جامعة البصرة

كلية الإدارة والاقتصاد

قسم الإحصاء

محاضرات

المرحلة الأولى – قسم الإحصاء

الفصل الأول – للعام الدراسي 2019-2020

اسم المقرر	**احصاء1**	**رمز المقرر**	STAT101	**عدد وحدات المقرر**	**3**
نوع المقرر	**إجباري كلية**	**المستوى**	**الأول**	**اسم التدريسي**	**م.م. احمد هشام محمد**

مصادر المقرر

1 - مبادئ الإحصاء التطبيقي ....تأليف : أ.د. رعد فاضل حسن التميمي – بغداد 2014

2 - مبادئ الإحصاء ..... د. محمود المشهداني .

3 - الإحصاء التطبيقي ....د. خاشع الراوي .

الأسبوع الأول :

1- نبذة تاريخية عن علم الإحصاء

تتطور علم الإحصاء ضمن مراحل متعددة مع التطور التكنولوجي والرياضي وزاد الاهتمام بهذا النوع من العلوم نظراً لما يوفره من اساليب وطرق علمية ورياضية لمعالجة البيانات وتحليلها، كما يعد علم الإحصاء من العلوم الأساسية التي تدخل في العديد من العلوم التطبيقية والاجتماعية لمعالجة العديد من الظواهر، أن كلمة (الإحصاء) في الماضي كانت تهدف إلى العد والحصر وسمي الإحصاء بعلم العد (The science of counting) كما أن لفظة إحصاء باللغة الانكليزية (Statistics) كانت تستعمل في بلاد أوربا للدلالة على أعمال وحسابات الدولة في شؤون الحرب والضرائب وعدد السكان والمواليد والوفيات والإنتاج الزراعي ...الخ، لذا فان مفهوم علم الإحصاء يختلف عند العامة فمنهم من يصفه بأنه العلم الذي يهتم بعملية جمع البيانات وحفظها، والبعض الأخر اتجه نحو وصفه بأنه العلم الذي يهتم بعملية جمع ووصف البيانات وتنظيمها ومن ثم تحليلها وإعادة تشكيلها مما يسهل قرأتها ومن ثم تهيئتها لغرض اتخاذ القرار المناسب، أما الآن فإن الإحصاء قد تطور كثيراً خاصة في القرن العشرين حيث اكتسب هذا العلم أهميةً كبيرة مع تطور النظرية الاحتمالية بحيث أصبح علماً مستقلاً وانتشر استخدامه من قبل العلماء في مجالات متعددة كـ(الطب، الهندسة، الزراعية، الاجتماعية، الجغرافيا، الفلك... الخ)، لذا أصبح له أهميته كوسيلة وأداة في البحث العلمي لجميع العلوم .

2- الهدف من دراسة علم الإحصاء

أن الهدف أو الأهداف الأساسية من دراسة علم الإحصاء

- التعريف بطبيعة علم الإحصاء وأهميته في مجالات تطبيقه .
- التعريف بأساليب وطرق جمع البيانات وسحب العينة وكيفية عرضها لغرض الوصول الى نتائج دقيقة بأقصر الطرق واقل الكلف .
- تخطيط المشاريع سواء كانت مشاريع فردية او مشاريع حكومية بما يسهم في تطوير العمل لتحقيق أفضل النتائج .

3- تعريف علم الإحصاء، مجالاته وأهميته

اولاً : تعريف علم الإحصاء

عرف علم الإحصاء بتعاريف عديدة اختلفت وتباينت من حيث المضمون والشمول باختلاف مراحل تطوير هذا العلم والفوائد المتوخاة منه، نذكر منها :

- هو العلم الذي يعتمد الطرق العلمية في عملية جمع وتلخيص وتحليل البيانات للوصول إلى نتائج دقيقة ومقبولة تسهم في اتخاذ قرارات سليمة في ضوء هذا التحليل .
- بأنه الطريقة العلمية التي تختص بجمع البيانات والحقائق بشكل يسهل عملية تحليلها وتفسيرها ومن ثم استخلاص النتائج واتخاذ القرار في ضوء التحليل
- بأنه العلم الذي يعمل على استخدام الأسلوب العلمي في طرق جمع البيانات وتبويبها وتلخيصها وعرضها وتحليلها بهدف الوصول إلى استنتاجات وقرارات مناسبة.
- هو احد فروع علم الرياضيات الذي يشمل النظريات والطرق الموجهة نحو جمع البيانات ووصفها، وتفسيرها ومن ثم صنع القرارات .

ثانياً: مجالات وأهمية علم الإحصاء وعلاقته بالعلوم الأخرى:

يعتبر علم الإحصاء من العلوم التطبيقية المهمة ووسيلة بحث علمي تسهم في دعم الدراسات والبحوث العلمية في جميع المجالات من خلال استعمال قوانينه وقواعده وأساليب جمع البيانات والمعلومات ومن ثم تحليلها بهدف الوصول إلى النتائج وأخيراً استخدام تلك النتائج في صنع القرار، ان من ابرز المجالات التي يدعمها علم الإحصاء هي :

1. التخطيط : أي وضع الخطط المستقبلية عن طريق التنبؤ بالنتائج، من هذا التعريف يتضح أن التخطيط يقوم على عملية التفكير و التقدير للمستقبل و النظر إلى البعد الزمني و التنبؤ بالمتغيرات و وضع الخطط لما يخفيه المستقبل و التأقلم مع الظروف المتغيرة.

2. الاقتصاد والاجتماع :يبرز دور الإحصاء في بحوث السكان متمثلاً بالتعداد السكاني اذ ان التخطيط السليم للتنمية الاقتصادية والاجتماعية لا ينفصلان ولا يمكن أن يتم إلا من خلال دراسات إحصائية للسكان .
3. الزراعية : يبرز دور الإحصاء بشكل مهم في المجالات الزراعية والتي تبدأ بملاحظة وجمع بيانات عن الطبيعة في الحقل أو المزرعة ثم يلي ذلك الدراسات العملية للتجارب الحقلية بعد ترتيب الملاحظات (المشاهدات) ومن تم تحليلها .
4. الصناعية
5. الطب والصحة العامة .
6. علم النفس .

4- الطريقة الإحصائية بالبحث العلمي :

استخدام الأسلوب الإحصائي في البحث العلمي يعني توفير البيانات والمعلومات عن الظاهرة المطلوب دراستها في ذلك البحث و هذا يعني ان أمكانية تطبيق الطريقة الإحصائية مرهوناً بإمكانية التعبير عن هذه الظاهرة أو تلك تعبيرا كمياً (رقمياً)، وتمتاز الطريقة الإحصائية بكونها تهيئ أسلوبا موضوعيا محايدا للبحث له قواعده وأصول التي يجب أن يلتزم بها الباحث حتى يتجنب التحيز الشخصي والوقوع في بعض الأخطاء . كما يساعد استخدام الطريقة الإحصائية إلى وصول الباحث إلى النتائج الدقيقة بأقصر طريق وأقل كلفة ، وعليه يمكن ان نحدد المراحل الرئيسية للطريقة الإحصائية :

- ✓ **تحديد المشكلة**
- ✓ **تحديد مجتمع الدراسة**
- ✓ **جمع البيانات والمعلومات**
- ✓ **تصنيف البيانات وتبويبها**
- ✓ **عرض البيانات**
- ✓ **حساب المؤشرات وتقدير المعالم**
- ✓ **التفسير والتنبؤ**

5- ينقسم علم الإحصاء الى نوعين هما

أولاً : الإحصاء الوصفي :

ويتضمن الطرق الإحصائية المستخدمة في جمع البيانات والمعلومات عن ظاهرة معينة أو مجموع من الظواهر وكيفية تنظيم وتصنيف وتبويب البيانات مع أمكانية عرضها في جداول ورسوم بيانية وحساب بعض المؤشرات الإحصائية .

ثانياً الإحصاء الاستدلالي (ألاستنتاجي) :

يستند الإحصاء الاستدلالي على مجموعة من النظريات الإحصائية أهمها نظرية الاحتمالات ونظرية العينات التي تمثلان حلقة الوصل بين الإحصاء الوصفي والاستنتاجي، ونسعى من خلال هذا النوع من الأساليب الإحصائية إلى تقدير معالم وخصائص المجتمعات من خلال ماهو متوفر من معلومات وعينات مختارة فضلا عن اختبار الفرضيات الإحصائية للمجتمع على أساس البيانات المتاحة .

6- طبيعة البيانات الإحصائية

ان جمع بيانات حول ظاهرة معينة نرمز لتلك الظاهرة بالرمز y او x أما مفردات (المشاهدة) لتلك الظاهرة فنرمز لها بالرمز y_i او x_i، فعند دراسة أطوال الطلبة في جامعة معينة فأننا نرمز لصفة الطول (y) بينما طول أي طالب (المشاهدة) لتلك الصفة نرمز لها بالرمز (y_i)، وعليه يمكن تعريف **المتغير** على انه الظاهرة التي تظهر الاختلافات بين مفردات (مشاهدات) تلك الظاهرة، وهناك نوعين من المتغيرات هما :

1- المتغيرات الوصفية أو النوعية : وهي الظواهر أو الصفات التي لا يمكن قياسها مباشرةً بالأرقام العددية ثل صفة اللون مثل صفة (لون العين) أو الحالة الاجتماعية (أعزب، متزوج، مطلق، أرمل) أو مستوى المعيشة (غني، متوسط الدخل، فقير) .

2- المتغيرات الكمية :هي الظواهر التي يمكن قياسها مباشرةً بأرقام عددية مثل صفة الطول، الوزن، العمر، كمية المحصول المنتج ...الخ، وهناك نوعن من المتغيرات الكمية، هما :

(المتغيرات المنفصلة أو المتقطعة) و(المتغيرات المتصلة أو المستمرة)

المتغيرات المنفصلة أو المتقطعة :

هي الظاهرة التي تكون مفرداتها (المشاهدات) عبارة عن قيم قابلة للعد (متباعدة او منقطعة)، على سبيل المثال عدد الثمار المنتجة من محصول معين في إحدى الحقول الزراعية عدد الطلبة في احدى الأقسام العلمية، أطوال أو أوزان الأطفال ...الخ .

المتغيرات المتصلة أو المستمرة

وهي الظاهرة التي تكون مفرداتها(المشاهدات) عبارة عن فترة زمنية أي بمعنى لمفردات التي تقع ضمن معين أي غير قابلة للعد على سبيل المثال دراسة أطوال طلبة جامعة معينة تتراوح أطوالهم بين (130.5 – 175.5) سم .

7- المجتمع والعينة

- ✓ **المجتمع الإحصائي:** عبارة عن جميع القيم أو المفردات التي يمكن أن يأخذها المتغير (الظاهرة)، ويكون المجتمع الإحصائي على نوعين هماً **مجتمعاً محدوداً** أي هو المجتمع الذي يمكن ملاحظة جميع مفرداته والسيطرة عليها، أو **مجتمعاً غير محدود** أي هو المجتمع الذي يكون من الصعب ملاحظة و حصر جميع مفرداته .
- ✓ **العينة الإحصائية** : هي الجزء المختار (المسحوب) من مجتمع إحصائي بشرط أن تكون ممثلة لجميع خصائص ذلك المجتمع، ويتم اختيار العينة بأسلوب او طريقة يطلق عليها (بالمعاينة الإحصائية)، ويعتبر أسلوب العينة هو الأسلوب الأخر في اختيار البيانات عندما يكون من الصعب الحصول على بيانات دقيقة من المجتمع الإحصائي وبالتحديد المجتمع غير المحدود .

8- أسلوب تصميم البحوث

هناك اعتبارات كثيرة يتوقف عليها تصميم البحث العلمي التي يجب ان يأخذها الباحث بنظر الاعتبار وهي مسألة الحصول على البيانات والمعلومات بأقصر وقت واقل جهد واقل كلفة ولكي يتم تصميم البحث بأسلوب علمي صحيح يجب مراعاة ما يلي :

1. **تحديد الغرض من البحث العلمي :**

بمعنى آخر تحديد الهدف الأساسي الذي يقوم عليه البحث العلمي والذي يجب ان يكون واضحاً، دقيقاً ومحدداً لجميع أوجه الاستفادة التي سيحققها البحث والاستفادة من نتائجه .

2. **أمكانية التنفيذ العلمي للبحث :**

ان من ضروريات تنفيذ البحث العلمي هو تحديد المتطلبات الضرورية التي تلتزمها عملية تنفيذ البحث كالموارد المالية، والإمكانات البشرية المتاحة لتحقيق بعض فقرات البحث، والتأكد من مدى توافر البيانات والمعلومات الدقيقة عن مشكلة البحث .

3. **تحديد إطار البحث :**

من الضروري والمهم تحديد إطار البحث من قبل الباحث من حيث تحديد نوع وطبيعة مجال البحث، أو المجتمع الإحصائي أي تحديد الوحدات او مفردات ذلك المجتمع ومعرفة صفته او صفات المشتركة، مع تحديد نوع ذلك المجتمع أي هل هو (مجتمع محدود) او (غير محدود) .

4. تحديد الأسلوب المناسب لجمع البيانات الإحصائية وهما أسلوبين (**أسلوب الحصر الشامل**) و (**أسلوب الحصر غير الشامل (أسلوب العينات)**) .

9- مصادر جمع البيانات

هناك مصدرين لجمع البيانات هما :

1- **المصادر التاريخية :** وهي البيانات والمعلومات المسجلة لدى أجهزة ومؤسسات ودوائر الدولة وكذلك شركات القطاع الخاص والمحفوظة لسنوات سابقة نتيجة إجراء عمليات المسح الميداني و الاحصائات من قبل الجهات الحكومية .

2- **المصادر الميدانية :** وهي المعلومات والبيانات التي يتم تجميعها ميدانياً من قبل الباحث ومن مصادرها الأصلية معتمداً أما أسلوب المراسلة، المقابلة الشخصية، أو توزيع استمارات الاستبيان .

10- اساليب جمع البيانات

ان عملية جمع البيانات والمعلومات تعتبر خطوة أساسية من خطوات البناء الإحصائي للبحث العلمي، لذا تتوفر اساليب يمكن من خلالها جمع البيانات والمعلومات ولكل من هذه الأساليب مزايا وعيوب، وهذه الأساليب هي :

- **أسلوب التسجيل الشامل (الحصر الشامل) :**

وهو الأسلوب الذي يلجأ اليه الباحث في حالة المجتمعات الإحصائية المحدودة التي يمكن ملاحظة جميع مفرداته على سبيل المثال (التعداد العام للسكان) ويعتبر هذا الأسلوب هو الأفضل في عملية جمع البيانات لأنه يعطي صورة حقيقة وكاملة حول مفردات الظاهرة المراد دراستها.

أما عيوب هذا الأسلوب

- ✓ يتطلب إمكانيات مادية وبشرية كبيرة.
- ✓ يتطلب وقت وجهد كبير .

- **أسلوب الحصر غير الشامل (أسلوب العينة):**

وهو الأسلوب الذي يتم اللجوء اليه في حالة المجتمعات الإحصائية غير المحدودة التي لا يمكن حصر وملاحظة جميع مفرداته ويسمى هذا الأسلوب بالعينة بشرط أن تحمل هذه العينة كل خصائص المجتمع الإحصائي، ويمتاز هذا الأسلوب بأنه لا يحتاج إلى إمكانات مادية وبشرية ووقت وجهد كما هو الحال في الحصر الشامل .

- **الأسلوب التجريبي :**

- **أسلوب السلاسل الزمنية :**

11- المعاينة الإحصائية

تُعد المعاينة الإحصائية إحدى اساليب وطرق اختيار المفردات من المجتمع والتي تؤلف العينة الإحصائية .

تعاريف ومصطلحات

- ✓ **العينة** : هي جزء من المجتمع يتم اختيارها بطريقة علمية محددة، ويشترط بالعين المختارة ان تكون ممثلة للمجتمع تمثيلاً جيداً وتحمل كل خواصه بما في ذلك الاختلافات في وحداته .
- ✓ **حجم العينة** : جميع وحدات العينة التي مؤلفة لها ويرمز لها بالرمز (n) .
- ✓ **وحدة المعاينة**: هي الجزء الصغير الذي تجمع منه البيانات، وهي كل وحدة من وحدات المجتمع وقد تكون متشابه من حيث الحجم أو مختلفة، ويجب تعريف وحدة المعاينة تعريفاً واضحاً عند تنفيذ البحوث الميدانية بحيث يشمل جميع مفردات الظاهرة وعدم تداخلها مع المفردات الأخرى .
- ✓ **المعاينة**: هي عملية اختيار جزء من مفردات المجتمع الإحصائي للاستدلال على خواص المجتمع بأكمله عن طريق تعميم نتائج العينة،وتقوم المعاينة على علم فن التحكم وقياس دقة المعلومات الإحصائية عن طريق استخدام بعض النظريات الرياضية .

(11-1) أنواع المعاينة الإحصائية

تنقسم المعاينة الإحصائية إلى قسمين رئيسين هما :

أولاً : المعاينة العشوائية (الاحتمالية)

ثانياً: المعاينة غير العشوائية (غير الاحتمالية)

أولاً : المعاينة العشوائية (الاحتمالية)

وهي الطريقة التي يتم من خلالها اختيار مفردات العينة بصورة عشوائية بدون تدخل الباحث في اختيارها بحيث تكون لكل مفردة من مفردات العينة نفس الفرصة في الظهور أو الاختيار . ومن أنواع المعاينة العشوائية

1- المعاينة العشوائية البسيطة :

هي طريقة اختيار عينة مكونة من (n) وحدة من بين وحدات المجتمع (N) بحيث يكون لكل عينة من العينات الممكن اختيارها فرصة متساوية (احتمال متساوٍ) في الظهور، أي أن احتمال سحب أي وحدة يكون متساوياً عند اختيار كل وحدة من وحدات العينة، ولتحديد عدد العينات العشوائية البسيطة الممكن سحبها من مجتمع إحصائي هناك طريقتين هما :

- **<u>السحب بدون إرجاع</u>** : ويقصد بها أن المفردة المسحوبة لا يتم اختيارها مرة أخرى، وفي هذه الطريقة نضمن أن كل مفردة تمتلك نفس الاحتمال بالاختيار، إضافة إلى ذلك أن تمتلك كل مفردة نفس عدد مرات الظهور .

فإذا كان لدينا مجتمع إحصائي مؤلف من (N) المفردات وتم سحب عينة عشوائية بسيطة بحجم (n)، فإن عدد العينات (m) الممكن سحبها من ذلك المجتمع تحسب على وفق الصيغة التالية :

$$\mathbf{m} = \binom{\mathbf{N}}{\mathbf{n}} \Longrightarrow \boldsymbol{M} = \frac{\boldsymbol{N!}}{\boldsymbol{n!(N-n)!}} \qquad \ldots(1)$$

$$\mathbf{P_{N_i}} = \frac{\mathbf{1}}{\mathbf{N}} \qquad \ldots(2)$$

مثال : مجتمع إحصائي مؤلف من (5) مفردات سحبت منه عينة عشوائية بسيطة بحجم (2)، فما هو عدد العينات الممكن سحبها وما هو احتمال اختيار كل مفردة .

$$\mathbf{m} = \binom{\mathbf{5}}{\mathbf{2}} \Longrightarrow \boldsymbol{m} = \frac{\mathbf{5!}}{\mathbf{2!(5-2)!}} \Longrightarrow \boldsymbol{m} = \frac{\mathbf{5!}}{\mathbf{2!3!}} \Longrightarrow \boldsymbol{m} = \mathbf{10}$$

2- <u>المعاينة العشوائية الطبقية Stratified Random Sampling :</u>

ان دقة تقدير معالم المجتمع تتوقف على حجم العينة كما تتوقف على عدم تجانس المجتمع، ويمكن وضع بعض القيود على المعاينة العشوائية البسيطة للتقليل من تأثير مشكلة عدم التجانس، وابسط هذه القيود هي تقسيم المجتمع إلى عدد من الطبقات، وهذا النوع من العينات يدعى بالمعاينة العشوائية الطبقية لذا يمكن تعريفها (وهي إحدى أنواع المعاينة العشوائية الاحتمالية التي تستخدم في المجتمعات الإحصائية الكبيرة وغير المتجانسة حيث يتم تقسيم المجتمع الذي حجمه N إلى عدد من الطبقات(h) المتجانسة التي أحجامها (N_1, N_2,N_3, …,N_h) ما ومن ثم سحب عينة عشوائية بسيطة من كل طبقة)) .

وعليه :

$$\mathbf{N = N_1 + N_2 + \ldots + N_h}$$

$$n = n_1+n_2+\ldots+n_h$$

حيث ان

N : حجم المجتمع **$N_1,N_2,..N_h$: حجم الطبقات**

h: عدد الطبقات

n : حجم العينة العشوائية **$n_1, n_2, \ldots, n_h$: حجم العينة العشوائية المسحوبة من كل طبقة**

مثال(1) : **اذا كان لدينا مجتمع إحصائي مؤلف من (12) موظفاً تم تحديد الدخل الشهري لكل منهم، ونريد سحب عينة عشوائية بسيطة مؤلفة من (4) أشخاص لتقدير متوسط الدخل، حدد أسلوب المعاينة المناسب لتقدير متوسط الدخل .**

الموظفين	A	B	C	D	E	F	G	H	I	J	K	L
الدخل الشهري	21	53	12	25	31	22	24	40	17	35	55	11

نلاحظ إن المجتمع الإحصائي غير متجانس وعليه لا يمكن سحب عينة عشوائية بسيطة مباشرة لتقدير متوسط الدخل للموظفين، وعليه فان الأسلوب المناسب لتقدير متوسط الدخل هو (**أسلوب المعاينة العشوائية الطبقية**)، وعليه يتم تقسيم لمجتمع (أعداد الموظفين إلى عدد من الطبقات) كل طبقة تمثل مجتمعاً متجانساً أو متجانساً إلى حد ما ومن ثم سحب عينة عشوائية بسيط من كل طبقة .

- ولتنفيذ ذلك يتم ترتيب المجتمع ترتيباً تصاعدياً أو تنازلياً وفقاً للدخل الشهري لكي نتمكن من تحديد الطبقات بشكل دقيق .
- فإذا تم ترتيب الدخل ترتيباً تصاعدياً سنحصل على الجدول التالي :

رمز الموظفين	الدخل الشهري
L	11
C	12
I	17
A	21
F	22
G	24
D	25
E	31
J	35
H	40
B	53
K	55

اذ نلاحظ ان عملية الترتيب أوضحت أمكانية تقسيم المجتمع إلى (4) طبقات متجانسة قدر الإمكان

م.م. احمد هشام محمد طاهر – قسم الإحصاء

تصنيف الطبقات	الطبقة	الدخل الشهري	رمز الموظفين
منخفض	h1	11	L
		12	C
		17	I
متوسط	h2	21	A
		22	F
		24	G
مرتفع	h3	25	D
		31	E
		35	J
مرتفع جدا	h4	40	H
		53	B
		55	K

فإذا تم اختيار موظفاً من كل طبقة وبشكل عشوائي سنتمكن من الحصول على عينة مؤلفة من (4) موظفين وعليه من خلال أسلوب المعاينة الطبقية لم يعد هناك احتمال الحصول على عينة جميع أفرادها من ذوي الدخول المرتفعة أو المنخفضة .

ان تقسيم المجتمع الإحصائي غير المتجانس إلى طبقات متجانسة قدر الإمكان ومن ثم سحب عدد من العينات من كل طبقة يتناسب مع حجم الطبقة فان هذا الأسلوب يطلق عليه بالمعاينة الطبقية المتناسبة (التوزيع المتناسب)، ويمكن التعبير عنه رياضياً بالصيغة التالية :

$$\frac{n_h}{n} = \frac{N_h}{N} \quad ...(3)$$

مثال(2) : في إحدى كليات الإدارة الاقتصاد تم إجراء استطلاع للرأي بمستوى الخدمات التنقية والتكنولوجية المطبقة في الكلية تطلب إجراء هذا النوع من الاستطلاع استخدام أسلوب العينات الاحتمالية، فإذا علمت ان حجم العينة هو (150) طالب وطالبة وان عدد الطلبة في الأقسام العلمية كان كما يلي :

الأقسام العلمية	الإدارة	المحاسبة	الاقتصاد	الإحصاء	العلوم المالية	نظم المعلومات
عدد الطلبة	600	570	420	400	620	140

<u>حدد الحجم العينة المناسب الممكن اختياره من كل قسم</u>

م.م. احمد هشام محمد طاهر - قسم الإحصاء

مثال(3): مجتمع إحصائي مقسم إلى (3) طبقات تم الحصول على البيانات التالية :

N_h حجم الطبقة	A	750	C
n_h الجزء المختار	a	70	50

حدد قيم حجوم الطبقات والأجزاء المختارة المفقودة في الجدول إذا علمت إن حجم المجتمع N=12000 .

مثال(4): مجتمع إحصائي غير متجانس تم تقسيمه إلى **(4)** طبقات حجم كل طبقة مرتبة على التوالي(**1000، 3500، 4000، 1500**) ويراد سحب عينة عشوائية بحجم **(400)** مفردة ، حدد حجوم الأجزاء المختارة من كل طبقة بحيث تشكل العينة تمثيلاً سليماً .

مثال(5): مجتمع إحصائي مقسم إلى(3) طبقات تم سحب عينة بحجم (300) وتم الحصول على البيانات التالية :

N_h حجم الطبقة	800	Y	Z
n_h الجزء المختار	x	y	60

حدد قيم حجوم الطبقات والأجزاء المختارة المفقودة في الجدول اذا علمت ان حجم المجتمع N=5000 .

3- المعاينة العشوائية المنتظمة (الأسلوبية) Systemic Random Sampling

هي طريقة اختيار عينة منتظمة وذلك عن طريق اختيار وحدة المعاينة (M) والتي تسمى أيضاً بنسبة المعاينة، ويطلق على العينة المنتظمة بالعينة ذات الفترات المتساوية وذلك لتساوي الفترات بين مفرداتها، وتعد العينة الأسلوبية شائعة الاستعمال لسهولتها وبساطتها ولكن طريقة اختيار مفرداتها بهذه الصورة قد تعرض العينة لبعض التحيز .

إذا كان مجتمع إحصائي حجمه N سحبت منه عينة عشوائية بسيطة بحجم n فإن وحدة المعاينة K تحسب $\mathbf{M = \frac{N}{n}}$، حيث يتم اختيار رقم عشوائي بين (1 و K) لتشكل رقم المفردة الأولى (K_1) ثم يتم إضافة K ومضاعفاتها على المفردة الأولى إلى أن يكتمل حجم العينة .

ويمكن حساب مفردات العينة المنتظمة باعتماد الصيغة التالية :

$$\mathbf{K_i = K_1 + (i - 1)M}$$

مثال(6) : تم اختيار عينة أسلوبية من مجتمع إحصائي حجمها (30) مفردة حيث تم الحصول على المفردات ($K_{10} = 200, K_{20} = 460$)، المطلوب جد حجم المجتمع والمفردة الأخيرة في العينة .

مثال(7) : تم سحب عينة عشوائية بسيطة بحجم (60) مفردة من مجتمع إحصائي متجانس فإذا علمت أن المفردة $K_{25} = 520$ و المفردة $K_{35} = 600$ ، المطلوب جد حجم المجتمع والمفردتين الأخيرتين في العينة، مع تحديد المفردات K_{55}, K_{48}, K_{30} .

4- <u>المعاينة متعددة المراحل :</u>

وهو أسلوب يهدف الى تقسيم المجتمع الى وحدات اولية ثم يتم اختيار عينة عشوائية من هذه الوحدات ومن ثم يتم تقسيم كل وحدة من الوحدات الأولية الى وحدات ثانوية ثم تؤخذ عينة عشوائية منها ثم يتم تقسيم كل وحدة من الوحدات الثانوية الى وحدات اصغر وتسحب من هذه الوحدات عينة عشوائية وهكذا نستمر الى ان نصل الى الجزء الذي سيتم جمع البيانات منه لتشكل عينة البحث .

ثانياً: المعاينة غير العشوائية (غير الاحتمالية) :

1 - المعاينة الحصصية : يتم تقسيم المجتمع إلى طبقات استناداً إلى معايير تقسيم معينة تتعلق بطبيعة الدراسة ثم يتم اختيار عدد من المفردات من كل طبقة وبشكل شخصي (غير عشوائي) بحيث أن عدد مفردات هذه العينات يمثل حجم العينة المطلوبة .

2 - المعاينة العمدية : وهو الأسلوب الذي يتم فيه اختيار مفردات العينة بشكل متعمد من قبل الباحث معتقداً ان هذه العينة تمثل المجتمع خير تمثيل .

(12) **تصنيف و تبويب البيانات**

بعد التعرف على مصادر جمع البيانات وأساليب جمعها فأن البيانات التي تم الحصول عليها تدعى (بالبيانات الأولية أو البيانات غير المصنفة، إذ أن البيانات بشكلها الأولي تكون غير منظمة مما يصعب التعامل معها واستخدامها لأغراض التحليل الإحصائي، ولغرض تحقيق الفائدة من هذه البيانات لا بد من التعريف بعملية تصنيف وجمع البيانات والتي تتم من خلال المراحل التالية :

1- **مراجعة البيانات** : بعد أتمام عملية جمع البيانات من مصادرها الأصلية وفق الأسلوب المناسب يجب على الباحث مراجعة وتدقيق البيانات للتأكد من مطابقتها وتكاملها .

2- **تصنيف البيانات** : تتحقق هذه المرحلة بعد أتمام عملية المراجعة والتدقيق حيث تتم عملية التصنيف على أساس الظواهر التي جمعت منها البيانات ، فقد يكون التصنيف على أساس ظاهرة (العمر ، الوزن، التحصيل الدراسي، المهنة، ...الخ) .

3- **تبويب البيانات** : تتم هذه العملية بعد أتمام عملية تصنيف البيانات ويقصد بها تفريغ البيانات المصنفة في جداول إحصائية خاصة بحيث كل جزء من هذه البيانات المصنفة يعبر عن ظاهرة معينة والهدف

من تبويب البيانات هو هو أبراز البيانات وتوضيحها في أضيق حيز ممكن، ويكون تبويب البيانات على انواع هي :

أ- **التبويب الزمني** : تجميع البيانات المصنفة وترتيبها في جداول إحصائية على أساس وحدات زمنية (يوم، شهر، سنة، ...الخ) مثال ذلك جدولة أعداد الطلبة الخريجين لكلية الإدارة والاقتصاد للسنوات من (2015-2019)

السنوات	عدد الطلبة الخريجين
2015	1500
2016	2000
2017	1600
2018	1400
2019	2000

ب- **التبويب الجغرافي** : تجميع البيانات المصنفة وجدولتها بالاعتماد على الوحدة الجغرافية (المكانية) او تقسيم أداري معين (قارات، بلدان، محافظات، أقضية، النواحي)، مثال تحديد أعداد النخيل في المحافظات الجنوبية والوسطى (البصرة، ذي قار، ميسان، كربلاء، النجف)

المحافظات	أعداد النخيل
البصرة	10000
ذي قار	6500
ميسان	3000
كربلاء	4750
النجف	6570

ت- **التبويب الكمي** : تجميع البيانات حسب صفة يمكن قياسها (الوزن، العمر، الطول، المساحة، الحجم، قيمة الأجرة اليومية للعاملين...الخ)

ث- **التبويب النوعي** (على أساس صفة معينة) :ويتم فيه تجميع البيانات المصنفة في جداول باعتماد صفة لا يمكن قياسها مثل (لون العين، الجنس، الحالة الاجتماعية ، المستوى العلمي، العنوان الوظيفي..الخ) .

م.م. احمد هشام محمد طاهر - قسم الإحصاء

(13) عرض البيانات الإحصائية

هناك أسلوبين لعرض البيانات الإحصائية هما :

1- **العرض الجدولي:** يتم في هذا النوع تلخيص البيانات وعرضها في جداول إحصائية أما بسيطة أو مركبة

- **الجداول البسيطة : يعبر عن الجداول الإحصائية بالجداول البسيطة إذا كانت تحتوي على صفة واحدة فقط .**
- **الجداول المركبة : وهي الجداول التي تحتوي على صفتين أو أكثر .**

2- **العرض البياني** : في هذا النوع تعرض البيانات الإحصائية بصورة أشكال هندسية أو رسوم تصويرية أو خرائط إحصائية ومن طرق العرض البياني هي :

- **المستطيل البياني**
- **الأشرطة البيانية**
- **الخط البياني**
- **الدائرة البيانية**

- **المستطيل البياني :**

ويستعمل المستطيل البياني لتمثيل ظاهرة معينة أو عدت ظواهر من خلال المقارنة فيما بينها، وتتلخص فكرة المستطيل البياني في رسم قاعدة للمستطيل ومن ثم تقسيمها الى مستطيلات جزئية تمثل بيانات الظاهرة، ولغرض حساب مساحة أي ظاهرة نعتمد الصيغة التالية :

$$Y_i = A * \frac{x_i}{\sum x_i}$$

حيث أن :

Y_i : تمثل مساحة الظاهرة

A : طول قاعدة المستطيل الكلية

x_i: بيانات الظاهرة

$\sum x_i$:المجموع الكلي للظواهر

م.م. احمد هشام محمد طاهر – قسم الإحصاء

مثال: بلغت كلفة إنتاج النفط الأبيض في إحدى المصافي (300) دولار للبرميل الواحد، وكما موضحة بالجدول :

مستلزمات الإنتاج	كلفة الإنتاج
أجور العاملين	130
مواد خام	40
مصاريف مباشرة	75
مصاريف غير مباشرة	55

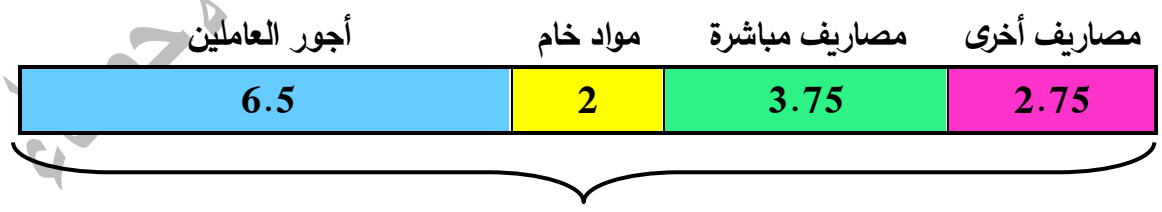
فإذا فرضنا ان طول قادة المستطيل هي (15)سم، مثل البيانات التالية باستخدام الرسم البياني

الحل

$$\text{اجور العاملين} = 15 * \frac{130}{300}$$
$$= 6.5$$

وهكذا يتم حساب مساحة مستلزمات الإنتاج المتبقية

مستلزمات الإنتاج	المساحة
أجور العاملين	6.5
مواد خام	2
مصاريف مباشرة	3.75
مصاريف أخرى	2.75
مجموع المساحة الكلية	15



المساحة الكلية 15 سم

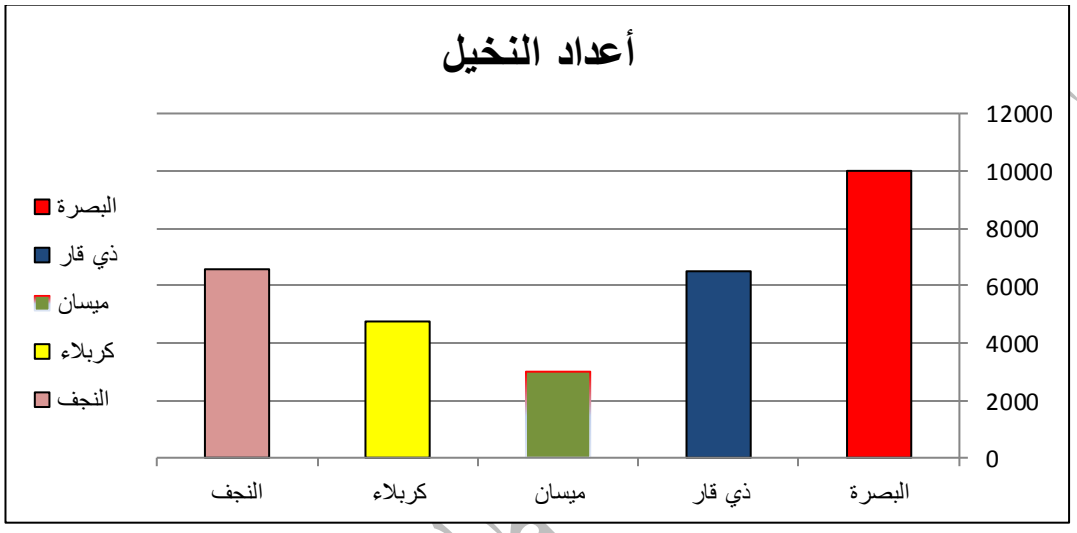
- **الأشرطة البيانية**

هي مجموعة من المستطيلات قواعدها تكون متساوية وارتفاعاتها مختلفة بحسب طبيعة البيانات ولغرض توضيح ذلك نأخذ المثال التالي :

مثال : في دراسة لتحديد أعداد النخيل في المحافظات الجنوبية والوسطى (البصرة، ذي قار، ميسان، كربلاء، النجف) تم الحصول على البيانات التالية :

المحافظات	أعداد النخيل
البصرة	10000
ذي قار	6500
ميسان	3000
كربلاء	4750
النجف	6570

المطلوب استخدم الرسم بالأشرطة البيانية لتوضيح أعداد النخيل حسب المحافظات



- **الدائرة الإحصائية**

هي أحد أشكال الرسوم البيانية التي تتناسب فيها زوايا الأشكال مع القيم المراد التعبير عنها)، ويسهل فهمه بنظرة سريعة إذ تعبر أجزاءه عن قيمة الأجزاء بالنسبة للكل (أي نسبتها المئوية) بشكل بسيط، يتم استخدام المخططات الدائرية في توضيح نتائج استطلاعات الرأي والإحصائيات والبيانات المعقدة ونسب الدخل أو الإنفاق. ويتم تقسيم الدائرة البيانية بتقسيم الزاوية المركزية لها والتي تساوي (360°)، كما في الصيغة التالية :

$$Y_i = 360^{\circ} * \frac{x_i}{\sum x_i}$$

حيث أن :

Y_i : زاوية القطاع

x_i: بيانات الظاهرة

$\sum x_i$: المجموع الكلي للظواهر

ويمكن توضح ذلك بالرجوع إلى بيانات المثال السابق الخاص بإعداد النخيل حسب المحافظات

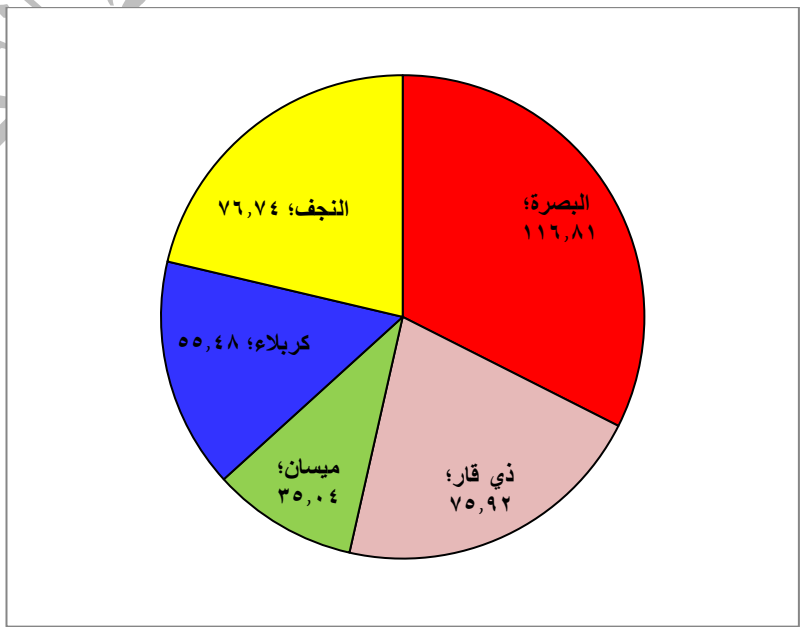
المحافظات	أعداد النخيل
البصرة	10000
ذي قار	6500
ميسان	3000
كربلاء	4750
النجف	6570
المجموع الكلي للنخيل	30820

الحل

$$\text{البصرة} = 360^{\circ} * \frac{10000}{30820}$$
$$= 116.81$$

وهكذا يتم حساب المساحة لجميع المحافظات

المحافظات	أعداد النخيل
البصرة	**116.81**
ذي قار	**75.92**
ميسان	**35.04**
كربلاء	**55.48**
النجف	**76.74**



م.م. احمد هشام محمد طاهر - قسم الإحصاء

(14) العرض الجدولي

(14-1) التوزيع التكراري :

تلخص البيانات الإحصائية سواء كانت وصفية او كمية في جداول تسمى (بالتوزيعات التكرارية) **Frequency Distribution** وهي عبارة عن جداول يلخص البيانات الخام (الأصلية) فيوزعها على عمودين الأول يمثل الفئات والثاني يمثل تكرارات تلك الفئات ويرمز لهذا العمود **(f)** .

ولغرض تكون جدول التوزيع التكراري يتطلب ذلك تكوين جدول أولي لتفريغ البيانات الأصلية وهو يتكون من ثلاث أعمدة الأول يمثل (الفئات) أو ما تسمى بصفات الظاهرة لتضمين صفة البيانات كمية كانت ام وصفية، العمود الثاني يمثل عمود العلامات ويشمل بحزمة من خمسة خطوط (𝍸) أما العمود الثالث فيمثل مجموع العلامات المتحققة في العمود الثاني .

بعض التعاريف والمصطلحات :

(1)البيانات غير المبوبة: وهي البيانات الأولية أو الأصلية التي تم جمعها من مصادرها ولكن لم يتم تبويبها .

(2)البيانات المبوبة : وهي البيانات التي رتبت ونضمت في جداول إحصائية تسمى (جداول التوزيع التكراري) .

(3)الفئات : وهي عدد من المجاميع التي تمثل صفة (ظاهرة معينة) والتي قسمت إليها قيم المتغير بحيث كل فئة تأخذ مدى معين من قيم المتغير .

أولاً: تكوين جدول توزيع تكراري لبيانات ذات صفة معينة أو بيانات رتبية

لتكوين جدول توزيع تكراري لبيانات تحمل صفة معينة (تقديرات الطلبة، لون العين، الحالة الاجتماعية، ...الخ) لا نحتاج في هذا النوع من البيانات أي قوانين رياضية وانما نحتاج فقط الى ترتيب هذه البيانات (تصاعدي او تنازلي) أو ترتيبها وفق اليه يحددها الباحث .

مثال/ إذا كانت لدينا بيانات تمثل تقديرات (60) طالب معرفة بالشكل التالي :

D	B	E	C	D	B	D	C	E	A
B	E	C	D	B	D	D	A	E	C
C	D	A	C	E	D	C	C	D	B
D	E	D	D	A	D	D	C	D	C
D	A	B	D	B	D	C	D	C	E
D	B	C	C	E	D	C	C	D	A

حيث تم توصيف التقديرات بالشكل التالي :

التقدير	A	B	C	D	E
وصفه	امتياز	جيد جدا	جيد	متوسط	مقبول

م// كون جدول توزيع تكراري لتقديرات الطلبة .

الفئات	العلامات	F (التكرارات)
A	𝍸 |	6
B	𝍸 |||	8
C	𝍸 𝍸 𝍸 |	16
D	𝍸 𝍸 𝍸 𝍸 ||	22
E	𝍸 |||	8
مجموع التكرارات		60

وعليه يكون جدول التوزيع التكراري

الفئات تقديرات الطلبة	التكرارات (f)
A	6
B	8
C	16
D	22
E	8

ثانياً : تكوين جدول التوزيع التكراري للبيانات الكمية

لتكوين جداول توزيع تكراريه تحمل فئاتها الصفة الكمية، يجب توضح بعض المفاهيم التالية :

الفئة : وهي عبارة عن مجاميع قسمت اليها قيم المتغير ضمن مدى معين، بحيث يكون لهذه الفئة حد ادني وحد أعلى .

المدى : وهو الفرق بين اكبر قيمة واقل قيمة بين مفردات الظاهرة، ويرمز له بالرمز R .

$$R=(X_{MAX}-X_{MIN})+1$$

$$\text{Or} \quad R=(X_{MAX}-X_{MIN})$$

مركز الفئة: وهي القيمة الواقعة في منتصف الفئة، وتحسب كما يلي :

$$X=\frac{L_c+U_c}{2}$$

L_c : الحد الأدنى للفئة

U_c : الحد الأعلى للفئة

عدد الفئات: هي عدد من المجاميع التي قسمت إليها قيم المتغير بحيث تحتوي هذه المجاميع جميع حالات الظاهرة ويكون عدد الفئات محصورة بين (5−20) حسب طبيعة البيانات ويجب أن يكون عدد الفئات اقل

م.م. احمد هشام محمد طاهر - قسم الإحصاء

من طول الفئة ويرمز لعدد الفئات بالرمز K، ويمكن تقدير عدد الفئات بشكل تقريبي باعتماد إحدى الصيغتين :

- طريقة سترجس

$$K = 1 + 3.3log(n)$$

- طريقة يول

$$K = 2.5 * \sqrt[4]{n}$$

حيث ان : (n) تمثل عدد مفردات الظاهرة .

طول الفئة : هو الفرق بين فئتين متتاليتين أو الفرق بين مركزي فئتين متتاليتين، ويرمز لها بالرمز L، كما يمكن حسابه على وفق الصيغة التالية :

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

الحدود الحقيقية للفئات :

- الحد الحقيقي الأدنى = مركز الفئة $-\frac{1}{2}$ (طول الفئة) .
- الحد الحقيقي الأعلى = مركز الفئة $+\frac{1}{2}$ (طول الفئة) .

مثال : كون جدول توزيع تكراري للأعداد التالية التي تمثل أوزان عدد من الطلبة في المعهد التقني في المنطقة الوسطى :

66	73	80	60	56	74
65	63	89	77	55	56
57	60	77	60	67	50
80	64	70	59	62	70
69	70	72	74	61	57

الحل :

نرتب البيانات ترتيباً تصاعديا أو تنازلياً او بطريقة أخرى نكتب البيانات في جدول مقسم كما في أعلاه ومن ثم نستخرج اقل وأكبر قيمة في كل صف وبعد ذلك يتم اختيار اقل القيم من عمود اقل قيمة واكبر القيم من عمود اكبر قيمة وكما يلي :

اقل قيمة	أوزان الطلبة/بالكيلو غرام						اكبر قيمة
56	66	73	80	60	56	74	80
55	65	63	89	77	55	56	89 ← أكبر القيم
50 ← أقل القيم	57	60	77	60	67	50	77
59	80	64	70	59	62	70	80
57	69	70	72	74	61	57	74

(1)حساب المدى

$R=(X_{MAX}-X_{MIN})+1$

X_{MAX}: اكبر قيمة = 89 X_{MIN} : اقل قيمة = 50

$R= (89-50)+1$

$R=40$

(2)عدد الفئات (K)

- استخدام طريقة سترجس

$$K = 1 + 3.3log(n)$$
$$K = 1 + 3.3log(30)$$
$$K = 1 + 3.3(1.48)$$
$$K = 5.88 \approx 6$$

- طريقة يول

$$K = 2.5 * \sqrt[4]{n}$$
$$K = 2.5 * \sqrt[4]{30}$$
$$K = 2.5 * (2.34)$$
$$K = 5.85 \approx 6$$

(3)طول الفئة

$$L = \frac{R}{K}$$
$$L = \frac{40}{6}$$
$$L = 6.67 \approx 7$$

(4)تكوين جدول العلامات

الفئات أوزان الطلبة/بالكيلوغرام	العلامات	التكرارات
50 – 56	𝍸	4
57 – 63	𝍸 𝍸	9
64 – 70	𝍸 𝍸	8
71 – 77	𝍸	6
78 – 84	𝍸	2
85 – 91	𝍸	1

ملاحظة / يمكن تحديد العلامات من خلال متابعة القيمة في الجدول الأصلي وتحديد موقعها في أي فئة وتأشيرها لضمان عدم الرجوع اليها .

م.م. احمد هشام محمد طاهر – قسم الإحصاء

66	73	80	60	56	74
65	63	89	77	55	56
57	60	77	60	67	50
80	64	70	59	62	70
69	70	72	74	61	57

(5) جدول التوزيع التكراري

الفئات أوزان الطلبة/بالكيلوغرام	التكرارات
50 – 56	4
57 – 63	9
64 – 70	8
71 – 77	6
78 – 84	2
85 – 91	1
المجموع الكلي للتكرارات	30

واجب / إذا كانت لديك البيانات التالية التي تمثل أطوال عدد من نباتات الظل في احد المشاتل مقاسه بالسنتمتر وهي كما يلي :

10	7	20	60	30
35	40	25	15	10
50	31	21	55	45
39	17	22	18	52
38	14	42	51	47

ويمكن تكوين جدولين من جدول التوزيع التكراري هما :

1- جدول التوزيع التكراري النسبي Relative Frequency Table

2- جدول التوزيع التكراري المئوي Percentage Frequency Table

م.م. احمد هشام محمد طاهر - قسم الإحصاء

(1)جدول التوزيع التكراري النسبي :

ويتكون هذا الجدول من خانتين مثل جدول التوزيع التكراري هما خانة الفئات أما خانة التكرارات (f_i) تسمى بأسم (خانة التكرار النسبي) ويرمز لها بالرمز (rf_i)، ويمكن حساب التكرار النسبي لكل فئة من خلال الصيغة التالية :

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

حيث ان :

f_i : التكرار الفعلي للكل فئة

$\sum_{i=1}^{k} f_i$: المجموع الكلي لتكرارات الفئات ، k: تمثل عدد الفئات في جدول التوزيع التكراري

rf_i : التكرار النسبي لكل فئة

$\sum_{i=1}^{k} rf_i$: المجموع الكلي للتكرار النسبي للفئات

بحيث يجب أن يكون مجموع التكرارات النسبية مساوية إلى الواحد الصحيح، أي أن $\sum_{i=1}^{k} rf_i = 1$

مثال / إذا كان لديك جدول التوزيع التكراري الذي يمثل اوزان طلبة المعهد التقني/ الجامعة التقنية الوسطى والمبينة في أدناه :

الفئات أوزان الطلبة/بالكيلوغرام	التكرارات f_i
50 – 56	4
57 – 63	9
64 – 70	8
71 – 77	6
78 – 84	2
85 – 91	1

المطلوب // كون جدول توزيع تكراري نسبي

الحل /

نحسب مجموع التكرارات الكلي للفئات $\sum f_i = 30$

ثم نطبق صيغة حساب التكرار النسبي لكل فئة $rf_i = \frac{f_i}{\sum_{i=1}^{k} f_i}$

نحسب التكرار النسبي للفئة الأولى $rf_1 = \frac{4}{30} \Rightarrow rf_1 = 0.13$

الفئات أوزان الطلبة/بالكيلوغرام	التكرارات f_i	rf_i
50 – 56	4	0.13
57 – 63	9	0.30
64 – 70	8	0.27
71 – 77	6	0.20
78 – 84	2	0.07
85 – 91	1	0.03

وعليه يكون جدول التكرار النسبي

م.م. احمد هشام محمد طاهر – قسم الإحصاء

الفئات أوزان الطلبة/بالكيلوغرام	rf_i
50 – 56	0.13
57 – 63	0.30
64 – 70	0.27
71 – 77	0.20
78 – 84	0.07
85 – 91	0.03
$\sum_{i=1}^{k} rf_i$	1

(2)جدول التوزيع التكراري المئوي :

ويمكن الحصول عليه من خلال جدول التوزيع التكراري النسبي وذلك بضرب عمود التكرار النسبي لكل فئة في (100)، ويعبر عنها رياضياً

$$pf_i = rf_i * 100$$

حيث أن :

pf_i: **هو التكرار المئوي لكل فئة**

ومن المثال السابق تم الحصول على جدول التوزيع التكراري النسبي لاوزان الطلبة/ بالكيلوغرام وكما مبين في ادناه :

الفئات أوزان الطلبة/بالكيلوغرام	rf_i
50 – 56	0.13
57 – 63	0.30
64 – 70	0.27
71 – 77	0.20
78 – 84	0.07
85 – 91	0.03

المطلوب // حساب جدول التكرار المئوي

الحل//

نسحب التكرار المئوي للفئة الأولى (50 – 56) :

$$pf_1 = rf_1 * 100$$
$$pf_1 = 0.13 * 100$$
$$pf_1 = 13$$

الفئات أوزان الطلبة/بالكيلوغرام	التكرار النسبي rf_i	التكرار المئوي pf_i
50 – 56	0.13	13
57 – 63	0.30	30
64 – 70	0.27	27
71 – 77	0.20	20
78 – 84	0.07	7
85 – 91	0.03	3

ملاحظة // يجب أن يكون المجموع الكلي للتكرارات المئوية مساوية (100)

م.م. احمد هشام محمد طاهر – قسم الإحصاء

وعليه يكون جدول التوزيع التكرار المئوي :

الفئات أوزان الطلبة/بالكيلوغرام	التكرار المئوي pf_i
50 – 56	13
57 – 63	30
64 – 70	27
71 – 77	20
78 – 84	7
85 – 91	3
$\sum pf_i$	100

مثال : الجدول التالي يمثل حجم المبيعات اليومية في إحدى المحلات التجارية بالدولار خلال فترة ثلاث أشهر، ولغرض الوقوف بشكل دقيق حول حجم المبيعات، كون ما يلي :

1- جدول توزيع تكراري لتحديد حجم المبيعات .

2- جدول توزيع تكراري نسبي .

3- جدول توزيع تكراري مئوي .

26.9	77.8	8.14	9.33	46.3	71.5	46	7.9	64.5	87.8
33.7	27	71.8	33.8	59.1	28.8	80.2	81.3	43.7	41.6
7.1	64	58.3	37.9	72.2	79.9	54.3	89.7	61.3	21.5
70.1	67	15.5	2.2	28.6	31.7	4.9	26.8	99.2	0.3
10.5	28.3	99.3	14.1	47	98.1	20.3	44.2	21.7	66.3
3.2	6.2	67	39.6	39.2	62.3	21.9	72.2	67.5	64
93.7	51	40	20.2	73.9	18.2	25.5	24.5	67.3	16.4
25.4	35.9	26.5	83.1	88	37.8	19.1	80	45.4	73.8
73.9	63.8	54.2	23.2	53	52.3	63.6	46.2	24.9	55.6

الحل//

نلاحظ حجم المبيعات (القيم) هي قيم ذات مراتب عشرية وهنا لا يجوز التقريب لأنه سيؤثر على البيانات وطبيعة الظاهرة المدروسة، وعليه سنوضح خطوات تكوين جدول توزيع تكراري لهكذا بيانات:

(1)حساب المدى R :

اصغر قيمة في كل صف	حجم المبيعات اليومية/ بالدولار										أكبر قيمة في كل صف
7.9	26.9	77.8	8.14	9.33	46.3	71.5	46	7.9	64.5	87.8	87.8
27	33.7	27	71.8	33.8	59.1	28.8	80.2	81.3	43.7	41.6	81.3
7.1	7.1	64	58.3	37.9	72.2	79.9	54.3	89.7	61.3	21.5	89.7
0.3	70.1	67	15.5	2.2	28.6	31.7	4.9	26.8	99.2	0.3	99.2
10.5	10.5	28.3	99.3	14.1	47	98.1	20.3	44.2	21.7	66.3	99.3
3.2	3.2	6.2	67	39.6	39.2	62.3	21.9	72.2	67.5	64	72.2
16.4	93.7	51	40	20.2	73.9	18.2	25.5	24.5	67.3	16.4	93.7
19.1	25.4	35.9	26.5	83.1	88	37.8	19.1	80	45.4	73.8	88
23.2	73.9	63.8	54.2	23.2	53	52.3	63.6	46.2	24.9	55.6	73.9

أقل القيم ←

→ أكبر القيم

$R=(X_{MAX}-X_{MIN})+1$

X_{MAX}: اكبر قيمة = 99.3 X_{MIN} : اقل قيمة = 0.3

$R= (99.3-0.3)+1$

$R=100$

(2)عدد الفئات (K)

- **استخدام طريقة سترجس**

$$K = 1 + 3.3log(n)$$
$$K = 1 + 3.3log(90)$$
$$K = 1 + 3.3(1.95)$$
$$K = 7.4 \approx 7$$

- **طريقة يول**

$$K = 2.5 * \sqrt[4]{n}$$
$$K = 2.5 * \sqrt[4]{90}$$
$$K = 2.5 * (3.08)$$
$$K = 7.7 \approx 8$$

ونظرا لوجود اختلاف في عدد الفئات بين الطريقتين نتجه نحو افتراض إن عدد الفئات مساوية إلى (K=5) فكلما كان عدد الفئات قليل كان هو الأفضل .

(3)طول الفئة

$$L = \frac{R}{K}$$
$$L = \frac{100}{5}$$
$$L = 20$$

(4)تكوين جدول العلامات

الفئات حجم المبيعات/بالدولار	حدود الفئات	العلامات	التكرارات
0 – 20	$0 \leq C_1 < 20$	𝍸 𝍸 ||||	14
20 – 40	$20 \leq C_2 < 40$	𝍸 𝍸 𝍸 𝍸 |	26
40 – 60	$40 \leq C_3 < 60$	𝍸 𝍸 𝍸 ||	17
60 – 80	$60 \leq C_4 < 80$	𝍸 𝍸 𝍸 𝍸 ||	22
80 – 100	$80 \leq C_5 < 100$	𝍸 𝍸 |	11

م.م. احمد هشام محمد طاهر – قسم الإحصاء

26.9	77.8	8.14	9.33	46.3	71.5	46	7.9	64.5	87.8
33.7	27	71.8	33.8	59.1	28.8	80.2	81.3	43.7	41.6
7.1	64	58.3	37.9	72.2	79.9	54.3	89.7	61.3	21.5
70.1	67	15.5	2.2	28.6	31.7	4.9	26.8	99.2	0.3
10.5	28.3	99.3	14.1	47	98.1	20.3	44.2	21.7	66.3
3.2	6.2	67	39.6	39.2	62.3	21.9	72.2	67.5	64
93.7	51	40	20.2	73.9	18.2	25.5	24.5	67.3	16.4
25.4	35.9	26.5	83.1	88	37.8	19.1	80	45.4	73.8
73.9	63.8	54.2	23.2	53	52.3	63.6	46.2	24.9	55.6

وعليه يكون جدول التوزيع التكراري

الفئات حجم المبيعات/بالدولار	التكرارات f_i
0 – 20	14
20 – 40	26
40 – 60	17
60 – 80	22
80 – 100	11
$\sum f_i$	90

وسيتم حساب التكرار النسبي والمئوي بنفس الجدول كما في أدناه :

الفئات حجم المبيعات/بالدولار	التكرارات f_i	التكرار النسبي rf_i	التكرار المئوي pf_i
0 – 20	14	$\frac{14}{90} = 0.16$	$0.16 * 100 = 16$
20 – 40	26	0.29	29
40 – 60	17	0.19	19
60 – 80	22	0.24	24
80 – 100	11	0.12	12
المجموع الكلي	$\sum f_i = 90$	$\sum rf_i = 1$	$\sum pf_i = 100$

م.م. احمد هشام محمد طاهر - قسم الاحصاء