x_i - a)}{\sum n_i}}$$

مثال4:

أحسب متوسط إن تاج مادة الحليب في المثال السابق بإتباع الطريقة المختصرة؟

الحل:

نختار وسط فرضي a=300 وهو مركز الفئة الوسطى ونتبع الخطوات التالية:

1. توجد قيم جديدة = y_i والتي تساوي مراكز الفئات ناقص الوسط الفرض.
2. نضرب هذه القيم الجديدة في تكرار الفئات n_i.
3. نحسب $\overline{Y}$ وهو المتوسط الحسابي لهذه القيم الجديدة.
4. نحسب المتوسط الحسابي للقيم الجديدة والذي يساوي $\overline{X} = \overline{Y} + a$

الإنتاج باللترات	عدد المزارع n_i	x_i	y_i	$n_i\, x_i$
240−200	5	220	−80	−400
280−240	6	260	−40	−240
320−280	8	300	0	0
360−320	4	240	40	160
400−360	2	380	80	160
المجموع	25			−320

$$\overline{Y}=\frac{\sum_{i=1}^{n} n_1 y_i}{\sum_{i=1}^{n} n_1}=\frac{-320}{25}=-12{,}8$$

$\overline{X}$ =300 − 12.8= لتر 287.2

المتوسط الحسابي المرجح (الموزون) Moyenne Arithmétique pondérée

في بعض الأحيان فإن القيم المراد حساب المتوسط الحسابي لها لا تكون لها نفس الأهمية بل أهميات نسبية مختلفة تختلف باختلاف عامل الترجيح الخاص بها. في مثل هذه الحالات فإن المتوسط الحسابي البسيط يمكن الاعتماد عليه في إيجاد المتوسط الصحيح والمنطقي، بل يتطلب الأمر استخدام صيغة أخرى تسمى بصيغة المتوسط الحسابي المرجح:

$$\text{المتوسط الحسابي المربح} = \frac{\text{مجموع حاصل ضرب مل قيمة في معاملها}}{\text{مجموع المعاملات}} = \overline{X}=\frac{\sum_{i=1}^{n} n_i x_i}{\sum_{i=1}^{n} x_i}$$

حيث: x_i = تمثل القيم.

n_i = تمثل المعاملات.

ويستخدم المتوسط الحسابي المرجع كذلك لإيجاد المتوسط الحسابي لمجموع البيانات أو أكثر في حالة دمجهم معا في مجموعة واحدة وبالتالي فإن متوسط الحسابي المرجح لمجموعتين من البيانات y,z يساوي:

$$\overline{X}=\frac{\sum_{i=1}^{n} z_i+\sum_{i=1}^{n} y_i}{n+Z}$$

وبما أن

$$\sum_{i=1}^{n} z_i = n\overline{Z}$$

$$\sum_{i=1}^{n} y_i = n\overline{Y}$$

فإن

$$\overline{X}=\frac{n\overline{Z}+Z\overline{Y}}{n+Z}$$

مثال5:

الجدول التالي يبين عدد العمال ومتوسط الأجر للعامل الواحد في الواحدات المختلفة التي تشكل الشركة الوطنية لإنتاج الأنابيب البلاستيكية. المطلوب حساب متوسط الأجور التي توزعها هذه الشركة؟.

الفرع	وحدة الشمال	وحدة الشرق	وحدة الجنوب
عدد العمال	130	110	80
متوسط الأجور	13000	14500	18500

الحل:

إذا اعتبرنا أن متوسط الأجر في شركة هو عبارة عن مجموع متوسط الأجر في الوحدات الثلاث مقسوما على ثلاثة فإن الإجابة تكون خاطئة فالإجابة الصحيحة هي تلك التي يمكن الحصول عليها من خلال علاقة المتوسط الحسابي المرجح.

$$\overline{X} = \frac{n_1\overline{x}_1 + n_2\overline{x}_2 + n_3\overline{x}_3 ....+ n_1\overline{x}_1}{n_1 + n_2 + n_3 + ...+ n_4}$$

$$\overline{X} = \frac{(130x13000) + (110x14500) + (80x18500)}{130+110+80} = \frac{1690000+1595000+1480000}{320}$$

$$\overline{X} = \frac{4765000}{320} = 14890{,}62 =$$ متوسط أجر عمال الشركة

خواص المتوسط الحسابي:

1 – يعتبر المتوسط الحسابي أبسط مقاييس المركزية حسابا وأكثرها استخداما.

2 – يأخذ المتوسط الحسابي بعين الاعتبار جميع قيم الظاهرة المدروسة.

3 – مجموع انحرافات القيم عن متوسطها الحسابي يساوي صفر = $\sum_{i=1}^{n}(x_i - \overline{X}) = 0$

وللتأكد من ذلك نقوم بتفكيك الطرف الأيمن:

$$\sum_{i=1}^{n}(x_i - \overline{X})$$
$$= \sum X_i - \sum X$$
$$= \sum X_i - n\overline{X}$$

(لأن :

$(n\overline{X} = ...\overline{X} + \overline{X} + \overline{X}$

نضرب الحد الأول في n ونقسمه على n

$$= \frac{n\sum X_i}{n} - n\overline{X}$$
$$= n\overline{X} - n\overline{X} = 0$$

4 – يتأثر المتوسط الحسابي بالقيم المتطرفة أو الشاذة.

مثال6:

إذا تحصل طالب على الدرجات التالية في خمس مواد: 20.9.8.7.6

فإن المتوسط درجاته $= \frac{20+9+8+7+6}{5} = \frac{50}{5} = 10$

أي أن الطالب ناجح، وفي الواقع فإن الطالب ناجح في مادة واحدة وراسب في أربع مواد، وظهور الطالب ناجح رغم رسوبه في أغلب المواد، راجع إلى أن المتوسط الحسابي قد تأثر بالنتيجة الأخيرة والتي تعتبر متطرفة في الكبر.

5 – لا يمكن إيجاد المتوسط الحسابي من الجداول التكرارية المفتوحة من البداية أو النهاية.

ثانيا: المتوسط الهندسي: Moyenne géométrique

في الحالات التي تكون فيها قيم الظاهرة المدروسة عبارة عن نسب أو معدلات، أي في الحالات التي نرغب فيها بدراسة معدل تغيرات ظاهرة ما فإن المتوسط الحسابي لن يصف هذه الظاهرة الوصف السليم، ولن يعطي أي فكرة صحيحة عن مثل هذه الظاهرة ولهذا دعت الضرورة إلى إيجاد متوسط آخر يصلح لوصف مثل هذه الظواهر سمي هذا المتوسط بالمتوسط الهندسي.

والمتوسط الهندسي واسع الاستخدام في الحياة الاقتصادية (كما سنرى لاحقا) لأن التركيز يكون غالبا منصبا على إيجاد متوسط نسب التغير لبعض الظواهر مثل: تطور الدخول، زيادة الأجور، والنمو السكاني ... إلخ.

تعريف:

إذا كانت لدينا القيم $x_1, x_2, x_3, x_4, ...x_n$ فإن متوسطها الهندسي يساوي الجــذر النــوني لحاصل ضرب هذه القيم في بعضها البعض.

$$G = \sqrt[n]{x_1.x_2x_3.....x_n}$$

وإذا كانت هذه القيم متكررة فإن:

$$G = \sqrt[N]{x_1^{n_1}.x_2^{n_2}.x_3^{n_3}.....x_3^{n_4}}$$

حيث $\Sigma ni = N$ (مجموع التكرارات)

نفس الشيء إذا كانت هذه القيم مبوبة في جداول توزيع تكراري فإن:

$$G = \sqrt[N]{x_1^{n_1}.x_2^{n_2}.x_3^{n_3}.....x_3^{n_4}}$$

حيث x_i = مراكز الفئات.

$\Sigma ni = N$ (مجموع التكرارات).

<u>مثال7:</u>

أوجد المتوسط الهندسي للأعداد: 2، 4، 5، 6.

<u>الحل:</u> $G = \sqrt[4]{2\times4\times5\times6} = \sqrt[4]{240} = 3{,}94$

في حالة كون البيانات كبيرة فإنه يفضل استخدام طريقة اللوغـــاريتم ويكـــون ذلـــك كالتالي:

في حالة بيانات مفردة: $G = \sqrt[n]{x_1.x_2.x_3.x_4...x_n}$

$$LogG = \frac{1}{N}\left[\log x_1 + \log x_2 + \log x_3 + ...+ \log x_4\right]$$

$$\mathbf{logG = \frac{1}{N}\sum_{i=1}^{n} logx_1}$$

وفي حالة البيانات المتكررة أو المبوبة في جداول توزيع تكرار فإن:

$$G = \sqrt[N]{x_1^{n_1}.x_2^{n_2}.x_3^{n_3}..........x_n^{n_n}}$$

$$LogG = \frac{1}{N}\left[n_1 \log x_1 + n_2 \log x_2 + n_3 \log x_3 + ...+ n_n \log x_4\right]$$

$$\mathbf{logG = \frac{1}{N}\sum_{i=1}^{n} n_i logx_1}$$

حيث: x_i تمثل القيمة أو مركز الفئة.

n_i تمثل التكرار.

مثال8:

أوجد المتوسط الهندسي للبيانات المبينة في الجدول أدناه باستخدام طريقة اللوغاريتم

القيم	2	4	5	6
التكرار	2	3	2	4

$$LogG = \frac{1}{N}\sum_{i=1}^{n} Ni \log X_i$$

$$= \frac{1}{11}\left[2\log 2 + 3\log 4 + 2\log 5 + 4\log 6\right]$$

$$= \frac{1}{11}\left[(2 \times 0,0301) + (3 \times 0,601) + (2 \times 0,699) + (4 \times 0,778)\right]$$

$$= \frac{1}{11}\left[0,602 + 1,806 + 1,398 + 3,112\right]$$

$$= \frac{1}{11} 6,918$$

$$\mathrm{LogG} = 0,629$$

ومنه المتوسط الهندسي = 4.24 = G

استخدام المتوسط الهندسي في الحياة الاقتصادي:

لتباين كيفية استخدام المتوسط الهندسي في الحياة الاقتصادية نأخذ المثال التالي:

الجدول التالي يبين تطور إنتاج مؤسسة ما خلال الفترة 2000–2003 .

السنة	2000	2001	2002	2003
الإنتاج	1000	1250	1875	2625

المطلوب :

5. إيجاد نسبة زيادة الإنتاج من سنة إلى أخرى؟
6. إيجاد متوسط نسبة الزيادة خلال الفترة؟

الحل:

السنة	200	2001	2002	2003
الإنتاج	1000	1250	1875	2625
نسبة الزيادة	–	0.25	0.50	0.40

إذا رمزنا لمتوسط نسبة الزيادة خلال الفترة بالرمز (t) و لنسبة الزيادة خلال السنوات الأولى والثانية والثالثة بالرموز (t_1)، (t_2) و (t_3) فإنه حتى يكون (t) ممثلا فعلا لمتوسط نسبة الزيادة خلال الفترة لابد من تحقق الشرط التالي: $(1+t)\,(1+t)\,(1+t) = (1+t_3)\,(1+t_2)\,(1+t_1)$.

ومنه $(1+t_3)\,(1+t_2)\,(1+t_1) = (1+t)^3$.

ونأخذ الجذر التكعيبي لطرفي المعادلة نحصل على

$$1+t=\sqrt[3]{(1+t_1)(1+t_2)(1+t_3)}$$

ومنه

$$t=\sqrt[3]{(1+t_1)(1+t_2)(1+t_3)}-1$$

وبصفة عامة إذا لدينا نسبة الزيادة لـــ (n) فترة متوسط نسبة هذه الزيـــادة (متوسط النسب) يعطي بالعلاقة:

$$\mathbf{t=\sqrt[N]{(1+t_1)(1+t_2)(1+t_3)\ldots\ldots\ldots(1+t_3)}-1}$$

وبتطبيق هذه العلاقة على بيانات المثال السابق فإن:

$$t=\sqrt[3]{(1{,}25)(1{,}50)(1{,}40)}-1$$

t = 0,3795 = سنويا = 37,95%

مثال9:

إذا كان عدد سكان مدينة ما في سنة 1995 يساوي 500000 نسمة وبلغ في سنة 2005، 609497 نسمة أوجد متوسط نسبة زيادة السكان خلال هذه الفترة؟.

الحل:

– نسبة زيادة السكان خلال الفترة = [عدد السكان في سنة 2005– عدد السكان في سنة 1995]/عدد السكان في 1995 = (609497 – 500000)/ 500000 = 21.19%

ومنه يمكن إيجاد متوسط نسبة الزيادة خلال الفترة بتطبيق العلاقة السابقة

$$t = \sqrt[10]{(1,2119)} - 1$$
$$t = 0,02 = 2\%$$

خواص المتوسط الهندسي:

من أهم خواص المتوسط الهندسي ما يلي:

1. يدخل في حساب جميع القيم ولكنه أقل تأثر بالقيم المتطرفة من المتوسط الحسابي.
2. لا يمكن حساب من الجداول التكرارية المفتوحة من البداية أو النهاية.
3. لا يمكن حسابه في حالة وجود قيمة سالبة أو معدومة.
4. يستخدم بأكثر واقعية عند وصف الظواهر النسبية.
5. قيمة المتوسط الهندسي لأي ظاهرة أصغر دائما من قيمة المتوسط الحسابي $G < \overline{X}$

ثالثا: المتوسط التوافقي: Moyenne Harmonique

المتوسط التوافقي هو من المقاييس الخاصة التي تستخدم لتحديد معدلات السرعة ومتوسط الأسعار، ومتوسط الكثافة السكانية. وكمثال على ذلك نقول أن سائق قطع المسافة الفاصلة بين مدينتين على أربع مراحل متساوية، المسافة المقطوعة في كل منها 100 كم.

فإذا قطع المرحلة الأولى بسرعة 100كم/ساعة والمرحلة الثانية بسرعة 120كم/ساعة والمرحلة الثالثة بسرعة 150كم/ساعة والمرحلة الرابعة بسرعة 80كم/ساعة، أوجد متوسط سرعة هذا السائق على طول المرحلة؟.

الحل:

إذا استخدمنا المتوسط الحسابي لتحديد متوسط السرعة لهذا السائق على طول المسافة فإننا سنجده يساوي: $\overline{X} = \frac{100 + 120 + 150 + 80}{4} =$ 112.5 كم/ساعة

وهذه النتيجة غير صحيحة ومضللة بدليل أن:

الزمن الذي إستغرقه السائق في المرحلة الأولى = 1 ساعة.

الزمن الذي إستغرقه السائق في المرحلة الثانية = 6/5 ساعة.

الزمن الذي إستغرقه السائق في المرحلة الثالثة = 3/2 ساعة.

الزمن الذي إستغرقه السائق في المرحلة الرابعة = 4/5 ساعة.

ومنه الزمن الذي استغرقه السائق على طول المسافة = $1 + \frac{5}{6} + \frac{2}{3} + \frac{5}{4} = \frac{45}{12}$ ساعة.

فإذا كان متوسط السرعة المتوصل إليه صحيحا فإن هذا السائق سيقطع خلال المدة $\frac{45}{12}$ ساعة المسافة 422 كم (112.5 × $\frac{45}{12}$ ساعة) وهذا غير صحيح لأن المسافة الكلية هي 400 كم فقط.

الآن إذا قسمنا المسافة الكلية على المدة الزمنية الفعلية فإننا سنحصل على متوسط السرعة الصحيح وهو 106.67 كم/ساعة = $\frac{400}{\frac{45}{12}} = \frac{4800}{45}$

وهذا المتوسط كما سنرى لاحقا يمكن إيجاده بالمتوسط التوافقي.

تعريف: المتوسط التوافقي لمجموعة من القيم هو مقلوب المتوسط الحسابي لمقاليب هذه القيم.

فإذا كانت لدينا القيم: $x_1, x_2, x_3, x_4 \ldots\ldots\ldots x_n$.

فإن مقاليب هذه القيم هو $\frac{1}{x_1}, \frac{1}{x_2}, \frac{1}{x_3}, \frac{1}{x_4}, \ldots\ldots\ldots \frac{1}{x_n}$

والمتوسط الحسابي لمقاليب هذه القيم هو $\frac{\frac{1}{x_1} + \frac{1}{x_2} + \frac{1}{x_3} + \ldots\ldots \frac{1}{x_n}}{n}$

ومقلـــوب المتوسـط الحســابي لمقاليـــب هـــذه القـــيم هـــو المتوســـط التـــوافقي.

$$H = \frac{n}{\frac{1}{x_1}+\frac{1}{x_2}+\frac{1}{x_3}+........\frac{1}{x_n}}$$

وباختصار $H = \frac{n}{\sum_{i=1}^{n}\frac{1}{x_i}}$

أما إذا كانت البيانات متكررة أو مبوبة في جداول توزيع تكراري فإن: $H = \frac{\sum_{i=1}^{n} N_i}{\sum_{i=1}^{n}\frac{n_i}{x_i}}$

حيث: n_i تمثل التكرار.

x_i تمثل القيم أو مراكز الفئات.

فإذا طبقنا علاقة المتوسط التوافقي على بيانات المثال السابق فإننا سنجد

متوسط السرعة

$$H = \frac{400}{\frac{100}{100}+\frac{100}{120}+\frac{100}{150}+\frac{100}{80}} = \frac{400}{\frac{1200+1000+800+1500}{1200}}$$

$H = \frac{48000}{4500} =$ 106.67 كم/ساعة

فإذا ضربنا هذه السرعة في زمن المرحلة $\frac{45}{12}$ ساعة فإننا نحصل على 400 كم هي المسافة المقطوعة فعلا.

<u>مثال10:</u>

أحسب المتوسطات الثلاث (الحسابي والهندسي والتوافقي) للبيانات التالية: 2، 4، 6، 8.

<u>الحل:</u>

7. المتوسط الحسابي $\overline{X} = \frac{2+4+6+8}{4} = \frac{20}{4} = 5$
8. المتوسط الهندسي $G = \sqrt[4]{2\times4\times6\times8} = \sqrt[4]{384} = 4,42$

9. المتوسط التوافقي $H = \dfrac{4}{\frac{1}{2}+\frac{1}{4}+\frac{1}{6}+\frac{1}{8}} = \dfrac{4}{\frac{24+12+8+6}{48}} = \dfrac{192}{50} = 3{,}84$

أي أن $\overline{X} > G > H$ وهذه العلاقة صحيحة في كل الحالات.

مثال11:

اشترى شخص من نفس السوق الكميات التالية من مادة البطاطس:

4كغ بقيمة 100 دينار ثم 5 كغ بقيمة 100 دينار أيضا ثم 8 كغ بقيمة 120 دينار.

المطلوب: إيجاد متوسط سعر البطاطس المشتراة؟.

الحل:

سعر الكيلو غرام في الحالة الأولى $= \dfrac{100}{4} =$ 25 دينار.

سعر الكيلو غرام في الحالة الثانية $= \dfrac{100}{5} =$ 20 دينار.

سعر الكيلو غرام في الحالة الثالثة $= \dfrac{120}{8} =$ 15 دينار.

و متوسط السعر للبطاطس المشتراة هو:

$$H = \frac{100+100+120}{\frac{100}{25}+\frac{100}{20}+\frac{120}{15}} = \frac{320}{\frac{1200+1500+2400}{300}} = \frac{96000}{1200+1500+2400}$$

$H = \dfrac{96000}{5100} =$ 18,8235 ≈ 18,82 DA دينار/كغ

خواص المتوسط التوافقي:

1 – يأخذ بعين الاعتبار جميع القيم وتأثره بالقيم الشاذة أقل من تأثر المتوسط الحسابي.

2 – لا يمكن حسابه في حالة وجود بيانات معدومة.

3 – يعطي نتائج أكثر واقعية في حالة حساب متوسطات الأسعار والسرعة.

4 – قيمته دائما أقل من قيمة المتوسط الهندسي.

مما سبق فإن $\overline{\mathbf{X}} > \mathbf{G} > \mathbf{H}$

مثال12:

أوجد المتوسط الحسابي والمتوسط الهندسي والمتوسط التوافقي للبيانات المبوبة في الجدول الإحصائي التالي ثم قارن بينها؟.

الفئة	3-1	5-3	7-5	9-7	11-9	المجموع
التكرار	2	4	6	4	2	18

الحل:

الفئة	التكرار	X_i	$N_i\ X_i$
3-1	2	2	4
5-3	4	4	16
7-5	6	6	36
9-7	4	8	32
11-9	2	10	20
المجموع	18		108

المتوسط الحسابي= $\overline{X} = \frac{\sum n_i x_i}{\sum n_i} = \frac{108}{18} = 6$

المتوسط الهندسي: $LogG = \frac{1}{N}\left[\sum N_i LogX_i\right]$

$$= \frac{1}{18}[2Log2 + 4Log4 + 6Log6 + 4Log8 + 2Log10]$$

$$= \frac{1}{18}[0{,}602 + 2{,}408 + 4{,}669 + 3{,}612 + 2]$$

$$= \frac{1}{18} \times 13{,}291$$

$$LogG = 0{,}7384$$

ومنه G= 5,48

المتوسط التوافقي:

$$H=\frac{\sum N_i}{\sum \frac{N_i}{X_i}}=\frac{18}{\frac{2}{2}+\frac{4}{4}+\frac{6}{6}+\frac{4}{8}+\frac{2}{10}}$$

$$H=\frac{18}{1+1+1+\frac{1}{2}+\frac{1}{5}}=\frac{180}{10+10+10+5+2}=\frac{180}{37}=4{,}86$$

ويلاحظ بوضوح: $\overline{X}>G>H$

رابعا: الوسيط Mediane

تبين لنا عند دراستنا للمتوسط الحسابي أن هذا المتوسط يعطي نتيجة صحيحة ومنطقية عندما تكون البيانات التي حسب منها متجانسة ومتقاربة، أما إذا كانت تحتوي على قيم متطرفة في الصغر أو الكبر فإن النتيجة التي يعطيها تكون غير واقعية، في مثل هذه الحالات فقد وجد متوسط آخر سمي بالوسيط الذي هو أكثر واقعية ودلالة وصحة للحصول على فكرة عامة عن حالة البيانات التي بها قيم متطرفة.

تعريف: وكتعريف فإن الوسيط هو القيمة التي تقسم مجموع البيانات إلى قسمين بحيث يكون نصف عدد البيانات أكبر منه ونصف عدد البيانات أصغر منه ويرمز له بالرمز: Me.

مثال13:

أوجد الوسيط للبيانات التالية: 8، 6، 8، 10، 12، 15، 9؟

الحل:

لإيجاد الوسيط نقوم بالتالي:

أ) نرتب البيانات تصاعديا أو تنازليا:

6، 8، 8، 9، 10، 12، 15.

ب) نبحث عن الوسيط لهذه البيانات، وهناك حالتين:

1- إذا كان عدد البيانات فردي فإن الوسيط هو القيمة التي ترتيبها $\frac{n+1}{2}$

2 – إذا كان عدد البيانات زوجي فإن الوسيط هو متوسط القيمتين اللتين ترتيبهما $\frac{n}{2}$ و $\frac{n}{2}+1$

.

وفي مثالنا فإن عدد البيانات المعطاة هو 7 أي فردي وبالتالي فإن الوسيط هو القيمة التي ترتيبها $\frac{7+1}{8}=4$ وهو **9.**

ويلاحظ جليا أن عدد البيانات أقل من 9 يساوي عدد البيانات أكبر من 9.

مثال14:

البيانات التالية تمثل الدرجات التي تحصل عليها 10 طلبة في امتحان معين:

16، 17، 17، 15، 14، 16، 15، 13، 4، 3.

المطلوب:أ) إيجاد متوسط درجات هؤلاء الطلبة؟.

ب) إيجاد وسيط الدرجات؟

ج) أيهما أكثر تعبيرا على نتائج الطلبة ولماذا؟.

الحل: أ) **متوسط الدرجات** $\overline{X}=\frac{16+17+17+15+14+16+15+13+4+3}{10}$ **=**

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

ب) الوسيط: نرتب أولا هذه البيانات ترتيبا تصاعديا: 3، 4، 13، 14، 15، 15، 16، 17، 17.

بما أن عدد البيانات زوجي فإن الوسيط هو المتوسط الحسابي للقيمتين اللتين ترتيبهما $\frac{n}{2}$ و $\frac{n}{2}+1$ وهما 15 و15 وبالتالي فإن الوسيط هو =15= Me.

ج) نلاحظ من النتائج المتحصل عليها أن المتوسط الحسابي $\left(\overline{X}=13\right)$ لم ينصف أغلب (8) الطلبة الذين تحصلوا على درجات أكبر بكثير من هذا المتوسط وأنحاز ناحية نتيجة الطالبين اللذين تحصلا على نتائج سيئة، في حين أن وسيط هذه الدرجات (Me = 15) قسم نتائج الطلبة إلى قسمين بحيث نصف عدد الطلبة تحصلوا على درجات أعلى منه ونصف عدد الطلبة تحصلوا على درجة أقل منه وهو هنا أصدق من المتوسط الحسابي:

الوسيط في حالة بيانات متكررة:

الوسيط لبيانات متكررة:

إذا كانت البيانات المراد حساب الوسيط لها متكررة (لها تكرارات) فإن الوسيط يوجد بإتباع الخطوات التالية:

10. نحسب التكرار المتجمع الصاعد لقيم الظاهرة.
11. نحدد ترتيب الوسيط $\frac{N}{2}$ (حيث N = مجموع التكرارات).
12. نبحث عن القيمة التي تكرارها المتجمع الصاعد أكبر من $\frac{N}{2}$ مباشرة وهي القيمة التي تمثل الوسيط.

مثال15:

الجدول التالي يمثل توزيع الطلبة فوج معين حسب الدرجات التي تحصلوا عليها في الفرض الفجائي في مادة الإحصاء. المطلوب إيجاد الوسيط لهذه البيانات؟.

الدرجة	4	5	6	7	8	9	المجموع
عدد الطلبة	2	4	8	6	5	4	**29**

1) نحسب التكرار المتجمع الصاعد للبيانات المعطاة:

الدرجة X_i	4	5	6	7	8	9	المجموع
عدد الطلبة N_i	2	4	8	6	5	4	**29**
التكرار المتجمع الصاعد	2	6	14	20	25	29	/

2) نحسب ترتيب الوسيط = $\frac{N}{2}$ = 14.5.

3) القيمة التي تكرارها المتجمع الصاعد أكبر مباشرة من $\frac{N}{2}$ هي7 وبالتالي فإن الوسيط = 7.

الوسيط لبيانات مبوبة في جداول توزيع تكراري:

لتحديد الوسيط لبيانات مبوبة في جداول توزيع تكراري فإننا نقوم بالتالي:

1) نحسب التكرار المتجمع الصاعد أو النازل.

2) نحدد ترتيب الوسيط وهو عبارة عن نصف مجموع التكرارات $\frac{N}{2}$.

3) نحدد الفئة الوسيطية أي الفئة التي يقع فيها الوسيط، وهي الفئة التي تقابل التكرار المتجمع الصاعد الذي يساوي ترتيب الوسيط أو أكبر منه مباشرة.

4) نحدد ونحسب الوسيط بتطبيق العلاقة الإحصائية للوسيط.

$$Me = L_1 + \frac{\frac{N}{2} - N_0}{n_e}.K$$

استنتاج العلاقة التي تحسب الوسيط:

لتبيان كيف تحصلنا على العلاقة التي تحسب الوسيط نأخذ المثال التالي:

<u>مثال16:</u>

يبين التوزيع التكراري التالي توزيع 50 طالب حسب الدرجة المتحصل عليها في امتحان مادة ما.

المطلوب: 1) استنتج المعادلة التي تحسب الوسيط؟.

2) أحسب قيمة الوسيط لهذه البيانات؟.

فئة الدرجات	4−2	6−4	8−6	10−8	12−10	14−12	16−14	**المجموع**
عدد الطلبة	2	6	8	10	14	6	4	**50**

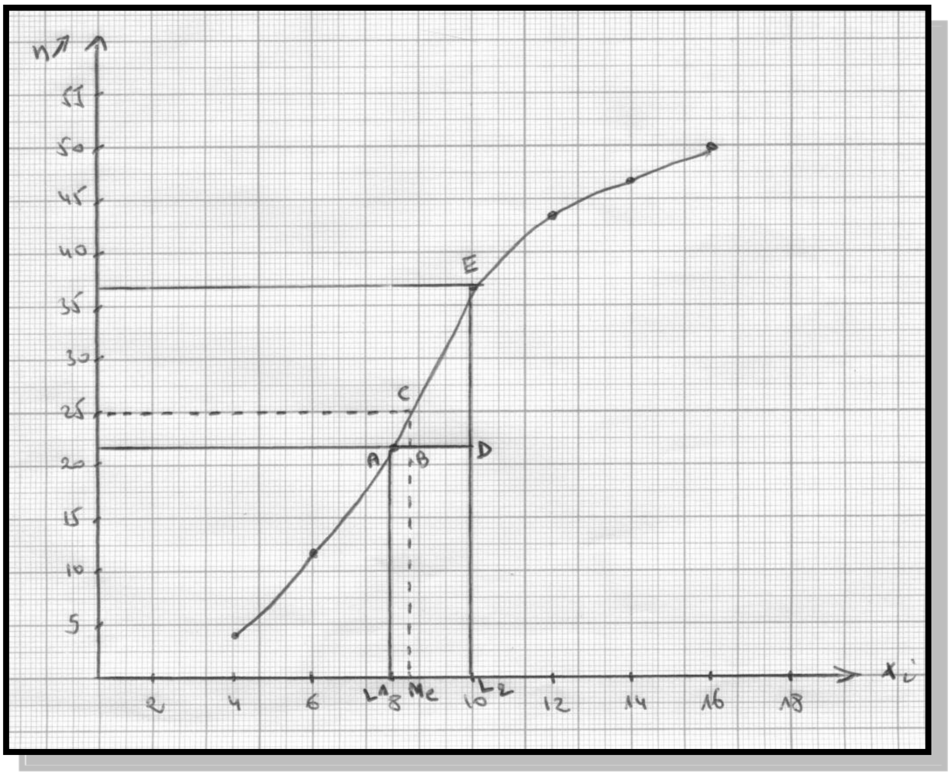
<u>الحل:</u>

أ) نحسب التكرار المتجمع الصاعد لبيانات التوزيع كما هو مبين في الجدول الآتي:.

ب) نحدد ترتيب الوسيط = $\frac{50}{2} = 25$

ج) نرسم منحنى التكرار المتجمع الصاعد لبيانات التوزيع.

الفئة	التكرار	التكرار المتجمع الصاعد
4−2	4	4
6−4	8	12
8−6	10	22
10−8	15	37
12−10	6	43
14−12	4	47<
16−14	3	50
المجموع	**50**	/



د) نحدد الفئة الوسيطية أي الفئة التي تكرارها المتجمع يساوي ترتيب الوسيط أو الأكبر منه مباشرة، (أي الفئة التي تكرارها المتجمع الصاعد ≤ 25 وهي هنا الفئة [8–10[).

هــ) نحدد التكرار المتجمع الصاعد للفئة الوسيطية والتكرار المتجمع الصاعد للفئة السابقة ونربط بينهما وبين الحد الأدنى والحد الأعلى للفئة الوسيطية على الرسم.

و) من النقطة التي تمثل ترتيب الوسيط نرسم خط مستقيم أفقي يقطع المنحنى في نقطة فاصلتها تساوي الوسيط.

وإذا افترضنا أن منحنى التكرار المتجمع الصاعد هو عبارة عن خط مستقيم عند الفئة [8–10[فإن الميل يكون ثابتا.

ي) نستخرج علاقة الوسيط كما يلي:

المثلثان ABC و ADE متشابهان، ومنه وحسب نظرية طاليس فإنه

$$\frac{\overline{\overline{AB}}}{\overline{\overline{BC}}} = \frac{\overline{\overline{AD}}}{\overline{\overline{DE}}}$$

وبالتعويض:

$$\frac{M_e - L_1}{\frac{N}{2} - N_0} = \frac{L_2 - L_1}{N_1 - N_0}$$

حيث: M_e = الوسيط.

L_1 = الحد الأدنى للفئة الوسيطية.

N = مجموع التكرارات.

N_0 = التكرار المتجمع الصاعد للفئة قبل الفئة الوسيطية.

L_2 = الحد الأعلى للفئة الوسيطية.

N_1 = التكرار المتجمع الصاعد للفئة الوسطية.

$$M_e - L_1 = \frac{(\frac{N}{2} - N_0)(L_2 - L_1)}{N_1 - N_0}$$

$$\mathbf{M_e = L_1 + \frac{\frac{N}{2} - N_0}{n_e}.K}$$

وذلك لأن (L_2-L_1) = K = طول الفئة و N_1-N_0 = ne = تكرار الفئة الوسيطية

2 – نحسب الآن قيمة الوسيط باستخدام العلاقة السابقة

$$M_e = 8 + \frac{25-22}{15}2 = 8 + 0,4$$

$M_e = 8,4$

<u>**خواص الوسيط:**</u>

يتصف الوسيط بعدة خصائص أهمها:

1 – لا يتأثر بالقيم المتطرفة وبالتالي فإنه يعتبر أصلح المقاييس عن وجود مثل هذه القيم.

2 – يمكن إيجاد الوسيط من الرسم.

3 – يمكن استخدامه في حالة البيانات النوعية.

4 – يمكن حسابه من الجداول التكرارية المفتوحة.

خامسا – أشباه الوسيط

الوسيط هو القيمة التي تقسم مجموع البيانات إلى قسمين متساويين، بحيث نصف عدد البيانات أقل منه ونصف البيانات أكبر منه، وما دام يمكن تقسيم بيانات أي ظاهرة إلى عدة أقسام متساوية وليس إلى قسمين فقط فإنه يمكن التعامل معه القيم التي تقسم هذه البيانات بنفس طريقة التعامل مع الوسيط.

فإذا تم تقسيم البيانات إلى أربعة أقسام فإن المقياس يسمى بالربيع.

وإذا تم تقسيم البيانات إلى عشرة أقسام فإن المقياس يسمى بالعشير.

أما إذا تم تقسيم البيانات إلى 100 قسم فإن المقياس يسمى بالمئين.

الربيعات Les quartils:

هي القيم التي تقسم مجموع البيانات إلى أربع أجزاء متساوية فمثلا:

الربيع الأول Q1: ويسمى كذلك بالربيع الأدنى وهو القيمة التي تقسم مجموع البيانات إلى قسمين بحيث ربع عدد البيانات أقل منه وثلاثة أرباع البيانات أكبر منه وترتيب الربيع الأول

هو $\frac{N}{4}$ ويحسب كالتالي:

$$Q_1 = L_1 + \frac{\frac{N}{4} - N_0}{n_{Q1}}.K$$

الربيع الثالث Q3: ويسمى كذلك الربيع الأعلى وهو القيمة التي تقسم مجموع البيانات إلى قسمين بحيث ثلاثة أرباع عدد البيانات أقل منه وربع عدد البيانات أكبر منه وترتيب الربيع الثالث هو $\frac{3N}{4}$.

ويحسب كالتالي: $$\mathbf{Q_3 = L_1 + \frac{\frac{3N}{4} - N_0}{nQ_3}.K}$$

وواضح أن الربيع الثاني هو نفسه الوسيط $Q_2 = L_1 + \frac{\frac{2N}{4} - N_0}{n_e}.K = M_e$

العشريات Les Diciles:

العشير الأول D_1: هو القيمة التي تقسم مجموع عدد البيانات إلى قسمين بحيث عشر عدد البيانات اقل منه وتسعة أعشار عدد البيانات أكبر منه وترتيبه هو $\frac{N}{10}$ ونحسب بالتالي:

$$\mathbf{D_1 = L_1 + \frac{\frac{N}{10} - N_0}{n_{d3}}.K}$$

المئينات: Les centiles :

إذا قسمت البيانات إلى مائة قسم متساوي فإن نقاط التقسيم هذه تسمى المئينات. **فالمئين الأول C_1** هو القيمة التي يسبقها 1% من البيانات ويليها 99% من البيانات ويحسب كالتالي:

$$\mathbf{C_1 = L_1 + \frac{\frac{N}{100} - N_0}{n_{C1}}.K}$$

والمئين العشرين C_{20} يحسب كالتالي:

$$\mathbf{C_{20} = L_1 + \frac{\frac{20N}{100} - N_0}{n_{C20}}.K}$$

مثال17:

أحسب الربيع الأول والعشير الثالث والمئين الستين للبيانات التاليـــة التـــي توضـــح درجات 50 طالب في مادة ما؟

الدرجة	4−2	6−4	8−6	10−8	12−10	14−12	16−14	المجموع
عدد الطلبة	4	8	10	15	6	4	3	50

الدرجة	عدد الطلبة	التكرار المتجمع الصاعد
4−2	4	4
6−4	8	12
8−6	10	22
10−8	15	37
12−10	6	43
14−12	4	47
16−14	3	50
المجموع	**50**	/

$$Q_1 = L_1 + \frac{\frac{N}{4} - N_0}{n_{Q1}}.K = 6 + \frac{12.5 - 12}{10} \times 2 = 6{,}1$$

$$D_3 = L_1 + \frac{\frac{3N}{10} - N_0}{n_{D3}}.K = 6 + \frac{15 - 12}{10} \times 2 = 6{,}6$$

$$C_{60} = L_1 + \frac{\frac{60N}{100} - N_0}{n_{C60}}.K = 8 + \frac{30 - 22}{15} \times 2 = 9{,}067$$

سادسا: المنوال Mode

المنوال هو القيمة الأكثر شيوعا أو تكرار في مجموعة القيم؛ والمنوال قد يكون وحيد القيمة كما قد يكون هناك أكثر من منوال لنفس التوزيع، وسنرمز له بالرمز **Mo**.

مثال18:

الجدول التالي يبين إنتاج مصنع للأحذية من المقاسات المختلفة؟ أوجد المنوال؟

المقاس	40	41	42	43	44	45	46
عدد الأزواج	300	800	500	300	200	100	10

الحل:

المنوال هنا هو المقاس 41.

مثال19:

البيانات التالية تمثل التقديرات التي تحصل عليها 10 طلاب في مادة الرياضيات.

المطلوب: إيجاد المنوال

ممتاز، جيد، جيد جدا، جيد، متوسط،، فوق المتوسط، جيد، ضعيف، جيد جدا، جيد.

الحل:

المنوال هنا هو جيد.

المنوال لبيانات مبوبة في جداول توزيع تكراري:

لإيجاد المنوال من الجداول التوزيع التكراري نبحث عن الفئة المنوالية وهي الفئة التي يقابلها أكبر تكرار. وهناك أكثر من طريقة لحساب المنوال:

أ) يمكن اعتبار مركز الفئة المنوالية كمنوال على وجه التقريب.

ب) يمكن حساب المنوال بالاعتماد على الفرق بين تكرار الفئة المنوالية والفئتين التي قبلها والتي بعدها طريقة بيرسون ويكون ذلك على النحو التالي:

إذا رمزنا للفرق بين تكرار الفئة المنوالية وتكرار الفئة التي قبلها بالرمز (d_1) وللفرق بين تكرار الفئة المنوالية التي بعدها بالرمز (d_2) ولطول الفئة بالرمز (K) ولحدها الأدنى بالرمز (L_1) فنحصل على العلاقة التالية:

$$M_0 = L_1 + \frac{d_1}{d_1 + d_2}.K^{*}$$

* ملاحظة: توجد طريقة حسابية أخرى لإيجاد المنوال تسمى بطريقة الرافعة ولكنها أقل دقة من طريقة بيرسون.

يمكن كذلك إيجاد المنوال من الرسم، وذلك من خلال رسم المدرج التكراري للفئة المنوالية وللفئتين التي قبلها والتي بعدها. نقوم بعد ذلك بإيصال نهاية المستطيل للفئة المنوالية من الناحية اليسرى بنهاية المستطيل للفئة التي بعدها من الناحية اليسرى. كذلك نهاية المستطيل للفئة المنوالية من الجهة اليمنى بنهاية المستطيل للفئة التي قبلها من الجهة اليمنى.

ومن نقطة التقاطع المستقيمين ننزل عمودا على المحور الأفقي فتكون نقطة تقاطع هذا العمود المحور مع المحور الأفقي هي قيمة المنوال.

<u>مثال20:</u>

الجدول التالي يبين توزيع عمال مؤسسة ما حسب الأجور الموزعة المطلوب:

1) تحديد قيمة المنوال؟

2) قارن بين قيمة المنوال والمتوسط الحسابي والوسيط؟.

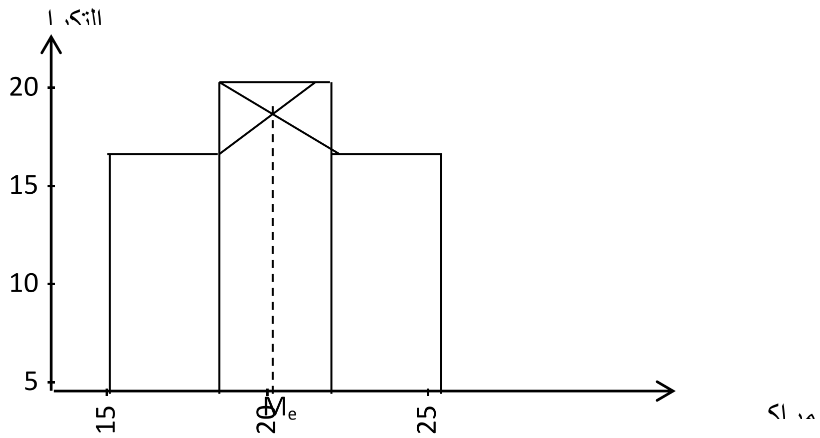
الأجور بآلاف الدينارات	15-10	20-15	25-20	30-25	35-30	-35 40	المجموع
عدد العمال	10	15	20	15	10	5	75

<u>الحل:</u>

الأجر بالآلاف	عدد العمال	X_i	n_iX_i	$n_i\nearrow$
15-10	10	12.5	125	10
20-15	15	17.5	262.5	25
25-20	20	22.5	450	45
30-25	15	27.5	412.5	60
35-30	10	32.5	325	70
40-35	5	37.5	187.5	75
المجموع	75		**1762.5**	

1) الفئة المنوالية هي الفئة [20–25[وبالتالي فإنه يمكن اعتبار مركز هــذه الفئــة (22.5) كمنوال تقريبي.

نحدد المنوال الآن بالرسم وذلك من خلال رسم المستطيلات التي تمثل كل من الفئة المنواليــة والفئتين التي قبلها والتي بعدها.



– نحسب المنوال من خلال العلاقة المتوصل إليها.

$$M_0 = L_1 + \frac{d_1}{d_1 + d_2}.K = 20 + \frac{5}{5+5}.5$$
$$M_0 = 22{,}5$$

2) لمقارنة بين $\overline{X}$,M_e ,M_1 نحس نحسب كل من المتوسط الحسابي والوسيط

$$\overline{X} = \frac{\sum n_i x_i}{\sum n_i} = \frac{1762{,}5}{75} = 23{,}5$$

$$M_e = L_1 + \frac{\frac{N}{2} - N_0}{n_2}.K = 20 + \frac{37{,}5 - 25}{20}.5 = 32{,}125$$

حيث أن الفئة الوسيطية هي نفسها الفئة المنوالية

نلاحظ من النتائج السابق أن $\overline{X} > M_e > M_0$

<u>خواص المنوال:</u>

1 – لا يأخذ بعين الاعتبار جميع البيانات المعطاة وبالتالي فهو لا يتأثر بالقيم المتطرفة.

2 – يمكن حسابه بيانيا.

3 – يمكن أن يوجد أكثر من منوال لتوزيع واحد.

4 – يمكن حساب من الجداول الإحصائية المفتوحة.

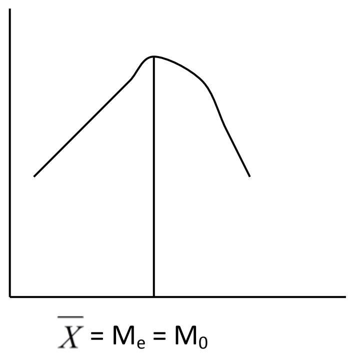
5 – يعتبر أفضل المتوسطات لوصف الظواهر النوعية.

1 – تنطبق هذه المقاييس على بعضها وتتساوى في حالة التوزيع التكراري المعتدل $\overline{X} = M_e = M_0$ وعندها يكون شكل منحنى التوزيع على شكل جرس.

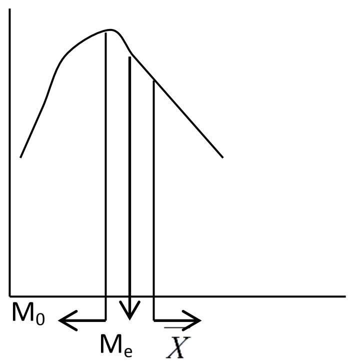
2 – عندما يكون التوزيع التكراري المدروس غير متناظر من اليمن، أي عندما تكون البيانات الصغيرة كثيرة (أكثر تكرارا) تكون المقاييس بالشكل التالي: $\overline{X} > M_e > M_0$ ويكون منحنى التوزيع ملتوي ناحية اليمين.

3 – عندما يكون التوزيع التكراري المدروس غير منتظرا من اليسار، أي عندما تكون البيانات الكبيرة كثيرة فإن المقاييس تكون بالشكل التالي: $\overline{X} < M_e < M_0$

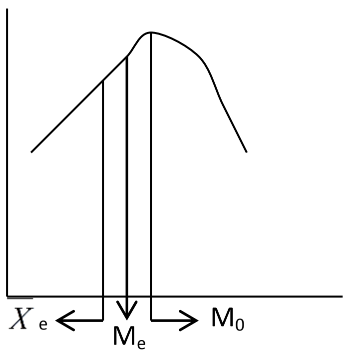
الحالة الأولى



الحالة الثانية



الحالة الثالثة



وفي كل الحالات فقد حدد كارل بيرسون علاقة تجريبية بين المقاييس الثلاث إذا كان التوزيع قريبا جدا من التناظر وهي

المتوسط الحسابي – المنوال = 3 (المتوسط الحسابي– الوسيط)

وبمقتضى هذه العلاقة يمكن استنتاج أي من المتوسطات بمعلومية المقياسين الآخرين.

سلسلة تمارين حول : مقاييس النزعة المركزية

التمرين 01:

- **الحالة الأولى:** اشترى شخص بقيمة 210000 دينار مجموعة من الأسهم بسعر 10 دينار للسهم واشترى مرة أخرى بنفس القيمة مجموعة أخرى من الأسهم بسعر 7 دينار للسهم. ما هو متوسط السعر للسهم.
- **الحالة الثانية:** الجدول التالي يبين معدل زيادة الأرباح التي حققتها شركة معينة خلال فترة رئاسة ثلاث مدراء:

الفترة	المدة بالسنوات	معدل زيادة الربح
المدير الأول	3	%5.8
المدير الثاني	1	%4.6
المدير الثالث	2	%11.2

- إيجاد معدل زيادة الأرباح خلال الفترة كاملة؟.

- **الحالة الثالثة:** قطع مسافر المسافة الكلية بين مدينتين على ثلاث مر احل كما هو مبين في الجدول أدناه:

المرحلـــة	المسافة	السرعة
الأولى	40	120
الثانية	35	100
الثالثة	25	80

- أوجد متوسط سرعة هذا المسافر على طول المسافة؟

- **الحالة الرابعة:** متوسط الدخل السنوي لموظفي أحد البنوك الجزائرية لسنة 2011 كان 65110 دينار، حيث أن متوسط الدخل السنوي للرجال هو 65380 دينار أما بالنسبة للنساء فهو 65000 دينار. ماهي نسبة العمال الرجال و النساء لهذا البنك في سنة 2011؟

التمرين 02: الجدول التالي يمثل توزيع عمال مؤسسة ما حسب السن والجنس:

الجنس / السن	الرجال	النساء
20-25	29	38
25-30	48	57
30-35	36	42
35-40	45	39
40-45	49	41
45-50	32	30
50-55	37	18
55-60	28	20

1) أحسب متوسط السن لكل من الرجال و النساء؟
2) أحسب متوسط السن لعمال المؤسسة؟

التمرين 03: أقيمت دراسة حول عدد أيام العطل المرضية على 100 عامل في شركة لسكك الحديدية، فكانت نتائج الدراسة كما يلي:

عدد الغيابات	0	1	2	3	4	5
عدد العمال	**9**	**15**	**11**	**8**	**4**	**3**

1) مثل بيانيا الجدول التالي باستخدام طريقة الأعمدة البسيطة؟ ثم استنتج قيمة المنوال بيانيا؟
2) أرسم منحنى التكرارات المتجمعة الصاعدة والنازلة؟ ثم استنتج قيمة الوسيط بيانيا؟
3) أحسب كل من المقاييس التالية: المتوسط الحسابي، المنوال و الوسيط، الربيع الأول والربيع الثالث، العشير الثالث و العشير السابع، بواسطة:

أ- التكرارات المتجمعة المطلقة. ب- التكرارات المتجمعة النسبية العادية والمئوية.

ب- أحسب عدد و نسبة العمال الذين لديهم: أ- غيابين أو أقل. ب- أقل من غيابين. ج- على الأقل غيابين.

التمرين 04: يبين التوزيع التالي بعد مساكن 60 عاملا عن مكان عملهم بالكلم:

الفئات –المسافة–	02–04	04 – 06	06 –08	08 –10	10 –12	12 – 14
التكرار –عدد العمال–	06	07	15	12	؟	05

4) ماهو تقدير عدد العمال الذين تبعد مساكنهم عن مكان عملهم بأكثر أو تساوي 03 كلم ولكنها تقل أو تساوي 9.5 كلم؟

5) أحسب كل من المتوسط الحسابي، الوسط الهندسي، التوافقي ؟ ماذا تلاحظ؟

6) أحسب الوسيط والمنوال مع تفسير معنى كل منهم ؟

التمرين 05: يمثل الجدول التالي تعداد الكريات البيضاء بالآلاف عند 50 شخص:

الفئات	1-3	3-5	7-9	9-11	11-13	13-15	15-17	17-19
التكرارات	03	05	17	14	07	02	01	01

1) أحسب المتوسط الحسابي، وأحسب قيمة المنوال وفسر معناه؟

2) حدد قيمة الوسيط و الربيع الثالث؟

3) ما هو عدد ونسبة الأشخاص الذين لهم أقل من 9000 كرية بيضاء؟ ما هو عدد ونسبة الأشخاص الذين لهم 11000 كرية بيضاء على الأقل؟

التمرين 06: الجدول التالي يبين توزيع العمال حسب عدد ساعات العمل الأسبوعية.

الفئات	38 –40	40 – 42	42 – 46	46 – 52	52 –56	56 –58
التكرارات	10	20	90	240	110	30

1) أرسم المدرج التكراري لهذا التوزيع واستنتج قيمة المنوال بيانيا ؟

2) أرسم منحنى المتجمع الصاعد والنازل ثم استنتج قيمة الوسيط بيانيا ؟

3) أحسب المتوسط الحسابي، الوسيط والمنوال ؟ ماذا تلاحظ ؟

4) أحسب عدد و نسبة العمال الذين يفوق أو يساوي عدد ساعات عملهم عن 40 ساعة ؟

التمرين 07: الجدول التالي يبين التوزيع التكراري لنفقات الاستهلاكية الشهرية بآلاف الدينارات لعينة من الأسر سحبت بطريقة عشوائية من مجتمع معين:

الفئات	0 – 4	4 – 8	8 – 12	12 – e_4	e_4 – 22	22 – 30	30 – 42	المجموع
التكرارات	6	n_2	n_3	17	14	11	3	100

1) أوجد قيمة كل من n_2 و n_3 علما ان العشير الرابع يساوي 9.5 ؟

2) أوجد قيمة e_4 علما أن المتوسط الحسابي يساوي 13؟

3) حدد الانفاق الاستهلاكي للأسرة ذات المرتبة 75%؟

التمرين 08

قام أستاذ بإجراء امتحان لطلبته المقسمين إلى مجموعتين فإذا توفرت لك المعلومات التالية.

عدد طلبة المجموعة الأولى = 248 المتوسط الحسابي لدرجاتهم = 09.5.

عدد طلبة المجموعة الثانية = 230 المتوسط الحسابي لدرجاتهم = 10.4.

أوجد المتوسط الحسابي لمجموع الطلبة؟

التمرين 09

إذا علمت بأن متوسط عدد الطلبة بالحافلات الصغيرة في إحدى رحلات الجامعة كان 23 طالبا، والمتوسط بالحافلات الكبيرة كان 65 طالبا. أوجد عدد الطلبة المشاركين في الرحلة إذا كان عدد الحافلات الصغير = 3 وعدد الحافلات الكبيرة =4؟

التمرين 10

ليكن التكرار المتجمع الصاعد للظاهرة (×) على الشكل التالي 10 30 70 90 100

فإذا علمت أن طول الفئة عبارة عن جداء التكرار الأول في التكرار الأخير وأن الحد الأعلى للفئة الثالثة عبارة عن جداء تكرار الفئة الثانية في تكرار الفئة الرابعة.

المطلوب:

- إعادة تكوين حساب الجدول
- حساب المتوسط الحسابي.

التمرين 11

الجدول التالي يمثل مبيعات 500 حذاء خلال أسبوع الدخول المدرسي

قياس الحذاء	35	36	37	38	39	40	41	42	43	44
العدد	50	60	55	35	60	65	α	β	45	10

المطلوب :

– تحديد طبيعة المتغير الإحصائي.

حساب α β إذا علمت أن $\alpha = \beta$ ماذا يمثل كل من α، β.

– إيجاد قيمة المنوال من الرسم أن أمكن؟

– حساب قيمة المنوال؟

التمرين 12

إذا كان مستهلكي سلعة ما موزعين حسب فئات العمر كما يظهر في الجدول التالي:

السن	15-20	20-25	25-30	30-35	35-40	40-50	أكبر من 50	المجموع
النسبة	13	26	28	15	10	5	3	100%

المطلوب:

- أرسم المدرج التكراري للتوزيع واستنتج منه المنوال
- أرسم منحنى التكرار النسبي المتجمع الصاعد والنازل واستنتج منه الوسيط
- أوجد حسابيا قيم المنوال، الوسيط، المتوسط الحسابي

التمرين 13

الجدول التالي يمثل توزيع إنتاج 300 عامل بإحدى المؤسسات:

الإنتاج	50-60	60-70	70-80	80-90	90-100	100-110	110-120	120-130	130-140	140-150
عدد العمال	4	6	10	15	20	40	80	50	45	30

- أوجد الربيع الأول والعشير الثاني والمئين الستين وما هي دلالات كل منهما؟.
- أوجد أقل إنتاج سجل بواسطة 25 بالمائة الأحسن إنتاجا؟.
- أعلى إنتاج سجل بواسطة 20 بالمائة الأقل إنتاجا في المؤسسة؟.

عرفنا في الفصل السابق أن مقاييس النزعة المركزية (من متوسط و وسيط ومنوال) تسمح لنا بالحصول على القيم المتوسطة للبيانات أو على تجمعها، غير أن هذه المقاييس لا تكفي لوحدها لمعرفة الصفات الإحصائية اللازمة لوصف الظواهر، لأن الفروق بين قيم الظواهر قد تزداد أو تنقص رغم تساوي المتوسطات لهذه الظواهر، ولتوضيح ما سبق نفترض أن طالبين تحصلا على النتائج التالية في خمس مواد دراسية:

الطالب x	10	11	13	14	15
الطالب y	8	9	13	15	18

فمتوسط درجات الطالب (x) يساوي 12.6 وكذلك متوسط درجات الطالب (y) يساوي 12.6 و وسيط درجات الطالب (x) يساوي 13 وكذلك وسيط درجات الطالب (y) يساوي 13. قد يفهم مما سبق أن الطالبين (x)و (y) لهما نفس المستوى غير أن التمعن الجيد في الدرجات التي تحصل عليها الطالبين تبين أن الطالب (x) ناجح في كل المواد المدروسة في حين أن الطالب (y) ناجح في ثلاث مواد فقط. إن هذه الحقيقة تبين أن مقاييس النزعة المركزية لا تعطي فكرة وافية عن اختلاف قيم الظواهر ولا تحقق كل الأغراض التي نرغب الوصول إليها من دراستنا لذلك فإن مقاييس النزعة المركزية لا بد أن تكون مصحوبة بمقاييس أخرى لقياس مدى تباعد أو تقارب البيانات من بعضها البعض أو من متوسطها، تسمى هذه المقاييس بمقاييس التشتت.

ما معنى التشتت؟.

تشتت بيانات ظاهرة ما يقصد به درجة أو مقدار التفاوت أو الاختلاف بين مفردات هذه الظاهرة، وتعتبر بيانات الظاهرة متجانسة عندما تكون قيمتها قريبة من بعضها البعض ونقول في هذه الحالة أنها غير مشتتة. أما إذا كانت بيانات الظاهرة متباعدة وغير متجانسة فنقول أن مفردات الظاهرة مشتتة وغير مركزة. ويقاس تشتت البيانات بعدة مقاييس منها:

أولا – المدى (المطلق)

المدى لمجموعة من البيانات هو الفرق بين أكبر قيمة وأصغر قيمة لها ويرمز له بالرمز R. المدى= أكبر قيمة – أصغر قيمة.

أما المدى للتوزيعات التكرارية فيحسب كما يلي: المدى = الحد الأعلى لفئة الأخيرة – الحد الأدنى للفئة الأولى

مثال1: أوجد المدى للبيانات التالية: 12،18،22،28،30.

مثال2: أوجد المدى للبيانات التالية 65،20،17،−4، 18، 19، 14

نلاحظ المدى في هذا المثال قد تأثر بشكل كبير جدا بالقيم المتطرفة، إذ نلاحظ أن معظم البيانات متقاربة باستثناء القيمة 65 والقيمة (−4)، فإذا استبعدنا هذه القيم المتطرفة فإن المدى يصبح: $R = x_{max} - x_{min} = 20 - 14 = 6$

وبسبب هذا العيب فإن المدى كمقياس للتشتت لا يستخدم إلا عندما نرغب في مقياس تقريبي وسريع لتشتت البيانات دون الاهتمام بالدقة في المقياس، أو عندما يكون للبيانات المتطرفة أهمية خاصة كتوزيعات درجات الحرارة على سبيل المثال، حيث تعلن درجات الحرارة اليومية بحدها الأقصى وحدها الأدنى خلال اليوم، كما يشيع استخدام هذا المقياس في حالات مراقبة جودة الإنتاج أو متابعة المبيعات التي يحققها رجال البيع لمؤسسة ما.

أما إذا أردنا أن نقلل من أثر القيم المتطرفة فإننا نقوم باستبعادها ويمكن أن يتم ذلك باستخدام الطرق التالية:

- المدى الربيعي = الربيع الثالث – الربيع الأول.
- المدى العشيري = العشير التاسع – العشير الأول.
- المدى المئيني = المئين 99 – المئين الأول.

خواص المدى:

1. يتصف المدى بسهولة حسابه.
2. يعتمد في حساب على قيمتين فقط هما القيمة الكبرى والقيمة الصغرى.
3. بسبب الخاصية الثانية فإن المدى شديد التأثر بالقيم المتطرفة.

ثانيا: الانحراف المتوسط L'écart moyen:

لأن مقاييس التشتت هي مقاييس لقوة تجمع البيانات حول بعضها، وحيث أن التجمع يكون حول القيم المتوسطة، فإنه إذا كان مقدار الاختلاف بين القيم ومتوسطها كبيرا دل ذلك على أن التشتت كبير والعكس صحيح. وحيث أن مجموع الاختلافات (الانحرافات) عن المتوسط يساوي صفرا فإنه لو حسبنا القيم المطلقة لمقدار الاختلاف عن المتوسط يكون متوسط هذه الاختلافات مقياسا مناسبا لمقدار التشتت، يسمى هذا المقياس بالانحراف المتوسط. ويعرف الانحراف المتوسط بأنه المتوسط الحسابي للقيم المطلقة لانحرافات القيم عن متوسطها الحسابي، وسوف نرمز للانحراف المتوسط في دراستنا بالرمز E_x وعليه:

إذا كانت لدينا القيم التالية: $x_1, x_2, x_3, \dots, x_n$ فإن الانحراف المتوسط لها هو:

$$E_x = \frac{|x_1-\bar{x}|+|x_2-\bar{x}|+|x_3-\bar{x}|+\cdots+|x_n-\bar{x}|}{n} = \frac{\sum_{i=1}^{n}|x_i-\bar{x}|}{n}$$

أما إذا كانت البيانات مكررة أو مبوبة في جدول توزيع تكراري فإن الانحراف المتوسط لها يعطي العلاقة:

$$E_x = \frac{\sum_{i=1}^{n} n_i|x_i-\bar{x}|}{\sum_{i=1}^{n} n_i}$$

<u>مثال3</u>: أوجد الانحراف المتوسط للبيانات التالية: 8،6،5،4،2؟

<u>مثال4</u>: أوجد تشتت البيانات المبوبة في جدول التوزيع التكراري الآتي باستخدام الانحراف المتوسط؟

الفئة	0−2	2−4	4−6	6−8	المجموع
التكرار	2	3	4	3	12

<u>خواص الانحراف المتوسط:</u>

1. يعتمد في حسابه على جميع القيم وليس على القيمة الكبرى والصغرى فقط.
2. يتأثر بالقيم المتطرفة، لأن انحرافها عن المتوسط الحسابي يكون كبيرا.

3. يعتبر الانحراف المتوسط أفضل من سابقه (المدى) لأنه أقل تأثر بالقيم المتطرفة غير أنه لا يستعمل بشكل واسع بسبب إعتماده على القيمة المطلقة لانحرافات القيم عن متوسطها الحسابي.

ثالثا – التباين والانحراف المعياري La variance et l'écart type:

أ – التباين La variance: وهو عبارة عن المتوسط الحسابي لمربعات الفروق بين قيم المتغير الإحصائي ومتوسطها الحسابي، ونستخدم مربعات الفروق هنا تفاديا لاستخدام القيم المطلقة كما هو الشأن في الانحراف المتوسط.

فإذا كانت لدينا البيانات التالية: $x_1, x_2, x_3, \dots, x_n$ فإن التباين لهذه البيانات يعطي بالعلاقة:

$$V_x = \frac{(x_1-\bar{x})^2+(x_2-\bar{x})^2+(x_3-\bar{x})^2+\cdots+(x_n-\bar{x})^2}{n} = \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{\sum_{i=1}^{n} n_i} = \frac{\sum_{i=1}^{n} x_i^2}{n} - \bar{x}^2$$

أما اذا كانت البيانات مبوبة او في جدول توزيع تكراري فيمكن حساب التباين بالعلاقة التالية:

$$V_x = \frac{\sum_{i=1}^{n} n_i(x_i-\bar{x})^2}{\sum_{i=1}^{n} n_i} = \frac{\sum_{i=1}^{n} n_i x_i^2}{\sum_{i=1}^{n} n_i} - \bar{x}^2$$

مثال5: أوجد التباين للبيانات التالية: 3،7،6،1،11،5،6،9.

ب – الانحراف المعياري: ويعتبر الانحراف المعياري من أهم المقاييس الإحصائية للتشتت، وهو أكثر استخداما في النظريات والقوانين الإحصائية، لأنه يعطي فكرة سليمة ومنطقية عن ظاهرة التشتت، ويعرف الانحراف المعياري بأنه الجذر التربيعي لمتوسط مجموع مربع انحراف القيم عن متوسطها، أي أنه الجذر التربيعي للتباين. و نرمز له عادة بالرمز: S_x. و منه: $S_x = \sqrt{V_x}$

مثال6: أوجد التباين والانحراف المعياري للبيانات المبوبة في الجدول الإحصائي التالي:

الفئة	4-8	9-13	14-18	19-23	24-28	المجموع
التكرار	3	4	6	2	4	19

خصائص الانحراف المعياري:

1. إذا كان الانحراف المعيار للقيم $x_1, x_2, x_3, \dots, x_n$ هو S_x فإنه إذا أضيفت أو طرحت قيمة ثابتة (a) إلى أو من جميع القيم فإن الانحراف المعياري لا يتغير
2. إذا كان الانحراف المعيار للقيم $x_1, x_2, x_3, \dots, x_n$ هو S_x فإنه إذا ضربت كل قيمة بالمقدار a (قسمت كل قيمة على المقدار a) فإن الانحراف المعياري يتأثر بالمقدار نفسه
3. يأخذ الانحراف المعياري نفس وحدة القياس للمتغير الأصلي (كلغ، متر، لتر ....) لذلك لا يمكن استخدامها كأساس للمقارنة بين تشتت توزيعين لهما وحدات قياس مختلفة.
4. بما أن الانحراف المعياري يتأثر بالمتوسط الحسابي لبيانات الظاهرة فإنه لا يمكن استخدامه للمقارنة بين تشتت بيانات توزيعين لهما متوسط حسابي مختلف ولو كان هذين التوزيعين من نفس النوعية
5. إذا كان لدينا مجموعة كلية متكونة من مجموعتين جزئتين أو أكثر فإنه يمكن حساب الانحراف المعياري لها من خلال العلاقة التالي:

$$s_x = \sqrt{\frac{1}{n}\left[n_1 S_{x1}^2 + n_2 S_{x2}^2\right] + \frac{1}{n}\left[n_1(\bar{x}_1 - \bar{x})^2 + n_2(\bar{x}_2 - \bar{x})^2\right]}$$

رابعا – معامل الاختلاف Coefficient de variation:

رأينا سابقا أن الانحراف المعياري هو مقياس واقعي ومؤشر صحيح عن مقدار التشتت غير أن الخاصيتين 3و4 السابقتين تبينان أنه إذا استخدمنا هذا المقياس للمقارنة بين تشتت ظاهرتين أو أكثر فإن المقارنة تكون واقعية فقط إذا كانت الظواهر من نوعية واحدة ولها متوسطات متساوية. أي يمكن مقارنة تشتت درجات مادة ما بدرجات مادة أخرى أو مقارنة تشتت دخل مجموعة من العمال بدخل مجموعة أخرى، وتكون المقارنة أكثر واقعية إذا كانت المتوسطات متساوية أو قريبة من بعضها. أما إذا كانت الظواهر من صفات مختلفة أو إذا كانت متوسطاتها متباعدة، فإن المقارنة اعتمادا على الانحراف المعياري ستكون غير منطقية وغير واقعية، ولهذا السبب وجدت مقاييس أخرى سميت مقاييس التشتت النسبي تعتمد على تمييز البيانات وتقيس التشتت كنسبة مئوية للمتوسط، أهم هذه المقاييس هو معامل الاختلاف. نرمز له عادة بالرمز: CV

$$CV = \frac{S_x}{\bar{X}} \times 100$$

مثال7: إذا كان متوسط درجات مجموعة من الطلبة في مادة ما هو 15 بانحراف معياري 3 ومتوسط درجاتهم في مادة أخرى هو 8 بانحراف معياري 2، فأي الدرجات في نظرك أكثر تشتتا؟

مثال8: ينتج مصنع نوعين من المصابيح الكهربائية، فإذا علمت أن المتوسط الحسابي والانحراف المعياري لعمر المصباح في كل نوع هما:

	المتوسط الحسابي	الانحراف المعياري
المصنع الأول	1500	300
المصنع الثاني	1800	325

أي المصابيح لها مدة حياة أكثر تشتت؟

سلسلة تمارين حول : مقاييس التشتت و الشكل

التمرين 01: برهن أن:

1) $V_x = \frac{\sum_{i=1}^{n} x_i^2}{n} - \bar{x}^2$.

2) إذا كان الانحراف المعياري للقيم $_1, x_2, x_3, \dots, x_n$ هو S_x فإنه إذا أضيفت أو طرحت قيمة ثابتة (a) إلى أو من جميع القيم فإن الانحراف المعياري لا يتغير.

3) إذا كان الانحراف المعياري للقيم $_1, x_2, x_3, \dots, x_n$ هو S_x فإنه إذا ضربت كل قيمة بالمقدار a (قسمت كل قيمة على المقدار a) فإن الانحراف المعياري يتأثر بالمقدار نفسه

التمرين 02: كان توزيع 50 طالب حسب عدد الغيابات كالتالي:

عدد الغيابات	0	1	2	3	4	5
عدد الطلبة	**9**	**15**	**11**	**8**	**4**	**3**

1) أحسب كل من المقاييس التالية: المدى العام، المدى الربيعي، الانحراف المتوسط والانحراف المعياري ؟

2) أدرس شكل منحنى التوزيع ؟

التمرين 03: البيانات التالية تمثل توزيع المستخدمين بالآلاف حسب فئات الأجور (الساعة/ دينار) وذلك في إحدى المزارع الموسمية:

الفئات	5-10	10-15	15−20	−25 20	25−30	30−35	35−40	40−45
n_i	**39**	**82**	**95**	**44**	**33**	**22**	**5**	**3**

1) أحسب كل من المقاييس التالية: المدى العام، المدى الربيعي، الانحراف المتوسط والانحراف المعياري ؟

2) أدرس شكل منحنى التوزيع التكراري ؟

التمرين 04: الجدول الآتي يبين أرباح الشركتين X، Y لفترة ما بملايين الدينارات: أي الشركتين أفضل بالنسبة للمستثمرين ولماذا؟

الشركة X	10	65	45	50	10
الشركة Y	35	40	35	30	40

التمرين 05: أخذت عينتان من مجتمعين مستقلين عن بعضهما البعض فأعطت النتائج التالية:

العينة الأولى	العينة الثانية
$\sum_{i=1}^{30} x_i = 450$	$\sum_{i=1}^{20} y_i = 100$
$\sum_{i=1}^{30} x_i^2 = 6900$	$\sum_{i=1}^{20} y_i^2 = 3500$

3) أحسب المتوسط الحسابي و الانحراف المعياري لكل عينة ؟

4) دمجت العينتان، أحسب كل من المتوسط الحسابي و الانحراف المعياري الناتج عن دمج العينتين؟

5) أحسب معامل الاختلاف لكل عينة، أي العينتان أكثر تشتتا ؟