Université Abou Bekr Belkaid جامعة أبي بكر بلقايد

Tlemcen Algérie

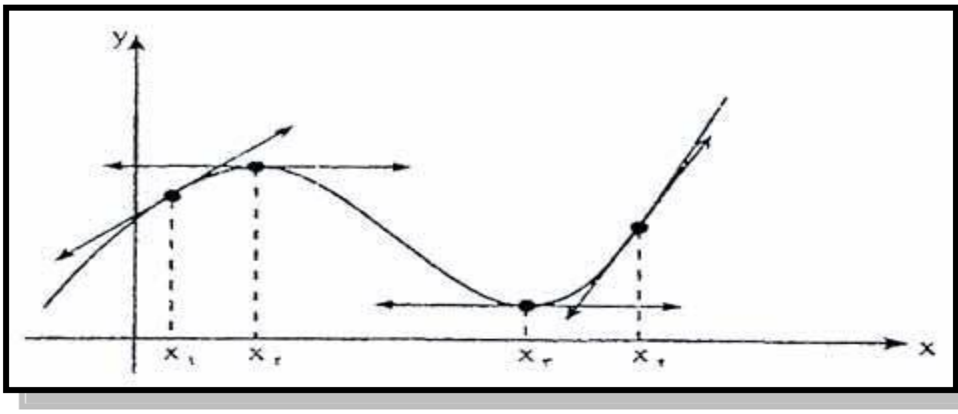
تلمسان الجزائر

كلية العلوم الاقتصادية، التجارية

وعلوم التسيير

السنة الأولى LMD

محاضرات في الاحصاء



الأستاذة المكلفة بالمحاضرات:

د.طالب دليلة

السنة الجامعية 2021-2022

بسم الله الرحمن الرحيم

تقديم

هذه المطبوعة هي عبارة عن محاضرات الإحصاء حسب البرنامج الوزاري لمقياس "إحصاء 1 و 2" للسنة الأولى علوم اقتصادية ، التجارية و علوم التسيير، و قد تم تقسيمه إلى قسمين : **الاحصاء الوصفي** الذي يعتبر المادة الأولية الأولى التي يكون الطالب بحاجة إليها قبل الخوض و التوسع في **الاحصاء الرياضي** الذي يعتبر كقسم ثاني، باعتبارها فرعا من الرياضيات.

هذه المطبوعة هي ثمرة تجربة سنوات عديدة في تدريس الإحصاء بكلية العلوم الاقتصادية لجامعة ابو بكر بلقايد بتلمسان، ولقد حاولنا أن نستفيد من هذه التجربة لصياغة محتوى المقياس بطريقة تلائم مستوى طلبة هذه الكلية و طبيعة التخصص، لتحقيق هذا الغرض حرصنا على ربط المفاهيم والقواعد النظرية باستخداماتها التطبيقية؛ فعملنا على إعطاء أمثلة محلولة عن كل مفهوم جديد. و لأن فهم القواعد الرياضية يكون أسهل إذا كان للمتلقي خلفية عن المشكلة التي يحتاج حلها إلى استخدام هذه القواعد، عملنا في كثير من الأحيان إلى التقديم لبعض الدروس أو النظريات بمسألة تكون بمثابة التمهيد، وأحيانا بمثابة مشكلة ننطلق منها لنتوصل إلى النظرية.

يتضمن البرنامج المقرر على قسمين كما أسلفنا سابقا فيتضمن القسم الأول على أربعة فصول، حيث يتناول الفصل الأول علم الإحصاء بصفة عامة إلى جانب عرض البيانات الاحصائية و التوزيعات التكرارية، ليتناول الفصل الثاني مقاييس النزعة المركزية والتي لها عدد من المتوسطات للتعبير عنها تختلف باختلاف الغرض الذي تستخدم فيه، وطبيعة البيانات المحسوبة منها، أهم هذه المقاييس: المتوسط الحسابي، الوسيط، المنوال، في حين يتضمن الفصل الثالث مقاييس التشتت التي تدرس مدى تباعد أو تقارب البيانات من بعضها البعض أو من متوسطها ، أما عن الفصل الرابع فيتضمن مقاييس الشكل إذا كانت مقاييس النزعة المركزية ومقاييس التشتت تسمح بتلخيص بيانات أي ظاهرة في صورة أرقام بسيطة تعطي فكرة عن خصائص توزيع هذه البيانات ودرجة تجانسها أو اختلافها فإن هذا الوصف الذي تعطيه يبقى تنقصه الدقة الكافية المطلوبة للتعرف على خواص التوزيع خاصة من حيث انتشار البيانات على المنحى البياني الممثل لها من حيث التوائه أو تفلطحه عن الوضع الطبيعي كذلك دعت الحاجة لاستخدام مقاييس أخرى لتحقيق هذا الغرض سميت هذه المقاييس بمقاييس الالتواء والتفلطح التي سنتعرف عليها في هذا الفصل.

اما القسم الثاني من هذه المطبوعة فيتضمن ثلاثة فصول ، يتضمن الفصل التمهيدي على ملخص حول التحليل العاملي من خلال التوفيقات ، الترتيبات و التبديلات، أما الفصل الأول فيتمحور حول نظرية الاحتمالات حيث يستخدم الكثير من الأشخاص وفي حالات مختلفة كلمة احتمال ولكن بدرجات مختلفة من الثقة، فإذا قال الطالب مثلا أنه سيحضر محاضرة الإحصاء الوصفي هذا الصباح فهذا يعني أنه متأكد من الحضور، أما إذا قال أنه يحتمل الحضور إلى هذه المحاضرة فهذا يعني أنه قد يحضر وقد لا يحضر، وهنا فإنه لا يمكننا معرفة درجة الثقة في الحضور،من أجل هذا جاءت الحاجة إلى وضع مقاييس رقمية تبين درجة التأكد وتعوض العبارات التي قد لا تعطي المعنى الدقيق، هذه المقاييس يمكن الحصول عليها من علم الاحتمالات، أما الفصل الثاني و الأخير فيتمحور حول التوزيعات الاحتمالية من التوزيعات الاحتمالية المتقطعة و التوزيعات الاحتمالية المستمرة إلى جانب التوزيعات الاحتمالية المتقطعة الاكثر استخداما.

فيما يتعلق بما تحتاجه متابعة وفهم هذا القسم ، من المهم التمييز بين الفصل الأول وبقية الفصول الأخرى من الاحصاء الرياضي، فالفصل الأول الذي يتضمن المفاهيم الأساسية لعلم الاحتمالات لا يحتاج استيعابه إلى مستوى عالي في الرياضيات، أما باقي الفصول فيتطلب فهمها أن يقوم الطالب بمراجعة عدد من المفاهيم الرياضية أغلبها متضمنة في برنامج الرياضيات للسنة الأولى. تتمثل هذه المفاهيم أساسا في الدوال، الاشتقاق، التكامل (خاصة التكامل بالتجزئة) والدوال الأسية.

د.طالب دليلة

نبذة تاريخية عن تطور علم الاحصاء

قبل الشروع في دراسة الطرائق المختلفة للإحصاء يستحسن أن يحيط الطالب بنظرة عن التطور التاريخي للإحصاء كممارسة وكعلم، وأن يطلع على مجموعة من أبرز من كتبوا في هذا العلم.

الفترة ما قبل الميلاد إلى غاية القرن 18: تدل الحفريات التي وجدت في أماكن متعددة على استخدام الإحصاء من قبل عدد من الحضارات القديمة عبر المعمورة. منذ القدم استخدم الحكام والأمراء الإحصاء كوسيلة للرقابة، و أداة لإدارة المملكة أو المدينة أو المقاطعة، واستخدموا في ذلك تعداد السكان وجرد السلع والموارد المختلفة. في الحضارة السومرية، التي سادت في بلاد ما بين النهرين 5 آلاف إلى ألفي سنة قبل الميلاد، والتي ازدهرت فيها التجارة بشكل كبير، كانت قوائم من السلع والأشخاص تدون على ألواح من الصلصال، وقد وجدت حفريات مشابهة تثبت استخدام الجرد في عهد الحضارة المصرية التي سادت 3 آلاف سنة قبل الميلاد. الحضارة المصرية التي قامت على التسيير والتقسيم الدقيق لمياه النيل اتسمت إدارتها بالمركزية الشديدة وهذا الذي أعطى الأهمية **للتدوين** كوسيلة للمراقبة، فقد كان للمصريين القدامى مدارس يتعلم فيها الموظفون القراءة والكتابة والقوانين المعمول بها، وكان مما يتعلمه الموظف أن لا اعتبار لأمر أو عقد ما لم يكن مكتوبا. واستخدم الجرد لدى جميع الحضارات القديمة تقريبا كالحضارة الصينية والهندية واليابانية واليونانية والرومانية، وكذا حضارة الإنكا في الساحل الغربي لأمريكا الجنوبية (ابتداءا من القرن 12 إلى غاية 1572) . في هذا العهد كان الإحصاء عبارة عن جرد المواد والأفراد وأحيانا نجد نظاما لتصنيف المعلومات لكن لم يوجد دليل على عمليات معالجة لهذه المعطيات.

في العهد الإسلامي كان الخليفة عثمان (ر) أول من أمر بالتدوين لإحصاء المستفيدين من عطايا بيت المال، أما في أوربا فنجد أن أول الآثار عن عمليات التعداد ترجع إلى 1086 فقط وبالتحديد في بريطانيا. أما في فرنسا فإن عمليات التعداد ترجع إلى القرن 14 الذي شهد ميلاد أول تسجيلات عقود الحالة المدنية وإجبارية تسجيل عقود الازدياد في عهد فرنسوا الأول. في فرنسا دائما تجدر الإشارة إلى أنه في القرن17 حين أراد "كولبيرت" - أب الإدارة الفرنسية – أن يدفع ببلاده إلى المستوى الصناعي الذي بلغته بريطانيا في ذلك الوقت، أسس إدارة مركزية قوية... وكان من منجزاته أن شهدت وزارته (1630-1660) عددا من عمليات التحقيق الكبرى، وشهدت ألمانيا وبريطانيا تطورا مشابها بالإضافة إلى دول أخرى. وقد كان "قرانت" (1620-1674 : GRANT) أول من استعمل في 1662 مصطلحات علم السكان مثل الخصوبة وطول مدة الحياة؛ كما قارن بين معدلات ولادة الإناث والذكور. وقد طور هذا العالم مع عالم آخر هو بيتي (PETTY) طريقة لتعداد السكان من خلال المعطيات الثانوية (عن عدد المساكن، عدد الوفيات...) تدعى "طريقة المضاعف" (Multiplicateur) عرفت بعد ذلك تحسينات متتالية على أيدي علماء آخرين منهم خاصة "لابلاس" (LAPLACE) في 1785.

ظهور نظرية الاحتمالات في قرن17 و18: تاريخيا ارتبط ظهور نظرية الاحتمالات بألعاب الحظ التي كانت سائدة بكثرة في أوربا في القرن السابع عشر وتنظمها البنوك بشكل خاص. لكن قلة انتشار طباعة الكتب والأجواء الدينية السائدة التي لا تبارك هذه الألعاب منعت انتشار الكتابات في هذا الشأن. وينسب البعض أول الكتابات في علم الاحتمالات إلى العالم "باسكال" (PASCAL 1662 -1623) الذي كتب عما أسماه آنذاك "هندسة الحظ" (La géométrie du hasard). وكان ذلك من خلال رسائل له مع زميله المعروف هو الآخر "فرمات" (1601 – 1665: FERMAT) . وتذكر في هذا الصدد بشكل خاص المسألة التي طرحها على باسكال أحد هواة الألعاب "كم ينبغي من رمية لمكعبي نرد حتى يمكن المراهنة بتفاؤل على الحصول على مجموع 12؟ ". ثم جاء علماء آخرون كانت لهم إضافات بارزة في هذه الفترة مثل هايجان (HUYGEN : 1629 – 1695)، جاك برنولي (JACQUES BERNOULLI)، موافر (MOIVRE) وكذا لايبنيتز (1646 – 1716: LEIBNIZ). كما

ساهم في هذه الفترة التي سبقت القرن 19 علماء كبار أمثال (GAUSSE, BAYES, LAPLACE)عرفت نظرية الاحتمالات على أيديهم إنجازات كبيرة.

القرن 19: في هذا القرن برزت إحدى أهم عناصر نظرية الاحتمالات وهي "التوزيع الطبيعي" وذلك لقياس نسبة الخطأ في مجال الحسابات الفلكية. كان هذا من ثمرة عمل العالمين لابلاس وقوس(GAUSSE وLAPLACE). في هذا القرن أيضا ظهرت حسابات الارتباط لقالتو (GALTOU) كما برزت أسماء مثل كتلت (QUETLET) وآخرون .

القرن العشرون: نظرية الاحتمالات كما نراها الآن، أي بصياغة رياضية ناضجة في شكل قوانين مبرهن عليها رياضيا، إنما تبلورت في القرن العشرين وبالضبط في بدايته. ومن الأسماء التي برزت في الفترة الأولى (1980 – 1920) من هذا القرن نجد من بريطانيا بيرسون (KARLE PEARSON) ومن روسيا ماركوف (MARKOV) ومن فرنسا بوريل (BOREL). في الفترة الثانية (1921 – 1932) درست مسائل التوقع، حيث كان لفيشر (FISHER) دورا بارزا.

في الفترة الممتدة من 1933 إلى نهاية الحرب العالمية الثانية برزت اختبارات الفروض على يد نايمان (NEYMAN) وإيقون بيرسون (EGON PEARSON) وبداية النظرية الحديثة للمعاينة لنايمان (NEYMAN) بالإضافة إلى خطط التجارب لفيشر. بداية من الخمسينات تكاثرت الكتابات في مجال الإحصاء حيث عرفت نظرية التقدير وتحليل البيانات. وبالتدريج انتشر استخدام الإحصاء في الميادين المختلفة والعلوم التجريبية والإنسانية.

و غني عن الذكر أن محتوى هذه المطبوعة من نظريات وقواعد ليس من إبداع مؤلفها، و إنما هي قواعد مبسوطة في المراجع جمعناها وعرضناها بأسلوب رأينا أنه الأنسب لمستوى طالب كلية العلوم الاقتصادية. وإذ نقدم لطلبتنا و زملائنا هذا العمل المتواضع، نهيب بهم أن لا يبخلوا علينا بملاحظاتهم وتعليقاتهم حتى نستفيد منها في طبعات مقبلة بحول الله.

البرنامج المقترح لمقياس الاحصاء

تقديم

القسم الأول: الاحصاء الوصفي

الفصل الأول: عرض البيانات الاحصائية ووصفها

أولا: تعريف الإحصاء

ثانيا: التعريف المصطلحات الإحصائية

ثالثا: العرض الجدولي للبيانات

رابعا: العرض البياني للبيانات

تمارين

الفصل الثاني: مقاييس النزعة المركزية

أولا: المتوسط الحسابي

ثانيا: الوسيط

ثالثا: أشباه الوسيط

رابعا: المنوال

تمارين

الفصل الثالث: مقاييس التشتت

أولا: المدى العام

ثانيا: مقاييس التشتت حول الوسيط

ثالثا: مقاييس التشتت حول الوسط الحسابي

تمارين

الفصل الرابع: مقاييس الشكل

أولا: العزوم

ثانيا: تحديد شكل التوزيع

تمارين

القسم الثاني: الاحصاء الرياضي

فصل تمهيدي: ملخص حول التحليل العاملي

أولا: المبدأ الأساسي للعد

القسم الاول

الاحصاء الوصفي

الفصل الأول

مقدمة

من المفاهيم الشائعة بين الناس عن الاحصاء، ماهي إلا أرقام و بيانات رقمية فقط، كأعداد السكان، أعداد المواليد و الوفيات وخلافه، من ثم ارتبط مفهوم الناس عن الاحصاء بأنه عد أو حصر الأشياء و التعبير عنها بأرقام و هذا هو المفهوم المحدود لعلم الاحصاء، و لكن الاحصاء كعلم، هو الدي يهتم بطرق جمع البيانات، تبويبها و تلخيصها بشكل يمكن الاستفادة منها في وصف البيانات وتحليلها للوصول إلى قرارات سليمة في ظل ظروف عدم التأكد.

أولا – تعريف علم الإحصاء

يجمع معظم العلماء على أن أصل مصطلح كلمة الاحصاء *Statistique* هو ألماني و الذي يعبر عنه باللغة الألمانية ب *Statistik* بحيث تعود جذوره إلى الكلمة اللاتينية *Status* و التي تعني الوضع أو الحالة.

وردت عدة تعاريف مختلفة لعلم الاحصاء و لكن التعريف الأشمل و الأمثل يعرف الإحصاء بأنه علم الذي يبحث في طريقة جمع البيانات عن الظواهر التي تحيط بنا سواء كانت علمية أو اقتصادية أو اجتماعية، وفي كيفية تسجيل هذه الحقائق والبيانات في صورة دقيقة، ثم وصفها بصورة سهلة تبين علاقات واتجاهات هذه الظواهر، وأخيرا يبحث في دراسة هذه العلاقات والاتجاهات ووضعها في صورة يسهل معها فهم الظواهر المراد دراستها.

كما يعرف علم الإحصاء بأنه علم اتخاذ القرارات في جميع نواحي الحياة، وذلك من خلال جمع ودراسة وتحليل البيانات المتوفرة واستخلاص النتائج عن الظواهر المدروسة مما سبق يمكن تصنيف الإحصاء كعلم إلى قسمين رئيسيين هما:

1 – الإحصاء الوصفي Statistique descriptive: وهو ذلك الفرع من الإحصاء الذي يتناول طرق تنظيم وتلخيص وعرض البيانات في صورة مبسطة.

2 – الإحصاء الرياضي :Statistique Mathématique وهو ذلك الفرع من الإحصاء الذي يهتم بطرق الوصول إلى نتائج معينة أو توقعات ما عن المجتمع من خلال دراسة عينة من ذلك المجتمع، و الهدف منه تشكيل قوانين انطلاقا من الملاحظات المأخوذة من العينات المدروسة.

مصادر البيانات:

مما لا شك فيه أن في الدراسات الإحصائية تعد البيانات المادة الأساسية الرئيسية، وعليها تتوقف دقة الوصف والتحليل وسلامة الاستنتاج ومنطقيته، فإذا كانت هذه البيانات والمعلومات دقيقة وشاملة وواقعية، كان الوصف والاستنتاج والقرار الذي نحصل عليه سليما وصحيحا، وعليه فالاهتمام التام والحرص الدقيق في الحصول على بيانات سليمة وواقعية حول الظواهر تحت الدراسة بعد أن العمود الفقري والحجر الأساسي في علم الإحصاء، وهناك عدة مصادر للحصول على البيانات تختلف باختلاف موضوع الدراسة والغرض منها، من أهم هذه المصادر ما يلي:

- النشرات والدوريات وسجلات.
- التجارب.
- الاستبيان.
- التعدادات العامة.

ثانيا – التعريف بالمصطلحات الإحصائية

1 – المجتمع الاحصائي Population: هو جميع العناصر المشتركة في الصفة التي تهم الباحث في دراسته، فقد يكون المجتمع مثلا عدد سكان مدينة، أو طلبة جامعة التكوين المتواصل، أو المساحات الزراعية في الجزائر أو إنتاج آلة معينة ... إلخ، و يرمز له بN.

2 – العينة Échantillon: هي جزء من المجتمع تحت الدراسة مثل مجموعة من سكان مدينة، أو مجموعة من طلبة جامعة التكوين المتواصل، أو بعض المساحات الزراعية في الجزائر ...الخ.

3 – الظاهرة Phénomène: هي صفة لعناصر تختلف من عنصر لآخر في شكل أو النوع أو الكمية، ويطلق على الصفة تحت الدراسة متغير variable مثل طول شخص ما، عدد الأخطاء الإملائية في بحث ما، سرعة سيارة بين مدنيين خلال أسبوع ... الخ.

4- **الميزة الاحصائية**: هي الصفة، أو الخاصية التي يتم اختيارها في الدراسة الاحصائيةX (كمية أو نوعية)

أ- الصفة الكمية **quantitative** : هي المشاهدات أو الصفات التي يمكن قياسها مباشرة بوسائل القياس المألوفة و يعتبر العمر، الطول و الوزن ... أمثلة لهذه الصفة، و تتميز بأنه يمكننا أن نصفها عدديا بأنها أكبر من أو أقل من قيمة عددية.

ب- الصفة النوعية **qualitative** (كيفية أو وصفية): و هي عبارة عن صفات أو أنواع معينة ليست عددية و تنقسم بدورها إلى: بيانات خاضعة للترتيب مثل المستوى التعليمي أو الرتب العسكرية، و بيانات غير خاضعة للترتيب: مثل الجنسية أو أنواع السيارات...

5 – المتغير variable: هو الصفة تحت الدراسة كما أشرنا أعلاه أو هو الشيء الذي يمكن أن يأخذ قيما مختلفة في الظروف المختلفة (زمنية، مكانية، سياسية، اقتصادية... الخ) فمثلا سعر التمر يختلف من يوم لآخر ويختلف في نفس السوق من سنة لأخرى، و تنقسم المتغيرات إلى نوعين:

يعبر عنها في صورة عددية وتنقسم إلى:

أ – متغير متقطع Variable discrète: وهو المتغير الذي يأخذ أعداد صحيحة، فمثلا إذا كان × متغير يمثل عدد أفراد الأسرة، فإنه لا يمكن أن يأخذ القيم 2، 3، 4، 5 ... ولا يمكن أن نأخذ × القيم 1.5، 3.25، 5.17، و يرمز له ب X_i

ب – متغير متصل (مستمر) Variable Continue: وهو المتغير الذي لا يمكن أن يأخذ أي قيمة بين قيمتين معنيتين، وكأمثلة عن المتغيرات المتصلة: الطول، الوزن، الزمن، السرعة ...الخ، فإذا كان C هو متغير الطول فمثلا فإن × يمكن أن تأخذ القيم 15متر، 11.3 متر، 14.75 متر، أي أن المتغير C يمكن أن نأخذ أي قيمة في فترة زمنية معينة، و يرمز له بC_i.

6– الأوضاع i: تسمى المواقع التي يمكن أن تأخذها أي مفردة بصفة معينة أي المواقع التي يمكن أن تأخذها XأوC .

7– التكرار المطلق ni: عدد المرات التي يتكرر فيها المتغير X_i هذا يعني أن $n_1+n_2+\ldots+n_k=\sum ni=N$

– **التكرار التجميعي الصاعد $N_i^{\nearrow}$:** عدد المرات أو المفردات التي تقل قيمتها عن قيمة المتغير الموافقة له.

– **التكرار التجميعي النازل $N_i^{\searrow}$:** عدد المرات أو المفردات التي تزيد قيمتها عن قيمة المتغير الموافقة له.

8– التكرار النسبي f_i: هي نسبة التكرارات المطلقة من مجموع التكرارات بحيث $fi=ni/\sum ni$، $\sum fi=1$

– **التكرار النسبي التجميعي الصاعد $F_i^{\nearrow}$:** نسبة التكرارات التي تقل قيمتها عن قيمة المتغير الموافقة له.

– **التكرار النسبي التجميعي النازل $F_i^{\searrow}$:** نسبة التكرارات التي تزيد قيمتها عن قيمة المتغير الموافقة له.

9- الجدول الاحصائي: يتكون من عمودين، يختص العمود الأول لأوضاع الصفة أو الظاهرة المدروسة، أما العمود الثاني للتكرارات المطلقة أو النسبية.

ثالثا – العرض الجدولي للبيانات

3-1 أنواع الجداول التكرارية:

تلخص الجداول التكرارية البيانات الكمية الكثيرة في وضعها على صورة جدول منتظم يوضح كيفية توزيع القيم التي حصلنا عليها من الظاهرة المدروسة و كما ذكرنا سابقا يدل العمود (السطر) الأول على قيم الظاهرة، ويدل العمود (السطر) الثاني على التكرار المقابل لهذه القيم، و تنقسم إلى الجدول البسيط أو المركب.

3-1-1: الجدول البسيط:

مثال1: (بيانات ذات طبيعة متقطعة)

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

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

لخص هذه البيانات في صورة جدول يوضح معالمها الأساسية؟.

الدرجة	4	5	6	7	8	9	11	12	**المجموع**
التكرار	7	1	5	2	5	4	4	2	30

إن وضع البيانات بهذه الصورة أصبح أكثر وضوحا لمعرفة عدة معلومات كانت غير واضحة في الصورة الخام. فمثلا من السهل الآن معرفة أكبر وأصغر درجة عليها هؤلاء الطلبة كما يمكن معرفة أن الناجحين أو عدد المقصيين ببساطة.

مثال 2: (بيانات ذات طبيعة مستمرة)

البيانات التالية تمثل توزيع العمال على حسب الأجر الأسبوعي

فئات الرواتب	**عدد العمال**
]1000-800]	26
]1100-1000]	33
]1200-1100]	64
]1300-1200]	07
أكثر من 1300	10
المجموع	**140**

3-1-2: الجدول المركب:

يستعمل الجدول المركب عند دراسة خاصيتين في نفس الوقت في مجتمع ما. وتوضح البيانات الإحصائية في هذا الجدول بالشكل التالي:

- نخصص الأسطر لبيانات الخاصية الأولى ونخصص الأعمدة لبيانات الخاصية الثانية.
- نرمز لقيم الخاصية الأولى بالرمز X_i حيث i تتراوح من 1 إلى n و

نرمز لقيم الثانية بالرمز y_i حيث j تتراوح من 1 إلى n.

وكمثال على ذلك عند دراسة مستوى معيشة العائلات يمكن أن نتطرق إلى خاصيتين هما: مهنة رب الأسرة والإنفاق الاستهلاكي، فإذا رمزنا للمتغير الأول مهنة رب الأسرة بالرمز Xi والمتغير الثاني الإنفاق الاستهلاكي بالرمز y_j يمكن أن نشكل الجدول المزدوج التالي:

ni	y_4	y_3	y_2	y_1	المتغير y / المتغير X
N_1	N_{14}	N_{13}	N_{12}	N_{11}	X_1
N_2	N_{24}	N_{23}	N_{22}	N_{21}	X_2
N_3	N_{34}	N_{33}	N_{32}	N_{31}	X_3
N_4	N_{44}	N_{43}	N_{42}	N_{41}	X_4
N_5	N_{55}	N_{53}	N_{52}	N_{51}	X_5
$\sum n_i = \sum n_j$	n_4	n_3	n_2	n_1	n_j

مثال 3: سحبت عينة عشوائية من مجتمع ما، تتكون من 183 أسرة قصد دراسة خاصيتين هما مهنة رب الأسرة والتركيب الأسرية من حيث عدد الأطفال، فكانت النتائج كالتالي:

المجموع	6	5	4	3	2	1	0	عدد الأطفال / المهنة
32	0	1	3	4	6	8	10	إطار المتوسط
28	0	0	1	2	3	7	15	مهنة حرة
57	7	9	15	10	8	5	3	عامل متخصص
66	10	17	13	11	9	4	2	عامل بسيط
183	17	27	32	27	26	24	30	**المجموع**

يمكن أن نقرأ من الجدول السابق ما يلي:

- من بين 183 أسرة المدروسة هناك 28 أسرة يشتغل رب الأسرة بالمهن الحرة.
- من بين 183 أسرة المدروسة هناك 32 أسرة لها أربع أطفال.
- أن هناك 8 عمال متخصصين لهم طفلين و 10 لهم 3 أطفال.
- أن هناك 17 أسرة فقط لها 6 أطفال منها 7 أسر يشتغل رب الأسرة كعامل متخصص و10 أسر يشتغل رب الأسرة كعامل بسيط.

3-2 بناء الجداول التكرارية:

تختلف طرق بناء الجداول التكرارية على حسب نوع و طبيعة المتغير المدروس.

- **العرض الجدولي في حالة البيانات الكمية**

- هناك الطبيعة المتقطعة و الطبيعة المستمرة

أ- **الطبيعة المتقطعة**: لبناء الجدول نتبع المراحل التالية

1- نرتب البيانات ترتيبا تصاعديا أو تنازليا

2- كتابة قيم المتغير داخل عمود الأوضاع

3- نحسب عدد المرات التي يتكرر فيها كل متغيرX_i ثم نكتب النتيجة أمامه.

مثال1: تمثل البيانات التالية النقاط التي تحصل عليها الطلبة في امتحان معين، ضعها في جدول تكراري

12	13	14	20	10
10	17	15	18	15
18	12	17	15	14
16	15	19	16	18

نقاط	10	12	13	14	15	16	17	18	19	20	∑
ni	2	2	1	2	4	2	2	3	1	1	**20**
fi	0.1	0.1	0.05	0.1	0.2	0.1	0.1	0.15	0.05	0.05	**1**

xi: قيم الظاهرة المدروسة،ni : التكرارات المطلقة، fi : التكرارات النسبية

ب-الطبيعة المستمرة: لبناء الجدول نتبع المراحل التالية

1- نرتب البيانات ترتيبا تصاعديا أو تنازليا

2- نحدد المجال (المدى) الذي تنتشر فيه البيانات، وهو الفرق بين أكبر قيمة للبيانات وأصغر قيمة لها، أي أن: **المدى = أكبر قيمة – أصغر قيمة**.

3- نحسب عدد الفئات G عدد الفئات=$\sqrt{\text{عدد البيانات}}$ وهناك عدة طرق لحساب عدد الفئات نذكر منها:

أ – معادلة ستيرجس *Sturges* التي تنص على أن عدد الفئات = 1 + 3.322 لغ عدد البيانات عندما يكون عدد الاجمالي للبيانات يساوي أو يفوق 1000.

ب – معادلة يول *yulle* التي تنص على عدد الفئات = 2.5 $\sqrt[4]{\text{عدد البيانات}}$ يكون عدد الاجمالي للبيانات يقل عن 1000.

4- نحسب طول الفئة ai وهو يساوي المدى مقسوما على عدد الفئات

$$\text{طول الفئة} = \frac{\text{المدى}}{\text{عدد الفئات}}$$

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

$$\boxed{\text{طول الفئة} \times \text{عدد الفئات} \geq \text{المدى}}$$

مثال2: تعبر السلسلة الإحصائية التالية عن أوزان 48 شخص

50	25	35	15	45	36	24	69	35	41	50	31	99	24
42	58	70	72	50	91	25	70	23	89	36	25	89	97
38	15	49	50	30	23	35	32	90	35	15	69	67	36
								38	25	77	79	69	62

المطلوب:

1- ماهي طبيعة المتغير الإحصائي المدروس؟ و لماذا؟

2- كون جدول التوزيع التكراري؟

3- أحسب التكرارات النسبية المئوية؟

4- أحسب التكرارات التجميعية الصاعدة و النازلة؟

الحل:

1) طبيعة المتغير: كمي مستمر لأنه يمكن أن يأخذ أي قيمة بين عددين صحيحين.

2) الجدول التكراري: مراحل بناء الجدول:

-1 حساب المدى: **المدى = أكبر قيمة – أصغر قيمة=e**

$e = 99-15=84$

-2حساب عدد الفئات:

G= √عدد المشاهدات=√48

$G= 6.928\approx 7$

-3طول الفئة: ai = $\frac{\text{المدى}}{\text{عدد الفئات}} = \frac{84}{7} = 12$

الوزن ci	xi	ni	fi	التكرار المتجمع الصاعد N_i	التكرار المتجمع النازل N_i
[15-27[	21	11	0.229	11	48
[27-39[	33	12	0.25	23	37
[39-51[	45	8	0.166	31	25
[51-63[	57	2	0.041	33	17
[63-75[	69	7	0.145	40	15
[75-87[	81	2	0.041	42	8
[87-99[	93	6	0.125	48	6
المجموع		**48**	**1**		

مركز الفئة: مركز الفئة (منتصف الفئة) و يرمز لها ب xi تحسب كالتالي:

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

- **التكرار النسبي**

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

$$\text{التكرارات النسبية}=\frac{\text{التكرارات المطلقة لكل فئة}}{\text{مجموع التكرارات}}=\frac{ni}{N}$$

3-3: جدول التوزيعات التكرارية المتجمعة:

إن التوزيعات التكرارية كما رأينا لحد الأن تظهر لنا فقط عدد مرات تكرار الفئة، و لمعرفة عدد التكرارات التي تقل أو تساوي أو تزيد عن حد معين من حدود الفئات، فإنه يتم اللجوء إلى جداول التوزيعات التكرارية المتجمعة(الصاعدة و النازلة) وكتعريف فإن:

– التكرار المتجمع **الصاعد** في حالة متغير مستمر أو متقطع هو تكرار هذه الفئة مضافا إليه مجموع تكرارات الفئات السابقة وهو يعبر عن عدد المفردات التي تقل قيمتها عن قيمة معينة و في حالة متغير مستمر فإنها تقل عن أقصى حد

– التكرار المتجمع **النازل** لأي فئة هو عبارة عن مجموع التكرارات مطروحا منه تكرارات الفئات السابقة وهو يعبر عن عدد المفردات التي تزيد قيمتها عن قيمة معينة و في حالة متغير مستمر فإنها تزيد عن أدنى حد.

مثال3:

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

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

الحل:

الفئة	التكرار n_i	التكرار المتجمع الصاعد $\overset{\nwarrow}{N_i}$	التكرار المتجمع النازل $\overset{\searrow}{N_i}$
]15 –10 [	4	4	40
]20 –15 [	6	10	36
]25 – 20]	12	22	30
] 30 – 25]	8	30	18
]30 – 35]	6	36	10
]35 – 40]	4	40	4
المجموع	**40**		

ملاحظات:

- التكرار المتجمع الصاعد للفئة الثالثة [20 – 25[= 22 ويعني أن عدد العمال الذي يقل دخلهم عن 25 ألف دينار يساوي 22 عامل أي 4 + 6 + 12.

- التكرار المتجمع النازل للفئة الرابعة [25 – 30 [= 18 ويعني أن عدد العمال الذي يساوي دخلهم أو يزيد عن 25 ألف دينار هو 18 عامل أي 40 – 4 – 6 – 12.

- نستعمل نفس الطريقة في حساب التكرار المتجمع الصاعد والتكرار المتجمع النازل في حالة الفئات غير متساوية.

- **العرض الجدولي في حالة البيانات النوعية**

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

مثال4:

أخذت عينة من طلبة جامعة بسكرة متكونة من 100 طالب، حيث كانت الخاصية المدروسة هي الانتماء إلى كلية من كليات الجامعة فكانت النتائج كما هي مبينة في الجدول أدناه

جدول توزيع مجموعة من طلبة بسكرة على كليات الجامعة

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

رابعا – العرض البياني للبيانات

بالإضافة إلى العرض الجدولي للبيانات هناك طريق أخرى تستخدم لتوضيح وتلخيص البيانات وهي طرق العرض البياني. وهذه الطرق أسهل وأبسط من سابقتها، فإذا كان العرض الجدولي لا يفهمه إلا المختصين فإن طرق العرض البياني يمكن أن تجلب غير المتخصصين، كما تمكن من القيام بتحليل سريع للظاهرة المدروسة. وتستخدم أنواع مختلفة للعرض البياني حسب نوعية المتغير المدروس.

1 – المتغير الكمي **المتقطع**: وهنا يمكن استخدام الأعمدة البسيطة، العرض البياني للتكرارات المتجمعة الصاعدة والنازلة.

2 –المتغير الكمي **المستمر**: هنا يمكن استخدام المدرج التكراري، المضلع التكراري، المنحنى البياني للتكرارات المتجمعة الصاعدة والنازلة.

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

- **العرض البياني في حالة متغير كمي متقطع:**

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

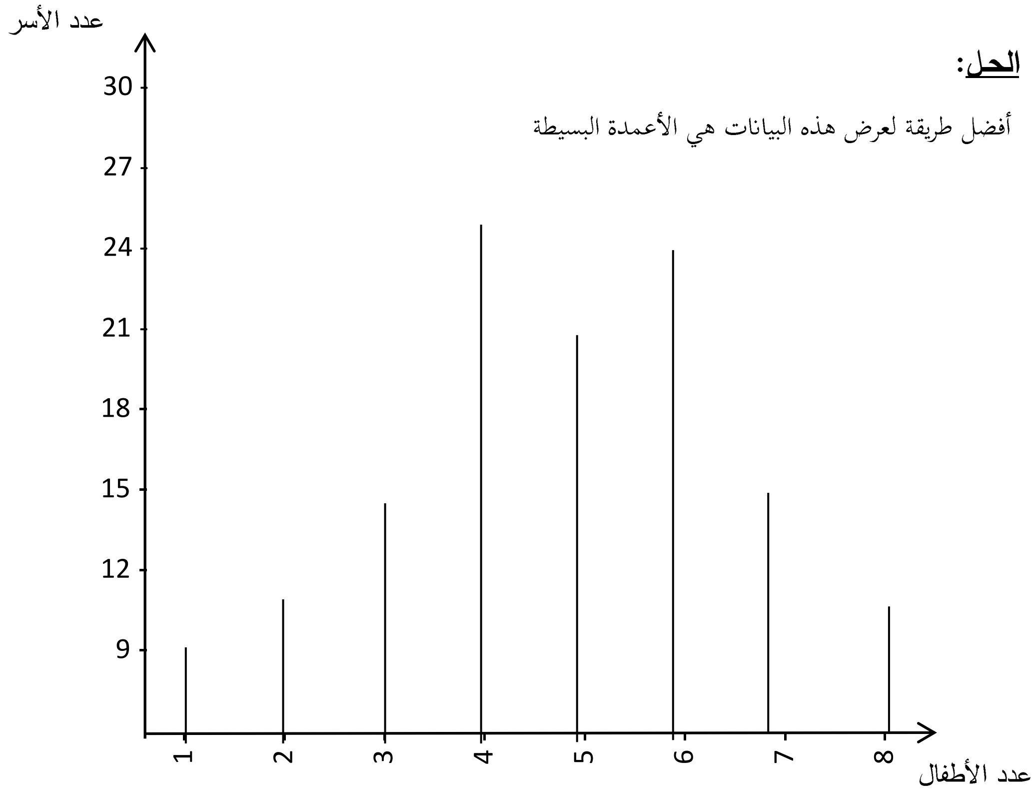
مثال 5:

يبين الجدول التالي عدد الأطفال في العالة لعينة تكون من 125 أسرة، المطلوب عرض هذه البيانات بطريقة العرض المناسب البسيطة.

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

<u>الحل</u>:

أفضل طريقة لعرض هذه البيانات هي الأعمدة البسيطة



نلاحظ بسهولة أن العمود الذي يقابل القيمة 4 هو أطول الأعمدة وتكراره يساوي 26 ويعني ذلك أن أغلب العائلات لها 4 أطفال.

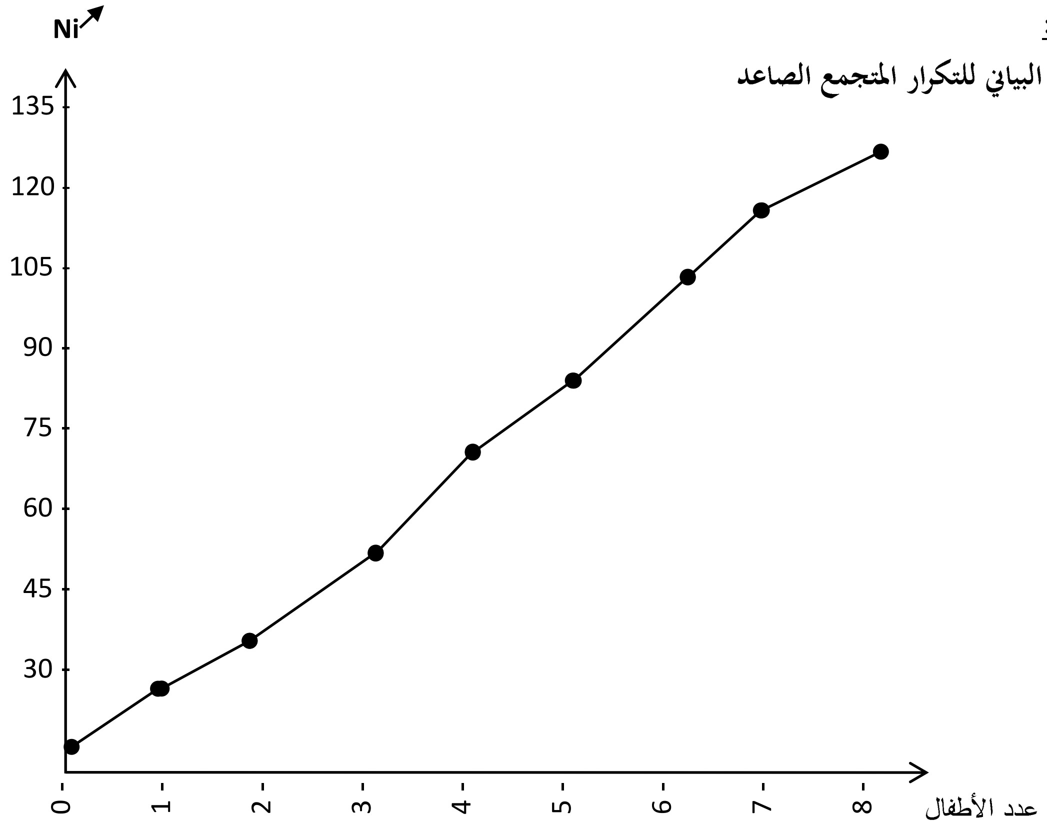
2 – العرض البياني للتكرار المتجمع الصاعد Effectifs cumulés croissants

مثال 6: مثل بيانات التكرار المتجمع الصاعد في المثال التالي؟.

$N_i \swarrow$	$N_i \nearrow$	N_i	X_i
125	6	6	0
119	15	9	1
110	25	10	2
100	39	14	3
86	55	16	4
70	75	20	5
50	100	25	6
25	115	15	7
10	125	10	8
		125	المجموع

الحل:

العرض البياني للتكرار المتجمع الصاعد

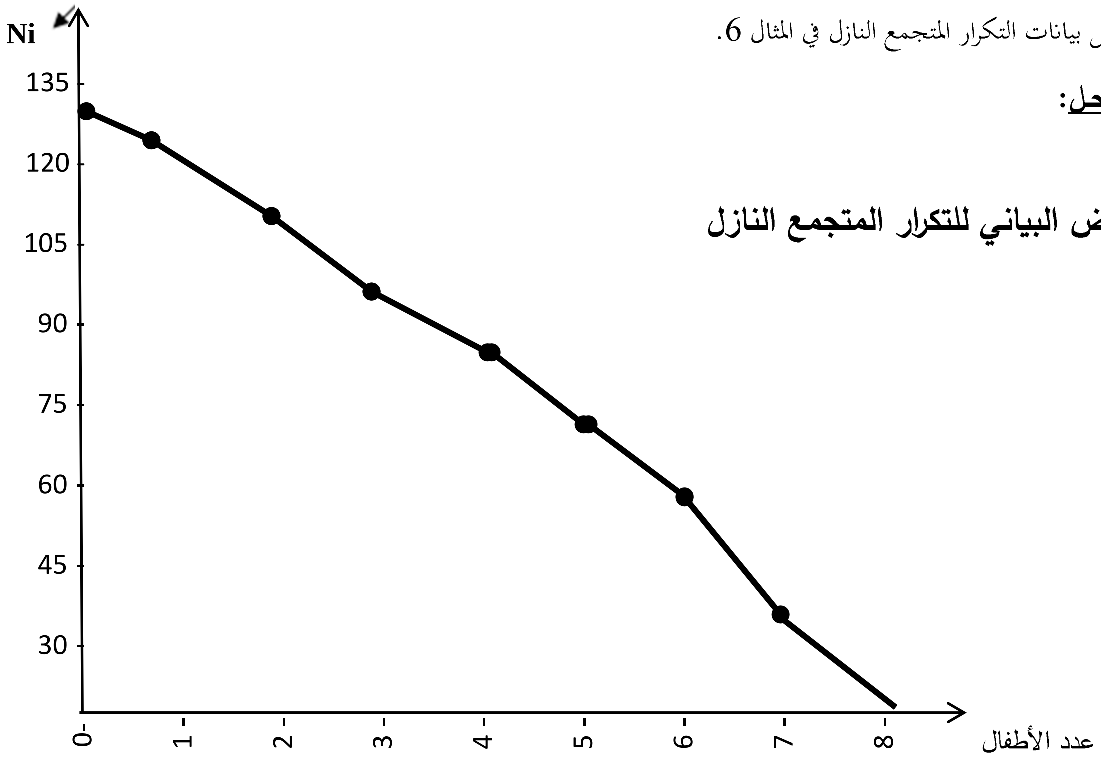


3 – العرض البياني للتكرار المتجمع النازل: Effectifs cumulé décroissants

مثل بيانات التكرار المتجمع النازل في المثال 6.

الحل:

العرض البياني للتكرار المتجمع النازل



– **العرض البياني في حالة متغير كمي متصل (مستمر):**

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

1 – المدرج التكراري Histogramme:

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

أ) **الحالة الأولى**: عندما تكون الفئات متساوية.

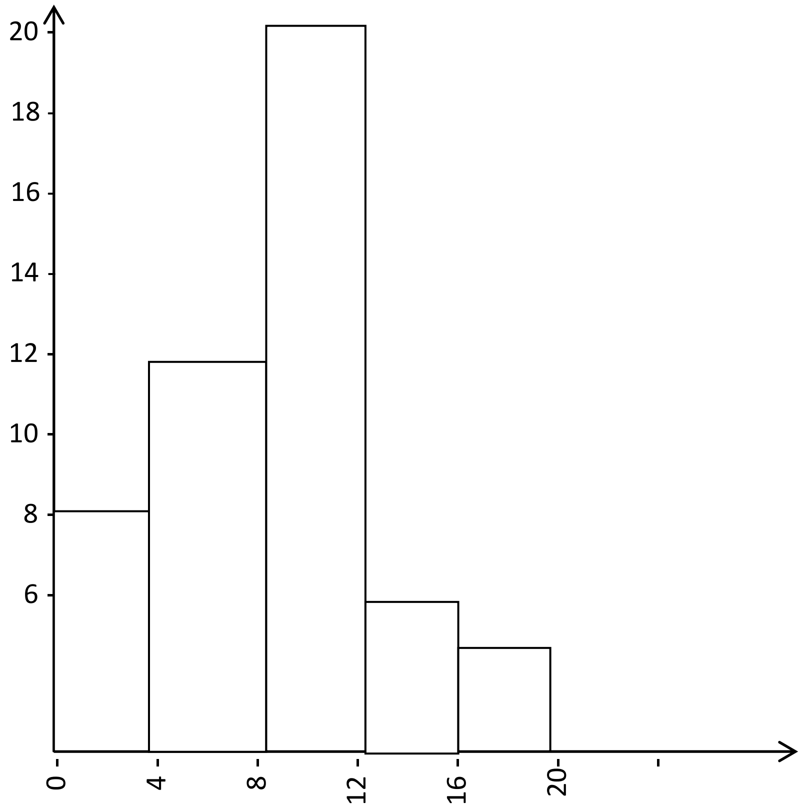
مثال07:

يبين الجدول التالي توزيع الدرجات التي حصل عليها 50 طالب في مادة الإحصاء.

أرسم المدرج التكراري الذي يمثل توزيع الدرجات؟.

الفئة	التكرار
4 – 0	8
8 – 4	12
12 – 8	20
16 – 12	6
20 – 16	4
المجموع	**50**

الحل:



ب) الحالة الثانية: عندما تكون الفئات غير متساوية.

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

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

مثال 08:

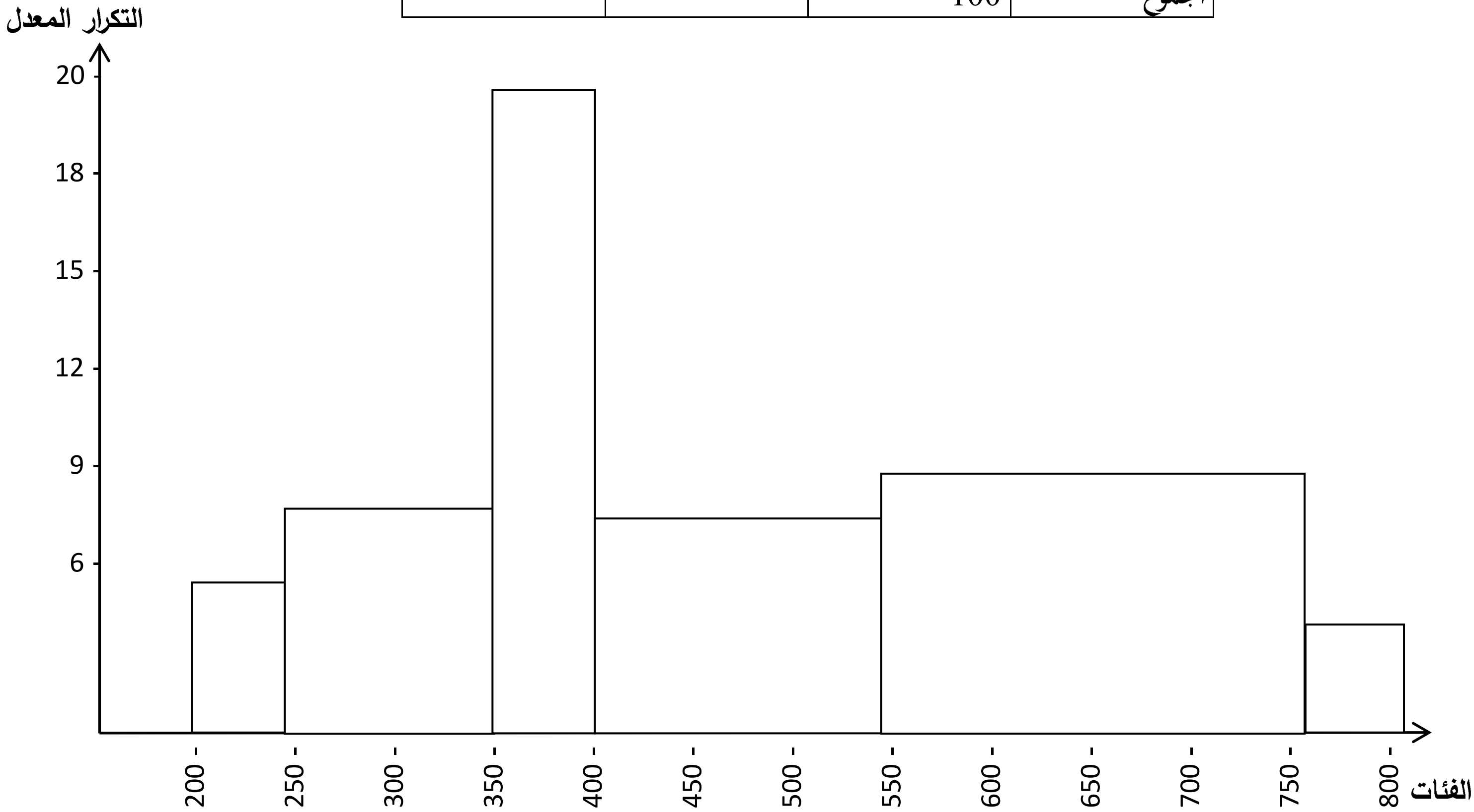
يبين الجدول التالي توزيع عينة من 100 عامل حسب الأجر اليومي المطلوب تمثيل هذه البيانات باستخدام المدرج التكراري؟.

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

الحل:

بما أن فئات التوزيع غير متساوية فإننا من أجل رسم المدرج التكراري نقوم بتعديل تكرار هذه الفئات وفقا للمعادلة السابقة.

فئة الأجر	عدد العمال	طول الفئة	التكرار المعدل
250 – 200	5	5	5
350 – 250	15	10	7.5
400 – 350	20	5	20
550 – 400	25	15	8.33
750 – 550	30	20	7.5
800 – 750	5	5	5
المجموع	100		



ملاحظة: نقوم بتعديل التكرارات في حالة الفئات غير المتساوية في الحالتين التاليتين:

- عند رسم المدرج التكراري.
- عند تحديد الفئة المنوالية وحساب المنوال.

2 – المضلع التكراري Polygone de fréquence:

هو مجموع من قطع مستقيمة متصلة ومنكسرة تتحدد بنقاط أحداثياتها مركز الفئة والتكرارات المقابلة.

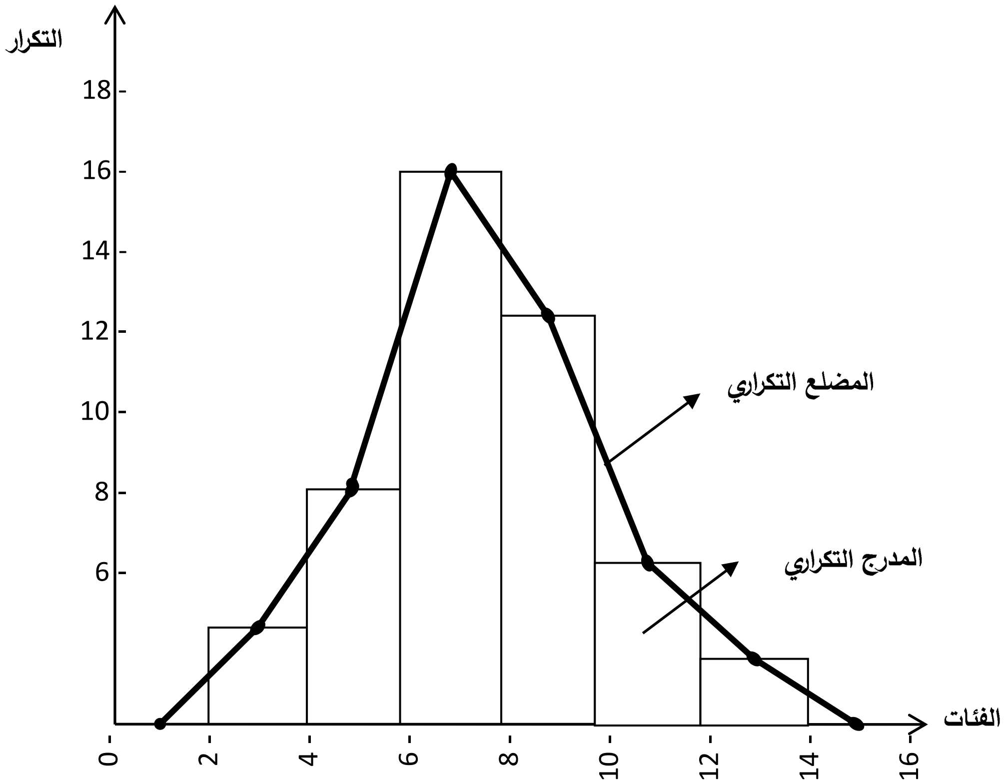
مثال 09:

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

الفئة	التكرار

4 – 2	4
6 – 4	8
8 – 6	16
10 – 8	12
12 – 10	6
14 – 12	2
المجموع	48

<u>**الحل**</u>



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

3– منحنى التكرارات المتجمعة الصاعدة والنازلة:

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

يبين كل من منحنى التكرار المتجمع الصاعد ومنحنى التكرار المتجمع النازل شدة أو ضعف تطور الظاهرة المدروسة عن مستوى معين من مجال الدراسة. إن فاصلة نقطة تقاطع المنحنيين تسمى بالوسيط.

مثال 10:

أرسم على نفس المعلم كل من منحنى التكرار المتجمع الصاعد ومنحنى التكرار المتجمع النازل لبيانات التكراري الآتي؟.

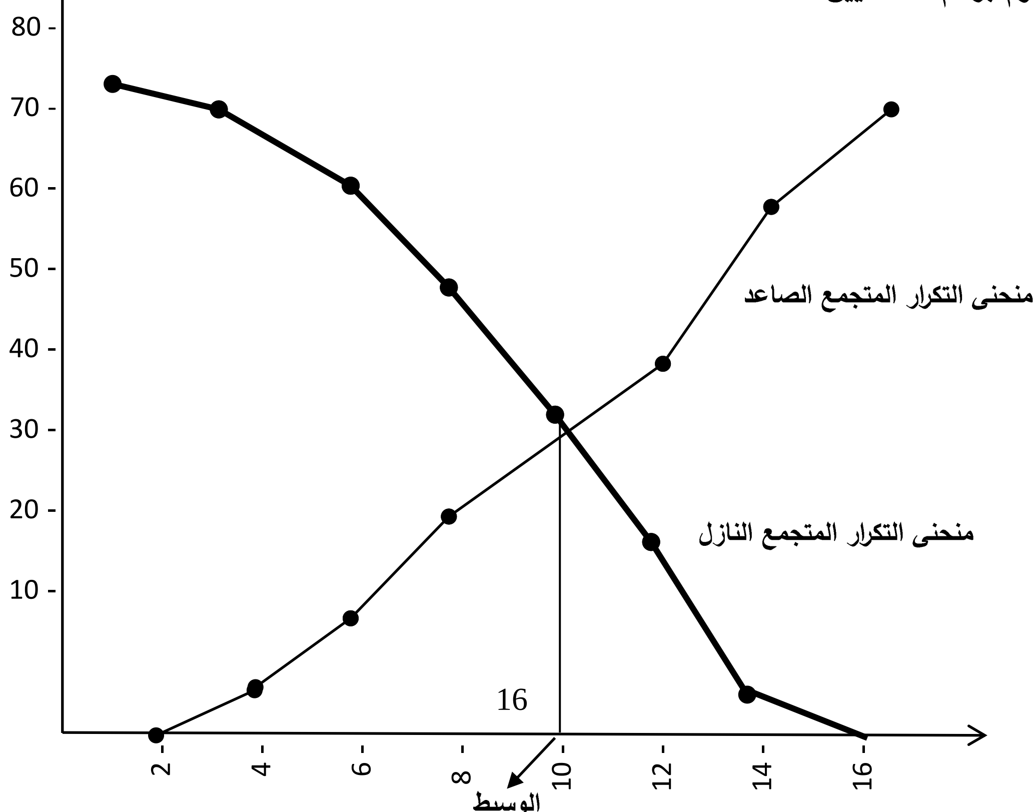
الفئة	التكرار
4 – 2	4
6 – 4	9
8 – 6	12
10 – 8	16
12 – 10	18
14 – 12	10
16 – 14	6
المجموع	75

الحل:

أولا نحسب كل من التكرار المتجمع الصاعد والتكرار المتجمع النازل.

الفئة	ni	Ni ↗	Ni ↘
4 – 2	4	4	75
6 – 4	9	13	71
8 – 6	12	25	62
10 – 8	16	41	50
12 – 10	18	59	34
14 – 12	10	69	16
16 – 14	6	75	6
المجموع	75		

ثانيا: نقوم برسم المنحنيين:



– العرض البياني في حالة متغير نوعي:

1) العرض الدائري: Diagramme circulaire:

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

مثال 11:

يبين الجدول التالي عدد طلبة كلية الحقوق والعلوم الاقتصادية سنة 2004 مقسمين على أقسام الكلية المختلفة.

القسم	الحقوق	الاقتصاد	التسيير	علوم سياسية	إعلام آلي تسيير	**المجموع**
عدد الطلبة	1200	1000	800	600	400	**4000**

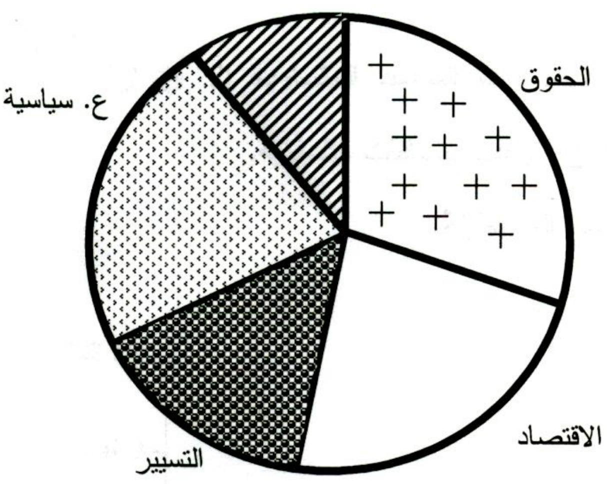
المطلوب: عرض البيانات باستخدام القطع الدائرية؟

الحل:

أولا: نحسب الزاوية المركزية.

القسم	**عدد الطلبة**	**الزاوية المركزية**
الحقوق	1200	°108
الاقتصاد	1000	°90
التسيير	800	°72
ع.سياسية	600	°54
إ. آلي تسيير	400	°36
المجموع	4000	°360

ثانيا: نرسم الدائرة.



ملاحظة: حسبت الزوايا المركزية بالطريقة التالية:

$$\frac{\text{تكرار الخاصية}}{\text{مجموع التكرارات}} \times 360°$$

2) العمود المجزأ Diagrammes en Barres:

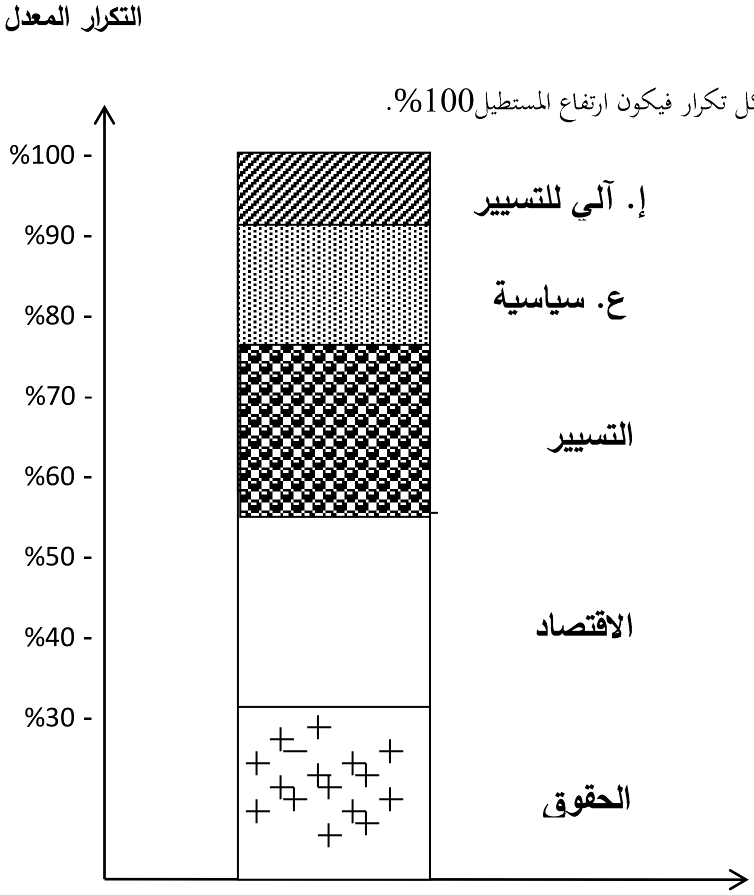
وهو عبارة عن مستطيل مقسم إلى عدة أجزاء كل جزء يقابل تكرار معين للخاصية المدروسة.

مثال 12:

أعرض بيانات المثال السابق باستخدام العمود المجزأ.

الحل:

لرسم هذا العمود نقوم بحساب النسبة المئوية المقابلة لكل تكرار فيكون ارتفاع المستطيل100%.



القسم	عدد الطلبة
الحقوق	30
الاقتصاد	25
التسيير	20
ع.سياسية	15
إ. آلي تسيير	10
المجموع	%100

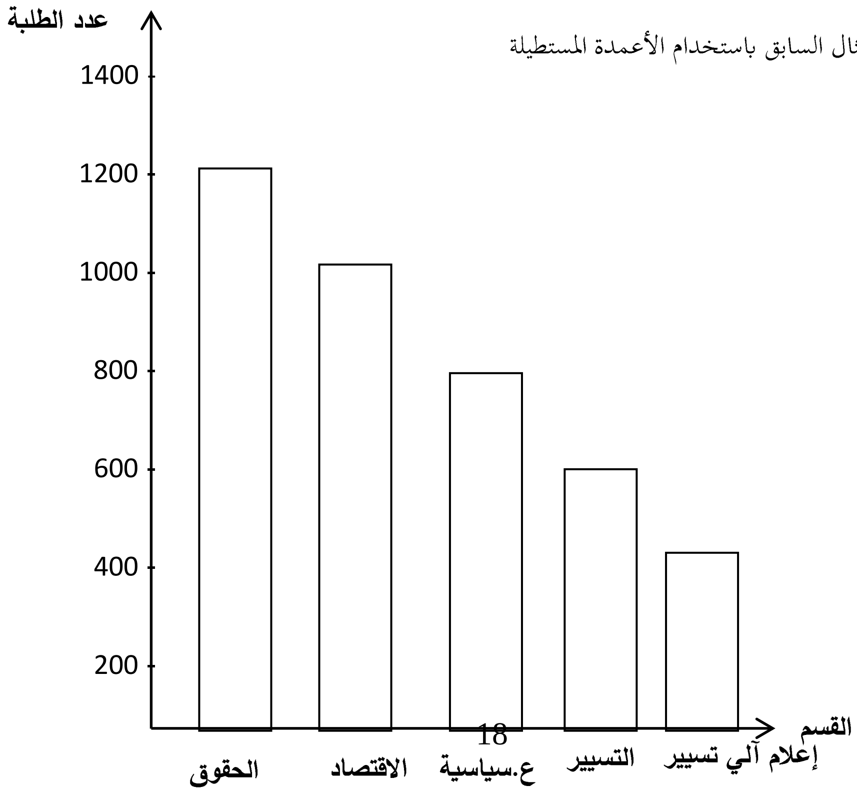
3 الأعمدة المستطيلة:

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

مثال 13:

أعرض بيانات المثال السابق باستخدام الأعمدة المستطيلة

الحل:



القائمة01: عرض البيانات الإحصائية و وصفها

تمرين 01:

من الأمثلة التالية، حدد كلا من المجتمع (أو العينة)، و مفرداته، المتغير الإحصائي، و تحديد نوعه و طبيعته.

1. في شركة معينة يتم مراقبة عدد حوادث العمل اليومية، التي وقعت للعمال خلال شهر نوفمبر **2020**.
2. مراقبة عدد الغيابات اليومية المسجلة من طرف عمال مصنع خلال شهر رمضان.
3. تصنيف دفعة من الطلبة حسب المعدل المحصل عليها في المواد العلمية.
4. نهتم بجنسيات المشاركين في منتدى دولي يتمحور موضوعه حول المقاولة.
5. مراقبة عدد الطلبة الذين يصلون متأخرين إلى حصة العمل الموجه لمقياس الإحصاء الوصفي خلال السداسي الأول.
6. مراقبة سعر بيع سيارة معينة في معرض تجاري يومياً لمدة **15** يوماً.
7. بحوزتنا **100** طفل عمرهم **6** سنوات نرغب في حساب أوزانهم ثم أطوالهم.
8. تسجيل درجة حرارة الجو لكل يوم في مدينة معينة خلال فصل الشتاء.
9. في مستشفى يحدد طبيب كمية الكولسترول في الدم لـ **200** مريض.

تمرين 02:

البيانات التالية تعبر عن عدد الغيابات عن العمل المسجلة من طرف موظفي شركة معينة خلال شهر جوان.

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

1. ما نوع و طبيعة المتغير المدروس ؟
2. ضع هذه البيانات داخل جدول تكراري.
3. مثل هذه الظاهرة بيانياً (التمثيل البياني الملائم).
4. ما هو عدد الأيام التي فيها تسجيل على الأكثر **غيابين** ؟ ثم حدد نسبتهم.
5. ما هو عدد الأيام التي تغيب فيها على الأقل **أربع** موظفين ؟ ثم حدد نسبتهم.

تمرين 03:

المعطيات التالية تتعلق بالإنتاج الفردي بالوحدات لـ **27** عامل.

142	136	142	125	135	123	138	126	125
128	130	131	141	135	136	136	141	120
135	128	122	135	132	143	133	121	142

1. أنقل هذه البيانات في جدول تكراري ذو طبيعة مستمرة.
2. مثل هذه الظاهرة بيانياً.
3. ما هو عدد العمال الذين قاموا بإنتاج أقل من **135** وحدة، ثم على الأقل **135** وحدة ؟
4. ما هي نسبة العمال الذين أنتجوا ما لا يقل عن **120** وحدة.
5. أرسم المدرج و المضلع التكراريين الذين يوضحون تطور هذه الظاهرة.
6. أرسم المضلع التكراري الصاعد ثم النازل على نفس المعلم.

تمرين 04:

يوضح الجدول التالي توزيع مجموعة من الطلبة حسب المستوى الاجتماعي و الاقتصادي.

1. حدد المجتمع الاحصائي.
2. حدد الصفة المدروسة و ما طبيعتها.
3. احسب التكرار النسبي.
4. ضع التمثيل البياني الملائم.

الفئة الاجتماعية و الاقتصادية	بسيطة جدا	بسيطة	متوسطة	راقية	راقية جدا
عدد الطلبة	43	76	51	12	04

تصحيح التمرين 01:

1. **المجتمع**: شهر نوفمبر **2013**. **مفرداته:** الأيام. **المتغير**: عدد حوادث العمل. **نوع المتغير**: كمي. **تكرار**: عدد الأيام.
2. **المجتمع**: شهر رمضان. **مفرداته:** أيام الشهر. **المتغير**: عدد الغيابات. **نوع المتغير**: كمي. **تكرار**: عدد الأيام.
3. **المجتمع**: دفعة الطلبة. **مفرداته:** الطلبة. **المتغير**: المعدل الذي تحصلوا عليه. **نوع المتغير**: كمي. **تكرار**: عدد الطلبة.
4. **المجتمع**: كل المشاركين في الملتقى. **مفرداته:** المشاركين. **المتغير**: جنسيات المشاركين. **نوع المتغير**: نوعي. **تكرار**: عدد المشاركين.
5. **المجتمع**: السداسي الأول. **مفرداته:** حصة المقياس. **المتغير**: الطلبة المتأخرين. **نوع المتغير**: كمي. **تكرار**: عدد الحصص.
6. **المجتمع**: **15** يوماً. **مفرداته:** أيام المعرض التجاري. **المتغير**: أسعار بيع السيارات. **نوع المتغير**: كمي. **تكرار**: عدد الأيام.
7. **المجتمع**: **100** طفل عمرهم **6** سنوات. **مفرداته:** الأطفال. **المتغير**: الأوزان، ثم الأطوال. **نوع المتغير**: كمي. **تكرار**: عدد الأطفال.
8. **المجتمع**: فصل الشتاء. **مفرداته:** الأيام. **المتغير**: درجة الحرارة. **نوع المتغير**: كمي. **تكرار**: عدد الأيام.
9. **المجتمع**: **200** مريض. **مفرداته:** المرضى. **المتغير**: كمية الكولسترول. **نوع المتغير**: كمي (مستمر). **تكرار**: عدد المرضى.

الفصل الثاني

مقاييس النزعة المركزية

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

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

أولا – المتوسط الحسابي Moyenne arithmétique

يعتبر المتوسط الحسابي من أسهل وأكثر متوسطات النزعة المركزية استخداما في الإحصاء، هو عبارة عن مجموع القيم مقسوما على عددها، فإذا كانت لدينا القيم $x_1,x_2,x_3,\dots\dots\dots x_n$ فإن متوسطها الحسابي يساوي

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

حيث: $\overline{X}$ = المتوسط الحسابي

X_i = تمثل قيم الظاهرة

n = تمثل عدد البيانات

نستعمل هذه العلاقة عندما يكون لدينا نفس العامل و بدون تكرار.

مثال1:

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

أحسب متوسط درجات هذا الطالب؟.

الحل:

$$\text{المتوسط} = \frac{\text{مجموع الدرجات}}{\text{عدد الدرجات}}$$

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

1– المتوسط الحسابي في حالة بيانات مبوبة (طبيعة متقطعة)

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

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

فإن المتوسط الحساب لها يعطي بالعلاقة

$$\overline{X} = \frac{n_1x_1 + n_2x_2 + n_3x_3 + ...n_nx_n}{n_1 + n_2 + n_3 + n_4 + ...n_n} = \frac{\sum_{i=1}^{n} n_ix_i}{\sum n_i}$$

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

وحيث: X_i = تمثل قيم الظاهرة.

n_i = تكرار مل قيمة.

Σn_i = مجموع التكرارات.

مثال2:

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

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

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

الحل:

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

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

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

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

2- المتوسط الحسابي في حالة طبيعة مستمرة:

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

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

حيث: x_i = تمثل مراكز الفئات.

n_i = تكرار الفئات.

Σn_i = مجموع التكرارات.

مثال3:

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

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

المطلوب: إيجاد متوسط إنتاج الحليب بهذه المزارع

الحل:

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

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

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

3- حساب المتوسط الحسابي بالطريقة المختصرة:

عندما تكون البيانات الموجودة في الجداول المعطاة كبيرة القيم فإن استخدام الطريقة السابقة يصبح صعب، كما يزداد احتمال الوقوع في الأخطاء بذلك فإنه في مثل هذه الحالات يفضل استخدام طرقة مختصرة الهدف منها تبسيط العمليات الحسابية الطويلة حتى يسهل التعامل معها.

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

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

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

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

مثال4:

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

الحل:

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

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

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

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

$\overline{X}$ =300 - 12.8= لتر 287.2

4- المتوسط الحسابي المرجح (الموزون) Moyenne Arithmétique pondérée

يستخدم المتوسط الحسابي المرجح لإيجاد المتوسط الحسابي لمجموع البيانات أو أكثر في حالة دمجهم معا في مجموعة واحدة وبالتالي فإن متوسط الحسابي المرجح لمجموع البيانات يساوي:

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

مثال5:

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

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

الحل:

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

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

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

5- خواص المتوسط الحسابي:

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

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

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

وللتأكد من ذلك نقوم بتفكيك الطرف الأيمن:

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

(لأن : $n\overline{X} = ...\overline{X} + \overline{X} + \overline{X}$)

نضرب الحد الأول في n ونقسمه على n

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

4 – يتأثر المتوسط الحسابي بالقيم المتطرفة أو الشاذة.

مثال6:

إذا تحصل طالب على الدرجات التالية في خمس مواد: 20.9.8.7.6

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

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

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

6- لا يمكن أن يكون للتوزيع التكراري أكثر من وسط حسابي واحد.

7- إذا أضيف لكل قيمة من قيم الظاهرة عدد ثابت a فإن الوسط الحسابي للقيم الجديدة يزداد بنفس المقدارa

$$y_i = x_i - a$$
$$\bar{y} = \frac{\sum y_i}{N} = \frac{\sum(xi+a)}{N} = \frac{1}{N}(\sum xi + \sum a) = \frac{\sum xi}{N} + \frac{Na}{N} = \bar{x} + a$$

8- إذا ضربنا كل قيمة من قيم الظاهرة عدد ثابت k فإن الوسط الحسابي للقيم الجديدة يضرب أيضا في k

$$z_i = x_i - a$$
$$\bar{z} = \frac{\sum z_i}{N} = \frac{\sum kxi}{N} = k\frac{\sum xi}{N} = k\bar{x}$$

ثانيا– الوسيط Médiane

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

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

مثال07:

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

الحل:

لإيجاد الوسيط نقوم بالتالي:

أ) نرتب البيانات تصاعديا أو تنازليا: 6، 8، 8، 9، 10، 12، 15.

ب) نبحث عن الوسيط لهذه البيانات، وهناك حاليتن:

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

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

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

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

مثال08:

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

16، 17، 17، 15، 14، 16، 15، 13، 4، 3.

المطلوب:أ) إيجاد متوسط درجات هؤلاء الطلبة؟.

ب) إيجاد وسيط الدرجات؟

ج) أيهما أكثر تعبيرا على نتائج الطلبة ولماذا؟.

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

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

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

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

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

الوسيط في حالة بيانات مبوبة:

1- الوسيط لبيانات مبوبة ذو طبيعة متقطعة:

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

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

مثال09:

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

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

1) نحسب التكرار المتجمع الصاعد للبيانات المعطاة:

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

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

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

2- الوسيط لبيانات مبوبة ذو طبيعة مستمرة:

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

1) نحسب التكرار المتجمع الصاعد.

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

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

4) نحدد ونحسب الوسيط بتطبيق العلاقة الإحصائية للوسيط.

$$Me = L_1 + \frac{\frac{N}{2} - N\uparrow_{i-1}}{n_e}.ai$$

استنتاج العلاقة التي تحسب الوسيط:

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

مثال10:

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

المطلوب:1) استنتج المعادلة التي تحسب الوسيط؟.

2) أحسب قيمة الوسيط لهذه البيانات؟.

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

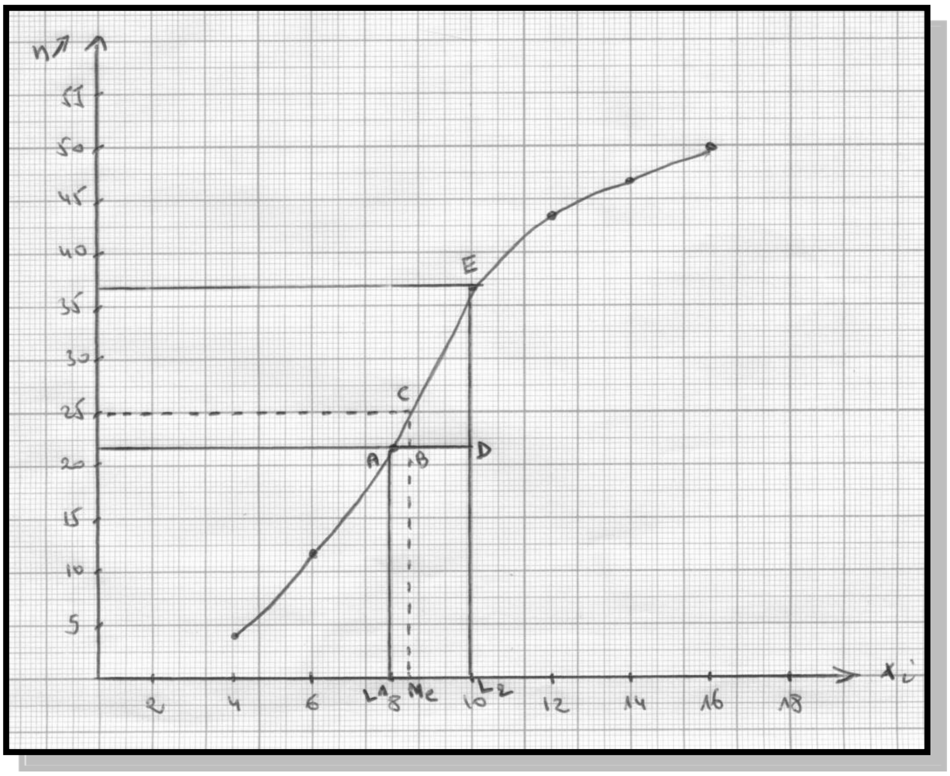
الحل:

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

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

ج) نرسم منحنى التكرار المتجمع الصاعد لبيانات التوزيع.

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



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

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

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

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

ي) نستخرج علاقة الوسيط كما يلي:

المثلثان ABC و ADE متشابهان، ومنه وحسب نظرية طاليس فإنه

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

وبالتعويض:

$$\frac{M_e - L_1}{Rg - N\uparrow_{i-1}} = \frac{L_2 - L_1}{N\uparrow_i - N\uparrow_{i-1}}$$

حيث: M_e = الوسيط.

L_1 = الحد الأدنى للفئة الوسيطية.

N = مجموع التكرارات.

N_0 = التكرار المتجمع الصاعد للفئة قبل الفئة الوسيطية.

L_2 = الحد الأعلى للفئة الوسيطية.

N_1 = التكرار المتجمع الصاعد للفئة الوسطية.

$$M_e - L_1 = \frac{(Rg - N\uparrow_{i-1})(L_2 - L_1)}{N\uparrow_i - N\uparrow_{i-1}}$$

$$M_e = L_1 + \frac{Rg - N\uparrow_{i-1}}{n_i} ai$$

وذلك لأن $(L_2-L_1) = ai$ = طول الفئة و $ni = N_i - N_{i-1}$ = تكرار الفئة الوسيطية

2 – نحسب الآن قيمة الوسيط باستخدام العلاقة السابقة

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

M_e = 8,4

خواص الوسيط:

يتصف الوسيط بعدة خصائص أهمها:

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

2 – يمكن إيجاد الوسيط من الرسم.

3 – يمكن استخدامه في حالة البيانات النوعية.

4 - يمكن حسابه من الجداول التكرارية المفتوحة.

ثالثا – أشباه الوسيط

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

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

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

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

-الربيعات Les quartils:

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

الربيع الأول Q1: ويسمى كذلك بالربيع الأدنى وهو القيمة التي تقسم مجموع البيانات إلى قسمين بحيث ربع عدد البيانات أقل منه وثلاثة أرباع البيانات أكبر منه وترتيب الربيع الأول هو $\frac{N}{4}$ ويحسب كالتالي:

$$Q_1 = L_1 + \frac{\frac{N}{4} - N\uparrow_{i-1}}{n_{Q1}}.ai$$

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

ويحسب كالتالي:

$$\mathrm{Q}_3 = \mathrm{L}_1 + \frac{\frac{3\mathrm{N}}{4} - \mathrm{N}\uparrow_{\mathrm{i-1}}}{\mathrm{nQ}_3}.\mathrm{ai}$$

وواضح أن الربيع الثاني هو نفسه الوسيط

$$Q_2 = L_1 + \frac{\frac{2N}{4} - N\uparrow_{i-1}}{nQ2}.ai = M_e$$

-العشريات Les Déciles:

العشير الأول D_1: هو القيمة التي تقسم مجموع عدد البيانات إلى قسمين بحيث عشر عدد البيانات اقل منه وتسعة أعشار عدد البيانات أكبر منه وترتيبه هو $\frac{N}{10}$ ونحسب بالتالي: $D_1 = L_1 + \frac{\frac{N}{10} - N\uparrow_{i-1}}{n_{D1}}.ai$

-المئينات: Les centiles :

إذا قسمت البيانات إلى مائة قسم متساوي فإن نقاط التقسيم هذه تسمى المئينات. **فالمئين الأول C_1** هو القيمة التي يسبقها 1% من البيانات ويليها 99% من البيانات ويحسب كالتالي: $C_1 = L_1 + \frac{\frac{N}{100} - N\uparrow_{i-1}}{n_{C1}}.ai$

والمئين العشرين C_{20} يحسب كالتالي: $C_{20} = L_1 + \frac{\frac{20N}{100} - N\uparrow_{i-1}}{n_{C20}}.ai$

مثال11:

أحسب الربيع الأول والعشير الثالث والمئين الستين للبيانات التالية التي توضح درجات 50 طالب في مادة ما؟

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

الحل:

الدرجة	عدد الطلبة	التكرار المتجمع الصاعد
4-2	4	4
6-4	8	12
8-6	10	22
10-8	15	37
12-10	6	43
14-12	4	47

50	3	16-14
	50	**المجموع**

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

$$D_3 = L_1 + \frac{\frac{3N}{10} - N\uparrow_{i-1}}{n_{D3}}.ai = 6 + \frac{15-12}{10} \times 2 = 6,6$$

$$C_{60} = L_1 + \frac{\frac{60N}{100} - N\uparrow_{i-1}}{n_{C60}}.ai = 8 + \frac{30-22}{15} \times 2 = 9,067$$

رابعا: المنوال Mode

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

مثال12:

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

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

الحل:

المنوال هنا هو المقاس 41.

مثال 13 :

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

المطلوب: إيجاد المنوال

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

الحل:

المنوال هنا هو جيد.

- **المنوال لبيانات مبوبة في جداول توزيع تكراري:**

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

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

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

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

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

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

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

مثال14 :

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

1) تحديد قيمة المنوال؟

2) قارن بين قيمة المنوال والمتوسط الحسابي والوسيط؟.

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

الحل:

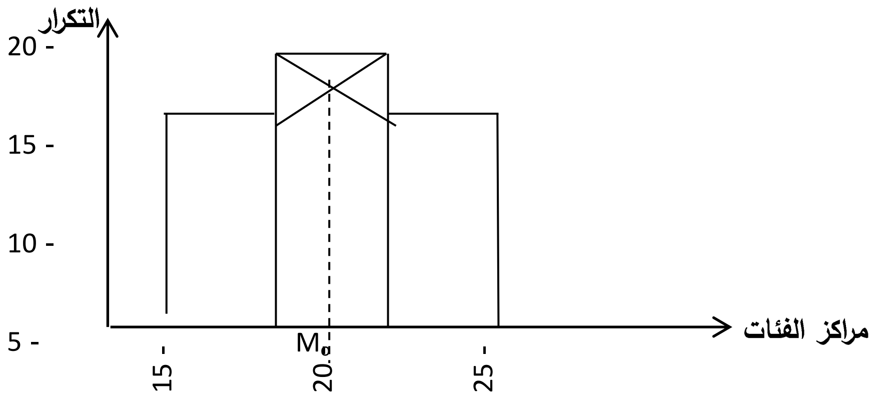
الأجر بالآلاف	عدد العمال	X_i	$n_i x_i$	$N_i \nearrow$
15-10	10	12.5	125	10

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

20-15	15	17.5	262.5	25
25-20	20	22.5	450	45
30-25	15	27.5	412.5	60
35-30	10	32.5	325	70
40-35	5	37.5	187.5	75
المجموع	**75**		**1762.5**	

1) الفئة المنوالية هي الفئة [20-25] وبالتالي فإنه يمكن اعتبار مركز هذه الفئة (22.5) كمنوال تقريبي.

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



- نحسب المنوال من خلال العلاقة المتوصل إليها.

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

2) لمقارنة بين $\overline{X}$,M_e ,M_0 نحسب كل من المتوسط الحسابي والوسيط

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

$$M_e = L_1 + \frac{\frac{N}{2} - N\uparrow_{i-1}}{n_2}.ai = 20 + \frac{37,5-25}{20}.5 = 32,125$$

حيث أن الفئة الوسيطية هي نفسها الفئة المنوالية

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

<u>خواص المنوال:</u>

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

2 – يمكن حسابه بيانيا.

3 – يمكن أن يوجد أكثر من منوال لتوزيع واحد.

4 – يمكن حساب من الجداول الإحصائية المفتوحة.

5 – يعتبر أفضل المتوسطات لوصف الظواهر النوعية.

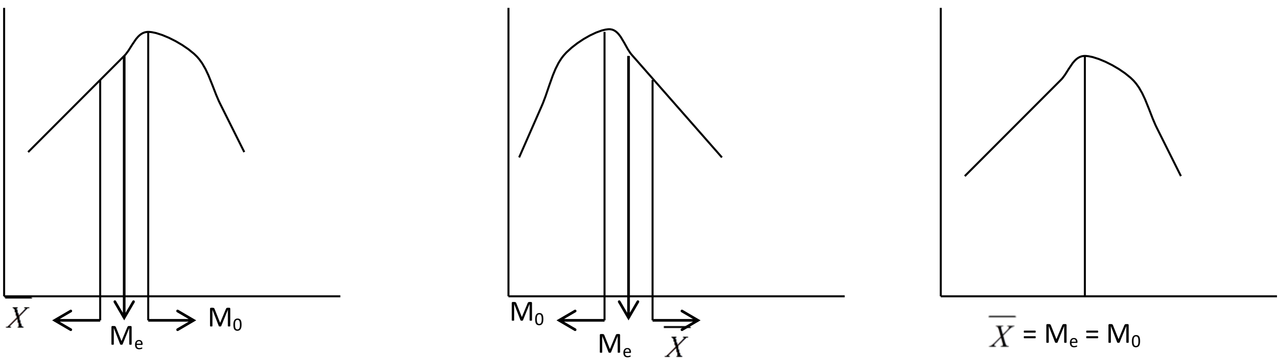
العلاقة بين المتوسط الحسابي والوسيط والمنوال

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

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

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

الحالة الأولى **الحالة الثانية** **الحالة الثالثة**



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

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

$$\overline{X} - M_0 = 3(\overline{X} - Me)$$

تمارين حول المتوسط الحسابي

تمرين01:

- إذا كان أجر عمل الساعة الواحدة بالنسبة لـ **12** عامل على الترتيب هو كالآتي: **14 – 15 – 18 – 16 – 18 – 15 – 10 – 11 – 12 – 9 – 8 – 8**. فما هو متوسط أجر ساعة العمل ؟
- إذا كانت عدد ساعات العمل لهؤلاء العمال في اليوم على الترتيب: **7 – 6 – 6 – 7 – 6 – 6 – 7 – 6 – 6 – 7 – 7 – 8**. فما هو متوسط عدد ساعات العمل لهؤلاء العمال ؟، و كم يصبح متوسط أجر ساعة العمل في هذه الحالة ؟

تمرين02:

تتألف مدرسة من 12 القاعة مربعة الشكل، طول ضلع خمس قاعات منها هو 8 أمتار، ثم أربع قاعات طول ضلع كل واحدة منها هو 9 أمتار، أما الثلاث قاعات الأخيرة هو 10 أمتار.

1 – أحسب متوسط طول ضلع قاعات هذه المدرسة.

2 – ما هي المساحة المتوسطة لقاعات هذه المدرسة ؟

تمرين04:

ينقسم مصنع إلى قسمين، يضم القسم الأول 200 عامل، أما القسم الثاني 270 عامل. خلال أسبوع معين؛ كان متوسط عدد ساعات العمل لعمال القسم الأول 40 سا، و القسم الثاني 43 سا.

المطلوب: ما هو متوسط عدد ساعات العمل لعمال هذا المصنع.

تمرين05:

في مؤسسة ما كان المتوسط العام لعدد ساعات العمل الأسبوعية لعمال المؤسسة هو 36.5 سا. إذا كان متوسط ساعات العمل للذكور هو 38 سا و الإناث هو 36 سا؛ فما هي نسبة الذكور و نسبة الإناث في هذه المؤسسة ؟

تمرين06:

إن المتوسط العام لأجور العمال في مؤسسة معينة هو 7000 دج، فإذا علمت أن متوسط أجور العمال ذكور هو 8000 دج و متوسط أجور العمال إناث هو 6500 دج، و أن مجموع عمال هذه المؤسسة هو 270 .

– فما هو عدد العمال ذكور، و عدد العمال إناث ؟

الفصل الثالث

مقاييس التشتت

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

الطالب (X): 10،11،13،14،15.

الطالب (Y):8،9،13،15،18.

فمتوسط درجات الطالب (X) يساوي 12،6 وكذلك متوسط درجات الطالب (Y) يساوي 12،6 ووسيط درجات الطالب (X) يساوي 13 وكذلك وسيط درجات الطالب (Y) يساوي 13.

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

ما معنى التشتت؟.

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

أولا – المدى العام (المطلق)

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

المدى= أكبر قيمة – أصغر قيمة.

أما المدى **للتوزيعات التكرارية** فيحسب بعدة طرق منها:

المدى = مركز الفئة الأخيرة – مركز الفئة الأولى

المدى = الحد الأعلى لفئة الأخيرة – الحد الأدنى للفئة الأولى

<u>مثال1</u>:

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

الحل: المدى = أكبر قيمة – أصغر قيمة

= 30-12=22.

مثال2:

أوجد المدى للبيانات التالية 65،20،17، -04،18،14.19

الحل:

المدى = 65 −(−4) = 69.

نلاحظ المدى في هذا المثال قد تأثر بشكل كبير جدا بالقيم المتطرقة، إذ نلاحظ أن معظم البيانات متقاربة باستثناء القيمة 65 والقيمة (−4)، فإذا استبعدنا هذه القيم المتطرفة فإن المدى يصبح

$$E = X_{\max} - X_{\min}$$ E = 20-14=6.

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

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

خواص المدى:

1 – يتصف المدى بسهولة حسابه.

2 – يعتمد في حساب على قيمتين فقط هما القيمة الكبرى والقيمة الصغرى.

3– بسبب الخاصية الثانية فإن المدى شديد التأثر بالقيم المتطرفة.

ثانيا – مقاييس التشتت حول الوسيط

* المدى الربيعي = الربيع الثالث – الربيع الأول. $IQ = Q_3 - Q_1$ و يضم 50% من الوحدات

* نصف المدى الربيعي = الربيع الثالث – الربيع الأول/2 $IQ/2 = {Q_3 - Q_1}/{2}$

* المدى العشري = العشر التاسع – العشر الأول. $ED = D_9 - D_1$

* المدى المئيني = المئين 99 – المئين الأول. $EC = C_{99} - C_1$

* نسبة التشتت (مقياس التشتت النسبي حول الوسيط)= نسبة المدى الربيعي على المدى العام $R = {Q_3 - Q_1}/{E} . 100$

فإذا كان:

R>50%: التوزيع قوي التشتت

R<50%: التوزيع قليل التشتت

R=50%: التوزيع متناظر تماما

* المدى الربيعي النسبي = $CQ = {Q_3 - Q_1}/{Me}$ (%50)

* المدى العشيري النسبي = $CD = {D_9 - D_1}/{Me}$ (%80)

ثالثا – مقاييس التشتت حول المتوسط الحسابي

1- الانحراف المتوسط L'écart moyen:

لأن مقاييس التشتت هي مقاييس لقوة تجمع البيانات حول بعضها، وحيث أن التجمع يكون حول القيم المتوسطة، فإنه إذا كان مقدار الاختلاف بين القيم ومتوسطها كبيرا دل ذلك على أن التشتت كبير والعكس صحيح.

وحيث أن مجموع الاختلافات (الانحرافات) عن المتوسط يساوي صفرا (*) فإنه لو حسبنا القيم المطلقة لمقدار الاختلاف عن المتوسط يكون متوسط هذه الاختلافات مقياسا مناسبا لمقدار التشتت، يسمى هذا المقياس بالانحراف المتوسط.

ويعرف الانحراف المتوسط بأنه المتوسط الحسابي للقيم المطلقة لانحرافات القيم عن متوسطها الحسابي، وسوف نرمز للانحراف المتوسط في دراستنا بالرمز E_x وعليه:

إذا كانت لدينا القيم التالية: $X_1, X_2, X_3, X_4 \ldots X_n$ فإن الانحراف المتوسط لها هو:

$$E_X = \frac{|X_1 - \overline{X}| + |X_2 - \overline{X}| + |X_3 - \overline{X}| + \ldots + |X_n - \overline{X}|}{n}$$

أو

$$E_X = \frac{\sum_{i=1}^{n} |X_i - \overline{X}|}{n}$$

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

$$E_X = \frac{\sum_{i=1}^{n} ni|X_i - \overline{X}|}{\sum_{i=1}^{n} ni}$$

مثال3:

أوجد الانحراف المتوسط للبيانات التالية: 2،4،5،6،8؟

الحل:

X_i	$\lvert X_I - \overline{X}\rvert$
2	3
4	1
5	0
6	1
8	3
	8

(*) أنظر خواص المتوسط الحسابي.

$$\overline{X}=\frac{\sum n_i}{n}=\frac{25}{5}=5$$

$$E_x=\frac{\sum\left|X_i-\overline{X}\right|}{n}=\frac{8}{5}=1.6$$

مثال4:

أوجد تشتت البيانات المبوبة في جدول التوزيع التكراري الآتي باستخدام الانحراف المتوسط؟

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

الحل:

الفئة	التكرار	X_i	n_iX_i	$\left\|X_I-\overline{X}\right\|$	$n_i\left\|X_i-\overline{X}\right\|$
0-2	2	1	2	3.33	6.66
4-2	3	3	9	1.33	4
6-4	4	5	20	0.67	2.68
8-6	3	7	21	2.67	8.01
المجموع	12		52		21.35

$$\overline{X}=\frac{\sum n_iX_i}{\sum n_i}=\frac{25}{12}=4.33$$

$$E_x=\frac{\sum n_i\left|X_i-\overline{X}\right|}{\sum n_i}=\frac{21.35}{12}=1.78$$

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

خواص الانحراف المتوسط:

1- يعتمد في حسابه على جميع القيم وليس على القيمة الكبرى والصغرى فقط.

2- لا يمكن حسابه في حالة التوزيعات التكرارية المفتوحة.

3- يتأثر بالقيم المتطرفة، لأن انحرافها عن المتوسط الحسابي يكون كبيرا.

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

أ – التباين La variance :

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

فإذا كانت لدينا البيانات التالية: $X_1,X_2,X_3,X_4 \ldots\ldots X_n$

فإن التباين لهذه البيانات يعطي بالعلاقة

$$V_X = \frac{(X_1-\overline{X})^2+(X_2-\overline{X})^2+(X_3-\overline{X})^2+........+(X_n-\overline{X})^2}{n}$$

$$V_X = \frac{\sum_{i=1}^{n} ni(X_1-\overline{X})^2}{\sum ni}$$

مثال5:

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

الحل:

$(X_i-\overline{X})^2$	$(X_i-\overline{X})$	X_i
9	-3	3
1	1	7
0	0	6
25	-5	1
25	5	11
1	-1	5
0	0	6
9	3	9
70		**48**

المتوسط الحسابي:

$$\overline{X} = \frac{\sum Xi}{n} = \frac{48}{8} = 6$$

التباين:

$$V_x = \frac{\sum(X_i-\overline{X})^2}{n} = \frac{70}{8} = 8.75$$

في بعض الأحيان عندما يكون المتوسط الحسابي للبيانات عبارة عن كسر، فإن عملية حساب التباين تكون مضنية وعرضة للأخطاء الحسابية لذلك فإنه تم تطوير طريقة مختصرة لحساب التباين.

طريقة مختصرة لحساب التباين:

انطلاقا من العلاقة المتوصل إليها سابقا:

$$V_x = \frac{\sum(X_i-\overline{X})^2}{N}$$

يمكن كتابة

$$V_x = \frac{\sum(X_i{}^2-2\overline{X}_i+\overline{X}^2)}{N}$$

$$V_X = \frac{\sum X_i^2}{N} - 2\overline{X}\frac{\sum X_i}{N} + \frac{N\overline{X}^2}{N}$$

$$V_X = \frac{\sum X_i^2}{N} - 2\overline{X}\,\overline{X} + \overline{X}^2$$

$$V_x = \frac{\sum X_i^2}{N} - 2\overline{X}^2 + \overline{X}^2$$

ومنه

$$\boxed{V_x = \frac{\sum X_i^2}{\sum ni} - \overline{X}^2} \quad (*)$$

أما في حالة البيانات المبوبة فإن العلاقة تصبح

$$\boxed{V_x = \frac{\sum n_i X_i^2}{\sum ni} - \overline{X}^2}$$

ب – الانحراف المعياري:

ويعتبر الانحراف المعياري من أهم المقاييس الإحصائية للتشتت، وهو أكثر استخداما في النظريات والقوانين الإحصائية، لأنه يعطي فكرة سليمة ومنطقية عن ظاهرة التشتت، ويعرف الانحراف المعياري بأنه الجذر التربيعي لمتوسط مجموع مربع انحراف القيم عن متوسطها، أي أنه الجذر التربيعي للتباين،سوف نرمز للانحراف المعياري في دراستنا بالرمز (δ_x).

الانحراف المعياري لبيانات مفردة

$$\delta x = \sqrt{\frac{\sum (X_i - \overline{X})^2}{n}} \quad (**)$$

الانحراف المعياري لبيانات متكررة أو مبوبة

$$\delta_x = \sqrt{\frac{\sum ni(X_i - \overline{X})^2}{\sum ni}}$$

أما الصيغة المختصرة للانحراف المعياري فتعطى بالعلاقات التالية:

(*) يفضل العلاقة المختصرة لسهولة الحسابات ولأنها تتعامل مع القيم الأصلية بدلا من انحرافاتها عن المتوسط الحسابي.

(**) في حالة العينات الصغيرة (n<30) فإنه يمكن استخدام العلاقة $\delta x = \sqrt{\frac{\sum (X_i - \overline{X})^2}{n-1}}$

الانحراف المعياري لبيانات مفردة
$$\delta_x = \sqrt{\frac{\sum X_i^2}{N} - \overline{X}^2}$$

الانحراف المعياري لبيانات متكررة أو مبوبة
$$\delta_x = \sqrt{\frac{\sum niX_i^2}{\sum ni} - \overline{X}^2}$$

مثال6:

أوجد التباين والانحراف المعياري بالصيغة الأصلية ثم بالصيغة المختصرة للبيانات المبوبة في الجدول الإحصائي التالي:

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

الحل:

الفئة	**التكرار**	X_i	n_{iXi}
8-4	3	6	18
13-9	4	11	44
18-14	6	16	96
23-19	2	21	42
28-24	4	26	104
المجموع	**19**		**304**

$(X_i - \overline{X})$	$n_i(X_i - \overline{X})$	$n_i(X_i - \overline{X})^2$
-10	-30	300
-5	-20	100
0	0	0
5	10	50
10	40	400
		850

$\mathbf{X_i^2}$	n_{iXi^2}
36	108
121	484
256	1536
441	882
676	2704
	5714

$$\overline{X} = \frac{\sum n_i X_i}{\sum n_i} = \frac{304}{19} = 16$$

التباين بالصيغة الأصلية
$$\overline{X} = \frac{\sum n_i (X_i - \overline{X})^2}{\sum n_i} = \frac{850}{19} = 44.74$$

الانحراف المعياري بالصيغة الأصلية
$$\delta_x = \sqrt{\frac{\sum n_i (X_i - \overline{X})^2}{\sum X_i}} = \sqrt{44.73} = 6.69$$

التباين بالصيغة المختصرة
$$V_x = \frac{\sum n_i X_i^2}{\sum n_i} - \overline{X}^2$$
$$= \frac{5714}{19} - (16)^2 = 300.74 - 256 = 44.74$$

الانحراف المعياري بالصيغة المختصرة

$$\delta_x = \sqrt{\frac{\sum n_i X_i^2}{\sum n_i} - \overline{X}^2}$$

$$S_x = \sqrt{\frac{5714}{19} - 256} = \sqrt{44.73} = 6.69$$

خصائص الانحراف المعياري:

1 – إذا كان الانحراف المعيار للقيم $X_1, X_2, X_3, X_4 \ldots\ldots X_n$ هو δ_x

فإنه إذا أضيفت أو طرحت قيمة ثابتة (a) إلى أو من جميع القيم فإن الانحراف المعياري لا يتغير

$$\delta_x = \sqrt{\frac{\sum X_i^2}{n} - \overline{X}^2}$$

بتربيع الطرفين

$$\delta_x^2 = V_x = \frac{\sum X_i^2}{n} - \overline{X}^2$$

$$n\delta^2 = \sum X_i^2 - n\overline{X}^2$$

بضرب الطرفين في (n)

نطرح قيمة ثابتة (a) من جميع القيم فنحصل على قيم جديدة $y_i = X_i - a$

$$\begin{aligned}
n\delta y_{ii}^2 &= \sum (X_i - a)^2 - n(\overline{X} -_i a)^2 \\
&= \sum \ (X_i^2 - 2aX_i + a^2) - n(\overline{X} - a)^2 \\
&= \sum X_i^2 - 2a\sum X_i + na^2 - n(\overline{X}^2 - 2a\overline{X} + a^2) \\
&= \sum X_i^2 - 2a\sum X_i + na^2 - n\overline{X}^2 + 2an\overline{X} - na^2 \\
&= \sum X_i^2 - 2an\frac{\sum X_i}{n} - n\overline{X}^2 + 2anX_i + na^2 - na^2 \\
&= \sum X_i^2 - 2a\sum X_i + na^2 - n\overline{X}^2 - 2an\overline{X} - na^2 \\
&= \sum X_i^2 - n\overline{X}^2 = n\delta_x^2
\end{aligned}$$

ومنه $\delta_x = \delta y$

2 – إذا كان الانحراف المعياري للقيم $X_1, X_2, X_3, X_4 \ldots\ldots X_n$ هو δ_x

فإنه إذا ضربت كل قيمة بالمقدار a (قسمت كل قيمة على المقدار a) فإن الانحراف المعياري يتأثر بالمقدار نفسه أي أنه إذا كان yi=ax فإن

$$n\delta_y{}^2 = \sum y_i{}^2 - n\overline{y}^2 = a\left[\sum X_i{}^2 - n\overline{X}^2\right]$$
$$n\delta_y{}^2 = \sum (aX_i)^2 - n\overline{aX_i}{}^2$$
$$n\delta_y{}^2 = a^2\sum X_i{}^2 - a^2 n\overline{X}^2$$
$$n\delta_y{}^2 = a^2(\sum X_i{}^2 - n\overline{X}^2)$$
$$\delta_y = a\delta_x$$

ومنه

3 – بالنسبة للتوزيع الطبيعي فإن:

* %68.27 من البيانات تقع في المجال $\overline{X} \pm \delta_x$
* %95.45 من البيانات تقع في المجال $\overline{X} \pm 2\delta_x$
* %99.73 من البيانات تقع في المجال $\overline{X} \pm 3\delta_x$

4 – يأخذ الانحراف المعياري نفس وحدة القياس للمتغير الأصلي (كلغ، متر، لتر ....) لذلك لا يمكن استخدامها كأساس للمقارنة بين تشتت توزيعين لهما وحدات قياس مختلفة.

5 – بما أن الانحراف المعياري يتأثر بالمتوسط الحسابي لبيانات الظاهرة فإنه لا يمكن استخدامه للمقارنة بين تشتت بيانات توزيعين لهما متوسط حسابي مختلف ولو كان هذين التوزيعين من نفس النوعية

6 – لا يمكن إيجاده بالنسبة للتوزيعات التكرارية المفتوحة من البداية أو النهاية.

7 – إذا كان لدينا مجموعة كلية متكونة من مجموعتين جزئتين أو أكثر فإنه يمكن حساب الانحراف المعياري لها من خلال العلاقة التالي:

$$S_x = \sqrt{\frac{1}{N}\left[n_1SX_1 + h_2SX_2\right] + \frac{1}{N}\left[h_1(\overline{X}_1 - \overline{X})^2 + h_2(\overline{X}_2 - \overline{X})^2\right]}$$

3– معامل الاختلاف Coefficient de variation:

رأينا في الصفحات السابقة أن الانحراف المعياري هو مقياس واقعي ومؤشر صحيح عن مقدار التشتت غير أن الخاصيتين 4و5 السابقتين تبينان أنه إذا استخدمنا هذا المقياس للمقارنة بين تشتت ظاهرتين أو أكثر فإن المقارنة تكون واقعية وواقعية فقط إذا كانت الظواهر من نوعية واحدة ولها متوسطات متساوية. أي يمكن مقارنة تشتت درجات مادة ما بدرجات مادة أخرى أو مقارنة تشتت دخل مجموعة من العمال بدخل مجموعة أخرى، وتكون المقارنة أكثر واقعية إذا كانت المتوسطات متساوية أو قريبة من بعضها.

أما إذا كانت الظواهر من صفات مختلفة أو إذا كانت متوسطاتها متباعدة، فإن المقارنة اعتمادا على الانحراف المعياري ستكون غير منطقية وغير واقعية، ولهذا السبب وجدت مقاييس أخرى سميت مقاييس التشتت النسبي تعتمد على تمييز البيانات وتقيس التشتت كنسبة مئوية للمتوسط، أهم هذه المقاييس هو معامل الاختلاف.

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

$$CV = \frac{\delta_x}{\overline{X}} \times 100$$

مثال7:

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

الحل:

إذا اعتمدنا على الانحراف المعياري فإننا نحكم على أن درجات المادة الأولى أكثر تشتتا ($\delta x = 3$) من درجات المادة الثانية ($\delta x = 2$)، وهذا غير صحيح لأننا إذا أدخلنا المتوسط الحسابي لدرجات الطلبة في المادتين في الحسبان سنحصل على النتائج التالية:

$$CV_1 = \frac{3}{15} \times 100 = 20\%$$

$$CV_2 = \frac{2}{8} \times 100 = 25\%$$

أي أن درجات المادة الثانية أكثر تشتتا

مثال8:

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

$\overline{X}_1 =$ 1500 ساعة		$\delta x_1 =$ 300 ساعة

$\overline{X}_2 =$ 1800 ساعة		$\delta x_2 =$ 325 ساعة

أي المصابيح لها مدة حياة أكثر تشتت؟

الحل:

معامل الاختلاف للنوع الأول: $CV_1 = \frac{300}{1500} \times 100 = 20\%$

معامل الاختلاف للنوع الثاني: $CV_2 = \frac{325}{1800} \times 100 = 18\%$

أي أن النوع الأول من المصابيح الكهربائية لها مدة حياة أكثر تشتتا.

تمارين حول مقاييس النزعة المركزية، و التشتت

تمرين1:

قامت مصلحة مكلفة بتعداد السكان بدراسة عدد الأطفال الذين تحتويهم كل عائلة من عائلات موجودة بمنطقة معينة، فتحصلوا على البيانات المدونة في الجدول المقابل.

عدد الأطفال	0	1	2	3	4	5	6 فأكثر
عدد العائلات	06	15	25	20	15	13	07

1- أحسب متوسط عدد أطفال هذه العائلات.

2- ما هو العدد الوسيطي للأطفال، ثم عدد الأطفال المتواجدين في أكبر عدد من العائلات.

3- حدد شكل التوزيع باستعمال مقاييس النزعة المركزية.

4- ما هو عدد العائلات التي تحتوي **على الأقل** على **4** أطفال ؟ ثم **أقل** من **4** أطفال ؟

5- ما هو عدد العائلات التي تحتوي **على الأكثر** على **3** أطفال ؟ ثم **أكثر** من **3** أطفال ؟

6- مثل هذه الظاهرة بيانياً.

تمرين2:

الأعمار	15 – 05	25 – 15	45 – 25	65 – 45	95 – 65
عدد الأشخاص	28	40	20	08	04

باستعمال العلاقة النظرية لبيرسون **[3(X - Me) = (X - Mo)]**، أحسب وسيط هذا التوزيع، ثم احسب أيضاً عدد الأشخاص الذي تقل أعمارهم عن **45** سنة، ثم الذين تزيد أو تساوي أعمارهم عن **25** سنة.

تمرين3:

يعبر التوزيع التكراري التالي عن توزيع **300** عامل في مؤسسة معينة على حسب أجرهم الأسبوعي بـ (**1000** دج).

أجر العمال	21	23	25	27	29	31	33
عدد العمال	12	36	57	81	63	42	09

1- أحسب متوسط الأجر الأسبوعي لعمال هذه المؤسسة.

2- ما هي قيمة الأجر الذي يتقاضاه أكبر عدد من العمال ؟

3- حدد شكل هذا التوزيع باستعمال مقاييس النزعة المركزية.

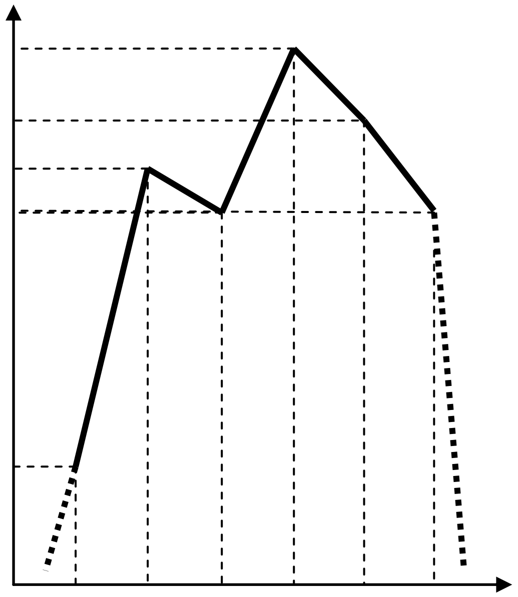
4- ما هو عدد العمال الذين يفوق أجرهم **26000** دج، ثم الذين يقل أجرهم عن **32000** دج ؟

5- كم هو عدد العمال الذين تتراوح أجورهم ما بين **23500** و **27500** دج ؟

6- أحسب المدى، المدى الربيعي النسبي ، ثم علل (ي) إجابتك (ي).

تمرين4:

يمثل العرض البياني المقابل، الأسعار التي بيع بها المنتوج A خلال أحد الأيام في معرض تجاري معين.



1- ضع معطيات هذا البيان داخل جدول تكراري

2- أحسب السعر المتوسط للمنتوج A.

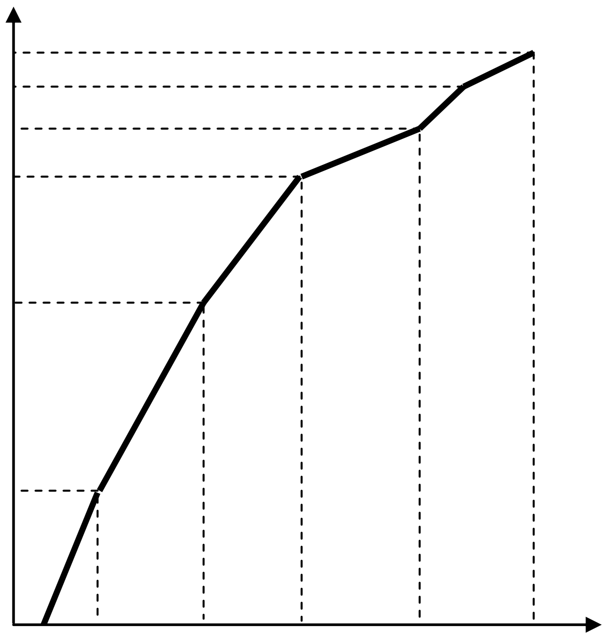
3- أحسب الوسيط، ثم حدده بيانياً.

4- أحسب السعر الذي بيع به المنتوج A بكثرة.

5- ما هو عدد المنتوجات التي بيعت بسعر محصور بين 400 و 600 دج.

تمرين5:

يعبر التمثيل البياني المقابل عن توزيع 200 سائح على حسب عدد الأيام التي مكثوها في فندق معين.



1- ضع المعطيات في جدول تكراري.

2- ما هي المدة التي بلغت فيها نسبة السياح **50**٪.

3- ما هي المدة التي قضاها أكبر عدد من السياح ؟

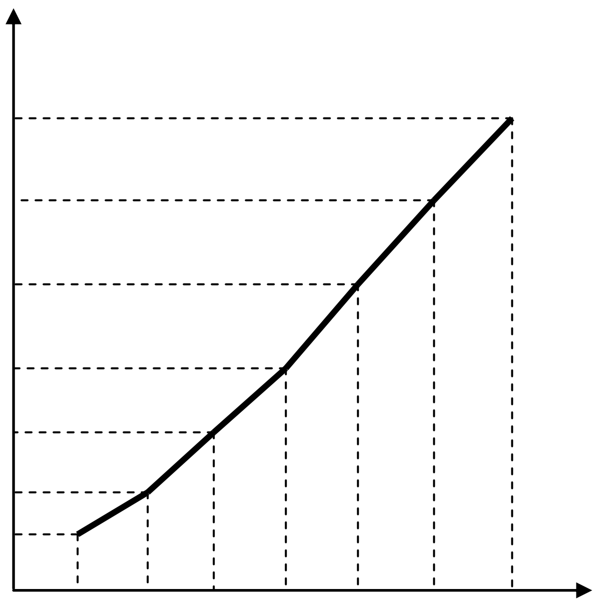
4- حدد شكل التوزيع باستعمال مقاييس النزعة المركزية.

5- أحسب الإنحراف الربيعي.

6- ما هو عدد السياح الذين تراوحت إقامتهم في الفندق بين 10 و 20 يوم ؟

تمرين6:

يمثل التمثيل البياني المقابل حجم المبيعات (بالوحدات) المحققة من طرف عينة من المحلات التجارية.



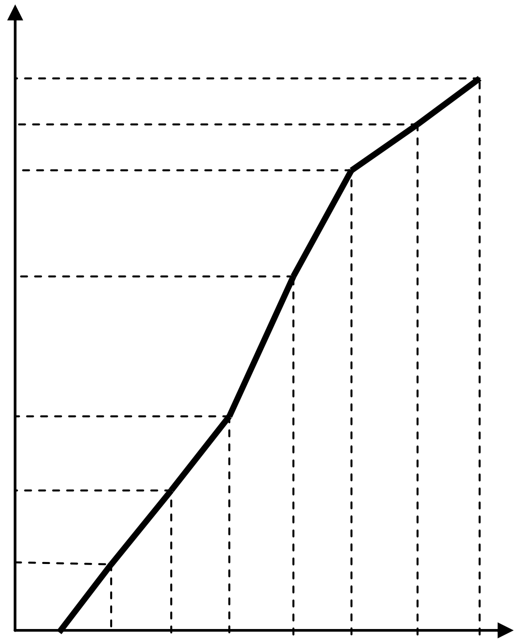
1- إستنتج عدد المحلات التجارية المكونة للعينة.

2- ضع معطيات البيان داخل جدول تكراري.

3- أحسب حجم المبيعات المتوسطة.

4- أحسب حجم المبيعات الوسيطية.

5- ما هو حجم المبيعات المحققة من طرف أكبر عدد من المحلات؟ ثم حدد شكل التوزيع.



تمرين7:

العرض البياني المقابل يبين التكرار المتجمع الصاعد لتوزيع عمال مؤسسة معينة على حسب فئات الأجور.(كل وحدة بـ **100** دج)

1- أرسم الجدول التكراري الذي يوضح تطور هذه الظاهرة.

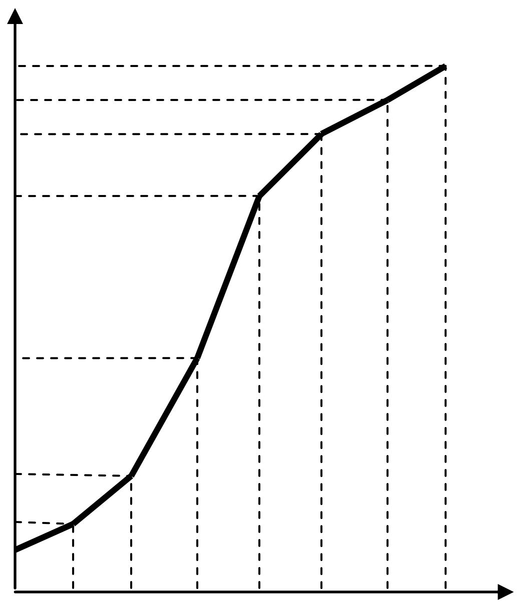
2- أحسب الأجر الوسيطي للعمال، ثم فسر نتيجته.

3- ما هو الأجر الذي يتقاضاه أكبر عدد من العمال.

4- حدد شكل هذا التوزيع باستعمال م النزعة المركزية.

5- ما هي نسبة العمال الذين يقل أجرهم عن 30000 دج ، ثم الذين يزيد أجرهم أو يساوي 15000 دج ؟

6- ما هو عدد العمال الذين يتراوح أجرهم بين **[18000 – 28000]** ؟



تمرين8:

الشكل المقابل يعبر عن الدراسة التي قام بها رئيس قسم **ل م د** عن المقاييس التي تحصل فيها طلبة دفعة معينة عن المعدل.

1- تحديد المتغير الإحصائي و نوعه و طبيعته.

2- تحديد المجتمع الإحصائي و العدد الإجمالي N مع العلم أن:

- $\sum f_i.x_i^2 = 18.95$
- $\sum n_i.x_i = 840$
- $V(x) = 1.31$

3- تحديد شكل التوزيع باستعمال مقاييس النزعة المركزية.

4- حساب المقاييس التالية مع تفسير نتائجهما: **D8، Q3، P28، و 1-F4↑**.

تمرين9:

كان متوسط معدلات الطلبة المتقدمين لإحدى الجامعات هو 10.2 بإنحراف معياري قدره 3.1، فيما أن متوسط معدلات الطلبة المتقدمين لجامعة أخرى هو 10.55، كان الإنحراف المعياري 2.8. إذا تقدم طالب معدله 12.5، في أي من الجامعتين تكون فرصة إدماجه و قبوله هي الأفضل ؟

تمرين10:

لدينا مجموعة من الأفراد ذوي دخل سنوي متوسط قدره 125000 دج بانحراف معياري قدره 35000 دج.

1- لو إقتطعنا ضريبة متساوية مقدارها 10000 لكل فرد فما هو الدخل السنوي المتوسط و الإنحراف المعياري الناتج سنوياً ؟

2- لو إقتطعنا نسبة من الدخل السنوي تقدر ب(1/10)، فما هو الدخل السنوي المتوسط و الإنحراف المعياري الناتج سنوياً ؟