البـــاب الأول

مدخـــل في علـــم الإحصاء

(Introduction to Statistics)

1.1 مقدمـــــة .

2.1 طبيعـــة البيانات الإحصائيـــة .

3.1 الوحدة الإحصائيـــة والمجتمعـــات الإحصائيـــة .

4.1 المعلومات الإحصائية .

5.1 خطوات البحث الإحصائي .

6.1 تصنيف البيانات الإحصائية .

1.6.1 مراجعة البيانات .

2.6.1 تصميم الجداول .

3.6.1 تبويب البيانات .

4.6.1 جداول التوزيعات التكرارية .

1.4.6.1 جداول التوزيعات التكرارية المزدوجة .

7.1 التمثيل والعرض البياني للبيانات .

8.1 التمثيـــل البيانـــــي للتوزيعـــــات التكراريـــة .

9.1 تمـــــارين .

obeikandl.com

1.1 مقدمـــــة (Introduction)

الإحصاء (Statistics) هو العلم الذي يدرس كيفية جمع المعلومات مـــن المجتمعات الإحصائية المختلفة سواء بالعد الشامل أو بالمعاينة وكيفية تحويـــل هذه المعلومات إلى بيانات رقمية في جداول إحصائية ، بالإضافة إلى الأساليب المختلفة التي يمكن استخدامها لتحليل هذه البيانات تحليلا رياضـــيا لاســـتنتاج المقاييس المختلفة مثل المتوسط والانحراف المعياري ، وعند حساب المقاييس أو المعاملات مثل معامل الارتباط ومعامـــل الانحـــدار ومعامـــل الاخـــتلاف وغيرها ، أو المؤشرات التي تدل على الاتجاهات الزمنية مثل الأرقام القياسية المختلفة ثم كيفية إجراء الاختبارات المختلفة علـــى المقـــاييس والمعـــاملات المستنتجة من عيناتٌ للحكم على معنوياتها وتحديد أخطائها عند درجات الثقـــة المختلفة . وأخيرا كيفية تفسير النتائج التي نصل إليها باستخدام هذه الأساليب في التحليل ثم توضيحها في تقرير نهائي عن موضوع الدراسة الـــذي أرادنـــا دراسته باستخدام الطريقة الإحصائية .

يتبين لنا بذلك أن علم الإحصاء يعنى بالأساليب الإحصائية التي يلجأ إليها الباحث سواء في العلوم الطبيعية أو الاجتماعية للتعرف على الحقائق الخاصة بالظواهر والمشاكل موضع البحث ، ولذلك نستطيع أن نعرفه بأنه العلم الـــذي يدرس إحدى طرق البحث العلمي ، حيـــث يوضح الخطوات التي تتبع عندما يتقرر اســـتخدام الطريقة الإحصائية كمنهاج للبحث في أي من الأبحاث العلمية أو بمعنى آخر هو العلم الذي يدرس كيفية جمع المعلومات ثم تبويبها وتحليلها إحصائيا ثم عرض النتائج في رسوم بيانية ومقاييس ومؤشرات ومعاملات مع التفسير الخاص بدلالة كل منها .

2.1 طبيعة البيانات الإحصائية (Nature of Statistical Data)

تقوم البيانات الإحصائية على المعالجة الرياضية للمعلومات الخاصة بعـــدد كبير من الوحدات الإحصائية ، أي الخاصة بمجتمعات إحصائية ضخمة عندما

تكون في شكل رقمي فقط . وبشكل عام نلاحظ أن المعلومات الخاصة بالنواحي التجارية تكون أساس معلومات رقمية مثل النقد المتداول ، المبيعات ، المخزون ، الودائع ، القروض ، رؤوس الأموال ، الكميات المنتجة ، الصادرات ، الواردات ، الدخل القومي ، أسعار الفائدة ، أسعار الصرف وغيرها .

إلاّ أن الكثير من المعلومات الأخرى تكون وصفية غير رقمية ، وبذلك يصبح من الضروري البدء بتحويلها إلى بيانات رقمية حتى يمكن إجراء التحليل الإحصائي الذي يتقرر إتباعه . ولاشك أن نجاح أية دراسة يتوقف إلى حد كبير على قدرة القائم على هذة الدراسة على صياغة الأسئلة التي تتضمنها الاستمارة الإحصائية التي تستخدم في جمع المعلومات الخاصة ، بحيث تساعد في الحصول على إجابات يمكن تحويلها إلى أرقام تصلح لأن تكون مادة لإجراء التحليل الإحصائي الذي يتفق وموضوع الدراسة والبحث .

فمثلا ، يمكن التعبير عن الصحة بعدد أيام المرض التي عاناها أفراد المجتمع الإحصائي موضوع الدراسة . كما يمكن التعبير عن الذكاء بإجراء اختبارات قياسية معينة . وكذلك يمكن التعبير عن آراء الناس بالنسبة لمشكلة ما بأرقام معينة . كما يمكن التعبير عن فئاتهم الوصفية بأرقام تدل عليها . ولا شك في أن هذه المشكلة تواجه الباحث بشكل خاص في أبحاث السوق حيث يتحتم عليه أن يوجه أسئلة إلى المستجوبين في المجتمع الإحصائي موضوع دراسته للحصول إلى إجابات وصفية ، وبذلك تظهر المهارة في كيفية صياغة هذه الأسئلة بالصيغة التي تؤدي إلى إجابات محددة يمكن تحويلها بسهولة وبدقة إلى أرقام تدل عليها وتكون في نفس الوقت قابلة للتحليل الإحصائي الذي يرغب في إجرائه . ويلاحظ في هذا الصدد أمرين ، الأمر الأول يتعلق بنتائج الدراسات والأبحاث الإحصائية التي نظرا لكونها تعتمد على القياس الرقمي والتحليل الرياضي تكون تبعا لذلك ذات مضمون

موضوعي ، إلاّ أن هذه النتائج من ناحية أخرى لابد وأن تتأثر في النهاية بالتفسيرات الشخصية للقائمين بهذه الدراسات والأبحاث . فعند دراسة الارتباط بين سعر سلعة ما والطلب عليها مثلا ، قد يعتمد الدارس والباحث نموذجا رياضيا بسيطا يتضمن هذين المتغيرين فقط ، وبذلك بالرغم من أنه قد يصل إلى مقياس يدل على وجود ارتباط شديد جدا بين هذين المتغيرين ، إلاّ أن هذه النتيجة قد تكون مضللة بعض الشيء نظرا لإهمال المتغيرات الأخرى الكثيرة التي يمكن أن يكون لها تأثير على طلب السلعة ، مثل دخول المستهلكين وأسعار السلع الأخرى البديلة وغير ذلك من المتغيرات .

أما الأمر الثاني فيتعلق بدرجة دقة نتائج الدراسات والأبحاث الإحصائية حيث أنه بالرغم من أنها تظهر بشكل رقمي إلاّ أننا لا نستطيع أن نضفي عليها صفة الدقة الكاملة مثل نتائج التحليل الرياضي البحت ، ذلك لأن النتائج الإحصائية تعتمد على معلومات نحصل عليها بتوجيه أسئلة معينة إلى وحدات المجتمع موضوع الدراسة ولا نستطيع أن نجزم بان جميع الإجابات تمثل الواقع تماما إذ لا بد أن يتسرب إليها شي من عدم الدقة .

أن هذا الأمر ذلك لا يعني عدم جدوى نتائج الدراسات التي تجرى باستخدام المنهاج الإحصائي ، حيث أن الباحث بمهارته وخبرته يستطيع أن ينقص الأخطاء التي يمكن أن تتسرب إلى المعلومات التي يبنى عليها التحليل الإحصائي إلى أدنى حد ممكن . وفي هذا المجال يلعب الوعي الإحصائي الذي يسود المجتمع العامل فيه الباحث دورا أساسيا في التأثير على مجرى الدراسات الإحصائية وعلى دقة نتائجها وأهميتها العملية تبعا لذلك . ومن ناحية أخرى يستطيع الباحث باستخدام أساليب إحصائية معينة أن يحدد مدى الخطأ في النتائج التي يصل إليها ودرجة الثقة في هذه النتائج ، خاصة عند استخدام المعاينة لجمع المعلومات التي يرغب في الحصول عليها .

3.1 الوحدة الإحصائية والمجتمعات الإحصائية
(Statistical Unit and Statistical Societies)

أن الدراسات العلمية التي تعتمد الأسلوب الإحصائي منهاجا للبحث تبدأ عادة بتحديد المجتمع الإحصائي موضوع الدراسة وبالتالي الوحدة الإحصائية التي تكون هذا المجتمع . ففي دراسة عن الصناعة مثلا ، يجب تحديد المؤسسات الصناعية التي تتضمنها الدراسة سواء من ناحية نوع نشاطها أو من ناحية حجمها قياسا بعدد العاملين فيها أو برؤوس الأموال الموظفة فيها وغيرها . كذلك عند دراسة مثلا ، ميزانية الأسرة فانه يجب تحديد الأسر التي تتضمنها الدراسة سواء من ناحية سكنها في الريف أو في المدن أو فيهما سوية ، وكذلك من ناحية مستوى دخلها وغيرها .

بذلك يتبين لنا أن المجتمع الإحصائي هو مجموع الوحدات مهما كان نوعها ، أفراد أو أسر أو مؤسسات أو مساكن أو مدارس والتي تكون موضوع الدراسة التي يرغب الباحث في القيام بها . ولاشك أن تحديد أي مجتمع إحصائي يستلزم حتما تعريف دقيق للوحدة الإحصائية التي يتكون منها هذا المجتمع . فإذا كنا بصدد دراسة المؤسسات الصناعية يجب قبل البدء بجمع المعلومات المطلوبة تحديد ما هو المقصود بمؤسسة صناعية ، وكذلك توضيح المشاكل التي يمكن أن نواجهها عند التعرف على المؤسسات الصناعية ميدانيا ، مثل مؤسسة تعمل في الزراعة والصناعة سويا ، أو مؤسسة لها فروع في أقسام إدارية مختلفة في الدولة موضوع الدراسة ، أو مؤسسة تعمل في نشاطات مختلفة ليست جميعا موضع البحث ، ووضع الحلول التي يجب أن يتبعها القانون عند مواجهة أي من هذه المشاكل .

وحتى إذا كان بالإمكان جمع المعلومات الميدانية على أسس موحدة فإنه لا يكفي للتحديد الوحدة الإحصائية والمجتمع الإحصائي الذي تستهدفه الدراسة ، بل يجب كذلك تعريف كل المصطلحات التي تتضمنها أسئلة البحث

تعريفا واضحا تماما ولا شك أن عجز الباحث عن إعداد هذه التعاريف الواضحة غير المبهمة يجعل تلك المعلومات التي تجمع من وحدات المجتمع الإحصائي موضوع الدراسة خاضعة للتفسيرات الشخصية من قبل المستجوبين أو من قبل القائمين بالعمل ميدانيا وبذلك لا يمكن المقارنة بينهما .

كذلك لا يمكن تصنيفها وتبويبها على أسس موحدة ، فإذا أجري مثل هذا التصنيف والتبويب على معلومات من هذا النوع يكون التحليل الإحصائي الذي يجري عليها فيما بعد عملا غير دقيق حيث يؤدي إلى نتائج مضللة وذلك لأن الإجابات التي نحصل فيها على أسئلة مبهمة غير واضحة تكون في الواقع إجابات لما يفهمه المستجوبون وهو بدون شك فهم يختلف من مستجوب لآخر نتيجة هذا الإبهام وعدم الوضوح .

ويجب الإشارة هنا إلى أن كثير من الكلمات قد تكون واضحة تماما في ذهن الباحث ولكنها تفهم بصورة مختلفة من قبل أولئك الأشخاص المستجوبين على الرغم من بساطة هذه الكلمات ، والأمثلة على ذلك كثيرة مثل الحالة الزوجية ، الجنسية ، المستوى التعليمي ، المهنة ، النشاط الاقتصادي ، إنتاج المؤسسة ، المشتغلون في المؤسسة وغيرها .

لذلك يجب التأكيد على الأهمية القصوى لتعريف جميع الألفاظ التي ترد في الأسئلة مهما كانت تبدو لنا واضحة وبسيطة ، وذلك لأن معنوية النتائج التي نصل إليها من الدراسات الإحصائية تتوقف أساسا على التعاريف الواضحة غير المبهمة لوحدات المجتمع الإحصائي موضوع أي من هذه الدراسات ولجميع المصطلحات والألفاظ التي تتضمنها الأسئلة التي تستخدم في جمع المعلومات الخاصة بهذه الدراسات .

4.1 المعلومـــات الإحصائيـــة (Statistical Information)

تسمى المعلومات التي تجمـــع خصيصـــا لدراســـــة إحصـــائية معينـــة بالمعلومات الأولية (Preliminary Information) ، وذلك مثل المعلومـــات التي تجمع في تعداد السكان أو في التعداد الصناعي أو في استقصاء ميزانيـــة الأسرة وغيرها. والميزة الأساسية لمثل هذه المعلومات هي إمكانية التحكم في صياغة الأسئلة التي تتضمنها اســـتمارة التعداد أو الاســـتقصاء وتوضيح كل ما تتضمنه هذه الأســـئلة من ألفاظ وتعابير بحيث يمكن تجنـــب ســـوء الفهـــم وتقليل الأخطاء التي تترتب على ذلك إلى حد كبير .

إلاّ أن هناك الكثير من المعلومات التي تجمع لأغـــراض أخـــرى غيـــر إحصائية ســـواء في مختلف الإدارات الحكومية أو في مختلـــف المؤسســـات الاقتصادية الخـــاصة والتي يمكن أن تظهر في ســـجلات هذه الإدارات وهذه المؤسسات ، وتكون بذلك مصدرا لكثير من الإحصاءات . هـــذه المعلومـــات تســـمى بالمعلومـــات الثانويـــة (Secondary Information) ، حيـــث أن استخدامها لأغراض إحصائية يكون مجرد اســـتخدام ثانوي ، فهـــي تـــنظم أســـاساً لأغراض إدارية أو قانونية . والأمثـــلة علـــى ذلـــك كثيـــرة منهـــا ســـجلات المواليد والوفيات والزواج والطلاق ، سجلات الجمارك ، ســـجلات الشرطة والقضاء ، السجلات المحاسبية للمؤسسات الاقـــتصادية المختلفـــة وغـــيرها .

وعـــند اســـتخدام هذه المعلومات كمصادر للإحصاءات الخاصة بها يجب أن يـــكون ذلك بعناية فائقة ، حيث أن هذه المعلومات قـــد لا تكـــون شـــاملة شمولاً كاملاً لجميع وحدات المجتمع الإحصائي موضوع الدراســـة ، كما قد لا تكون مسجلة بالأســـلوب الصحيح الذي يتفق مع اعتبارهـــا مصدرا إحصائيا جيدا . ولذلك قد يتعين علينا إجراء بعض التعديلات على المعلومات من هـــذا

النوع حتى تصبح صالحة لأن تكون مادة أولية جيدة للدراسة الإحصائية المطلوب القيام بها . مثلاً ، لمعرفة التغيرات الموسمية في مبيعات إحدى المؤسسات أو في النقد المتداول يكون من الواجب إعداد البيانات الإحصائية عن المبيعات أو النقود المتداولة في فترات شهرية أو ربع سنوية بدلا من إعدادها على أساس سنوي .

لذلك يستحسن تنظيم السجلات المستخدمة في مختلف الإدارات الحكومية وفي مختلف المؤسسات الاقتصادية بالتعاون بين القائمين بالعمل فيها وبين المسئولين عن إعداد الإحصاءات حتى يمكن الاتفاق مقدما على طرق الحصول على البيانات التي تتضمنها هذه السجلات ، وعلى الأسئلة التي توجه وعلى تعريف الألفاظ والتعابير التي ترد في هذه الأسئلة ، وعلى جميع الصعوبات التي يمكن أن تظهر عند العمل وكيفية معالجة تلك الصعوبات .

لذا يتعين علينا قبل استخدام أي إحصاءات في أي من الأبحاث العلمية ومهما كان مصدر هذه الإحصاءات أولياً أو ثانوياً ، إن نتأكد من كيفية الحصول عليها ، ومن درجة شمولها للمجتمع الإحصائي موضوع البحث ، ومن الفترة الزمنية التي تتعلق بها ، ومن التعاريف المختلفة التي استخدمت بالنسبة للألفاظ والتعابير الخاصة بموضوعها ، وإن أي تهاون في التأكد من هذه القواعد الأساسية التي يقوم عليها أي إحصاء مهما كان مصدره أولياً أو ثانوياً يمكن أن يؤدي إلى نتائج مضللة ، وبذلك يكون البحث عديم الجدوى ، بل مؤدياً إلى أخطاء فادحة في مسيرة الأعمال التي تعتمد على نتائجه .

5.1 خطوات البحث الإحصائي (Statistical Research Steps)

يقصد بالبحث الإحصائي دراسة أي موضوع ما أو مشكلة ما في أي من المجالات العلمية ، الاجتماعية أو الطبيعية ، باستخدام الطريقة الإحصائية منهاجا للبحث وأداة للتوصل إلى إجابات عن الأسئلة المختلفة

التي يطرحها الباحث في دراسته,وإلى التأكد من الفرضيات النظرية التي بدأ بها بحثه لقبولها إذا تبين من البحث معنوياتها أو استبعادها إذا تبين عدم معنوياتها . ولإجراء مثل هذا البحث تجرى دراسة استقصائية ميدانية لجمع المعلومات التي يرى الباحث ضرورتها . وعند إجراء مثل هذه الدراسة تتبع الخطوات التالية :

1- تحديد موضوع البحث تحديداً واضحاً من جميع النواحي بحيث يكون مفهوماً مقدماً المشكلة التي يرغب الباحث في دراستها أو الفرضية النظرية التي يرغب في اختبارها والحكم عليها . فإذا كنا بصدد دراسة الأجور في المؤسسات الصناعية ، مثلاً يكون من الواجب أن نحدد مقدماً هل نحن بصدد دراسة معدلات الأجور أو كسب العمل . وهل تشمل الدراسة جميع المشتغلين أو بعض الفئات منهم وما هي هذه الفئات . وهل تتضمن الدراسة المكافآت الإضافية التي حصل عليها المشتغلون أو تقتصر على الأجور الأساسية فقط . ومن الواضح أن تحديد موضوع البحث تحديداً شاملاً يرشد الباحث في تحديد المعلومات المختلفة التي يسأل عنها ويحاول الوصول عليها .

2 - إذا لم يكن من الممكن جمع المعلومات من جميع وحدات المجتمع الإحصائي الذي يشمله البحث بسبب ضخامة التكاليف ، واتساع الجهاز الإداري والفني الذي يحتاجه العد الشامل ، وقد يكون إعداد مثل هذا الجهاز أمراً متعذراً بالإضافة إلى طول الوقت الذي يمضي قبل الوصول إلى نتائج الدراسة ، وكثرة الأخطاء التي يمكن أن تترتب على العمل على نطاق واسع ، لذلك تجمع المعلومات من عينة من الوحدات على أن تتم معالجة نتائج العينة معالجة رياضية معينة ، بحيث يمكن التوصل إلى تقدير المقاييس والمعاملات الخاصة بالمجتمع الذي تمثله العينة موضوع الدراسة وحساب أخطاء هذه التقديرات ودرجة الثقة فيها .

في هذه الحالة لا بد أن يتقرر مقدماً نـوع العينة وحجمها وطريقة سـحبها من المجتمع المستهدف في الدراسة . ولا شك في أن مثل هذه القرارات تتوقف على مدى تجانس وحدات المجتمع ودرجة الدقة التي يقبل بها الباحث ونـوع المعلومات التي سوف يسأل عنها والمبالغ المخصصة للدراسة .

3 – إعداد استمارة البحث ، وهي عملية شاقة حيث تحتاج إلى خبـرة وفهـم دقيق لموضوع الدراسـة بالإضافة إلى الخبـرة الإحصـائية والممارسـة الطويلة لهذا الـنوع من الأعـمال ، إذ لا بد أن تـأتي الأسـئلة واضـحة ومحددة من حيث ألفاظها وصياغتها فلا تتعرض إلى أي نوع من الغمـوض والتأويل . كما لابد أن تأتي الأسئلة في صورة تـؤدي إلـى إجابـات قابلـة للمعالجة الإحصائية ، أي قابلة للتبويب ضمن أرقام في جداول يمكن أن يجري عليها التحليل الإحصائي الرياضي . وبالرغم من ضرورة الخبـرة بهذا النوع من الأعمال ، إلاّ أن تجربة الاسـتمارة ميدانياً أمـر حتمـي للتأكـد قـبل اسـتخدامها من أن فهم المستجوبين لأسئلتها يطابق تـماماً فهم الباحثين لها ، ولتعديل أي لـفظ أو تعبير يتبين من التجربة سوء الفهم له .

وبالرغم من أن الكثير من المعلومات الخاصة بالمؤسـسات الاقتصـادية تتوفر في سـجلات هذه المؤسـسات ، مثل المعلومات الخاصـة بالتكـاليف والمشتغلين والمبيعات ، إلاّ أن الحاجة تـظهر أحيانـاً لإجـراء دراسـات استقصائية ميدانية للتعرف على أمور معينة تتعلق بأعمالها ، وذلك مثل أبحاث السوق .

4 – جمع المعلومات ويكون ذلك بإحدى الطرق الآتية :

a) إرسال الاستمارات إلى المستجوبين بالبريد ومطالبتهم بإجابة أسئلتها وفـق التعليمات المرفقة ثم إعادتها إلى الدائرة القائمة بالدراسـة في ملف مرفق عليه العـنوان الخاص بهذه الإدارة.ولا شـك أن نجاح هذه الطريقة يتوقف إلى حد

كبير على الوعي الإحصائي للمستجوَبين ، حيث أنَّ هذا الوعي يجعلَهم يدركون أهمية المعلومات التي تطلب منهم ولذلك يَستجيبون بإعطاء الإجابة الدقيقة .

b) إرسال فريق من الموظفين الذين تم تدريبهم على مقابلة المستجوَبين وكيفية إقناعهم بإعطاء المعلومات الدقيقة وكيفية طرح الأسئلة وتدوين إجاباتها وأخيراً كيفية مراجعة الاستمارات بعد تعبئتها للتحقق من عدم وجود أي خطأ وأن جميع الأسئلة قد أجيبت . ومن الواضح أن العمل بهذه الطريقة يحتاج إلى تنظيم وأشراف دقيق للتأكد من قيام الموظفين بجمع المعلومات ولمساعدتهم في مواجهة المشاكل التي قد تظهر أثناء العمل ولتدقيق الاستمارات بعد تعبئتها حتى يمكن تصحيح أي خطأ يظهر فيها وهي لا تزال قيد التداول ميدانياً .

c) إرسال فريق من الموظفين إلى أماكن معينة لملاحظة ما يحدث وتدوين هذه الملاحظات على بطاقات معدة لذلك الغرض.وتتبع هذه الطريقة عند إجراء دراسات عن حركة المرور أو عن سلوك الأطفال أثناء لعبهم.

d) توجيه أسئلة محدودة بواسطة جهاز الهاتف,إلاّ أن هذه الطريقة تكون في الغالب متحيزة حيث أن الأسر التي لديها أجهزة هواتف لا يمكن اعتبارها عينة تمثل المجتمع موضوع البحث تمثيلاً صادقاً.ويمكن اتباع هذه الطريقة في بعض الدراسات الخاصة بآراء الناس بالنسبة لبرامج المذياع أو التلفاز.

e) إجراء التجارب أو القياسات المحددة ، وتتبع هذه الطريقة عند مراقبة الإنتاج كماً وكيفاً للتأكد من سير العمل وفق المواصفات والمعدلات المحددة بالنسبة للماكينات أو للعمال ، أو عند إجراء التجارب الزراعية وتسجيل نتائجها التي تكون بعد ذلك موضع التحليل الإحصائي .

5 – المراجعة الأخيرة للمعلومات المدونة في الاستمارات للتأكد من تماسكها، أي عدم تناقضها فيما بينها أو عدم تناقضها مع ما متوقع لها . ولغرض الكشف عن السهو أي عن الأسئلة التي ليس هناك أي إجابة لها . بعد ذلك تجرى بعض العمليات الحسابية لكي يستنتج من واقع المعلومات المعطاة في كل استمارة مقياساً مطلوباً ، مثل حساب القيمة الصافية المضافة لكل مؤسسة صناعية من واقع المعلومات التي أعطتها المؤسسة عن مبيعاتها خلال عام معين والمخزون المتوفر لديها في أوائل وأواخر العام المعني ، وكذلك مشترياتها من المواد الخام والوقود خلال العام والمخزون من هذه المواد في أول العام وفي آخره .

بعد ذلك تبدأ العمليات الخاصة بالتوبيب الآلي وهي ترميز المعلومات غير الرقمية ، ثم تثقيب البطاقات وبعد ذلك تصنيفها وفق الفئات المطلوبة ، ثم التبويب الآلي أي وضع البيانات بعد تصنيفها في جداول باستخدام الآلات الخاصة بذلك . ونشير هنا إلى أن جميع العمليات بعد الترميز تجري آلياً ويراجع إجرائها آلياً كذلك .

6 – تحليل البيانات الإحصائية أي المعلومات بعد تصنيفها وتبويبها ، وتفسير النتائج التي تترتب على هذا التحليل ، ومن ثم إعداد تقرير عن الدراسة معززاً بالجداول والرسوم البيانية والمقاييس والمعاملات والمؤشرات التي أمكن التوصل إليها ، مع ملاحظة أن يتضمن التقرير تعريف موجز بجميع الخطوات التي أتبعت وتوضيح أسباب وخلفيات أي قرار أتخذ بصدد هذه الخطوات ، مثلاً لماذا تقرر استخدام هذا النوع من العينات ولم يتبع أي من العينات الأخرى ، وعلى أي أساس حدد حجم العينة بعدد معين من المفردات وغيرها .

6.1 تصنيف البيانات وتبويبها

(Classification and Tabulation of Data)

1.6.1 مراجعة البيانات (Revision of Data)

قبل البدء في تصنيف البيانات ووضعها في جداول مناسبة يتم عادةً مراجعتها بهدف اكتشاف بعض الاستمارات التي قد تحتوي بيانات متناقضة أو نقص في الإجابات ، والتي يجب إعادتها إلى الميدان لاستيفائها أو إلغائها في حالة عدم التمكن من تصحيحها . فمثلاً إذا كان البيان المطلوب هو عن أجور بعض العمال الصناعيين فيجب أن يكون أساس الأجر موحداً حيث أن بعض المصانع تعطي أجراً يومياً والأخرى أسبوعياً أو شهرياً ، فيجب إرجاع جميع الأجور إلى أساس موحد أسبوعي مثلاً حتى تكون جميع الوحدات المستخدمة في البحث متجانسة ومن نوع واحد .

2.6.1 تصميم الجداول (Design of Tables)

بعد الانتهاء من مراجعة البيانات ، يتم إعادة تنظيمها بطريقة تسهل من دراستها وذلك للاستفادة منها على نحو أفضل . من ذلك يتم تصنيفها أي تقسيمها إلى مجموعات متجانسة وجدولتها أي وضعها بصورة جداول تلخيصية . وعادةً يتوقف هذا التقسيم على طبيعة البيانات وعلى الغرض الذي نسعى إليه من عمل البحث والدراسة . وبشكل عام لا توجد طريقة موحدة واحدة لعمل هذه الجداول إلاّ أن هناك قواعد عامة يجب مراعاتها وأخذها بعين الاعتبار عند تصميم الجداول ومنها :

a) أن يكون عنوان الجدول المقصود واضحاً ومختصراً ومحدداً لما يحتويه من معلومات .

b) أن تكون عنوين الصفوف والأعمدة مختصرة وواضحة وضوحاً تاماً .

c) أن ترتب البيانات في الجداول وفق تسلسلها الزمني أو حسب أهميتها من الناحية الوصفية .

d) يجب ترقيم الصفوف أو الأعمدة لتسهل الإشارة إلى بيانات الجداول .

e) أن يوضح المصدر الذي أخذت منه بيانات الجداول .

f) أن توضع وحدات القياس المستخدمة بدقة .

3.6.1 تبويب البيانات (Grouping of Data)

ويقصد بالتوبيب في علم الإحصاء وضع البيانات الإحصائية على شكل جداول (Tables) ، بحيث تمكننا من عرضها بصورة تلخص معالمها، وتساعد على استخلاص النتائج منها . ويتم عادةً تبويب البيانات على أساس تقويم زمني أو نوعي أو كمي أو جغرافي . أو على أساس خليط من هذه الأسس المختلفة التقسيم . فمثلاً الجدول (1-1) يبين قيمة المنتجات الصناعية في إحدى الدول العربية مبوبة على أساس أحد الأصناف السابقة وهو التبويب الزمني ، حيث يتم العرض الخاص بكل وحدة زمنية وفي هذا المثال هي السنة .

وهكذا يتم تبويب البيانات على أساس الأسس الأخرى المذكورة سابقاً. وقد يشمل التبويب على أكثر من واحد من هذه الأسس ، فقد يكون التبويب نوعياً وكمياً مثلاً في نفس الوقت وذلك إذا تم تصنيف المنتجات حسب حجم المصنع المنتج لها وذلك بالنسبة لكل نوع من أنواع الصناعة .

جـــدول (1-1)

قيمة المنتجات الصناعية في إحدى الدول

العربية خلال السنوات (1980 – 1985)

السنـــة	قيمـــة المنتجـــات (بالآلف الدنانير)
1980	482359
1981	519697
1982	538723
1983	622451
1984	734523
1985	754291

4.6.1 جـــداول التوزيعـــات التكراريـــــة

(Frequency Distribution Tables)

في كثير من الأحيان نلاحظ أنه عـــند الحصول على بيانات حول ظواهر أو تجارب معينة تكون هذه البيانات في وضع عشوائي أي غير خاضـــعة لأي نوع من الترتيب أو التصنيف ، وعادة ما تكون البيانات في مثل هذه الحالات كثيرة العدد ، وفي هذه الحالات يصعب فهم وتوضيح أي معالم أو صفات لهذه البيانات بسهولة . ولكن عند وضعها في صورة منظمة يمكـــن بهـــا عندئـــذ توضيح معالمها الأساسية بسهولة ، فمثلاً لو كان لدينا درجات 30 طالباً فـــي مادة ما مبينة على النحو التالي :

21 , 32 , 54 , 75 , 43 , 75 , 86 , 54 , 86 , 32 , 24 , 32 , 63 , 75
96 , 43 , 51 , 21 , 63 , 96 , 21 , 86 , 46 , 32 , 34 , 54 , 63 , 51
67, 21.

من الواضح وبالرغم من أن البيانات قليلة العدد إلاّ أن هناك عدة معلومات مبعثرة قد يصعب استنتاجها وتوضيحها من هذه الصورة للبيانات ببساطة ، ولكن وبمجهود بسيط نستطيع الحصول على معلومات عديدة توضح المعالم الأساسية الخاصة بهذه البيانات . فمثلاً إذا تم ترتيب هذه البيانات ترتيباً تصاعدياً أو ترتيباً تنازلياً مع توضيح تكرار كل قيمة ، لتمكنا من توضيح هذه البيانات على النحو التالي :

جـــدول (1 - 2)

الدرجة	21	24	32	43	46	51	54	63	67	75	86	96	المجموع
التكرار	4	1	4	2	1	2	3	3	1	3	3	2	30

بهذه الطريقة أصبح من الواضح معرفة معلومات عديدة كانت غير واضحة في الوضع السابق للبيانات ، فمثلاً من السهل الآن معرفة أكبر درجة وأصغر درجة كما يمكن معرفة عدد الناجحين ببساطة ، أي بهذه الطريقة استطعنا توضيح المعالم الأساسية لهذه البيانات وإعطاء فكرة عامة عن درجات الطلبة . ومن الواضح أنه لو كانت الدرجات كثيرة ومنتشرة داخل مجال متسع ووضعت على الشكل السابق والذي يبين درجات الطلبة مفردة مع تكرار كل درجة فسوف نجد أن الجدول (1-2) لن يفي بالغرض المطلوب.

من هذه الطريقة وهي تلخيص ووصف البيانات بطريقة أبسط توضح المعالم الرئيسية أي وضع البيانات في صورة مختصرة ومنظمة تعطي فكرة عامة عن هذه البيانات ، حيث نجد أن الجدول الذي نحصل عليه لن يختلف كثيراً عن وضع البيانات الأصلي لذلك فقد وجدت طريقة أخرى أكثر اختصاراً وأشمل فائدة من الوضع الأصلي للبيانات يمكن بواسطتها وضع البيانات في جدول يبين ويوضح الخصائص العامة لهذه البيانات .

وتتلخص هذه الطريقة في الواقع ، في أنه بدلاً من التعامل مع البيانات بالطريقة السابقة فإن المفردات تقسم إلى مجموعات أو فئات متجانسة بحيث تشمل كل مجموعة أو فئة عدداً من القيم المتقاربة من بعضها ، بحيث لا تنتمي كل مفردة من المفردات إلاّ إلى واحدة فقط من هذه الفئات (Sets or Classes) .

ويتوقف عدد هذه المجموعات أو الفترات (Intervals) على المدى (Range) بين أكبر وأصغر قيمة من قيم المفردات التي لدينا حيث يقسم المدى إلى عدد مناسب من الفترات أو الفئات لكي تضم كل فئة من الفئات مجموعة من القيم المتقاربة . إن تحديد أطوال هذه الفئات وعددها يتوقف عادة على طبيعة البحث ودرجة البحث ودرجة التلخيص المطلوبة .

وقد تكون الفترات أو الفئات متساوية الطول أو غير متساوية وذلك حسب طبيعة البيانات والغرض المعمول من أجله الجدول ، وبعد ذلك تعامل البيانات كمجموعات جديدة وليس كأعداد منفصلة كما كانت عليه قبل التبويب والتجميع ، وبالرغم من أنه لا توجد طريقة وحيدة لوضع جداول التوزيعات التكرارية إلاّ أنه هناك بعض النقاط الأساسية والخطوات الرئيسية التي يجب مراعاتها حتى تكون المعلومات التي نحصل عليها من الجداول أقرب إلى واقع البيانات وأكثر دقة ومن أهم هذه الخطوات والنقاط :

1. نحدد المدى (Range) الذي تنتشر فيه البيانات ويمثل القيمة الكبرى للبيانات أي القصوى (Maximum) مطروحاً منه القيمة الصغرى للبيانات أي (Minimum) أي أن المدى يساوي :

$$R = \text{Maximum (Max.)} - \text{Minimum (Min.)} \quad \ldots\ldots\ldots\ldots \quad (1\text{-}1)$$

2. نقسم المدى إلى فترات أو فئات متساوية الطـول بحيـث يكـون عـددها مناسباً. وقد أعتبر العدد من 5 إلى 25 عدداً مناسباً للفترات. وقد وضع العالم (استيرجس) علاقة يمكن الاستعانة بها لتحديد عدد الفترات وهي :

عدد الفترات = 1 + 3.3 Log n ..... (2 - 1)

حيث إن :

n – هو عدد البيانات . وليس من الضروري إتباع هذه العلاقة في تحديد عدد الفترات ولكن الأمر متروك للباحث وخبرته في وضع الجداول .

وهنا يجب الإشـارة إلى أن عدد الفترات يجب أن يكون مناسـباً ولـيس بالعدد الكبير حتى تكون معظم البيانات مشتتة بين فترات عديدة ، وبذلك لـن يكون هناك فرق كبير بين البيانات بوضعها الأصلي وبين الجـداول ولـن يوضح الجدول المعالم الأساسية التي وضع من أجلها ، كذلك يجب أن لا يكون عدد الفترات صغير جداً حتى لا نضطر إلى دمج معظم المعلومات معاً وفـي هذه الحالة لن نستطيع توضيح وإعطاء فكرة عامة عن البيانات .

3. تحديد الحدود العليا والدنيا الفعلية للفـئات حتى لا يـكون هناك فجوات أو تداخل ما بين هذه الفئات .

4. يكون اختيار طول الفئة من الأعداد التي يمكن التعامل معها بسهولة .

ويجب الإشارة إلى أنه عند تبويب جداول التوزيعـات التكراريـة فـإن البيانات الأصلية تفقد في الواقع ، ولا يمكن الرجوع من الجداول إلـى تلـك البيانات الأولية ولكن مقابل ذلك يتم الحصول من الجداول علـى المعلومـات

والاستنتاجات التي توضح المعالم الأساسية للبيانات بسهولة ، لذا يجب عند تكوين الجداول إتباع الخطوات الأساسية التي تم التعرض لها حتى تكون المعلومات التي سنحصل عليها من الجداول قريبة من الواقع الأولي لهذه البيانات .

وقد ذكرنا أيضاً أن طول الفترات في الجداول يجب أن يكون متساوياً وذلك لسهولة التعامل معها ، ولكن في بعض الحالات قد تستخدم فترات غير متساوية الطول لوجود غرض معين من وراء ذلك . فمثلاً إذا كان الغرض من الدراسة الاهتمام ببعض الفترات والتركيز عليها مع عدم الاهتمام لباقي الفترات الأخرى ، فإنه يتم دمج الفترات التي لا تهم الباحث في فترة واحدة ويكون الجدول في هذه الحالة غير متساوي الفترات . كذلك إذا كان التكرار (Frequency) لبعض الفترات صغيراً جداً مقارنة بباقي الفترات يمكن وضع هذه الفترات معاً .

كذلك هناك نوع آخر من جداول التوزيعات التكرارية تجدر بنا الإشارة إليه وهي الجداول ذات الفترات المفتوحة (Open Intervals) ، حيث يصادفنا أن يكون الحد الأدنى للفترة الأولى أو الحد الأعلى للفترة الأخيرة غير محدد ، مثل أن نقول الفترة الأولى أقل من 100 أو الفترة الأخيرة من 400 فأكثر . وهذا النوع من الجداول عادةً ما يكون قليل الأهمية حيث إن مثل هذا النوع لا يمكننا منه حساب مقاييس إحصائية هامة يجب إيجادها كما سيتم شرحه في الأبواب القادمة من هذا الكتاب . وبالرغم من ذلك نجد أن لهذا النوع من الجداول أي الجداول المفتوحة بعض الاستخدامات في بعض الحالات .

ولتوضيح طريقة تكوين جداول التوزيعات التكرارية سنقوم بدراسة المثال التالي والذي سيوضح الخطوات الأساسية التي يجب مراعاتها عند تبويب البيانات في جداول .

مثـــال (1-1)

أرادت شركة صناعية كبرى متخصصة بصـناعة الحاسـوب ورقـائق الكومبيوتر المختلفة أن تدرس كفاءة المهندسين العاملين لديها للتعـرف علـى مدى ملائمتهم لإعمالهم الحالية واختيارهم لتمكن تمناسـب صـمم مجـال عملهم . لأجل ذلك تـم اختيار عينة عشـوائية (Random Sample) مـن (50) مهندس من بين المهندسـين العاملين بها والـبالغ عـددهم (1500) مهندس حيث أجري لهم اختبارين ، الاختبار الأول يقيس درجة ذكاء المهندس وقدرته على التصرف في مواقف معينة ، والاختبار الثـاني يقـيس درجـة المهارة اليدوية في سرعة الحركة . فكانت البيانات للدرجات التي حصل عليها المهندسون الخمسون كما مبين في الجدول (1- 3) .

جـــدول (1 - 3)

العينـــة العشوائيـــة للخمسين مهندس

الرقم	الرقم العشوائي للمهندس في العينة	الدرجة في أختبار الذكاء	الدرجة في أختبار المهارة اليدوية
1	0624	110	74
2	1434	123	68
3	0753	109	59
4	1101	104	46
5	1481	111	82
6	0962	131	46
7	0416	132	70
8	0045	127	50
9	0629	91	64
10	0483	126	49
11	1121	135	87
12	0085	116	59
13	0817	111	61
14	1383	113	59
15	1129	119	89

56	114	0612	16
48	118	1290	17
69	101	0849	18
56	118	1473	19
73	119	1082	20
56	102	0895	21
63	106	0302	22
48	124	0797	23
44	101	1311	24
57	97	0942	25
47	95	1499	26
68	121	0097	27
58	105	0009	28
70	107	0557	29
66	115	0380	30
55	110	0262	31
72	115	0589	32
54	101	0018	33
56	118	1327	34
70	121	0318	35
53	122	1337	36
75	128	0717	37
53	144	1299	38
58	119	0022	39
52	121	0609	40
77	101	0939	41
72	107	0778	42
61	141	1288	43
46	120	0089	44
69	102	0203	45
68	121	0967	46
63	133	1355	47
69	107	1023	48
60	103	0067	49
61	119	0581	50

لتوضيح كيفية تلخيص البيانات الخاصة بدرجات الذكاء للمهندسين مثلاً في جدول توزيع تكراري ، نتبع الخطوات والمراحل التالية التي توضح لنا كيفية عمل ذلك :

1. نقوم باختيار الفئات وكما شرحنا سابقا مع أنه لا توجد قاعدة ثابتة لتحديد عدد وأطوال الفئات ، حيث إن ذلك يعتمد على طبيعة البيانات المستخدمة وعلى مستوى الدقة المطلوبة . عند النظر إلى بيانات ومفردات الجدول (1-3) ، نجد المدى لهذه البيانات حيث إن أكبر درجة في اختبار الذكاء هي 144 وأقل درجة هي 91 إذاً المدى يساوي :

$$R = 144 - 91 = 53$$

فلو قسّمنا هذا المدى إلى 11 فئة نجد أن طول الفئة سيكون 5 درجات ، وإذا قسّمناه إلى 6 فئات يكون طول الفئة 10 ... وهكذا .

2. بعد اختيارنا لعدد الفئات وأطوالها يجب علينا توضيح حدود هذه الفئات بحيث لا تتداخل مع بعضها البعض ، فمثلاً لو قسّمنا الدرجات إلى 11 فئة طول كل منها 5 درجات فقد يتسرع أحدنا ويخطئ بكتابة الفئات على الصورة التالية :

$$\begin{array}{c} 90 - 95 \\ 95 - 100 \\ 100 - 105 \\ \vdots \\ 140 - 145 \end{array}$$

والخطأ هنا واضح في كتابة هذه الفئات حيث لا يمكن معرفة فيما إذا كان المهندس الذي حصل على 95 درجة ينتمي إلى الفئة الأولى أو الفئه الثانية ، والمهندس الذي حصل على 105 درجة ينتمي إلى الفئة الثالثة أو الرابعة وهكذا . وللتغلب على ذلك يمكن كتابة الفئات كما يلي :

$$90 - 94$$
$$95 - 99$$
$$100 - 104$$
$$\vdots$$
$$140 - 144$$

إن توزيع حدود الفئات بهذه الصورة يكون صحيحاً فقط إذا كان المتغير الذي ندرسه متغيراً غير مستمر(Discrete Variable) ، أي لا يأخذ قيماً كسرية أما إذا كان المتغير الذي ندرسه مستمراً (Continuous Variable) فإنه يكون من الخطأ كتابة حدود الفئات بهذا الشكل حيث يجب ألاّ نترك أي ثغرات بين الفئات .

فالمهندس الذي حصل على 99.5 درجة لا نعرف ما إذا كان ينتمي إلى الفئة الثانية أو الثالثة ، لذلك في هذه الحالة ، يجب توضيح جميع حدود الفئات وكتابتها بصورة واضحة وصالحة لدراسة جميع المتغيرات المستمرة وذلك بأن نجعل كل فئة تبدأ مباشرة حيث تنتهي الفئة السابقة لها دون أن يحدث تداخل بين الفئات ودون أن نترك أية ثغرات بينها كما يلي :

95 – 90
100 – 95
105 – 100

140 وأقل من 145

وللاختصار يمكن أن نحدد فقط بداية الفئة ونترك نهايتها لتحدد ضمناً من الفئة التـــالية لها , وفي هذه الحالة ينبغي أن نحدد نـــهاية الفئة الأخيرة على الشكل التالي :

– 90
– 95
– 100

140 وأقل من 145

ومعنى هذا أن الفئة الأولى تشمل كل القيم التي تساوي 90 أو تزيد عنها بحيث تقل عن 95 وعليه فإن القيمة 95 تدخل ضمن الفئة الثانية والقيمة 140 ضمن الفئة الأخيرة وهكذا .

3. بعد اختيار الفئات وتحديد حدودها نكوّن جدولاً (Table) ، يحتوي علـــى عدد من الصفوف مســـــاوي لعـــدد الفئـــات ثـــم نقـــوم بتفريـــغ البيانـــات الأولية ، مفردة بعد الأخرى كل في الفئة التي تنتمي إليها وذلك بوضع علامة (/) أمام هذه الفئة ونستمر في عملية التبويب هذه حتى ننتهي من تفريغ جميع

مفردات البيانات الأولية وذلك بوضع عدد من العلامات مساوٍ لعدد المفـــردات الأولية التي لدينا ويلاحظ عادة أثناء التفريغ أننا نقوم بوضع كل أربعة علامات بجوار بعضها على الشكل التالي (////) أما العلامة الخامسة فتشطب الأربعة السابقة وتكون حزمة على الشكل التالي (////) والغرض من ذلك هو تسهيل عد هذه العلامة التي تم رصدها أمام كل فئة (Class) ، والذي سوف نطلـــق عليه أسم التكرار أو ما يعرف (Frequency) .

أن الجدول (1-4) ، يبين التبويب الخاص بـــدرجات الـــذكاء للمهندســـين الخمسين في العينة العشوائية للمثال السابق . ويلاحظ فـــي هـــذا الجـــدول أن الفئات تكتب في العمود الأول ثم أخذت البيانات الأوليـــة الخاصـــة بـــدرجات الذكاء مفردة بعد الأخرى ، وقد وضعت علامة لكل مفردة أمام الفئـــة التـــي تنتمي إليها هذه المفردة . وبعد الانتهاء من عملية التفريغ نسجل عدد العلامات في العمود الأخير من الجدول والذي يمثل تكرار الفئة (f_i) .

إن مجموع التكرارات يجب أن يساوي عدد المفردات التي تم تفريغها . إن الجداول على شكل الجدول (1-4) ، والتي تتكون مـــن الفئـــات والتكـــرارات تسمى بجداول التوزيعات التكرارية (Frequency Distribution Tables) .

وكما أشرنا سابقاً أنه عند تكوين الجداول التكرارية تضـــيع معـــالم القـــيم الأولية الأصلية ، ولا نعرف شيئاً عن أي مفردة من المفردات الأصلية حيـــث أنها سوف تنتهي إلى فئة معينة محددة بحدين معلومين مثل 100 وأقـــل مـــن 105 . ولكن وبفضل هذه الجداول نستطيع معرفة الدرجة التي حصل عليهـــا أي من المهندســـين الثـــمانية المنتمين إلى هـــذه الفئة ، وما إذا كانـــت هـــذه الدرجة تقع بالقرب من بداية الفئة 100 أو بالقرب من نهايتهـــا 105, لـــذلك نفترض أن جميع المهندسين في كل فئة حصلوا على درجات متساوية وأن كل منها يساوي مركز هذه الفئـــة أي منتصـــف المـــدى بـــين الحـــدين الأدنـــى والأعلى للفئة .

فمثلاً مركز الفئة 100 وأقل من 105 هو:

$$a.m.=\frac{105+100}{2}=102.5$$

حيث أن :

a.m. – هو مركز الفئة (Arithmetic Mean)

جـــدول (1 - 4)

جدول التوزيع التكراري لدرجات الذكاء للمهندسين الخمسين

الفئـــات (Sets)	العلامـــات	التكـــرار (f_i)
- 90	/	1
- 95	//	2
- 100	/// ~~////~~	8
- 105	/ ~~////~~	6
- 110	/ ~~////~~	6
- 115	~~////~~ ~~////~~	10
- 120	// ~~////~~	7
- 125	////	4
- 130	///	3
- 135	/	1
140 إلى أقل من 145	//	2
المجموع (Total)		50

وقد يكون من المناسب عرض البيانات في بعض الأحيان على شكل توزيع تكراري نسبي (Relative Frequency Distribution R.F.D.) أي إظهار تكرار كل فئة كنسبة من المجموع الكلي للتكرارات . والجـــدول(1-5) يبـــين

تبويب البيانات الخـــاصة بدرجات الـــذكاء للخمســـين مهندس المختارين في العينة العشوائية السابقة في شكل جدول توزيع تكراري نسبي مكوناً من ســـت فئات طول كل فئة منها عشرة درجات .

جـــــدول (1 - 5)

الفئات (Sets)	التكرار (f_i)	التكرار النسبي (p)	التكرار المئوي (P%)
- 90	3	0.06	% 6
- 100	14	0.28	% 28
- 110	16	0.32	% 32
- 120	11	0.22	% 22
- 130	4	0.08	% 8
140 وأقل من 145	2	0.04	% 4
المجموع (Total)	50	1.00	% 100

نلاحظ من الجدول (1-4) والجدول (1-5) أنه تـــم اســـتخدام فئـــات ذات أطوال متساوية ، يسمى التوزيع التكراري من هذا النوع بالتوزيع التكـــراري المنتظم أو ما يدعى (Uniform Frequency Distribution) .

وتجدر الإشارة هنا إلى أنه من الأفضل دائماً استخدام الفئـــات متســـاوية الطول ، وذلك تسهيلاً للعمليات الحسابية غير أنه في بعض الأحيان قد نضطر لاستخدام فئات غير متساوية الطول وهنا يسمى التوزيع التكراري بـــالتوزيع التكراري غير المنتظم (Non-Uniform Frequency Distribution) .

كما ويلاحظ من الجدول (1-5) أن بداية الفئة الأولى ونهاية الفئة الأخيرة محددتين لذلك يسمى هذا النوع من بالجداول المقفلة (Closed Tables) أمـــا

الجداول التي تكون فيها نهاية الفئة الأخيرة غير محـددة فتسـمى بالجـداول المفتوحة Open Tables.

1.4.6.1 جـداول التوزيعـات التكراريـة المزدوجـة
(Double Frequency Distribution Tables)

لاحظنا في الجدول(1-3) كيفية تكوين الجداول التكرارية البسيطة وذلـك بتبويب مجموعة من المفردات التي تخص ظاهـرة واحـدة ، أمـا إذا كـان لدينا مجموعة من أزواج القـيم لمتغيرين تربطهمـا علاقـة معينـة مثـل مجموعة من الأزواج وأعـمار زوجاتهم أو أوزان أشـخاص ما وأطـوالهم وغـيرها ، فإننا نستطيع تبويب مثل هذه البيانات في جدول تكراري مـزدوج أو ما يصطلح على تسـميته (Double Frequency Table) ، وهو جدول ينقسم أفقياً ورأسياً إلى عـدد من الصفـوف والأعمدة بحيث يبين التقسـيم الرأسـي فئات إحدى الظاهرتين ويبين التقسيم الأفقي فئات الظاهرة الأخرى .

فمثلاً إذا أردنا تبويب البيانات الخاصة بكل من درجات اختبـار الـذكاء ودرجات اختبار المهارة للمهندسين الخمسين في العينة العشوائية السابقة فـي شكل جدول تكراري مزدوج ، فإننا نقسم مدى كل من المتغيرين إلـى عـدد مناسـب من الفئات ثم نقوم بتفريغ البيانات ، بأن نضع لكل قيمتين متناظرتين علامة (/) في الخلية التي تقابل فئتيهما . فلو قسمنا المدى لدرجات الذكاء إلى ست فئات طول كل منها 10 درجات ، وقسـمنا المدى لدرجات المهارة إلـى خمسـة فئات طول كل منها 10 درجات أيضاً ، بأن نضع لكـل زوج مـن أزواج القيم التي لدينا والتي تخص أحد المهندسين علامة في الخلية التي ينتمي إليها هذا المهندس .

فالمهنـدس الأول حصل على 91 درجة في اختـبار الذكاء و56 درجـة في اختبار المهارة فهو إذن ينتمي إلى الخلية التي تـقع فـي الصـف الأول والعمود الثاني حيث أن درجـة ذكـاء هـذا المهنـدس تقـع فـي الفئـة

(90 – أقل من 100) ودرجة مهارته تقع في الفئة(50 – أقل من 60) ونستمر في عملية التفريغ هذه مكونين عدداً من الحزم والعلامات كما تم ذكره سابقاً عند توضيح تكوين الجداول التكرارية البسيطة .

جـــدول (1 - 6)

تفريغ درجات اختبار الذكاء والمهارة للخمسين مهندس

درجات اختبار المهارة / درجات اختبار الذكاء	40 -	50 -	60 -	70 -	80 وأكثر من 90
90 -	//	/			
100 -	/ ###	/// ###			
110 -		////	### ###		
120 -		///	////	////	
130 -				///	/
140 وأكثر من 150					//

وبعد الانتهاء من التفريغ يتم استبدال العلامات بعددها فنحصل على الجدول التكراري المزدوج الذي يبين توزيع المهندسين حسب الدرجات التي حصلوا عليها في اختباري الذكاء والمهارة كما يبين الجدول (1-5) ، حيث يلاحظ في العمود الأول والأخير يعطيان التوزيع التكراري لدرجات الذكاء بينما يعطي الصف الأول والصف الأخير التوزيع التكراري لدرجات المهارة . لذلك يجب أن يكون مجموع كل من العمود الأخير والصف الأخير مساوياً لعدد المهندسين في العينة العشوائية المدروسة .

جـــدول (1- 7)

جدول التوزيع التكراري لدرجات اختباري الذكاء والمهارة

فئات درجات اختبار الذكاء / فئات اختبار درجات المهارة	40-	- 50	- 60	- 70	80 وأقل من 90	المجموع
- 90	2	1				3
- 100	6	8				14
- 110		4	10	2		16
- 120		3	4	4		11
- 130				3	1	4
140 وأقل من 150					2	2
المجموع	8	16	14	9	3	50

وفي كثير من الأحيان قد نحتاج لمعرفة عدد من المفردات التي تقل قيمتها عن حد معين أو تلك التي تساوي قيمتها أو تزيد عن حد معين ففي الجـــدول (1-3) ، قد نحتاج لمعرفة عدد المهندسين الذين حصلوا على درجـــات فـــي اختبار الذكاء تقل عن 110 درجة ، وهؤلاء عددهم 17 مهندس لأنهم يمثلون مجموع تكرار الفئتين الأولى والثانيـــة وهمـــا(90 إلـــى أقـــل مـــن 100) و (100 إلى أقل من110) وهو 3 + 14 = 17 مهنـــدس ، بينمـــا إذا أردنـــــا معرفة عدد المهندسين الذين حصلوا على درجات تساوي أو تزيد عـــن 130 درجة فنجدهم 6 مهندسين لأنهم يمثـــلون مجموع تكراري الفئتين الأخيـــرتين وهما(130 وأقل من 140) والفئة(140 وأقـــل مـــن 150) وهـــو4 + 2 = 6 مهندسين .

ولمعرفة هذا النوع من البيانات نكون ما يسمى بجدول التكرار المتجمع (Cumulative Frequency Table) ، جدول التكرار المتجمع النسبي (Relative Cumulative Frequency Table) ، وهناك نوعين من الجداول المتجمعة المستخدمة بكثرة هي جدول التكرار المتجمع الصاعد (Ascending Cumulative Frequency) ، وجدول التكرار المتجمع النازل (Descending Cumulative Frequency) .

إن الجدول (1-8) ، يمكننا من معرفة عدد المهندسين الذين حصلوا على درجات تقل عن الحدود العليا للفئات ونحصل عليه بتجميع بيانات التكرارات من الجدول(1-3) من جهة الفئات الصغيرة إلى الكبيرة .

جـــدول(1 - 8)

التكرار المتجمع الصاعد لدرجات الذكاء للخمسين مهندس

الحدود العليا للفئات	التكرار المتجمع الصاعد	التكرار المتجمع الصاعد النسبي
أقل من 90	0	0
أقل من 100	3	0.06
أقل من 110	17	0.34
أقل من 120	33	0.66
أقل من 130	44	0.88
أقل من 140	48	0.96
أقل من 150	50	1.00

أما الجدول (1-9) فيبين لنا عدد المهندسين الذين حصلوا على درجات تساوي أو تزيد عن الحدود الدنيا للفئات ونحصل عليه بتجميع بيانات

التكرارات من الجدول (1-3) ابتداً من الفئات الكبيرة إلى الفئات الصغيرة أي يرمز للفئة المعنية بحدها الأعلى فأكثر وهكذا .

جـدول (1 - 9)

التكرار المتجمع النازل لدرجات اختبار الذكاء للخمسين مهندس

الحدود الدنيا للفئات	التكرار المتجمع النازل	التكرار المتجمع النازل النسبي
90 فأكثر	50	1.00
100 فأكثر	47	0.94
110 فأكثر	33	0.66
120 فأكثر	17	0.34
130 فأكثر	6	0.11
140 فأكثر	3	0.04
150 فأكثر	0	0

7.1 التمثيـل والعـرض البيـاني للبيانـات
(Graphical Representation of Data)

يفيد العرض البياني في إظهار البيانات العددية وتتبـع المتغيـرات فيهـا بطريقة تجلب الانتباه وتتسم بالبساطة والسهولة في تذكرها ، كما تفيد أيضا في توضيح العلاقات بين المتغيرات التي ندرسها .

وتختلف الرسائل التي نستخدمها تبعاً لنوع البيان الإحصـائي والحقـائق المطلوب إبرازها، ونذكر فيما يلي بعض الطرق التي تستخدم فـي العـرض والتمثيل البياني :

1- الخـــط البيـــاتـــي (Graphical Line)

يســـتخدم الخط البياني لتمثيل العلاقة بين ظاهرتين أو متغيرين ، بحيـــث يبين كيفية تغير إحدى الظاهرتين مع الظاهرة الأخرى أو تبعاً لها . فإذا كـــان الزمن هو أحد المتغيرين ، يكون الغرض من الرسم هو معرفة مدى التغيـــر الذي يحدث في الظاهرة التي ندرسها خلال فترة زمنية محددة . وفـــي هـــذه الحالة يسمى الخط البياني الذي نحصل عليه بـــالمنحنى التـــاريخي للظـــاهرة (Historical Curve) .

2- الأعمـــدة البيـــانيـــة (Bar Charts)

وهي أعمدة أو مستطيلات رأسية قواعدها متساوية لأنها تعتمد على طول الفئة أو الفترة التي غالباً ما تكون متســـاوية ، أما ارتفاعاتها فتتناســـب مـــع قيمة الظاهرة التي ندرســـها أي التكرار (f_i) فإذا كان المحور الأفقي يمثـــل الزمن فإن التغير فـــي ارتفـــاع المســـتطيلات يمثـــل التطـــوّر التـــاريخي للظاهرة ، وقد تستخدم الأعمدة البيانية للمقارنة بين أكثر من ظـــاهرة وذلـــك برســـم أعمدة متلاصقة للظواهر المراد مقارنتها في الســـنوات المختلفة على أن يســـتخدم لكل ظاهرة لون مختلف أو ظل مختلف .

وفي بعض الأحيان قد يوضع فرق بين الأعمدة لتسهيل المقارنة وذلك إذا كانت البيانات التي لـــدينا هـــي بيانـــات إحصـــائية إجماليـــة مقســـمة إلـــى مكوناتها ، مثل بيانات عدد السكان في السنوات المختلفة مقسمة إلـــى ذكـــور وإناث وهنا نرسم عموداً يمثل عدد السكان لكل ســـنة والجزء الأسفل منه يمثل الذكور والجزء الأعلى يمثل الإناث ، وإذا اســـتخدمنا الأعمدة البيانية لدراسة ظاهرة معينة أو عدد من الظواهر بحيث لا تشتمل على عنصر زمنـــي كـــأن نقسم الطلبة حسب المراحل الدراسية المختلفة أو العمال حســـب الصـــناعات

المختلفة ، فيستحسن في هذه الحالة ترتيب الأعمدة حسب قيمها ترتيباً تصاعدياً أو ترتيباً تنازلياً لأن ذلك يبسط الشكل ويسهل من دراسته . ولاحظ أيضاً أنه إذا كان أحد الأعمدة أطول بكثير من الأعمدة الأخرى بحيث لا تتسع مساحة الرسم فإنه يمكن كسر العمود قبل نهايته مع توضيح القيمة التي تمثله في أعلى العمود ويمكن استخدام قضبان أفقية بدلاً من الأعمدة الرأسية وفي هذه الحالة يستبدل المحورين الأفقي والرأسي في جميع الحالات التي أشرنا إليها .

3- البيانات أو القطاعات الدائرية (Pie Charts)

تستخدم الدوائر عادةً للمقارنة بين المكونات المختلفة لظاهرة معينة ببعضها البعض وبين كل منها والمجموع . وبذلك تظهر الأهمية النسبية لهذه المكونات ويتم ذلك بتقسيم الدائرة إلى عدد من القطاعات تتلاقى في المركز بحيث تتناسب مساحة هذه القطاعات مع القيم المختلفة للمكونات الجزئية للظاهرة . وحيث أن الزاوية المركزية الرئيسية للدائرة هي $360°$ ، فإن القطاع الذي تكون زاويته $3.6°$ سوف يمثل 1% من مساحة الدائرة وبذلك يمكن تحديد أي قطاع بواسطة زاويته المركزية ، فمثلاً القطاع الذي تكون مساحته 30% من مساحة الدائرة تكون زاويته المركزية هي :

$$360° \times \frac{30}{100} = 72°$$

والقطاع الذي تمثل مساحته 40% من مساحة الدائرة ستكون زاويته المركزية $169.2°$ وهكذا .

ولتوضيح كيفية التمثيل بهذه الطريقة نقوم بدراسة المثال التالي .

مثـــال (1–1)

الجدول (1-10) يبين مســـاحات القارات المختلفة في العالم مقدرة بملايين الكيلومترات المربعة والمطلوب تمثيل هذه البيانات بالآتي :

a) القطاعات الدائرية .

b) الأعمدة البيانية .

جـــدول (1-10)

أفريقيا	أستراليا	آسيا	أوروبا	أمريكا الشمالية	أمريكا الجنوبية
30	8	44	11	24	18

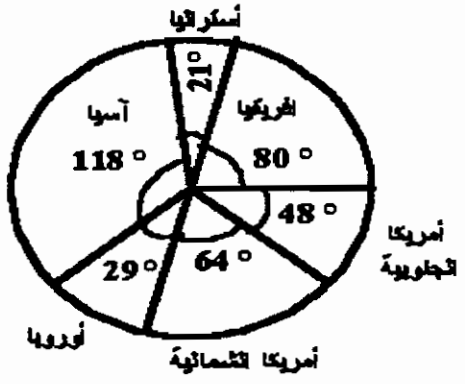
الحـــل :

نقوم أولاً بإيجاد الزوايا المقابلة والتي تمثل مساحة كل قطاع والموضـــحة فـــي الجـــدول (1-11) ، بمـــا أن الزاويـــة المركزيـــة الرئيســـية للـــدائرة هي 360° إذن :

جـــدول(1-11)

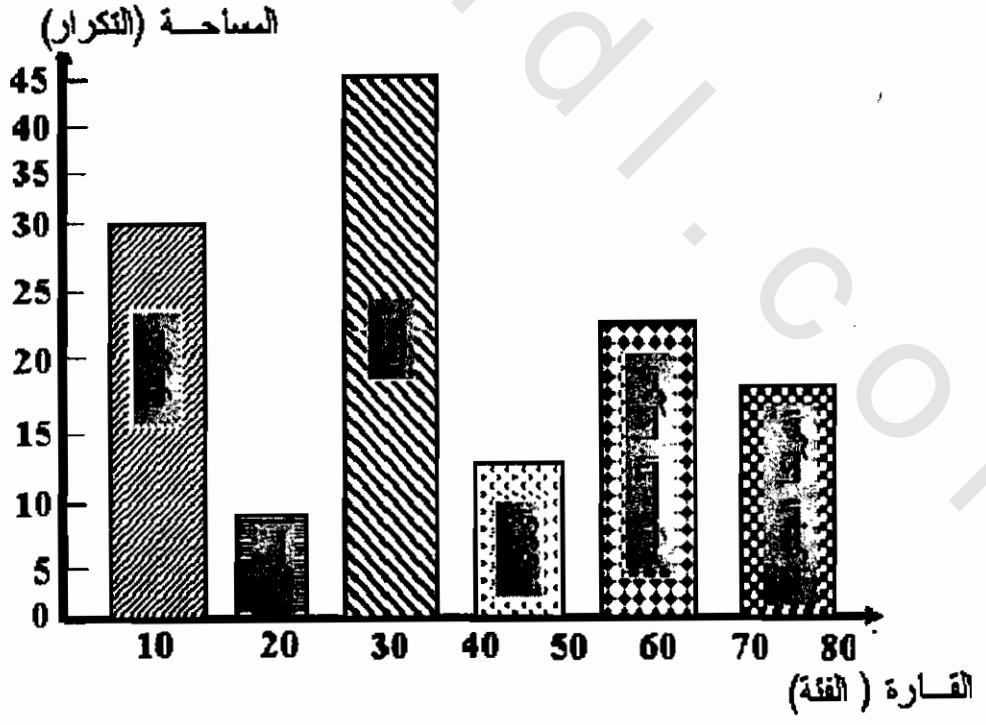
ت	الفئة (القارة)	المساحة (التكرار)	زاوية القطاع
1	أفريقيا	30	$80° = 360° \; x \frac{30}{135}$
2	أستراليا	8	$21° = 360° \; x \frac{8}{135}$
3	آسيا	44	$118° = 360° \; x \frac{44}{135}$
4	أوروبا	11	$29° = 360° \; x \frac{11}{135}$
5	أمريكا الشمالية	24	$64° = 360° \; x \frac{24}{135}$
6	أمريكا الجنوبية	18	$48° = 360° \; x \frac{18}{135}$
المجموع		135	360 °

بعد ذلك نقوم نرسم دائرة بنصف قطر مناسب ونقسم زاويتها المركزية حسب نسبة القطاعات المختلفة والزوايا المقابلة لها فنحصل على الشكل (1-1) :



الشكل (1-1)

أما بالنسبة لتمثيل البيانات بطريقة الأعمدة البيانية فتكون كما يوضح الشكل (1-2) ، حيث تمثل القارات على المحور الأفقي بفترات متساوية ويمثل التكرار بمستطيل .



الشكل (1-2)

8.1 التمثيل البياني للتوزيعات التكرارية
(Graphical Representation of Frequency Distributions)

يمكن تمثيل التوزيعات التكرارية بيانيا بعدة طرق مختلفة أهمها المدرج التكراري(Histogram) ، والمضلع التكراري(Frequency Polygon) ، وكذلك المنحني التكراري (Frequency Curve) ، وسوف نقوم بعرض طرق التمثيل البياني لهذة التوزيعات بالتفصيل :

1- المدرج التكراري (The Histogram)

في هذا التمثيل يتم إنشاء محورين وهما المحور الأفقي والمحور الرأسي ، وتمثل الفئات أو الفترات على المحور الأفقي والتكرارات على المحور الرأسي ، ويراعى عند التمثيل البياني اختيار مقياس رسم مناسب لكل محور ، بحيث يكفي المحور الأفقي لتمثيل جميع الفئات ويسمح المحور الرأسي بتمثيل جميع التكرارات ثم نكوّن المدرج التكراري وذلك برسم عدد من المستطيلات المتلاصقة بحيث يكون مساحتها متناسبة مع التكرارات .

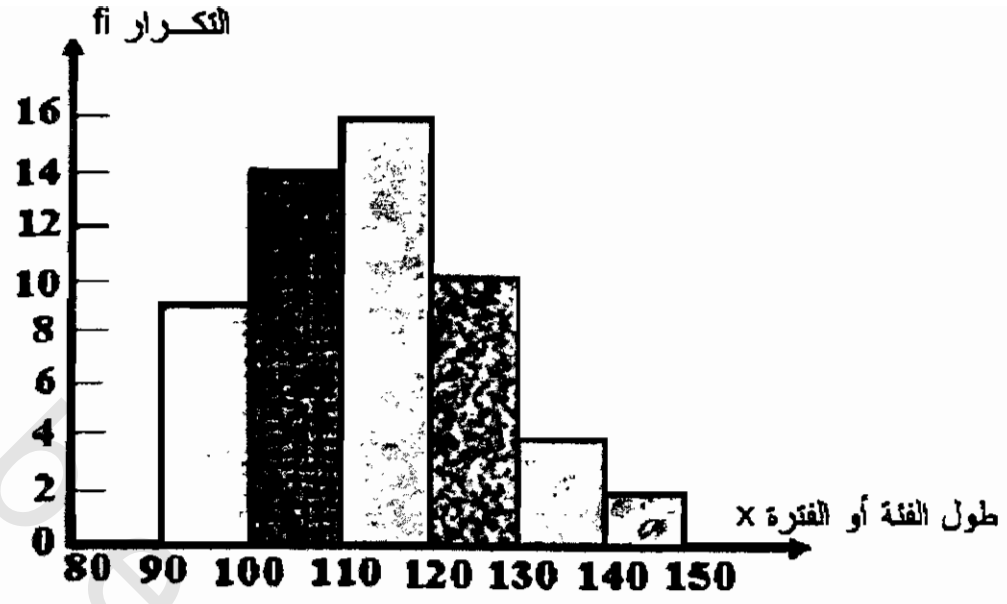
إذا كانت أطوال الفئات متساوية فإننا ننشى عدداً من المستطيلات قاعدة كل مستطيل تمثل طول الفئة ، وارتفاعه هو تكرار تلك الفئة وهنا سوف نجد بأن مساحة المستطيلات تتناسب مع التكرارات لأن أطوال قواعد المستطيلات جميعها متساوية ولأن مساحة كل مستطيل تساوي طول القاعدة في الارتفاع .

ويلاحظ هنا أن المحور الرأسي لا بد وأن يبدأ من الصفر حتى يمكننا مقارنة التكرارات المختلفة ببعضها البعض ، أما المحور الأفقي فليس من الضروري أن يبدأ من الصفر حيث لو فعلنا ذلك لاحتجنا إلى مسافة كبيرة جداً لتمثيل الفئات بينما يكون المحور الأفقي من الصفر حتى بداية الفئة الأولى غير مستعمل .

ولو عدنا إلى الجدول (1-4) والذي يمثل التوزيع التكراري لدرجات الذكاء للخمسين مهندس لوجدنا بأنه من الممكن تمثيل تلك البيانات على شكل مدرج تكراري ، حيث يتبين أن أطوال الفئات هي متساوية لأن طول كل فئة يساوي 10 درجات ، لذا فأننا نقسم المحور الأفقي إلى ستة أقسام متساوية تمثل فئات الدرجات .

أما المحور الرأسي فيتم تقسيمه إلى أقسام متساوية بحيث يسمح بتمثيل أكبر التكرارات وهو 16 في مثالنا هذا ثم نرسم ستة مستطيلات طول قاعدة كل منها 10 درجات وارتفاع كل منها يساوي تكرار الفئة المقام فوقها هذا المستطيل ، وبذلك نحصل على المدرج التكراري الموضح في الشكل(1-3) والذي يبين أن مساحة المستطيلات تتناسب مع التكرارات .

وفي حالة الجداول غير المنتظمة أي ذات الفئات غير المتساوية الأطوال فإنه يلزم تعديل التكرارت قبل رسم المدرج التكراري ، وذلك بقسمة تكرار كل فئة على طول هذه الفئة وبذلك نحصل على التكرار المعدّل ثم نكوّن المدرج التكراري بإنشاء عدد من المستطيلات قواعدها هي أطوال الفئات غير المتساوية ارتفاعاتها هي التكرارات المعدّلة وفي هذه نجد أن مساحة كل مستطيل تساوي تماماً تكرار الفئة التي يمثلها .



الشكـــل (1 - 3)

تمثيل المدرج التكراري لدرجات الذكاء للخمسين مهندس

وكمثال على ذلك دعنا نرســـم المدرج التكراري لتوزيع المصانـــع المبين في الجدول(1 -12) .

جـــدول (1 - 12)

توزيع 200 مصنع حسب عدد العمال الذين يستخدمهم كل مصنع

فئات العمال x	- 1	- 3	- 5	- 10	- 20	50 وأقل من100
عدد المصانع f	8	22	80	48	27	15

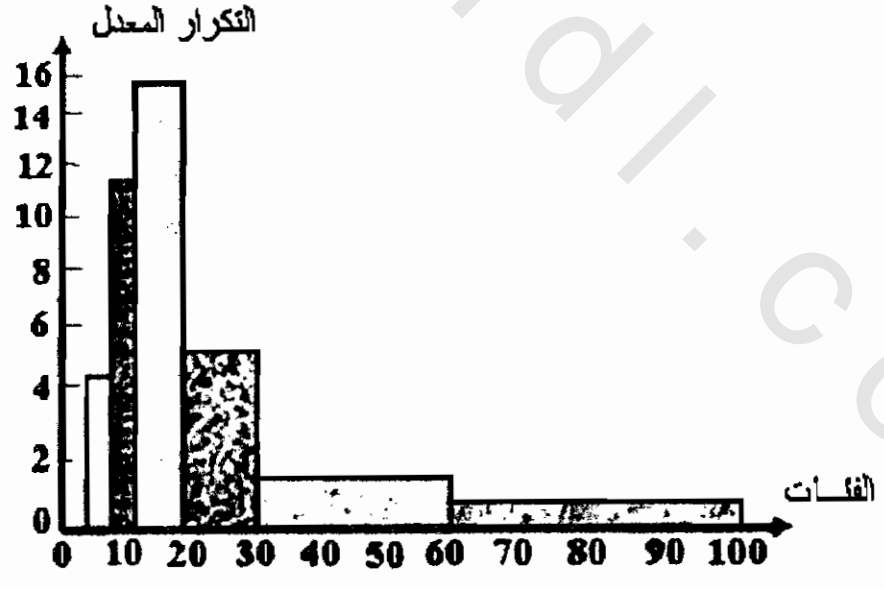
ويلاحظ في هـــذا التوزيع عـــدم تســـاوي أطـــوال الـــفئات ولذلك يلزم تعديل التكرارات قبل رســـم المدرج التكراري كـــما موضح فـــي الجـــدول(1 -13)

وبتمثيل الـــفئات على المحور الأفـــقي والتكرار المعدّل على المحور الرأسي نحصل على المدرج التكراري الموضح في الشكل(1 - 4).

جدول (1-13)

ت	الفئات	التكرار	طول الفئة	التكرار المعدّل
1	- 1	8	2	4
2	- 3	22	2	11
3	- 5	80	5	16
4	- 10	48	10	4.8
5	- 20	27	30	0.9
6	50 وأقل من 100	15	50	0.3
المجموع Total		200		——

وعند النظر للشـــكل(1-4) سنجد أن مساحة كل مـــن المســـتطيلات تســـاوي حاصل ضرب القاعدة أي طول الفئة في الارتفاع أي التكرار المعدّل والتـــي تساوي التكرار الأصلي .

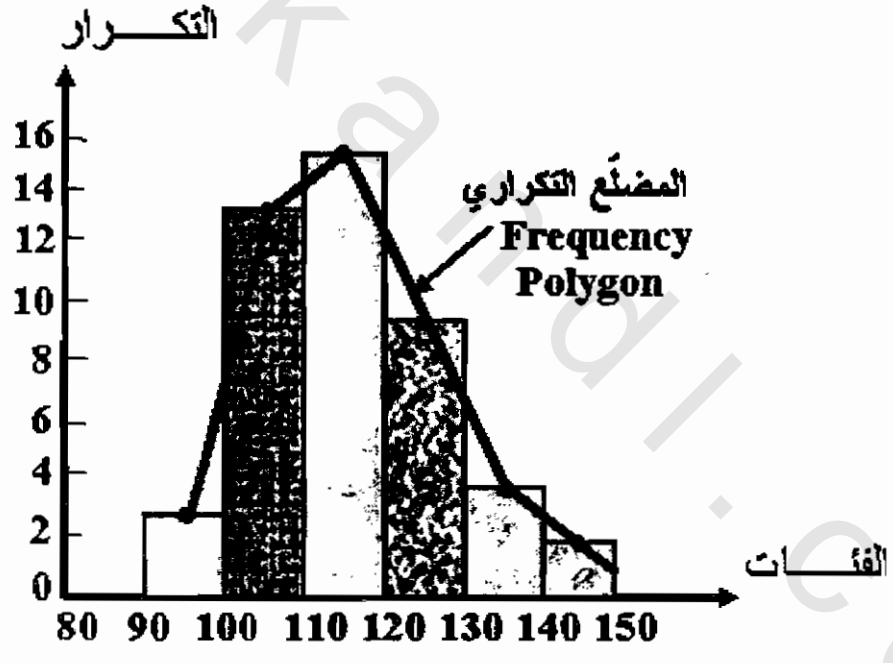


الشكـــل (1 - 4)

المدرج التكراري لفئات عمال المصانع غير المتساوية

2-المضلّــع التكــراري (The Frequency Polygon)

سبق أن أوضحنا عند تكوين الجداول التكرارية انه يمكن اعتبــار مركــز الفئة ممثلاً لجميع القيــم الواقعة في هذه الــفئة ، فلو مثلنا مراكز الفئات على المحور الأفــقي والتكرارات عــلى المحور الرأســي فإنه يمكن تمثيل بيانات الجدول التكراري بعدد من النقاط تقع كل منها أمام مركز الفئة على بعد رأسي تكرار هذه الفئة ، وبتوصيل هذه النقاط بخطوط مستقيمة نحصل على المضلّع التكراري كما موضح في الشكل(1-5) ، الذي يبين المضلع التكراري لتوزيع درجات الذكاء للخمسين مهندس والمدرجة في الجدول(1 - 4) .

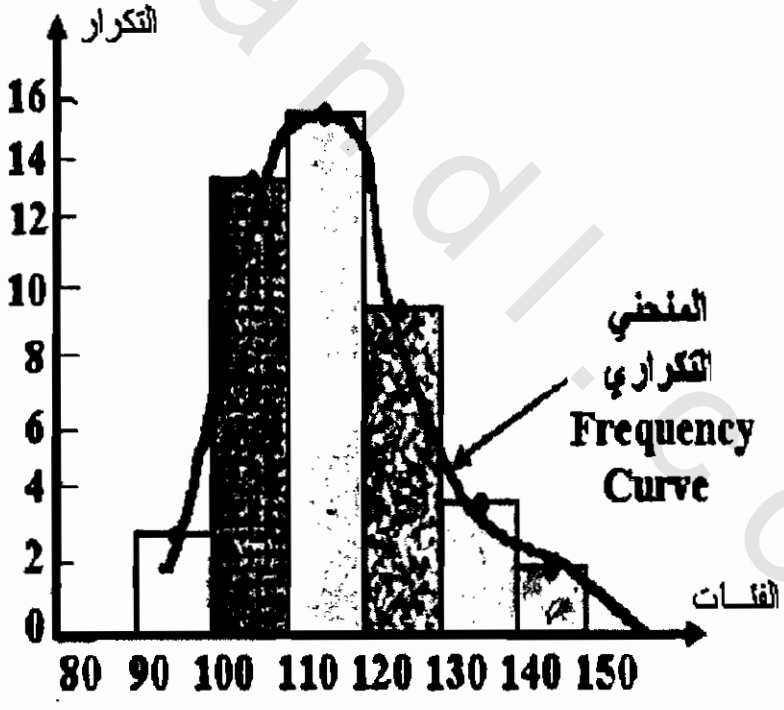


الشكل (1 - 5)

المضلع التكراري لتوزيع درجات الذكاء للخمسين مهندس

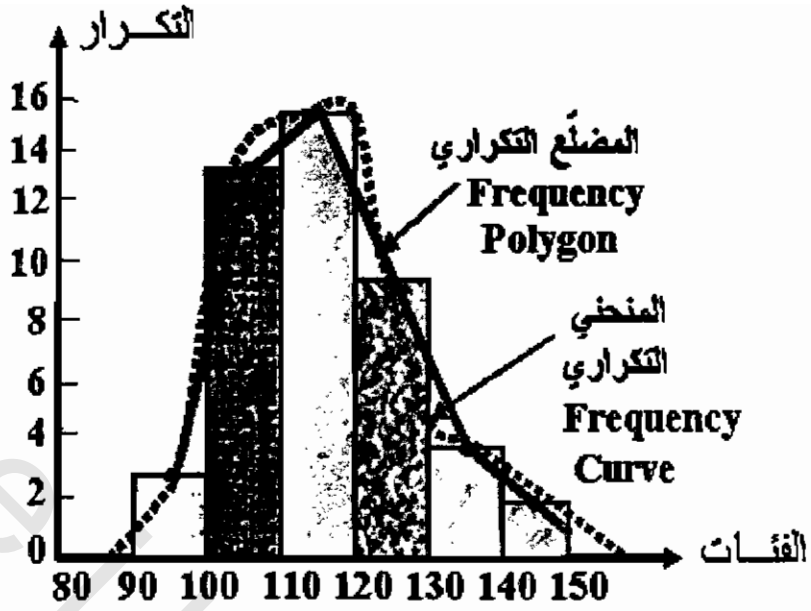
3- المنحنـــــى التكـــراري (The Frequency Curve)

بتمثيل مراكز الفئات على المحور الأفقي والتكرارات على المحور الرأسي نعين عدداً من النقاط ولتكن مراكز الفئات ، ونحاول أن نرسم منحنى بينها بدلاً من الخط المنكســر في حالة المضلــع التكراري عــلى أن يمر بمعظم النقاط ويتوســط بقيتهــا خيــر توســط ، وهــذا المنحنــي يعــرف بــالمنحني التكراري . والشكل (1-6) يبين منحني التوزيع التكراري لــدرجات الــذكاء للخمسين مهندس أما الشكل (1-7) فيبين رســم كل من المضــلع التكــراري والمنحني التكراري في نفس الإحداثيات حتى يمكن المقارنة بينهما حيث يتضح أن المنحني التكراري لا يمر بجميع نقاط المضلّع .



الشكل (1 - 6)

المنحنى التكراري لدرجات الذكاء للخمسين مهندس



الشكـــل (1 - 7)

مقارنة بين المضلع التكراري والمنحنى التكراري

لتوزيع درجات الذكاء للخمسين مهندس

ويلاحظ أنه كلما زاد عدد المفردات وقصرت أطوال الفئات فإن الخط الذي يمر برؤوس المضلع يكون أقـــل انكســـاراً وأكـــثر تمهيداً وسيقترب المدرج التكراري من المنحني التكراري .

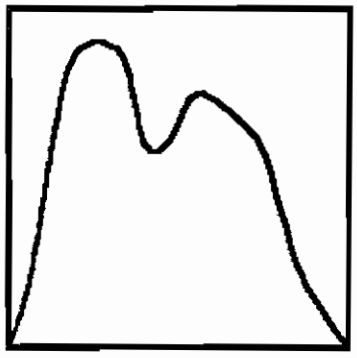
4- أشكـــال المنحنيـــات التكراريـــة (Frequency Curves Types)

تأخذ المنحنيات التكرارية أشكالاً مختلفة كما يتضح من الشكل(1-8) وأكثر هذه المنحنيات شـــيوعاً هو المنحني المعتـــدل(Normal Curve) ، وهـــو منحني متماثل (Symmetric) يشـــبه الـــناقوس وله نهايـــة عـــظمى فـــي منتصفة تماماً ويتماثل حول محور رأسي يمر بهذه النهاية العظمى بحيث يقسم المساحة تحت المنحني إلى قســـمين متطابقين ولهـــذا المنحني أهميـــة كبـــيرة في الدراسات الإحصائية ولذلك سندرسه بالتفصيل فيما بعد .

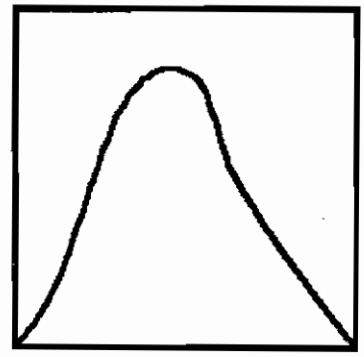
والمنحنى المعتدل هو الشكل الذي نتوقع الحصول عليه من دراسة كثير من الظواهر التي تتغير تبعاً لأسباب طبيعية مثل الطول والوزن وغيرها ، وفي بعض الحالات قد نحصل على أشكال قريبة من المنحنى المعتدل ولكنها ليست متماثلة تماماً فيكون للمنحنى قمة واحدة ولكن أحد طرفيه أطول من الطرف الآخر ويسمى ذلك بالالتواء(Skewness) ويكون الالتواء موجباً إذا كان الطرف الأيمن أو الذيل الأيمن للمنحني هو الأطول(Positive Skewness) ، أما عندما يكون الطرف الأيسر أو الذيل الأيسر للمنحنى هو الأطول فيسمى ذلك الالتواء سالباً (Negative Skewness) .

وفي حالات أخرى قد نجد أن المنحنى التكراري يكون ذا " شعبة واحدة " أو ما يدعى (Reverse J - Shaped Curve) ، ويحدث ذلك عند وجود أكبر التكرارات ومن أمثلة ذلك توزيع ملكية الأراضي حسب فئات المساحة المملوكة حيث أن عدد الملكيات الصغيرة كبير جداً بينما يقل عدد الملاك كلما كبرت المساحة المملوكة . وبذلك يكون للمنحنى فرع واحد أي شعبة واحدة كما في الشكل (1-8) .

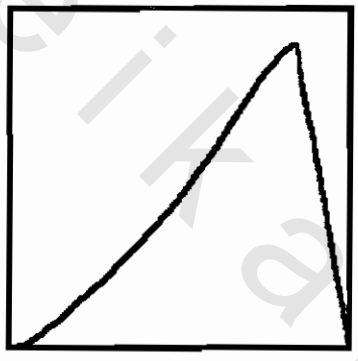
أما إذا وجدت أكبر التكرارات عند طرف المنحنى فإن المنحنى يكون ذو شعبتين(U - Shaped Curve) ويشابه في شكله حرف U ، ومن أمثلة ذلك توزيع الوفيات حسب فئات العمر فنجد أن عدد الوفيات يكون كبيراً في سنوات العمر الأولى ثم يقل في فترة الشباب ثم يرتفع مرة أخرى في فترة الشيخوخة وبذلك نحصل على منحني ذو شعبتين .



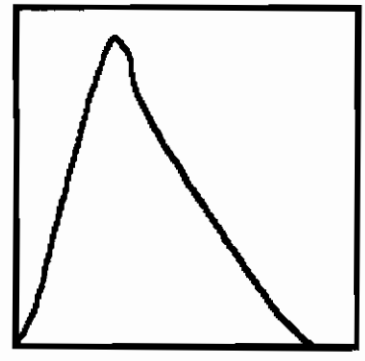
منحنى ذو قيمتين
Bimodal Curve



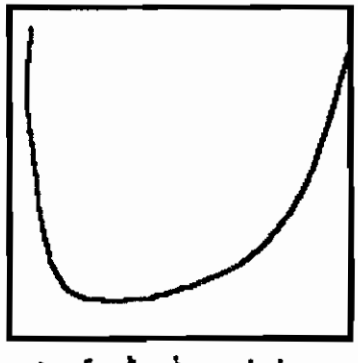
منحنى تكراري معتدل
Symmetric Frequency Curve



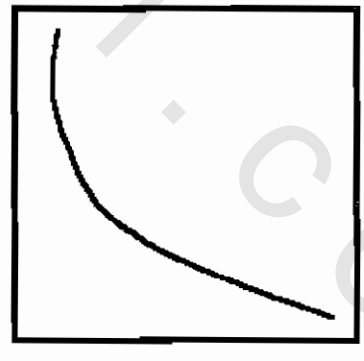
منحنى سالب الالتواء
Negative Skewness Curve



منحنى موجب الالتواء
Positive Skewness Curve



منحنى ذو شعبتين
U-Shaped Curve



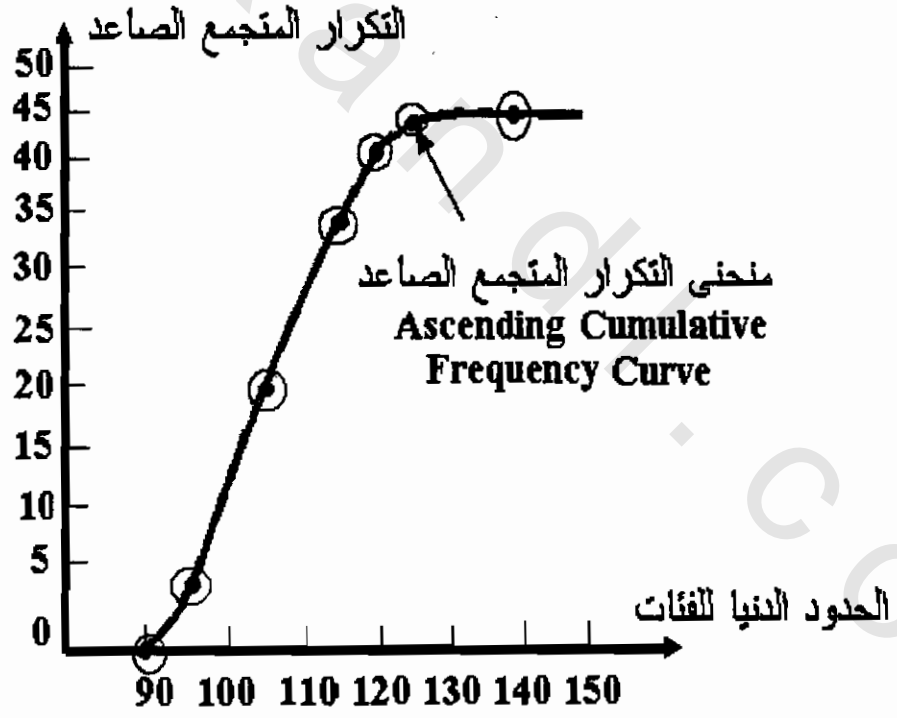
منحنى ذو شعبة واحدة
Reverse J-Shaped Curve

الشكل (1 - 8)

أشكال المنحنيات التكرارية

5- المنحنـــى التكـــراري المتجمـــــع
(Cumulative Frequency Curve)

عند دراســـتنا لتبويب البيانات أوضحنا كيفية تكوين جداول التكرار المتجمع الصاعد والنازل . ولتمثيل هذه الجداول بيانياً ، نمثل حدود الفئات على المحور الأفقي والتكرارات المتجمعة على المحور الرأســـي فنحصل على كـــل مـــن المنحنى المتجمع الصاعد والمنحني المتجمع النازل وفي الشكل(1-9) تم تمثيل منحنى التكرار المتجمع الصاعد لتوزيعات درجات الذكاء للخمســين مهنـــدس وفي الشكل(1-10) تم تمثيل منحنى التكرار المتجمع النازل لدرجات الـــذكاء لنفس الخمسين مهندس .



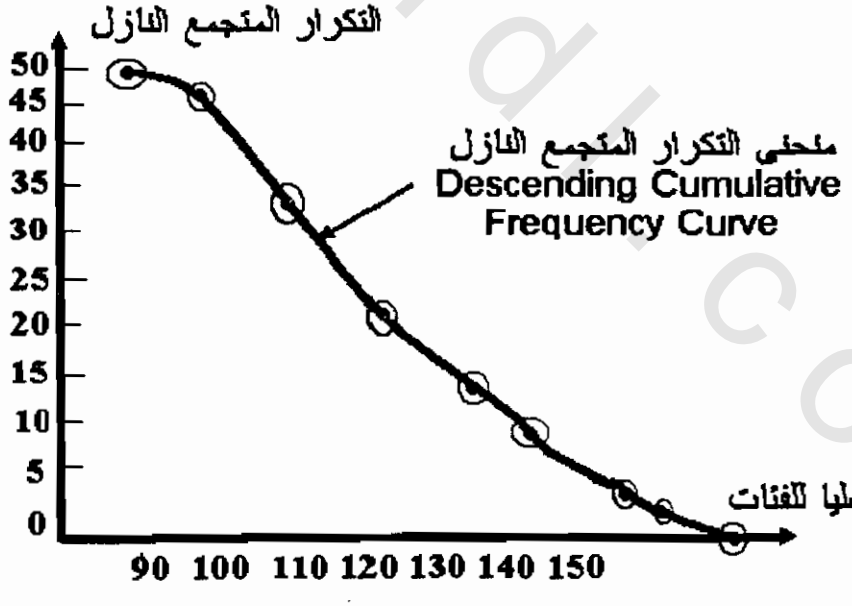
الشكـــل (1 - 9)

منحنى التوزيع المتجمع الصاعد لدرجات الذكاء للخمسين مهندس

ويتضح من الشكل أعلاه أن عدد المهندسين الذين حصلوا على درجات أقل من 110 هو 17 مهندس أي بنسبة 34% من مجموع المهندسين وأن عدد المهندسين الذين حصلوا على درجات أقل من 130 هو 33 مهندس أي بنسبة 66% من مجموع المهندسين وهكذا . وبذلك يمكننا استخدام الشكل في معرفة عدد المهندسين الذين حصلوا على درجات أقل من حد معين وبالتالي معرفة النسبة المئوية لهؤلاء المهندسين ، فلو أردنا معرفة النسبة المئوية للمهندسين الذين حصلوا على درجات أقل من 135 درجة نحسب أولاً عددهم وذلك بإقامة عمود على المحور الأفقي عند القيمة 135 حتى يلاقي المنحني المتجمع الصاعد في نقطة ثم نرسم منها خطاً أفقياً موازياً للمحور الأفقي حتى يقابل المحور الرأسي في نقطة تحدد لنا عدد هؤلاء المهندسين فنجده مثلاً 46 وهولاء تكون نسبتهم :

$$100 \times \frac{46}{50} = 92\ \%$$

مساوية إلى % 92 من مجموع المهندسين .



الشكل (1 - 9)

منحنى التوزيع المتجمع النازل لدرجات الذكاء للخمسين مهندس

9.1 تمارين (Exercises)

س1: الجدول (1-14) يبين عدد الوحدات التي تم سحبها من مخزون أحد المصانع خلال فترة 40 يوماً :

جدول (1-14)

99	83	97	87	82	88	81	91	80	83
98	93	78	87	90	84	92	73	85	75
82	89	101	82	83	88	86	93	80	86
94	103	81	76	92	84	89	80	95	85

المطلوب :

a) تفريغ هذه البيانات في جدول تكراري مناسب .

b) رسم المدرج التكراري والمضلّع التكراري للبيانات .

س2: الجدول (1-15) يبين عينة عينة عشوائية مكونة من 25 عاملاً أخذت من عمال أحد المصانع فوجد أن متوسط الوحدات التي أنتجها هؤلاء العمال في الأسبوع كانت كما يلي :

جدول (1-15)

132	143	171	151	134
160	123	144	145	141
135	156	114	167	140
155	138	147	130	168
138	150	159	130	150

المطلوب :

a) إدراج البيانات أعلاه في جدول تكراري مناسب .

b) رسم المدرج التكراري والمضلّع التكراري للبيانات .

c) مثل منحني التكرار المتجمع الصاعد للبيانات ومنه أستنتج نسبة العمال الذين تقل إنتاجية كل منهم عن 135 وحده أسبوعياً .

س3: الجدول (1-16) يبين عدد الأسر لأقرب مائة والمقيمة في البلديات التالية حسب النتائج الأولية لتعداد 1973 للسكان في ليبيا .

جدول (1-16)

البلدية	أجدابيا	جالو	بني وليد	سرت	الكفرة	المجموع
عدد الأسر	91	15	12	39	32	179

المطلوب تمثيل البيانات أعلاه بطريقة القطاعات الدائرية .

س4: الأعداد المبينة في الجدول (1-17) تمثل درجات 75 طالباً في امتحان لمادة مقاومة المواد في الفصل الثالث لإحدى الكليّات الهندسية .

جدول (1-17)

60	73	85	79	73	93	76	88	62	90	68	82	75	84	68
72	63	78	91	62	74	87	75	65	61	75	85	59	71	83
75	63	79	78	96	60	68	74	69	77	84	75	82	75	66
75	71	65	76	85	78	97	69	62	79	71	83	79	60	90
82	81	73	67	76	74	65	62	66	78	68	57	73	53	80

المطلوب :

a) تمثيل البيانات في جدول تكراري مكون من 8 فئات .

b) تمثيل المدرج التكراري والمنحني التكراري لهذه البيانات .

c) رسـم منحني التكرار المتجمع الصاعـد ومنحني التكرار المتجمع النازل لهذه البيانات .

س5: البيانات في الجدول (1-18) تبين كمية الكهرباء المستهلكة يومياً بآلاف الكيلووات في إحدى المدن خلال فترة زمنية مقدارها 100 يوم .

جـدول (1-18)

25	32	28	26	30	21	36	27	32	32	33	23	31
29	33	31	29	33	28	31	27	29	29	33	27	28
32	23	30	25	29	32	21	24	30	31	28	32	31
34	31	26	30	30	34	26	26	29	32	22	24	31
28	32	28	37	30	35	27	33	28	31	28	32	30
31	30	33	30	28	29	26	29	32	27	27	34	29
31	33	39	30	33	28	33	32	25	31	31	30	31
28	30	30	33	30	31	30	32	28				

المطلوب :

a) أكتب جدول التوزيع التكراري للبيانات ، أختر عدد الفئات 7 .

b) أنشىء المدرج التكراري والمضلّع التكراري للبيانات .

c) أنشىء المدرج التكراري النسبي والمضلع التكراري النسبي للبيانات .

d) أكتب جـدول التكرار المتجمع الصاعد ثم جدول التكرار المتجمع النـازل لهذه البيانات .

e) مثل منحني التكرار المتجمع الصاعد والنازل للبيانات .

س6: أراد عالم نفسي أن يجري تجربة على القردة ليتعرف من خلالها على إمكانية تعلم القردة أن الضوء الأحمر يعني قف والضوء الأخضر يعني تقدّم ، فصار يكافئ القرد بموزة إذا استجاب الاستجابة الصحيحة للإشارة ويعاقبه بصدمة كهربائية إذا هو أخطأ ، أجرى التجربة على 100 قرد وسجل عدد المحاولات التي بذلها كل قرد حتى تعلّم الدرس ، فحصل على النتائج المبينة في الجدول (1-19) :

جـدول (1-19)

1	6	17	11	4	18	24	34	2	26
4	2	7	3	12	5	10	1	15	3
16	5	3	8	13	9	20	1	14	2
6	2	8	4	14	9	5	1	2	10
1	7	7	6	10	5	5	9	3	15
6	3	8	8	13	1	6	13	19	10
2	7	4	9	20	2	7	12	9	18
1	6	3	5	11	8	3	19	8	10
21	2	5	4	30	7	4	13	9	23
1	4	3	6	10	5	16	22	28	32

المطلوب :

a) أكتب جدول التوزيع التكراري للبيانات أعلاه (أختر عدد الفئات 7) .

b) أنشئ المدرج التكراري والمضلّع التكراري للبيانات .

c) أنشئ المدرج التكراري النسبي والمضلع التكراري النسبي للبيانات .

d) أكتب جـدول التكرار المتجمع الصاعد ثم جدول التكرار المتجمع النـازل لهذه البيانات .

e) مثل منحني التكرار المتجمع الصاعد والنازل للبيانات .

س7: جدول (1-20) يبين التوزيع التكراري لعدد الهكتارات التـي يزرعهـا 150 مزارعاً في إحدى المناطق الزراعية :

جـدول (1-20)

الفئة (بالهكتار)	50 -	56 -	62 -	68 -	74 -	80 -	86 -	90 - 98
المزارعين (التكرار)	2	12	24	30	44	20	12	6

المطلوب :

a) مثل المدرج التكراري والمضلع التكراري والمنحني التكراري للبيانات .

b) أنشئ جدول التكرار المتجمع الصاعد للبيانات ومثله بالرسم .

c) أنشئ جدول التكرار المتجمع النازل للبيانات ومثله بالرسم .

س8: الجدول (1-21) يبين العمر بالشـهر الذي تـعلّم فيه أربعـون طفـلاً ركوب الدراجة ذات العجلات الثلاثة :

جـدول (1-21)

21	23	31	35	18	19	22	25	27	16
26	29	28	25	25	21	28	32	25	21
23	26	24	20	23	24	26	20	28	20
32	24	25	28	29	31	26	32	29	24

المطلوب:

a) نظّم البيانات السابقة في جدول توزيع تكراري مكون من 5 فترات .

b) مثل التوزيع التكراري في(a) بمدرج تكراري ومضلع ومنحني تكراري .
c) على نفس ورقة الرسم مثل التكرار المتجمع الصاعد والنازل للبيانات .

س9: الجدول (1-22) يبين الأجور الأسبوعية بالدنانير للعاملين بأحد شركات البترول :

جـدول (1-22)

فئات الأجور (بالدنانير)	50 -	60 -	70 -	80 -	90 -	100 -	110 - 120
عدد العاملين (التكرار)	8	10	16	14	10	5	2

المطلوب :

a) أرسم منحني التكرار المتجمع الصاعد للبيانات ومنه أستنتج نسبة العاملين الذين يحصلون على أجـر أقـل من 85 دينار أسبوعياً وعدد العاملين الـذين يحصلون على أجر يتراوح بين 63 وأقل من 75 دينار أسبوعياً .

b) أرسم منحني التكرار المتجمع النازل للبيانات ومنه أستنتج نسبة العـاملين الذين يحصلون على أجر يساوي أو يزيد عن 98 دينار أسبوعياً .

س10: الجدول (1-23) يبين المسـاحة المزروعـة قمحـاً($\bar{x}$) بـآلاف الهكتارات وكمية الإنتاج من القمح($\bar{y}$) بآلاف القنطارات والتي جمعت من 10 مناطق .

جـــدول (1-23)

المنطقة	1	2	3	4	5	6	7	8	9	10
$\bar{x}$	146	138	126	130	142	128	148	148	138	132
$\bar{y}$	999	993	996	987	993	1005	981	1008	1002	993

المطلوب :

تمثيل البيانات أعلاه في جدول تكراري مزدوج مناسب .

opeikandl.com