الفصل الأول

مقدمـــة في علم الإحصـــاء

1.1 مقدمــة:

يُعدّ علم الإحصـاء علماً مستقلاً بقوانينه وأُسسـه الرياضـية، وتتنوع مجالات تطبيقه لتشمل العلوم الطبيعية والاجتماعية والتطبيقية، وقد بدأ في بداياته بالعدّ والحصـــر فقط، حيث كان الإنســـان يسـتخدم الحصـــى كوسـيلة بدائية لإحصـــاء الأشـياء الكثيرة، ومن هنا جاءت تسـميته العربية "إحصـاء"، أما تسـميته بالإنجليزية (Statistics) فترجع إلى كلمة State، لارتباطه قديماً بتعداد السكان والموارد والمواليد والوفيات، بهدف مساعدة الدولة في التخطيط ورسم السياسات.

تطور علم الإحصـاء عبر الزمن من مجرد التعدادات إلى جمع البيانات عن مختلف الظواهر في شـتى المجالات، وأصـبح يعتمد على الرياضـيات كالتفاضـل والتكامل والجبر والهندسـة التحليلية لتفسير أساليبه وتطوير نظرياته، وبجهود العلماء تحول إلى علم مستقل له قواعده الخاصـة، ومع ظهور الحاسبات الآلية ازداد تطوره بشكل ملحوظ، مما سهّل تحليل كميات ضـخمة من البيانات بسرعة ودقة.

2.1 تعريف علم الإحصاء:

علم الإحصـاء هو العلم الذي يبحث في المبادئ والقوانين والطرق العلمية المختلفة لجمع البيانات وتبويبها وعرضـها وتحليلها، ثم استخلاص النتائج والتعميمات والتوصـل إلى أحسن القرارات لحل المشاكل المختلفة على أساس من التحليل العلمي للبيانات المتاحة.

3.1 فروع علم الإحصاء:

تطوّر علم الإحصاء وأصبح يضم عدة فروع رئيسية يمكن تقسيمها إلى فرعين أساسيين:

1. الإحصاء الوصفي:

يهتم هذا الفرع بطرق جمع البيانات الإحصـائية وتنظيمها ووصـفها وتلخيصـها، وذلك من خلال استخدام الجداول التكرارية والرسوم البيانية، بالإضافة إلى بعض المقاييس الإحصائية التي تساعد في توضيح الخصائص الأساسية لهذه البيانات.

2. الإحصاء الاستنتاجي:

يركّز هذا الفرع على استنتاج القرارات المتعلقة بالظواهر المختلفة، وتعميم النتائج المستخلصة من العينات إلى المجتمع الإحصـــائي الأكبر، مع تحديد وقياس درجة الثقة المرتبطة بهذه القرارات والاستنتاجات.

4.1 أهمية علم الإحصاء وعلاقته بالعلوم الأخرى:

اكتسب علم الإحصاء أهميته من قدرته على تطبيق نظرياته ومبادئه وأساليبه في أي مجال يمكن التعبير عن ظواهره ببيانات عددية، أو حتى غير عددية يمكن تحويلها إلى شكل عددي، وبفضل ذلك أصبح بالإمكان استخدام الأساليب الإحصائية في مختلف العلوم.

ففي علم الاقتصاد، يُستخدم الإحصاء لتفسير الظواهر الاقتصادية مثل نظريات العرض والطلب، وعلاقة مستويات الدخل بالإنفاق الاستهلاكي، وقياس هذه العلاقات، بالإضافة إلى مراقبة إنتاج المصانع من حيث الكمية والجودة ومدى ملاءمتها لاحتياجات السوق وأذواق المستهلكين.

أما في علم النفس، فتُستخدم الطرق الإحصائية في قياس مستوى ذكاء الأفراد ودراسة العلاقة بين الذكاء والمهارات المختلفة.

وفي علم الأحياء، يُعتمد على الإحصـاء لدراسـة الأجناس البشـرية والفصـائل الحيوانية والنباتية وخصائصها، بالإضافة إلى دراسة العلاقات بينها.

أما في علم الوراثة، فقد اسـتُخدمت الأسـاليب الإحصـائية لدراسـة انتقال الصـفات من الآباء إلى الأبناء، وتمييز الصفات الوراثية عن الصفات المكتسبة.

كما لعب الإحصاء دوراً مهماً في علم الفلك، من خلال الدراسات المتعلقة بتحديد مدارات الكواكب والنجوم والأجرام السماوية الأخرى.

وفي الجغرافيا، سواء الطبيعية أو البشرية، تُستخدم الأساليب الإحصائية لدراسة أشكال سطح الأرض، الجغرافيا المناخية، جغرافيا البحار والمحيطات، بالإضـــافة إلى تطبيقاتها في الجغرافيا السكانية، جغرافيا المدن، وعلم الخرائط.

أما في العلوم الطبية، فيُطبق علم الإحصاء في معظم الدراسات لمقارنة الأمراض المختلفة وسبل علاجها، وتحديد العلاقة بين بعض الأمراض وأسبابها.

وبذلك يتضح أن الإحصاء أصبح في الوقت الحاضر علماً ذا أهمية كبيرة وفائدة واسعة في شتى المجالات العلمية. وبشكل عام، لا يمكن لأي تخطيط اقتصادي أو اجتماعي أن يكون صحيحاً دون الاستعانة بالأساليب والمبادئ والنظريات الإحصائية في تنفيذه وتقييم نتائجه.

5.1 أهداف علم الإحصاء:

1. تبسيط البيانات الإحصائية من خلال عرضها في جداول أو رسومات بيانية لتسهيل فهمها وتحليلها.
2. التعبير عن الحقائق بطريقة عددية دقيقة بدلاً من الأسلوب الإنشائي.
3. مقارنة المجموعات المختلفة والكشف عن العلاقات بينها.
4. التنبؤ بالبيانات المستقبلية بما يدعم عملية التخطيط واتخاذ القرارات.
5. استخلاص النتائج واتخاذ القرارات المناسبة بدقة بعد تحليل البيانات في مختلف فروع العلوم.

6.1 مصادر البيانات الإحصائية:

تتنوع مصادر الحصول على البيانات حسب موضوع الدراسة والغرض منها، وأهمها:

1. النشرات والدوريات والسجلات الرسمية.
2. التجارب العلمية.
3. الاستبيانات والمسوحات الميدانية.
4. التعدادات العامة للسكان والموارد.

7.1 بعض المصطلحات الإحصائية:

أ. المجتمع الإحصائي (قيد الدراسة):

هو مجموعة من العناصر أو الوحدات التي تشترك في الصفة التي يهتم بها الباحث.

ب. العينة:

هي جزء من المجتمع الإحصائي يتم اختيار عناصره غالباً بطريقة عشوائية، دون تدخل مباشر من الباحث.

ت. الظاهرة (المتغير):

هي صفة تتميز بها عناصر المجتمع، تختلف من عنصر لآخر من حيث الشكل أو النوع أو الكمية، وتعرف أيضاً بالمتغير.

ث. المعلمة والإحصاء:

المقادير الخاصة بالمجتمع تُسمى معلمات، بينما المقادير الخاصة بالعينة تُسمى إحصاءات.

8.1 أنواع البيانات الإحصائية:

تنقسم البيانات الإحصائية إلى نوعين رئيسيين:

أ. البيانات النوعية (الوصفية):

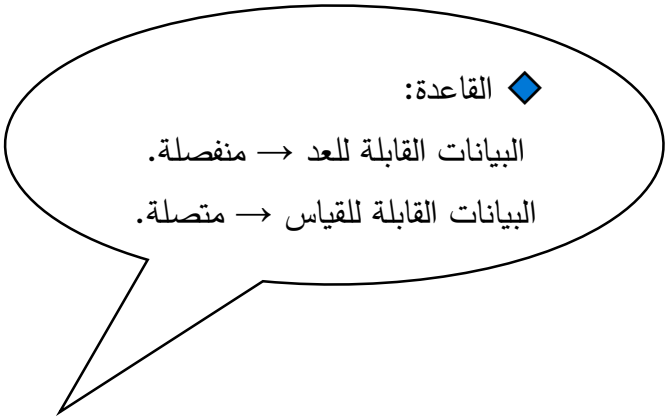
هي بيانات تصف الظواهر من حيث الصفات أو الأنواع أو الأزمنة، وتنقسم إلى:

- بيانات نوعية قابلة للترتيب: مثل مستوى الخبرة لدى المهندس، تقديرات النجاح....
- بيانات نوعية غير قابلة للترتيب: مثل نوع المعدن، اللون...

ب. البيانات الكمية (العددية):

هي بيانات يتم التعبير عنها بالأرقام، وتنقسم إلى:

- بيانات كمية منفصلة: تأخذ أعداداً صحيحة فقط، مثل عدد التروس، عدد المسامير، عدد الطلبة في قسم الهندسة الكهروميكانيكية.
- بيانات كمية متصلة: يمكن أن تأخذ أي قيمة عددية (صحيحة أو كسرية) من وحدة القياس، مثل أبعاد قطعة ميكانيكية، وزن جسم، زمن تشغيل آلة.



9.1 جمع البيانات الإحصائية:

توجد عدة طرق لجمع البيانات الإحصائية، ومن أهمها:

1. طريقة المسح (الحصر) الشامل:

في هذه الطريقة يتم جمع البيانات من جميع الوحدات المكوّنة للمجتمع محل الدراسة.

وقبل البدء في عملية جمع البيانات، ينبغي على الباحث تحديد الهدف من الدراسة بدقة، وعلى ضوء ذلك يحدد البيانات المطلوبة والضرورية للبحث، ثم يقوم بتصميم استمارة إحصائية، وهي قائمة تضم أسئلة تؤدي إجاباتها إلى توفير البيانات المطلوبة.

2. طريقة العينات:

عندما يتعذر استخدام طريقة المسح الشامل في التحليل الاقتصادي أو أي عملية بحثية، يتم اختيار جزء من المجتمع محل الدراسة بطريقة عشوائية دون تدخل مباشر من الباحث، يمكن بعد ذلك تعميم نتائج هذه العينة باستخدام الأساليب الإحصائية على المجتمع ككل، مع الأخذ في الاعتبار أنه كلما زاد حجم العينة واتبِع الأسلوب الإحصائي المناسب، كانت نتائجها أقرب إلى حقائق المجتمع وأكثر موثوقية.

1.9.1 أنواع الاستمارة الإحصائية

تنقسم الاستمارات الإحصائية إلى نوعين رئيسيين:

1. كشف البحث (Research Schedule):

استمارة يدوّن فيها الباحث أو الشخص المكلّف عملية جمع البيانات بنفسه.

يتم ذلك إمّا من خلال:

- الملاحظة المباشرة للظاهرة محل الدراسة (مثال: قياس أبعاد قطعة ميكانيكية في الورشة).
- المقابلة الشخصية مع المبحوث (مثال: سؤال المهندسين عن المعدات المستخدمة وجودتها).
- أو الاتصال الهاتفي.

في هذه الحالة يقوم الباحث بطرح الأسئلة مباشرة على المبحوث ثم يدوّن الإجابة كما هي.

يستخدم هذا النوع غالبًا في الدراسات التي تشمل مجتمعات ذات نسب أمية مرتفعة، أو في بحوث تتعلق بالأطفال حيث يصعب عليهم تعبئة الاستمارات بأنفسهم.

2. صحيفة الاستبيان (Questionnaire):

استمارة الاستبيان هي استمارة يعبئها المبحوث بنفسه دون تدخل مباشر من الباحث. يتم تسليم الاستمارة للمبحوث يدويًا أو إرسالها عبر البريد العادي أو الإلكتروني، وبعد تعبئتها يُعاد تسليمها إلى الباحث أو الجهة المشرفة على الدراسة. ولتشجيع المبحوثين على إعادة الاستمارة، يُفضل إرفاق ظرف مختوم يتضمن عنوان الباحث أو الجهة المشرفة، أو توفير خيار التعبئة الإلكترونية وإعادة إرسالها عبر البريد الإلكتروني.

أمثلة:

- استبيان لقياس آراء المهندسين حول كفاءة ماكينات جديدة.
- استبيان لتقييم تصميم محطة طاقة شمسية أو نظام تحكم ميكانيكي.

وهناك مجموعة من الاعتبارات الواجب مراعاتها عند تصميم الاستمارة الإحصائية، من أبرزها:

1. أن تحتوي الاستمارة على أقل عدد ممكن من الأسئلة، مع ضمان الحصول على جميع المعلومات المطلوبة.
2. أن تكون صياغة الأسئلة بسيطة وواضحة، مع استخدام الألفاظ الشائعة.
3. يُفضل أن تُصاغ الأسئلة بطريقة تسمح بإجابات محددة مثل: (نعم/لا) أو الاختيار من ثلاث أو أربع إجابات يحددها المصمم، ويقوم المبحوث باختيار ما يناسبه.
4. تجنّب الأسئلة المحرجة أو المتعلقة بأمور شخصية وسرية مثل الأرباح.
5. تجنّب الأسئلة الإيحائية: أي الأسئلة التي توحي بإجابة معيّنة مسبقًا. مثل: "هل تغيبت عن العمل بسبب المرض؟"، إذ غالبًا تكون الإجابة هنا بالإيجاب. والصياغة الأفضل: "لماذا تغيبت عن العمل؟".
6. تجنّب الأسئلة ذات الطابع النسبي: مثل السؤال عن مستوى الثقافة أو التدين، لأنها أمور تختلف من شخص لآخر ولا يمكن ضبطها بشكل موضوعي.

7. تخصـيص نقطة واحدة لكل سـؤال: ينبغي أن يطرح كل سـؤال موضـوعًا واحدًا فقط، مع تجنّب دمج أكثر من سـؤال في صـياغة واحدة. كما يجب ترتيب الأسـئلة بشـكل منطقي وتسلسل واضح، مع إعطاء كل سؤال رقمًا ليسهُل الرجوع إليه لاحقًا.
8. تكرار الأسئلة المهمة: إذا اعتبر الباحث أن إجابة معيّنة محورية في دراسته، يمكنه إعادة طرح السـؤال بصـيغ مختلفة وفي أماكن متفرقة من الاسـتمارة، وذلك للتحقق من دقة الإجابات وثباتها.

تمـــارين

1. ما المقصود بعلم الإحصاء؟

2. وضّـح بإيجاز كيف تطوّر علم الإحصـاء من العدّ اليدوي إلى اسـتخدام الحاسـوب في تحليل البيانات.

3. ما الفرق بين كلمة إحصـاء في اللغة العربية وكلمة Statistics في اللغة الإنجليزية من حيث الأصل والمعنى؟

4. عدّد فروع علم الإحصاء واذكر وظيفة كل فرع بإيجاز.

5. اذكر ثلاثة أسباب تبرز أهمية علم الإحصاء في المجالات الهندسية.

6. حدّد أهداف علم الإحصاء كما وردت في الفصل.

7. ميّز بين المجتمع الإحصائي والعينة مع إعطاء مثال هندسي مناسب.

8. ما المقصود بالمعلمة والإحصاء؟ وما العلاقة بينهما؟

9. صنّف البيانات الآتية إلى نوعية أو كمية، واذكر السبب:

- عدد التروس في ناقل الحركة
- درجة حرارة فرن صهر معدني
- سرعة دوران عمود الماكينة
- نوع الوقود المستخدم في المصنع

10. ما الفرق بين البيانات الكمية المنفصلة والبيانات الكمية المتصلة؟ أعط مثالين هندسيين.

11. في مصـــنع ميكانيكي، تم جمع بيانات عن عدد العيوب في 50 قطعة منتجة. وضّـــح الخطوات التي ستتبعها لتجميع هذه البيانات وتحليلها إحصائيًا.

12. إذا أراد باحث دراسـة العلاقة بين كفاءة العمال وسـاعات التدريب، فهل يسـتخدم الإحصـاء الوصفي أم الاستنتاجي؟ ولماذا؟

13. أعط مثالاً على موقف هندسي يمكن استخدام طريقة المسح الشامل فيه، وآخر تستخدم فيه طريقة العينة فقط.

14. من خلال خبرتك كطالب هندسـة، كيف يمكن أن يسـاهم الإحصـاء في رفع جودة الإنتاج الصناعي؟

الفصل الثــاني

عــرض البيــانات

العرض الجدولي للبيانات الإحصائية

1.2 جداول التوزيعات التكرارية:

البيانات التي يتم الحصول عليها في البداية من مفردات المجتمع الإحصائي تُسمّى بيانات خام غير مبوبة، هذه البيانات في صورتها الأولية لا تساعد الباحث على التعرف على الاتجاهات العامة للظاهرة محل الدراسة، كما يصعب استخدامها لإجراء التحليلات الرياضية أو لاستخراج المقاييس الإحصائية المختلفة، على سبيل المثال: إذا جمعنا قياسات الطول لعينة كبيرة من أعمدة خرسانية في موقع بناء وتركناها كما هي، فلن نتمكن من ملاحظة ما إذا كانت أغلب الأعمدة متقاربة في الطول أو متباينة بشكل كبير.

لذلك، تأتي المرحلة الثانية بعد جمع البيانات وهي تنظيمها وتبويبها وعرضها في جداول حتى يسهل التعامل معها، سواء للفهم أو للتحليل أو لاستخلاص النتائج، ويُطلق على الجدول الذي يُستخدم لهذا الغرض اسم **جدول التوزيع التكراري**، وعند تنظيم البيانات وعرضها بهذا الشكل، تتحول من بيانات خام إلى بيانات مبوبة، أي أنها أصبحت مصنفة ومرتبة بشكل يعكس خصائصها ويوضح أنماطها.

2.2 تعريف جدول التوزيع التكراري:

يمكن تعريف جدول التوزيع التكراري بأنه جدول يتكوّن من عمودين رئيسيين:

العمود الأول: تُقسَّم فيه البيانات إلى فئات محددة، إما حسب النوع (نوعية) أو حسب القيمة العددية (كمية). مثلًا، يمكن تقسيم نتائج اختبار هندسي إلى فئات درجات (50–59، 60–69، 70–79…).

العمود الثاني: يُسجَّل فيه عدد القيم التي تقع داخل كل فئة من الفئات، ويُطلق على هذا العدد اسم تكرار الفئة.

بهذا التنظيم، يستطيع الباحث أن يرى بوضوح كيف تتوزع البيانات بين الفئات المختلفة، مما يسهّل تفسيرها وإجراء العمليات الإحصائية عليها.

مثال هندسي:

إذا أردنا دراسة توزيع أقطار المواسير المستخدمة في شبكة مياه صناعية، فبدلًا من عرض الأقطار لكل ماسورة على حدة (وهو ما قد يكون مربكًا وصعب التفسير)، نقوم بتقسيمها إلى فئات مثل: (2–2.9 بوصة، 3–3.9 بوصة، 4–4.9 بوصة...) ثم نسجّل عدد المواسير التي ينتمي قطرها لكل فئة. هذا التبويب يوضح بشكل مباشر أي الأقطار أكثر شيوعًا في الاستخدام، ويساعد المهندسين في اتخاذ قرارات تتعلق بالتصميم أو الشراء.

3.2 كيفية إعداد الجداول التكرارية:

تعتبر الخطوة الأولى في تبويب البيانات الخام وإعدادها في جداول التوزيع التكراري هي إعداد ما يُسمّى جدول تفريغ البيانات. هذا الجدول يتكوّن عادة من ثلاثة أعمدة، وكل عمود له وظيفة محددة كما يلي:

العمود الأول – الفئات:

يُدوَّن في هذا العمود الفئات التي تم تصنيف البيانات عليها، سواء كانت حسب النوع أو حسب القيمة العددية. وإذا كان من الممكن ترتيب البيانات، يُستحسن أن تكون الفئات مرتبة ترتيبًا تصاعديًا أو تنازليًا لتسهيل متابعة التوزيع وفهم الاتجاهات.

العمود الثاني – العلامات:

يُخصص هذا العمود لتسجيل القيم المشاهدة أو المشاهدات، وهي القيم التي تم الحصول عليها من أو عن مفردات المجتمع محل الدراسة أثناء جمع البيانات، عند تسجيل هذه القيم، يتم وضع علامة (/) أمام الفئة التي تنتمي إليها القيمة.

ولتسهيل عملية العدّ، يُستحسن أن تُجمّع العلامات بحيث تُكتب أربع علامات متجاورة (////)، أما العلامة الخامسة فتُكتب بشكل مائل أو في اتجاه معاكس للعلامات الأربع السابقة، على شكل (////). بهذه الطريقة نحصل على حزم من العلامات كل منها يحتوي على خمس علامات، مما يُسهّل العدّ السريع والدقيق للتكرارات.

العمود الثالث – التكرارات:

في هذا العمود يُســـجَّل عدد العلامات التابعة لكل فئة، هذا العدد يُمثّل تكرار الفئة، أي عدد المرات التي ظهرت فيها القيم المشـاهدة ضـمن هذه الفئة، بعد تسـجيل التكرارات لجميع الفئات، يجب التأكد من أن مجموع التكرارات يســـاوي العدد الكلي للقيم المشـــاهدة في البيانات، وذلك لضمان دقة جدول التفريغ وصحة النتائج.

بعد الانتهاء من إعداد جدول التفريغ، يمكننا اسـتخلاص جدول التوزيع التكراري النهائي، ويُعرف هذا الجدول بأنه يتكوّن من العمود الأول (الفئات) والعمود الثالث (التكرارات)، بينما يُحذف العمود الخاص بالعلامات.

وفي الخطوات القادمة، سـنقوم بدراسـة كيفية إعداد الجداول التكرارية لكل نوع من أنواع البيانات، سـواء كانت نوعية (وصـفية) أو كمية (عددية)، مع أمثلة توضـيحية تسـاعد على تطبيق هذه الطريقة عمليًا.

أولا: جدول التوزيع التكراري لبيانات وصفية (كيفية)

عند التعامل مع البيانات الوصـفية مثل مسـتوى التعليم أو الجنسـية، تُصـنّف البيانات إلى فئات تمثل كل صفة أو نوع، ثم يُعد جدول تفريغ البيانات لتسجيل كل مشاهدة باستخدام علامات (//) لتســـهيل العدّ، وبعد ذلك يُســـتخرج جدول التوزيع التكراري النهائي الذي يضـــم فقط الفئات والتكرارات، مع حذف عمود العلامات الوسيطة، ويمكن عرض مثال عملي يوضّح هذه الطريقة.

مثال (1.2)

البيانات الخام التالية توضـح نوع المواد المسـتخدمة في تصـنيع 20 قطعة ميكانيكية في ورشـة هندسية:

حديد	حديد	حديد	نحاس	حديد
ألومنيوم	ألومنيوم	نحاس	حديد	نحاس
حديد	نحاس	ألومنيوم	ألومنيوم	ألومنيوم
نحاس	حديد	حديد	نحاس	حديد

الحل

البيانات السابقة بيانات وصفية تمثل متغيرًا وصفياً وهو نوع المادة، والتي تنقسم إلى 3 أنواع: حديد، ألومنيوم، نحاس، يمكن وضعها في جدول تفريغ البيانات بحيث تمثل كل فئة نوع مادة، مع ترتيب الفئات بطريقة منطقية (مثلاً أبجدياً: ألومنيوم، حديد، نحاس).

بعد ذلك، نسجل أمام كل فئة العلامات التي تمثل المفردات التابعة لها، ويكون عدد العلامات هو تكرار كل فئة، ويجب أن يكون مجموع التكرارات مساوياً للعدد الكلي للقطع، كما يوضّح جدول (1.2) جدول تفريغ البيانات.

الجدول (1.2): جدول تفريغ البيانات

نوع المعدن (الفئة)	العلامات	عدد العلامات (التكرار)
ألومنيوم	𝍸	5
حديد	𝍸 ////	9
نحاس	𝍸 /	6
المجموع		20

ومن ثم يأتي جدول تفريغ البيانات على جدول التوزيع التكراري للتنقيح عن طريق الاكتتاب بالعمود الأول وحذف العمود الخاص بالعلامات والتوضيح فيما يلي:

الجدول (2.2): جدول التوزيع التكراري

نوع المعدن (الفئة)	عدد العلامات (التكرار)
ألومنيوم	5
حديد	9
نحاس	6
المجموع	20

وبعد وضع البيانات في جدول تفريغ البيانات أصبح الوضع واضحًا، حيث نفهم من الجدول أن القطع الموجودة في الورشة والمستخدمة في التصنيع هي 5 قطع ألومنيوم، و 9 قطع حديد، و6 قطع نحاس.

ثـانيا: جدول التوزيع التكراري لبيانات كمية (عددية)

1. جدول التوزيع التكراري للبيانات المنفصلة (المتقطعة)

إذا كانت البيانات كمية وتمثل متغيرًا منفصلاً، فهناك طريقتان لإعداد الجداول التكرارية:

أ. في حال كان عدد القيم المختلفة التي يأخذها المتغير محدودًا وصـغيرًا، نقوم بإنشـاء جدول تكراري يمثل كل فئة قيمة واحدة من القيم التي يأخذها المتغير، مع ترتيب هذه القيم ترتيبًا تصـــاعديًا. بعد ذلك، يمكن إعداد جدول تفريغ البيانات لتســـجيل المشاهدات ومنه استنتاج جدول التوزيع التكراري النهائي.

مثال (2.2)

في تجربة على جودة الصـبغ في تصـنيع 30 قطعة معدنية في ورشـة هندسـية، قمنا بفحص كل قطعة وعدد العيوب السطحية التي تظهر عليها، وكانت البيانات كما يلي:

0،1،0،2،1،0،1،3،0،1،2،1،0،1،2،1،0،1،2،1،0،1،0،1،2،1،3،0،1،2

الحل

المتغير محل الدراسة في هذا المثال هو عدد العيوب السطحية في كل قطعة معدنية، وهو متغير منفصل. القيم المختلفة التي يأخذها المتغير هي:

0: لا توجد عيوب على القطعة

1: عيب واحد على القطعة

2: عيبان على القطعة

3: ثلاثة عيوب على القطعة

حيث أن عدد القيم المختلفة يســاوي 4، وهو عدد محدود وصــغير، لذا نخصــص لكل فئة في جدول التوزيع التكراري قيمة واحدة فقط، يمكن بعد ذلك إعداد جدول تفريغ البيانات لتســـجيل العلامات (/) أمام كل فئة، ثم اســـتخراج التكرارات لكل فئة لإظهار التوزيع الكامل للعيوب على جميع القطع.

الجدول (3.2): جدول تفريغ البيانات

عدد العيوب في القطعة (الفئة)	العلامـــات	عدد العلامـــات (التكرار)
0	//// ⧸⧸⧸⧸	9
1	/// ⧸⧸⧸⧸ ⧸⧸⧸⧸	13
2	/ ⧸⧸⧸⧸	6
3	//	2
المجموع		30

ومن جدول تفريغ البيانات نحصل على جدول التوزيع التكراري:

الجدول (4.2): جدول التوزيع التكراري

عدد العيوب في القطعة (الفئة)	عدد العلامـــات (التكرار)
0	9
1	13
2	6
3	2
المجموع	30

ومن جدول التوزيع التكراري يتضـــح بســـهولة أن عدد القطع التي لم تظهر عليها أي عيوب ســطحية هو 9 قطع من أصــل 30 قطعة، وعدد القطع التي ظهرت عليها عيب واحد فقط هو 13 قطعة، بينما عدد القطع التي ظهرت عليها عيبان هو 6 قطع، وعدد القطع التي ظهرت عليها ثلاثة عيوب هو قطعتين؛ وبهذه الطريقة يصــبح من الســهل فهم توزيع العيوب الســطحية على القطع المعدنية بالكامل.

عند وضـــع البيانات الخام في جدول التوزيع التكراري، يصـــبح من الســـهل فهم المعلومات واســتخلاصــها، كما يمكن حســاب مقاييس إحصــائية مهمة مثل المتوســط والانحراف المعياري بسهولة.

ب. إذا كانت البيانات المنفصلة تمثل متغيراً يأخذ عدداً كبيراً من القيم، فيفضل أن تمثل كل فئة في الجدول مجموعة من القيم بدل أن تمثل قيمة واحدة فقط، وذلك لتجنب طول الجدول وتشـــتيت المعلومات، ولتكوين الجدول التكراري في هذه الحالة، نتبع الخطوات التالية:

خطوات إنشاء جداول التوزيعات التكرارية للبيانات المنفصلة:

1. تحديد المدى (Range) للبيانات:

أول خطوة هي معرفة المدى الذي تتوزع فيه القيم المشـــاهدة، ويتم ذلك بطرح أصـــغر قيمة في البيانات (الحد الأدنى) من أكبر قيمة (الحد الأعلى). ويرمز للمدى بالرمز R.

المدى = الحد الأعلى للبيانات – الحد الأدنى للبيانات

$$R = Max\ value - Mine\ value$$

2. تقسيم المدى إلى فئات:

بعد تحديد المدى، نقسـمه إلى عدد مناسـب من الفئات، عادةً بين 5 و15 فئة، ويُفضل أن تكون جميع الفئات متساوية الطول.

طول الفئة (Class Length) هو عدد القيم التي تمثلها كل فئة، ويحسب بالمعادلة:

طول الفئة = المدى ÷ عدد الفئات المطلوبة

$$L = \frac{R}{N}$$

حيث:

L هو طول الفئة.

N هو عدد الفئات المطلوبة.

إذا نتج عن القسمة كسر، يُستخدم العدد الصحيح الأعلى لتجنب ظهور حدود فئات كسورية، مما يجعل الجدول أكثر وضوحًا وسهولة في الاستخدام.

3. تحديد حدود كل فئة:

بعد حســاب طول الفئة، نحدد الحد الأدنى والحد الأعلى لكل فئة، بحيث يكون الحد الأدنى للفئة الأولى مساوياً لأصغر قيمة مشاهدة في البيانات الخام، بينما الحد الأعلى للفئة الأخيرة يكون مساوياً لأكبر قيمة مشاهدة في البيانات الخام، بهذه الطريقة، نضمن أن جميع القيم المشاهدة ستقع ضمن فئات الجدول التكراري.

4. إعداد جدول تفريغ البيانات:

نقوم بإنشاء جدول التفريغ الذي يحتوي على جميع الفئات، مع تسجيل العلامات لكل قيمة كما شرحنا سابقاً.

بعد ذلك، يمكننا اســـتخراج جدول التوزيع التكراري من جدول التفريغ عن طريق الاكتفاء بالعمود الأول (الفئات) والعمود الأخير (التكرار)، مع حذف العمود الخاص بالعلامات، ليصـــبح الجدول جاهزاً للتحليل الإحصائي.

مثـال (3.2)

البيانات التالية تمثل عدد المنتجات المعيبة التي تم تســجيلها يومياً في أحد خطوط الإنتاج خلال 30 يومًا:

22, 25, 28, 30, 20, 24, 35, 40, 38, 30, 27, 26, 29, 33, 31, 25, 23, 37, 42, 45, 41, 39, 28, 32, 36, 44, 46, 50, 48, 30

المطلوب: عرض هذه البيانات في جدول توزيع تكراري يتكون من 6 فئات متساوية الطول.

الحل

لتكوين جدول التوزيع التكراري لهذه البيانات، نتبع الخطوات التالية:

- تحديد الحد الأعلى والحد الأدنى للبيانات وحساب المدى:

الحد الأدنى = 20 منتج ، الحد الأعلى = 50 منتج

المدى = الحد الأعلى للبيانات – الحد الأدنى للبيانات

$$R = 50 - 20 = 30$$

- تحديد عدد الفئات وطول كل فئة:

نريد تقسيم البيانات إلى 6 فئات متساوية الطول.

طول الفئة = المدى ÷ عدد الفئات المطلوبة

$$L = \frac{R}{N}$$

$$L = \frac{30}{6} = 5$$

تحديد حدود الفئات:

الحد الأدنى للفئة الأولى = 20 منتج، وبم إن طول الفئة يكون مســـاوياً لـــــــ 5 بيانات، فهي كالـتالي: 20،21،22،23،24. فيكون الحد الأعلى للفئة الأولى = 24 .

وبنفس الطريقة نحدد باقي الفئات:

الفئة الأولى: 20–24

الفئة الثانية: 25–29

الفئة الثالثة: 30–34

الفئة الرابعة: 35–39

الفئة الخامسة: 40–44

الفئة السادسة: 45–50

ملاحظة: الفئة الأخيرة تُغلق عادةً لتشــمل أكبر قيمة في البيانات، لذلك قد تبدو أطول عدديًا من باقي الفئات.

الجدول (5.2): جدول تفريغ البيانات لعدد المنتجات المعيبة.

(الفئة)	العلامـــات	عدد العلامـــات (التكرار)
24-20	////	4
29-25	//𝍸	7
34-30	/𝍸	6
39-35	𝍸	5
44-40	////	4
50-45	////	4
المجموع		30

ومن جدول تفريغ البيانات وبحذف العمود الثاني نحصـــل على جدول التوزيع التكراري الخاص بعدد المنتجات المعيبة يومياً.

الجدول (6.2): جدول التوزيع التكراري للمنتجات المعيبة يومياً.

(الفئة)	عدد العلامـــات (التكرار)
24-20	4
29-25	7
34-30	6
39-35	5
44-40	4
50-45	4
المجموع	30

الفئة الأولى: 20 – 24 ← تمثل القيم 20،21،22،23،24

الفئة الثانية: 25 – 29 ← تمثل القيم 25،26،27،28،29

الفئة الثالثة: 30 – 34 ← تمثل القيم 30،31،32،33،34

الفئة الرابعة: 35 – 39 ← تمثل القيم 35،36،37،38،39

الفئة الخامسة: 40 – 44 ← تمثل القيم 40،41،42،43،44

الفئة السادسة: 45 – 50 ← تمثل القيم 45،46،47،48،49،50

(الفئة)	عدد العلامـــات (التكرار)
24-20	4
29-25	7
34-30	6
39-35	5
44-40	4
50-45	4
المجموع	30

وبسهولة نفهم من هذا الجدول أنه في 4 أيـــــــام كان عدد المنتجات المعيبة بين 20 إلى 24 قطعة، وفي 7 أيـــــــام كان عدد المنتجات المعيبة بين 25 إلى 29 قطعة، وهكذا ... وهذه المعلومات بالطبع لم تكن واضـــحة عندما كانت البيانات خاماً وغير مصـــنفة في جدول، وإنما أصبحت مفهومة بعد تبويبها في جدول التوزيع التكراري.

مثـال (4.2)

عدد الأجهزة المصنعة يوميًا في ورشة هندسية خلال 30 يومًا هي كالتـالي:

3, 7, 12, 15, 9, 18, 20, 22, 10, 8

14, 16, 17, 19, 5, 6, 13, 21, 24, 11

7, 9, 15, 12, 8, 10, 16, 18, 23, 14

المطلوب: عرض هذه البيانات في جدول توزيع تكراري.

الحل

لتكوين جدول التوزيع التكراري لهذه البيانات، نتبع الخطوات التالية:

- تحديد الحد الأعلى والحد الأدنى للبيانات وحساب المدى:

الحد الأدنى = 3 أجهزة ، الحد الأعلى = 24 جهاز.

المدى = الحد الأعلى للبيانات – الحد الأدنى للبيانات

$$R = 24 - 3 = 21$$

- تحديد عدد الفئات وطول كل فئة:

نفترض تقسيم البيانات إلى 5 فئات متساوية الطول.

طول الفئة = المدى ÷ عدد الفئات المطلوبة

$$L = \frac{R}{N}$$

$$L = \frac{21}{5} = 4.2 = 5$$

تحديد حدود الفئات:

الحد الأدنى للفئة الأولى = 3 أجهزة، وبم إن طول الفئة يكون مساوياً لـ5 بيانات، فهي كالـتالي: 3،4،5،6،7 فيكون الحد الأعلى للفئة الأولى = 7 .

وبنفس الطريقة نحدد باقي الفئات:

الفئة الأولى: 3–7

الفئة الثانية: 8–12

الفئة الثالثة: 13–17

الفئة الرابعة: 18–22

الفئة الخامسة: 23–27

ملاحظة: الفئة الأخيرة تُغلق عادةً لتشــمل أكبر قيمة في البيانات، لذلك قد تبدو أطول عدديًا من باقي الفئات.

الجدول (7.2): جدول تفريغ البيانات لعدد الأجهزة المنتجة.

(الفئة)	العلامـــات	عدد العلامـــات (التكرار)
7-3	𝍸	5
12-8	//// 𝍸	9
17-13	/// 𝍸	8
22-18	/ 𝍸	6
27-23	//	2
المجموع		30

ومن جدول تفريغ البيانات وبحذف العمود الثاني نحصـــل على جدول التوزيع التكراري الخاص بعدد الأجهزة المنتجة يومياً.

الجدول (8.2): جدول التوزيع التكراري للأجهزة المنتجة يومياً.

(الفئة)	عدد العلامـــات (التكرار)
7-3	5
12-8	9
17-13	8
22-18	6
27-23	2
المجموع	30

الفئة الأولى: 3 – 7 → تمثل القيم 7،6،5،4،3

الفئة الثانية: 8 – 12 → تمثل القيم 12،11،10،9،8

الفئة الثالثة: 13 – 17 → تمثل القيم 17،16،15،14،13

الفئة الرابعة: 18 – 22 → تمثل القيم 22،21،20،19،18

الفئة الخامسة: 23 – 27 → تمثل القيم 27،26،25،24،23

(الفئة)	عدد العلامـــات (التكرار)
3-7	5
8-12	9
13-17	8
18-22	6
23-27	2
المجموع	30

وبسهولة نفهم من هذا الجدول كذلك أنه في 5 أيـــــــام كان عدد الأجهزة المنتجة بين 3 إلى 7 أجهزة، وفي 9 أيام كان عدد الأجهزة المنتجة بين 8 إلى 12 جهاز، وهكذا...

مركز الفئة:

مركز الفئة هو القيمة الواقعة في منتصف كل فئة، ويُحسب باستخدام الصيغة التالية:

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

ونطبق هذه الصـيغة سـواء كانت البيانات منفصـلة أو متصـلة، فمثلاً، إذا قسـمنا عدد الأجهزة المنتجة يوميًا إلى فئات متساوية الطول مثل:

الفئة الأولى: 30 – 33

الفئة الثانية: 34 – 37

الفئة الثالثة: 38 – 41

الفئة الرابعة: 42 – 45

الفئة الخامسة: 46 – 50

فإن مركز الفئة الأولى يُحسب كالتالي:

$$\text{مركز الفئة} = \frac{30+33}{2} = 61.5$$

وهكذا بالنسبة لبقية الفئات، حيث يعطي مركز الفئة نقطة وسطية تمثل كل فئة لتسهيل التحليل والرسم البياني.

ويمثل مركز الفئة القيمة النظرية التي تمثل كل القيم الموجودة داخل الفئة، أي أنه ليس بالضرورة أن يأخذها المتغير محل الدراسة فعليًا، ففي العمليات الحسابية لبعض المقاييس الإحصائية، يُستخدم مركز الفئة كممثل عن كل القيم التابعة للفئة.

في المثال الخاص بعدد الأجهزة المنتجة يوميًا في المصنع، الفئة الأخيرة تتراوح بين 46 و 50 جهاز، وتكرارها يساوي 3، أي أن هناك ثلاثة أيام كان فيهما عدد القطع المنتجة ضمن هذه الفئة، لكن الجدول لا يوضح عدد القطع في كل يوم على حدة.

بافتراض توزيع الإنتاج بشكل متساوٍ داخل الفئة، يُعتبر عدد القطع في كل يوم مساويًا لمركز الفئة، وهو هنا 48 قطعة. وبالتالي، يكون مجموع القطع المنتجة في هذه الفئة:

$$48 \times 3 = 144 \text{ قطعة}$$

بنفس الطريقة يمكن حساب مجموع القطع لجميع الفئات بضرب مركز الفئة لكل فئة في تكرارها، ثم جمع هذه المجاميع للحصول على المجموع الكلي لإجمالي القطع المنتجة خلال الفترة.

تُستخدم هذه المجاميع لاحقًا في حساب العديد من المقاييس الإحصائية، كما سنوضح في الفصول القادمة من الكتاب.

الجداول غير المنتظمة:

يُسمّى جدول التوزيع التكراري غير منتظم إذا كانت الفئات غير متساوية الطول، بينما إذا كانت الفئات متساوية الطول فيُسمّى جدولاً منتظماً، عادةً ما يتم تفضيل الجداول المنتظمة لأنها أسهل في الحسابات الإحصائية، مثل حساب الوسط أو الانحراف المعياري، ولكن في بعض الحالات الصناعية، تكون البيانات مركّزة في جزء ضيق من النطاق (مثلاً: أبعاد منتج تتكرر ضمن حدود محدودة) بينما تكون متفرقة أو نادرة في الأجزاء الأخرى.

إذا استعملنا فئات متساوية الطول هنا، سنجد كثيراً من الفئات فارغة تقريباً، مما يقلل من وضوح الصورة؛ لذلك يُفضل استخدام فئات قصيرة في المناطق التي تتركز فيها القيم (لإظهار الدقة في تلك المنطقة)، وفئات أوسع في المناطق المبعثرة أو النادرة.

مثال: عند تحليل توزيع أقطار مواسير معدنية منتجة، قد نجد أن أغلب المواسير قطرها بين 38–42 مم (قيمة مقاربة للمواصفات القياسية)، بينما توجد بعض المواسير الشاذة بأقطار أكبر أو أصغر، ففي هذه الحالة، نُنشئ جدولاً غير منتظم:

الفئات بين 38–42 مم تكون ضيقة (لكشف التفاصيل بدقة)، والفئات خارج هذا النطاق (مثلاً: 30–35 مم أو 45–50 مم) تكون أوسع لأنها قليلة التكرار.

مثال آخر:

29 – 0	74 – 65
تقدير ضعيف جداً	تقدير جيد
49 – 30	84 – 75
تقدير ضعيف	تقدير جيد جداً
64 – 50	100 – 85
تقدير مقبول	تقدير ممتاز

الجداول المفتوحة:

يُسمّى جدول التوزيع التكراري مفتوحاً إذا لم يُحدد الحد الأدنى للفئة الأولى أو الحد الأعلى للفئة الأخيرة (أو كلاهما)، ويكون السبب في ذلك عادةً:

- عدم توافر بيانات دقيقة عن البداية أو النهاية.
- أو لوجود عدد قليل من القيم المتطرفة البعيدة جداً عن باقي البيانات.

الأنواع:

1. مفتوح من الطرف الأدنى: إذا لم يُحدد الحد الأدنى للفئة الأولى (مثلاً: أقل من 30).
2. مفتوح من الطرف الأعلى: إذا لم يُحدد الحد الأعلى للفئة الأخيرة (مثلاً: 50 فأكثر).
3. مفتوح من الطرفين: إذا كانت الفئة الأولى والأخيرة غير محددتي الحدود معاً.

مثال: عند تحليل أعمار الآلات في مصـــنع، قد نجد الفئة الأولى مكتوبة: "أقل من 3 ســـنوات" (مفتوحة من الأسفل). والفئة الأخيرة: "15 سنة فأكثر" (مفتوحة من الأعلى).

أقل من 5	5 – 1	أقل من 5
10 – 6	10 – 6	10 – 6
15 – 11	15 – 11	15 – 11
20 – 16	20 – 16	20 – 16
25 – 21	21 أو أكثر	21 أو أكثر
مفتوح من طرفه الأدنى	مفتوح من طرفه الأعلى	مفتوح من الطرفين

مثــال (4.2)

في مصنع لإنتاج مسامير معدنية، تم قياس طول كل مسمار (بالمليمتر) لعينة من 40 مسمارًا. نظرًا لوجود مسامير قصيرة جدًا أو طويلة جدًا خارج النطاق المعتاد، فيطلب منك تسجيل البيانات التالية في جدول تكراري مفتوح:

18, 22, 35, 41, 28, 37, 19, 50, 44, 23, 36, 30, 32, 45, 48, 29, 21, 40,
55, 34, 31, 46, 38, 42, 20, 47, 33, 39, 43, 51, 26, 24, 49, 36, 28, 37,
52, 27, 35, 44

الحـل

الجدول (9.2): جدول التوزيع التكراري المفتوح لأطـوال المسامير المصنعة.

الفئة (ملم)	التكرار
أقل من 20	2
29 – 20	10
39 – 30	13
49 – 40	11
50 وأكبر	4

4.2 جداول التوزيعات التكرارية المتجمعة:

التوزيعات التكرارية العادية تُبيّن عدد المفردات في كل فئة على حدة، لكن في بعض الحالات نحتاج صورة أشمل، مثل:

- معرفة عدد المفردات التي قيمتها أقل من قيمة معينة.
- معرفة عدد المفردات التي قيمتها أكبر من أو تساوي قيمة معينة.

على سبيل المثال: في بيانات رواتب الموظفين، قد يهمنا أن نعرف:

- كم عدد الموظفين الذين تقل رواتبهم عن 50 دينارًا أسبوعياً؟
- كم عدد الموظفين الذين تبلغ رواتبهم 50 دينارًا فأكثر؟

ولعرض هذا النوع من المعلومات، نستخدم التوزيع التكراري المتجمع، وهو شكل آخر للجداول الإحصائية.

ويوجد منه نوعان أساسيان:

1. التوزيع التكراري المتجمع الصاعد (Less than).
2. التوزيع التكراري المتجمع الهابط (More than).

افترض أن لدينا بيانات عن عدد الساعات اللازمة لإنهاء تصنيع قطعة معدنية في خط إنتاج، وكانت القيم تتراوح بين 30 و 60 دقيقة، بعد تنظيمها في جدول توزيع تكراري عادي، قد نرغب في معرفة:

- كم عدد القطع التي استغرق تصنيعها أقل من 45 دقيقة؟
- كم عدد القطع التي استغرق تصنيعها 45 دقيقة فأكثر؟

هنا نستخدم التوزيع التكراري المتجمع:

- التوزيع المتجمع الصاعد (Less than): يوضح عدد القطع التي لم تتجاوز زمناً معيناً في التصنيع (مثلاً أقل من 40 دقيقة، أقل من 50 دقيقة...).
- التوزيع المتجمع الهابط (More than): يوضح عدد القطع التي استغرقت زمناً أكبر من أو يساوي قيمة معينة (مثلاً 40 دقيقة فأكثر، 50 دقيقة فأكثر...).

1. التوزيع التكراري المتجمع الصاعد (Less than Cumulative Frequency):

يُظهر التوزيع التكراري المتجمع الصاعد عدد الوحدات أو العناصر التي تقع قيمتها أقل من الحد الأعلى لكل فئة في جدول التوزيع التكراري العادي.

على سبيل المثال، افترض أن لدينا بيانات عن عدد دقائق التصنيع لكل قطعة معدنية في خط إنتاج، إذا كانت الفئة الأولى تمثل القطع التي استغرق تصنيعها من 30 إلى أقل من 40 دقيقة، نجد أن عدد القطع التي استغرق تصنيعها أقل من 30 دقيقة يساوي صفرًا، أي لا توجد أي قطعة أسرع من 30 دقيقة. أما عدد القطع التي استغرق تصنيعها أقل من 40 دقيقة فيساوي عدد القطع التابعة للفئة الأولى، ولنفترض أنه قطعتان.

عند النظر إلى الفئة الثانية (40–50 دقيقة)، يكون عدد القطع التي استغرق تصنيعها أقل من 50 دقيقة مساويًا مجموع القطع في الفئة الأولى والثانية، أي 2 + 3 = 5 قطع. وهكذا نتابع جمع التكرارات للفئات التالية حتى نصل إلى مجموع جميع القطع المصنعة، وهو ما يُعطي التوزيع التكراري المتجمع الصاعد.

يُسمى هذا التوزيع "صاعدًا" لأن التكرار يزيد مع كل فئة، ويبدأ دائمًا من الصفر وينتهي بالعدد الكلي للمفردات أو العناصر محل الدراسة، هذه الطريقة تساعد المهندس أو مدير الإنتاج على معرفة كمية القطع التي اكتملت ضمن أطر زمنية معينة بسرعة وسهولة.

مثال (5.2)

كون جدول التوزيع التكراري المتجمع الصاعد الخاص بالمرتبات اليومية لثلاثين موظف بشركة معينة، والتي كانت كالتالي:

32, 39, 30, 33, 42, 40, 37, 48, 34, 39,

41, 35, 33, 46, 37, 38, 41, 39, 50, 38,

44, 36, 38, 39, 42, 45, 33, 36, 41, 41

الحـل

الجدول (10.2): جدول التوزيع التكراري لمرتبات الموظفين بالشركة.

(الفئة)	عدد العلامـــات (التكرار)
33-30	5
37-34	6
41-38	12
45-42	4
50-46	3
المجموع	30

نبدأ من الفئة الأولى في الجدول التكراري العادي، ونجمع التكرارات بشكل تصـاعدي بحيث يمثل كل صــف مجموع عدد المفردات التي قيمتها أقل من الحد الأعلى للفئة الســابقة. نســتمر بهذه الطريقة لجميع الفئات حتى تصل التكرارات في الصف الأخير إلى العدد الكلي لجميع المفردات.

الجدول (11.2): جدول التوزيع التكراري المتجمع الصاعد لمرتبات الموظفين بالشركة.

(الفئة)	التكرار المتجمع الصاعد
أقل من 30	0
33 فـأقل	5
37 فـأقل	11
41 فـأقل	23
45 فـأقل	27
50 فـأقل	30

2. التوزيع التكراري المتجمع الهابط (More than Cumulative Frequency):

يُظهر التوزيع التكراري المتجمع الهابط عدد الوحدات أو العناصـر التي تسـاوي أو تزيد عن الحد الأدنى لكل فئة في جدول التوزيع التكراري العادي.

على سـبيل المثال، افترض أن لدينا بيانات عن عدد دقائق التصــنيع لكل قطعة معدنية في خط إنتاج، إذا كانت الفئة الأولى تمثل القطع التي استغرق تصـنيعها من 30 إلى أقل من 40 دقيقة،

فإن عدد القطع التي اســـتغرق تصـــنيعها 30 دقيقة أو أكثر يســـاوي مجموع جميع القطع في الجدول، أي العدد الكلي للمفردات، لأنه يشمل كل القطع.

عند النظر إلى الفئة الثانية (40–50 دقيقة)، يكون عدد القطع التي استغرق تصنيعها 40 دقيقة أو أكثر مساويًا مجموع التكرارات للفئة الثانية والفئات الأعلى منها، أي نطرح أي قطع كانت أقل من 40 دقيقة، وهكذا نستمر لكل فئة لاحقة بحيث يشمل كل تكرار الفئة نفسها والفئات الأعلى، حتى نصل إلى الفئة الأخيرة، حيث يمثل عدد القطع التي استغرق تصنيعها الحد الأعلى أو أكثر فقط.

ويُسمى هذا التوزيع "هابطًا" لأن التكرار ينخفض تدريجيًا مع كل فئة لاحقة، ويبدأ دائمًا من العدد الكلي للمفردات وينتهي بالصفر، وهذه الطريقة تساعد المهندس أو مدير الإنتاج على معرفة عدد القطع التي استغرقت زمن تصنيع معين أو أكثر، لتقييم أداء خط الإنتاج بسرعة وسهولة.

مثـــال (6.2)

كون جدول التوزيع التكراري المتجمع الهابط للمثـال 5.2

الحـــل

من الجدول 10.2

(الفئة)	عدد العلامـــات (التكرار)
33–30	5
37–34	6
41–38	12
45–42	4
50–46	3
المجموع	30

لتحويل جدول التوزيع التكراري العادي إلى التكراري المتجمع الهابط، نبدأ من العدد الكلي للمفردات ثم نطرح تكرارات الفئات واحدة تلو الأخرى من الأعلى إلى الأدنى. أي: لكل فئة، يكون التكرار المتجمع الهابط مســـاويًا لإجمالي عدد المفردات ناقص مجموع تكرارات كل الفئات التي تسبقها (أعلى منها). بهذه الطريقة، يبدأ التكرار من العدد الكلي للمفردات وينخفض تدريجيًا حتى يصل إلى صفر عند الفئة الأخيرة.

الجدول (12.2): جدول التوزيع التكراري المتجمع الهابط لمرتبات الموظفين بالشركة.

(الفئة)	التكرار المتجمع الهابط
30 فأكثر	30
34 فأكثر	25
38 فأكثر	19
42 فأكثر	7
46 فأكثر	3
أكثر من 50	0

يمكننا دمج التوزيع التكراري العادي والتكراري المتجمع الصـــاعد والهابط في جدول واحد، بحيث يحتوي كل صـــف على: الفئة، التكرار العادي، التكرار المتجمع الصـــاعد، والتكرار المتجمع الهابط، والجدول التـالي يوضح ذلك.

الجدول (13.2): جدول التوزيع التكراري العادي والمتجمع الصاعد والهابط لمرتبات الموظفين بالشركة معاً.

(الفئة)	(التكرار)	(الفئة)	التكرار المتجمع الصاعد	(الفئة)	التكرار المتجمع الهابط
33-30	5	أقل من 30	0	30 فأكثر	30
37-34	6	33 فأقل	5	34 فأكثر	25
41-38	12	37 فأقل	11	38 فأكثر	19
45-42	4	41 فأقل	23	42 فأكثر	7
50-46	3	45 فأقل	27	46 فأكثر	3
المجموع	30	50 فأقل	30	أكثر من 50	0

5.2 جداول التوزيعات التكرارية النسبية:

يمكن حساب التكرار النسبي في جميع أنواع الجداول التكرارية، سواء كانت خاصة بالبيانات النوعية (الوصفية) أو الكمية، وسواء كانت البيانات منفصلة أو متجمعة، وبوجه عام، يُحسب التكرار النسبي لأي فئة

من خلال قسـمة تكرار تلك الفئة على العدد الكلي للمشـاهدات (مجموع التكرارات)، وذلك وفق العلاقة التالية:

$$\text{التكرار النسبي} = \frac{\text{تكرار الفئة}}{\text{العدد الكلي للمشاهدات (التكرارات)}}$$

ويمكن التعبير عن التكرار النسبي بطريقتين:

- على شكل كسر عشري (مثل: 0.25)
- أو على شكل نسبة مئوية (مثل: 25%)

مثـال (7.2)

أوجد الجدول التكراري النسبي لجدول التوزيع التكراري في المثـال (1.2) والذي يمثل بيانات لنوع المواد المستخدمة في تصنيع 20 قطعة ميكانيكية في ورشة هندسية.

الحـل

من المثـال 1.2 الخاص بالمواد المستخدمة في التصنيع نجد أن:

التكرار النسبي للفئة الأولى $\frac{5}{20} = 0.25$، أي أن 25% من المواد المستخدمة في التصنيع كانت ألومنيوم، وبنفس الطريقة يمكن حساب التكرار النسبي لجميع الفئات بالجدول.

الجدول (14.2): جدول التوزيع التكراري النسبي للمواد المستخدمة في التصنيع.

نوع المعـدن (الفئة)	التكرار النسبي
ألومنيوم	0.25
حديد	0.45
نحاس	0.3
المجموع	1

ومن الجدول (14.2) نلاحظ أن 25% من المواد المسـتخدمة كانت ألومنيوم، و 45% لمعدن الحديد، و30% كانت لمعدن النحاس. ويُلاحظ هنا أنه تم التعبير عن التكرارات النسـبية بصـيغة

النسـب المئوية بدلًا من الكسـور العشـرية، وفي هذه الحالة يكون مجموع النسـب المئوية لجميع الفئات مساويًا لـ 100%.

مثــال (8.2)

أوجد الجدول التكراري النسبي لجدول التوزيع التكراري في المثـــــــــال (3.2) والذي يمثل عدد المنتجات المعيبة لأحد خطوط الإنتاج في مصنع.

الحـل

من المثـال 2.2 الخاص بعدد المنتجات المعيبة في أحد خطوط الانتاج نجد أن:

التكرار النســـبي للفئـة الأولى $\frac{4}{30} = 0.13$، أي أن 13% من الفترة الزمنيـة كـان عدد القطع المنتجة معيبة يتراوح بين 20 إلى 24 قطعة، وبنفس الطريقة يمكن حساب التكرار النسبي لجميع الفئات بالجدول.

الجدول (15.2): جدول التوزيع التكراري النسبي للقطع المعيبة خلال 30 يوماً.

(الفئة)	التكرار النسبي
24−20	0.13
29−25	0.23
34−30	0.2
39−35	0.16
44−40	0.13
50−45	0.13
المجموع	1

مثــال (9.2)

كوّن جدول التكرار المتجمع الصـاعد النسبي للمثال 5.2، وجدول التكرار المتجمع الهابط النسبي للمثال 6.2.

الحـل

من الجدول 11.2 في المثال 5.2 نجد أن:

التكرار المتجمع الصـــاعد لعدد الموظفين الذين راتبهم أقل من 33 هو 5، وبالتالي فإن: التكرار المتجمع الصـــاعد النســـبي لعدد الموظفين الذين راتبهم أقل من 33 هو $\frac{5}{30} = 0.16$، وبنفس الطريقة يمكن حساب التكرار النسبي لجميع الفئات بالجدول.

الجدول (16.2): جدول التوزيع التكراري المتجمع الصاعد النسبي لمرتبات الموظفين.

(الفئة)	التكرار المتجمع الصاعد النسبي
أقل من 30	0
33 فـأقل	0.16
37 فـأقل	0.36
41 فـأقل	0.76
45 فـأقل	0.9
50 فـأقل	1

من الجدول 12.2 في المثال 6.2 نجد أن:

التكرار المتجمع الهابط لعدد الموظفين الذين راتبهم أعلى من 30 هو 30، وبالتالي فإن: التكرار المتجمع الهابط النســبي لعدد الموظفين الذين راتبهم أعلى من 30 هو $\frac{30}{30} = 1$، وبنفس الطريقة يمكن حساب التكرار النسبي لجميع الفئات بالجدول.

الجدول (17.2): جدول التوزيع التكراري المتجمع الهابط النسبي لمرتبات الموظفين.

(الفئة)	التكرار المتجمع الهابط النسبي
30 فأكثـر	1
34 فأكثـر	0.83
38 فأكثـر	0.63
42 فأكثـر	0.23
46 فأكثـر	0.1
أكثـر من 50	0

العرض البيـاني للبيانات الإحصائية

في الجزء السـابق من هذا الفصـل اسـتعرضـنا كيفية عرض عدد كبير من البيانات في صـورة جداول تكرارية، مما يُسـهل تبسـيط هذه البيانات وفهمها بشـكل أفضـل، ومع ذلك قد لا تكفي الجداول التكرارية أحيانًا لتوضـيح المعلومات بشـكل كامل، خاصـةً أن بعض الأشـخاص يجدون صعوبة في استيعاب مدلولات الأرقام المعروضة عليهم في الجداول.

ولذلك، يلجأ الإحصـائي إلى عرض البيانات بطريقة بصـرية باسـتخدام الرسـومات البيانية، التي تساعد على توضيح التوزيع واتجاهات البيانات بسرعة وسهولة، وتوجد عدة طرق رئيسية لتمثيل البيانات إحصائيًا بطريقة بيانية، ومن أهمها:

- **القطاعات الدائرية (Pie Chart):** تُسـتخدم لعرض النسـب المئوية لكل فئة من البيانات.
- **الأعمدة البيانية (Bar Chart):** تُظهر الفروقات بين الفئات بشـكل واضـح من خلال ارتفاع الأعمدة.
- **المدرج التكراري (Histogram):** يُسـتخدم للبيانات الكمية المتصـلة لتوضـيح توزيع التكرارات.
- **المضـلع التكراري (Frequency Polygon):** يربط منتصـف كل فئة بنقاط متجاورة لإظهار شكل التوزيع.
- **المنحنى التكراري (Ogive / Cumulative Curve):** يُظهر التوزيع التراكمي للبيانات.
- **المنحنى التكراري المتجمع الصـاعد والهابط (Ascending & Descending Cumulative Curves):** يُسـتخدم لمقارنة عدد المفردات الأقل أو الأعلى من قيم معينة، ويعطي صورة أوضح عن تراكم البيانات على طول النطاق.
- **6.2 القطاعات الدائرية (Pie Chart):**

في هذه الطريقة، نُمثّل العدد الكلي للمفردات (أي مجموع التكرارات) بدائرة كاملة، ثم نُمثل عدد المفردات التابعة لكل فئة بواسطة قطاع من الدائرة، بحيث تكون مساحة كل قطاع متناسبة مع تكرار الفئة التي يمثّلها، وبما أن مسـاحة القطاع تتناسـب مع زاويته المركزية في الدائرة، فإن

حســـاب زاوية كل قطاع يكون عن طريق تقســـيم الدائرة كاملة (360°) إلى زوايا تتناســـب مع تكرارات الفئات.

لذلك زاوية القطاع لكل فئة تُحسب كالآتي:

$$\text{زاوية القطاع} = \frac{\text{تكرار الفئة}}{\text{مجموع التكرارات}} \times 360$$

بهذه الطريقة، يعطي المخطط الدائري صـــورة واضـــحة عن توزيع المفردات بين الفئات المختلفة ويسهّل المقارنة بين أحجام الفئات بصريًا.

يعتمد المخطط الدائري على تمثيل كل فئة بقطاع من الدائرة تكون مســـاحته متناســـبة مع تكرار الفئة أو أهميتها. يمكن تظليل كل قطاع أو اسـتخدام ألوان مختلفة لتمييز الفئات بصـريًا، ويُكتب اسم الفئة أو قيمتها بجوار كل قطاع لتوضيح ما يمثله.

تُعد هذه الطريقة مناســـبة جدًا للبيانات الوصـــفية، حيث تســـهّل على القارئ فهم نســـب الفئات المختلفة مقارنة بالعدد الكلي للمفردات.

مثـال (10.2)

الجدول التالي يوضـح عدد القطع المعدنية حسـب نوع المعالجة السـطحية في مصـنع ميكانيكي خلال أسبوع:

نوع المعالجة	عدد القطع
صقل (Polishing)	12
دهان (Painting)	18
طلاء كهربائي (Electroplating)	6
تبريد سريع (Quenching)	4

المطلـوب: تمثيل هذه البيانات في مخطط دائري.

الحـل

في هذا المثـــال الفئـــات عبارة عن نوع المعالجة والتكرار هو عدد القطع المعالجة، ولتمثيل هذه البيانات بقطاعات دائرية نقوم بحساب زاوية كل فئة "كل قطاع":

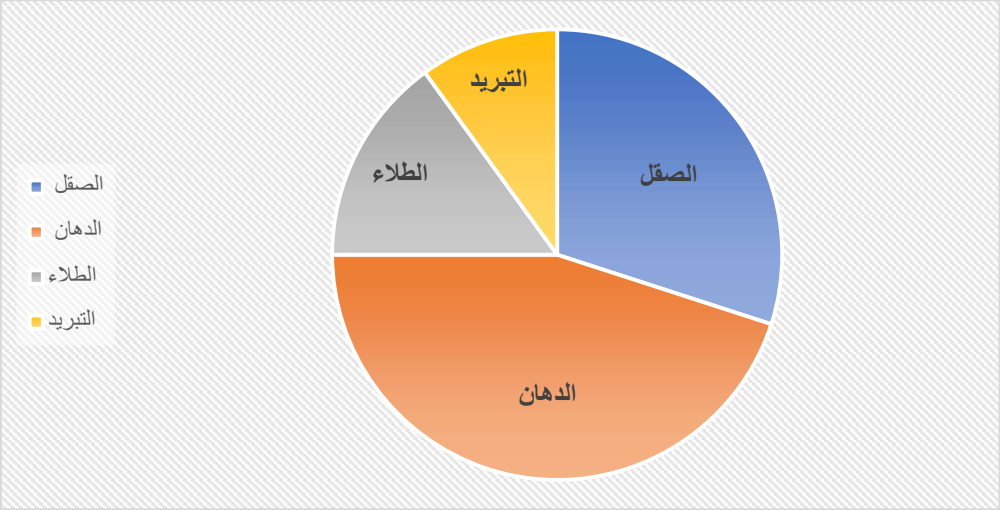
زاوية القطاع الذي يمثل المعالجة بالصقل: $108 = 360x\frac{12}{40}$

زاوية القطاع الذي يمثل المعالجة بالدهان: $162 = 360x\frac{18}{40}$

زاوية القطاع الذي يمثل المعالجة بالطلاء الكهربائي: $54 = 360x\frac{6}{40}$

زاوية القطاع الذي يمثل المعالجة بالتبريد السريع: $36 = 360x\frac{4}{40}$

ثم نقوم برسم دائرة بنصف قطر مناسب، ونقسم زاويتها المركزية التي تساوي 360 إلى الزوايا السابقة (36،54،162،108). والموضحة في الشكل (1.2)



الشكل (1.2): يوضح القطاعات الدائرية لنوع معالجة المعدن.

مثـال (11.2)

الجدول التالي يوضـــح كمية الانتاج شـــهرياً بالكرتونة لأحد خطوط الإنتاج داخل مصـــنع مواد غذائية خلال أشهر السنة، المطلوب عرض هذه البيانات باستخدام القطاع الدائري.

الشهر	عدد الكرتونات	الشهر	عدد الكرتونات
يناير	1100	يوليو	1235
فبراير	955	أغسطس	1150
مارس	1021	سبتمبر	982
أبريل	1005	أكتوبر	1204
مايو	864	نوفمبر	1200
يونيو	1420	ديسمبر	1320

الحـل

لكي يتم حساب الزوايا يجب حساب اجمـــــــالي التكرارات وهو ما يساوي اجمالي الإنتاج خلال السنة كاملاً: 13456 كرتونة.

زاوية القطاع الذي يمثل انتاج شهر يناير: $29.42 = 360x\frac{1100}{13456}$

زاوية القطاع الذي يمثل انتاج شهر فبراير: $25.54 = 360x\frac{955}{13456}$

زاوية القطاع الذي يمثل انتاج شهر مارس: $27.31 = 360x\frac{1021}{13456}$

زاوية القطاع الذي يمثل انتاج شهر أبريل: $26.88 = 360x\frac{1005}{13456}$

زاوية القطاع الذي يمثل انتاج شهر مايو: $23.12 = 360x\frac{864}{13456}$

زاوية القطاع الذي يمثل انتاج شهر يونيو: $37.99 = 360x\frac{1420}{13456}$

زاوية القطاع الذي يمثل انتاج شهر يوليو: $33.04 = 360x\frac{1235}{13456}$

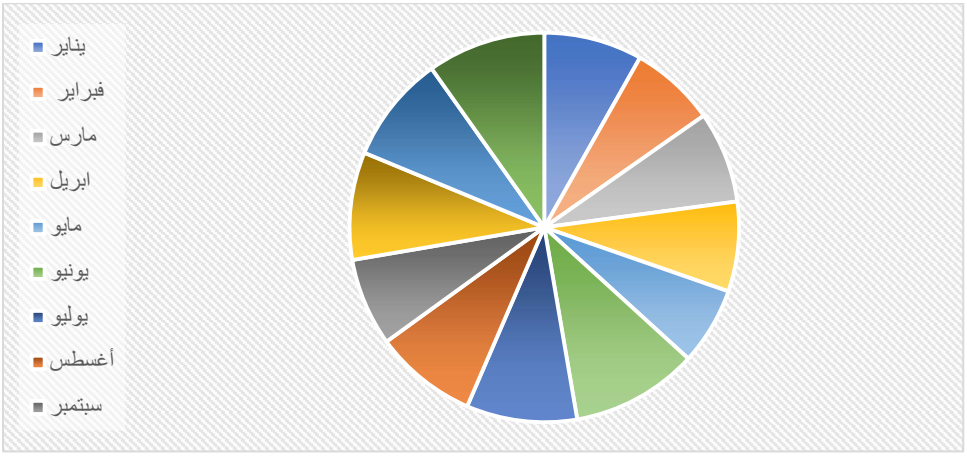
زاوية القطاع الذي يمثل انتاج شهر أغسطس: $30.76 = 360x\frac{1150}{13456}$

زاوية القطاع الذي يمثل انتاج شهر سبتمبر: $26.27 = 360x\frac{982}{13456}$

زاوية القطاع الذي يمثل انتاج شهر أكتوبر: $32.21 = 360x\frac{1204}{13456}$

زاوية القطاع الذي يمثل انتاج شهر نوفمبر: $32.1 = 360x\frac{1200}{13456}$

زاوية القطاع الذي يمثل انتاج شهر ديسمبر: $35.31 = 360x\frac{1320}{13456}$



الشكل (2.2): يوضح القطاعات الدائرية للإنتاج الشهري لمصنع مواد غذائية.

7.2 الأعمدة البيانية (Bar Charts):

الأعمدة البيانية وهي عبارة عن مســـتطيلات قائمة، ذات قواعد متســـاوية ومرتبة على المحور الأفقي، حيث يمثل كل مســـتطيل فئة محددة من فئات جدول التوزيع التكراري، ويعكس ارتفاع المســـتطيل تكرار الفئة أو مقدارها، كما يجب أن تكون المســـافات بين المســـتطيلات متســـاوية لضمان وضوح الرسم وسهولة قراءة البيانات.

عادةً تُســـتخدم الأعمدة البيانية لتمثيل البيانات الوصـــفية أو البيانات الكمية المتعلقة بمتغيرات منفصلة، حيث تمثل كل فئة قيمة واحدة من قيم المتغير المدروس.

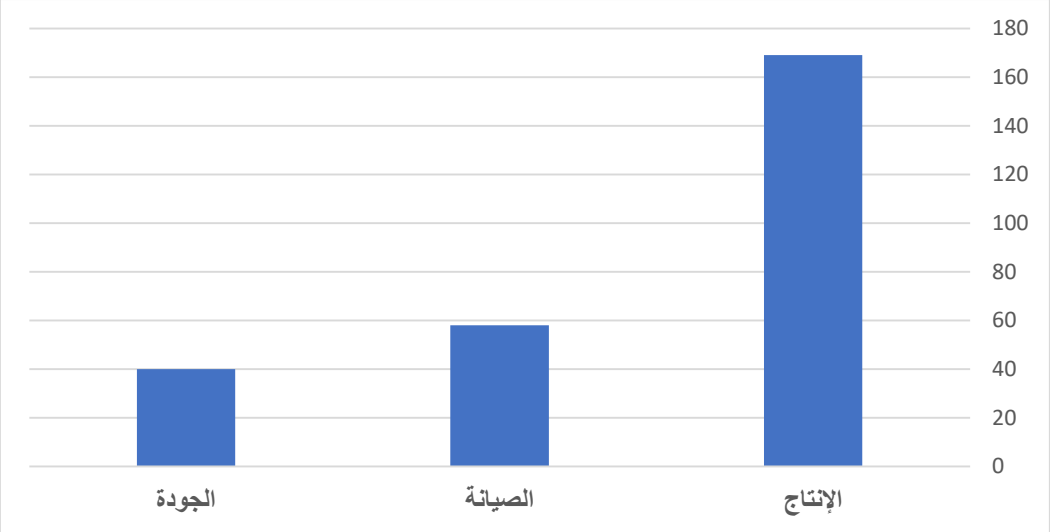
مثـال (12.2)

يبين الجدول التـــــــالي توزيع العمال الليبيين والغير ليبيين في أحد المصـانع داخل مدينة مصراتة.

الإدارة	عدد الليبيين	الغير ليبيين	العدد الكلي
الإنتـاج	124	45	169
الصيـانة	24	34	58
الجـودة	25	15	40

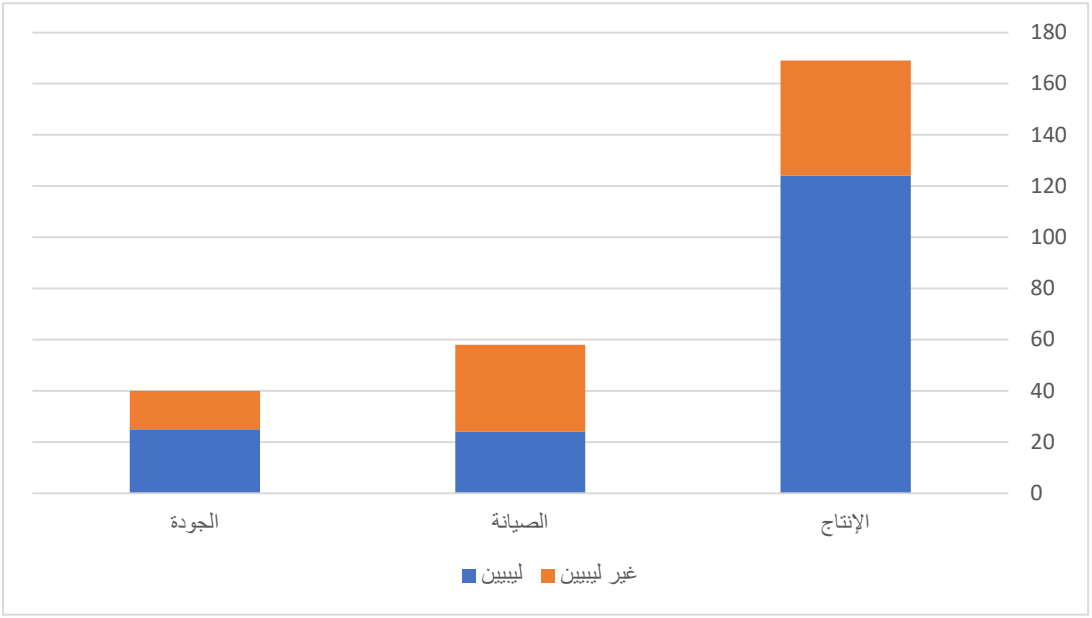
الحـل

إذا رغبنا في تمثيل الفئات المختلفة (الإدارات) دون التمييز بين الليبيين وغير الليبيين، فإن كل فئة تُمثل بمستطيل مستقل على الرسم البياني. يجب أن تكون قواعد المستطيلات متساوية الطول (في هذا المثال طول كل قاعدة يساوي 1 سم)، بينما يعكس ارتفاع كل مستطيل حجم المجموع الكلي للقوى العاملة في القطاع الذي يمثله. هذا الأســـلوب يتيح مقارنة أحجام الفئات بشـــكل مباشر، كما هو موضح في الشكل (3.2).



الشكل (3.2): يوضح تمثيل الأعمدة البيانية لإجمالي العمال داخل إدارات المصنع.

كما يمكننا باستخدام الأعمدة المجزأة (Stacked Bar Charts) مقارنة مكونات ظاهرة معينة مع إجماليها، على سبيل المثال يمكن مقارنة عدد العمال الغير الليبيين في كل إدارة بإجمالي العمال في نفس الإدارة. يتم ذلك عن طريق تقسيم كل مستطيل إلى جزأين بواسطة خط يوازي القاعدة، بحيث يمثل الجزء السفلي عدد العمال الغير الليبيين، والجزء العلوي يمثل بقية العمال، والشكل (4.2) يوضح هذا النوع من العرض البياني، والذي يتيح رؤية الحصة النسبية لكل مكون ضمن المجموع الكلي بطريقة واضحة وسهلة الفهم.

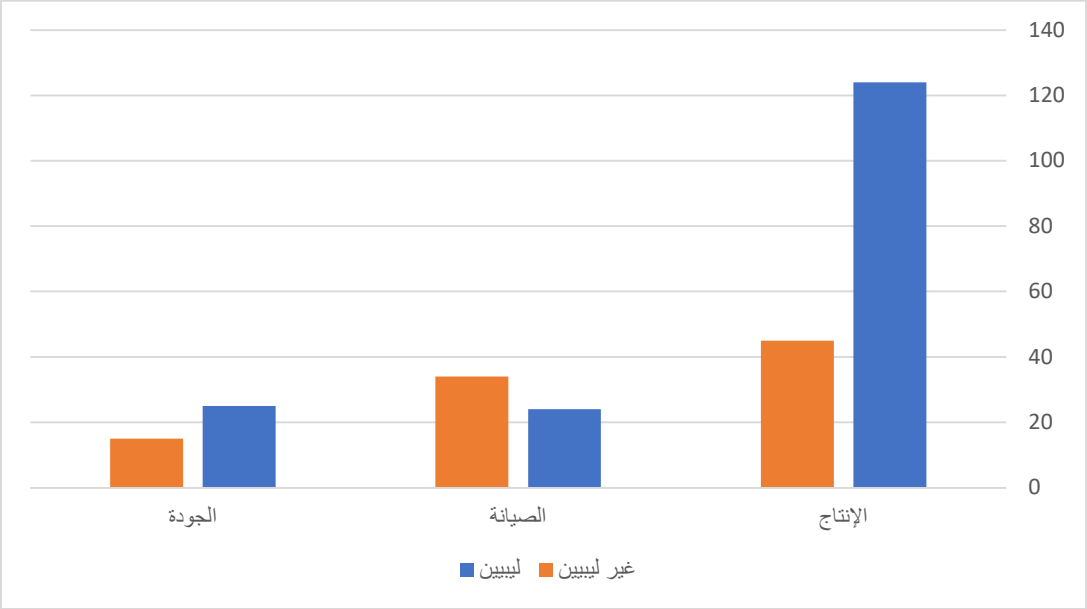


الشكل (4.2): يوضح تمثيل الأعمدة البيانية لتوزيع العمال بين ليبيين وغير ليبيين داخل إدارات المصنع.

كما يمكن استخدام الأعمدة المتلاصقة (Clustered Bar Charts) لمقارنة ظاهرتين أو أكثر ضمن نفس الفئات، فعلى سبيل المثال يمكن مقارنة عدد العمال الليبيين بغير الليبيين في كل قطاع من القطاعات.

يتم ذلك برسم مستطيلات متجاورة تمثل كل ظاهرة (مثل العمال الليبيين وغير الليبيين) داخل الفئة الواحدة (القطاع)، بحيث يكون عرض المستطيلات متساويًا والمسافة الفاصلة بينها ثابتة لسهولة المقارنة البصرية بين القيم.

يوضح الشكل (5.2) هذا النوع من التمثيل البياني، والذي يُعرف باسم طريقة الأعمدة المتلاصقة.



الشكل (5.2): يوضح تمثيل الأعمدة البيانية لتوزيع العمال بين ليبيين وغير ليبيين داخل إدارات المصنع.

كما ذكرنا سابقًا، يمكن استخدام الأعمدة البيانية أيضًا لتمثيل الجداول التكرارية الخاصة بالبيانات المنفصلة، حيث تمثل كل فئة في الجدول قيمة واحدة للمتغير المنفصل قيد الدراسة.

المثال التالي يوضح كيفية استخدام هذه الطريقة في عرض بيانات كمية منفصلة، بحيث يعكس ارتفاع كل مستطيل مجموع القيم لكل فئة، ويمكن تقسيمه إلى مكونات مختلفة عند الحاجة.

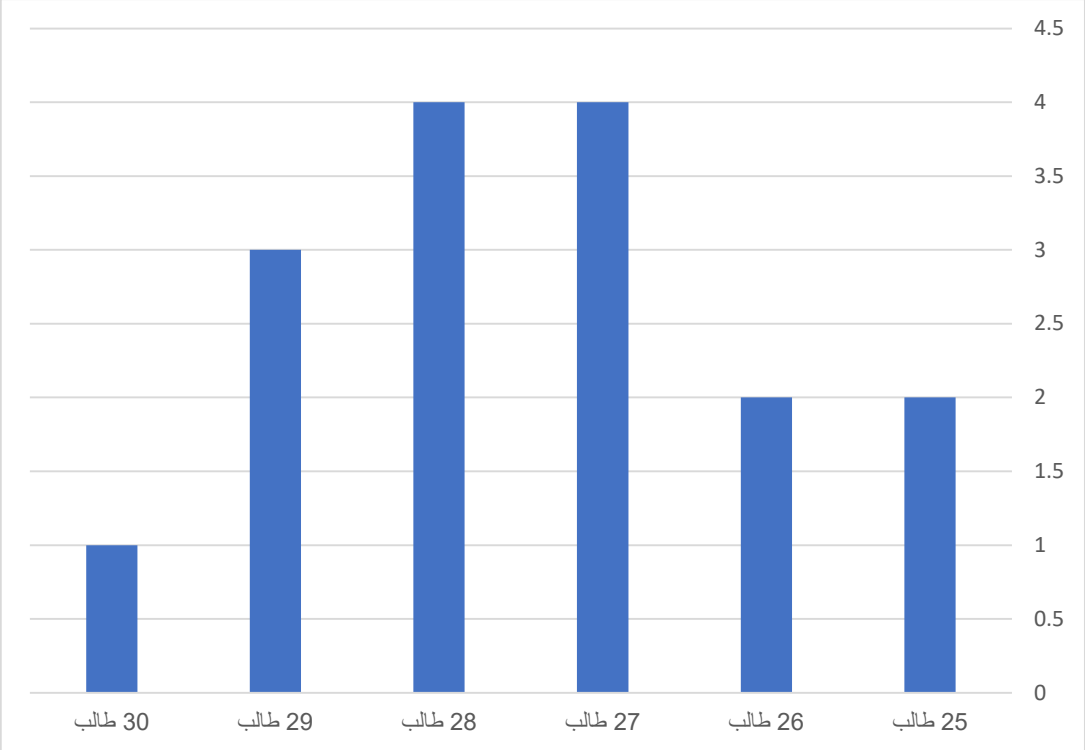
مثـال (13.2)

مدرسـة اعدادية بها 16 فصـلا دراسـيا، وجدول التوزيع التكراري التالي يوضـح توزيع الفصـول الدراسية حسب أعداد الطلبة بمدرسة بها 16 فصلا.

عدد الطلبة "الفئات"	25	26	27	28	29	30	المجموع
عدد الفصول "التكرار"	2	2	4	4	3	1	16

الحـل

ونفهم من هذا الجدول أنه يوجد فصلان بالمدرسة كل واحد يشمل 25 طالباً، وفصلان يشمل كل واحد منهم 26 طالباً، وأربعة فصـــول تشـــمل كل واحد منها 27 طالباً، وبالتالي نلاحظ أن هذه البيانات تمثل بشكل علمي بشكل منفصل، وكل فئة في الجدول تحتوي على قيمة واحدة، وبالتالي نرغب في أن نمثل هذه البيانات بأعمدة بيانية كما في الشكل (6.2).



الشكل (6.2): يوضح تمثيل الأعمدة البيانية لتوزيع عدد الطلبة على الفصول.

8.2 المدرج التكراري (Histogram) :

المدرّج التكراري هو أحد أهم الأدوات البصـــرية في التحليل الإحصـــائي، ويُســـتخدم بكثرة في المجالات الهندســـية لتمثيل البيانات العملية بطريقة دقيقة وواضـــحة، ويتكوّن من مجموعة من المستطيلات المتلاصقة، بحيث يمثل كل مستطيل فئة من فئات التوزيع التكراري، وتكون مساحة كل مستطيل متناسبة مع تكرار الفئة التي يعبّر عنها.

الفكرة ببساطة: كلما زاد عدد المفردات (القياسات أو القيم) في فئة معينة، زاد ارتفاع المستطيل المقابل لها، مما يجعل من السهل على المهندس أو الباحث رؤية أماكن التركّز في البيانات.

يُستخدم المدرج التكراري مثلاً في تحليل نتائج اختبارات المتانة أو تفاوت أبعاد القطع الميكانيكية أو درجات الحرارة في نظام تبريد صناعي.

فمثلاً، إذا كانت لدينا بيانات عن سُمك صفائح معدنية تم إنتاجها في مصنع، يمكن تمثيل هذه القيم في مدرّج تكراري لمعرفة ما إذا كان معظم سُمك هذه الصفائح تقع ضمن حدود الجودة المطلوبة أم أن هناك انحرافاً في عملية الإنتاج، ولرسم المدرج التكراري نقوم بالخطوات الآتية:

- نرسم محورين متعامدين:

المحور الأفقي (X) يمثل الفئات مثل نطاقات سمك الصفائح أو الزمن أو درجة الحرارة.

المحور الرأسي (Y) يمثل التكرارات أي عدد المشاهدات في كل فئة.

- نرسم مستطيلات متلاصقة فوق كل فئة بحيث يكون ارتفاع كل مستطيل متناسباً مع تكرارها.

لا تُترك أي مسافات بين المستطيلات لأن الفئات تمثل قيماً متصلة (أو متجاورة في حالة المتغيرات المنفصلة).

وبذلك نحصل على المدرج التكراري الذي يتيح للمهندس أو الإحصائي تصور شكل توزيع البيانات — هل هو متمركز، مائل، أو متعدد القمم — وهي خطوة أساسية في فهم العمليات الصناعية وجودتها.

مثـال (14.2)

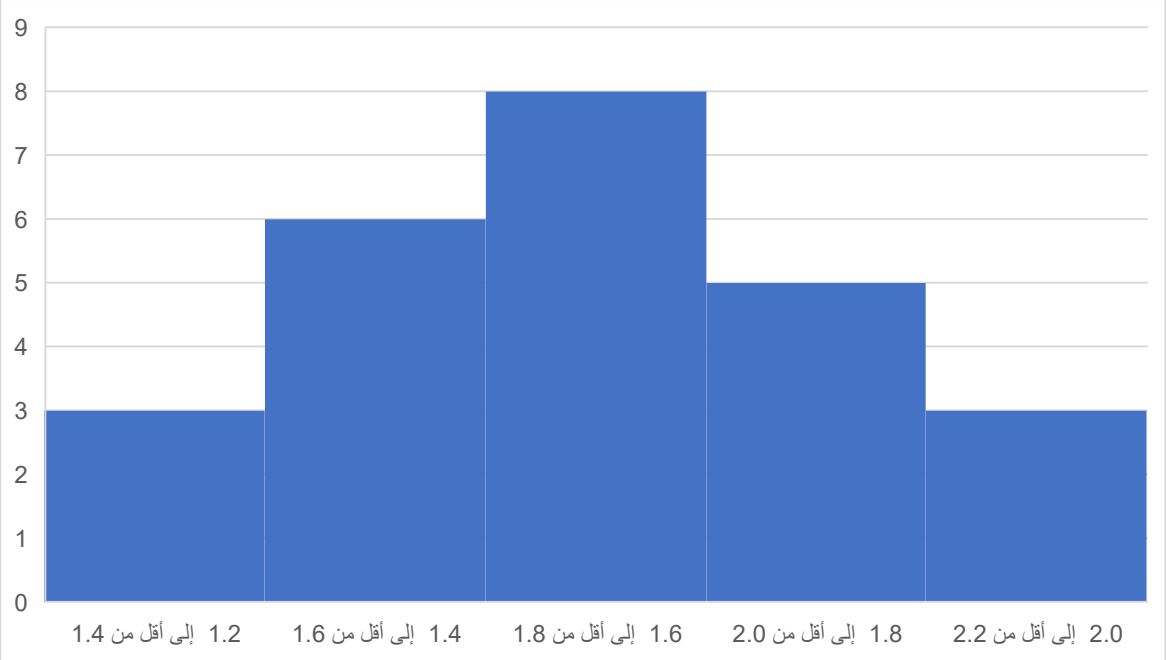
الجدول التالي يوضـح التوزيع التكراري لزمن الإشـارة (بالميلي ثانية) الذي تسـتغرقه حزمة بيانات للانتقال من محطة إرســال إلى محطة اســتقبال في أحد خطوط الألياف الضــوئية، خلال 25 عملية قياس متتالية.

المطلوب: تمثيل هذا التوزيع باستخدام المدرج التكراري.

زمن الإشارة (مللي ثانية)	التكرار (عدد القياسات)
1.2 إلى أقل من 1.4	3
1.4 إلى أقل من 1.6	6
1.6 إلى أقل من 1.8	8
1.8 إلى أقل من 2.0	5
2.0 إلى أقل من 2.2	3
المجموع	25

الحـل

نمثل حدود الفئات (زمن الإشـــــارة) على المحور الأفقي، ونمثل التكرارات (عدد القياسات) على المحور الرأسـي، باسـتخدام مقياس رسـم مناسـب، وبما أن الفئات متسـاوية الطول (أي أن الجدول منتظم)، فإننا نرسـم فوق كل فئة مسـتطيلاً ارتفاعه يسـاوي تكرارها، بحيث تتلاصـق المسـتطيلات دون فواصل.



الشكل (7.2): يوضح تمثيل المدرج التكراري لتمثيل زمن الإشارة لحزمة بيانات خلال 25 عملية قياس متتالية.

من خلال المدرج يمكن لمهندس الاتصـــالات أن يلاحظ أن أغلب الإشـــارات تســـتغرق زمنًا بين 1.6 و1.8 مللي ثانية، مما يشـــير إلى أن هذا هو النطاق المثالي للأداء، بينما القيم الأعلى (مثل 2.0 مللي ثانية فأكثر) قد تدل على بطء أو فقد في الكفاءة في أجزاء معينة من الشبكة.

مثـــال (15.2)

الجدول التالي يوضح التوزيع التكراري لتيار التشغيل (بالأمبير) المسجَّل أثناء اختبار 30 محركًا كهربائيًا في ورشة الصيانة بقسم الهندسة الكهروميكانيكية.

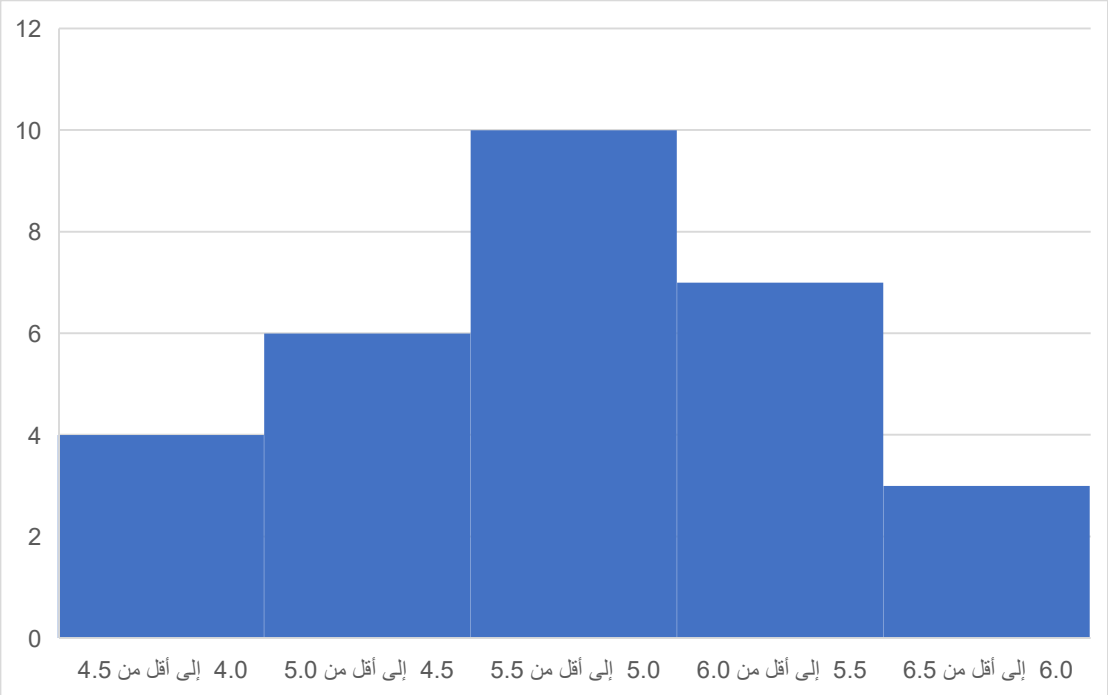
المطلوب: تمثيل هذا التوزيع باستخدام المدرج التكراري.

تيار التشغيل(A)	التكرار (عدد المحركات)
4.0 إلى أقل من 4.5	4
4.5 إلى أقل من 5.0	6
5.0 إلى أقل من 5.5	10
5.5 إلى أقل من 6.0	7
6.0 إلى أقل من 6.5	3
المجموع	30

الحـل

نرسـم المحور الأفقي (X) لتمثيل فئات تيار التشـغيل (A)، ومن ثم نرسـم المحور الرأسـي (Y) لتمثيل التكرارات (عدد المحركات).

بما أن الفئات متساوية الطول (0.5 أمبير)، نرسم مستطيلات متلاصقة، ارتفاع كل منها يساوي تكرار الفئة المقابلة.



الشكل (8.2): يوضح تمثيل المدرج التكراري لتوزيع تيار التشغيل (بالأمبير) لـثلاثين محركًا.

حيث يمثل ارتفاع المستطيل عدد المحركات التي يمر بها تيار تشغيل ضمن المدى المحدد، كما يمكن ملاحظة أن أعلى تكرار عند الفئة 5.0 – 5.5 أمبير، ما يعني أن أغلب المحركات تعمل في هذا المدى، وهو المدى الاسمي أو المثالي للتشغيل.

أما المحركات التي تعمل عند تيار أعلى من 6 أمبير فهي قليلة (3 فقط)، وقد تكون مؤشراً على مشاكل في ملفات المحرك أو زيادة في الحمل الميكانيكي.

9.2 المضلع التكراري:

المضلع التكراري هو خط منكسر أو منحنى ولكنه مكوَّن من أضلاع مستقيمة تصل بين نقاط تمثل فئات التوزيع التكراري، وتمثل كل نقطة من هذه النقاط مركز فئة معينة على المحور الأفقي، بينما يعبّر موقعها على المحور الرأسي عن تكرار تلك الفئة.

لكي يكون المضلع مغلقًا من الطرفين، نضيف فئة قبل الأولى وأخرى بعد الأخيرة، بحيث يكون تكرارهما صفراً، فتقع النقطتان المقابلتان لهما على المحور الأفقي، مما يُغلق الشكل تمامًا، ولرسح المضلع التكراري نقوم بالتالي:

- نرسم محورين متعامدين:

  الأفقي (X) لتمثيل الفئات.

  الرأسي (Y) لتمثيل التكرارات.

- نضيف فئة قبل الأولى وفئة بعد الأخيرة بتكرار يساوي صفرًا.
- نحسب مراكز الفئات (بجمع الحد الأدنى والحد الأعلى لكل فئة ثم قسمتهما على 2).
- نضع نقاط التمثيل بحيث:

  الإحداثي السيني (X) = مركز الفئة.

  الإحداثي الصادي (Y) = تكرار الفئة.

- نصل بين هذه النقاط بخطوط مستقيمة متتالية، فنحصل على المضلع التكراري.

مثـال (16.2)

في مصـنع لإنتاج قطع معدنية، تم تسـجيل عدد القطع المنتجة يوميًا على مدى أسـبوع، ووزعت البيانات في جدول تكراري

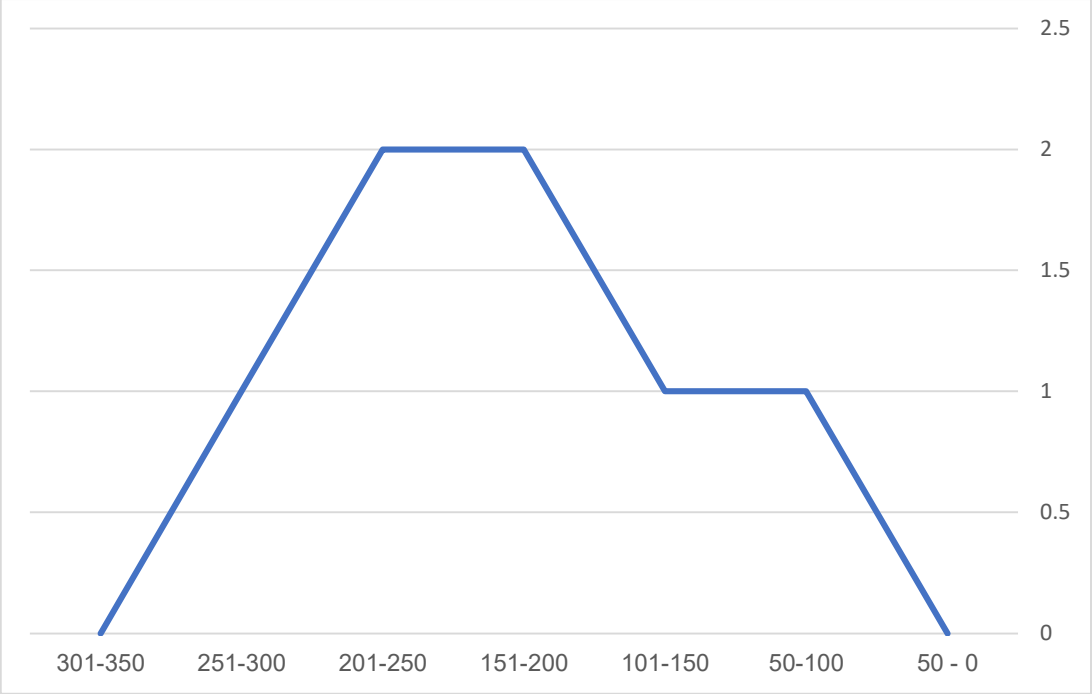
عدد القطع المنتجة (قطع)	عدد الأيام (التكرار)
51 – 100	1
101 – 150	1
151 – 200	2
201 – 250	2
251 - 300	1

الحـل

لرسـم المضـلع التكراري، نفترض وجود **فئة قبل الأولى وفئة بعد الأخيرة** بحيث يكون تكرار كل منهما صـفرًا. بعد ذلك، نحدد **مراكز الفئات** ثم نحدد **نقاط المضـلع** كما هو موضـح في الجدول (18.2)، ترسم هذه النقاط على الرسم البياني وتوّصل بخطوط مستقيمة، نحصل على **المضلع التكراري** الموضح في الشكل (9.2).

الجدول (18.2): جدول يوضح التوزيع التكراري ومركز الفئة ونقاط المضلع.

الفئة	مركز الفئة	التكرار	نقاط المضلع
0 – 50	25	0	(0،25)
51 – 100	75.5	1	(1،75.5)
101 – 150	125.5	1	(1،125.5)
151 – 200	175.5	2	(2،175.5)
201 - 250	225.5	2	(2،225.5)
251 – 300	275.5	1	(1،275.5)
301 - 350	325.5	0	(0،325.5)



الشكل (9.2): يوضح المضلع التكراري لإنتاج القطع المعدنية لمدة أسبوع.

مثـال (17.2)

تم تسجيل أوزان المنتجات النهائية المعبأة خلال أسبوع، وكانت البيانات موزعة كالتالي:

فئة الوزن (كغ)	عدد المنتجات (التكرار)
10 – 6	3
15 – 11	5
20 – 16	4
25 – 21	2

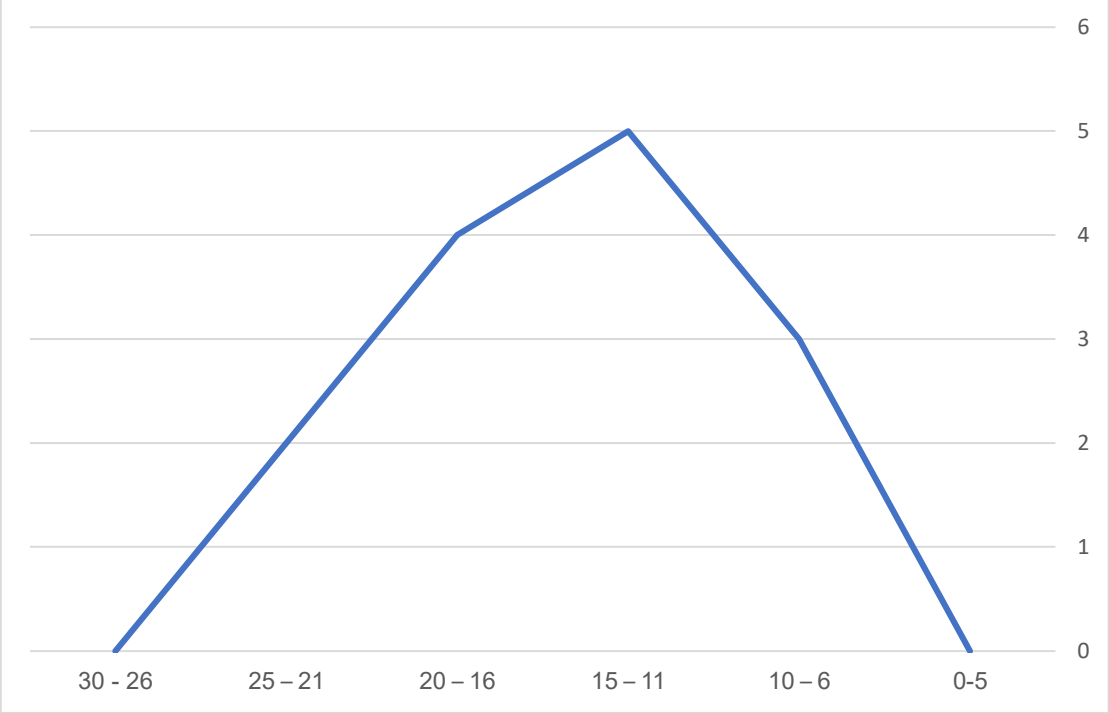
الحـــل

كما العادة نفترض وجود **فئة قبل الأولى وفئة بعد الأخيرة** بحيث يكون تكرار كل منهما صـــــفرًا.

بعد ذلك، نحدد **مراكز الفئات** ثم نحدد **نقاط المضلع**.

الجدول (19.2): جدول يوضح التوزيع التكراري ومركز الفئة ونقاط المضلع.

الفئة	مركز الفئة	التكرار	نقاط المضلع
5 – 1	3	0	(0،3)
10 – 6	8	3	(3،8)
15 – 11	13	5	(5،13)
20 – 16	18	4	(4،18)
25 – 21	23	2	(2،23)
30 – 26	28	0	(0،28)



الشكل (10.2): يوضح المضلع التكراري لأوزان المنتجات النهائية المعبأة خلال أسبوع.

10.2 المنحنى التكراري:

يُرسم المنحنى التكراري باتباع نفس خطوات المضلع التكراري، مع اختلاف طريقة توصيل النقاط. ففي المنحنى التكراري، تُوصل النقاط بمنحنى مستمر وممهّد يمر بأكبر عدد ممكن من النقاط، ويحقق توازنًا عند النقاط المتبقية.

في هذا النوع من الرسوم، لا نضيف فئة قبل الأولى أو بعد الأخيرة، وبسبب التقريب في الرسم وعدم الالتزام بالنقاط بدقة، فإن المساحة تحت المنحنى لا تساوي المساحة تحت المضلع أو

المدرج التكراري، كما يكون المنحنى التكراري ممهّدًا بشكل أفضل كلما زاد عدد الفئات وصغرت أطوالها، أي عندما تكون النقاط قريبة من بعضها.

مثـــال (18.2)

في مختبر هندسي، تم تسجيل أوقات التشغيل المتواصل لمحركات كهربائية خلال تجارب اختبار استمرارية التشغيل، وكانت النتائج موزعة كالتالي:

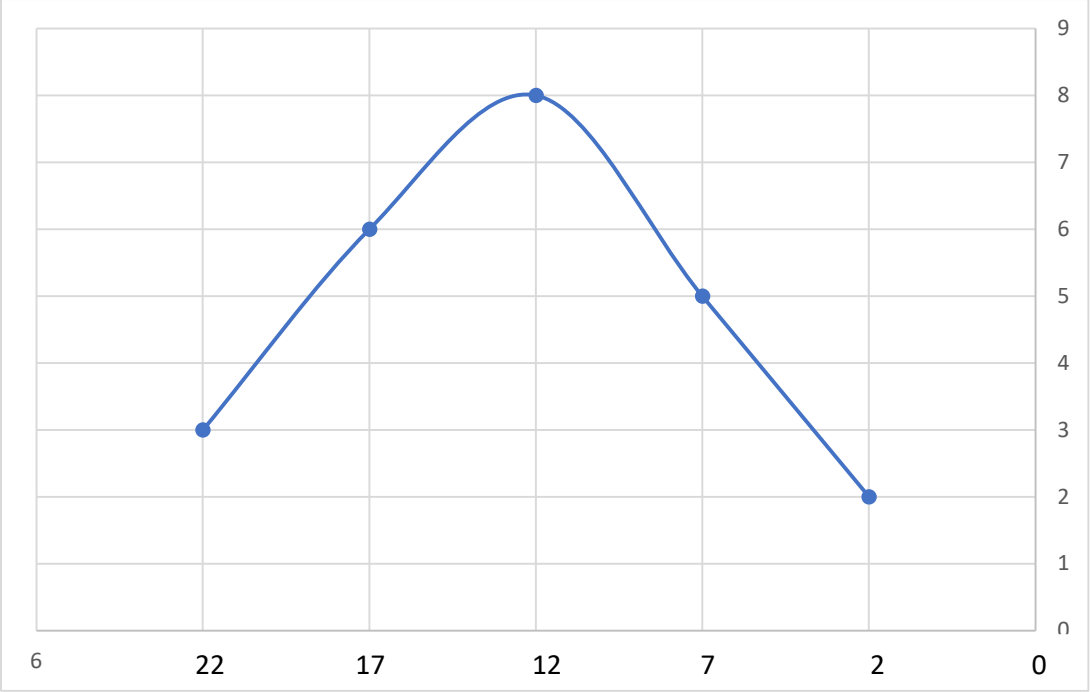
فترة التشغيل (ساعات)	عدد المحركات (التكرار)
4 – 0	2
9 – 5	5
14 – 10	8
19 – 15	6
24 – 20	3

الحـــل

نحدد مراكز الفئات ثم نحدد نقاط المنحنى كما هو موضح في الجدول (20.2)، ترسم هذه النقاط على الرسم البياني وتوّصل بخطوط ممهدة ومنحنية، نحصل على المنحنى التكراري الموضح في الشكل (12.2).

الجدول (20.2): جدول يوضح التوزيع التكراري ومركز الفئة ونقاط المنحنى.

الفئة	مركز الفئة	التكرار	نقاط المنحنى
4 – 0	2	2	(2،2)
9 – 5	7	5	(5،7)
14 – 10	12	8	(8،12)
19 – 15	17	6	(6،17)
24 – 20	22	3	(3،22)

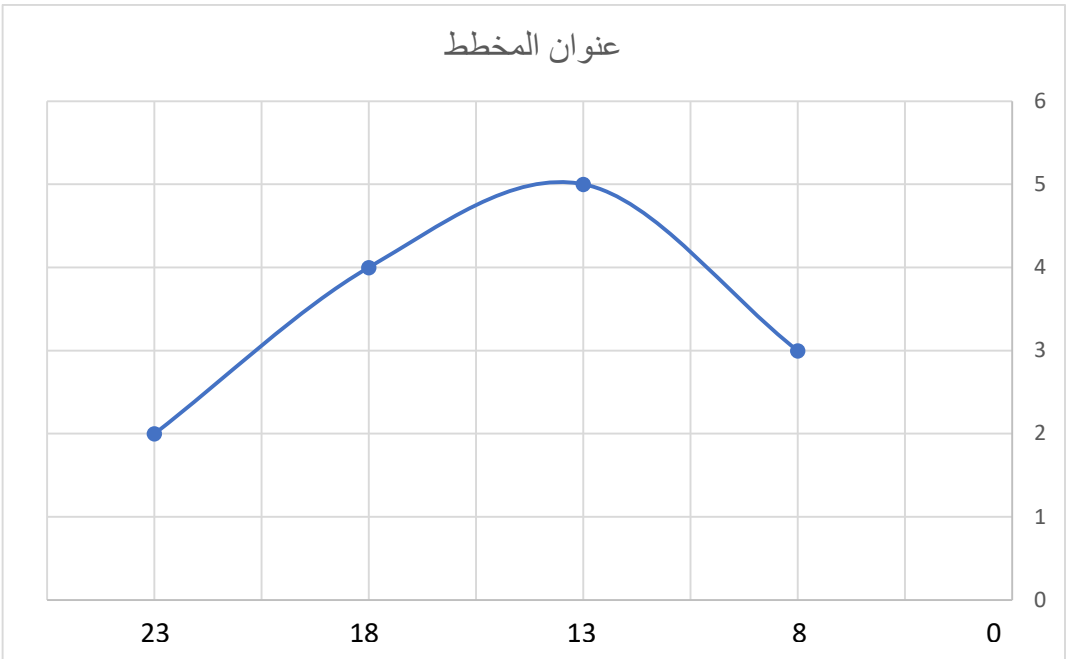


الشكل (11.2): يوضح المنحنى التكراري لساعات تشغيل عدد من المحركات.

مثـــال (19.2)

مثل الجدول التكراري الوارد في المثال (17.2) بيانياً باستخدام المنحنى التكراري.

الحـــل



الشكل (12.2): يوضح المنحنى التكراري لأوزان المنتجات النهائية المعبأة خلال أسبوع.

11.2 المنحنى التكراري المتجمع "الصاعد والهابط":

لتمثيل جدول التوزيع التكراري المتجمع (سواء كان صاعدًا أو هابطًا) بيانيا نتبع الخطوات التالية:

- نرســـم محورين متعامدين؛ المحور الأفقي لتمثيل الفئات، والمحور الرأســـي لتمثيل التكرارات.

  يجب اختيار مقياس رسم يتناسب مع أكبر قيمة للتكرار المتجمع (أي التكرار الكلي).

- نمثل كل فئة بنقطة يكون:

  الإحداثي الأفقي = الحد الأدنى للفئة،

  الإحداثي الرأسي = التكرار المتجمع لتلك الفئة.

- بعد تحديد النقاط، نصل بينها بمنحنى منتظم ومائل بسلاسة.

  إذا كانت البيانات تمثل التكرار المتجمع الصاعد، نحصل على منحنى صاعد.

  وإذا كانت تمثل التكرار المتجمع الهابط، نحصل على منحنى هابط.

ويوضّح المثال التالي تطبيق هذه الخطوات عمليًا.

مثــال (20.2)

ارســـم المنحنى التكراري المتجمع الصـــاعد والمنحنى التكراري المتجمع الهابط للتوزيع التكراري التالي واذ1ي يمثل أوزان 100 قطعة غيار سيارات متنوعة.

الوزن (بالجـرام)	عدد القطع
600 إلى أقل من 750	5
750 إلى أقل من 1000	15
1000 إلى أقل من 1250	40
1250 إلى أقل من 1500	28
1500 إلى أقل من 1750	12
المجموع	100

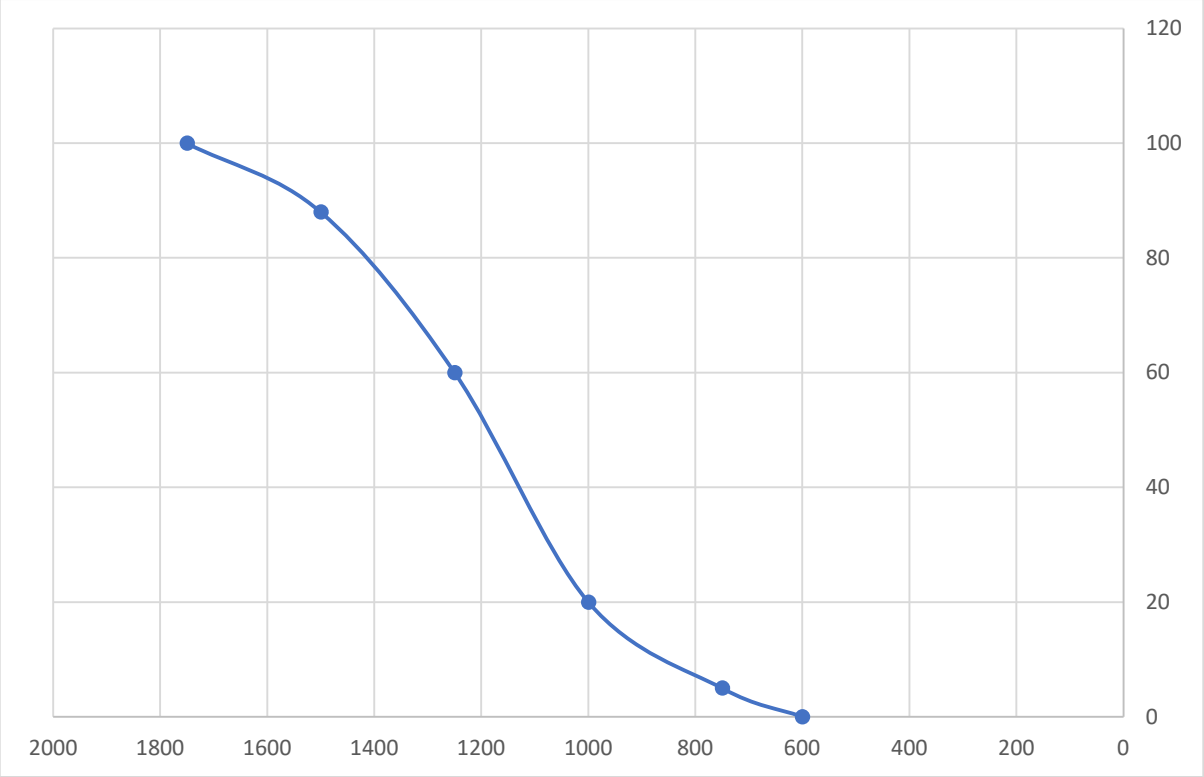
الحـــل:

نكون جدول التكرار المتجمع الصـاعد والموضـح في الجدول (21.2)، ثم نقوم برصـد النقط على الرســم ، ونصـــل بين النقط بمنحنى فنحصـــل على المنحني التكراري المتجمع الصـــاعد الموضح في الشكل (13.2).

ونكون جدول التكرار المتجمع الهابط والموضـــح في الجدول (22.2)، ثم نقوم برصـــد النقط على الرسم ، ونصل بين النقط بمنحنى فنحصل على المنحني التكراري المتجمع الهابط الموضح في الشكل (14.2).

الجدول (21.2): جدول يوضح التوزيع التكراري المتجمع الصاعد لأوزان القطع.

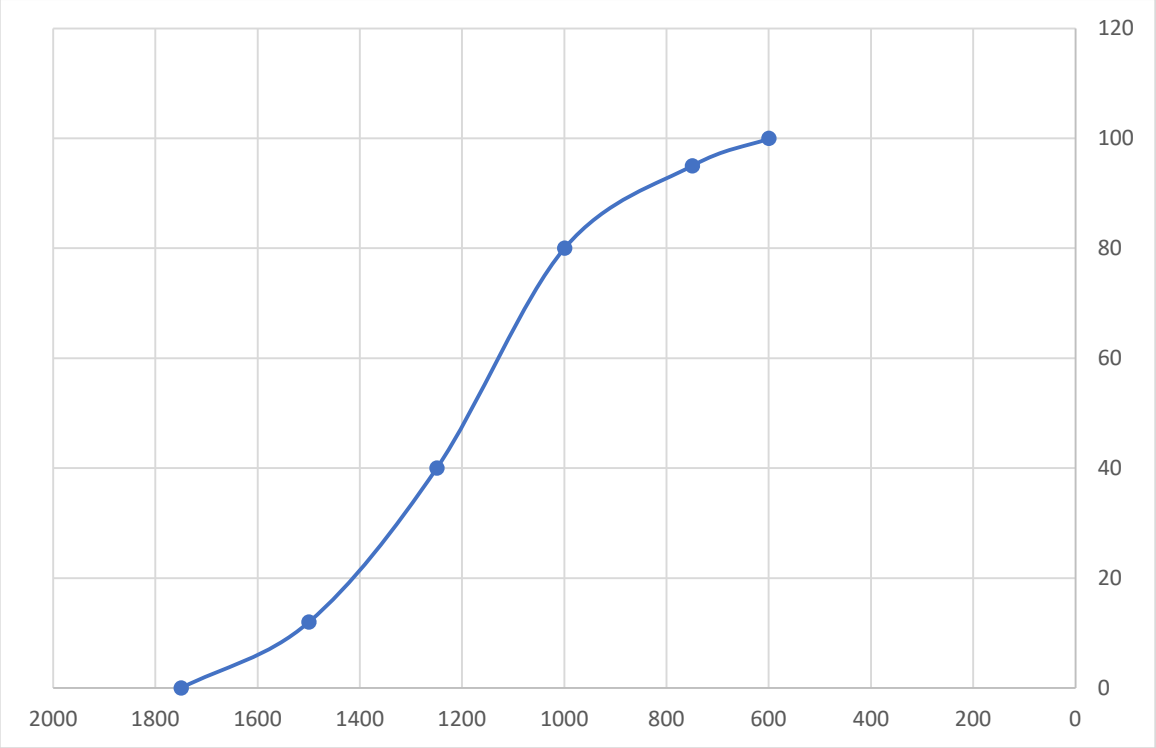
الفئة	التكرار
أقل من 600	0
أقل من 750	5
أقل من 1000	20
أقل من 1250	60
أقل من 1500	88
أقل من 1750	100



الشكل (12.2): يوضح المنحنى التكراري الصاعد لأوزان قطع غيار السيارات.

الجدول (22.2): جدول يوضح التوزيع التكراري المتجمع الهابط لأوزان القطع.

الفئة	التكرار
600 فأكثر	100
750 فأكثر	95
1000 فأكثر	80
1250 فأكثر	40
1500 فأكثر	12
1750 فأكثر	0



الشكل (12.2): يوضح المنحنى التكراري الهابط لأوزان قطع غيار السيارات.

على الرغم من اختلاف شكل المنحنيين المتجمعين الصاعد والهابط، إلا أنهما يؤديان الغرض نفسه، فمن خلال المنحنى المتجمع الصاعد يمكن معرفة عدد المفردات الأقل من قيمة معينة، وبطرح هذا العدد من المجموع الكلي للمفردات نحصل على عدد المفردات الأكبر من أو تساوي تلك القيمة، وهو ما يقدمه أيضًا المنحنى المتجمع الهابط.

وعند رسم المنحنيين في الشكل نفسه، نلاحظ أنهما يلتقيان في نقطة واحدة تمثل القيمة التي يكون عندها عدد المفردات الأصغر مساويًا لعدد المفردات الأكبر.

أي عندما يتساوى التكرار المتجمع الصاعد مع الهابط، ويشكل كل منهما 50٪ من العدد الكلي للمفردات، وتُسمى هذه النقطة الوسيط لأنها تمثل القيمة الوسطى في مجموعة البيانات.

مثـــال (21.2)

الجدول التالي يوضح أوزان 100 قطعة منتجة بالجرامات.

الوزن	عدد القطع
20 إلى أقل من 25	6
25 إلى أقل من 30	14
30 إلى أقل من 35	43
35 إلى أقل من 40	32
40 إلى أقل من 45	5

ارسم المنحنى التكراري المتجمع الصاعد للأوزان، ومن الرسم أوجد:

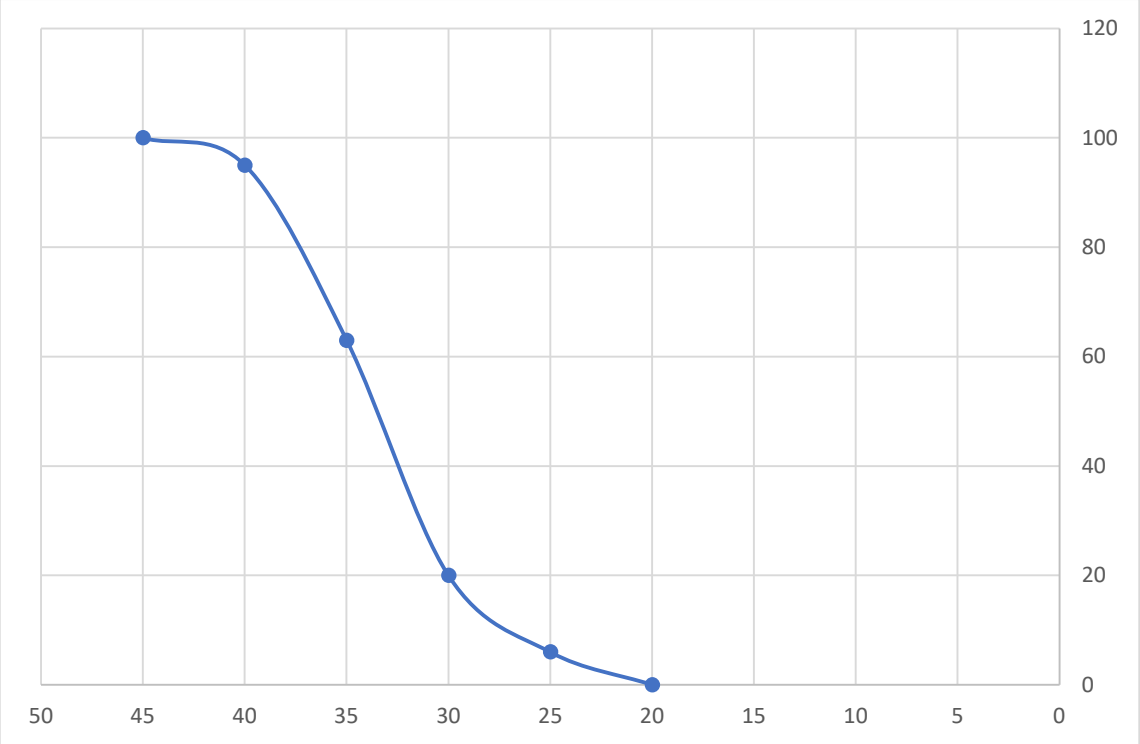
عدد القطع التي أوزانها من 24 إلى أقل من 32.

الحـــل

نكون جدول التكرار المتجمع الصـاعد والموضـح في الجدول (23.2) من ثم نقوم برصـد النقاط التي تمثل الفئات ونصل بينها بمنحنى فنحصل على المنحنى التكراري المتجمع الصاعد.

الجدول (23.2): جدول يوضح التوزيع التكراري المتجمع الصاعد لأوزان القطع.

الفئة	التكرار
أقل من 20	0
أقل من 25	6
أقل من 30	20
أقل من 35	63
أقل من 40	95
أقل من 45	100



الشكل (13.2): يوضح المنحنى التكراري الصاعد لأوزان القطع.

يمكننا تحديد عدد القطع التي تتراوح أوزانها بين 24 جرامًا وأقل من 32 جرامًا باستخدام المنحنى التكراري المتجمع الصاعد كما يلي:

نرسم خطًا عموديًا من النقطة التي تمثل الوزن 24 على المحور الأفقي حتى يلتقي بالمنحنى، ثم نرسم من نقطة التقاطع خطًا أفقيًا نحو المحور الرأسي، وتدل نقطة التقاطع مع المحور على عدد القطع التي أوزانها أقل من 24 جرامًا، وهي تقريبًا 5 قطع من الشكل (13.2).

وبالطريقة نفسها، نرسم خطًا عند الوزن 32 جرامًا لنجد أن عدد القطع التي أوزانها أقل من 32 جرامًا هو تقريبًا 35 قطعة.

إذن، عدد القطع التي تتراوح أوزانها بين 24 و32 جرامًا يُحسب كالتالي:

عدد القطع = 35 − 5 = 30 قطعة.

تمـــارين

أسئلة اختيار من متعدد:

1. أي مما يلي يُعتبر جدول توزيع تكراري للبيانات الوصفية؟
   أ- جدول يحتوي على فئات تقديرات الاختبار وعدد الطلاب في كل فئة.
   ب- جدول يحتوي على متوسط الأعمار لكل مجموعة.
   ت- جدول يحتوي على درجات الحرارة اليومية بالفهرنهايت.
   ث- جميع ما ذكر.
2. عند إعداد جدول توزيع تكراري للبيانات الكمية المنفصـــلة، إذا كان عدد القيم المختلفة صغيرًا، نقوم بـ:
   أ- دمج كل القيم في فئة واحدة.
   ب- تخصيص فئة لكل قيمة منفصلة.
   ت- تجاهل القيم المتكررة.
   ث- تحويل البيانات إلى نسب مئوية.
3. إذا كانت البيانات موزعة في جدول توزيع تكراري بفئات متساوية الطول، فإن مركز الفئة يُحسب بـ:
   أ- مجموع القيم داخل الفئة.
   ب- الحد الأعلى للفئة ÷ الحد الأدنى للفئة.
   ت- (الحد الأدنى + الحد الأعلى) ÷ 2.
   ث- الفرق بين أكبر وأصغر قيمة في الفئة.
4. أي مما يلي يُمثل جدول توزيع تكراري مفتوح من الطرف الأعلى؟
   أ- أقل من 10.
   ب- 10 – 20.
   ت- 50 فأكثر.
   ث- كلا من أ و ت.

5. أي العبارات صحيحة عن التكرار النسبي؟

أ- يُحسب بقسمة تكرار الفئة على مجموع التكرارات.

ب- يكون دائمًا عددًا صحيحًا.

ت- لا يمكن التعبير عنه بنسبة مئوية.

ث- يستخدم فقط للبيانات النوعية.

أسئلة الصواب والخطاء:

1. يتم استخدام علامات (//) في جدول تفريغ البيانات لتسهيل العدّ. ( )
2. الفئات في الجداول التكرارية يجب أن تكون دائمًا متساوية الطول، ولا يجوز أن تكون غير منتظمة. ( )
3. التوزيع التكراري المتجمع الصاعد يوضح عدد العناصر التي قيمتها أقل من قيمة معينة. ( )
4. التوزيع التكراري المتجمع الهابط يبدأ دائمًا من الصفر وينتهي بالعدد الكلي للمفردات. ( )
5. يمكن تحويل جدول توزيع تكراري عادي إلى جدول التكرار النسبي بسهولة بقسمة كل تكرار على العدد الكلي للمشاهدات. ( )

أسئلة تطبيقية:

1. أنشئ جدول توزيع تكراري عادي للبيانات التالية والتي تمثل أعمار 20 طالبًا في الصف، جدولاً موضحاً فيه كل عمر وعدد مرات تكراره.

15، 16، 17، 16، 18، 17، 15، 19، 17، 16،

18، 17، 16، 19، 18، 17، 16، 15، 18، 17

2. البيانات التالية تبين عدد ساعات الدراسة أسبوعيًا لـ 25 طالبًا:

أنشئ جدول تكراري نسبي يوضح كل فئة وعدد الطلاب ونسبتهم المئوية، وكذلك احسب التكرار النسبي التراكمي الصاعد لكل فئة.

5، 10، 15، 10، 20، 15، 25، 10، 15، 5، 20، 15،

10، 25، 20، 15، 5، 10، 15، 20، 25، 15، 10، 5، 15

3. في أحد المقررات الدراسية كانت البيانات التالية تمثل النتيجة النهائية لثلاثين طالب:

56، 72، 68، 85، 90، 78، 65، 88، 92، 70

60, 75, 80, 68, 74, 82, 77, 69, 66, 71

73, 79, 81, 67, 84, 76, 63, 62, 89, 87

المطلوب: أنشئ جدول تكراري صاعد.

4. تمثل البيانات التالية عدد ساعات التدريب أسبوعيًا لـ 35 متدرباً في شركة هندسية:

15، 20، 18، 25، 30، 22، 28، 26,

24, 19, 21, 27, 23, 25, 20, 22,

24, 18, 21, 26, 29, 19, 23, 25, 28,

20, 22, 27, 24, 26, 30, 18, 21, 22, 25

أنشـــئ جدول تكراري هابط، ثم حدّد ما إذا كان معظم المتدربون يتدربون ســـاعات كثيرة أم قليلة.

5. في إحدى الشركات قامت إدارة الموارد البشرية بعمل استبانة لجمع أعمار الموظفين وهم 50 موظفًا وكانت كالتالي:

23، 35، 29، 40، 33، 31، 28، 42, 36, 30, 27, 38, 41, 34, 32, 29, 37,
26, 39, 30, 31, 33, 35, 28, 36, 34, 29, 32, 37, 30, 33, 39, 31, 28,
40, 35, 34, 29, 38, 27, 36, 32, 30, 33, 37, 31, 29, 34, 35, 30

المطلوب منك عمل الجدول التكراري العادي، الجدول التكراري المتجمع الصـــاعد، الجدول التكراري المتجمع الهابط، الجدول التكراري النسبي.

6. البيانات: عدد الطلاب المسجلين في 4 تخصصات خلال السنة:

التخصص	الفصل الأول	الفصل الثاني	الفصل الثالث
ميكانيكا	130	120	140
كهرباء	120	100	110
مدني	160	150	155
كمبيوتر	210	200	220

ارسم أعمدة بيانية لمقارنة عدد الطلاب بين التخصصات في كل فصل.

7. تمثل البيانات التالية نسبة استخدام مصادر الطاقة في مصنع:

المصدر	النسبة المئوية
كهرباء	40
غاز	25
نفط	20
طاقة شمسية	15

ارسم رسمًا دائريًا يوضح النسب المئوية لمصادر الطاقة.

8. قاس 30 طالبًا أوزانهم (بالكيلوغرام) وحُدِّدت الفئات التالية:

الوزن (كغ)	عدد الطلاب
49 – 40	2
59 – 50	5
69 – 60	10
79 – 70	8
89 – 80	5

ارســـم المدرج التكراري للبيانات، ثم أوجد ما الفئة التي تحتوي على أعلى عدد من الطلاب؟ وهل توزيع الأوزان يميل إلى الوزن الأعلى أم الأدنى؟ ولماذا؟

9. يمثل الجدول التكراري التالي درجات 50 طالبًا في اختبار إحصاء:

الفئة	عدد الطلاب
10 – 0	2
20 – 11	5
30 – 21	8
40 – 31	12
50 – 41	10
60 – 51	8
70 – 61	5

ارسم المضلع التكراري ووضح أي فئة لها أكبر تكرار.

10 . في إحدى معارض السيارات بمدينة مصراتة كان عدد السيارات المباعة في 6 أشهر كالتالي:

الشهر	عدد السيارات
يناير	50
فبراير	70
مارس	65
أبريل	80
مايو	75
يونيو	90

ارسم المنحنى التكراري العادي للبيانات، وحدد الشهر الذي كان فيه البيع الأعلى. من ثم ارسم المنحنى التكراري الصاعد والهابط لنفس الجدول التكراري.

الفصل الثالث

مقاييس النزعة المركزيــة

1.3 مقدمــة:

تقوم فكرة مقاييس النزعة المركزية على مبدأ بسيط لكنه جوهري في علم الإحصاء، وهو أنه يمكن تمثيل مجموعة كبيرة من البيانات بقيمة واحدة تعبّر عنها وتعكس اتجاهها العام، وهذه القيمة عادة ما تقع في منتصـــف البيانات، وتكون بمثابة النقطة التي تميل نحوها القيم وتتجمع حولها، ولهذا سُمّيت بالنزعة المركزية.

تمثل هذه المقاييس القلب النابض للتوزيع التكراري، فهي تمنحنا رؤية واضـــحة لمكان تمركز البيانات، وتســـاعدنا على فهم طبيعتها ومدى تشـــتتها، ويُنظر إلى القيمة المتوســـطة باعتبارها أفضل ممثل للبيانات، لأنها لا تكون متطرفة ولا شاذة، بل تعبّر عن الاتجاه العام الذي تميل إليه معظم القيم.

إن تحديد هذه القيمة الوسطى يمكّن الباحث من دراسة خصائص التوزيعات التكرارية، كما يسهّل المقارنة بين مجموعات مختلفة من البيانات للظواهر نفسها، وسنتناول في هذا الفصل أهم مقاييس النزعة المركزية، وهي:

1. الوسط الحسابي (Arithmetic Mean)
2. الوسيط (Median)
3. المنوال (Mode)

2.3 الوسط الحســابي:

يُعرَّف الوســط الحســابي بأنه القيمة التي تمثّل مركز ثقل البيانات، أي النقطة التي يتحقق عندها التوازن الإحصــائي للقيم، فلو تخيّلنا مجموعة من البيانات تمثل بأثقال متســاوية موزعة على لوح مدرّج، فإن اللوح ســيصـــل إلى حالة الاتزان عندما يُعلّق من نقطة تمثل مركز ثقله، وهذه النقطة هي ما يُعبّر عنه إحصائياً بالوسط الحسابي.

ويمثل الوسط الحسابي القيمة التي توازن بين القيم الصغيرة والكبيرة، فهو أشبه بخط التوازن الذي تستقر حوله البيانات، ويُحسب الوسط الحسابي لأي مجموعة من القيم وفق العلاقة التالية:

$$\text{الوسط الحسابي} = \frac{\text{مجموع القيم}}{\text{عدد القيم}}$$

ويُطلق على الوسط الحسابي أيضاً اسم المتوسط الحسابي ويرمز له في حالة العينة بالرمز $\bar{X}$ وفي حالة المجتمع الإحصائي بالرمز μ.

وسيتم لاحقًا توضيح كيفية تطبيق مفهوم الوسط الحسابي في حالتين أساسيتين:

الأولى عند التعامل مع البيانات غير المبوبة، أي تلك التي تُعرض في صورتها الأصلية دون تنظيم في جداول تكرارية، والثانية عند التعامل مع البيانات المبوبة، أي التي تُرتب داخل جداول تكرارية محددة الفئات والتكرارات، وسنرى كيف يختلف أسلوب الحساب في كل حالة مع بقاء الفكرة الأساسية للوسط الحسابي واحدة، وهي تمثيل مركز التوازن الإحصائي للبيانات.

1- حساب الوسط الحسابي للبيانات الغير مبوبة:

إذا اعتبرنا المتغير محل الدراسة بالرمز X وكانت القيم المرصودة لهذا المتغير هي:

$$X_1, X_2, X_3, X_4, , , X_n$$

حيث n يرمز إلى عدد القيم المرصودة، فإن الوسط الحسابي يحسب كالتالي:

$$\bar{X} = \frac{X_1 + X_2 + X_3 + X_4, , , + X_n}{n}$$

وباستخدام الرمز $\sum$ للتعبير عن مجموع القيم، يمكن تبسيط الصيغة إلى:

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

وهذه الصيغة تُستخدم لحساب الوسط الحسابي لبيانات غير مبوبة، أي عندما تكون كل قيمة مرصودة منفردة ومتاحة للمعالجة مباشرة.

مثال (1.3)

مهندس يقوم بقياس أطوال خمسة أعمدة معدنية (بوحدة السنتيمتر) في مصنع، وكانت النتائج كما يلي:

112 ، 120 ، 125 ، 118 ، 122 ، 121

المطلوب: حساب الوسط الحسابي لأطوال الأعمدة.

الحل

المتغير محل الدراســة هنا هو أطوال الأعمدة وســنرمز له بالرمز X بينما القيم المقاســة والتي عددها 5 سنرمز لها بالرمز $X_1, X_2, , , , X_n$.

نقوم بحساب الوسط الحسابي باستخدام المعادلة (2.3).

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

$$\bar{X} = \frac{121 + 122 + 118 + 125 + 120 + 112}{6} = 119.66$$

وهذا يعني أن مركز الثقل، أي النقطة التي يحصـــل عليها عند التوازن هي القيمة 119.66 سنتيمتر كما مبين في الشكل.

مثال (2.3)

البيانات التالية توضح قيمة المبيعات لشـــركة استيراد وتصدير بمئات الألاف في السنوات من عام 2016م إلى عام 2020 وذلك وفقاً لسجلات الشركة.

السنــة	2016	2017	2018	2019	2020
قيمة المبيعات	450	375	550	622	604

المطلوب: حساب الوسط الحسابي لقيمة مبيعات الشركة لهذه السنوات الخمس.

الحــل

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

$$\bar{X} = \frac{450 + 375 + 550 + 622 + 604}{5} = 520.2$$

2- حساب الوسط الحسابي للبيانات المبوبة:

- عندما تكون البيانات مبوّبة في جدول توزيع تكراري، بحيث تمثل كل فئة قيمة واحدة فقط (أي أن المتغير المنفصل يأخذ قيمًا محددة)، يمكن حساب الوسط الحسابي باتباع الخطوات التالية:

- نحسب مجموع القيم لكل فئة:

وذلك بضرب القيمة X_i التي تمثل الفئة في تكرارها f_i أي نحسب:

$$f_i * X_i$$

- نحسب المجموع الكلي للقيم:

بجمع نتائج الخطوة السابقة لجميع الفئات، أي:

$$\sum_{i=1}^{K}(f_i * X_i)$$

حيث k هو عدد الفئات.

- نحسب المجموع الكلي للتكرارات:

وهو يعبّر عن العدد الكلي للمشاهدات، ويحسب كالآتي:

$$\sum_{i=1}^{K} f_i$$

- نحسب الوسط الحسابي:

بقسمة مجموع القيم على مجموع التكرارات:

$$\bar{X} = \frac{\sum_{i=1}^{K}(f_i * X_i)}{\sum_{i=1}^{K} f_i}$$

- عندما تكون البيانات مبوّبة في جدول توزيع تكراري بحيث تمثل كل فئة مجموعة من القيم (أي أن كل فئة تغطي مدى محددًا من القيم وليس قيمة واحدة فقط)، فإننا لا نعرف القيم الدقيقة داخل كل فئة، بل نعرف تكرار كل فئة فقط، ولذلك، نفترض أن القيم داخل كل فئة موزعة توزيعًا متجانسًا حول مركز الفئة، أي أن مركز الفئة

يُمثل القيمة النموذجية لجميع القيم الواقعة داخلها، ولحساب الوسط الحسابي في هذه الحالة نتبع الخطوات التالية:

- حساب مركز كل فئة، ويرمز له بالرمز X_i، ويُحسب بالعلاقة:

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

- نحسب المجموع لكل فئة، وذلك بضرب مركز الفئة (X_i) في تكرارها (f_i):

$$f_i * X_i$$

- نحسب المجموع الكلي للقيم بجمع نواتج الخطوة السابقة لجميع الفئات:

$$\sum_{i=1}^{K} (f_i * X_i)$$

- نحسب المجموع الكلي للتكرارات:

$$\sum_{i=1}^{K} f_i$$

- وأخيرًا نحسب الوسط الحسابي بالعلاقة:

$$\bar{X} = \frac{\sum_{i=1}^{K} (f_i * X_i)}{\sum_{i=1}^{K} f_i}$$

حيث X القيمة التي تمثلها الفئة (عندما تكون الفئة ممثلة بقيمة فقط)، وهي عبارة عن مركز الفئة (عندما تكون الفئة تمثل أكثر من قيمة واحدة).

مثـال (3.3)

تمثل البيانات المنفصلة الموضحة في الجدول التكراري التالي توزيع الفصول على حسب عدد الطلبة في مدرسة ابتدائية بها 20 فصلاً، فما هو الوسط الحسابي لعدد الطلبة في هذه المدرسة.

عدد طلبة الفصل	30	31	32	33	34	35
عدد الفصول	1	2	4	5	6	2

الحـل

بما أن كل فئة في هذا الجدول التكراري تمثل قيمة واحدة فقط، فإنّ X_i تُمثل القيمة المحددة الواردة في الفئة نفسها، أي أنّها ليست مركز فئة بل القيمة الأصلية للمتغير محل الدراسة.

ويحسب الوسط الحسابي من العلاقة $\bar{X} = \frac{\sum_{i=1}^{K}(f_i * X_i)}{\sum_{i=1}^{K} f_i}$ ويمكن إيجاده من خلال تطبيق خطوات المعادلة والموضّـــحة في الجدول (1.3) مباشـــرة على هذا النوع من البيانات دون الحاجة إلى حساب مراكز فئات جديدة.

جدول (1.3): جدول يوضح حساب الوسط الحسابي للبيانات المبوبة.

عدد طلبة الفصول (X_i)	عدد الفصول (f_i)	$f_i * X_i$
30	1	30
31	2	62
32	4	128
33	5	165
34	6	204
35	2	70
المجمـوع	20	659

$$\bar{X} = \frac{\sum_{i=1}^{K}(f_i * X_i)}{\sum_{i=1}^{K} f_i} = \frac{659}{20} = 32.95$$

مثـال (4.3)

سجلت إدارة الجودة داخل مصنع لإنتاج الأنابيب عدد العيوب السطحية في 30 قطعة.

عدد العيوب	التكرار
0	9
1	13
2	6
3	2
المجموع	**30**

الحـل

القيم الملاحَظة مرتبة في فئات قيمة واحدة كما يلي (عدد العيوب = X، التكرار = f).

جدول (2.3): جدول يوضح حساب الوسط الحسابي للبيانات المبوبة.

عدد العيوب	عدد القطع (f_i)	$f_i * X_i$
0	9	0
1	13	13
2	6	12
3	2	6
المجموع	30	31

$$\bar{X} = \frac{\sum_{i=1}^{K}(f_i * X_i)}{\sum_{i=1}^{K} f_i} = \frac{31}{30} = 1.033$$

مثـال (5.3)

تمثل البيانات التالية درجات 50 طالباً في مقرر الاحصاء

الدرجـة	عدد الطلبة
39−30	4
49−40	6
59−50	8
69−60	12
79−70	9
89−80	7
100−90	4

احسب الوسط الحسابي لهذه الدرجـات.

الحـل

بما أن كل فئة في هذا الجدول التكراري تعبر عن مجموعة من القيم وليس قيمة واحدة فستكون X_i عبارة عن مركز الفئة، والجدول (2.3) يوضح خطوات الوصول إلى الحل.

جدول (3.3): جدول يوضح حساب الوسط الحسابي للبيانات المبوبة.

الدرجة	عدد الطلبة (f_i)	مراكز الفئات (X_i)	$f_i * X_i$
39−30	4	34.5	138
49−40	6	44.5	267
59−50	8	54.5	436
69−60	12	64.5	774
79−70	9	74.5	670.5
89−80	7	84.5	591.5
100−90	4	95	380
المجموع	50		3257

$$\bar{X} = \frac{\sum_{i=1}^{K}(f_i * X_i)}{\sum_{i=1}^{K} f_i} = \frac{3255}{50} = 65.14$$

مثـال (6.3)

من المثـــال (12.2) احسب المتوسط الحسابي لتيار التشغيل (بالأمبير) والمسجَّل أثناء اختبار 30 محركًا كهربائيًا في ورشة الصيانة بقسم الهندسة الكهروميكانيكية.

تيار التشغيل(A)	التكرار (عدد المحركات)
4.0 إلى أقل من 4.5	4
4.5 إلى أقل من 5.0	6
5.0 إلى أقل من 5.5	10
5.5 إلى أقل من 6.0	7
6.0 إلى أقل من 6.5	3
المجموع	30

الحـل

باستخدام معادلة الوسط الحسابي $\bar{X} = \frac{\sum_{i=1}^{K}(f_i * X_i)}{\sum_{i=1}^{K} f_i}$ وتطبيق الخطوات في الجدول التـالي.

جدول (4.3): جدول يوضح حساب الوسط الحسابي للبيانات المبوبة.

تيار التشغيل(A)	عدد المحركات (f_i)	مراكز الفئات (X_i)	$f_i * X_i$
4.0 إلى أقل من 4.5	4	4.25	17
4.5 إلى أقل من 5.0	6	4.75	28.5
5.0 إلى أقل من 5.5	10	5.25	52.5
5.5 إلى أقل من 6.0	7	5.75	40.25
6.0 إلى أقل من 6.5	3	6.25	18.75
المجموع	30		157

$$\bar{X} = \frac{\sum_{i=1}^{K}(f_i * X_i)}{\sum_{i=1}^{K} f_i} = \frac{157}{30} = 5.23$$

مثـــال (7.3)

في مصـــــنع تُقاس أطوال المنتجات الحديدية (بالمليمتر) وجُمعت في فئات متســـــاوية الطول، والنتائج كالتالي:

طول المنتج (مم)	عدد القطع
22 – 18	4
27 – 23	7
32 – 28	10
37 – 33	6
42 – 38	3
المجموع	30

المطلوب: حساب الوسط الحسابي لأطوال المنتجات الحديدية.

الحــل

جدول (5.3): جدول يوضح حساب الوسط الحسابي للبيانات المبوبة.

طول المنتج	عدد القطع (f_i)	مراكز الفئات (X_i)	$f_i * X_i$
18 – 22	4	20	80
23 – 27	7	25	175
28 – 32	10	30	300
33 – 37	6	35	210
38 – 42	3	40	120
المجموع	30		885

$$\bar{X} = \frac{\sum_{i=1}^{K}(f_i * X_i)}{\sum_{i=1}^{K} f_i} = \frac{885}{30} = 29.5$$

1.2.3 خواص الوسط الحسابي

1. أكثر مقاييس النزعة المركزية استخدامًا وأهميةً:

يُعَدّ الوسـط الحسـابي أهم مقاييس النزعة المركزية وأكثرها اسـتعمالًا، نظرًا لسـهولة حسـابه ومرونته في المعالجة الرياضية، ولذلك يُعتمد عليه اعتمادًا كبيرًا في التحليل الإحصائي، إذ يدخل في حساب العديد من المقاييس والمؤشرات الإحصائية الأخرى.

2. يشمل جميع القيم ويُمكّن من استعادة مجموعها:

يأخذ الوسط الحسابي في اعتباره جميع القيم المشاهدة دون استثناء، وبذلك يُعدّ مقياسًا شاملًا يعبّر عن مركز ثقل البيانات. كما يمكن من خلاله استنتاج مجموع القيم بسهولة، وذلك بالعلاقة:

مجموع القيم = الوسط الحسابي × عدد القيم

3. قيمة نظرية تعبّر عن التوازن:

الوسـط الحسـابي هو قيمة نظرية تمثل نقطة التوازن بين القيم، وليس بالضـرورة أن تكون إحدى القيم الفعلية في مجموعة البيانات.

فعلى سبيل المثال، إذا كان عدد الطلبة في عدة فصول كما في المثال (3.3)، فوجدنا أن الوسط الحسـابي لعدد الطلبة هو 32.95، فهذه القيمة لا يمكن أن تكون عددًا حقيقيًا للطلبة (لأن عدد الطلبة لا يكون كسريًا)، لكنها تعبّر بدقة عن المتوسط العام للفصول.

4. تأثر الوسط الحسابي بالقيم المتطرفة:

يتأثر الوسط الحسابي بشكل كبير بوجود القيم المتطرفة في البيانات، وقد ينحاز لها.

مثال: تبرعات خمسة عشر فردًا لعمل خيري بالمبالغ التالية (بالدينار):

10,15,15,20,15,15,12,20,10,10,18,20,10,1000,80

$$\text{الوسط الحسابي} = \frac{80+1000+10+20+18+10+10+20+12+15+15+20+15+15+10}{15}$$

الوسط الحسابي = 84.66

نلاحظ أن الوسط الحسابي أصبح كبيرًا جدًا بسبب القيمة المتطرفة 1000، ولا يعكس ما دفعه معظم الأفراد، وبالتالي يُعد مقياسًا مضللاً في هذه الحالة، أما إذا استُبعدت القيمة المتطرفة، فسيصبح الوسط الحسابي قريبًا من المبالغ الفعلية لمعظم المتبرعين.

5. الوسط الحسابي لا يُحسب للبيانات النوعية (الوصفية):

لأنه يعتمد على القيم العددية، فلا يمكن تطبيقه على البيانات مثل الألوان أو العلامات الوصفية.

6. الوسط الحسابي لا يُستنتج مباشرة من الرسم البياني:

لا يمكن وضعه أو تحديده بشكل دقيق من خلال الرسم البياني فقط، لأنه يمثل قيمة عددية نظرية.

7. الوسط الحسابي لا يُحسب للبيانات المفتوحة في جداول التكرار:

في حالة وجود فئات مفتوحة (مثل "أعلى من 50")، لا يمكن تحديد مركز الفئة بدقة، وبالتالي لا يمكن حساب الوسط الحسابي.

8. مجموع الانحرافات عن الوسط الحسابي يساوي صفرًا:

حيث يُعرف انحراف كل قيمة عن الوسـط بأنه القيمة مطروحًا منها الوسـط الحسـابي. بالتالي، لدينا:

$$\sum_{i=1}^{n}(X_i - \bar{X}) = 0$$

الوسـط الحسـابي للبيانات 6 ، 8 ، 5 ، 11 ، 10 هو 8، وهو مجموع هذه القيم على عددها، فسنجد كما هو مبين في الجدول (6.3):

جدول (6.3): جدول يوضح حساب الانحرافات عن الوسط الحسابي.

X_i القيمة	$X_i - \bar{X}$ الانحرافـات
6	6−=8−2
8	8−8=0
5	5−=8−3
11	11−8=3
10	10−8=2
المجمـوع	صفر

9. الوسط الحسابي يقلل مجموع مربعات الانحرافات:

مجموع مربعات الانحرافات عن الوسط الحسابي أقل ما يمكن مقارنة بأي قيمة أخرى، أي أن:

$$\sum_{i=1}^{n}(X_i - \bar{X})^2 = \text{نهاية صغرى}$$

مثال: أوزان قطع معدنية (بالغرام):

50,55,60,65,70

المتوسط الحسابي لها هو 60 غرام، والجدول (7.3) يوضح هذه الخاصية.

جدول (7.3): جدول يوضح حساب الانحرافات عن الوسط الحسابي ومربع هذه الانحرافات.

الوزن	انحرافات القيم عن الوسط الحسابي	مربع الانحرافات عن الوسط الحسابي
50	10−	100
55	5−	25
60	0	0
65	5	25
70	10	100
المجموع	صفر	250

ويعني ذلك أنه لو اخترنا أي قيمة أخرى — سواء كانت أقل من الوسط الحسابي أو أكبر منه — فإن مجموع مربعات انحرافات القيم عن تلك القيمة سيكون دائمًا أكبر من مجموع مربعات انحرافات القيم عن الوسط الحسابي نفسه.

بمعنى آخر، الوسط الحسابي هو النقطة المثالية للتوازن الإحصائي، إذ يجعل تباعد القيم عنه في أدنى حد ممكن، ولهذا السبب يُستخدم الوسط الحسابي بكثرة في التطبيقات الهندسية والتحكم الصناعي، لأنه يُمثّل القيمة التي تقلل الخطأ الكلي في البيانات إلى الحد الأدنى.

فلو جربنا قيمة أخرى بدلاً عن الوسط الحسابي مثل 62 نجد في الجدول (8.3) الآتي:

جدول (8.3): جدول يوضح حساب الانحرافات عن قيمة أخرى ومربع هذه الانحرافات.

الوزن	انحرافات القيم عن 62	مربع الانحرافات عن 62
50	12−	144
55	7−	49
60	2−	4
65	3	9
70	8	64
المجموع	صفر	270

نلاحظ أن مجموع مربعات الانحرافات أكبر من مجموع مربعات الانحرافات عند الوسط الحسابي (250 < 270).

3.3 الوسط الحسابي المرجح:

عندما تكون القيم المشاهدة ليست متساوية في الأهمية أو التأثير، فإن استخدام الوسط الحسابي العادي لا يعطي تمثيلاً دقيقاً للبيانات، في هذه الحالة، يجب أن نُعطي لكل قيمة وزنًا يعكس أهميتها أو تكرارها النسبي، ومن هنا يأتي مفهوم الوسط الحسابي المرجَّح، والذي يُرمز له عادة بالرمز (W).

بصيغة أخرى، الوسط المرجَّح هو المتوسط الذي يأخذ في الاعتبار وزن كل قيمة، بحيث تكون القيم ذات الأهمية الأكبر أكثر تأثيراً في النتيجة النهائية.

فإذا كانت لدينا القيم:

$$X_1, X_2, X_3, X_4 \ldots\ldots X_n$$

ولكل قيمة وزن يقابلها:

$$W_1, W_2, W_3, W_4 \ldots\ldots W_n$$

فإن صيغة الوسط الحسابي المرجّح تكتب على النحو التالي:

$$\bar{X} = \frac{\sum_{i=1}^{n}(X_i * W_i)}{\sum_{i=1}^{n} W_i}$$

ونلاحظ أن صيغة الوسط الحسابي المرجَّح تشبه تماماً صيغة الوسط الحسابي للبيانات المبوبة في جدول تكراري، إذ يعتمد كلاهما على مبدأ الوزن النسبي لكل فئة أو مجموعة من القيم.

كما أنه في حال كانت لدينا عدة مجموعات من البيانات، بحيث تحتوي كل مجموعة على عدد مختلف من القيم، وكان معروفاً الوسط الحسابي لكل مجموعة، فإننا لا نستطيع حساب الوسط الحسابي العام لجميع البيانات بمتوسط بسيط لهذه القيم، لأن عدد المشاهدات في كل مجموعة يختلف.

لذلك، عند دمج هذه المجموعات معاً للحصول على الوسط الحسابي العام، يجب أن يُعامل عدد القيم في كل مجموعة باعتباره وزنها أو أهميتها النسبية.

ويُحسب الوسط الحسابي العام لجميع المجموعات بالصيغة التالية:

الوسط الحسابي للمجموعات:

$$\bar{X}_1, \bar{X}_2, \bar{X}_3 \ldots\ldots \bar{X}_n$$

عدد مفردات المجموعات (ألوزن):

$$n_1, n_2, n_3 \ldots\ldots n_k$$

مثـــال (8.3)

في أحد المصانع، تم إنتاج ثلاثة أنواع من القطع الميكانيكية تستخدم في خطوط التجميع:

تكلفة تصنيع القطعة من النوع الأول = 26 دينارًا.

تكلفة تصنيع القطعة من النوع الثاني = 20 دينارًا.

تكلفة تصنيع القطعة من النوع الثالث = 6 دينار.

علمًا بأن أهمية النوع الثاني في العملية الإنتاجية تســاوي ضــعف أهمية النوع الأول، بينما النوع الثالث يمثل خمسة أمثال أهمية النوع الأول من حيث الاستخدام أو الطلب في خط الإنتاج.

احسب الوسط الحسابي المرجح لتكلفة القطعة الواحدة لجميع الأنواع.

الحـــل

من صــيغة المثال يتضــح أن أهمية كل منتج تختلف عن الآخر، لذلك ســنلجأ لحســاب الوســط الحسابي المرجَح والموضح في الجدول (9.3).

جدول (9.3): جدول يوضح حساب الوسط الحسابي المرجّح.

	النوع الأول	النوع الثاني	النوع الثالث
تكلفة التصنيع X_i	26	20	6
الأهمية W_i	1	2	5

وبتطبيق صيغة الوسط الحسابي المرجّح:

$$\bar{X} = \frac{\sum_{i=1}^{n}(X_i * W_i)}{\sum_{i=1}^{n} W_i} = \frac{(26x1) + (20x2) + (6x5)}{1 + 2 + 5} = 12$$

مثـــال (9.3)

أحد الطلبة في قســم الهندســة الصــناعية حصــل على الدرجات التالية في أربعة مقررات خلال الفصل الدراسي، كما هو موضح في الجدول:

اسم المقرر	درجة الطالب X_i	عدد الوحدات W_i
رياضيات هندسية	80	4
انتقال حرارة	75	3
إحصاء هندسي	90	2
رسم هندسي باستخدام الحاسوب	95	1

الحــل

$$\bar{X} = \frac{\sum_{i=1}^{n}(X_i * W_i)}{\sum_{i=1}^{n} W_i} = \frac{(80x4) + (75x3) + (90x2) + (95x1)}{4 + 3 + 2 + 1} = 82$$

4.3 الوسيط:

الوسـيط هو القيمة التي تقع في منتصـف مجموعة البيانات بعد ترتيبها تصـاعديًا (من الأصـغر إلى الأكبر) أو تنازليًا (من الأكبر إلى الأصـغر)، ويمثل الوسـيط النقطة التي تقسـم البيانات إلى نصـفين متساويين؛ حيث يكون نصـف القيم أقل من الوسيط والنصـف الآخر أكبر منه ويرمز له عادة بالرمز m .

1- في حالة البيانات غير المبوبة (البيانات الخام):

لحساب الوسيط في هذه الحالة نتبع الخطوات التالية:

- ترتيب البيانات ترتيبًا تصاعديًا، بحيث تبدأ بأصغر قيمة وتنتهي بأكبر قيمة، أو بالعكس "تنازلياً".
- تحديد ترتيب الوسيط باستخدام العلاقة التالية:

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

حيث (n) تمثل عدد القيم في مجموعة البيانات.

ملاحظة: الرتبة هنا لا تعني القيمة نفسها، بل موقع الوسيط بين القيم بعد ترتيبها.

- تحديد قيمة الوسيط:

إذا كان عدد القيم فرديًا، فإن الوسيط هو القيمة التي ترتيبها (n+1/2) .

أما إذا كان عدد القيم زوجيًا، فإن الوسيط سيقع بين قيمتين ويكون هو الوسط الحسابي للقيمتين الواقعتين في الترتيبين.

مثـــال (10.3)

أوجد قيمة الوسيط للبيانات التـالية:

25 ، 29 ، 26 ، 23 ، 27 ، 28 ، 30

الحـــل

القيم مرتبة ترتيباً تصاعدياً:

23 ، 25 ، 26، 27 ، 28 ، 29 ، 30

الوسيط

ترتيب الوسيط = (عدد البيانات + 1 \ 2) = (7+1 \ 2) = 4.

أي أن ترتيب أو رتبة الوسيط هي القيمة الرابعة في البيانات بعد ترتيب البيانات وبالتالي فإن الوسيط هو 17.

مثـــال (11.3)

أوجد قيمة الوسيط للبيانات التـالية:

21 ، 15 ، 16 ، 19 ، 18 ، 13

الحـــل

القيم مرتبة ترتيباً تصاعدياً:

21 ، 19 ، 16،18 ، 15 ، 13

ترتيب الوسيط = (عدد البيانات + 1 \ 2) = (6+1 \ 2) = 3.5.

أي أن قيمة الوســـيط تقع بين القيمة الثالثة والرابعة في ترتيب هذه البيانات، إذا فإن الوسيط هو الوسط الحسابي لهاتين القيمتين

$$m = \frac{16 + 18}{2} = 17$$

2- حساب الوسيط في حالة البيانات المبوبة:

عند التعامل مع البيانات المبوبة (أي التي تُعرض في جداول تكرارية بفئات مرتبة تصـــاعديًا أو تنازليًا)، نحسب الوسيط باتباع الخطوات التالية:

- تحديد رتبة الوسيط:

نحسب رتبة الوسيط من خلال العلاقة:

$$\frac{\sum_{i=1}^{k} f_i}{2}$$

حيث f_i تمثل مجموع التكرارات الكلي.

- تحديد الفئة الوسيطية:

الفئة الوسيطية هي الفئة التي يقع فيها الوسيط، ويتم تحديدها باستخدام التكرار المتجمع الصـاعد، بحيث تكون هذه الفئة هي أول فئة تكرارها المتجمع الصـاعد أكبر من أو يسـاوي رتبة الوسيط.

- تطبيق قانون الوسيط للبيانات المبوبة:

بعد تحديد الفئة الوسيطية، نحسب قيمة الوسيط باستخدام القانون التالي:

$$m = L + \left(\frac{\frac{\sum_{i=1}^{k} f_i}{2} - F}{f_m} \right) x\ C$$

حيث:

L : الحد الأدنى للفئة الوسيطية، وفي حالة التوزيعات الخاصة بمتغير منفصل، نستخدم الحد الأدنى الحقيقي للفئة الوسيطية، وهو عبارة عن الحد الأدنى للفئة مطروحاً منه نصف وحدة من وحدات القياس المستخدمة فمثلاً:

إذا كانت الفئة الوسيطية تبدأ بالقيمة 20، فإن الحد الأدنى الحقيقي يكون 19.5 وهكذا.

F : التكرار المتجمع الصاعد للفئة السابقة للفئة الوسيطية.

f_m: تكرار الفئة الوسيطية نفسها.

C : طول الفئة الوسيطية (أي الفرق بين الحد الأعلى والحد الأدنى للفئة).

الأساس الرياضي لقانون الوسيط:

يعتمد قانون الوسيط على افتراض أساسي مفاده أن القيم داخل كل فئة موزعة توزيعًا منتظمًا حول مركزها، أي أن البيانات داخل الفئة الوسيطية تتوزع بشكل متساوٍ، وبناءً على هذا الافتراض، يمكن القول إن هناك تناسبًا بين:

المسافة التي يبعدها الوسيط عن الحد الأدنى للفئة الوسيطية L، وطول الفئة الوسيطية C من جهة، وبين عدد القيم الواقعة قبل الوسيط داخل الفئة الوسيطية (رتبة الوسيط $- F$)، والتكرار الكلي للفئة الوسيطية f_m، من الجهة الأخرى، وبتطبيق هذا التناسب، نحصل على الصيغة الرياضية العامة لقانون الوسيط.

مثال (12.3)

احسب الوسيط للبيانات التالية التي تمثل أوزان قطع غيار ميكانيكية (بالجرام).

الوزن (بالجرام)	عدد القطع
200 إلى أقل من 250	15
250 إلى أقل من 300	12
300 إلى أقل من 350	18
350 إلى أقل من 400	15
المجموع	60

الحـــل

الخطوة الأولى هي إيجاد ترتيب الوسيط باستخدام المعادلة التالية:

$$\frac{\sum_{i=1}^{k} f_i}{2} = \frac{60}{2} = 30$$

من ثم تقوم بتحديد الفئة الوسيطية " أي الفئة التي تحتوي على الوسيط" وهي القيمة التي ترتيبها 30 وذلك من خلال التكرار المتجمع الصاعد وهو كما موضح في الجدول التالي:

جدول (10.3): جدول يوضح حساب التكرار المتجمع الصاعد وتحديد الفئة الوسيطية.

الوزن (بالجرام) الفئة	عدد القطع (التكرار)	التكرار المتجمع الصاعد	
أقل من 250	15	15	
أقل من 300	12	27	
أقل من 350	18	45	← الفئة الوسيطية
أقل من 400	15	60	

بما أن التكرار المتجمع للفئة الثانية يســاوي 27، وللفئة الثالثة يســاوي 45، فهذا يعني أن القيم الثمانية عشـــر (18) الواقعة ضـــمن الفئة الثالثة تمثل الترتيبات من 28 إلى 45 في مجموعة البيانات الكلية.

وحيث إن ترتيب الوسيط = 30، فإن هذه القيمة تقع ضمن الفئة الثالثة، أي أن الفئة الثالثة هي الفئة الوسيطية.

بعد تحديد الفئة الوسيطية، نطبّق قانون حساب الوسيط وفق المعادلة التالية:

$$m = L + \left(\frac{\frac{\sum_{i=1}^{k} f_i}{2} - F}{f_m} \right) x\ C$$

$$f_m = 18 \quad , \quad C = 50 \quad , \quad L = 300 \quad , \quad F = 27$$

$$m = 300 + \left(\frac{30 - 27}{18} \right) x\ 50 = 308.33$$

وهذا يعني أن نصف هذه القطع غيار المنتجة وزنها أقل من 308.33 جرام، والنصف الآخر منها أكثر من 308.33 جرام.

مثـــال (13.3)

يُبين الجدول التالي الزمن اللازم (بالدقائق) لإتمام اختبار متانة مجموعة من القطع المعدنية المصنعة في ورشة هندسية، وذلك لعينة مكونة من 100 قطعة:

الفئة (زمن الاختبار بالدقائق)	عدد القطع (التكرار)
10 إلى أقل من 20	8
20 إلى أقل من 30	12
30 إلى أقل من 40	20
40 إلى أقل من 50	35
50 إلى أقل من 60	15
60 إلى أقل من 70	10
المجمـوع	100

احسب الوسيط لزمن الاختبـار.

الحـــل

الخطوة الأولى هي إيجاد ترتيب الوسيط = (100\2) = 50

من ثم تقوم بتحديد الفئة الوسيطية من خلال التكرار المتجمع الصاعد.

جدول (11.3): جدول يوضح حساب التكرار المتجمع الصاعد وتحديد الفئة الوسيطية.

زمن الاختبار (بالدقائق) الفئة	عدد القطع (التكرار)	التكرار المتجمع الصاعد	
أقل من 20	8	8	
أقل من 30	12	20	
أقل من 40	20	40	
أقل من 50	35	75	← الفئة الوسيطية
أقل من 60	15	90	
أقل من 70	10	100	

بما أن التكرار المتجمع للفئة الثالثة يســاوي 40، وللفئة الرابعة يســاوي 75، فهذا يعني أن القيم الخمسة والثلاثون (35) الواقعة ضمن الفئة الرابعة تمثل الترتيبات من 41 إلى 75 في مجموعة البيانات الكلية.

وحيث إن ترتيب الوسيط = 50، فإن هذه القيمة تقع ضمن الفئة الرابعة، أي أن الفئة الرابعة هي الفئة الوسيطية، من ثم نطبّق قانون حساب الوسيط وفق المعادلة التالية:

$$m = L + \left(\frac{\frac{\sum_{i=1}^{k} f_i}{2} - F}{f_m} \right) x\ C$$

$$f_m = 15 \quad , \quad C = 10 \quad , \quad L = 40 \quad , \quad F = 40$$

$$m = 40 + \left(\frac{50 - 40}{15} \right) x\ 10 = 56.6$$

مثــال (14.3)

الجدول التكراري التالي يبين التوزيع التكراري لدخل ستين موظفاً في شركة:

الدخل	عدد الموظفين
1500 إلى أقل من 1750	8
1750 إلى أقل من 2000	15
2000 إلى أقل من 2250	20
2250 إلى أقل من 2500	12
2500 فأكثر	5

فاحسب الوسيط لهذه البيانات.

الحــل

لاحظ أن هذا الجدول التكراري يحتوي على فئة أخيرة مفتوحة، ومع ذلك يمكن حســاب الوســيط بنفس الطريقة المتبعة للفئات الأخرى، وذلك كما يلي:

جدول (12.3): جدول يوضح حساب التكرار المتجمع الصاعد وتحديد الفئة الوسيطية.

الدخـل (الدينار) الفئة	عدد الموظفين (التكرار)	التكرار المتجمع الصاعد
أقل من 1750	8	8
أقل من 2000	15	23
أقل من 2250	20	43 ← الفئة الوسيطية
أقل من 2500	12	55
2500 فأكثر	5	60

ترتيب الوسيط هو = (2\60) = 30

من ثم تقوم بتحديد الفئة الوسـيطية من خلال التكرار المتجمع الصـاعد في الجدول (12.3)، حيث أن ترتيب الوسيط 30 سيكون ضمن الفئة الثالثة أي " أقل من 2250"، والتي تحتوي في الأسـاس على عدد بيانات يبدأ من القيمة الرابعة والعشـرون (24) حتى تصـل إلى الثالثة والأربعون (43).

$$f_m = 20 \quad , \quad C = 250 \quad , \quad L = 2000 \quad , \quad F = 23$$

$$m = L + \left(\frac{\frac{\sum_{i=1}^{k} f_i}{2} - F}{f_m} \right) x\ C$$

$$m = 2000 + \left(\frac{30 - 23}{20} \right) x\ 250 = 2087.5$$

مثـــال (15.3)

في فصـل هندسـة ميكانيكية، سـجل عدد الطلاب الذين حصـلوا على درجات مختلفة في اختبار الرسم الهندسي كما يلي:

الدرجة	عدد الطلاب	التكرار المتجمع الصاعد
59–50	5	5
69–60	8	13
79–70	12	25
89–80	10	35
99–90	5	40

المطلوب :حساب الوسيط للدرجات.

الحـــل

ترتيب الوسيط في هذه البيانات هو (20=2\40).

نبحث في الجدول التكراري المتجمع الصاعد عن الفئة الوسيطية والتي تحتوي على الترتيب 20، أي التكرار الصاعد 20، وهي ستكون الفئة الثالثة "70−79".

وباستخدام المعادلة:

$$m = L + \left(\frac{\frac{\sum_{i=1}^{k} f_i}{2} - F}{f_m} \right) x\ C$$

$$f_m = 12 \quad , \quad C = 10 \quad , \quad L = 69.5 \quad , \quad F = 13$$

$$m = 69.5 + \left(\frac{20 - 13}{12} \right) x\ 10 = 75.33$$

تحديد الوسيط بيانياً:

يمكن تحديد الوسيط بيانياً باستخدام المنحنى التكراري المتجمع الصاعد أو المنحنى التكراري المتجمع الهابط، أو باستخدامهما معاً كما يلي:

- نحدد ترتيب الوسيط على المحور الرأسي الذي يمثل التكرارات المتجمعة.
- من هذه النقطة، نرسم خطاً أفقياً موازياً لمحور السينات حتى يلتقي بالمنحنى المتجمع الصاعد أو المنحنى المتجمع الهابط.
- من نقطة التلاقي، نسقط عموداً عمودياً ليقطع محور السينات، فتكون نقطة القطع هي قيمة الوسيط.
- يمكن أيضاً استخدام المنحنيين الصاعد والهابط معاً في رسم واحد؛ ففي هذه الحالة يكون الإحداثي السيني لنقطة تقاطع المنحنيين هو قيمة الوسيط، أي القيمة الوسطى بين القيم التراكمية، وهو ما يُعرف بالوسيط عند دراسة المنحنى المتجمع الصاعد والمتجمع الهابط معاً.

مثـال (16.3)

لنفترض أن لدينا بيانات حول عدد الساعات الأسبوعية التي يقضيها 30 طالب في دراسة مقرراته هذا الفصل، موزعة في جدول تكراري متجمع صاعد كما يلي:

عدد الساعات	عدد الطلبة
0 إلى أقل من 5	4
5 إلى أقل من 10	6
10 إلى أقل من 15	10
15 إلى أقل من 20	8
20 إلى أقل من 25	2

المطلوب إيجاد الوسيط بيانياً

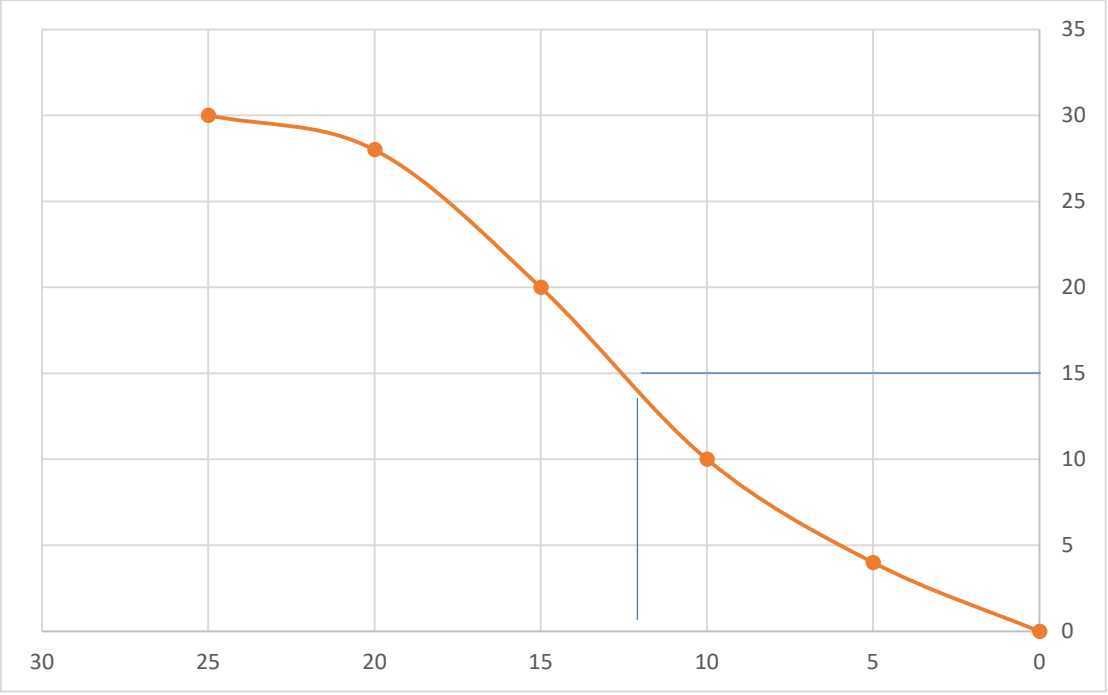
الحـل

يتم أولاً إنشاء جدول التكرار المتجمع الصاعد أو المتجمع الهابط للبيانات، كما هو موضح في جدول (13.3)، من ثم رسم المنحنى المتجمع، وبعدها نقوم بتحديد ترتيب الوسيط، والذي في هذا المثال يساوي (30\2=15) على المحور الرأسي (محور التكرارات المتجمعة).

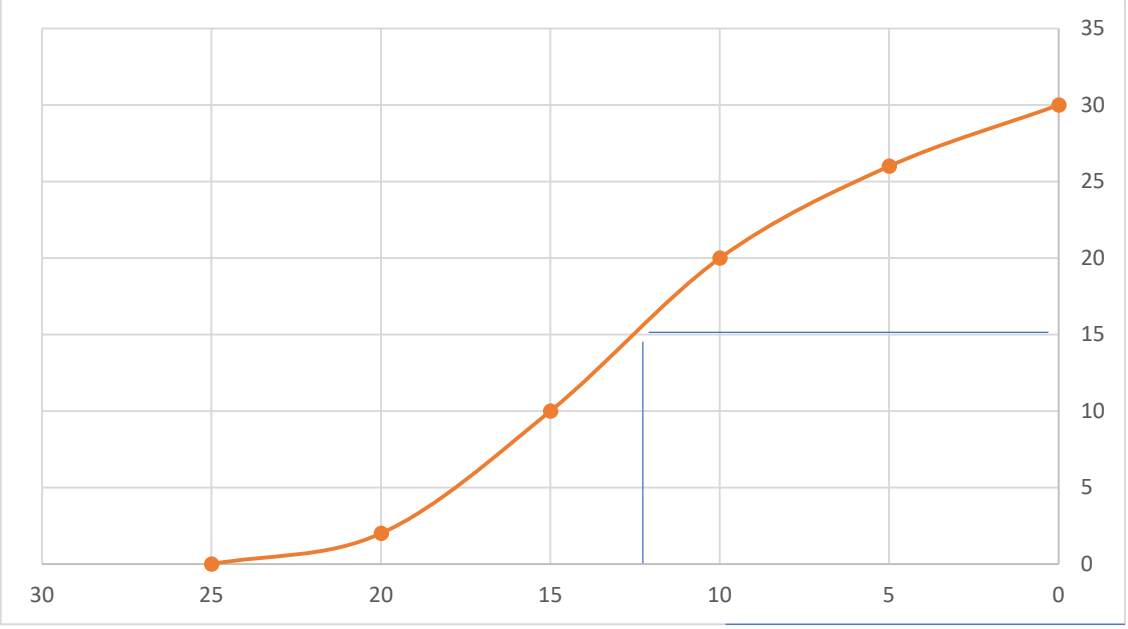
ثم نرسم من هذه النقطة خطاً أفقياً يوازي محور السينات حتى يلتقي بالمنحنى المتجمع، ومن نقطة التلاقي نسقط خطاً عمودياً على محور السينات لتحديد قيمة الوسيط، وهكذا نحصل على قيمة الوسيط بيانياً.

جدول (13.3): جدول يوضح حساب التكرار المتجمع الصاعد والهابط.

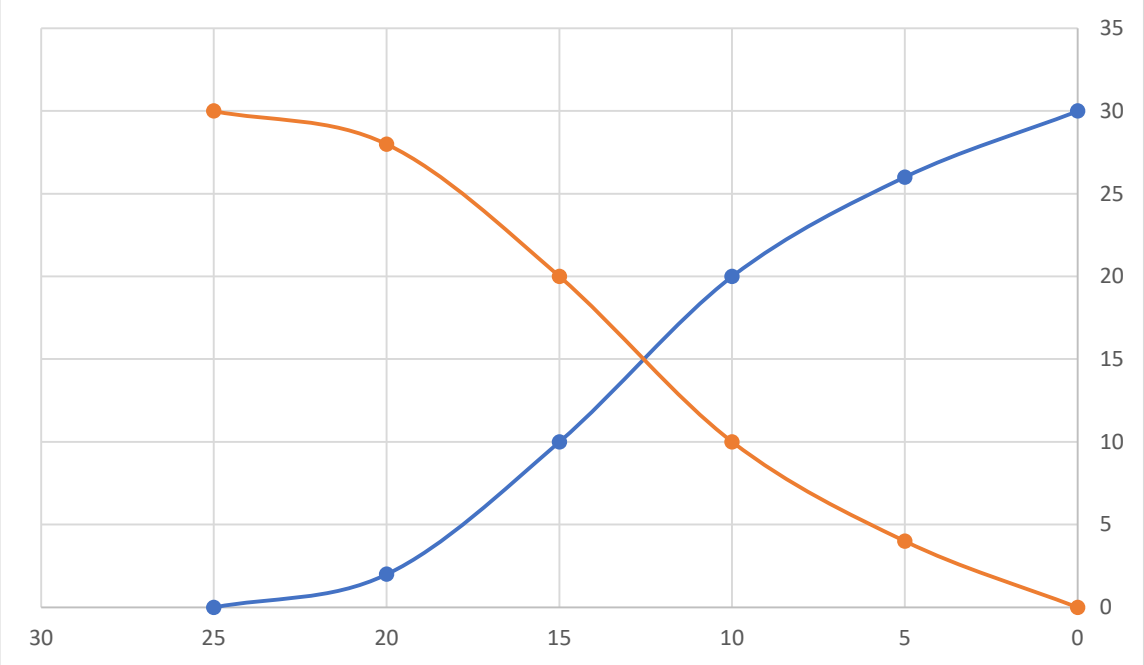
عدد الساعات الفئة	التكرار المتجمع الصاعد	عدد الساعات الفئة	التكرار المتجمع الهابط
أقل من صفر	0	صفر أو أكثر	30
أقل من 5	4	أكثر من 5	26
أقل من 10	10	أكثر من 10	20
أقل من 15	20	أكثر من 15	10
أقل من 20	28	أكثر من 20	2
أقل من 25	30	أكثر من 25	0



الشكل (1.3): يوضح المنحنى التكراري الصاعد وإيجاد الوسيط بيانياً.



الشكل (2.3): يوضح المنحنى التكراري الهابط وإيجاد الوسيط بيانياً.



الشكل (3.3): يوضح المنحنى التكراري الصاعد والهابط وإيجاد الوسيط بيانياً.

من الأشكـــــال السابقة (1.3) (2.3) (3.3) نجد أن قيمة الوسيط بيانياً هي 12.5 تقريبا، كما يمكن تحديدها باستخدام المعادلة الرياضية للتأكيد عليها.

1.4.3 خواص الوسيط:

1. سهولة التعريف والحساب:

الوسيط من أبسط مقاييس النزعة المركزية حيث يمكن تحديده بسرعة من البيانات.

2. يعتمد فقط على القيمة أو القيم الوسطية:

إذا كان عدد القيم فردياً، فالوسيط هو القيمة الوسطى.

إذا كان عدد القيم زوجياً، فالوسيط هو متوسط القيمتين الوسطيتين، وباقي القيم لا تؤثر في حساب الوسيط.

3. مقاومة للقيم المتطرفة:

الوسيط لا يتأثر بالقيم الشاذة أو الكبيرة جدًا مقارنة ببقية البيانات، على عكس الوسط الحسابي.

مثال توضيحي:

لدينا بيانات عدد ساعات العمل على مشروع هندسي لمجموعة طلبة:

9,1,10,4,6,7,2

- ترتيب البيانات تصاعدياً:

10،9،7،6،4،2،1

- عدد القيم فردي (7)، إذن الوسيط هو القيمة الوسطى = 6

لاحظ كيف أن القيم الصغيرة جدًا (1) والكبيرة (10) لم تؤثر على الوسيط، إذا عدلنا القيمة الكبرى في المثال السابق من 10 إلى 84، فسنلاحظ أن الوسيط لم يتغير، وظل 6، هذا يوضح أن الوسيط لا يتأثر بالقيم الشاذة أو المتطرفة.

4. يمكن حسابه للبيانات النوعية (الوصفية):

شرط إمكانية ترتيب البيانات تصاعدياً أو تنازلياً وأن يكون عددها فردياً.

مثال: تقديرات 9 طلبة في مادة الإحصاء: مقبول، جيد، ضعيف، ممتاز، جيد جداً، مقبول، مقبول، جيد، مقبول.

- ترتيب البيانات تصاعدياً:

ضعيف، مقبول، مقبول، مقبول، مقبول، جيد، جيد، جيد جداً، ممتاز

- ترتيب الوسيط هو (9+1\2 = 5)
- الوسيط هو التقدير الخامس "مقبول".

5. يمكن حسابه للجداول التكرارية المفتوحة كما هو موضح في المثال (14.3).
6. يمكن إيجاده بيانياً باستخدام المنحنيات التكرارية المتجمعة الصاعدة أو الهابط.
7. يُستخدم الوسيط في حالة وجود بيانات ناقصة.

مثال: إذا أردنا معرفة المدة التي يقضيها العامل لإنتاج سلعة معينة، يمكن تسجيل مدة 50% من العاملين فقط؛ لأن المدة التي سيستغرقها النصف الآخر ستكون أكبر من الوسيط، وبالتالي يوفر الوسيط تقديراً وسطياً حتى مع البيانات الناقصة.

5.3 المنوال:

المنوال هو القيمة أو الفئة التي تتكرر أكثر عدد من المرات في مجموعة البيانات، أي القيمة الأكثر شيوعاً بين القيم المشاهدة.

- للبيانات النوعية أو الوصفية: هو القيمة الأكثر حدوثاً.
- للبيانات الكمية: هو القيمة أو الفئة التي تحتوي على أكبر تكرار.

1- في حالة البيانات الغير مبوبة:

في البيانات غير المبوبة، أي أن كل قيمة مسجلة بشكل فردي، ولإيجاد المنوال في هذه الحالة نتبع الخطوات التالية:

- نعد عدد مرات ظهور كل قيمة في مجموعة البيانات.
- نحدد القيمة التي ظهرت أكثر من غيرها، فهي المنوال.
- إذا تكررت أكثر من قيمة بنفس العدد الأعلى، يكون لدينا منوال ثنائي " في حالة عددين" أو متعدد " في حالة أكثر من عددين"، وقد يكون أحياناً لا يوجد منوال للبيانات "عديمة المنوال".

مثـال (17.3)

احسب المنوال للبيانات التاليـة:

6 ، 8 ، 6 ، 5 ، 7 ، 6 ، 9 ، 5 ، 6 ، 8 ، 5 ، 6

الحـل

إن المنوال لهذه القيم هي 6، حيث أنها القيمة الأكثر تكراراً في هذه البيانات، وهي في هذه الحالة وحيدة المنوال.

مثـال (18.3)

احسب المنوال للبيانات التاليـة:

6 ، 9 ، 7 ، 9 ، 6 ، 9 ، 9 ، 6 ، 8 ، 5 ، 6

الحـل

في هذه الحالة نجد أن القيمتين 6،9 هما الأكثر تكراراً بأربع مرات، وبالتالي هما المنوال لهذه البيانات "منوال ثنائي".

مثـال (19.3)

أوجد المنوال لكل مجموعة من المجموعات التالية:

المجموعة الأولى: 22، 18، 20، 17، 19، 21

المجموعة الثانية: 14، 12، 14، 16، 12

المجموعة الثالثة: 11، 11، 13، 12، 13، 12

المجموعة الرابعة: 15، 18، 17، 16، 15، 19، 15

المجموعة الخامسة: 8، 7، 10، 9، 7، 8، 10

الحل

المجموعة الأولى:

القيم: 22، 18، 20، 17، 19، 21

كل قيمة تظهر مرة واحدة " لا يوجد منوال ".

المجموعة الثانية:

القيم: 14، 12، 14، 16، 12

التكرارات: 14 ← 2 مرات، 12 ← 2 مرات، 16 ← 1 مرة

المنوال = 12 و 14 " منوال ثنائي ".

المجموعة الثالثة:

القيم: 11، 11، 13، 12، 13، 12

التكرارات: 11 ← 2 مرات، 12 ← 2 مرات، 13 ← 2 مرات

كل القيم تكررت بنفس التكرار، " منوال متعدد "

المجموعة الرابعة:

القيم: 15، 18، 17، 16، 15، 19، 15

التكرارات: 15 ← 3 مرات، 16 ← 1 مرة، 17 ← 1 مرة، 18 ← 1 مرة، 19 ← 1 مرة

المنوال = 15 " منوال وحيد ".

المجموعة الخامسة:

القيم: 8، 7، 10، 9، 7، 8، 10

التكرارات: 7 ← 2 مرات، 8 ← 2 مرات، 10 ← 2 مرات، 9 ← 1 مرة

المنوال = 7، 8، 10 " منوال متعدد ".

مثــال (20.3)

لدى 12 طالباً تقييم في مادة ما، والتقييمات كانت كالتالي:

جيد جداً، جيد، ممتاز، ضعيف، جيد، مقبول، ممتاز، جيد جداً، جيد، ضعيف، مقبول، جيد

الحـل

نقوم بعمل جدول تكراري لهذه البيانات كما في الجدول (14.3):

جدول (14.3): جدول تكراري يوضح التكرار لتقديرات الطلبة.

التقدير	عدد الطلاب
ممتاز	2
جيد جدا	2
جيد	4
مقبول	2
ضعيف	2

المنوال هو الفئة التي تكررت أكثر عددًا من المرات، وهي "جيد" تكررت 4 مرات.

المنوال = جيد.

2- في حالة البيانات المبوبة:

لحساب المنوال في الجداول التكرارية التي تكون فئاتها متساوية الطول، نتبع الخطوات التالية:

- تحديد الفئة المنوالية:

نبحث عن الفئة التي تحتوي على أكبر تكرار، وتسمى هذه الفئة الفئة المنوالية.

- حساب قيمة المنوال:

بعد تحديد الفئة المنوالية، يمكن إيجاد قيمة المنوال بدقة باستخدام الصيغة التالية:

$$M = L + \frac{\Delta_1}{\Delta_1 + \Delta_2}\ x\ C$$

حيث:

L : الحد الأدنى للفئة المنوالية، (أو الحد الأدنى الحقيقي في حالة البيانات التي تمثل متغيراً منفصلاً).

Δ_1 : تكرار الفئة المنوالية – تكرار الفئة السابقة لها.

Δ_2 : تكرار الفئة المنوالية – تكرار الفئة اللاحقة لها.

C : طول الفئة المنوالية (أي الفرق بين الحد الأعلى والحد الأدنى للفئة).

في حالة كان الجدول التكراري غير منتظم فإنه يجب تعديل التكرارات قبل تطبيق الخطوات السابقة.

مثـــال (21.3)

احسب قيمة المنوال للبيانات التالية التي تمثل أوزان قطع غيار ميكانيكية (بالجرام).

الوزن (بالجرام)	عدد القطع
200 إلى أقل من 250	15
250 إلى أقل من 300	12
300 إلى أقل من 350	18
350 إلى أقل من 400	15
المجموع	60

الحـــل

الفئة المنوالية هي الفئة التي لها أكبر تكرار وهي الفئة الثالثة (300 إلى أقل من 350) ونجد أن:

$$L = 300 \quad , \quad C = 50 \ , \quad \Delta_1 = 18 - 12 = 6 \quad , \quad \Delta_2 = 18 - 15 = 3$$

$$M = L + \frac{\Delta_1}{\Delta_1 + \Delta_2} \, x \, C$$

$$M = 300 + \frac{6}{6 + 3} \, x \, 50 = 333.33$$

مثــال (22.3)

في فصــل هندســة ميكانيكية، ســجل عدد الطلاب الذين حصــلوا على درجات مختلفة في اختبار الرسم الهندسي كما يلي:

الدرجة	عدد الطلاب
59–50	5
69–60	8
79–70	12
89–80	10
99–90	5

احسب قيمة المنوال لهذه البيانات.

الحــل

الفئة المنوالية في هذا المثـال هي الفئة (70–79) وهي التي لها أكبر عدد من التكرارات.

هنا البيانات تمثل قيم منفصلة.

$$L = 69.5 \quad , \quad C = 10 \ , \quad \Delta_1 = 12 - 8 = 4 \quad , \quad \Delta_2 = 12 - 10 = 2$$

$$M = L + \frac{\Delta_1}{\Delta_1 + \Delta_2}\, x\, C$$

$$M = 69.5 + \frac{4}{4 + 2}\, x\, 10 = 76.166$$

مثــال (23.3)

احسب المنوال للزمن المستخدم لتجهيز 30 قطعة ميكانيكية في ورشة خراطة هندسية.

عدد الدقائق	عدد القطع
2 – 0	9
5 – 3	13
8 – 6	6
11 – 9	2
المجموع	**30**

الحــل

الفئة المنوالية هي الفئة الثانية والتي يقابلها أكبر تكرار "13".

هنا البيانات تمثل قيم منفصلة.

$$L = 2.5 \quad , \quad C = 3 \ , \ \Delta_1 = 13 - 9 = 4 \quad , \quad \Delta_2 = 13 - 6 = 7$$

$$M = L + \frac{\Delta_1}{\Delta_1 + \Delta_2} \, x \, C$$

$$M = 2.5 + \frac{4}{4 + 7} \, x \, 3 = 3.59$$

حساب المنوال بيانياً:

يمكن تقدير قيمة المنوال بيانياً من خلال المنحنى التكراري أو المضـلع التكراري، حيث تمثل قمة المنحنى النقطة التي يقابلها المنوال، لأن هذه القمة تعبّر عن أعلى تكرار في البيانات، ولأن رسم المنحنى عادة يتم يدوياً وبطريقة تقديرية، فإن القيمة المســتنتجة للمنوال تكون قريبة من الحقيقية ولكن غير دقيقة تماماً.

كما يمكن أيضًا تحديد المنوال باستخدام المدرج التكراري، إذ تقع قيمة المنوال عند منتصف الفئة ذات العمود الأعلى (أي الفئة المنوالية).

مثــال (24.3)

من المثــال (22.3) والخاص بسجل درجات عدد من الطلبة الذين حصلوا على درجات مختلفة في اختبار الرسم الهندسي كما يلي:

الدرجة	عدد الطلاب
59–50	5
69–60	8
79–70	12
89–80	10
99–90	5

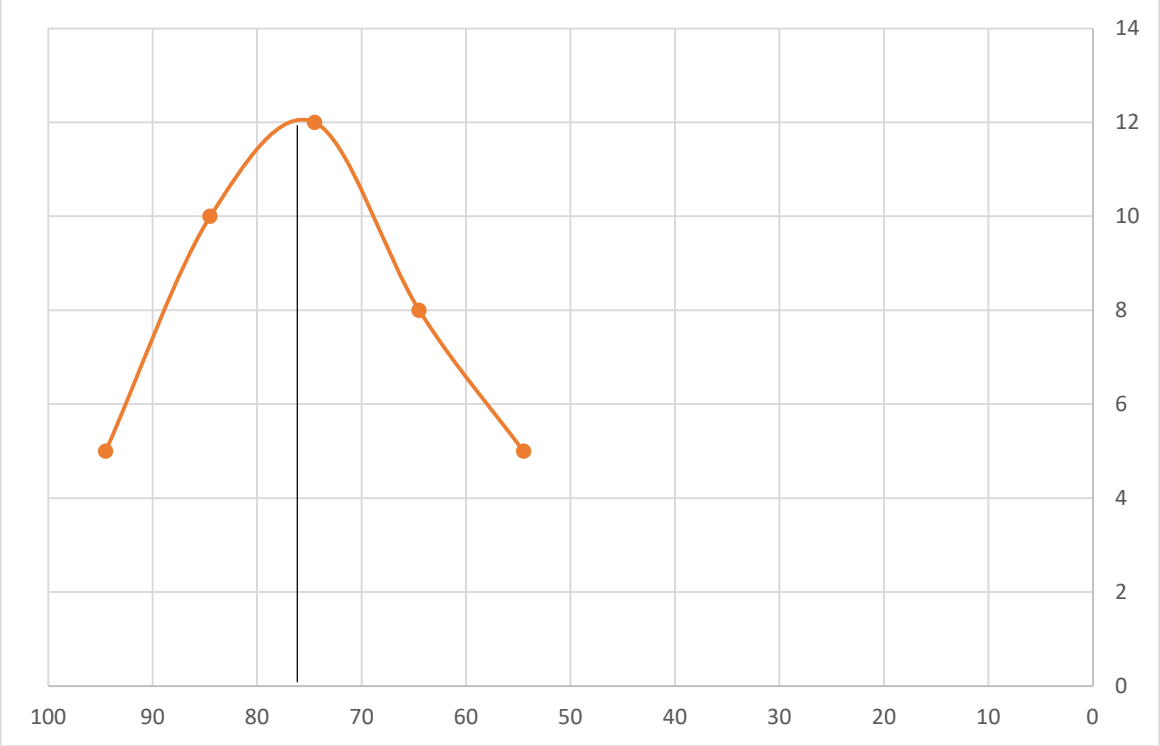
احسب قيمة المنوال بيانياً مستخدماً المنحنى التكراري.

الحـــل

يتم رســـم المنحنى التكراري وذلك باســـتخدام مراكز الفئات، حيث أنها هي النقاط التي تمثل هذه الفئات على الرسم البياني، والجدول (15.3) يوضح هذه الخطوة.

الدرجة	عدد الطلاب	مركز الفئة	النقاط (y,x)
59–50	5	54.5	5 ، 54.5
69–60	8	64.5	8 ، 64.5
79–70	12	74.5	12 ، 74.5
89–80	10	84.5	10 ، 84.5
99–90	5	94.5	5 ، 94.5

من ثم نقوم برسـم المنحنى التكراري والموضـح في الشـكل (4.3) وتحدد منه قيمة المنوال والتي تساوي 76.16 كما تم حلها رياضياً.



الشكل (4.3): يوضح المنحنى التكراري وإيجاد المنوال بيانياً.

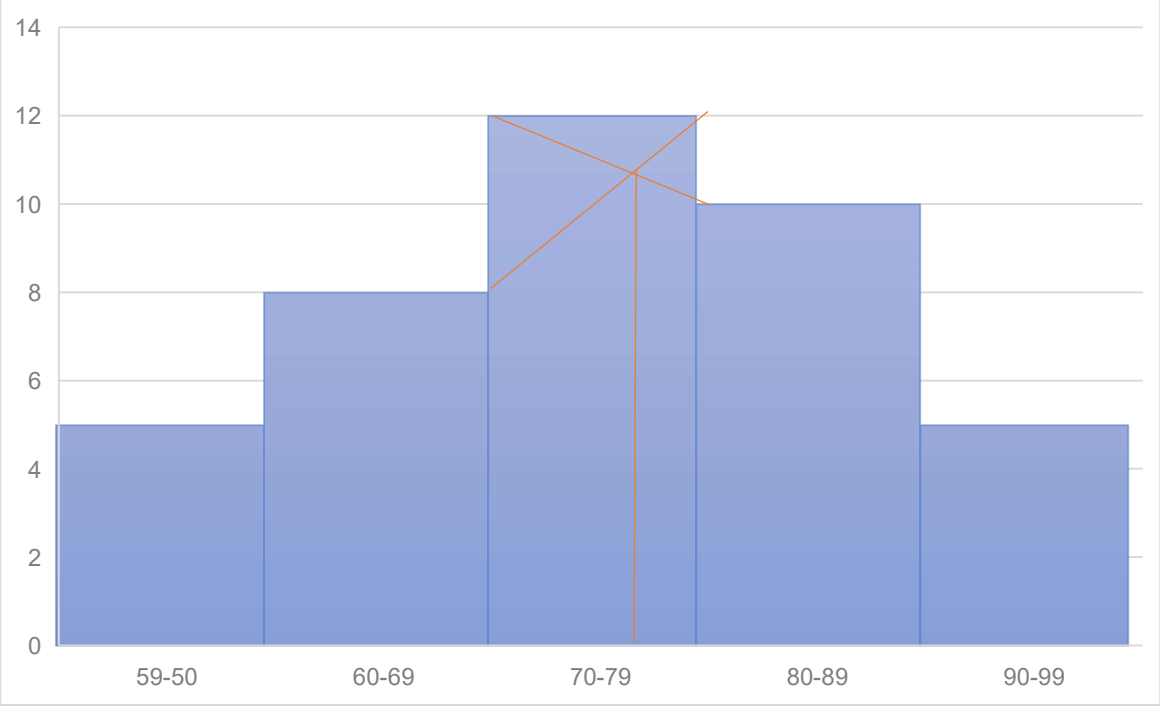
مثـال (25.3)

لنفس المثال السـابق والخاص بسـجل درجات عدد من الطلبة الذين حصـلوا على درجات مختلفة في اختبار الرسم الهندسي، احسب المنوال بيانياً باستخدام المدرج التكراري.

الحـل

نرســـم المدرج التكراري أولاً، ثم نحدد الفئة المنوالية (أي الفئة ذات العمود الأعلى)، والفئتين المجاورتين لها من الجهتين.

بعد ذلك نصـــل الحد الأعلى لعمود الفئة الســـابقة مع الحد الأعلى لعمود الفئة المنوالية بخط مستقيم، ونفعل الأمر نفسه بين الحد الأدنى لعمود الفئة المنوالية والحد الأدنى لعمود الفئة التالية، ثم من نقطة تقاطع الخطين المائلين نسـقط عموداً على المحور الأفقي، فتكون النقطة التي يلتقي عندها العمود بالمحور الأفقي هي قيمة المنوال تقريباً، كما هو موضح في الشكل (5.3).



الشكل (5.3): يوضح المدرج التكراري وإيجاد المنوال بيانياً.

وهي تقريبا من خلال الرسم = 76 .

1.5.3 خـواص المـنوال:

1. يُعدّ أسهل مقاييس النزعة المركزية حسابًا، ولا يحتاج إلى عمليات رياضية معقدة.
2. لا يتأثر بالقيم المتطرفة، لأن حسابه يعتمد فقط على أكثر القيم تكرارًا دون النظر لبقية القيم.
3. يمكن حســـابه في حالة الجداول التكرارية المفتوحة بشـــرط ألا تكون الفئة المنوالية نفســـها مفتوحة.
4. يمكن استخدامه مع البيانات النوعية (الوصفية)، كما هو موضح في المثـال (20.3).
5. يفقد معناه في حالة العينات الصغيرة، إذ قد لا يوجد منوال أصلًا، أما في البيانات الكبيرة فله أهمية عملية كبيرة، خاصة في مجالات التسويق.

فمثلًا: شــركات بيع الأحذية لا تهتم بالوســط أو الوســيط، بل تهتم بالقياس الأكثر طلبًا أي المنوال.

6. يمكن تحديد المنوال بيانيًا باستخدام المدرج التكراري أو المنحنى التكراري.
7. قد يكون للبيانـات أكثر من منوال واحد (ثنـائي أو متعدد المنوال)، وقد لا يكون لهـا منوال إطلاقًا.
8. يتأثر المنوال بطريقة تقســـيم الفئات في الجدول التكراري، فاختلاف حدود الفئات يؤدي إلى تغير التكرارات، وبالتالي قد تتغير الفئة المنوالية، مما ينتج قيمًا مختلفة للمنوال.

تمـــارين

أسئلة اختيار من متعدد:

1. أيّ من مقاييس النزعة المركزية يتأثر أكثر بالقيم الشاذة (المتطرفة)؟

أ. الوسيط ب. المتوسط الحسابي
ج. المنوال د. لا شيء منها

2. الوسيط هو القيمة التي:

أ. تتكرر أكثر من غيرها ب. تساوي مجموع القيم مقسومًا على عددها
ج. تقع في منتصف البيانات بعد ترتيبها د. تمثل أعلى قيمة في البيانات

3. إذا كانت قيم البيانات مرتبة كالتالي: 5، 7، 9، 11، 13، فإن الوسيط يساوي:

أ. 7 ب. 11
ج. 9 د. 5

4. إذا كانت البيانات 2، 3، 3، 5، 5، 5، 8، 9، فإن المنوال هو:

أ. 3 ب. لا يوجد
ج. 5 د. 8

5. إذا كان مجموع القيم = 400 وعددها = 8، فإن المتوسط الحسابي =

أ. 40 ب. 80
ج. 50 د. 60

6. في البيانات النوعية، يمكن حساب:

أ. المنوال فقط ب. الوسيط فقط
ج. المتوسط فقط د. الوسيط والمنوال ولهما شروط.

7. في الجدول التكراري، الفئة الوسيطية هي الفئة التي:

أ. لها أكبر تكرار ب. تسبق الفئة المنوالية

ج. تحتوي على قيمة المنوال د. يحتويها ترتيب الوسيط

أسئلة صح أو خطأ (مع التعليل):

1. المتوسط الحسابي يمكن حسابه لجميع أنواع البيانات. (.....)
2. المنوال قد يكون أكثر من قيمة واحدة. (.....)
3. الوسيط يتأثر بالقيم المتطرفة مثل المتوسط. (.....)
4. المنوال يصلح لقياس بيانات الأسماء مثل "النوع – التخصص". (.....)
5. يمكن أن تكون قيمة الوسيط مساوية لأحد قيم البيانات. (.....)
6. في البيانات المبوبة، الوسيط لا يعتمد على التكرار المتجمع. (.....)
7. المنوال هو القيمة التي لا تتكرر أبدًا. (.....)
8. إذا كان لدينا توزيع ثنائي المنوال، فإن القيم تتركز حول قيمتين. (.....)
9. المتوسط الحسابي لا يُستخدم مع القيم الفئوية أو الرمزية. (.....)

أسئلة تطبيقية:

1. احسب المتوسط الحسابي للقيم الآتية:

15، 18، 22، 25، 30، 40

2. احسب الوسيط والمنوال للقيم التالية:

8، 10، 12، 12، 15، 18، 20

3. الجدول التالي يوضح أوزان قطع معدنية في مصنع هندسي:

الفئة (جرام)	أقل من 20	أقل من 30	أقل من 40	أقل من 50	أقل من 60	أقل من 70
التكرار	8	12	20	35	15	10

احسب الوسيط مبينًا خطوات الحل.

4. الجدول الآتي يمثل عدد الأخطاء في خطوط إنتاج مختلفة:

عدد الأخطاء	0	1	2	3	4	5
عدد الخطوط	2	5	10	8	3	2

احسب المنوال وعدد الأخطاء الأكثر شيوعًا.

5. في أحد المعامل، تم تسجيل إنتاج خمس آلات خلال يوم واحد كما يلي:

45، 50، 55، 60، 90

احسب المتوسط، ثم ناقش هل تعتبر القيمة الأخيرة (90) مؤثرة على المقياس؟ وأي مقياس أنسب لتمثيل البيانات؟

6. في توزيع تكراري لدرجات اختبار طلاب، كانت النتائج كالتالي:

الفئة	49–40	59–50	69–60	79–70	89–80
التكرار	3	5	9	7	6

احسب المتوسط الحسابي التقريبي باستخدام الفئات.

أسئلة تأملية وتحليلية:

1. إذا كان متوسط دخل مجموعة من العاملين هو 2000 دينار، بينما أغلبهم يتقاضى بين 1000 و1500 دينار، فما دلالة ذلك.
2. في تحليل جودة الإنتاج، وُجد أن منوال سماكة الصفائح المعدنية = 5 مم، بينما الوسيط = 5.5 مم، والمتوسط = 6 مم. ماذا يمكن أن يعني هذا بالنسبة للجودة؟
3. إذا زادت جميع القيم في مجموعة بيانات بمقدار ثابت، فكيف يتأثر كل من المتوسط، الوسيط، والمنوال؟ وضّح بالأمثلة.

الفصل الرابع

مقاييس التشـــتت

1.4 مقدمــة:

لقد تعرفنا في الباب الســابق على مقاييس النزعة المركزية (المتوســطات)، وهي التي تُمكننا من تحديد القيمة التي تميل إليها البيانات وتتجمع حولها، ولكن هذه المقاييس – رغم أهميتها – لا تكشـف لنا عن مدى انتشـار القيم أو تباعدها عن بعضـها البعض أو عن مركزها، أي لا تُظهر درجة تفاوت البيانات داخل المجموعة، فقد نجد مجموعتين من البيانات لهما نفس المتوســط الحسابي، ومع ذلك تكون القيم في إحداهما متقاربة جدًا، بينما في الأخرى متباعدة بدرجة كبيرة.

ومن هنا تظهر أهمية مقاييس التشـتت، فهي التي تعبر عن مدى تنوع القيم وتفرقها حول مقياس النزعة المركزية، وتُعطينا فكرة أوضح وأعمق عن طبيعة التوزيع الإحصائي للظاهرة المدروسة.

بمعنى آخر، مقياس النزعة المركزية يخبرنا "أين يتمركز التوزيع"، وأما مقياس التشــتت فيخبرنا "كيف تتوزع القيم حول هذا المركز".

مثــال (1.4)

لدينا المجموعات التالية من البيانات:

المجموعة الأولى: 10، 10، 10، 10، 10

المجموعة الثانية: 12، 11، 10، 9، 8

المجموعة الثالثة: 0، 20، 5، 15، 10

نلاحظ أن جميع المجموعات لها نفس الوسط الحسابي والوسيط، حيث كلاهما يساوي 10، رغم وجود اختلاف واضح بين المجموعات في انتشار القيم وتباعدها.

في المجموعة الأولى، لا يوجد اختلاف أو تشــتت بين القيم، فجميعها متســاوية وتســاوي قيمة الوسط الحسابي 10.

في المجموعة الثانية، القيم تختلف قليلاً عن بعضـها وعن الوسـط الحسـابي، أي أن تشـتت القيم صغير نسبيًا.

في المجموعة الثالثة، القيم تختلف بشكل كبير عن بعضـها وعن الوسط الحسابي، أي أن تشتت القيم كبير، مما يوضح أن الوسط الحسابي لا يعكس التباعد بين القيم.

عدم تســـاوي توزيع القيم بين المجموعات لا يعني أن البيانات متكافئة أو متشـــابهة فيما بينها. لذلك، للوصف الدقيق للبيانات محل الدراسة، لا يكفي معرفة القيمة التي تتجمع حولها القيم (مثل الوسط الحسابي أو الوسيط)، بل يجب أيضًا دراسة كيفية انتشار هذه القيم وتباعدها عن بعضها البعض وعن المركز.

لقياس هذا الترابط أو التشتت، نستخدم مقاييس إحصائية خاصة بالانتشار، والتي تساعد في فهم طبيعة البيانات أكثر، وأهم هذه المقاييس:

- **المدى**.
- **التباين**.
- **الانحراف المعياري**.
- **معامل الاختلاف** للمقارنة بين مجموعتين أو أكثر من البيانات من حيث تشتتها.

كلما زاد تشتت القيم داخل مجموعة البيانات، زادت قيمة مقياس التشتت الخاص بها، ومن المهم ملاحظة أن مقاييس التشـتت تعتمد على القيم العددية للبيانات، وبالتالي فهي لا تصـلح للبيانات النوعية أو الوصفية.

2.4 المــدى:

المدى هو الفرق بين أكبر قيمة وأصــغر قيمة في مجموعة البيانات، ويعبّر عن الفترة التي تتغير فيها القيم للمتغير محل الدراسة. يُرمز له بالرمز R. Range.

في حـالة البيانات الغير مبوبة يحسب كما يلي:

المدى = أكبر قيمة – أصغر قيمة

أما في حالة البيانات المبوبة، يُحسب المدى كذلك باعتباره الفرق بين أكبر قيمة وأصغر قيمة في البيانات، مع الأخذ بعين الاعتبار أن الجدول التكراري مرتب ترتيبًا تصاعديًا.

ففي هذه الحالة، نعتبر الحد الأعلى للفئة الأخيرة هو أكبر قيمة، والحد الأدنى للفئة الأولى هو أصغر قيمة. وعليه، يمكن حساب المدى بالطرح بين هذين الحدين.

المدى = الحد الأعلى للفئة الأخيرة – الحد الأدنى للفئة الأولى

يشـــير المدى إلى مدى انتشـــار القيم داخل البيانات؛ فإذا كان صـــغيرًا، دل ذلك على أن القيم متقاربة من بعضـها وتشتتها قليل، أما إذا كان كبيرًا، فهذا يعكس أن القيم موزعة على فترة واسعة وأن التشتت بينها كبير.

مثـــال (2.4)

قارن بين تشتت المجموعات الثلاث المذكورة في المثال 1.4

مدى المجموعة الأولى: 10 – 10 = 0

مدى المجموعة الثانية: 12 – 8 = 4

مدى المجموعة الثالثة: 20 – 0 = 20

بما أن مدى المجموعة الأولى يساوي صفرًا، فهذا يدل على غياب أي اختلاف أو تشتت بين قيم هذه المجموعة، أي أن جميع القيم متســـاوية تمامًا، أما بالنســـبة للمجموعتين الثانية والثالثة، فنلاحظ أن مدى المجموعة الثانية أصـــغر من مدى المجموعة الثالثة، مما يعني أن القيم في المجموعة الثانية أكثر تقاربًا من بعضـها، وبالتالي يكون تشـتت القيم فيها أقل مقارنة بالمجموعة الثالثة التي تظهر قيمها موزعة على نطاق أوسع وتشتت أكبر.

مثـــال (3.4)

افترض أن لدينا درجتين لمجموعتين من الطلاب في مادة "الإحصاء".

المجموعة الأولى: 15، 15، 15، 15، 15

المجموعة الثانية: 12، 14، 16، 18، 20

قارن بين تشتت درجات المجموعتين.

الحـــل

المجموعة الأولى:

المدى = 15 – 15 = 0

المجموعة الثانية:

$$\text{المدى} = 20 - 12 = 8$$

مدى المجموعة الأولى = 0 يدل على أن جميع درجات الطلاب متساوية ولا يوجد تشتت بين القيم.

مدى المجموعة الثانية = 8 يدل على أن الدرجات متباعدة أكثر، أي أن التشتت في هذه المجموعة أكبر مقارنة بالمجموعة الأولى.

مثـــال (4.4)

احسب المدى للبيانات التالية التي تمثل درجات اختبار لمجموعة من الطلاب مقسمة في جدول تكراري:

الفئة (الدرجات)	التكرار
14 – 10	3
19 – 15	5
24 – 20	4
29 – 25	2

الحـــل

أكبر قيمة = الحد الأعلى للفئة الأخيرة = 29 درجة.

أصغر قيمة = الحد الأعلى للفئة الأولى = 10 درجات.

المدى هو = 29 – 10 = 19 درجة R

1.2.4 خواص المدى:

1. المدى مقياس بسيط وسهل الحساب، ومفهومه ودلالته واضحة.
2. يعتمد المدى على قيمتين فقط من البيانات: أكبر قيمة وأصـــغر قيمة، ويتجاهل بقية القيم.

3. قد يكون المدى مضـللاً إذا كانت إحدى القيمتين أو كلاهما شـاذة أو متطرفة، إذ يعطي انطباعاً بتشتت أكبر من الواقع.

مثال: البيانات 6، 8، 9، 10، 11، 12، 14، 55

المدى = $55 - 6 = 49$ ← يبدو كبيراً رغم أن معظم القيم متقاربة

بعد حذف القيمة المتطرفة 55: المدى = $14 - 6 = 8$ ← أكثر واقعية

4. لا يمكن حســـاب المدى في حالة التوزيعات التكرارية المفتوحة، لأنه يكون الحد الأدنى للفئة الأولى أو الحد الأعلى للفئة الأخيرة أو كلاهما مجهولاً.

3.4 التبـــاين:

هو مقياس لتشتت البيانات حول وسطها الحسابي، ويعرف كذلك بأنه الوسط الحسابي لمربعات انحرافات القيم عن وسطها الحسابي ويرمز له بالرمز S^2 .

أ- التباين في حالة البيانات غير المبوبة:

لحساب التباين للبيانات غير المبوبة، نتبع الخطوات التالية:

- نحسب أولاً الوسط الحسابي للبيانات.
- نحسب انحراف كل قيمة عن الوسط الحسابي، أي:

انحراف القيمة عن الوسط الحسابي = القيمة − الوسط الحسابي

- نربع كل انحراف للحصول على مربع الانحراف لكل قيمة.
- نجمع جميع مربعات الانحرافات للحصول على مجموع مربعات الانحرافات.
- نحسب التباين بقسمة مجموع مربعات الانحرافات على عدد القيم − 1، أي باستخدام الصيغة:

للمجتمع:

$$\sigma^2 = \frac{\sum_{i=1}^{n}(X_i - \mu)^2}{n}$$

للعينة:

$$S^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}$$

ملاحظة: عند حســـاب التباين، يجب الانتباه إلى ما إذا كانت البيانات تمثل كامل المجتمع الإحصائي أو مجرد عينة منه:

- إذا كانت البيانات تمثل المجتمع الكامل، نقســـم مجموع مربعات الانحرافات على عدد القيم. n (σ^2)
- أما إذا كانت البيانات تمثل عينة فقط من المجتمع، فنقســـم على $n-1$ n لتصـــحيح الانحياز، ويعرف هذا بـ تباين العينة. (S^2)

مثــال (5.4)

احســـب التباين في أوزان 8 قطع إنتاجية بالجرامات داخل مصـــنع ميكانيكي، وكانت القيم كما يلي:

20,22,19,25,21,23,20,24

الحــل

المتوسط الحسابي = $\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n} = \frac{174}{8} = 21.75$

والجدول (1.4) يوضح العمليات الحسابية.

جدول (1.4): يوضح العمليات الحسابية للحصول على قيمة التباين.

القيمة X_i	$X_i - \bar{X}$	$(X_i - \bar{X})^2$
20	−1.75	3.0625
22	0.25	0.0625
19	−2.75	7.5625
25	3.25	10.5625
21	−0.75	0.5625
23	1.25	1.5625
20	−1.75	3.0625

5.0625	2.25	24
31.5	$\sum_{i=1}^{n}(X_i - \bar{X})^2$	

$$S^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n} = \frac{31.5}{8} = 3.9375$$

مثـــال (6.4)

احسب التباين لأوزان خمس عبوات من منتج غذائي "بالكيلوغرامات" والأوزان كالتـالي:

10.2 ،10.3 ،10.0 ،10.4 ،9.7

الحـــل

المتوسط الحسابي = $\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n} = \frac{50.6}{5} = 10.12$

والجدول (2.4) يوضح العمليات الحسابية.

جدول (2.4): يوضح العمليات الحسابية للحصول على قيمة التباين.

القيمة X_i	$X_i - \bar{X}$	$(X_i - \bar{X})^2$
9.7	−0.42	0.1764
10.4	0.28	0.0784
10.0	−0.12	0.0144
10.3	0.18	0.0324
10.2	0.08	0.0064
$\sum_{i=1}^{n}(X_i - \bar{X})^2$		0.3062

$$S^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n} = \frac{0.3062}{5} = 0.06124$$

ب-التباين في حالة البيانات المبوبـة:

عند التعامل مع بيانات مبوبة في جداول تكرارية، نحسب التباين باتباع الخطوات الآتية:

- نحسب مراكز الفئات لكل فئة من فئات الجدول.

- نحسب الوسط الحسابي للتوزيع التكراري.
- نحسب انحراف كل مركز فئة عن الوسط الحسابي.
- نحسب مربع انحراف كل مركز فئة عن الوسط الحسابي.
- نضرب مربع الانحراف في تكرار الفئة المقابلة.
- نجمع نواتج الضـــرب لجميع الفئات للحصـــول على مجموع مربعات الانحرافات الموزونة.
- نحسب قيمة التباين باستخدام الصيغة التالية:

$$S^2 = \frac{\sum_{i=1}^{k} f_i (X_i - \bar{X})^2}{\sum_{i=1}^{k} f_i - 1}$$

وهنا تمثل X_i مركز الفئة، f_i تكرار الفئة، $\bar{X}$ الوسط الحسـابي.

مثـــال (7.4)

لدى مصنع قياسات زمن تشغيل آلة (بالدقائق) مُجمَّعة في جدول تكراري كما يلي:

الفئة (دقيقة)	التكرار (f)
14 – 10	3
19 – 15	7
24 – 20	10
29 – 25	6
34 – 30	4
المجموع	**30**

احسب التباين لهذه البيانات.

الحـــل

أولا نقوم بحساب المتوسط الحسابي للبيانات

$$\bar{X} = \frac{\sum_{i=1}^{k} X_i f_i}{\sum_{i=1}^{k} f_i} = \frac{665}{30} = 22.16$$

الجدول التالي يوضح العمليات الحسابية اللازمة.

جدول (3.4): يوضح العمليات الحسابية للحصول على قيمة التباين.

الفئة	التكرار	مركز الفئة	$X_i f_i$	$X_i - \bar{X}$	$(X_i - \bar{X})^2$	$f_i(X_i - \bar{X})^2$
14 – 10	3	12	36	−10.16	103.2256	309.6768
19 – 15	7	17	119	−5.16	26.6256	186.3792
24 – 20	10	22	220	−0.16	0.0256	0.256
29 – 25	6	27	162	4.84	23.4256	140.5536
34 – 30	4	32	128	9.84	96.8256	387.3024
المجمـوع	30		665			1024.168

$$S^2 = \frac{\sum_{i=1}^{k} f_i(X_i - \bar{X})^2}{\sum_{i=1}^{k} f_i} = \frac{1024.168}{30} = 34.1388 \text{ (دقيقة)}^2$$

مثـال (8.4)

يبين الجدول التالي عدد السـاعات التي يقضـيها مجموعة من الطلبة أسـبوعيًا في اسـتخدام الإنترنت:

فئة الساعات	التكرار (f)
4 − 0	2
9 − 5	6
14 − 10	8
19 − 15	3
24 − 20	1

احسب التباين لهذه البيانات.

الحـــل

أولا نقوم بحساب المتوسط الحسابي للبيانات

$$\bar{X} = \frac{\sum_{i=1}^{k} X_i f_i}{\sum_{i=1}^{k} f_i} = \frac{215}{20} = 10.75$$

الجدول التالي يوضح العمليات الحسابية اللازمة.

جدول (3.4): يوضح العمليات الحسابية للحصول على قيمة التباين.

الفئة	التكرار	مركز الفئة	$X_i f_i$	$X_i - \bar{X}$	$(X_i - \bar{X})^2$	$f_i(X_i - \bar{X})^2$
4 – 0	2	2	4	−8.75	76.56	153.12
9 – 5	6	7	42	−3.75	14.06	84.36
14 – 10	8	12	96	1.25	1.56	12.48
19 – 15	3	17	51	6.25	39.06	117.18
24 – 20	1	22	22	11.25	126.56	126.56
المجمـوع	20		215			493.7

$$S^2 = \frac{\sum_{i=1}^{k} f_i (X_i - \bar{X})^2}{\sum_{i=1}^{k} f_i - 1} = \frac{493.7}{20} = 24.685 \text{ (ساعة)}^2$$

ملاحظة:

لاحظ أن وحدة قياس التباين تكون مربّع وحدة القياس الأصـــلية، وهذا يجعلها في الغالب غير ذات معنى مباشـر، فمثلاً في المثال السـابق كانت قيمة التباين = 25.98 سـاعة²، وواضـح أن عبارة "ساعة تربيع" لا تُعبّر عن مفهوم واقعي يمكن تفسيره عملياً.

ولهذا الســـبب يُعتبر هذا أحد عيوب التباين، وللتغلب عليه نقوم بأخذ الجذر التربيعي للتباين لإرجاع الوحدات إلى أصلها، ويسمّى الناتج الجديد الانحراف المعياري.

4.4 الانحـراف المعياري:

الانحراف المعياري هو الجذر التربيعي الموجب لمتوســـط مربعات انحرافات القيم عن وســـطها الحســـابي، أي أنه الجذر التربيعي الموجب للتباين، ويرمز له بالرمز (S)، ويُحســـب باســـتخدام العلاقة التالية:

أ- في حالة البيانات الغير مبوبة:

$$S = \sqrt{S^2} = \sqrt{\frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}}$$

ب- في حالة البيانات المبوبة:

$$S = \sqrt{S^2} = \sqrt{\frac{\sum_{i=1}^{k} f_i(X_i - \bar{X})^2}{\sum_{i=1}^{k} f_i - 1}}$$

ويكون الانحراف المعياري في الأمثلة السابقة كالتالي:

- للمثـال (5.4) $S = \sqrt{S^2} = \sqrt{3.9375} = 1.9843$
- للمثـال (6.4) $S = \sqrt{S^2} = \sqrt{0.06124} = 0.2474$
- للمثـال (7.4) $S = \sqrt{S^2} = \sqrt{34.1388} = 5.8428$
- للمثـال (8.4) $S = \sqrt{S^2} = \sqrt{24.685} = 4.9684$

1.4.4 خــواص الانحــراف المعياري:

1. يُعدّ الانحراف المعياري من أهم مقاييس التشـــتت وأكثرها اســـتخدامًا في التحليل الإحصائي.
2. يتميز بأنه يأخذ في الحسبان جميع القيم في مجموعة البيانات دون إهمال أي منها.
3. يمتاز بقدرته العالية على المعالجة الجبرية وإمكانية استخدامه في مختلف التطبيقات الرياضية والإحصائية.
4. لا يمكن حسابه في حالة التوزيعات التكرارية المفتوحة، نظرًا لعدم معرفة حدود بعض الفئات.

5.4 معامـــل الاختلاف:

جميع مقاييس التشـــتت تعتمد على وحدات القياس المســـتخدمة في البيانات، ولذلك لا يمكن المقارنة بينها عندما تختلف وحدات القياس، كأن نحاول مقارنة تشتت الأطوال مع تشتت الأوزان مثلاً.

ولتفادي هذه المشكلة، نلجأ إلى استخدام مقياس نسبي لا يعتمد على وحدات القياس، ويُعرف هذا المقياس باسم معامل الاختلاف، ويتم حسابه بالصيغة التالية:

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

وبالطبع كلما زاد التشتت أخذ معامل الاختلاف نسبة أكبر.

مثـــال (9.4)

لدينا مجموعتان من البيانات الأولى لأطوال أشـــجار بالســـنتيمتر، والثانية لأوزان فاكهة التفاح بالكيلو جرام، قارن بين تشتت المجموعتين باستخدام معامل الاختلاف.

فئة الأطوال	التكرار
119 – 100	3
139 – 120	4
159 – 140	2
179 – 160	1

فئة الأوزان	التكرار
0.9 – 0.5	2
1.4 – 1.0	3
1.9 – 1.5	4
2.4 – 2.0	1

الحـــل

أولا – لأطوال الشجر:

نقوم بحساب المتوسط الحسابي للبيانات

$$\bar{X} = \frac{\sum_{i=1}^{k} X_i f_i}{\sum_{i=1}^{k} f_i} = \frac{1315}{10} = 131.5$$

الجدول التالي يوضح العمليات الحسابية اللازمة.

جدول (4.4): يوضح العمليات الحسابية للحصول على قيمة التباين.

الفئة	التكرار	مركز الفئة	$X_i f_i$	$X_i - \bar{X}$	$(X_i - \bar{X})^2$	$f_i(X_i - \bar{X})^2$
119 – 100	3	109.5	328.5	−22	484	1452

16	4	−2	518	129.5	4	139 – 120
648	324	18	299	149.5	2	159 – 140
1444	1444	38	169.5	169.5	1	179 – 160
3560			1315		10	المجمـوع

$$S^2 = \frac{\sum_{i=1}^{k} f_i(X_i - \bar{X})^2}{\sum_{i=1}^{k} f_i - 1} = \frac{3560}{9} = 395.5 \ \left(\text{سم}\right)^2$$

الانحراف المعياري لمجموعة أطوال الأشجار هو

$$S = \sqrt{S^2} = \sqrt{395.5} = 19.887$$

معـامل الاختلاف هو

$$\frac{19.887}{131.5} x100 = 15.12\%$$

ثـانياً – لأوزان التفـاح:

نقوم بحساب المتوسط الحسابي للبيانات

$$\bar{X} = \frac{\sum_{i=1}^{k} X_i f_i}{\sum_{i=1}^{k} f_i} = \frac{13}{10} = 1.3$$

الجدول التالي يوضح العمليات الحسابية اللازمة.

جدول (5.4): يوضح العمليات الحسابية للحصول على قيمة التباين.

$f_i(X_i - \bar{X})^2$	$(X_i - \bar{X})^2$	$X_i - \bar{X}$	$X_i f_i$	مركز الفئة	التكرار	الفئة
1.6562	0.8281	−0.91	1.4	0.7	2	0.9 – 0.5
0.03	0.01	−0.1	2.6	1.2	3	1.4 – 1.0
0.64	0.16	0.4	6.8	1.7	4	1.9 – 1.5
0.81	0.81	0.9	2.2	2.2	1	2.4 – 2.0
3.1362			13			المجمـوع

$$S^2 = \frac{\sum_{i=1}^{k} f_i(X_i - \bar{X})^2}{\sum_{i=1}^{k} f_i - 1} = \frac{3.1362}{9} = 0.348 \ \left(\text{كيلوجرام}\right)^2$$

الانحراف المعياري لمجموعة أطوال الأشجار هو

$$S = \sqrt{S^2} = \sqrt{0.348} = 0.589$$

معامل الاختلاف هو

$$\frac{0.589}{1.3} x100 = 45.30\%$$

معامل اختلاف أوزان التفاح أكبر من أطوال الشجر، إذن تشتت الأوزان أكبر من تشتت الأطوال.

تمـــارين

1. ما المقصود بمقاييس التشتت؟ ولماذا نحتاج إليها رغم وجود مقاييس النزعة المركزية؟
2. وضّح الفرق بين مقاييس النزعة المركزية ومقاييس التشتت من حيث المعنى والاستخدام.
3. ما المقصود بمدى القيم في مجموعة البيانات؟ وما دلالته الإحصائية؟
4. متى يكون التشتت معدومًا؟ ومتى يكون في أعلى حالاته؟
5. لماذا يُعد المدى مقياسًا غير دقيق في تمثيل التشتت أحيانًا؟
6. وضّح المقصود بكل من:
   - التباين (Variance)
   - الانحراف المعياري (Standard Deviation)
   - معامل الاختلاف (Coefficient of Variation)
7. ما العلاقة بين الانحراف المعياري والتباين؟
8. كيف يُســـتخدم معامل الاختلاف في المقارنة بين تشـــتت مجموعتين تختلفان في وحدات القياس؟
9. هل يمكن أن يكون الانحراف المعياري صفرًا؟ وماذا يعني ذلك؟
10. البيانات التالية تمثل عدد ساعات تشغيل محرك كهربائي خلال 6 أيام:

    7، 9، 8، 6، 10، 7

    احسب المدى والانحراف المعياري.

11. الجدول التالي يوضح عدد مرات توقف آلة في مصنع خلال أسابيع مختلفة:

عدد مرات التوقف	0 – 2	2 – 4	4 – 6	6 – 8	8 – 10

عدد الأسابيع	5	8	12	9	6

احسب التباين والانحراف المعياري لعدد مرات التوقف.

12. في اختبارين هندسيين حصل طالبان على نفس المتوسط الحسابي للدرجات، ولكن أحدهما كان انحرافه المعياري أكبر.

ماذا يعني ذلك من حيث استقرار الأداء؟

13. إذا كان متوسط زمن معالجة الإشارة في دارة إلكترونية هو 40 ميلي ثانية، والانحراف المعياري 2 ميلي ثانية، فما معامل الاختلاف؟ وماذا نستنتج منه؟
14. في مصنع لتصنيع الموصلات المعدنية، تم تسجيل أطوال عشر عينات (بالمليمتر):

48، 50، 52، 50، 49، 51، 50، 52، 53، 49

احسب المدى والانحراف المعياري.

15. مجموعتان من البيانات لهما نفس المتوسط الحسابي، ولكن المدى في الأولى أكبر من الثانية. أي المجموعتين أكثر تشتتًا؟ ولماذا؟
16. ما الموقف الذي تفضل فيه استخدام الانحراف المعياري بدلًا من المدى كمقياس للتشتت؟
17. إذا كان معامل الاختلاف لمجموعة بيانات هو 5% ولأخرى 25%، فماذا تستنتج عن استقرار القيم في كل مجموعة؟
18. وضّح بمثال هندسي كيف يمكن أن يساعد حساب التشتت في تحسين جودة الإنتاج الصناعي.
19. ناقش أهمية معرفة درجة تشتت البيانات في تقييم كفاءة الأنظمة الكهروميكانيكية.