وزارة التعليم العالي والبحث العلمي

الجامعة التقنية الشمالية

المعهد التقني الموصل

قسم تقنيات صحة المجتمع

الإحصاء الحياتي

العام الدراسي ٢٠٢٢ ـ ٢٠٢٣

الفصل الاول

تعاريف ومصطلحات

الفصل الأول

مقدمة :-

اشتقَ مصطلح الإحصاء باللغة الإنجليزية (Statistics) من الكلمة الإيطالية(Statista) ، والكلمة الألمانية(Statistik) ، والكلمة اللاتينية(Status) ، حيث كانت بداية استخدام هذا المصطلح لجمع البيانات التي تخص أفراد الدولة .

١.١ تعريف علم الإحصاء

عُرف الإحصاء على أنه "**العلم الذي يهتم بجمع البيانات الرقمية او النصية ، ومن ثم تنظيمها، وترتيبها، وتحليلها، بهدف الوصول إلى نتائج معينة لتوضيح ظاهرة أو حالة ما** " ، أو بأنه "**العلم الذي يهتم بالطريقة التي يتم من خلالها جمع البيانات والمعلومات وتحويلها إلى صورة عددية (استبدال البيانات النصية بأرقام)، حيث تُجمَع البيانات من خلاله بشكل منتظم باستخدام احدى طرق جمع البيانات**" .

وفيما يخص استخدامات علم الإحصاء فهي كثيرة؛ كاستخدامه في العلوم الطبية، وعلم الاجتماع، والاقتصاد، والصناعة، والكيمياء، والرياضة، والإدارة، وغيرها العديد من المجالات .

٢.١ أهمية علم الإحصاء

يعتبر علم الإحصاء احد الوسائل المهمة في البحث العلمي من خلال استخدام قواعده وقوانينه وطرقه في جمع البيانات والمعلومات اللازمة للبحث العلمي وتحليل هذه البيانات والمعلومات بغية التوصل إلى النتائج التي يهدف إليها البحث . كما وان للإحصاء دورا بارزا في وضع الخطط المستقبلية عن طريق التنبؤ بالنتائج ولكافة القطاعات سواء كانت إنتاجية أم خدمية . وحيث أن الإحصاء بحد ذاته يعتبر وسيلة وليس غاية فذلك يعني إن مجالات استخدامه أينما وجد البحث العلمي

ونتيجة لكثرة استخدامات هذا العلم في الكثير من المجالات فقد برزت وخلال فترة ليست بالقصيرة من الزمن مسميات للإحصاء مقرونة باسم علم آخر ولو أن ذلك لا يغير من مضمون وهدف علم الإحصاء إلا أن ذلك حدث بهدف إيجاد تخصصات دقيقه للإحصاء مثل **الإحصاء السكاني ، الإحصاء الحيوي ، الإحصاء الاقتصادي** ، وغيرها من المسميات الأخرى التي بمجملها تعني استخدام الإحصاء والطريقة الإحصائية في هذه المجالات ، فالأسلوب واحد وان اختلف مجال التطبيق .

١.٣ الطريقة الإحصائية في البحث العلمي

إن استخدام الأسلوب الإحصائي في البحث العلمي يعني توفر بيانات ومعلومات عن الظاهرة أو الظواهر المطلوب دراستها في ذلك البحث . وهذا يعني إن إمكانية تطبيق الطريقة الإحصائية مرهون بإمكانية التعبير عن هذه الظاهرة **تعبيرا كميا** .

وفيما يلي المراحل الرئيسية للطريقة الإحصائية في البحث العلمي :

١- تحديد مشكلة أو فرضية البحث أو الدراسة .

٢- جمع البيانات والمعلومات عن الظاهرة أو الظواهر ذات العلاقة بالبحث أو الدراسة .

٣- تصنيف البيانات وتبويبها وعرضها .

٤- حساب المؤشرات الإحصائية (كالوسط الحسابي , التباين, المنوال ...) كتقديرات لمعالم مجتمع البحث أو الدراسة .

٥- تحليل معطيات الدراسة والتوصل للنتائج على ضوء فرضية أو فرضيات البحث أو الدراسة .

٦- تفسير النتائج وعملية اتخاذ القرار بشأن فرضيات البحث .

١.٤ أنواع البيانات الإحصائية

البيانات هي مجموعة من المشاهدات أو الملاحظات التي تؤخذ أو تجمع أثناء أخذ العينة ، وقد تكون البيانات رقمية (كمية) مثل أطوال وأوزان ، أو الدخل الشهري الأسرة ، وقد تكون البيانات غير رقمية (وصفية) مثل الجنس (ذكر او انثى) او لون العين (اسود او ازرق ....) او زمرة الدم (A,B,AB,O) وغيرها ... .

كلما كان جمع البيانات دقيقا زادت ثقة الباحث في الاعتماد عليها، ولن يكون تحليل البيانات صحيحا أو مفيدا إذا كان هناك أخطاء في جمع البيانات .

١.٥ أساليب جمع البيانات الإحصائية: يتم جمع البيانات الإحصائية بأحد الأساليب التالية :

١- المسح الشامل: جمع البيانات من جميع مفردات المجتمع دون إستبعاد أي مفردة فمثلا إذا أردنا التعرف على مستوى طلاب الجامعة المستنصرية في مادة الإحصاء نقوم برصد درجات جميع طلاب الجامعة في مادة الإحصاء وهكذا... وهذه الطريقة عادة تكون طويلة ومكلفة وتحتاج إلى الكثير من الوقت ناهيك عن عدم إمكانية تطبيقها في الحالات التي تؤدي فيها جمع البيانات عن مفردات البحث إلى فناء هذه المفردات.

٢- العينة : وفيها يتم اختيار عينة تمثل المجتمع وتجرى عليها الدراسة وتعمم النتائج على المجتمع وكلما كانت العينة مختارة بطريقة صحيحة وممثلة تمثيلا صادقا المجتمع كلما كانت النتائج صادقة ودقيقة.

١.٦ مصادر جمع البيانات

إن أي بحث علمي يستند في تحليله إلى الطريقة الإحصائية يحتاج إلى بيانات ومعلومات حول موضوع البحث قيد الدراسة . ويمكن للباحث الحصول على هذه البيانات من احد المصدرين التاليين :

أ – المصادر التأريخية (Historical Sources)

وهي البيانات والمعلومات المحفوظة والمتجمعة لدى أجهزة ومؤسسات ودوائر الدولة المختلفة نتيجة لاستقصاءات أو مسوحات قامت بها هذه الجهات ، أو هيئات معينة لأغراض خاصة بها ، أو تجمعت لديها بحكم وظائفها الإدارية والفنية مثال على ذلك البيانات المتجمعة عن تعدادات السكان في العراق ، إحصاءات الطلبة المتخرجين من الجامعات العراقية ، وغيرها .

ب – مصادر الميدان (Field Sources)

بيانات ومعلومات يمكن الحصول عليها من مصادرها الأصلية بطريقة المراسلة أو المواجهة أو أي طريقة اتصال أخرى .

إن اختيار هذا المصدر دون ذاك في جمع البيانات والمعلومات يعتمد بالأساس على طبيعة البحث والنتائج المتوخاة منه .

١.٧ طرق جمع البيانات: Collecting Data

هناك عدة طرق لجمع البيانات نذكر منها:

١ـ المقابلة الشخصية : (Personal Interview)

وهي أن تقوم بمقابلة أفراد العينة والتحدث إليهم عن الموضوع الذي يتم إجراء البحث فيه , وبذلك فإن المعلومات التي سنقوم بجمعها ستكون دقيقة إلى حد ما، إلا أن تحليلها سيكون صعبا، وعليك أن تنتبه إلى تدوين البيانات أثناء المقابلة لان أي خطأ في تدوين هذه البيانات يؤدي إلى خطأ في النتائج .

٢ـ الملاحظة المباشرة (Direct Observation)

عندما لا يكون هناك أفراد للعينة ، فانك تستخدم هذه الطريقة أي الملاحظة المباشرة ومن الأمثلة عليها أن تقف على تقاطع طرق ، وتعد السيارات التي تمر من هذا التقاطع من الساعة الثامنة وحتى التاسعة بهدف حصر كثافة السير في وقت ذهاب الموظفين إلى أعمالهم أو أن تقوم بمراقبة تصرف مجموعة من الأطفال أثناء اللعب وتدوين الملاحظات بهدف التعرف على سلوكيات الأطفال في بعض المواقف.

٣ـ الإستبيان (Questionnaire)

الأستبيان يُعرَف الاستبيان بأنه أداة لجمع البيانات المتعلقة بموضوع بحث محدَد عن طريق استمارة يجري تعبئتها من قبل المستجيب، ويستخدم لجمع المعلومات بشأن معتقدات ورغبات المستجيبين، ولجمع حقائق هم على علمٍ بها , ولهذا يستخدم بشكلٍ رئيس في مجال الدراسات التي تهدف إلى استكشاف حقائق عن الممارسات الحاليّة واستطلاعات الرأي العام وميول الأفراد، وإذا كان الأفرادُ الذين يرغب الباحث في الحصول على بيانات بشأنهم في أماكن متباعدة فإنّ أداة الاستبيان تمكّنه من الوصول إليهم جميعا بوقت محدود وبتكاليف معقولة .

من الملاحظ أنّ أداة الاستبيان **منتشرة في الدراسات الابتكاريّة والتطبيقيّة**، وذلك لأسباب منها :ـ

١ـ أنها أفضل طريقة للحصول على معلومات وحقائق جديدة لا توفرها مصادر أخرى.

٢ـ أنها تتميَّز بالسهولة والسرعة في توزيعها بالبريد على مساحة جغرافيَّة واسعة.

٣ـ أنها توفر الوقت والتكاليف.

٤ـ أنها تعطي للمستجيب حريَّة الإدلاء بأيَّة معلومات يريدها.

أنواع الاستبيان

للاستبيان بحسب إجاباته المتوقفة على طبيعة أسئلة الاستبيان ثلاثة أنواع هي : ـ

١ـ الاستبيان المفتوح : وفيه فراغات يتركها الباحث ليدوّن فيها المستجيبون إجاباتهم، وهذا النوع يتميَّز بأنه أداة لجمع حقائق وبيانات ومعلومات كثيرة غير متوفرة في مصادر أخرى، ولكنّ الباحث يجد صعوبة في تلخيص وتصنيف النتائج , لتنوُّع الإجابات، ويجد إرهاقا في تحليلها ويبذل وقتا طويلا لذلك، كما أنّ كثيرا من المستجيبين قد يغفلون عن ذكر بعض الحقائق في إجاباتهم بسبب أنّ أحدا لم يذكّرهم بها وليس لعدم رغبتهم بإعطائها.

٢ـ الاستبيان المقفول : وفيه الإجابات تكون باختيار (نعم أو لا)، أو بوضع علامة (صحّ أو خطأ) ، أو تكون باختيار (إجابة واحدة من إجابات متعدّدة) ، وفي مثل هذا النوع ينصح الباحثون أن تكونَ هناك إجابة أخرى مثل : غير ذلك، أو لا أعرف ، وليحافظ الباحث على الموضوعيَّة يجب عليه أن يصوغ عبارات هذا النوع من الاستبيان بكلّ دقة وعناية بحيث لا تتطلَّب الإجابات تحفُّظات أو تحتمل استثناءات ، ويتميَّز هذا النوع من الاستبيانات بسهولة تصنيف الإجابات ووضعها في قوائم أو جداول إحصائيَّة يسهل على الباحث تلخيصها وتصنيفها وتحليلها، ومن ميزاته أنه يحفز المستجيب على تعبئة الاستبانة لسهولة الإجابة عليها وعدم احتياجها

إلى وقتٍ طويل أو جهدٍ شاق أو تفكيرٍ عميق بالمقارنة مع النوع السابق، ولهذا تكون نسبة إعادة الاستبانات في هذا النوع أكثر من نسبة إعادتها في النوع المفتوح.

٣- الاستبيان المفتوح – المقفول : يحتوي هذا النوع على أسئلة النوعين السابقين، ولذلك فهو أكثرُ الأنواع شيوعا، ففي كثير من الدراساتِ يجد الباحث ضرورة أن تحتوي استبانته على أسئلة مفتوحة الإجابات وأخرى مقفلة الإجابات، ومن مزايا هذا النوع أنه يحاول تجنب عيوب النوعين السابقين وأن يستفيد من ميزاتهما .

تصميم الاستبيان وصياغتها

ممّا يجب على الباحث مراعاته عند تصميم الاستبيان الآتـي :

١- الإيجاز بقدر الإمكان .

٢- حسن الصياغة ووضوح الأسلوب والترتيب وتخطيط الوقت .

٣- استخدام المصطلحات الواضحة البسيطة، وشرح المصطلحات غير الواضحة .

٤- إعطاء المبحوث مساحة حرّة في نهاية الاستبانة لكتابة ما يراه من إضافة أو تعليق .

٥- الابتعاد عن الأسئلة الإيحائيّة الهادفة إلى إثبات صحّة فرضيّات دراسته .

٦- تجنّب الأسئلة التي تستدعي تفكيرا عميقا من المبحوثين أو المتعاونين مع الباحث .

٧- تزويد الاستبانة بما يشرح أهداف الدراسة وقيمتها التطبيقيّة بما يعود على الأفراد المبحوثين أو المجتمع المبحوث بالخير .

٨- تزويد الاستبانة بتعليمات وبإرشادات عن كيفيّة الإجابة، وحفز المبحوثين ليستجيبوا بكلٍ دقة وموضوعيّة .

٩- احتواء الاستبيان على أسئلة مراجعة للتأكُد من صدق البيانات وانتظامها .

مزايا وعيوب الاستبيان

تعرّضت أداة الاستبيان إلى نقد شديد من المهتمّين بأساليب البحث العلميّ، ومعظم انتقاداتهم تركّزت على مدى دقة وصحّة البيانات والمعلومات التي يجمعها الباحث بهذه الأداة، وبالرغم من وجود عيوب أداة الاستبيان فلها مزايا تجعلها من أهمّ أدوات جمع البيانات وأكثرها شيوعا .

مزايا الاستبيان

١ـ تمكّن أداة الاستبيان من حصول الباحثين على بيانات ومعلومات من وعن أفراد ومفردات يتباعدون وتتباعد جغرافيًا بأقصر وقتٍ مقارنة مع الأدوات الأخرى .

٢ـ يعدُ الاستبيان من أقل أدوات جمع البيانات والمعلومات تكلفة سواءً أكان ذلك بالجهد المبذول من قبل الباحث أم كان ذلك بالمال المبذول لذلك .

٣ـ تعدُ البيانات والمعلومات التي تتوفر عن طريق أداة الاستبيان أكثر موضوعيَّة ممّا يتوفر بالمقابلة أو بغيرها، بسبب أنَّ الاستبيان لا يشترط فيه أن يحملَ اسم المستجيب ممّا يحفزه على إعطاء معلومات وبيانات موثوقة .

٤ـ توفر طبيعة الاستبيان للباحث ظروف التقنين أكثر ممّا توفره له أدوات أخرى، وذلك بالتقنين اللفظيّ وترتيب الأسئلة وتسجيل الإجابات .

٥ـ يوفّر الاستبيان وقتا كافيا للمستجيب أو المتعاون مع الباحث للتفكير في إجاباته ممّا يقلّلُ من الضغط عليه ويدفعه إلى التدقيق فيما يدوّنـه من بيانات ومعلومات .

عيوب الاستبيان

١ـ قد لا تعود إلى الباحث جميعُ نسخ استبيانه , ممّا يقلل من تمثيل العيّنة لمجتمع البحث .

٢ـ قد يعطي المستجيبون أو يدوّن المتعاونون مع الباحث إجابات غير صحيحة، وليس هناك من إمكانيَّة لتصحيح الفهم الخاطئ بسبب الصياغة أو غموض المصطلحات وتخصُصها .

٣ـ قد تكون الانفعالات من المعلومات المهمّة في موضوع الدراسة، وبالاستبيان لا يتمكّن الباحث من ملاحظة وتسجيل ردود فعل المستجيبين لفقدان الاتّصال الشخصيّ معهم .

٤ـ لا يمكن استخدام الاستبيان في مجتمع لا يجيد معظمُ أفراده القراءة والكتابة .

٥ـ لا يمكن التوسُع في أسئلة الاستبيان خوفا من ملل المبحوث أو المتعاون مع الباحث حتى ولو احتاجت الدراسة إلى ذلك .

التالي نموذج لإستمارة استبيان من النوع المقفول :ـ

استمارة إستبيان

تقييم أولي لمستوى الطالب في بعض إستخدامات الحاسب الآلي

#	الفئة	ممتاز	جيد	مقبول	لا أعرف
١	نظام تشغيل مايكروسوفت ويندوز 7	☐	☐	☐	☐
٢	برنامج معالج النصوص Microsoft Word	☐	☐	☐	☐
٣	برنامج الجداول الإلكترونية Microsoft Excel	☐	☐	☐	☐
٤	برنامج العروض التقديمية Microsoft PowerPoint	☐	☐	☐	☐
٥	برنامج قواعد البيانات Microsoft Access	☐	☐	☐	☐
٦	برنامج عرض الشرائح Microsoft Powerpoint	☐	☐	☐	☐
٧	الإنترنت Internet	☐	☐	☐	☐
٨	كيفية البحث عن طريق الإنترنت Electronic Search	☐	☐	☐	☐
٩	كيفية بناء موقع إلكتروني Website	☐	☐	☐	☐
١٠	التواصل عن طريق البريد الإلكتروني Email	☐	☐	☐	☐
١١	برنامج التصاميم الهندسية AutoCAD	☐	☐	☐	☐
١٢	برنامج SPSS	☐	☐	☐	☐
١٣	برامج أخرى، أذكرها إن وجد:				

ملاحظات أخرى (إن وجد):

____

____

____

للتواصل مع الباحث :-

١٨ المتغيرات العشوائية (Random Variables)

يُعرف المتغير العشوائي بأنه داله ذات قيمة حقيقية (عدد حقيقي) معرفة على فضاء يدعى فضاء العينه ، وغالبا ما يرمز للمتغير العشوائي بأحد الأحرف الكبيرة مثل X ، Y، Z ،...الخ ولقيم المتغير العشوائي عند تنفيذ التجربة بأحد الأحرف الصغيرة x ، y ، z ، ..... الخ .

(مثال)

تأمل تجربة رمي زهر نرد وملاحظة العدد الذي سوف يظهر على وجه الزهر بعد رميه . هنا المتغير X هو العدد الذي سوف يظهر على وجه الزهر بعد رميه ، افرض أن العدد الذي ظهر بعد تنفيذ الرمية هو العدد 3 (عدد حقيقي) ، وان الحالات الكلية الممكنة الظهور على وجه الزهر هي مجموعة الأعداد الحقيقية {6،5،4،3،2،1} ولا يمكن إطلاقا ظهور غيرها ، أي أنها القيم الممكنة إلى X التي نتوقع ظهور احدها بعد عملية الرمي . وهذا ما نطلق عليه فضاء العينه للمتغير X . وحيث أن العدد 3 هو قيمه من قيم المتغير X الممكنة الذي ينتمي لفضاء هذا المتغير . وحيث أن تجربة رمي الزهر هي تجربة عشوائية تتم دون تحيز لهذا الوجه أو ذاك ، عليه فان X متغير عشوائي .

وغالبا ما يرمز إلى مجموعة القيم الممكنة (فضاء العينه) بالرمز Ω ويلفظ اوميكا ، ففي المثال السابق فان Ω ستعرف بالشكل التالي $\Omega=\{X:x=1,2,3,4,5,6\}$.

وتقسم المتغيرات العشوائية إلى قسمين رئيسيين هما :-

١.١٨ المتغيرات العشوائية النوعية (الوصفية) Qualitative Variable

وهي المتغيرات التي لا يمكن قياسها بوسائل قياس مألوفة كالعد أو التقييس إنما تشكل صفات لذلك المتغير . مثل لون العين (اسود . عسلي . ازرق) ، الحالة الاجتماعية (أعزب . متزوج . مطلق . أرمل) , الجنس (ذكر .أنثى) وغيرها من الأمثلة .

وتستخدم عدة مقاييس لقياس البيانات النوعية منها:

أ.التدرج الاسمي : Nominal Scale

هذا المقياس يصنف عناصر الظاهرة التي تختلف في النوعية لا في الكمية ، وكثيرا ما نستخدم الأعداد لتحديد هوية المفردات ، وفي هذه الحالة لا يكون للعدد ذلك المدلول الكمي الذي يفهم منه عادة . فمثلا يمكن استعمال العددين (0 ،1) ليدلا على التصنيف حسب الجنس فيجعل الصفر(0) يدل على الذكر والـ (1) يدل على الأنثى، لاحظ أن (0 ، 1) لا يدلان على قيم عددية أي لا يخضعان للعمليات الحسابية

لأنه يمكن تعيين أي عددين بدلهما ليدلا على نوع الجنس .

أمثلة أخرى على المقياس الاسمي : الحالة الاجتماعية (أعزب- متزوج) ، ونـوع العمـل (إداري - فني - تدريسي) .
ومن الجدير بالذكر أن هذا المقياس لا يعطي الأفضلية لإحدى طبقات المجتمع على الأخرى .

ب- التدرج الترتيبي : Ordinal Scale

يقع هذا التدرج في مستوى أعلى من التدرج الاسمي، فبالإضافة إلى خواص التدرج الاسمي فان التدرج الترتيبي يسمح بالمفاضلة ، أي بترتيب العناصر حسب سلم معين: مثل الرتب الأكاديمية (أستاذ (1) ، أستاذ مساعد(2)، مدرس(3)، مدرس مساعد (4)) . وتقديرات الطلاب (ممتاز(6)، جيد جدا(5)، جيد(4)، متوسط (3) ، مقبول(2)، راسب(1)) وكذلك درجة التأييد لإجابة السؤال (موافق جدا (5) ، موافق (4) ، متردد(3) ، لا أوافق (2) ، لا أوافق جدا (1)) . نذكر أن هذا المقياس لا يحدد الفرق بدقة بين قيم الأفراد المختلفة

١.٨.٢ المتغيرات الكمية Quantitative Variable

وهي المتغيرات التي يمكن قياسها بوسائل قياس مألوفة مثل عدد الطلبة في صف معين , عدد أشجار البرتقال في بستان , طول الشخص بالسنتمتر , وزن حمولة من الاسمنت بالطن , وهذا القسم من المتغيرات على توعين رئيسيين هما :-

أ – المتغيرات المتقطعة Discrete Variable

إذا كانت مجموعة القيم الممكنة للمتغير X مجموعة قابلة للعد سواء كانت مجموعة محدودة أم غير محدودة عندئذٍ يقال أن X متغير عشوائي متقطع .

(مثال ١)

إن مجموعة القيم الممكنة إلى X في تجربة رمي زهر النرد هي المجموعة $\Omega = \{X : x = 1, 2, 3, 4, 5, 6\}$. وحيث انه من الممكن عد العناصر هذه المجموعة (أي أنها مجموعة قابلة للعد) بالرغم من كونها محدودة (أي لها بداية , العدد 1 ، ونهاية ، العدد 6) عليه فأن Xمتغير عشوائي متقطع .

(مثال ٢)

افترض أن X متغير عشوائي يشير إلى عدد النداءات الهاتفية المستقبلة من قبل بدالة هاتف خلال فترة زمنية محددة، واضح هنا إن مجموعة القيم الممكنة إلى Xهي مجموعة الأعداد الحقيقية $\{X : x = 0, 1, 2, 3, \cdots\} = \Omega$ وحيث انه من الممكن عد عناصر هذه المجموعة بالرغم من كونها مجموعة غير محدودة (لها بداية وليس لها نهاية) فإذا X متغير عشوائي متقطع .

ب – المتغيرات المستمرة Continuous Variables

إذا كانت مجموعة القيم الممكنة للمتغير X مجموعة غير قابلة للعد سواء كانت مجموعة محدودة أم غير محدودة عندئذٍ يقال أن X متغير عشوائي مستمر .

(مثال ١)

افترض أن X متغير عشوائي يشير إلى الزمن المستغرق لقطع المسافة بين بغداد ونينوى البالغة 400 كم بسيارة تتراوح سرعتها ما بين 80 إلى 100 كم / ساعة . واضح وفق قانون السرعة إن الزمن المستغرق لقطع هذه المسافة يتراوح ما بين 4 ساعات إلى 5 ساعات . عليه فأن مجموعة القيم الممكنة إلى X هي مجموعة الأعداد الحقيقية $\Omega = \{X : 4 \le X \le 5\}$. وحيث انه لا يمكن عد عناصر هذه المجموعة كونها واقعة ضمن فترة مستمرة ، أي وجود عدد غير منتهي من القيم الواقعة ضمن هذه الفترة على الرغم من كونها مجموعة محدودة (لها بداية ونهاية) فإذن X متغير عشوائي مستمر .

(مثال ٢)

افترض أن X متغير عشوائي يشير إلى كمية المواد المتدفقة مقاسة بالمتر المكعب (m^3) من انفجار بركاني محتمل الوقوع . واضح هنا ان مجموعة القيم الممكنة إلى X تتراوح ما بين الصفر وعدد كبير جدا (نظريا ما لا نهايه ∞) . هذه المجموعة يمكن ترميزها بالشكل $\Omega = \{X : 0 \le X \le \infty\}$. وحيث انه لا يمكن عد عناصر هذه المجموعة على الرغم من كونها مجموعة غير محدودة . فإذن X متغير عشوائي مستمر .

الشكل التالي يوضح أنواع المتغيرات العشوائية :-

المتغيرات العشوائية

المتغيرات الكمية
Quantitative Variable

المتغيرات العشوائية النوعية (الوصفية)
Qualitative Variable

المتغيرات المتقطعة
Discrete Variable

المتغيرات المستمرة
Continuous Variables

شكل (1 – 1) أنواع المتغيرات العشوائية

الفصل الثاني

عرض البيانات

الفصل الثاني

سوف نركز الاهتمام في هذا الفصل على أساليب تبويب البيانات في جداول خاصة تدعى بجداول التوزيعات التكرارية . كذلك استعراض لأساليب عرض البيانات هندسيا . وفيما يلي تعاريف لبعض المفاهيم والمصطلحات المطلوبة في بعض فقرات هذا الفصل .

٢.١ العرض الجدولي للبيانات

سوف تخصص هذه الفقرة لدراسة أساليب عرض البيانات المصنفة في جداول خاصة تدعى بالتوزيعات التكرارية التي تتخذ أشكالا متعددة حسب نوع المتغير العشوائي الذي صنفت على أساسه البيانات الخام .

٢.٢ التوزيع التكراري Frequency Distribution

التوزيع التكراري عبارة عن تلخيص وترتيب لبيانات المتغير العشوائي , التي سبق ان جمعت وصنفت , مقسمة إلى عدد من المجاميع كل منها تسمى (الفئه Class) . هذه الفئات قد تكون مرتبة تصاعديا أو تنازليا حسب طبيعة البيانات . ويسمى توزيع عدد القيم لـ X حسب الفئات " التوزيع التكراري " . **وقد تكون فئات التوزيع متساوية في الطول أم غير متساوية وذلك يعتمد على طبيعة الدراسة ومتطلباتها .**

ولكي نتمكن من استيعاب ما تقدم نأتي إلى إعطاء شرح لمكونات التوزيع التكراري من خلال المثال الافتراضي التالي :-

لتكن X1,X2,X3,…,Xn تمثل بيانات المتغير العشوائي X من عينة عشوائية من المفردات قوامها n مفردة . ونرغب في تلخيص هذه البيانات في توزيع تكراري عدد فئاته هو m .

لنفرض إن (X_S) تعني اصغر قيمة و(X_L) تعني اكبر قيمة في مجموعة البيانات هذه عندئذٍ تعرف مكونات التوزيع على النحو التالي :

أ. المدى الكلي للتوزيع Total Range

يعرف المدى الكلي للتوزيع بأنه الفرق ما بين اكبر قيمة واصغر قيمة في المجموعة . فإذا رمزنا للمدى الكلي بالرمز T. R عندئذٍ فإن :

$$T.R = X_L - X_s$$

X_L اكبر قيمة X_s أصغر قيمة

٢ـ عدد فئات التوزيع Number of Classes

تمثل عدد المجاميع التي يتألف منها التوزيع التكراري . وهنالك صيغ تقريبية يمكن من خلالها تحديد عدد فئات التوزيع أهمها :

أ – صيغة يول (Yule) وهي : $m = (2.5) . \sqrt[4]{n}$

(حيث إن n تمثل عدد المشاهدات)

ب – صيغة ستر جس وهي (Sturges) $m = 1+(3.322).(\log.n)$

(log.n) :- تمثل لوغاريتم عدد المشاهدات

وعند التطبيق يتم تقريب الناتج إلى اقرب عدد صحيح .

٣ـ طول الفئة Length of a Class

ويسمى أحيانا بالمدى الفئوي ويمثل مقدار سعة الفئة , أي مقدار المسافة ما بين الحد الأدنى للفئة وحدها الأعلى .وإذا رمزنا لطول الفئة بالرمز(L) فإنه يمكن تحديد قيمة L من خلال الصيغة التالية :

$$L = \frac{T.R}{m}$$

....

٤ـ الحد الأدنى والحد الأعلى للفئة Lower and Upper Bound of a Class

لكل فئة من فئات التوزيع التكراري بداية ونهاية . فالبداية تعني الحد الأدنى للفئة والنهاية تعني الحد الأعلى لها . ويمكن تكوين حدود الفئات على النحو التالي :

تسلسل الفئة	الحد الأدنى	الحد الأعلى
١	X_s (اصغر قيمة)	X_s +L -1 (اصغر قيمة+طول الفئة)
٢	X_s +L (اصغرقيمة+ طول الفئة)	X_s +2L -1 (اصغرقيمة+اثنين من طول الفئة)
٣	X_s +2L (اصغر قيمة+اثنين من طول الفئة)	X_s +3L -1 (اصغرقيمة+ثلاثة من طول الفئة)
.	.	.
.	.	.
.	.	.
M	X_s +(m-1)L (اصغرقيمة+(m-1)من طول الفئة)	X_s +(m)L -1 (اصغر قيمة+(m)من طول الفئة)

٥ـ مركز الفئة Center of a Class

يمثل مركز الفئة قيمة من قيم المتغير العشوائي X التي تتوسط المسافة بين الحد الأدنى والحد الأعلى للفئة .

فإذا رمزنا للحد الأدنى بالرمز L.L , والحد الأعلى بالرمز U.L , ولمركز الفئة بالرمز X فإن :

$$X = \frac{L.L + U.L}{2}$$

٦ـ تكرار الفئة Class Frequency

يمثل تكرار الفئة جزء من مفردات العينة التي تتصف بكونها تقع من حيث القيمة العددية ما بين حدي الفئة بحيث إن مجموع هذه الأجزاء يشكل عدد مفردات العينة (n) , فإذا رمزنا لتكرارات الفئات بالرمز f1,f2,f3,…,fm فإن f1+f2+f3+…+fm=n (أي أن مجموع تكرارات الفئات يساوي عدد مشاهدات هذه العينة) .

وفيما يلي الطرق المختلفة لكتابة حدود فئات التوزيع التكراري استنادا إلى نوع المتغير العشوائي :

أـ في حالة المتغيرات العشوائية المتقطعة :ـ

نكتب حدود الفئات حسب التوزيع التكراري التالي بحيث نضمن أن كل قيمة من قيم البيانات تضمن في فئة واحدة من فئات التوزيع دون أي تكرار قد يحصل في هذه الفئة أو تلك وبحيث أن طول الفئة (L) يكون مساويا للفرق مابين الحد الأعلى والحد الأدنى .

(مثال ١)

البيانات التالية تمثل عدد المراجعين المرضى تم تسجيلها في مجموعة مراكز صحية عددها (٦٠ مركز صحي) , يطلب تبويب هذه البيانات في جدول توزيع تكراري .

72	68	65	90	82	132	120	80	76	60
110	103	101	98	90	88	164	157	142	150
116	111	122	120	114	109	126	120	116	119
119	121	113	120	104	98	95	93	78	90
154	150	142	120	118	136	131	130	126	125
136	110	137	136	154	156	125	139	123	122

الحل : واضح إن المتغير العشوائي (عدد المراجعين المرضى) هو من النوع المتقطع .

إن اصغر قيمة في المجموعة هو العدد 60 , وان اكبر قيمة فيها هو العدد 164 وبذلك فإن المدى الكلي لهذه المجموعة هو

$T.R = \chi_L - \chi_s$,

$T.R= 164-60 = 104$ المدى الكلي للبيانات

وباستخدام صيغة يول في تحديد عدد فئات التوزيع نلاحظ إن

$m = (2.5) . \sqrt[4]{n}$

$m = 2.5 . \sqrt[4]{60} = 2.5 . (2.783)$

$m = 6.958$ (بالتقريب) ≈ 7 عدد الفئات جدول التوزيع التكراري

وحيث أن عدد الفئات يجب أن يكون عدد صحيح لذا يتم تقريب قيمة m إلى اقرب عدد صحيح وهو 7 .

وبذلك فإن طول الفئة يمكن تحديده وفق الصيغة التالية

$L = \frac{T.R}{m}$

$L = \frac{104}{7} = 14.857$ (بالتقريب) ≈ 15 طول الفئة في جدول التوزيع التكراري

واستنادا إلى طول الفئة نبدأ بتحديد حدي كل فئة على النحو التالي :

ت . الفئه	الحد الأدنى للفئه	الحد الأعلى للفئه	الفئات
1	60	60 + 15 - (1)=74	60 - 74
2	60 + 15=75	60 + 30 - (1)=89	75 - 89
3	60 + 30=90	60 + 45 - (1)=104	90 - 104
4	60 + 45=105	60 + 60 - (1)=119	105 - 119
5	60 + 60=120	60 + 75 - (1)=134	120 - 134
6	60 + 75=135	60 + 90 - (1)=149	135 - 149
7	60 + 90=150	60 + 105 - (1)=164	150 - 164

ت . الفئه	الفئات	التكرارات (f)	مركز الفئات (X)
1	60 - 74	4	67
2	75 - 89	5	82
3	90 - 104	10	97

4	105 - 119	12	112
5	120 - 134	16	127
6	135 - 149	7	142
7	150 - 164	6	157
المجموع	60		

وهذا يعني أن أربعة مراكز صحية كان عدد المراجعين المرضى يتراوح ما بين 60 إلى 74 مراجع , وخمس مراكز صحية كان عدد المراجعين المرضى يتراوح ما بين 75 إلى 89 مراجع وهكذا .

العمود الأخير من الجدول (مركز الفئات) يعني أن أربعة مراكز صحية كان عدد المراجعين المرضى بالمتوسط 67 مراجع , وخمس مراكز صحية كان عدد المراجعين المرضى بالمتوسط 82 مراجع , وهكذا

ويلاحظ أن الفرق بين مركز الفئة اللاحق ومركز الفئة السابق ماهو إلا طول الفئة . فالفرق ما بين مركز الفئة الثانية ومركز الفئة الأولى 15=67-82. وهو نفس الفرق ما بين مركز الفئة الثامنة ومركز الفئة السابعة.

<u>(مثال ٢)</u>

الأتي درجات امتحان لمادة ما من 25 درجة (٤٠ درجة امتحانية فقط)

13	14	12	11	8	11	12	10	13	14
13	16	14	17	19	21	22	23	24	24
25	25	17	15	14	15	18	17	19	18
20	17	16	15	14	20	22	21	23	24

المطلوب :- نظم هذه الدرجات في جدول توزيع تكراري ؟

الحل : نلاحظ أن المتغير العشوائي (درجات الامتحان) من النوع المتقطع .

إن اصغر قيمة في المجموعة هو العدد 8 , وان اكبر قيمة فيها هو العدد 25 وبذلك فإن المدى الكلي لهذه المجموعة هو

$T.R = X_L - X_s$,

$T.R = 25 - 8 = 17$ المدى الكلي للبيانات

وباستخدام صيغة سترجس في تحديد عدد فئات التوزيع نلاحظ إن

$m = 1+(3.322).(\log.n)$

$m = 1+ (3.322) (\log 40) = 1+ 5.322 = 6.322$

وحيث أن عدد الفئات يجب أن يكون عدد صحيح لذا يتم تقريب قيمة m إلى اقرب عدد صحيح وهو 6 .

$$L = \frac{T.R}{m}$$

$$L = \frac{17}{6} = 2.833 \text{ (بالتقريب) } \approx 3$$ طول الفئة في جدول التوزيع التكراري

لذا واستنادا إلى طول الفئة نبدأ بتحديد حدي كل فئة على النحو التالي :

ت . الفئه	الحد الأدنى للفئه	الحد الأعلى للفئه	الفئات
1	8	8 + 3 - (1)=10	8 - 10
2	8 + 3 =11	8 + 6 - (1)=13	11 - 13
3	8 + 6 =14	8 + 9 - (1)=16	14 - 16
4	8 + 9 =17	8 + 12 - (1)=19	17 - 19
5	8 + 12 =20	8 + 15 - (1)=22	20 - 22
6	8 + 15 =23	8 + 18 - (1)=25	23 - 25

ت . الفئه	الفئات	التكرارات (f)	مركز الفئات (X)
1	8 - 10	2	9
2	11 - 13	7	12
3	14 - 16	10	15
4	17 - 19	8	18
5	20 - 22	6	21
6	23 - 25	7	24
المجموع		40	

وهذا يعني أن اثنين من الطلبة كانت درجاتهم مابين 10 – 8 درجة , وسبعة طلاب كانت درجاتهم ما بين 11ـ13 درجة , وهكذا .

ب- في حالة المتغيرات العشوائية المستمرة :-

في هذه الحالة نكتب حدود الفئات كما هو موضح في التوزيع التكراري التالي بحيث نضمن أن كل قيمة من قيم البيانات تضمن في فئة واحدة من فئات التوزيع دون أي تكرار قد يحصل في هذه الفئة أو تلك , وبحيث أن طول الفئة L يكون مساويا للفرق مابين الحد الأعلى والحد الأدنى .

تسلسل الفئة	الحد الأدنى	اقل من الحد الأعلى
١	x_s (اصغر قيمة)	$L + x_s$ (اصغر قيمة+طول الفئة)
٢	$L + x_s$ (اصغر قيمة + طول الفئة)	$2L + x_s$ (اصغر قيمة+اثنين من طول الفئة)
٣	$2L + x_s$ (اصغر قيمة +اثنين من طول الفئة)	$3L + x_s$ (اصغر قيمة +ثلاثة من طول الفئة)
.	.	.
.	.	.
.	.	.
M	$(m-1)L + x_s$ (اصغرقيمة +(m-1)من طول الفئة)	$(m)L + x_s$ (اصغرقيمة +(m)من طول الفئة)

(مثال ١)

البيانات التالية تمثل أوزان عينه من طلبة إحدى الكليات قوامها (١٠٠ طالب) يطلب تفريغ هذه البيانات في جدول توزيع تكراري , ومن ثم حساب مراكز الفئات ؟

77.2	62.3	47.8	46	55	94	83	70.5	65.3	101
82.4	60.2	73.8	73.2	70	68.3	66.9	66.5	62.9	61.3
78.3	75.1	80.2	74.1	51.8	52.6	58.7	54.4	59.9	58.2
81.5	76.3	80.1	89.1	88.2	67.1	65.2	66.3	62	58.5
67	73.6	72.1	65.3	66.2	64.9	63.1	65	61.1	96.3
88.1	85	79.3	68.1	62.1	55	49.2	60.1	71.2	69.2
59.8	58.1	62.3	69.1	75	84	82	81.3	88.2	95.1
74.5	81.3	79.2	75.1	69.3	65.2	67	66	55.1	52.9
77.8	72.1	68.4	66.2	65.1	63	59.4	58.6	54.8	51.9
59.6	54.8	66.2	65.7	64.8	72.1	73.9	71.2	69.1	78.3

الحل

واضح أن المتغير العشوائي (وزن الطالب /بالكغم) هو من النوع المستمر وعليه سوف نستخدم اسلوب تكوين التوزيع التكراري الخاص بالمتغيرات المستمرة .

إن اصغر قيمة في المجموعة تمثل العدد 46 , وان اكبر قيمة فيها تمثل العدد 101 وعليه فأن المدى الكلي لهذه المجموعة هو

$$T.R = x_L - x_s$$

المدى الكلي للبيانات $T.R = 101 - 46 = 55$

وباستخدام صيغة سترجس في تحديد عدد فئات التوزيع نلاحظ إن

$$m = 1+(3.322).(\log.n)$$

$$m = 1+ (3.322)(\log 100) = 1+ 6.644 = 7.644$$

وحيث أن عدد الفئات يجب أن يكون عدد صحيح لذا يتم تقريب قيمة m إلى اقرب عدد صحيح وهو 8 .

عليه فإن طول الفئة يمكن تحديده وفق الأتي :-

$$L = \frac{T.R}{m}$$

$$L = \frac{55}{8} = 6.875 \quad (\text{بالتقريب}) \approx 7$$ طول الفئة في جدول التوزيع التكراري

لطانواستنادا إلى طول الفئة نبدأ بتحديد حدي كل فئة وعلى النحو التالي :-

ت . الفئة	الحد الأدنى للفئة	الحد الأعلى للفئة(اقل من)	الفئات
1	46	46 + 7 =53	من 46 إلى اقل من 53
2	46 + 7 =53	46 + 14 =60	من 53 إلى اقل من 60
3	46 + 14 =60	46 + 21 =67	من 60 إلى اقل من 67
4	46 + 21 =67	46 + 28 =74	من 67 إلى اقل من 74
5	46 + 28 =74	46 + 35 =81	من 74 إلى اقل من 81
6	46 + 35 =81	46 + 42 =88	من 81 إلى اقل من 88
7	46 + 42 =88	46 + 49 =95	من 88 إلى اقل من 95
8	46 + 49 =95	46 + 56 =102	من 95 إلى اقل من 102

لطانولإغراض السهولة في كتابة الفئات بشكلها النهائي , يفضل اعتماد الشكل التالي :-

ت . الفئة	الفئات	التكرارات (f)	مركز الفئات (X)
1	46 -	7	49.5
2	53 -	15	56.5
3	60 -	27	63.5
4	67 -	21	70.5
5	74 -	14	77.5
6	81 -	8	84.5
7	88 -	5	91.5
8	95 – 102	3	98.5
	المجموع	100	

من هذا الجدول نلاحظ أن 7 طلاب أوزانهم تتراوح مابين 46 كغم وأقل من 53 كغم, وأن 15 طالب تتراوح أوزانهم ما بين 53 كغم وأقل من 60 كغم , وهكذا .

ونلاحظ أيضا أن 7 طلاب أوزانهم بالمتوسط هي 49.5 , وان 15 طالب أوزانهم بالمتوسط 56.5 كغم وهكذا.

(مثال ٢)

البيانات التالية تمثل معدلات مجموعة من الطلبة خريجي الدراسة الإعدادية / الفرع العلمي (٦٧ معدل) مقربة لمرتبة عشية واحدة . يطلب تفريغ هذه البيانات في جدول توزيع تكراري ؟

70.8	68	64.2	53	82.3	85.9	95	79	68.1	55.2
58.3	89	92.1	55	71.3	70.7	89	86	63.2	75.1
63.1	66.2	81.6	82.5	94.1	93	80	64.1	71.3	62.2
86.6	82.2	83.1	86.9	91	55.3	54	69.2	74.2	71.8
81.7	69	62.9	61.3	60.1	87.7	62.5	66.1	76.3	74
80.2	75.3	71.1	76	87	92.1	66.1	60.2	54.1	88
			72.7	70.2	61.3	62.9	58	59.2	60.1

الحل

واضح أن المتغير العشوائي (معدل الطالب) هو من النوع المستمر وعليه سوف نستخدم اسلوب تكوين التوزيع التكراري الخاص بالمتغيرات المستمرة .

إن اصغر قيمة في المجموعة تمثل العدد 53 , وان اكبر قيمة فيها تمثل العدد 95 وعليه فأن المدى الكلي لهذه المجموعة هو

$$T.R = \chi_L - \chi_s$$

$T.R = 95 - 53 = 42$ المدى الكلي للبيانات

وباستخدام صيغة يول في تحديد عدد فئات التوزيع نلاحظ أن

$$m = (2.5) \ . \ \sqrt[4]{n}$$

$$m = (2.5) \ . \ \sqrt[4]{67} \quad = (2.5) * (2.861)$$

$$m = 7.15$$

وحيث أن عدد الفئات يجب أن يكون عدد صحيح لذا يتم تقريب قيمة m إلى اقرب عدد صحيح وهو 7 .

وبذلك فإن طول الفئة يمكن تحديده وفق الآتي

$$L = \frac{T.R}{m}$$

$$L = \frac{42}{7} = 6$$ طول الفئة في جدول التوزيع التكراري

واستنادا إلى طول الفئة نبدأ بتحديد حدي كل فئة وعلى النحو التالي :-

ت . الفئة	الحد الأدنى للفئة	الحد الأعلى للفئة(اقل من)	الفئات
1	53	53 + 6 =59	من 53 إلى اقل من 59
2	53 + 6 =59	53 + 12 =65	من 59 إلى اقل من 65
3	53 + 12 =65	53 + 18 =71	من 65 إلى اقل من 71
4	53 + 18 =71	53 + 24 =77	من 71 إلى اقل من 77
5	53 + 24 =77	53 + 30 =83	من 77 إلى اقل من 83
6	53 + 30 =83	53 + 36 =89	من 83 إلى اقل من 89
7	53 + 36 =89	53 + 42 =95	من 89 إلى اقل من 95

ولإغراض السهولة في كتابة الفئات بشكلها النهائي , يفضل اعتماد الشكل التالي :-

ت . الفئة	الفئات	التكرارات (f)	مركز الفئات (X)
1	53 -	8	56
2	59 -	14	62
3	65 -	10	68
4	71 -	11	74
5	77 -	8	80
6	83 -	8	86
7	89 – 95	8	92
المجموع		67	

من هذا الجدول نلاحظ أن 8 طلاب معدلاتهم تتراوح مابين 53 وأقل من 59 , وأن 14 طالب تتراوح معدلاتهم ما بين 59 وأقل من 65 , وهكذا .

ونلاحظ أيضا أن8 طلاب معدلاتهم بالمتوسط هي 56 , وان 14 طالب معدلاتهم بالمتوسط 62 وهكذا.

٢ – ٢ العرض الهندسي للبيانات

بغية إعطاء فكرة واضحة وسريعة عن البيانات فأنه يتم في أحوال كثيرة عرض هذه البيانات بهيئة رسوم بيانية وأشكال هندسية متعددة الأشكال والتصاميم والبعض منها بهيئة رسوم تصويرية, هذه الرسوم والأشكال **للبيانات الغير مبوبة** هي :-

٢-٢-١ الخط البياني Line – Chart

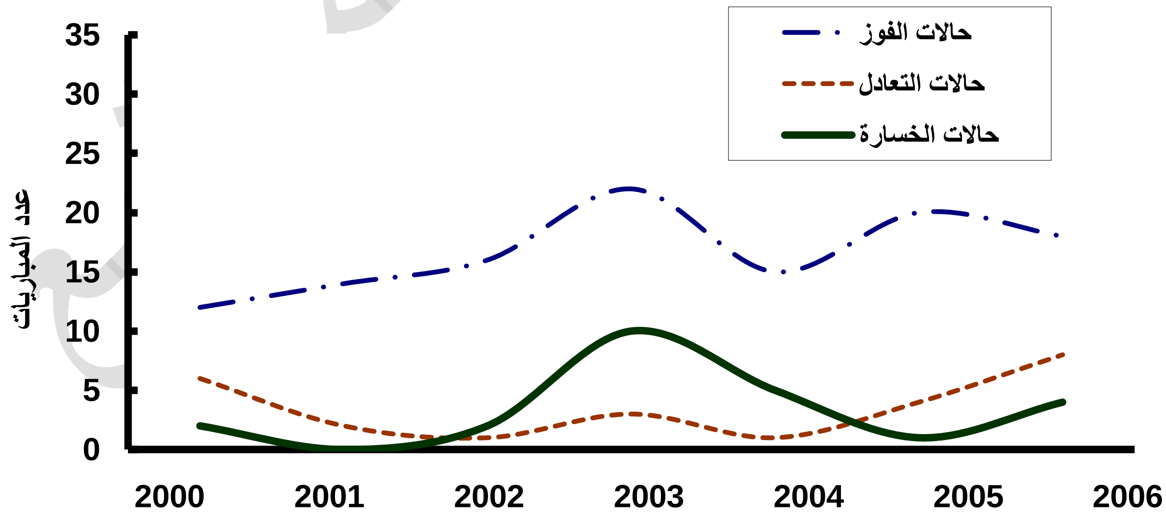
عبارة عن شكل بياني يوضح التغيرات الحاصلة في ظاهرة معينة عبر فترة معينة من الزمن . وهو شكل نافع في حالة إجراء المقارنة بين ظاهرتين أو أكثر , مقاسة بنفس وحدات القياس .

على سبيل المثال مقارنة التغيرات الحاصلة بين كميات النفط المنتجة والمصدرة , مقارنة التغيرات الحاصلة بين تكاليف إنتاج سلعة معينة والأرباح المتحققة من مبيعاتها خلال فترة زمنية معينة .

مثال) الأتي توزيع يمثل عدد مباريات كرة القدم التي خاضها فريق معين وعدد حالات الفوز والتعادل والخسارة التي تحققت من هذه المباريات خلال الفترة ٢٠٠٠-٢٠٠٦ .يطلب تمثيل هذه البيانات بخطوط بيانيه ؟

السنوات	2000	2001	2002	2003	2004	2005	2006
عدد المباريات	20	16	19	35	21	25	30
حالات الفوز	12	14	16	22	15	20	18
حالات التعادل	6	2	1	3	1	4	8
حالات الخسارة	2	0	2	10	5	1	4

الحل :- نحدد ثلاثة أنواع من الخطوط لتمثيل حالات الفوز والخسارة والتعادل وكما يبين الشكل أدناه :-



شكل (٢-١) توزيع عدد المباريات التي خاضها فريق معين بكرة القدم ونتائج هذه المباريات للفترة من سنة (٢٠٠٠- ٢٠٠٦)

٢-٢-٢ الأشرطة البيانية Bar – Chart

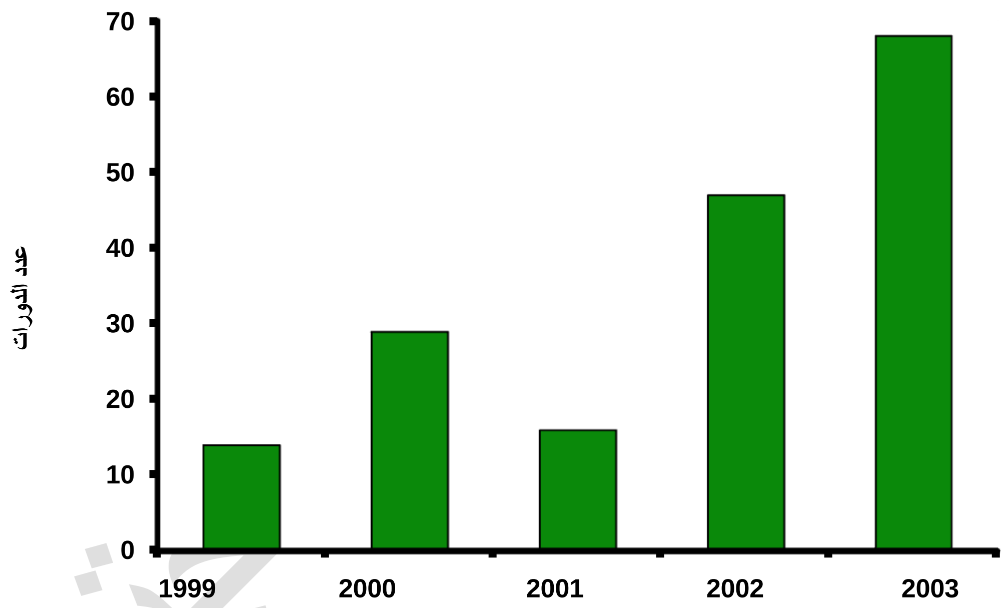
عبارة عن مجموعة من المستطيلات الرأسية أو الأفقية قواعدها متساوية وتمثل الصفة التي تم على أساسها التبويب (سنة , شهر , محافظة , صنف دم , ...الخ) وارتفاعاتها تمثل البيانات المقابلة لتلك الصفة (درجات الحرارة , كميات محصول الحنطة حسب المحافظات , عدد المرضى حسب صنف الدم , ... الخ).

(مثال) بلغ عدد الدورات التدريبية التي نفذت من قبل كليات إحدى الجامعات للعاملين في مؤسسات الدولة للفترة (١٩٩٩-٢٠٠٣) كما يلي :-

السنة	1999	2000	2001	2002	2003
عدد الدورات	14	29	16	47	68

المطلوب تمثيل هذه البيانات بأشرطة بيانية ؟

الحل : نختار قاعدة لكل مستطيل (شريط) بطول مناسب , ونقسم المحور العمودي على نحو ملائم للبيانات (عدد الدورات) وكما موضح أدناه :-



شكل (٢-٢) توزيع عدد الدورات التدريبية المنفذة من قبل كليات جامعة معينه للفترة من سنة (١٩٩٩-٢٠٠٣)

ملاحظة :- هنالك نوع آخر من الأشرطة البيانية تدعى "الأشرطة البيانية المركبة" تخص صنفين أو أكثر من البيانات , مثل عدد الطلبة المقبولين في التعليم العالي للسنوات (٢٠٠٠-٢٠٠٥) مصنفين حسب جنسهم (ذكر, أنثى) ,عدد سكان العراق حسب التعدادات السكانية المنفذة مصنفين حسب الحالة الاجتماعية (أعزب, متزوج, أرمل , مطلق).

٣-٢-٢ الدائرة البيانية Bie – Chart

وهي عبارة عن شكل هندسي يستخدم لتمثيل بيانات ظاهرة معينة يمكن تجزئتها إلى عدد من الأصناف مثل عدد الطلبة موزعين حسب المراحل الدراسية , تكاليف إنتاج سلعة معينة .وفكرة هذا الشكل هو اختيار دائرة ذات قطر مناسب , هذه الدائرة تمثل مجموع البيانات الكلية , أما تصنيفات هذه البيانات فيتم تمثيلها بقطاعات داخل هذه الدائرة بحيث أن مجموع مساحات القطاعات تمثل مساحة الدائرة . وبهدف تحديد كل قطاع فانه يتوجب تحديد زاوية كل منها وفق ما يلي :-

$$\text{زاوية القطاع} = \frac{\text{عدد بيانات الصنف}}{\text{مجموع البيانات الكلية}} \times ٣٦٠$$

مثال) بلغ عدد الطلبة في إحدى الكليات 2000 طالب وطالبه , منهم 800في الصف الأول , 500 طالب في الصف الثاني , 400 في الصف الثالث , 300 في الصف الرابع . يطلب تمثيل هذه البيانات بدائرة بيانيه .

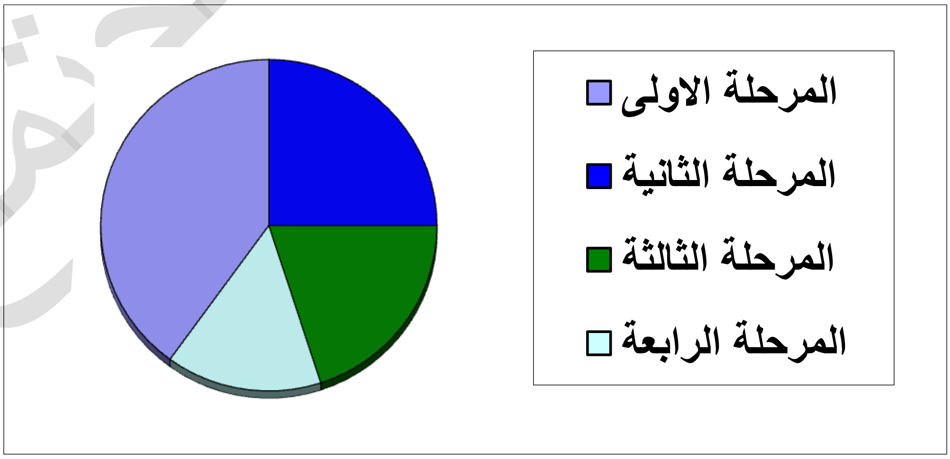
الحل : نختار دائرة ذات قطر مناسب كأن يكون 10 سم . ونبدأ بتحديد زاوية كل قطاع الذي يمثل مرحلة معينة .

$$\text{زاوية قطاع المرحلة الأولى} = \frac{800}{2000} \times 360 = 144$$

$$\text{زاوية قطاع المرحلة الثانية} = \frac{500}{2000} \times 360 = 90$$

$$\text{زاوية قطاع المرحلة الثالثة} = \frac{400}{2000} \times 360 = 72$$

$$\text{زاوية قطاع المرحلة الرابعة} = \frac{300}{2000} \times 360 = 54$$



شكل (٣-٢) توزيع عدد الطلبة حسب المراحل في إحدى كليات جامعة معينة

٢-٢-٤ المستطيل البياني Rectangular – Chart

عبارة عن شكل هندسي يستخدم لتمثيل بيانات ظاهرة معينة يمكن تجزئتها إلى عدد من الأصناف مثل عدد الطلبة موزعين حسب المراحل الدراسية , تكاليف إنتاج سلعة معينة .وفكرة هذا الشكل هو اختيار مستطيل ذو قاعدة مناسبة , هذه المستطيل يمثل مجموع البيانات الكلية , وبعد ذلك يتم تمثيل كل صنف من البيانات بمستطيل جزئي داخل المستطيل الكبير بحيث إن مجموع مساحات المستطيلات الجزئية تمثل مساحة المستطيل الكبير.

وتتم عملية اختيار المستطيلات الجزئية وفق ما يلي :-

طول قاعدة المستطيل الجزئي = (عدد بيانات الصنف ÷ مجموع البيانات الكلية) X طول قاعدة المستطيل الكبير

(مثال)

بلغ عدد الطلبة في إحدى الكليات 2000 طالب وطالبه , منهم 800في الصف الأول , 500 طالب في الصف الثاني , 400 في الصف الثالث , 300 في الصف الرابع . يطلب تمثيل هذه البيانات بمستطيل بياني .

الحل :

نختار مستطيل ذو قاعدة مساوية إلى 10 cm .ونبدأ بتحديد طول قاعدة كل مستطيل جزئي الخاص بذلك الصف .

قاعدة مستطيل المرحلة الأولى = $\frac{800}{2000}$ x 10 سم = 4 cm

قاعدة مستطيل المرحلة الثانية = $\frac{500}{2000}$ x 10 سم = 2.5 cm

قاعدة مستطيل المرحلة الثالثة = $\frac{400}{2000}$ x 10 سم = 2 cm

قاعدة مستطيل المرحلة الرابعة = $\frac{300}{2000}$ x 10 سم = 1.5 cm

شكل (٢-٤) المستطيل البياني لتوزيع عدد الطلبة حسب المراحل في إحدى كليات جامعة معينة

أما الرسوم والأشكال **للبيانات المبوبة** فهي :-

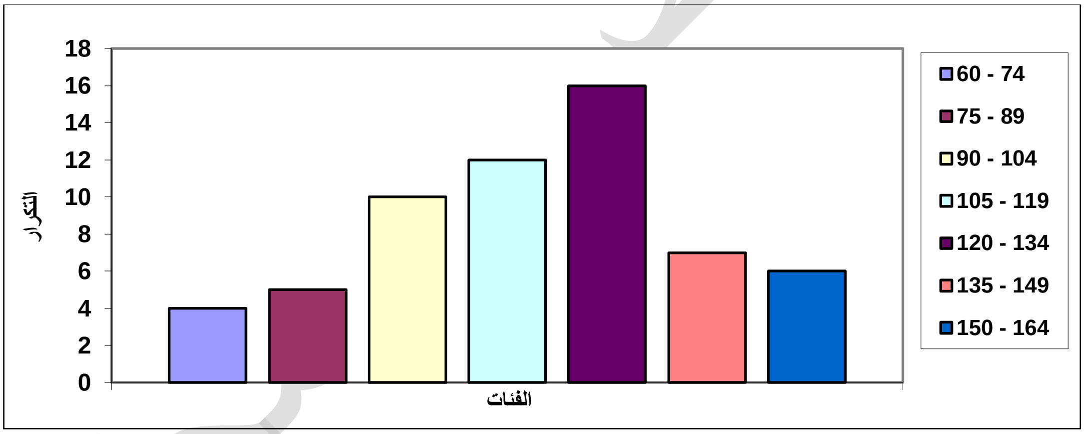
٢-٥ المدرج التكراري Histogram

عبارة عن مجموعة من المستطيلات قاعدة كل منها تمثل طول الفئة في التوزيع التكراري وارتفاع كل منها يمثل قيمة التكرار المقابل لتلك الفئة . هذه المستطيلات تكون **منفصلة** في حالة المتغيرات المتقطعة وتكون **متصلة مع بعضها** في حالة المتغيرات المستمرة .

مثال ١) :- البيانات التالية تمثل عدد المراجعين المرضى تم تسجيلها في مراكز صحية عددها (٦٠ مركز صحي) . يطلب رسم المدرج التكراري لهذا التوزيع ؟

الفئات	التكرارات (f)
60 - 74	4
75 - 89	5
90 - 104	10
105 - 119	12
120 - 134	16
135 - 149	7
150 - 164	6

الحل :- واضح أن المتغير العشوائي من النوع المتقطع , وعليه فأن المدرج التكراري سيكون كالأتي :-

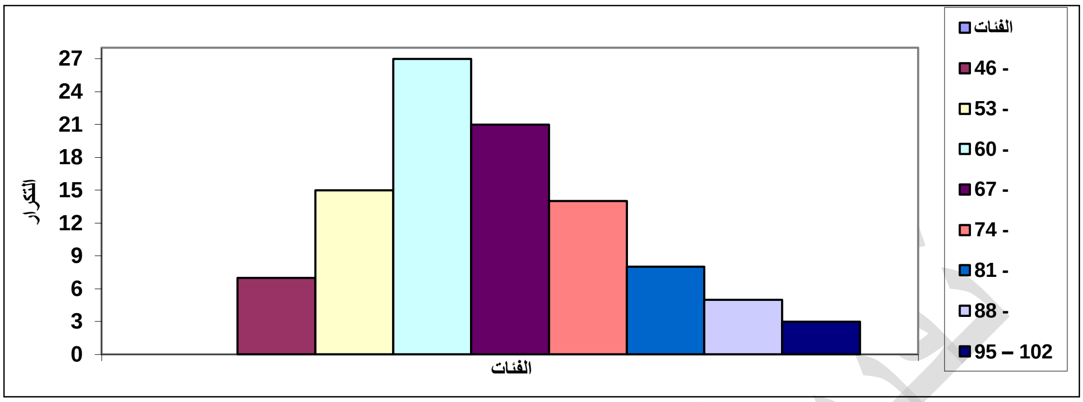


شكل (٢-٥) المدرج التكراري لعدد المراجعين المرضى (متغير متقطع)

مثال ٢) :- الأتي توزيع تكراري لأوزان عينه من طلبة إحدى الكليات قوامها ١٠٠ طالب . يطلب رسم المدرج التكراري لهذا التوزيع ؟

الفئات	التكرارات (f)
46 -	7
53 -	15
60 -	27
67 -	21
74 -	14
81 -	8
88 -	5
95 – 102	3

الحل :ـ واضح أن المتغير العشوائي من النوع **المستمر** , وعليه فأن المدرج التكراري سيكون كالأتي :ـ



شكل (٦-٢) المدرج التكراري لأوزان عينه من الطلبة قوامها ١٠٠ طالب(**متغير مستمر**)

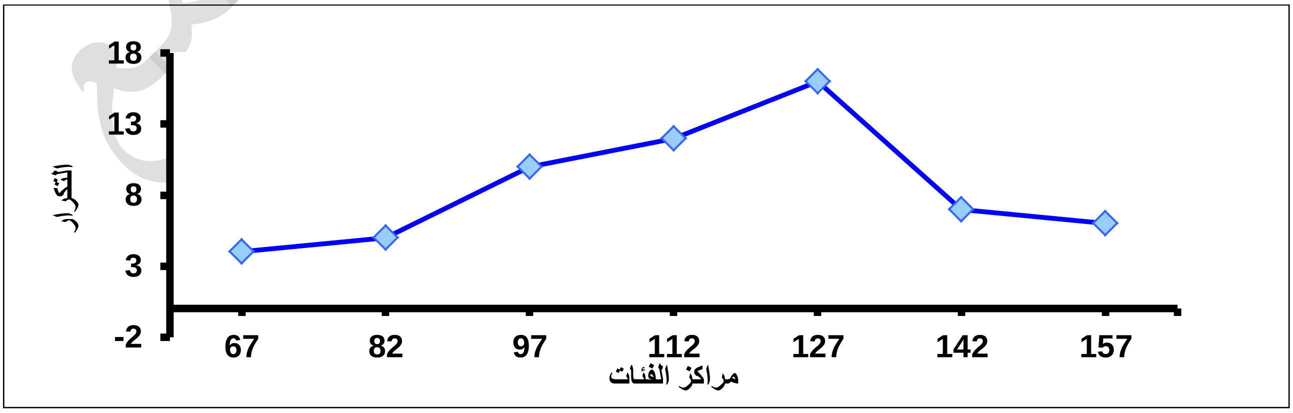
٢-٢-٦ المضلع التكراري Frequency polygon

عبارة عن عدد من المستقيمات المتصلة مع بعضها على شكل سلسلة ونقطة اتصال المستقيم بالآخر تقابل مركز الفئة . وهذا يعني انه عند رسم مضلع تكراري يستوجب الأمر إيجاد مراكز الفئات ومن ثم رسم المضلع على أساس أزواج القيم (مركز الفئة,التكرار) .

(**مثال**) :ـ البيانات التالية تمثل عدد المراجعين المرضى تم تسجيلها في مراكز صحية عددها (٦٠ مركز صحي) . يطلب رسم المضلع التكراري لهذا التوزيع ؟

مركز الفئات (X)	التكرارات (f)	الفئات
67	4	60 - 74
82	5	75 - 89
97	10	90 - 104
112	12	105 - 119
127	16	120 - 134
142	7	135 - 149
157	6	150 - 164

الحل :ـ



شكل (٢-٧) مضلع تكراري يوضح توزيع المراجعين المرضى للمراكز الصحية

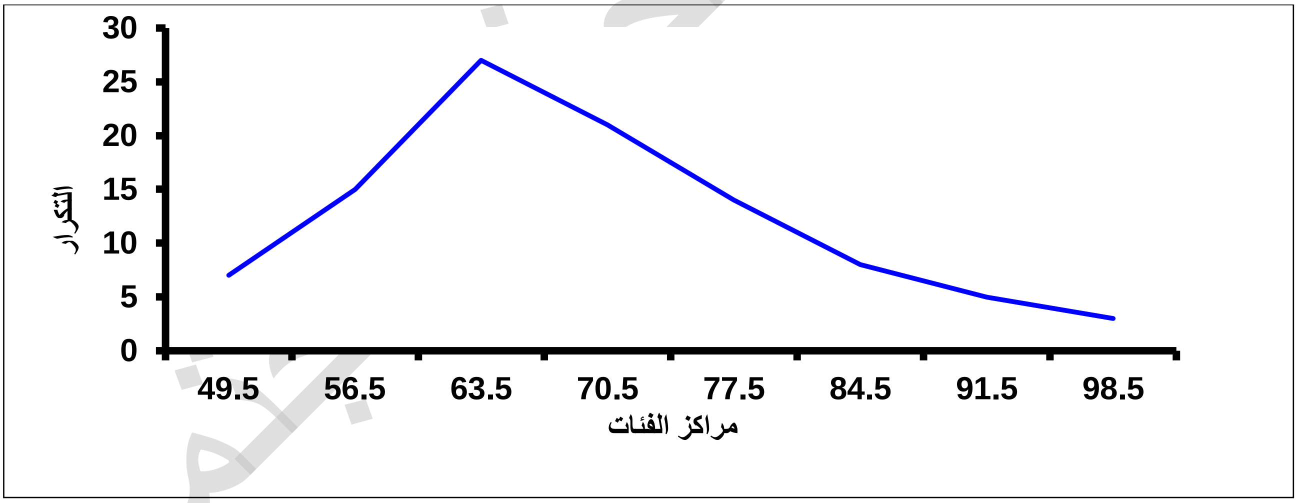
٢-٢-٧ المنحني التكراري Frequency curve

لا تختلف فكرة رسم المنحني التكراري عن المضلع التكراري من حيث الأسلوب ولكن الفرق الوحيد بينهما هو انه بدلا من توصيل النقاط (مركز الفئة , التكرار) بمستقيمات فأنه يتم تمرير منحنى مابين هذه النقاط هذا المنحنى يمثل المنحنى التكراري للتوزيع . إن المنحنى التكراري يتم رسمه للتوزيعات التكرارية الخاصة بالمتغير من النوع المستمر فقط وذلك لاعتبارات تتعلق بموضوع " الاستمرارية " في حساب التفاضل والتكامل .

(**مثال**) جدول التوزيع التكراري التالي يمثل توزيع أوزان عينه من طلبة إحدى الكليات قوامها (١٠٠ طالب) , يطلب رسم المنحني التكراري لهذا التوزيع ؟

الفئات	التكرارات (f)	مركز الفئات (X)
46 -	7	49.5
53 -	15	56.5
60 -	27	63.5
67 -	21	70.5
74 -	14	77.5
81 -	8	84.5
88 -	5	91.5
95 – 102	3	98.5

الحل :-



شكل (٢-٨) المنحني التكراري يوضح توزيع أوزان ١٠٠ طالب في إحدى الكليات

الفصل الثالث

مقاييس النزعة المركزية

الفصل الثالث

مقاييس النزعة المركزية

لاحظنا في الفصلين الأول والثاني استعراض لأهم أساليب جمع وتصنيف وتبويب البيانات وكيفية تمثيلها في جداول ورسوم هندسية وبيانية الهدف من ذلك إعطاء صورة سريعة توضح ماهية هذه البيانات , في هذا الفصل سوف نتطرق للحديث عن كيفية تمثيل مجموعة من البيانات بقيمة واحدة فقط من خلال مقياس يدعى " مقياس نزعة مركزية " .

مفهوم المتوسطات والهدف من احتسابها

يمكن تمثيل مجموعة من البيانات بقيمة واحدة فقط الهدف من ذلك إعطاء صورة واضحة سريعة عن ماهية تلك المجموعة من خلال إيجاد عدد يمثل قيمها . إن العدد الذي يختص بتحديد هذا العدد يسمى مقياس نزعة مركزية أو مقياس متوسط , هذا العدد يميل لأن يقع في وسط تلك المجموعة من البيانات في حال ترتيبها حسب صغرها أو كبرها وهذا ما جعلنا نطلق على هذا النوع من المقاييس (مقاييس نزعة مركزية) .

أهمية المتوسطات تتوضح في موضوع الاستدلال الإحصائي من خلال تقدير قيم عددية لبعض مؤشرات المجتمع تحت الدراسة والبحث التي غالبا ما تكون غير معلومة , أي دراسة خصائص مجتمع ما من خلال خصائص العينة وان المتوسطات هي إحدى هذه الخصائص .

وفيما يلي أهم مقاييس النزعة المركزية

٣-١ الوسط الحسابي Arithmetic Mean

ويسمى في بعض الأحيان الوسط أو المتوسط أو المعدل الحسابي وهو احد أهم مقاييس النزعة المركزية على الإطلاق لما يمتاز به من خصائص جيده وسهولة في الحساب وهو متداول كثيرا في حياتنا اليومية , فمثلا عند تخرج طالب من الدراسة الإعدادية فأن أول ما يسأل عنه هو المعدل "ما معدلك ؟" فيجيب إن معدلي هو 85 على سبيل المثال والذي هو في الحقيقة هو محصلة لسبعة أعداد تمثل درجات هذا الطالب في دروس الصف السادس الإعدادي مقسومة على عددها (7) وهذا يعني إننا تمكنا من تمثيل مجموعة من الأعداد بقيمة واحدة هي معدل درجاته .

طرق حساب الوسط الحسابي

أولا - حساب الوسط الحسابي لبيانات غير مبوبة

لتكن $X1,X2,\dots,Xn$ تمثل البيانات المستحصل عليها من المتغير العشوائي X على أساس عينه من المفردات قوامها n مشاهدة , يرمز عادة للوسط الحسابي بالرمز ($\bar{x}$) ويعرف على انه دالة بدلالة قياسات مفردات العينة , أي أن

$$\bar{X} = f(X_1, X_2, \dots, X_n)$$

بحيث أن هذه الدالة تأخذ الشكل التالي :-

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

ملاحظه :- الوسط الحسابي ($\bar{x}$) هو تقدير للوسط الحسابي لقياسات مفردات المجتمع الذي اختيرت منه هذه العينة , حيث أن متوسط قياسات مفردا المجتمع غالبا ما يرمز له بالرمز (μ) ويعطى بالصيغة التالية :

$$\mu = \frac{\sum_{i=1}^{N} x_i}{N}$$

حيث أن N عدد مفردات المجتمع

....

مثال (١) البيانات التالية تمثل أوزان عينه من الطلبة قوامها (١٥) طالب , يطلب إيجاد متوسط وزن الطالب في هذه العينه .

50.2	60.9	68.3	59.2	58.1	62.3	65.3	52.9
61.5	63.2	59.1	69.3	64.2	65.2	56.6	

الحل :-

$$\bar{x} = \frac{\sum_{i=1}^{n} xi}{n} = \frac{50.2 + 60.9 + 68.3 + ... + 56.6}{15} = 61.087\,\text{kg}$$

....

مثال (٢) البيانات التالية تمثل عدد أفراد عينه من الأسر قوامها 12 أسرة (بضمنها الوالدين) . يطلب إيجاد متوسط عدد أفراد الأسرة ؟

3	4	7	8	10	9	2	5	6	9	7	5

الحل :-

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{3+4+7+8+10+9+2+5+6+9+7+5}{12} = 6.25$$

وحيث إن عدد أفراد الأسرة متغير متقطع لذا يتم تقريب الناتج إلى اقرب عدد صحيح لأنه لايوجد تقييس لجزء من الفرد . وعليه فان متوسط عدد أفراد الأسرة في هذه العينة هو تقريبا ستة أفراد .

ثانيا - حساب الوسط الحسابي لبيانات مبوبة

لتكن X1,X2,…,Xm تمثل مراكز فئات توزيع تكراري عدد فئاته m , وان f1,f2,…,fm تمثل التكرارات المقابلة لهذه الفئات عندئذٍ يحسب الوسط الحسابي لهذا التوزيع وفق مايلي :-

$$\overline{x} = \frac{\sum_{i=1}^{m} fi.xi}{\sum_{i=1}^{m} fi}$$

حيث أن :- $\sum_{i=1}^{m} fi.xi$ تمثل مجموع حاصل ضرب تكرار كل فئة في مركز هذه الفئة .

$\sum_{i=1}^{m} fi$ تمثل مجموع تكرارات الفئات .

<u>مثال (١)</u> :- الأتي توزيع تكراري لدرجات الحرارة في مدينة معينة المسجلة لمدة ٩٥ يوما متتاليا . يطلب حساب متوسط الحرارة في هذه المدينة خلال هذه الفترة .

الفئات	التكرارات (fi)	مركز الفئات (Xi)	Fi . xi
0 –	4	0.5	2
1 -	8	1.5	12
2 -	12	2.5	30
3 -	16	3.5	56
4 -	20	4.5	90
5 -	25	5.5	137.5
6 -	6	6.5	39
7 - 8	4	7.5	30
المجموع			396.5

الحل :-

$$\overline{x} = \frac{\sum_{i=1}^{m} fi.xi}{\sum_{i=1}^{m} fi} = \frac{396.5}{95} = 4.174$$ درجة مئوية

<u>مثال (٢)</u> :- الأتي توزيع تكراري لعينة من الأسر قوامها 75 أسرة حسب عدد , أفراد الأسرة (بضمنها الوالدين) .

يطلب حساب متوسط عدد أفراد الأسرة في هذه العينة ؟

عدد الأفراد (الفئات)	2 - 4	5 - 7	8 - 10	11 - 13	14 - 16	17 - 19	20 - 22
عدد الأسر (التكرارات fi)	8	12	20	13	10	8	4

الحل :- نعمل الجدول التالي :-

الفئات	التكرارات (fi)	مركز الفئات (Xi)	fi.Xi
2 - 4	8	3	24
5 - 7	12	6	72
8 - 10	20	9	180
11 - 13	13	12	156
14 - 16	10	15	150
17 - 19	8	18	144
20 - 22	4	21	84
المجموع	75	-	810

$$\overline{X} = \frac{\sum_{i=1}^{m} fi.xi}{\sum_{i=1}^{m} fi} = \frac{810}{75} = 10.8 \text{ فرد}$$

وحيث إن عدد أفراد الأسرة متغير من النوع المتقطع وانه لا يوجد تقييس لجزء من الفرد عليه يتم تقريب الناتج لأقرب عدد صحيح , وعليه فان متوسط عدد أفراد الأسرة في هذه العينة هو تقريبا احد عشر فردا .

٢.٣ الوسيط Median

يعتبر الوسيط احد مقاييس النزعة المركزية المهمة في التطبيقات الإحصائية , ويعرف الوسيط بأنه تلك القيمة من قيم المتغير العشوائي X التي تقسم مجموعة قيم المتغير إلى قسمين متساويين . أي أنها قيمة X التي تجعل عدد القيم قبلها مساوٍ لعدد القيم بعدها .

طرق حساب الوسيط (حسابيا)

أولا- حساب الوسيط لبيانات غير مبوبة

لتكن X1,X2,...,Xn تمثل قياسات مفردات عينه قوامها n مفردة . وافترض أن هذه القياسات رتبت على نحو تصاعدي , وافرض أن $y_1<y_2<y_3<\dots<y_n$ تمثل قيم x المرتبة تصاعديا (وقد يكون الترتيب تنازليا) عندئذٍ :

١- إذا كان عدد القيم n فردي عندئذٍ فأن قيمة الوسيط تمثل قيمة x بعد الترتيب (أي y) التي تسلسلها هو

$$Me = \frac{n+1}{2}$$

مثال (١) :ـ الأتي درجات عينه من الطلبة قوامها تسعة طلاب في امتحان معين , جد الوسيط لهذه المجموعة .

55 ,62 ,53 ,70 ,68 ,65 ,63 ,79 ,80

الحل : نرتب هذه القيم ترتيب تصاعدي وكما يلي :

53 ,55 ,62 ,63 ,65 ,68 ,70 ,79 ,80

وعليه فأن ترتيب الوسيط هو

$$Me = \frac{n+1}{2} = \frac{9+1}{2} = 5$$

وهذا يعني أن القيمة الخامسة هي قيمة الوسيط أي الدرجة 65

٢ـ إذا كان عدد القيم n زوجي عندئذٍ فأن قيمة الوسيط تمثل الوسط الحسابي لقيمتي x بعد الترتيب

اللتين تسلسلهما على التوالي هو($\frac{n}{2}, \frac{n}{2}+1$) .

مثال (١) :ـ الأتي أعمار عينه من الأفراد قوامها 12 فرد . جد الوسيط لعمر الفرد في هذه المجموعة .

20 ,22 ,19.5 ,26 ,24.5 ,27 ,28 ,29 ,18 ,20 ,23 ,25

الحل : نرتب هذه القيم وفق ترتيب تنازلي وكما يلي

29 ,28 ,27 ,26 ,25 ,24.5 ,23 ,22 ,20 , 20 ,19.5 ,18

وعندئذٍ فان القيمتان اللتان تحددان الوسيط هما (23, 24.5) اللتين تسلسلهما على التوالي هو

($\frac{12}{2}=6$, $\frac{12}{2}+1=7$) أي القيمة السادسة والسابعة بعد الترتيب . وبذلك فأن الوسيط لهذه

المجموعة يمثل الوسط الحسابي لهاتين القيمتين أي :ـ

$$Me = \frac{24.5+23}{2} = 23.75$$ سنه

ثانيا – إيجاد الوسيط لبيانات مبوبة

١ـ إيجاد الوسيط لبيانات مبوبة لمتغير متقطع

افرض وجود توزيع تكراري لبيانات متغير متقطع عدد فئاته m وان f1,f2,...,fm تمثل التكرارات المقابلة لفئات هذا التوزيع . وافرض أن F1,F2,...,Fm تمثل التكرارات المتجمعة الصاعدة المقابلة للحدود العليا لفئات التوزيع . من تعريف الوسيط هو القيمة التي تقسم مجموعة القيم إلى قسمين متساويين وهذا يعني أن

تسلسل (ترتيب) الوسيط في حالة التوزيعات التكرارية يمثل نصف التكرارات أي ($\frac{\sum_{i=1}^{m} fi}{2}$) أي أن الوسيط هنا يمثل القيمة التي تترك نصف مجموع التكرارات قبلها والنصف الأخر بعدها .

ولغرض تحديد الوسيط نقوم بمقارنة ترتيب الوسيط مع التكرار المتجمع الصاعد فإذا كان

$$F_K \langle \frac{\sum fi}{2} \langle \ F_{K+1}$$

, F_{K+1} يمثل التكرار المتجمع الصاعد المقابل للفئة K+1

F_K يمثل التكرار المتجمع الصاعد المقابل للفئة K

عندئذٍ يقال أن فئة الوسيط هي الفئة التي تسلسلها هو (K+1) وبذلك فان قيمة الوسيط تمثل مركز هذه الفئة .

مثال (١) :- الأتي توزيع تكراري لعينة من الأسر حسب عدد أفراد الأسرة يطلب حساب الوسيط لعدد الأفراد .

عدد الأفراد (الفئات)	2 - 4	5 - 7	8 - 10	11 - 13	14 - 16	17 - 19	20 - 22
عدد الأسر (التكرارات fi)	6	9	12	20	14	11	8

الحل :- نقوم بعمل الجدول التالي :-

الفئات	التكرار	الحدود العليا للفئات	التكرار المجمع الصاعد
2 - 4	6	4	6
5 – 7	9	7	15
8 - 10	12	10	27
11 – 13	20	13	47
14 – 16	14	16	61
17 – 19	11	19	72
20 - 22	8	22	80

إن ترتيب الوسيط هو $\frac{\sum fi}{2} = \frac{80}{2} = 40$

وبملاحظة ترتيب الوسيط ضمن التكرار المتجمع الصاعد نلاحظ أن

$$F_K \langle \frac{\sum fi}{2} \langle \ F_{K+1} \Longrightarrow 27 < 40 < 47$$

وعليه فأن فئة الوسيط هي الفئة الرابعة من التوزيع أي (13- 11) وعندئذ فأن الوسيط لهذا التوزيع يمثل مركز هذه الفئة أي

$$X_4 = \frac{11+13}{2} = \frac{24}{2} = 12$$ قيمة الوسيط مساوية لمركز الفئة الرابعة(Me= 12)

مثال (٢) :ـ الأتي توزيع تكراري لعدد النداءات الهاتفية العاجلة التي استقبلتها بدالة مستشفى خلال فترة شهر واحد على أساس عدد النداءات المستقبلة خلال ساعة واحدة . جد الوسيط لعدد النداءات ؟

عدد النداءات (خلال ساعة)	0 - 4	5 - 9	10 - 14	15 - 19	20 - 24
(التكرارات fi)	12	36	56	51	45

الحل :ـ نقوم بعمل الجدول التالي :ـ

الفئات	التكرار	الحدود العليا للفئات	التكرار المجمع الصاعد
0 - 4	12	4	12
5 – 9	36	9	48
10 - 14	56	14	104
15 – 19	51	19	155
20 – 24	45	24	200

إن ترتيب الوسيط هو $\frac{\sum fi}{2} = \frac{200}{2} = 100$

وبملاحظة ترتيب الوسيط ضمن التكرار المتجمع الصاعد نلاحظ أن

$$F_K \langle \frac{\sum fi}{2} \langle F_{K+1} \Longrightarrow 48 < 100 < 104$$

وعليه فأن فئة الوسيط هي الفئة الثالثة من التوزيع أي (10 -14) وعندئذ فأن الوسيط لهذا التوزيع يمثل مركز هذه الفئة أي

$$X_4 = \frac{10+14}{2} = \frac{24}{2} = 12$$ قيمة الوسيط مساوية لمركز الفئة الثالثة(Me= 12)

٢ـ إيجاد الوسيط لبيانات مبوبه لمتغير مستمر

افرض وجود توزيع تكراري لبيانات متغير مستمر عدد فئاته m وان f1,f2,…,fm تمثل التكرارات المقابلة لفئات هذا التوزيع.وافرض أن F1,F2,…,Fm تمثل التكرارات المتجمعة الصاعدة المقابلة للحدود العليا لفئات التوزيع . ليكن ($\sum_{i=1}^{m} fi \;/\; 2$) يمثل ترتيب الوسيط في هذا التوزيع . فإذا كان

F_{K-1} يمثل التكرار المتجمع الصاعد المقابل للفئة K-1 , $F_{K-1} \langle \frac{\sum_{i=1}^{m} fi}{2} \langle F_K$

F_K يمثل التكرار المتجمع الصاعد المقابل للفئة K

عندئذٍ يقال أن فئة الوسيط هي الفئة التي تسلسلها هو k وعندئذٍ يتم حساب قيمة الوسيط وفق الصيغة التالية :-

$$Me = L_K + \frac{h_k}{f_k}\left(\frac{\sum_{i=1}^{m} fi}{2} - F_{K-1}\right)$$

حيث أن :- L_K : الحد الأدنى لفئة الوسيط , f_k : تـكـرار فئـة الـوسـيـط .

h_k : طـول فئـة الـوسـيـط . , F_{k-1} : التكرار المتجمع الصاعد السابق لفئة الوسيط .

مثال (١) :- الأتي توزيع تكراري للدخل الشهري (بآلاف الدنانير) لعينة من الأسر قوامها 80 أسرة جد الوسيط للدخل الشهري للأسرة في هذه العينة .

الفئات	100 -	120 -	140 -	160 -	180 -	200 -	220 - 240
عدد الأسر	3	7	14	20	18	12	6

الحل :- نقوم بعمل الجدول التالي :-

الفئات	التكرار	الحدود العليا للفئات	التكرار المجمع الصاعد
100 -	3	اقل من 120	3
120 -	7	اقل من 140	10
140 -	14	اقل من 160	24
160 -	20	اقل من 180	44
180 -	18	اقل من 200	62
200 -	12	اقل من 220	74
220-240	6	اقل من 240	80

ثم نجد ترتيب الوسيط وهذا مساوٍ إلى نصف التكرارات أي $40 = \frac{80}{2} = \frac{\sum fi}{2}$.

وبملاحظة ترتيب الوسيط ضمن التكرار المتجمع الصاعد نلاحظ أن

$$F_{K-1} \langle \frac{\sum fi}{2} \langle \; F_K \quad \Longrightarrow \quad 24< 40 < 44$$

وعليه فان فئة الوسيط عي الفئة الربعة من التوزيع أي الفئة (180 – 160) وبذلك فأن :

$L4 = 160$, $h4 = 20$, $f4 = 20$, $F3 = 24$

وعندئذٍ فأن :-

$$Me = L_K + \frac{h_k}{f_k}\left(\frac{\sum_{i=1}^{m} fi}{2} - F_{K-1}\right)$$

$$Me = 160 + \frac{20}{20}(40-24) = 160+16 = 176$$ ألف دينار

مثال (٢) :- الأتي توزيع تكراري لأعمار عينه من تلاميذ إحدى المدارس الابتدائية قوامها 90 تلميذ . جد الوسيط لعمر التلميذ في هذه العينة

فئات العمر	5 -	6 -	7 -	8 -	9 -	10 -	11 -	12 -
عدد التلاميذ	3	6	9	12	20	18	17	5

الحل : - نجد التوزيع التكراري المتجمع الصاعد وكما موضح بالجدول أدناه :-

الفئات	التكرار	الحدود العليا للفئات	التكرار المجمع الصاعد
5 -	3	اقل من 6	3
6 -	6	اقل من 7	9
7 -	9	اقل من 8	18
8 -	12	اقل من 9	30
9 -	20	اقل من 10	50
10 -	18	اقل من 11	68
11 -	17	اقل من 12	85
12 - فأكثر	5	اقل من مجهول	90

ثم نجد ترتيب الوسيط وهذا مساوٍ إلى نصف التكرارات أي $\frac{\sum fi}{2} = \frac{90}{2} = 45$.

وبملاحظة ترتيب الوسيط ضمن التكرار المتجمع الصاعد نلاحظ أن

$$F_{K-1} \langle \frac{\sum fi}{2} \langle \ F_K \qquad \Longrightarrow \qquad 30 < 45 < 50$$

وعليه فان فئة الوسيط عي الفئة الخامسة من التوزيع أي الفئة (– 9) وبذلك فأن :

$L5 = 9$, $h5 = 1$, $f5 = 20$, $F4 = 30$

وعندئذٍ فأن :ـ

$$Me = L_K + \frac{h_k}{f_k}\left(\frac{\sum_{i=1}^{m} fi}{2} - F_{K-1}\right)$$

$$Me = 9 + \frac{1}{20}(45-30) = 9 + \frac{15}{20} = 9.75$$

٣ـ٣ المنوال The Mode

يعرف المنوال بأنه تلك القيمة التي تتكرر أكثر من غيرها من بين مجموعة من القيم ,أو أنها القيمة الشائعة من بين مجموعه من القيم.

أـ حساب المنوال لبيانات غير مبوبة

لتكن $X1,X2,\ldots,Xn$ تمثل قياسات مفردات عينه قوامها n مفردة.وافترض أن Xj قيمه من قيم هذه المجموعة لوحظ أنها تكررت أكثر من غيرها ,عندئذٍ وحسب تعريف المنوال فإن Xj تمثل المنوال لهذه المجموعة .

<u>مثال (١)</u> :ـ للبيانات التالية جد المنوال

2,3,2,4,2,5,4,4,5,4,6,8,9,4,7,3,7,6

الحل : واضح من المجموعة أن العدد 4 قد تكرر خمس مرات وهو اكثر من تكرار أي عدد آخر . عليه فان المنوال لهذه المجموعة هو (Mo= 4) .

<u>مثال (٢)</u> :- جد المنوال للبيانات التالية (2,4,3,6,8,7,10,12)

الحل : واضح من هذه المجموعة انه لا يوجد عدد متكرر أكثر من غيره . وعليه فأنه **<u>لا يوجد منوال</u>** لهذه المجموعة .

<u>ملاحظة</u> :ـ من الممكن وجود أكثر من منوال واحد لمجموعة من البيانات في حالة تساوي التكرار .

ب – إيجاد المنوال لبيانات مبوبة

١ـ إيجاد المنوال لبيانات مبوبة لمتغير متقطع

لتكن X1,X2,...,Xm تمثل مراكز فئات توزيع تكراري عدد فئاته m , وان f1,f2,...,fm تمثل التكرارات المقابلة لفئات هذا التوزيع . عندئذٍ فأن المنوال لهذا التوزيع يمثل قيمة مركز الفئة التي تقابل اكبر تكرار في التوزيع . وفي حالة وجود فئتين أو أكثر يقابلهما نفس التكرار عندئذٍ فان توزيع من هذا النوع سوف يمتلك أكثر من قيمة واحده للمنوال كل منها تمثل مركز الفئة التي تقابل ذلك التكرار .

مثال :ـ الأتي توزيع تكراري توزيع 40 عائلة فلاحيه حسب ملكيتها من أشجار البرتقال .يطلب تحديد المنوال لهذا التوزيع .

الفئات	التكرارات (f)
60 - 74	2
75 - 89	6
90 - 104	14
105 - 119	10
120 - 134	8

الحل :ـ نبحث عن اكبر تكرار موجود في هذا التوزيع وهو العدد 14 المقابل للفئة الثالثة (90-104) . وعليه فأن المنوال في هذا التوزيع هو مركز الفئة أي ($X_3 = \frac{90+104}{2} = \frac{194}{2} = 97$) 97 شجرة برتقال .

٢ـ إيجاد المنوال لبيانات مبوبة لمتغير مستمر ـ طريقة بيرسون (حسابيا)

افرض وجود توزيع تكراري لبيانات متغير مستمر عدد فئاته m وأفرض أن (f_k) يمثل اكبر تكرار في هذا التوزيع .وهذا يعني أن الفئة التي تحتوي المنوال هي الفئة المقابلة إلى f_k . وافرض أن (f_{k-1}) يمثل التكرار السابق لتكرار فئة المنوال وان (f_{k+1})يمثل التكرار اللاحق لتكرار فئة المنوال وهذا يعني ($f_{k-1} < f_k < f_{k+1}$) , وان (h_k) تمثل طول فئة المنوال وان(L_k) تمثل الحد الأدنى لفئة المنوال .
عندئذٍ ووفق هذه المعطيات يمكن حساب قيمة المنوال بالصيغة التالية :

$$mo = L_k + \frac{(f_K - f_{k-1})}{(f_k - f_{k-1}) + (f_k - f_{k+1})} \cdot h_k$$

مثال (١) :ـ الأتي توزيع تكراري لأطوال عينه من الأشخاص البالغين قوامها 50 شخص . يطلب حساب القيمة الشائعة لطول الشخص في هذه العينة .

فئات الطول	150 -	160 -	170 -	180 -	190 - 200
عدد الأشخاص	8	12	15	9	6

الحل : واضح أن اكبر تكرار في التوزيع هو 15 عليه فأن فئة المنوال هي الفئة (170 -) أي الفئة الثالثة .

$f_3 = 15 \quad , \quad f_2 = 12 \quad , \quad f_4 = 9 \quad , \quad L_3 = 170 \quad , h_3 = 10$

$$mo = L_k + \frac{(f_K - f_{k-1})}{(f_k - f_{k-1}) + (f_k - f_{k+1})} \cdot h_k$$

$$mo = 170 + \frac{(15-12)}{(15-12)+(15-9)} .10 \quad = 170 + \frac{30}{9} = 173.33$$ سنتمتر

مثال(٢) :- الأتي توزيع تكراري لأعمار عدد من المرضى الراقدين في إحدى المستشفيات. يطلب إيجاد العمر الشائع للمريض في هذه المجموعة.

فئات العمر	10 -	20 -	30 -	40 -	50 -	60 - فأكثر
عدد المرضى	2	8	16	17	23	14

الحل :- بما أن اكبر تكرار في التوزيع هو 23 عليه فأن فئة المنوال هي الفئة الخامسة (-50) وبذلك فأن :-

$F_5 = 23 \quad , \quad f_4 = 17 \quad , \quad f_6 = 14 \quad , \quad L_5 = 50 \quad , h_3 = 10$

$$mo = L_k + \frac{(f_K - f_{k-1})}{(f_k - f_{k-1}) + (f_k - f_{k+1})} \cdot h_k$$

$$mo = 50 + \frac{(23-17)}{(23-17)+(23-14)} .10$$

$$Mo = 500 + \frac{60}{15} = 54$$ سنه

العـــيــنـــات

العينات الاحتمالية : وهي العينات التي يمكن استخدامها في الطريقة الإحصائية لتمدنا بتقديرات دقيقة عن المجتمع موضوع الدراسة وهي تعطي نفس الفرصة للاختيار لكل مفردة , للوصول إلى تقديرات دقيقة عن المجتمع الأصلي , ومن هذه العينات :-

أ -العينة العشوائية البسيطة:

تختار العينة العشوائية البسيطة في حالة توفر شرطين أساسيين هما **الأول** : إن يكون جميع أفراد المجتمع الأصلي معروفين ، **والثاني** :أن يكون هناك تجانس بين هؤلاء الأفراد ، ففي مثل هذه الحالة يعمد الباحث إلى اختيار عينة عشوائية بسيطة وفق الأساليب التالية:

١ -**القرعة**: حيث يتم ترقيم أفراد المجتمع الأصلي ووضع الأرقام في صندوق خاص ويتم سحب الأرقام حتى نستكمل العدد المناسب للعينة.

٢-**جدول الأرقام العشوائية** : وهي عبارة عن جداول يوجد بها أرقام عشوائية كثيرة يختار الباحث منها سلسلة من الأرقام العمودية أو الأفقية ، ثم يختار من المجتمع الأصلي الأفراد الذين لهم الأرقام نفسها التي اخترناها من جدول الأرقام العشوائية ، ويكون هؤلاء الأفراد هم العينة المختارة. من الواضح إن اختيار هذه العينة العشوائية البسيطة يبدو سهلا ولكن ذلك يتطلب جهدا ووقتا طويلين، كما لا نضمن أن يكون هذه العينة ممثلة بدقة للمجتمع الأصلي.

ب -العينة الطبقية:

عرفنا ان العينة العشوائية تختار في حالة واحدة هي تجانس جميع أفراد المجتمع الأصلي وبذلك نضمن تمثيل هذه العينة لمجتمعها الأصلي ،ولكن هذا التجانس بين أفراد المجتمع الأصلي قد لا يكون دائمي ، وان أفراد المجتمع قد يكون متباينين ، فإذا كان باحث ما يريد ان يدرس اتجاهات الطلاب الملتحقين بالمهن التعليمية نحو دراستهم فان بامكانه إن يعتبر المجتمع الأصلي هنا – وهو الطلاب الملتحقين بالمهن التعليمية – هو مجتمع يضم أفرادا متجانسين ، لأن نظريتهم إلى دراستهم والمساقات التي يدرسونها أو يحتاجون إليها تكون متقاربة وبالتالي يمكن إن يختار الباحث عينة عشوائية بسيطة تمثلهم جميعا .أما إذا أراد هذا الباحث أن يدرس مشكلات الطلاب الملتحقين بالمهن التعليمية فانه هنا أمام مجتمع الطلاب الملتحقين بالمهن التعليمية وهو غير متجانس لان مشكلات الطلاب في هذه الحالة تتأثر بالجنس –ذكورا وإناثا وتتأثر بالعمر ، اقل من عشرين عاما وأكثر من عشرين عاما ، وتتأثر بالمستوى الاجتماعي للطلاب ،كما تتأثر بعوامل اجتماعية واقتصادية متعددة ، فالمجتمع في هذه الحالة لا يضم أفرادا متجانسين بل يضم طبقات أو فئات متعددة ومتباينة حيث يمكن أن نلاحظ الفئات التالية:

- طلاب السنة الأولى وطلاب السنة الثانية.

- الطلاب الذكور والطلاب الإناث.

- الطلاب المتفوقين وغير المتفوقين.

- الطلاب من مستويات اجتماعية مختلفة.

وفي مثل هذه الحالة لا بد ان تكون العينة ممثلة لجميع هذه الطبقات وبذلك نختار عينة عشوائية طبقية ، فكيف يتم الاختيار ؟ إن على الباحث إن يقوم بما يلي :

اولا - ان يحدد الفئات المختلفة في المجتمع الأصلي.

ثانيا - ان يحدد عدد الطلاب في كل فئة.

ثالثا - ان يختار من كل فئة عشوائية بسيطة تمثلها مراعيا في ذلك نسبة ثابتة من كل فئة بحيث تمثل كل فئة بعدد من الأفراد متناسبا مع حجم هذه الفئة.

ج -العينة المنتظمة Systematic

وهي شكل من أشكال العينة العشوائية يتم اختيارها في حالة تجانس المجتمع الأصلي , فإذا كان المجتمع الأصلي مكونا من ٢٠٠ طالب ونريد أن نختار عينة عشوائية منتظمة مكونة من ١٠, تكون المسافة بين الارقام = ٢٠ عشرين , نقسم ٢٠٠ العدد الكلي للبيانات على عدد العينة١٠ , ثم نختار الرقم الأول عشوائيا وليكن الرقم ٦ وبذلك تكون العينة مكونة من الطلاب الذين يحملون الأرقام التالية

... ،٣٦، ٢٦ ،٦،١٦، فهذه العينة تسمى منتظمة لأننا اخترنا مسافة ثابتة بين كل رقم والرقم الذي يليه ولكن يعاب على هذه العينة بأن تمثيلها ليس دقيقا خاصة إذا أجريت في مجال البحوث الاجتماعية ، فلو افترضنا أننا نجري دراسة على سكان المنازل المكونة من شقق لها أرقام خاصة ،فقد لا تحوي العينة أية أرقام للشقق الأرضية أو الشقق العليا ،وهذا ما يبعد هذه الفئة عن التمثيل الدقيق.

مثال: يراد اختيار عينة بحجم ١٥ وحدة من مجتمع متجانس .

ـ نحدد المسافة على أساس سحب الوحدات على بعد ١٠وحدة

ـ نختار رقم عشوائي من بين الأرقام الأولى لمفردات المجتمع مثلا الرقم ٧

ـعندئذ بقية الوحدات ستحمل الأرقام(١٧,٢٧,٣٧,٤٧,٠٠٠)

ء- العينة المتعددة المراحل

وتستخدم عندما يكون وحدات المجتمع متباعدة في منطقة جغرافية واسعة , فنقسم المجتمع الى وحدات أولية كل منها تقسم الى عدد من الوحدات الثانوية , ثم تقسم هذه الى وحدات اصغر وهكذا الى ان نصل الى الوحدات المطلوبة التي توصف بأنها متجانسة.

مثال : عند دراسة غلة الدونم الواحد لمحصول الحنطة في محافظة بابل تقسم المحافظة الى عدد من الاقضية ويتم اختيار عينة من هذ ه الاقضية وتقسم الاقضية المختارة الى نواحي نأخذ منها عينة عشوائية (الوحدات الثانوية) , وتقسم النواحي المختارة الى قرى ونختار منها العينة المطلوبة , ويتم إعداد قوائم بالمزارع الموجودة في القرى التي تم اختيارها لجمع المعلومات عنها , من مزايا هذه الطريقة انها تقلل التكاليف .

الفصل الرابع

الاحصاء الحيوي

الإحصاءات الحيوية

١. أهمية الإحصاءات الحيوية

تعتبر الإحصاءات الحيوية من الإحصاءات الضرورية نظرا لأنها تستخدم في العديد من الأغراض والتي منها على سبيل المثال لا الحصر التقديرات السكانية التي تستند أساسا إلى أعداد المواليد والوفيات إلى جانب حجم الهجرة. ولا يقتصر مفهوم الإحصاءات الحيوية على المواليد والوفيات فقط ، بل يشمل أيضا عقود الزواج وواقعات الطلاق التي تشكل حوادث لها أهميتها في التزايد السكاني . كما تعتبر هذه الإحصاءات جزءا مكملا للإحصاءات التي تنتجها الأجهزة الإحصائية والتي كثيرا ما يتم استخدامها لحساب العديد من المؤشرات الديموغرافية وغيرها من المؤشرات التي تعتبر ضمن مفهوم المؤشرات الاجتماعية.

وتخدم الإحصاءات الحيوية العديد من مستخدمي البيانات، حيث أنها تعتبر بيانات على قدر كبير من الأهمية لتبرير الاحتياجات التي تعمل الحكومات لتوفيرها. ويمثل وجود سجل مدني للحوادث الحيوية أهمية كبيرة ليس للحكومة فقط، وإنما للمواطنين أيضا، وذلك لأن وجود هذا السجل يزود ويتيح للمواطنين الحصول على الإثبات القانوني لهويتهم بالدرجة الأولى وكذلك بالمعلومات الدقيقة الخاصة بهم كالخصائص المختلفة.

يتم تعريف **<u>السجل المدني من قبل الأمم المتحدة</u>** على النحو التالي:

<u>" التسجيل المستمر، والدائم، والإلزامي، والعالمي لوقوع , وخصائص الأحداث الحيوية (المواليد الأحياء والوفيات، والوفيات الجنينية، والزواج، والطلاق) وأحداث الأحوال المدنية الأخرى المتعلقة بالسكان على النحو المنصوص عليه بموجب القانون أو التنظيم، وفقا للمتطلبات القانونية في كل بلد " .</u>

تشكل البيانات المتعلقة بالإحصاءات الحيوية التي يتم الحصول عليها من السجلات المدنية الأدوات الضرورية للتحليل الديموغرافي اللازم في عمليات التخطيط الاقتصادي والاجتماعي والتي تشتمل أيضا على الاتجاهات في النمو السكاني والتوزيع السكاني.

ويتوقف تحديث البيانات المتعلقة بالسكان خلال الفترات الفاصلة بين التعدادات على توافر الإحصاءات الحيوية. وبالإضافة لأهمية الإحصاءات الحيوية للأفراد وللأغراض الرسمية الأخرى هناك أهمية أخرى تتمثل في استخدامها في الأغراض العلمية.

وتشمل استخدامات الإحصاءات الحيوية على المجالات الديموغرافية والصحية والتي تعتمد على المعلومات التي يتم جمعها من خلال نظام السجل المدني لتطوير أنظمة بيانات موازية تحتوي على معلومات على قدر كبير من الأهمية حول مواضيع مختلفة كالخصوبة والوفيات ووفيات الرضع والأطفال وبناء جداول الحياة التي تعتمد

أساسا على معدلات الوفاة التفصيلية، وكذلك لدراسة تأثير الأمراض المزمنة أو المؤقتة. وتخدم البيانات المتعلقة بالحوادث الحيوية العديد الأغراض الإدارية والحكومية، حيث تستخدم السجلات الخاصة بالمواليد في الأنشطة الصحية العامة كبرامج الرعاية الصحية ما بعد الولادة للأمهات والرضع والتطعيم وبرامج التغذية.

كما تستخدم المعلومات المتعلقة بالوفيات لتحديد الأمراض المعدية التي تحتاج إلى متابعة من قبل السلطات المختصة بالصحة العامة، وكذلك لتنقيح جداول الناخبين وغيرها من الاستخدامات الإدارية. ولا شك أن الانعكاسات التي تنتج عن البحث العلمي تتعدى نطاق ذلك البحث، حيث يستند إليها في صياغة السياسات وإعداد الخطط والبرامج الموجهة للفئات السكانية المختلفة.

٢- الإحصاءات الحيوية واحتياجات مستخدمي البيانات

يشهد الوقت الحالي تزايدا في الطلب على البيانات الإحصائية بكافة أنواعها ومن ضمنها الإحصاءات الحيوية ومؤشراتها. وتضم قائمة البيانات التي تزايد الطلب عليها العديد من البيانات المتعلقة بالمواليد والوفيات والزواج والطلاق. وتعتبر البيانات المتوافرة حول هذه الجوانب مكتملة إلى حد بعيد لدى دائرة الإحصاءات العامة ودائرة الأحوال المدنية والجوازات.

ويعتبر مسح السكان والصحة الأسرية المصدر الرئيسي للبيانات الحيوية والتي تشتمل على بيانات المواليد والوفيات وما يتفرع عنها من بيانات ثانوية. التي يمكن استخدامها في حساب العديد من المؤشرات الحيوية المتعلقة بالخصوبة والوفاة، حيث يمكن حساب المؤشرات التالية:

CBR	Crude Birth Rate	معدل المواليد الخام
GFR	Glomerular Filtration Rate	معل الإنجاب العام
ASFR	Age-specific fertility Rates	معدلات الخصوبة التفصيلية
CMR	Child Mortality Rate	معدل وفيات الأطفال
NMR	Neonatal Mortality Rate	معدل وفيات حديثي الولادة
IMR	Infant Mortality Rate	معدل وفيات الأطفال الرضع
UFMR	Under Five Mortality Rate	معدل وفيات دون الخامسة من العمر

كما يتيح المسح حساب المؤشرات المبينة أعلاه على مستويات إدارية مختلفة بالإضافة إلى المستوى الوطني كالأقاليم وبعض المحافظات الكبيرة والمدن الكبرى والحضر والريف وكذلك حسب خصائص ومتغيرات رئيسية تتعلق بالأم . وتتيح البيانات المتوفرة لدى دائرة الأحوال المدنية المتعلقة بالمواليد والوفيات حساب المؤشرات المشار إليها أعلاه على مستويات إدارية مختلفة.

كما تتيح البيانات المتوفرة لدى دائرة الأحوال المدنية حول عقود الزواج والطلاق والتي يتم تقديمها إلى الدائرة

من قبل المحاكم المختصة خلال ثلاثين يوما من حدوثها، لمعرفة أعداد عقود الزواج وواقعات الطلاق المسجلة ، وخصائص المتزوجين والمطلقين التي تشتمل على عدد من المتغيرات كالعمر والجنس والمستوى التعليمي والحالة الزوجية السابقة ومهنة المتزوج والمتزوجة. وتستخدم هذه المعلومات في حساب معدلات الزواج المختلفة كمعدل الزواج العام ومعدل الطلاق العام .

وتوفر دائرة الإحصاءات العامة بيانات معلومات حول الحالة الزوجية للسكان الذين أعمارهم ١٥ سنة فأكثر من التعدادات السكانية التي تنفذها مرة كل عشر سنوات ، وكذلك من المسوح الديموغرافية والأسرية.

مما سبق فأن الإحصاءات الحيوية تشتمل على العديد من الإحصاءات , سيتم التركيز على البعض منها وهي :-

١ – إحصاءات المواليد .

٢ – إحصاءات الوفيات .

٣ – إحصاءات الزواج والطلاق (الخصوبة) .

٤ – إحصاءات الأمراض .

وفيما يلي تفصيل هذه الإحصاءات :-

أولا : إحصاءات المواليد

يتم الحصول على البيانات المتعلقة بالمواليد من المستشفيات وكذلك السجل المدني الذي يفرض على المواطنين التسجيل والتبليغ عن كل ولادة جديدة (خاصة إذا كانت الولادة خارج المستشفى).

و يجب التفريق بين المواليد الأحياء والمواليد الموتى .

تعريف **المولود الحي** :-"هو كل مولود تظهر عليه بعد ولادته أي علامة من علامات الحياة بعد انفصاله عن امه حتى ولو توفى بعد ذلك فورا " .

تعريف **المولود الميت** :- " من ولد ميتا بعد الشهر السادس من الحمل سواءً أحدثت الوفاة قبل الوضع أو أثناءه ولم يظهر على الجنين بعد الانفصال التام أي علامة من علامات الحية .

ومن أهم إحصاءات المواليد :-

$$\textbf{معدل المواليد الخام} = \frac{\text{عدد المواليد أحياء في السنة}}{\text{عدد السكان في منتصف السنة}} * 1000$$

وقد سمي هذا المعدل بالمعدل الخام أو العام لأنه لا يأخذ في الاعتبار اختلافات الترتيب السكاني بين المجتمعات .

مثال : كان عدد المواليد أحياء خلال السنة 3254 ولادة حسب سجلات شعبة الولادات والوفيات , بينما كان عدد السكان في منتصف العام 31587 نسمة , المطلوب إيجاد معدل المواليد الخام ؟

الحل :

$$\textbf{معدل المواليد الخام} = \frac{\text{عدد المواليد أحياء في السنة}}{\text{عدد السكان في منتصف السنة}} * 1000$$

معدل المواليد الخام $= \frac{3254}{31587} * 1000 = 103.01$ ولادة حية لكل 1000

ثانيا : إحصاءات الوفيات

معدل الوفيات

هو مقياس لعدد الوفيات (بشكل عام أو لسبب محدد) بالنسبة لتعداد السكان في السنة الواحدة. وعادة ما يتم التعبير عنه بصيغة (لكل 1000 شخص بالسنة) .معدل الوفيات أيضا يمكن تعريفه على انه يعبر عن عدد الوفيات خلال فترة معينة من الزمن بين نوع أو مجموعة معينة من الناس .

تشير بيانات الوفيات لأعداد الوفيات في مكان وزمان محددين مع معرفة السبب. السبب الكامن وراء الوفاة يعرف على انه "المرض أو الإصابة التي شرعت في سلسلة من الأحداث المرضية المؤدية مباشرة إلى الموت ، أو ملابسات الحادث أو العنف الذي أسفر عن إصابة قاتلة" ، وفقا لقواعد التصنيف الدولي للأمراض .

و هناك عدة مصطلحات تندرج تحت معدل الوفيات :ـ

معدل الوفيات الخام

هو عدد الوفيات بين سكان منطقة جغرافية معينة خلال سنة معينة , ويمثل "عدد المتوفين في كل 1000 شخص في منتصف العام لسكان في منطقة جغرافية معينة خلال العام نفسه .

$$\text{معدل الوفيات الخام} = \frac{\text{عدد الوفيات/(عدا المواليد موتى)}}{\text{عدد السكان في منتصف السنة}} * 1000$$

وبشكل عام يمكن تلخيص العدد الكبير من المعدلات التي يمكن استخراجها وحسب الحاجة اليها بالمعادلة التالية :ـ

$$\text{معدل الوفيات الخاص} = \frac{\text{عدد الوفيات الذين لهم صفة خاصة}}{\text{عدد السكان في منتصف السنة}} * 1000$$

ويقصد بالصفة الخاصة الجنس أو الجنسية أو اللون أو فئة السن أو الحالة الاجتماعية .... الخ

مثال :ـ الجدول التالي يبين فئات السكان وفئات الوفيات في بلد ما .

فئة السن	الحالة الاجتماعية **للسكان**					
	لم يتزوج مطلقا		متزوج		مطلق	
	ذكر	أنثى	ذكر	أنثى	ذكر	أنثى
25-18	781.700	222.100	987.500	1,812,800	8.000	120.000
35-25	450.000	150.000	1.800.000	2.700.000	18.000	22.000
35- فأكثر	300.000	2.000	1.500.000	3.000.000	5.000	2.000

فئة السن	الحالة الاجتماعية **للمتوفين**					
	لم يتزوج مطلقا		متزوج		مطلق	
	ذكر	أنثى	ذكر	أنثى	ذكر	أنثى
25-18	8.200	13.100	27.000	20.000	2.000	1.500
35-25	22.000	18.000	50.000	30.000	9.000	5.000
35- فأكثر	50.000	10.000	100.000	50.000	1.000	500

المطلوب : ١ – حساب معدل الوفاة الخاص بفئة السن (18 – 25) .

٢ – حساب معدل الوفاة الخاص بالمتزوجين .

الحل ١ / معدل الوفاة الخاص بفئة السن (18 – 25)

معدل الوفيات الخاص بفئة السن (18 ـ 25) = $\dfrac{\text{عدد وفيات الذين هم بفئة السن (18 – 25)}}{\text{عدد السكان (18-25) في منتصف السنة}} * 1000$

$$= \frac{1.500+2.000+20.000+27.000+13.100+8.200}{120.000+8.000+1.812.800+987.500+222.100+781.700} * 1000$$

معدل الوفيات الخاص بفئة السن (18 ـ 25) = $\dfrac{71.800}{3.932.100}$ = 18.26 وفاة لكل 1000

الحل ٢ / معدل الوفاة الخاص بالمتزوجين

$$\text{معدل الوفيات المتزوجين} = \frac{\text{عــدد وفيات المتزوجين}}{\text{عــدد السكان(المتزوجين) في منتصف السنة}} * 1000$$

$$= \frac{50.000+100.000+30.000+50.000+20.000+27.000}{3.000.000+1.500.000+1.800.000+2.700.000+1.812.800+987.500} * 1000$$

$$\text{معدل الوفيات المتزوجين} = \frac{277.000}{11.800.300} * 1000 = 23.47$$ وفاة لكل 1000

__تمرين__ : - احسب المعدلات التالية بالاعتماد على البيانات من الجدول أعلاه

١ – معدل الوفاة للذين لم يتزوجوا وهم من فئة السن (25 – 35) .

٢ – معدل الوفاة للمطلقين الذكور .

٣ – معدل الوفاة للإناث الذين لم يتزوجوا وهم من فئة السن (18 – 25) .

معدل وفيات الأمومة

(معدل وفيات النساء بسبب الحمل إلى عدد الولادة الحية)

هو مقياس يعبر عن وفاة امرأة أثناء الحمل أو خلال 42 يوما من انتهاء الحمل، بصرف النظر عن مدة وموقع الحمل، او أي قضية متعلقة أو متفاقمة من جراء الحمل أو إدارتها ولكن ليس من أسباب عارضة أو طارئة. لتسهيل التعرف على وفيات الأمهات في ظروف سبب إسناد الموت غير كاف .

$$\text{معدل وفيات الأمومة} = \frac{\text{عــدد وفيات النساء بسبب الحمل والولادة}}{\text{عــدد المواليد أحياء}} * 1000$$

وفيات الفترة المحيطة بالولادة والطفولة

الوفيات في فترة ما حول الولادة تشير إلى مزيج من وفيات الأجنة والمواليد الأحياء فقط مع بقاء فترة قصيرة (أيام أو أسابيع)، ويتم تجميعها على افتراض أن هناك عوامل مماثلة تربط هذه الخسائر.عدد وفيات حديثي الولادة (أقل من ٢٨ يوم) و الإملاص (ولادة وليد ميت) بالنسبة لجميع الولادات على الرغم من كون الأطفال حديثي الولادة ليسوا مرضى، فإن أعدادا كبيرة من الأطفال يموتون بعد الولادة مباشرة وكثير منهم في الأسابيع الأربعة الأولى من الحياة (وفيات الأطفال حديثي الولادة)، ومعظم هؤلاء خلال الأسبوع الأول (في وقت مبكر وفيات المواليد الجدد). لكل طفل يموت في الأسبوع الأول بعد الولادة، وآخر يولد ميتا (موت الجنين أو موت الجنين داخل الرحم).

من أهم معدلات الوفيات الفترة المحيطة بالولادة والطفولة :-

- معدل وفيات الأطفال حديثي الولادة

الطفل حديث الولادة أو الوليد هو طفل تحت 28 يوما من العمر. خلال هذه الـ 28 يوما الأولى من عمر الطفل يكون أكثر عرضة لخطر الموت, بالتالي فمن الأهمية أن يتم توفير التغذية والرعاية المناسبة خلال هذه الفترة، وذلك لتحسين فرص الطفل في البقاء على قيد الحياة ، ووضع أسس لحياة صحية .

معدل وفيات الأطفال حديثي الولادة(اقل من 28 يوم) = (**عــدد وفيات الأطفال اقل من 28 يوم** / **عــدد المواليد أحياء**) * 1000

- معدل وفيات الأطفال الرضع

وهو نسبة وفيات الأطفال الرضع (عدا المواليد موتى) إلى عدد الولادات الحية , يعبر عنه بعدد الوفيات في ١٠٠٠ ولادة حية .

معدل وفيات الأطفال الرضع = (**عــدد وفيات الأطفال الرضع عدا (المواليد موتى)** / **عــدد المواليد أحياء**) * 1000

- معدل وفيات الطفولة المبكرة

وهو نسبة وفيات الأطفال (أعمارهم أقل من سنة واحدة وأكثر من 28 يوم) إلى عدد الولادات الحية , احتمال الوفاة بين 28 يوم وسنة واحدة من العمر، على مدى فترة محددة، يعبر عنه بعدد الوفيات في 1000 ولادة حية .

عـــدد الوفيات من (28 يوم إلى 12 شهر)

مـعـدل وفــيــات = ———————————————— * 1000

الطفولة المبكـرة عـــدد المواليد أحياء – عدد الوفيات اقل من 28 يوم

ـ معدل وفيات الأطفال دون سن الخامسة

يعــبر عــن احتمــال الوفــاة بــين ســن الــولادة وبالضــبط ســن الخامســة لكــل 1000 ولادة حيــة (نســبة وفيات الأطفال أعمارهم أقل من 5 سنوات) إلى عدد الولادات الحية

عـــدد الوفيات الأطفال (دون سن الخامسة)

معدل الوفيات الأطفال = ———————————————— * 1000

دون سن الخامسة عـــدد المواليد أحياء

مثال :ـ إذا كان عدد الوفيات النساء أثناء الحمل والولادة (42.600) , وعدد المواليد أحياء **مليون** طفل , وعدد المواليد موتى (3.000) , وعدد وفيات الأطفال الرضع الأقل من سنة يساوي (15.000) منهم (1.000) طفل حديث الولادة , وعدد وفيات الأطفال دون سن الخامسة (24.000) , المطلوب حساب معدلات الوفاة التالية :

١ – معدل وفيات الأمومة .

٢ – معدل وفيات الأطفال حديثي الولادة .

٣ – معدل وفيات الأطفال الرضع .

٤ – معدل وفيات الطفولة المبكرة .

٥ – معدل وفيات الأطفال دون سن الخامسة .

الحل ١ :ـ

عـــدد وفيات النساء بسبب الحمل والولادة

معدل وفيات الأمومة = ———————————————— * 1000

عـــدد المواليد أحياء

$$\text{معدل وفيات الأمومة} = \frac{42.600}{1.000.000} * 1000 = 42.6$$

وفاة لكل 1000 ولادة

الحل ٢ :-

معدل وفيات الأطفال حديثي الولادة (اقل من 28 يوم) = (عــدد وفيات الأطفال اقل من 28 يوم ÷ عــدد المواليد أحياء) * 1000

معدل وفيات الأطفال حديثي الولادة (اقل من 28 يوم) = $\frac{1.000}{1.000.000}$ * 1000 = 1 وفاة لكل 1000 ولادة

الحل ٣ :-

مـعدل وفــيــات الأطفال الرضع = (عــدد وفيات الأطفال الرضع عدا (المواليد موتى) ÷ عــدد المواليد أحياء) * 1000

مـعدل وفــيــات الأطفال الرضع = $\frac{3.000 - 15.000}{1.000.000}$ * 1000 = 12 وفاة لكل 1000 ولادة

الحل ٤ :-

مـعدل وفــيــات الطفولة المبكرة = (عــدد الوفيات من (28 يوم إلى 12 شهر) ÷ عــدد المواليد أحياء – عدد الوفيات اقل من 28 يوم) * 1000

مـعدل وفــيــات الطفولة المبكرة = $\frac{1.000 – 15.000}{1.000 - 1.000.000}$ * 1000 = 14.01 وفاة لكل 1000 ولادة

الحل ٥ :-

معدل الوفيات الأطفال دون سن الخامسة = (عــدد الوفيات الأطفال (دون سن الخامسة) ÷ عــدد المواليد أحياء) * 1000

$$\text{معدل الوفيات الأطفال دون سن الخامسة} = \frac{24.000}{1.000.000} * 1000 = 24 \text{ وفاة لكل 1000 ولادة}$$

تمرين : من بيانات المثال أعلاه احسب معدل الوفاة للأطفال دون سن الخامسة وفوق عمر الطفولة المبكرة .

ثالثا : إحصاءات الخصوبة

معدل الخصوبة العام:

هو النسبة بين عدد المواليد الأحياء في السنة إلي إجمالي عدد الإناث في سن الحمل (من 15 إلى 49 عام) وهو يستبعد كل الذكور وكل الإناث اللواتي خارج فترة الحمل.

$$\text{معدل الخصوبة العام} = \frac{\text{عدد المواليد الأحياء في سنة ما}}{\text{عدد الإناث في سن الحمل في منتصف السنة}} *1000$$

معدل الخصوبة العمري النوعي الخاص (حسب فئة السن) :

هو النسبة بين جملة عدد المواليد لأمهات في أعمار معينة إلي عدد الإناث في كل فئة عمرية، ويعتبر هذا المعدل أدق من المعدل السابق وذلك لأن عدد المواليد يختلف باختلاف أعمار الأمهات بدرجة كبيرة.

$$\text{معدل الخصوبة لفئة عمرية معينة} = \frac{\text{عدد المواليد الأحياء في فئة عمرية معينة}}{\text{عدد الأمهات في نفس الفئة العمرية في منتصف السنة}} *1000$$

مثال : الجدول التالي يبين توزيع الإناث (في سن الحمل) حسب فئات السن وعدد المواليد أحياء المقابلة لفئات سن الأمهات في سجلات وحدة الرعاية الأسرية في مجموعة مراكز صحية في مدينة معينة .

فئات السن	عدد الإناث في منتصف السنة	عدد المواليد أحياء
20 - 15	**38.615.780**	85.610
25 - 20	**17.865.162**	279.685
30 - 25	**25.686.921**	781.765
35 - 30	**29.686.198**	5.380.076
40 - 35	**21.485.000**	426.970
45 - 40	**16.692.400**	166.350
45 فأكثر	**21.820.000**	49.270
المجموع	**171.788.000**	2.327.726

احسب ما يلي :

١ – معدل الخصوبة للفئة العمرية (20 – 25) .

٢ – معدل الخصوبة للفئة العمرية (45 فأكثر) .

٣ – معدل الخصوبة العام .

الحل ١ /

$$\text{معدل الخصوبة للفئة العمرية} (20-25) = \frac{\text{عدد المواليد أحياء لإمهات في الفئة العمرية} (25\text{ـ}20)}{\text{عدد الأمهات في الفئة العمرية} (20-25)} * 1000$$

$$\text{معدل الخصوبة للفئة العمرية} (20-25) = \frac{279.685}{17.865.162} * 1000 = 15.55 \text{ ولادة لكل } 1000$$

الحل ٢ /

$$\text{معدل الخصوبة للفئة العمرية} (45\text{فأكثر}) = \frac{\text{عدد المواليد أحياء لإمهات في الفئة العمرية} (45\text{فأكثر})}{\text{عدد الأمهات في الفئة العمرية} (45\text{ فأكثر})} * 1000$$

$$\text{معدل الخصوبة للفئة العمرية} (٤٥\text{فأكثر}) = \frac{49.270}{21.820.000} * 1000 = 2.26 \text{ ولادة لكل } 1000$$

الحل ٣ /

$$\text{معدل الخصوبة العام} = \frac{\text{عدد المواليد الأحياء في سنة ما}}{\text{عدد الإناث في سن الحمل في منتصف السنة}} * 1000$$

$$\text{معدل الخصوبة العام} = \frac{2.327.726}{171.788.000} * 1000 = 13.55 \text{ ولادة لكل } 1000$$

تمرين :

١ـ احسب معدل الخصوبة للنساء اللاتي أعمارهن بين (20 – 30) .

٢ـ احسب معدل الخصوبة للنساء اللاتي أعمارهن فوق (40 سنة) .

رابعا : إحصاءات الأمراض

من الإحصاءات التي تهم العاملين في المجال الصحي وتحليل الوضع الصحي في المجتمع هو موضوع إحصاءات المرض , وفيما يلي بعض المعدلات الخاصة بالإحصاءات المرضية .

١ـ معدل الإصابة.

٢ـ معدل الانتشار.

٣ـ نسبة حالات الهلاك.

معدل الإصابة: ويمثل نسبة عدد المصابين الجدد بمرض معين الى عدد السكان في منتصف السنة .

$$\text{معدل الإصابة} = \frac{\text{عدد الإصابات الجديدة في مرض معين خلال السنة}}{\text{عدد السكان في منتصف السنة}} * 1000$$

معدل الانتشار: يمثل نسبة عدد الإصابات الموجودة في لحظة معين إلى عدد السكان في منتصف السنة

$$\text{معدل الانتشار} = \frac{\text{عدد الإصابات الموجودة في لحظة معينة}}{\text{عدد السكان في منتصف السنة}} * 1000$$

نسبة حالات الهلاك : يمثل نسبة عدد حالات الوفاة بسبب مرض ما الى عدد المصابين بهذا المرض .

$$\text{نسبة حالات الهلاك} = \frac{\text{عدد حالات الوفاة بسبب مرض معين}}{\text{عدد حالات الإصابة بهذا المرض}} * 1000$$

مثال : في مجتمع مكون من 10.000 شخص وجد أن 60 حالة إصابة بمرض الحمى النزفية توفي منهم 10 أشخاص والباقون تعافوا من المرض , وان 120 حالة إصابة بمرض الأنفلونزا خلال الأسبوع الأول من شهر كانون الثاني .

المطلوب : ١ – اوجد معدل الإصابة بمرض الحمى النزفية .

٢ – احسب معدل انتشار مرض الأنفلونزا .

٣ – ما هي نسبة حالات الهلاك لمرض الحمى النزفية .

الحل ١ /

$$\text{معدل الإصابة بمرض الحمى النزفية} = \frac{\text{عدد الإصابات بمرض الحمى النزفية خلال السنة}}{\text{عدد السكان في منتصف السنة}} * 1000$$

$$\text{معدل الإصابة بمرض الحمى النزفية} = \frac{60}{10.000} * 1000 = 6 \text{ إصابة لكل 1000 شخص}$$

الحل ٢ /

$$\text{معدل الانتشار لمرض الأنفلونزا} = \frac{\text{عدد الإصابات الموجودة في الأسبوع الأول لشهر ك ٢}}{\text{عدد السكان في منتصف السنة}} * 1000$$

$$\text{معدل الانتشار لمرض الأنفلونزا} = \frac{120}{10.000} * 1000 = 12 \text{ إصابة لكل 1000 شخص}$$

الحل ٣ /

$$\text{نسبة حالات الهلاك بسبب مرض الحمى النزفية} = \frac{\text{عدد حالات الوفاة بسبب مرض الحمى النزفية}}{\text{عدد حالات الإصابة بمرض الحمى النزفية}} * 1000$$

نسبة حالات الهلاك بسبب مرض الحمى النزفية= $\frac{10}{60}$ *1000 = 166.66 وفاة لكل 1000 مصاب

<u>تمرين</u>: هل يمكن حساب معدل المصابين بمرض الحمى النزفية والذين تعافوا من المرض ؟