المحاضرة 01: مدخل إلى المقياس

إن طرح الأسئلة التي ترمي إلى الحصول على المعلومات قد يتخذ في كثير من الأحيان طابعا ((إمبريقيا))، بهدف الكشف عن الطريقة التي تحدث بها التغيرات والظواهر على أرض الواقع، وهذا ما جعل العلماء والباحثون على اختلاف تخصصاتهم ومنهم علماء الاجتماع ينشغلون بالطرق المنهجية المناسبة لتحليل ومعالجة المعطيات والمعلومات كمجال جدير بالدراسة والبحث، بأساليب ومنهجيات ميدانية وتجريبية مختلفة لهذا سنحاول في هذه المحاضرة التعرف على معنى تعريف تحليل البيانات أو المعطيات ومن ثم تعريف البيانات وخصائصها في علم الاجتماع وأهم المصادر وطرق جمعها، و كذلك البحث عن العلاقة بين هذه المعطيات ونوع البحث، و وجبات الباحث اتجاه المعطيات الاجتماعية.

1- تعريف تحليل البيانات أو المعطيات (Data analysis) : تحليل البيانات هو عملية الفحص والتدقيق للبيانات، وتمشيطها لتكون أكثر دقة، وإعادة تشكيلها، وتخزينها أيضا لنحصل ونستنبط في النهاية علي معلومات يمكن علي أساسها اتخاذ وتحديد القرارات. وتعرف أيضا بأنها عبارة عن مجموعة من التقنيات الإحصائية تهدف إلى تقليص واختزال جدول بأبعاد (متغيرات) متعددة إلى جدول ذو هيكل بسيط. هذه التقنيات تهدف إلى وصف، تقليص، تصنيف البيانات .فمن خلال هذين التعريف يتبين لنا أن تحليل البيانات أو المعطيات الاجتماعية هو عبارة عن استخدام للوسائل الحسابية والرياضية في تجميع البيانات والمعلومات المختلفة ومن ثم تنظيمها وتبويبها بغرض وصفها وتفسيرها وفهم العلاقات المختلفة فيما بينها بشكل يساعد في الوصول إلى عدد من الاستنتاجات التي توصل إلى حقيقة الأهداف المنشودة. ولتحليل البيانات طرق عديدة تختلف باختلاف المجال المستخدمة فيه. غير أننا سنركز على الطريقة التي تستخدم في العلوم الإنسانية والعلوم الاجتماعية فقط.

2 – تعريف البيانات: Data : هي المادة الخام المسجلة كرموز أو أرقام أو جمل أو عبارات يمكن للإنسان تفسيرها أو تعليلها، فالبيانات هي عناصر من الحقيقة، مازالت تفتقر إلى المعنى العام، لذا يتوجب تحويلها إلى معلومات. مثال على ذلك: بيانات الطلبة (الأسماء – الأرقام المستوى – الشعبة – التخصص) بدون ترتيب، وينتج عن هذه البيانات بعد المعالجة ما يطلق عليه مصطلح معلومات. فالبيانات هي مجموعة من المشاهدات، أو الملاحظات التي تؤخذ أثناء دراسة معينة. وتجمع لغرض اختبار الفروض، وتقديم دليل تجريبي للتفسيرات والتنبؤات.

3- أنواع البيانات: تتباين طبيعة المعطيات الضرورية للبحث بين حقل معرفي وحقل معرفي آخر، فمعطيات الرياضيات معادلات وعلاقات رقمية ، ومعطيات الفيزياء ، ربما تكون رياضية لكنها في الآن نفسه تجريبية إلى حد ما، أما معطيات حقلي بحثنا العلوم الإنسانية والاجتماعية ، فهي متجانسة ومتباينة في آن واحد هي عامة معطيات وصفية ، لكنها تتباين بين ما هو كمي مثل ما هو الحال في علوم الجغرافيا الديموغرافيا وبعض فروع علم الاجتماع وعلم النفس (بيانات ، نسب وإحصائيات) ، وبين ما هو نوعي كحال الفلسفة والبحوث الأدبية والجمالية والأخلاقية (وقائع، أدلة ، بنى ، عناصر ، مفردات ، حالات محددة وسواها)،حيث تكثيف المادة المجمعة ، وتنويعها ومقاطعة مصادرها ، والنقد الصارم لها ،أي التحقق المتعدد الشكل منها ، والتحليل المنطقي ، ثم الاستنتاج ، وإعادة امتحان النتائج سلبا ، قبل الإيجاب كفيل بتوفير ضمانات موضوعية موثوقة علميا .بشكل عام تصنف البيانات إلى عدة تصنيفات حسب الاستعمال و المعنى وكذلك حسب مجال الدراسة ويمكن تصنيف البيانات إلى نوعين رئيسيين هما:

أ- **البيانات النوعية Qualitative Data**: وتنقسم إلى نوعين هما البيانات الاسمية Nominal Data و البيانات الترتيبية Ordinal Data.، أما البيانات الاسمية تكون في صورة غير عددية أي لا يمكن قياسها وتتكون من فئات لا يتم التفاضل بينهما مثل الجنس يتكون من طبقتين ذكور ونرمز لهما الرقم (1) والإناث ونرمز لهما بالرقم (2) أو سؤال تكون إجابته "نعم " ونرمز له بالرقم (1) و" لا" ونرمز له بالرقم (0) والخطأ الذي يقع فيه الباحث هو إجراء عمليات حسابية على البيانات الاسمية.

أما البيانات الترتيبية هي أيضا تكون في صورة غير عددية ولا يمكن إجراء العمليات الحسابية عليها والفرق بينها وبين البيانات الاسمية هي عملية المفاضلة والترتيب بين طبقات المتغير مثل المستوى التعليمي (ابتدائي (1)، إعدادي(2)، ثانوي (3)، جامعي فأكثر (4)) في هذه الحالة المتوسط الحسابي أو أي عملية حسابية على تلك البيانات ليس لها معنى.

2-البيانات الرقمية أو الكمية Quantitative Data :تنقسم أيضا إلى نوعين هما البيانات الفترية Interval Data والبيانات النسبية Rational Data، أما البيانات الفترية تكون في صورة عددية ويمكن إجراء العمليات الحسابية عليها مثل المتوسط الحسابي والانحراف المعياري وغيرها ويمتاز هذا المقياس بتساوي المسافات بين الرتب حيث أنه يسمي أحيانا " بمقياس المسافة " ، ويستخدم هذا المقياس كثيراً في العلوم التربوية والنفسية والاجتماعية مثل قياس الذكاء وغيرها ، والجدير بالذكر أن هذا المقياس لا يعني الصفر فيه عدم وجود الخاصية فدرجة طالب تساوي صفر مثلا لا يعني أنه لا يعرف شيء في المقرر. وأخيرا البيانات النسبية تكون أعلى مستوي من أنواع البيانات السابقة حيث يمتاز بالتصنيف (بيانات إسمية) و الترتيب (بيانات ترتيبية) والمسافات المتساوية (بيانات فترية) وخاصية النسبية والتي يعني إن للصفر خاصية العدم أي خاصية انعدام الظاهرة مثل سرعة سيارة تساوي صفر تعني أن السيارة متوقفة ، أو أن وزن شخص يساوي 60 كيلو جرام هو ضعف وزن شخص وزنه 30 كيلو جرام..

4- طرق جمعها البيانات Data : ومن أهم طرق جمع البيانات نذكر ما يلي:

أ . طريقة المسح الشامل: والتي تجمع فيها المعطيات من جميع أفراد المجتمع الإحصائي، والمجتمع الإحصائي (Statistical population) هو عبارة عن جميع عناصر موضوع الدراسة والتي نرغب في معرفة حقائق عنها سواء كانت على شكل إنسان أو حيوان أو جماد أو درجات امتحان أو منازل أو مزارع أو سفن ... الخ. وقد يتكون المجتمع من عدد محدود من العناصر مثل عدد أفراد مدينة ما أو عدد المنازل بهذه المدينة ... الخ، أو يتكون المجتمع من عدد غير محدود مثل الأسماك في الخليج العربي أو عدد النجوم ... الخ ، وتعتبر هذه الطريقة من أفضل طرق جمع البيانات لأنها تعطي بيانات كاملة حول موضوع الدراسة، أين تمتاز بالدقة العالية والوضوح والتفصيل والمصداقية، و يعاب عليها ارتفاع التكاليف ماديا و بشريا و زمنيا في حالة اتساع مجتمع الدراسة. إلا أنها مكلفة ومجهدة وتحتاج عددا كبيرا من الباحثين.

ب . طريقة العينة الإحصائية (Statistical sample): تستخدم هذه الطريقة في حال تعذر إجراء الدراسة بأسلوب المسح الشامل، وفي هذه الطريقة يمكن الاكتفاء ببيانات عن جزء من المجتمع تسمى العينة، والعينة هي جزء من عناصر المجتمع الكلي يتم اختيارها بحيث تكون تمثل هذا المجتمع الكلي، والتي يراعى فيها ضرورة التمثيل الصادق للمجتمع بعد تعيينه وتحديد هدف الدراسة، ويتم اللجوء إلى أسلوب أخذ العينات عند إجراء الدراسات والبحوث الإحصائية لأن تكاليفه أقل، وبواسطته يمكن الحصول على نتائج سريعة، مقارنة بأسلوب الحصر الشامل الذي يتم فيه جمع البيانات من كل مفردات المجتمع.

5- مصادر جمع البيانات: تحدد مصادر جمع المعلومات عادة من مصدرين:

أ . مصدر مباشر: ويحصل من خلال النزول للميدان وجمع المعلومات مباشرة.

ب- مصدر غير مباشر: ويندرج تحته:

. السجلات والوثائق التاريخية

. الاستبيان

. المقابلات الشخصية

. الاختبارات الخاصة مثل اختبارات الذكاء.

6- خصائص المعطيات الاجتماعية وأهميتها: للمعطيات الاجتماعية مجموعة من الخصائص تجعلها مهمة في البحوث الاجتماعية نلخصها فيما يلي:

- المعطيات هي حلقة وصل بين النظرية والمفاهيم في مستوى التصور، وبين المنهجية المستخدمة في البحث من حيث مناهجها وأدواتها المستخدمة.

- المعطيات الرقمية في علم الاجتماع تعكس التصورات النظرية المطروحة في البحث، أو تتجه نحو تأسيس نظرية جديدة.

- الأرقام المتحصل عليها في أي بحث ميداني، هي معطيات تعبر عن الواقع الذي تكون عليه الظاهرة في حدودها الزمنية (تطور البحث)، الجغرافية (ريف ، مدينة ...)، وفي حجم انتشارها (ضيقة ، واسعة)، وفي استمراريتها(متكررة أو ثابتة)، بالإضافة إلى أن هذه الأرقام قد تتجه نحو دحض النظرية (قوانين ونتائج) ، أو إثباتها و تطويرها في حال صدق النتائج.

– البيانات تفسح المجال للتحليل من خلال تطوير المداخل النظرية.

– المعطيات تدعم البحث الاجتماعي نحو الاستقلالية عن العلوم الطبيعية والدقيقة أو التلاقح معها، خاصة في مجال القوانين الإحصائية والرياضية، والتفاعل مع الحواسيب والوسائل التكنولوجية.

– المعطيات تبحث في مجال المقارنة والموازنة بين الدراسات السابق بعضها بعضا.

– للمعطيات الاجتماعية أثرها في توجيه الرأي العام وأصحاب القرار نحو سن القوانين واتخاذ القرارات التي تعكسها البحوث المبنية عليها.

7– علاقة المعطيات الاجتماعية بنوع البحث: لنوع البحث علاقة مهمة بالمعطيات التي يسعى الباحث نحو جمعها وتحليلها. خاصة في البحوث الكمية و النوعية.

. البحث الكمي: ترتبط فيه المعطيات الكمية بإجراءات رياضية، فتكون المعلومات المطلوب الحصول عليها قابلة للقياس، لأن الظواهر هنا قابلة للتقييم.

. البحث الكيفي أو النوعي: يصعب فيه تقييم الظواهر كميا. فيكون جمع المعطيات هنا لا لقياسها، وإنما لتأويلها وفق تصورات الباحث. فالبحث الكيفي لا يسمح بنفس الدقة التي يسمح بها البحث الكمي.

ملاحظة: يراعى أثناء جمع المعطيات بالضرورة حجم البحث الذي تتواجد فيه. فحينما يكون بحثا بسيطا أو تقريرا لغرض التدريب، يختلف الأمر عنه في البحوث الشاملة والمتخصصة. حيث لا بد من تلاؤم وسائلها وحجمها للأغراض التي تجمع لأجلها.

8– واجبات الباحث اتجاه المعطيات الاجتماعية: يتطلب البحث العلمي في العلوم الاجتماعية مجموعة من القواعد التي تعكس اقترابه من التفسير السليم لمسببات الظواهر الاجتماعية ونتائجها. وللمعطيات في خضم ذلك أكبر الأثر نحو تشخيص الواقع بما يحمله من تفاعلات في مختلف الميادين الاجتماعية والاقتصادية والسياسية وغيرها. و في كثير من الأحيان تكون النتائج المتحصل عليها من معطيات البحوث موجها لعدد من المسارات البحثية من جهة وللتفسيرات والتأويلات المكونة لثقافة المجتمع من جهة ثانية. ولذلك فالباحث مطالب بمجموعة من الواجبات اتجاه المعطيات التي يجمعها في بحثه، نذكر منها:

أ– الدقة: فكلما كانت المعطيات دقيقة كانت الدراسة البحثية أكثر قربا من تفسير الواقع، وأقوى حجة للباحث في الوصول إلى النتائج الصحيحة. كما قد تساهم في تطوير البحث العلمي من خلال التواصل في الدراسات الاجتماعية.

ب– الأمانة: فالمعطيات قد ينتج عنها قرارات مهمة لدى الساسة وأصحاب القرار، كما أنها تساهم في تكوين الوعي الثقافي وتطويره نحو الظواهر الحاصلة في المجتمع. إضافة إلى كونها تراث علمي تتبني عليه مواقف و واتجاهات لدى الباحثين في مجالات مختلفة.

ج– الموضوعية: فيجب أن تتلاءم المعطيات والأهداف التي وُجد من أجلها البحث. وعليها أن تتجه نحو التشخيص المبني على الأسس العلمية التي تجعلها متوافقة مع الحاجات المجتمعية من جهة، والرغبة في إشباع الفضول العلمي لدى الباحثين.

د– الواقعية: فيجب أن تعكس الواقع الذي تحصل فيه الظاهرة دون توجيه من الباحث نحو أغراض غير نزيهة، وعلى كل فالشروط الموضوعية تتعلق بالباحث ذاته، وما تتوفر فيه من مميزات، تجعله أكثر قربا من تحقيق ما سبق من خصائص.

المحاضرة 02: تبويب وتفريغ المعطيات

إن وجود البيانات بصورتها الأولية لا يتيح للباحث أن يستنتج شيئا منها، لاسيما عندما تكون بكميات كبيرة. ولهذا فان البيانات الأولية بصورتها غير المبوبة تمر بمراحل، قصد تلخيصها وتوضيحها للتعرف على ما تحويه من أغراض. وعملية تبويب البيانات تعد من العمليات المهمة في البحث العلمي، وذلك لتوفيرها الجهد والزمن عبر مراجعة البيانات ومعرفة مدى فائدتها للباحث، وتوافقها مع أهداف البحث. وهكذا تأتي عملية فرز البيانات الصحيحة عن البيانات الخاطئة، ثم تتم منها عملية ترقيمها و ترميزها بغية تفريغها إما بالطريقة اليدوية أو بالطريقة الآلية حيث يستخدم برنامج الحزمة الإحصائية للعلوم الاجتماعية لتفريغ وتحليل البيانات، ومن ثم يسهل على الباحث جدولتها . أي وضعها بأصغر حيّز ممكن، ولا أفضل من الجداول التي تساعد على التحليل استخلاص النتائج.

أولا– تبويب البيانات وتصنيفها: يمثل تبويب البيانات، الخطوة الأولى لبناء الأدوات البحثية. فيعمد الباحث إلى تصنيف المعطيات إلى فئات، معتمدا على نوعها وخصائصها. وموظفا مختلف الشروط الموضوعية والمنهجية التي تربط تساؤلات الدراسة وفروضها بالأهداف التي وضعت

من أجلها البيانات. بمعنى أن التبويب يتجه نحو تصنيف المعطيات، وفق ما هو شائع في قسمين من البيانات نوعية (كيفية)، وأخرى كمية(عددية).

ومن أهم المراحل التي تمر بها البيانات الأولية قصد تلخيصها وتبويبها، الآتي :

1- مرحلة مراجعة البيانات : ما إن تنتهي عملية جمع البيانات وفق الوسيلة المناسبة لذلك البحث، حتى تأتي مرحلة المراجعة، والتي تفحص فيها البيانات وتدقق لغرض التأكد من مطابقتها وتكاملها مع متطلبات الدراسة، فعلى سبيل المثال لو كانت الوسيلة المستخدمة في جمع البيانات هي الاستمارة عندئذ يتوجب عليا مراجعة وتدقيق كل الاستمارات التي تم جمعها، من أجل أن تستبقى الاستمارات ذات الإجابة الصحيحة الكاملة، وتستبعد الاستمارات الناقصة غير الكاملة، أو الاستمارات ذات الإجابات غير الصحيحة أو الواضحة، وكذلك عزل الاستمارات التي يعتقد الباحث أنها غير مطابقة لما هو مطلوب (لا تخدم البحث).

2- مرحلة تصنيف البيانات : تعد عملية تصنيف البيانات من الخطوات المهمة في عملية التبويب فبعد الانتهاء من عملية جمع الاستمارات المعنية بالبيانات المطلوبة ومراجعتها تأتي عملية فرز البيانات إلى مجاميع وأصناف صغيرة مختصرة تساعدنا على فهمها وتيسر تحليلها إحصائيا، بحيث توحدها قاعدة معينة كأن تشترك كل مجموعة في بعض من الصفات أو الخصائص (الطول، العمر، الوزن، الحالة الأسرية، مكان العمل، ...) بحسب ما يتطلبه البحث، و يعتمد هذا النظام على فكرتين أساسيتين هما: نوع البيانات (نوعية أو كمية)، وتصميم البحث(نوع المتغيرات مستقلة، تابعة، وسيطة).

هناك عدة شروط لابد من مراعاتها أثناء عملية التصنيف منها:

- إن يقوم التصنيف على أساس واحد وبموجب خاصية واحدة، في كل مستوى وذلك حتى لا تتداخل الفئات وتتشابك. مثال: تصنيف الطلاب على أساس الجنس : ذكور إناث، ثم في مستوى ثاني حسب المستوى (الليسانس، الماستر، ماجستير، دكتوراه) وهكذا....

- إن تكون فئات التصنيف شاملة بحيث تغطي كل الوحدات التي تتضمنها الخاصية، وأن تتطابق مع الفئات الموجودة في الواقع مثال: تصنف تخصصات قسم علم الاجتماع إلى : تربوي - تنظيم وعمل- حضري...إلخ، وفي بعض الحالات قد تكون هناك بعض الفئات الشاذة وفي هذه الحالة يمكن إضافة فئات تصنيف مثل: بيانات أخرى، أقل من، أكبر من... لتغطية هذه الحالات الشاذة.

3 - مرحلة الجداول : الترتيب الذي توضع فيه البيانات بعد فرزها وتنظيمها يسمى بـ (الجدول). والجداول على أشكال مختلفة ومتنوعة، وكل منها يصلح للاستخدام في حالات معينة إلاّ أنها جميعا تهدف إلى إبراز البيانات وتوضيحها في حجم مكثف ومصغر ، كما سنتطرق لذلك في المحاضرة التالية بإسهاب.

ثانيا – تفريغ البيانات:

بعد مراجعة البيانات يتم صبها وفق التصميم الذي أعده الباحث في برامج الترميز في نظام التفريغ اليدوي (المصفوفة)، أو بالاعتماد على برامج الإعلام الآلي. ويكون ذلك بالاعتماد بشكل كبير على الجداول التكرارية المختلفة بحسب نوع الدراسة ونوع البيانات المجمعة. واختيار أدوات جمع البيانات له علاقة مهمة بأساليب الترميز المتبعة من طرف الباحث والمؤدية إلى تفريغ سليم لها. ولذلك فالبناء المنهجي للاستبيانات وبطاقات الملاحظة وغيرهما من أدوات جمع البيانات تستدعي التفكير المسبق لمرحلة التفريغ، بمعنى أن الفئات المرتبة ترتيبا منطقيا للبيانات العامة ثم بيانات الفرضيات المقترحة في البحث يسهل صبها في برامج التفريغ حينما تكون مرتبة ومرقمة بشكل محضر مسبقا لها.

ومن أهم المراحل التي تمر بها البيانات الأولية لتسهيل عملية التفريغ الآتي :

1- الترقيم : ويتم عادة على ثلاث فترات هي:

أ- نقوم أولا بمنح رقما لكل استمارة أسئلة، لكل مقابلة، لكل ورقة ترميز، لكل بطاقة وثائقية، لكل شخص تمت ملاحظته أو لكل عنصر من عناصر التجربة.

ب- نمنح رقما لكل سؤال في الاستمارة أو في مخطط أو دليل المقابلة، لكل فئة من فئات التحليل، لكل جانب تمت ملاحظته، لكل سلسلة رقمية وكذا لأي متغير مأخوذ بعين الاعتبار.

ج- نعطي رقما لكل اختيار من الإجابات، وكذا لكل سؤال من الأسئلة الفرعية على مخطط أو دليل المقابلة، لكل سلوك محتمل للأشخاص موضوع الملاحظة ، لكل وحدة ذات دلالة أو معنى ... إلخ

2- الترميز: يعرف الترميز على أنه عملية تصنيف الاستجابات إلى فئات، كما يعبر على إحلال الرموز أو الدلائل مكان الكلمات والصفات والأسماء، إنه يسمح بمنح رمزا عادة ما يكون رقما أو حرفا و قد يكون حرف ورقما معا لمجموعة من المعطيات أو لمعلومة تم الحصول عليها انه يبدأ بأداة الجمع التي تستخدم كقاعدة لهذا الترميز لأنه يتضمن عادة ترقيما.

مثال: ترميز الحالات المرتبطة بالمستوى الدراسي، لنفرض انه لدينا المستويات التالية:

- متوسط ونرمز له بالرمز م أما السنة فنعبر عنها بالأرقام حيث يصبح لدينا: 1 م والذي يعني الأولى متوسط، 2 م ويعني الثانية متوسط ، 3 م ، 4 م ، وهكذا في المستويات الأخرى، ثانوي ورمزه ثا، أما الجامعي فنرمز له بـ ج وهكذا...

أ- خصائص الترميز:

- يجب أن لا يكون غامضا (رمز وحيد لكل معلومة)
- يجب أن يتكيف مع حاجيات المستعملين (سهولة الاستعمال)
- إمكانية التمديد والإدراج
- يجب أن يكون الرمز قصيرا قدر الإمكان.
- يجب أن يكون ذا معنى بقدر الإمكان (معبر)

ب- قواعد الترميز:

- يجب أن تتبع الأعداد الرمزية الإدراك الحدسي للمتغيرات الترتيبية على سبيل المثال، تخصص الأعداد الرمزية الأعلى للعلامات الأعلى.
- يجب أن تقع كل وحدة تحليل في فئة واحدة ووحيدة فقط.
- يجب أن يكون برنامج الترميز شاملا.
- يجب أن تكون الفئات محددة بشكل كاف لتضم جميع الفروقات الموجودة في إجابات المبحوثين.

ج- مدلول الترميز: يمكن للتوضيحات الآتية أن تساهم في إعطاء انسجاما ومنطلقا لمعنى الرمز:

- عندما يتعلق الأمر بمتغير الشدة، مثل الرضي أو درجة الاهتمام ، فإننا نرقم في اتجاه يذهب من الأقل شدة إلى الأكثر شدة، و ذلك انطلاقا من الرقم 1 إلى العدد الضروري المطلوب .
- عندما تكون هناك فئتين فقط بالنسبة إلى متغير واحد وأن هاتين الفئتين تتكرران لمرات عديدة، فترميزهما حينئذ سيكون دائما هو نفسه مستعملين الرقم 1 كدليل على الحضور والرقم 2 كدليل على الغياب مثلا وهكذا ...
- عندما لا يكون لفئات منطقا خاصا في علاقاتها ببعضها البعض، وذلك مثل الانتماء العرقي أو نوع الموسيقى المفضلة، فيمكن للأرقام أن توزع دون تمييز.
- أخيرا فإن بعض المتغيرات ترجع مسبقا إلى أرقام وأن الترميز يمنح من طرف الطبيعة العددية للمتغير، مثل تكرار ظاهرة ما، كالسن، الدخل .....إلخ .

د- ترميز الأجوبة المفتوحة:

القاعدة الأولى: نختار بالصدفة عددا معينا من الاستمارات بهدف الحصول على مجموعة متنوعة من الإجابات (ثلث الإجابات فيما يخص أقل من مئة استمارة ، و الربع بالنسبة إلى مئة فما فوق) .

القاعدة الثانية: انطلاقا من هدف السؤال نقوم بمقارنة الإجابات بعضها ببعض للنظر إن كانت تترجم مختلف ردود أفعال السؤال مرة ، مرتين ، ثلاث ، ونحاول إرجاعها إلى بعض ردود الأفعال الأساسية .

القاعدة الثالثة: نقوم باستخراج الأفكار الرئيسية في الإجابات المقدمة، وإعادة دمجها وتركيبها لغرض تحديد الفئات النهائية.

مثال : يمكن إبراز خمس أفكار رئيسية من ثلاث إجابات لمجموعة خاصة من التلاميذ لماذا تريد أن تلتحق بالشرطة ؟

*(كان لدي زملاء في الثانوية سجلوا في الشرطة، أعرف بعض رجال الشرطة). تظهر في إجابة التلميذ فكرتان أساسيتا هما :1 – الحفاظ على العلاقات الودية. 2 – معرفة أشخاص في المهنة.

*(قال لي الموجه في الثانوية إنني أصلح لذلك، ولم أكن أعرف شيء أخر يمكنني القيام به) وهنا أيضا تظهر فكرتان أساسيتان هما:1 – نصائح الخبير المختص . 2 – جهل بالاختيارات الأخرى .

*(لقد سبق لي وأن كنت في نادي الكشافة ولدي تجربة في المخيمات العسكرية) هنا تبرز لنا فكرة رئيسية واحدة هي: التجربة الملائمة للمهنة.

القاعدة الرابعة: يبقى بعد ذلك جمع هذه الأفكار الرئيسية الخمس في بعض ردود الأفعال الأساسية برجوعنا إلى مؤشر السؤال أو بصفة عامة إلى الإطار المفهومي لتعريف المشكلة، إذا ما كن نهدف إلى معرفة من هم الأشخاص الذين أثروا في الشخص ، فعندئذ يمكن إبراز ثلاث فئات و التمسك بالأفكار الخمس الرئيسية :

1 – الأشخاص المحيطين به . 2– أشخاص أجانب أو خارجين مثل الموجه. 3– ليس هناك أشخاص على الخصوص.

أما إذا كنا نهدف بالأحرى إلى معرفة مصدر التأثير فيما يخص الاختيار ، فسوف نحتفظ بفئتين يمكن أن تتضمنهما الأفكار الخمس الرئيسية و هما :1– التأثير المرتبط بالأشخاص . 2 – التأثير المرتبط بطبيعة العمل .

ثم نواصل بعد ذلك عملية التفريغ محاولين حصر الإجابات الجديدة في الفئات التي أقمناها ، فإن كانت بعض الإجابات لا تقبل هذه التفيئة نقوم بإدخالها في مراجعة التفيئة من أجل إضافة فئة جديدة.

ملاحظة: يجب أن نحتفظ بمجموعة إجراءات الترميز في كراسة خاصة مع التأكد من احترامنا للقواعد الخاصة بالترقيم، والدلالة المعطاة للرموز ولعملية الترميز الخاصة بالإجابات عن الأسئلة المفتوحة .

ثالثا– تحويل المعطيات و المراجعة:

بعد عملية التحقق من المعطيات تأتي مرحلة تحويلها إلى سند، بالطريقة اليدوية (مصفوفة)، وغالبا ما يكون بالطريقة الآلية بهدف معالجتها، حيث تنظم المعطيات الكمية بعد ترميزها في مصفوفة خاصة بواسطة جهاز الإعلامي الآلي عن طريق البرنامج الإحصائي spss ، ثم سيكون بإمكاننا القيام بطلبات المعالجة الآلية بهدف مراجعة البيانات الخاطئة لتصحيحها أو حذفها، حيث يتطلب الأمر فحص كل جدول تجميع لكل متغير، و بمجرد ما ينتهي ترتيب هذه المعطيات، لابد علينا من تقديمها في شكل يسمح لنا بتحليلها و ذلك بالطرق التالية :

(جداول و رسومات – تقليص فئات متغير– إنشاء متغيرات جديدة – اختبارات إحصائية – القياسات الوصفية ...إلخ)

مصفوفة تفريغ البيانات:هي عبارة جدول التفريغ البياني للمعطيات، بحيث يقسم هذا الجدول إلى صفوف و أعمدة فيكتب على العمود الأول أرقام الاستمارات أو المبحوثين و في الصف الأول أرقام الأجوبة كما هو مبين في الجدول التالي:

رقم ج / المبحوث	ج1 الجنس	ج2 السن	ج3 السكن	ج4 ...	ج5 ...	ج6 ...	ج7 ...	ج8 ...	ج9 ...	ج10 ...	ج11 ...	... ...	إلخ جواب
01													
02													
03													
04													
05													
06													
07													
08													
09													
10													
...إلخ													

إذا كان حجم العينة صغيرا يتم تفريغها يدويا، أما إذا كان حجمها كبيرا فيمكن الاستعانة بالكمبيوتر و بالبرامج الاحصائية مثل الحقبة الإحصائية للعلوم الاجتماعية SPSS ، SAS، ... MINITAB الخ.

رابعا– علاقة التبويب والتفريغ بالمفاهيم والفروض:

يرتبط تبويب البيانات وترميزها وتفريغها بمستويات متعددة من مراحل البحث الاجتماعي. حيث أن البناء المنهجي لفروض الدراسة المشتقة من مفاهيمها يستدعي البحث منذ البداية على مجموعة من القواعد التي تجعل من البيانات ذات دلالة إحصائية للصيغ المقدمة في فرضيات الدراسة. ولذلك يتوجب مراعاة الشروط التالية:

* يجب أن تكون الفروض قابلة للقياس من طريق إحصائي يعكس ضرورة اللجوء إلى البيانات لتحقيقها أو دحضها.

* يجب توضيح الأبعاد المستخرجة من مفاهيم الدراسة وتخطيطها في مسودات الباحث كي يسهل استخراج المؤشرات منها. وتحويلها إلى رموز وأرقام قابلة للتنظيم في جداول التحليل.

* الفروض تحمل ارتباطات بين متغيرات، ولذلك يحتاج الباحث إلى تصميم أساليب مناسبة لكل نوع من المتغيرات حتى يسهل إحداث التقاطعات الضرورية، خاصة منها في الجداول المركبة.

خامسا– علاقة التبويب والتفريغ بأدوات الدراسة ونتائجها:

يعد البناء السليم للأدوات البحثية والاستبيان منها خاصة مرحلة أساسية يعتمد عليها الترميز وتفريغ البيانات. ويستدعي ذلك التفكير في نوع الأسئلة الممكن تقديمها للمبحوثين وعلاقتها بالبحث عن نتائج منطقية تعكس الحقيقة التي هي عليها الظاهرة أو توصل إلى أقرب مسافة من التفسيرات الحقيقية لمسبباتها. وتظهر تلك العلاقة في مستويات ثلاث:

المستوى الأول: البيانات التي ينطلق منها الباحث في محل الظاهرة تستثيره نحو التفكير في مشكلة الدراسة وتحويلها إلى المستوى البحثي. ولذلك فهي تعبير أولي يستدعي التساؤل وفرض الفروض، وهنا تظهر البيانات بشكل خفي محفز يتطلب البحث في دلالات انتشاره لدى أفراد المجتمع. وهذا هو المستوى الأول الذي يشغل الباحث ويضعه في موقف الحيرة العلمية والاستفهام الموضوعي نحو الظاهرة.

مثال على ذلك بروز ظاهرة اختطاف الأطفال كظاهرة منتشرة في المجتمع الجزائري في السنوات الأخيرة. فمستوى الانتشار تثيره وسائل الإعلام، ويثيره حديث العامة من أفراد المجتمع، ويصبح التفكير في الظاهرة من جانبها الرقمي أكثر تأثيرا رغم أنها ليست ظاهرة جديدة في المجتمع. فهي تاريخيا موجودة في كل المجتمعات، إلا أن الرقم الخفي الذي يبرز الظاهرة يستفز الباحث بكمونه نحو البحث فيها وفي أسبابها. وهنا تظهر العلاقة التي أشرنا إليها للبيانات غير المعلنة بالتفكير في وضع تساؤلات وفروض لها.

المستوى الثاني: ينتقل فيه الباحث إلى جمع البيانات وترتيبها وترميزها وتحويلها إلى مستويات التحليل بالجداول والرسومات البيانية. وهنا تبرز العلاقة التي كانت كامنة في بداية البحث والتي تستدعي التنظيم والبناء المنطقي في عنوان الدراسة وتساؤلاتها وفروضها، بحيث تحمل ضمنها ضرورة البيانات كطريق منهجي يبعث على إيجاد التفسيرات المنطقية للفروض بالتدليل على صحتها أو دحضها. وفي هذه المرحلة تبرز أهمية المناهج والأدوات البحثية كوسائل ضرورية لتحويل البيانات إلى مستوى القابلية للتحليل والتأويل والمعالجة.

المثال الذي يوضح ذلك توضحه ظاهرة العنوسة في الجزائر، حيث ومن خلال رقم واحد تثيره الجهات الرسمية في سنة ما (15% في 2016 لم هن أكثر من 35 سنة) تجعل من الباحث يتجه نحو البحث في دلالة هذه الرقم بتحويله إلى بحث كامل تكون فيه الأرقام ضرورة لتفسير أسباب انتشار الظاهرة. لكن المستوى المشار إليه هنا هو في المجتمع الإحصائي أو العينة المقترحة للدراسة حيث يبني الباحث تصميما تكون فيه البيانات المجموعة بأدوات الدراسة ومرمزة ومفرغة في جداول إحصائية طريقا لتفسير أبعاد الظاهرة وأسبابها أو أي زاوية أخرى للعلاج.

المستوى الثالث: الذي يظهر في الأهداف التي وضعها الباحث لمسار البيانات التي استدعاها من محل وقوع الظاهرة، وهنا تبرز أهمية النظريات المفسرة التي ينتهجها الباحث في البحث عن الحقائق الموصلة إلى تحقيق الفروض أو دحضها في مرحلة النتائج. أين تظهر البيانات بصورة أخرى كامنة تحمل في دلالاتها أهمية الدراسة ككل وأهدافها التي وجدت من أجلها.

للتوضيح بمثال عن هذه العلاقة نثير قضية التسرب المدرسي كظاهرة تعتمد على بحث إمبريقي لتفسيرها، فالبيانات هنا تتجه نحو طرائق جمع البيانات وترتيبها وترميزها وتفريغها في جداول التحليل. فالأرقام التي قد يجمعها الباحث من مصادر أو وثائق رسمية لن تكون دلالاتها

الإحصائية ذات جدوى ما لم ترسم لها أهداف مسبقة للوصول إلى نتائج تشخص الظاهرة من زاويتها السوسيولوجية. بمعنى أن الأرقام الموجودة مسبقا في المؤسسات المعنية ليس لها أي دلالة سوى الاعتبار أن الظاهرة في تزايد أو أنها تسبب مشاكل معينة. لكن الطرق الإحصائية المبنية على خطوات البحث والمعتمدة على بيانات أخرى مجموعة من محل حصول الظاهرة يحول الظاهرة إلى مستوى آخر من التفسير وهو الدلالات الإحصائية لمعطيات اجتماعية أخرى مسببة لها أو ناتجة عنها أو محيطة بها. فإذا كانت الإحصاءات تشير مثلا إلى أن التسرب المدرسي في مرحلة التعليم الثانوي هي 10% فإن ذلك لا يستدعي سوى محاولة تقليص هذه النسبة. أما الدلالات المقصودة في المستوى الأخير للعلاقة بالنتائج فيظهر في بحث المتغيرات التي تسبب الظاهرة وتفسر أبعادها، وتفسح مجالات واسعة لبيانات أخرى لها دلالاتها الاجتماعية.

المحاضرة 03:العرض ألجدولي للمعطيات (الجداول البسيطة والجداول المركبة).

تعد الجدولة الإحصائية إحدى التقنيات المنهجية التي تمكن الباحث من التعبير عن البيانات الكمية والكيفية تعبيرا دقيقا، حتى تعكس صورة الواقع الذي أخذت منه من جهة، ومن جهة ثانية حتى تكون قابلة للتحليل بشكل يناسب احتياجات البحث الذي أعدت لأجله، وذلك من خلال تنظيمها بصيغة مجاميع متجانسة وبشكل ملخص ومختصر ودقيق. ومثل هذا التنظيم يسمى أحيانا التمثيل أو (التكوين الجدولي)، أي أن عملية العرض الجدولي للبيانات، عملية مشابهة لعملية تبويب البيانات، إذ توضع فيها البيانات على شكل جداول.

1- تعريف الجداول: بصفة عامة فإن الجداول هي وسيلة مختصرة لعرض الإحصائيات ونتائج التجارب، والتعبير عن البيانات. وتعرف الجدول على أنها شكل تقني مختصر تجمع فيه البيانات بشكل متناسب مع خصائصها البحثية التي جمعت من أجلها، فالجدول هنا عبارة عن ترتيب يضم مجموعة قيم متقاربة تسمى (الفئات) ويقابلها عدد المفردات التي تقع ضمن الفئة الواحدة تسمى (التكرارات)، كما أنه الطريق المنهجي نحو تحويل المعطيات إلى دلالاتها الإحصائية التي تنقل الظاهرة المدروسة من مستوى وقوعها إلى مستوى تفسيرها واستخلاص النتائج منها.

2 – شروط إعداد الجداول: ولتنظيم الجداول لابد من توافر جملة من الشروط هي :

أ- أرقام الجداول: ترقم الجداول ترقيما تسلسليا، من أول جدول رقم 01، ثم جدول رقم 02، ثم جدول رقم 03،...الخ.

ب-عناوين الجدول: يوضع عنوان مناسب للجدول، بحيث يكون هذا العنوان بسيطا وواضحا ومعبرا عن محتويات، وخصائص المتغير في حالة الجدول البسيط، أو عن العلاقات بين المتغيرات في الجدول المركب، ويكتب فوق الجدول.

ج- بيانات الجدول: يجب أن تكون بيانات الجدول بسيطة وواضحة، كما يجب أن تكون كافية حتى يتمكن القارئ من فهم محتوياتها بسهولة و دون الحاجة إلى نص توضيحي.

هـ- استعمال الرموز والاختصارات: يمكن استعمال الرموز والاختصارات في أعمدة الجدول على أن يوضح معناها في آخر الجدول إن لم تكن معروفة.

و- مصادر الجداول: في نهاية كل الجدول يجب أن نشير إلى مصدر معلوماته.

3- أنواع الجداول: والجداول على أشكال مختلفة ومتنوعة، وكل منها يصلح للاستخدام في حالات معينة إلا أنها جميعا تهدف إلى إبراز البيانات وتوضيحها في حجم مكثف ومصغر، ومن أهم الجداول المستعملة في البحوث الاجتماعية ما يلي:

أ- الجدول التكراري البسيط: يشير الجدول البسيط إلى ذلك النموذج الذي تجمع فيه البيانات بشكل بسيط، فتعرض في العموده الأول صفات أو خصائص مرتبة ترتيبا منطقيا يليه عمود أو عمودين، حيث يخصص الأول للتكرارات والثاني للنسب المئوية. مقابل الصفوف التي تعبر على تلك الصفات والمجاميع، ولهذا فالجداول البسيطة هي الجداول التي تصف المبحوثين بمقتضى متغير واحد فقط، وتسمى متغيراتها بالمتغيرات المستقلة الثابتة، هذا ما اعتاد عليه الباحثين في العلوم الاجتماعية حيث أنهم يقومون بتفريغ البيانات العامة في جداول بسيطة، سواء كان سؤال مغلق، أو اختياري أو مفتوح مبوب، وذلك للتعمق أكثر في واقع الموضوع المعالج. وفي هذا الصدد يقول موريس أنجرس "...من الممكن بناء جدول لكل متغير، وكذلك الأمر بالنسبة إلى كل سؤال في الاستمارة أو إلى كل فئة ناتجة عن سحب كمي بتقنية أخرى. عندما يتضمن الجدول متغيرا واحدا فقط فيتعلق الأمر عندئذ بالجدول ذو المدخل الواحد وتعتبر بمثابة وصف أولي للبيانات وبالطبع يمكن تكوين جداول أولية بعدد أسئلة الاستمارة .

مثال1: تقديرات علامات التلاميذ المتحصلين على شهادة البكالوريا لثانوية ما خلال الموسم الدراسي 20/19:

جدول رقم (01) يبين تقديرات الناجحين في البكالوريا دورة 2020 لثانوية ما

التقدير	التكرار (fi)
مقبول	65
قريب من الجيد	55
جيد	41
جيد جدا	34
ممتاز	5
المجموع	200

مثال2: جدول يمثل توزيع الناجحين لهذه الثانوية حسب الجنس كما يلي:

الجنس	التكرارات (fi)	النسبة المئوية %
ذكر	116	% 58
أنثى	84	% 42
المجموع	200	% 100

ب . الجدول التكراري ذو الفئات: المقصود بالفئات هي البيانات المتشابهة في خصائصها والتي تختصر عرض عدد كبير منها في أقسام متجانسة مقابل عدد. فإذا رجعنا للمثال01 فإن متغير التقدير و الذي يغبر في الأصل عن فئات تحصر فيها علامات الطلبة والتي نستطيع تمثيلها في الجدول التالي:

الفئات	التكرار (fi)
مقبول من 10.00 إلى 11.99	65
قريب من الجيد من 12.00 إلى 13.99	55
جيد من 14.00 إلى 15.99	41
جيد جدا من 16.00 إلى 17.99	34
ممتاز من 18.00 إلى 19.99	5
المجموع	200

جـ . الجدول التكراري المتجمع الصاعد: يقصد به تجميع تكرار كل فئة على جميع التكرارات السابقة لها بحيث يكون مجموع التكرار التصاعدي للفئة الأخيرة مساوٍ لمجموع التكرارات.

مثال: حساب التكرار المتجمع الصاعد للمعطيات التالية:

حدود الفئات	التكرار (fi)
05 - 10	5
11 - 20	7
21 - 30	12
31 - 40	6
41 - 50	18
51 - 60	2
المجموع	50

وذلك بإضافة عمود نسجل فيه الحدود العليا لكل فئة كما يلي:

حدود الفئات	التكرار (fi)	أقل من
05 - 10	5	10
11 - 20	7	20

30 - 21	12	30
40 - 31	6	40
50 - 41	18	50
60 - 51	2	60
المجموع	50	

ثم نضيف عمود آخر لحساب ت م الصاعد ابتداء بأول تكرار على أساس أنه ممثل لجميع من هم اقل من 10 يعني (5)

نضيف لها تكرار الفئة الموالية (5 + 7) فنحصل على تكرار الفئة الثانية الأقل من 20 نحصل على (12)

نضيف لها تكرار الفئة الموالية (12+12) فنحصل على تكرار الفئة الثانية الأقل من 30 فنحصل على (24) وهكذا مع بقية الفئات بحيث نحصل في نهاية الفئة تكرار متجمع صاعد مساوٍ لمجموع التكرارات.

حدود الفئات	التكرار (fi)	أقل من	التكرار المتجمع الصاعد ↑N
10 - 05	5	10	5
20 - 11	7	20	12
30 - 21	12	30	24
40 - 31	6	40	30
50 - 41	18	50	48
60 - 51	2	60	50
المجموع	50		

وفي هذا الصدد يجب مراقبة التأثير المفترض للمتغير المستقل على المتغير التابع، و يتعلق الأمر بتموقع المتغيرين حيث يوضع المتغير المستقل أفقيا بينما التابع عموديا.

د . الجدول التكراري المتجمع النازل(الهابط): نقصد به تجميع تكرار كل فئة على جميع التكرارات التالية لها بحيث يكون مجموع التكرار التنازلي للفئة الأولى مساوٍ لمجموع التكرارات.

مثال: حساب التكرار المتجمع النازل لنفس المثال السابق

حدود الفئات	التكرار (fi)
10 - 05	5
20 - 11	7
30 - 21	12
40 - 31	6
50 - 41	18
60 - 51	2
المجموع	50

نضيف عمود نحدد فيه الحدود الدنيا للفئات لكل فئة

حدود الفئات	التكرار (fi)	أكبر من
10 - 05	5	0
20 - 11	7	11
30 - 21	12	21
40 - 31	6	31
50 - 41	18	41
60 - 51	2	51

المجموع	50	

نضيف عمود آخر لنحسب من خلاله التكرار المتجمع النازل ننطلق من الخلية الأولى لأول فئة (أكبر من 0) فنسجل فيها قيمة مجموع التكرارات أي (50) على أساس أنها هي التكرار المتجمع النازل لها.

نحسب قيمة التكرار المتجمع النازل للفئة الموالية (أكبر من 21) بطرح تكرارها من التكرار المتجمع الهابط السابق أي (50 – 5) فنحصل على 45.

نحسب قيمة التكرار المتجمع النازل للفئة الموالية (أكبر من31) بطرح تكرارها من التكرار المتجمع الهابط السابق أي (45–7) فنحصل على 38.

وهكذا مع بقية الفئات حتى نحصل على التكرار المتجمع النازل للفئة الأخيرة بحيث يكون مساوٍ لتكرارها. (وينتهي بالصفر كقيمة غير مسجلة)

حدود الفئات	التكرار (fi)	أكبر من	التكرار المتجمع النازل N↓
05 – 10	5	0	50
11 – 20	7	11	45
21 – 30	12	21	38
31 – 40	6	31	26
41 – 50	18	41	20
51 – 60	2	51	2
المجموع	50		

هـ- الجدول المزدوج أو المركب: يأخذ هذا النوع من الجداول في الحسبان أكثر من متغير واحد، حيث يلجأ الباحث أحيانا في الجدول المركب إلى الأسلوب الذي يساعد على التعرف على الارتباط أو العلاقة بين خاصيتين أو متغيرين، حيث أنها تعمل على تسهيل تجميع الحالات التي يحدث أن ترتبط ببعضها في اثنين أو أكثر من الفئات. وفي هذا الصدد يجب مراقبة التأثير المفترض للمتغير المستقل على المتغير التابع، و يتعلق الأمر بتموقع المتغيرين حيث يوضع المتغير المستقل أفقيا بينما التابع عموديا.

يتم بناء الجدول الذي يربط بين متغيرين في نفس الوقت بالخطوات التالية:

- تحديد المتغير المستقل والمتغير التابع
- تحديد فئات كل متغير
- تحديد المتغير المستقل في السطر الأفقي الأول والمتغير التابع في العمود الأول.
- وضع العلامات التي تمثل التكرار ومن ثمة الأرقام.

مثال: توزيع مشترك بين التحصيل الدراسي والحوار بين الأساتذة و التلاميذ:

الحوار بين الأستاذ والتلميذ / التحصيل الدراسي	نعم		لا		المجموع	
	التكرار	النسبة	التكرار	النسبة	التكرار	النسبة
يؤثر	15	%30	45	%56.25	60	%46.15
لا يؤثر	35	%70	35	%43.45	70	%53.85
المجموع	50	100	80	100	130	100

المحاضرة 04:التمثيل البياني للمعطيات (الرسوم البيانية).

إن الطرائق المستعملة في تمثيل المعطيات الإحصائية في البحوث الاجتماعية بيانيا كثيرة جدا ومتنوعة، وهي تتبع تصورات واضعها وهدفه من عرضها، فمنها التمثيل البياني باستخدام أعمدة أو خطوط أو رسوم بيانية أو دوائر نسبية، يمكن من خلال إلقاء نظرة عليها تكوين فكرة

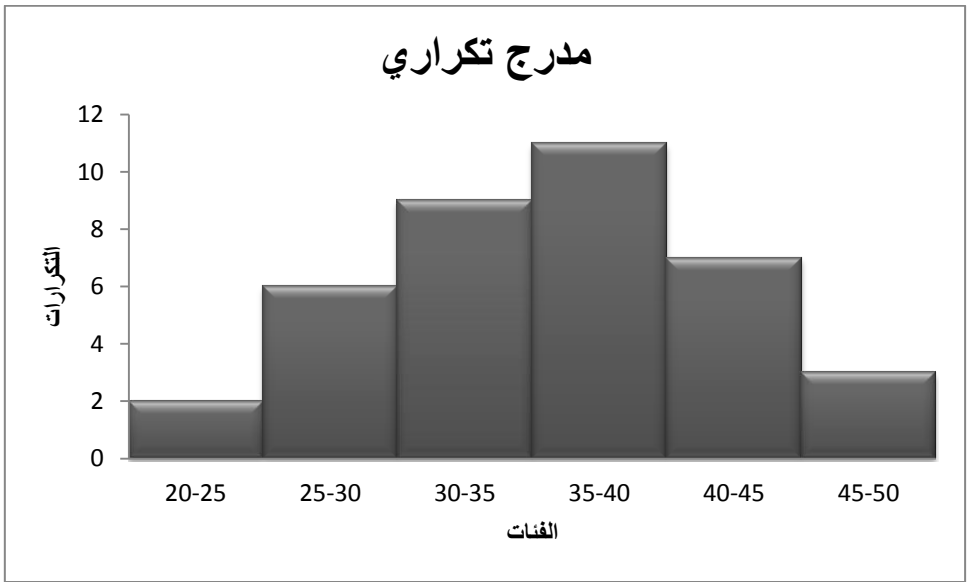
واضحة عن المسألة المدروسة بعد جدولتها، وهنا يراعى أن تكون المساحات متناسبة مع الأرقام التي تمثلها. وتعد الدائرة النسبية أبسط الأشكال وأحسنها بيانيا في عرض نسب مئوية من كمية واحدة.

1- المدرج التكراري: هو تمثيل بياني يعرض البيانات العددية المنظمة في فئات متساوية، في شكل سلسلة من المستطيلات المتلاصقة ممثلة في رسم، بحيث يمثل عرض المستطيل على المحور الأفقي (X) طول الفئة (i) ويمثل طول المستقيمات على المحور العمودي تكرار الفئة (ni).

مثال: ليكن لدينا التوزيع التكراري لأعمار مجموعة من العمال والممثلة في الجدول التالي:

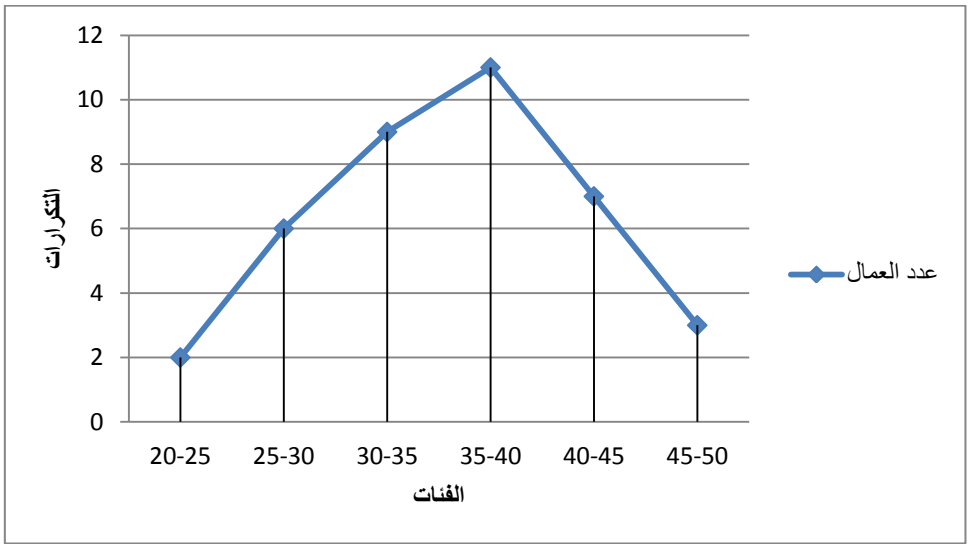
فئات العمر	25-20	30-25	35-30	40-35	45-40	50-45
عدد العمال	2	6	9	11	7	3

وعلى ضوء البيانات في الجدول التكراري يمكن أن نرسم المدرج التكراري التالي:



2- المضلع التكراري: هو عبارة عن مستقيمات تصل منصفي القاعدتين العلويتين لمستطيلين متتاليتين لمدرج تكراري (مراكز الفئات)، و يسمح بمعرفة توزيع قيم الظاهرة المدروسة كما يوضح الرسم التالي للجدول السابق:

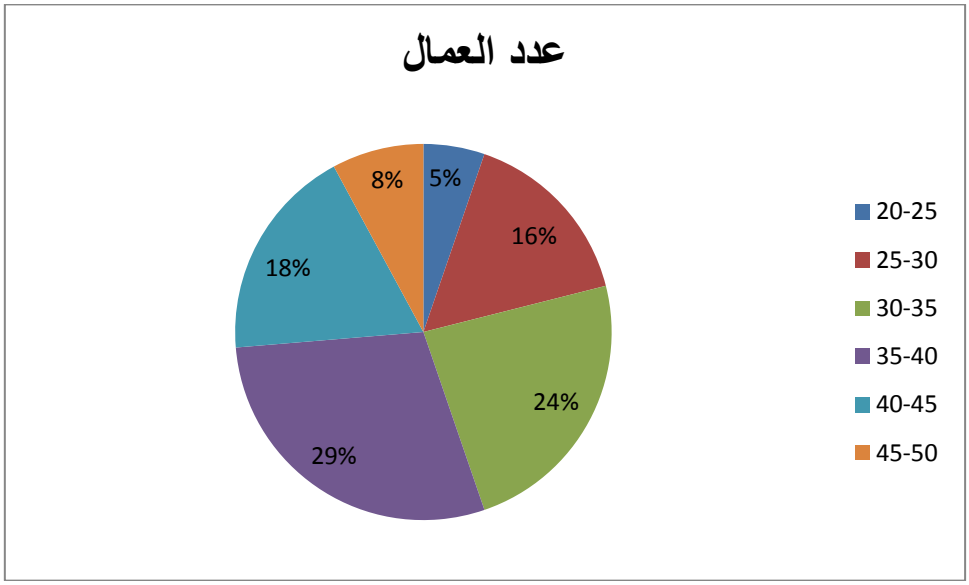
بحيث أن مركز الفئة = (الحد الأدنى + الحد الأعلى) / 2



3- المنحنى التكراري: يعتبر المنحى التكراري من الوسائل المهمة لتمثيل التوزيعات التكرارية بيانيا ومن الضروري أن ندرك أن المساحة المحصورة بين المنحنى والمحور الأفقي مساوية بالضبط للمساحة بين المدرج وذلك المحور. لذلك لا يختلف المنحنى التكراري عن المضلع التكراري في الرسم، والاختلاف الوحيد بين المحنيين هو طريقة التوصيل بين النقاط، إذ توصل النقاط في المنحنيات التكرارية بخطوط منحنية تمر بأكبر عدد ممكن من النقاط، بينما نوصل النقاط في المضلع التكراري بخطوط مستقيمة تمر بمراكز الفئات.ويمكن أن نوضح ذلك من خلال الشكل التالي للمثال السابق :

4- التمثيل بالدوائر النسبية: هذا النوع من التمثيل هو عبارة عن قطاعات مقسمة داخل دائرة كل قطعة تمثل جزء من الظاهرة الكلية ويستعمل في هذه العملية التكرارات النسبية.

مثال : نمثل معطيات الجدول السابق في دائرة حسب الشكل التالي:



ملاحظات مهمة حول الرسوم البيانية:

. تعرض الرسومات البيانية في البحوث الاجتماعية لإضفاء صورة أوضح عن البيانات المقدمة في الجداول. إلا أن تقديمها يحتم على الباحث مراعاة الأساليب الإحصائية التي تربط بعض أنواع الاختبارات .

. علاقة الرسوم البيانية بالفروض المقدمة في البحث، تعكس نفس الصورة التي تقدمها الجداول. ولذلك يطلب من الباحث توظيفها عند التحليل بنفس المستوى الذي تقدمه المعطيات الجدولية.

. علاقة الرسوم البيانية بنتائج البحث هي أيضا نفسها التي تعكسها الجداول، ولذلك يطلب من الباحث توظيفها أيضا عند التحليل مثلما تقدمها المعطيات الجدولية.

ـ تغني البرامج التي تعالج المعطيات الاجتماعية في الحواسيب والتي يمكن من خلالها اختيار ما يناسب الباحث من الرسومات البيانية مختلف الجهود اليدوية التي يمكن أن تستغرق وقتا أطول،ومن بينها برنامج الإكسيلExcel و برنامج الحزم الإحصائية للعلوم الاجتماعية spss .

المحاضرة 05: طرق و تحليل المعطيات.

يمثل التحليل محورا رئيسيا ترتكز عليه البحوث الاجتماعية في الانتقال من مستوى التصورات النظرية والافتراضات، إلى مستوى التجريب والاختبار، ومنه إلى استخلاص النتائج التي يستهدفها الباحث في دراسته. لهذا السبب فإن تحليل المعطيات يعتمد بشكل كبير على قدرة الباحث لتجزئة الواقع، وذلك بعرض كل الملاحظات ثم العلاقات السببية أو التبعية المتبادلة بين المتغيرات، والأهمية النسبية للروابط بين مختلف الظواهر. وفي هذه المرحلة سيكون من الممكن التحقق من الفرضية وهدف الدراسة من خلال المعطيات المهيئة، إننا نتقصى وبطرق مختلفة المعطيات التي نهدف من خلال الحصول عليها إلى بلوغ أكبر قدر ممكن من المعاني على أساس ما كنا نبحث عنه منذ البداية، والتحليل أنواع ومستويات تتعلق بنوع الدراسة ومجال تخصصها، نجد من بينها ما يلي :

1- التحليل الوصفي للمعطيات: التحليل الوصفي للمعطيات هو طريقة لدراسة الظواهر أو المشكلات العلمية من خلال القيام بالوصف بطريقة علمية ودقيقة، من خلالها يمكن للباحث من تحليل المعطيات والبيانات التي تم جمعها تحليلا منطقيا وواقعيا لتأثير متغيرات متنوعة على ظاهرة معينة. ومن ثم الوصول إلى تفسيرات منطقية لها دلائل وبراهين تمنح الباحث القدرة على وضع أطر محددة للمشكلة، ويتم استخدام ذلك في تحديد نتائج البحث. وقد يتصور البعض أن التحليل الوصفي أو الكيفي ما هو إلا مجرد وصف بيانات وخصائص ما هو قيد الدراسة، من خلال الإجابة عن الأسئلة : من وماذا وأين ومتى وكيف، بل الفكرة الكامنة وراء هذا النوع من التحليل هو دراسة الترددات، والمتوسطات وغيرها من الحسابات الإحصائية. حيث يلجأ الباحث في حالة هذا النوع من التحليل إلى استخدام الأساليب الإحصائية الوصفية التي تصف الظاهرة أو المشكلة موضوع البحث ولا تتطرق إلى التعمق بدراسة الظاهرة والتعرف على بعض نواحي العلاقة أو الارتباط بمتغيرات الدراسة.

ومن أهم الأساليب الإحصائية الوصفية: التكرار، الوسط الحسابي، المنوال، والوسيط والانحراف المعياري والتشتت.أما في حالة التعمق في الدراسات الارتباطية ووضع الفروض قصد فحصها إحصائيا، يلجأ الباحث إلى استخدام مقاييس الارتباط مثل اختبار سيبرمان والانحدار وغيرها من المقاييس. فالتحليل الوصفي يرتبط كثيرا بالإحصاء الوصفي الذي يعتمد هو بدوره على القوانين والحسابات الرياضية التي تطبق قوانينها في مستويات ومجالات مختلفة بحسب خصائص الظاهرة و طرق علاجها.

فالمسار الإحصائي هو الطريق الذي يقود الباحث نحو توظيف الأساليب الإحصائية في توصيف الظواهر من منطلق كمي، تكون فيها البيانات المجموعة من المبحوثين أو الملاحظة لديهم هي الركيزة التي يبنى عليها تصميم البحث، بداية من وضع الفروض وتساؤلاتها ومرورا بتنظيمها في الجداول والرسومات البيانية ووصولا إلى استخلاص النتائج. فالتحليل الوصفي كما أشرنا ينبني على مختلف الأساليب الإحصائية التي تحول الظاهرة من موقعها الحاصل في محلها إلى مستوى الدلالات الرقمية التي تجعلها ملاحظة و مقاسة بأبعاد رياضية، ويتفق كثير من المختصين في تقسيم الإحصاء في هذا المجال إلى قسمين (وصفي واستدلالي).

أ - **الإحصاء الوصفي:** يهدف إلى إدماج وتلخيص البيانات الرقمية بغية تحويلها من مجرد كمّ من الأرقام إلى شكل أو صورة أخرى يمكن فهمها واستيعابها بمجرد النظر. ومن أغلب الأساليب المستخدمة مقاييس النزعة المركزية ومقاييس التشتت ومقاييس الارتباط والانحدار. ويتوقف استخدام أيّا منها على نوعية البيانات ومستوى القياس، سواء أكان اسميا أو وصفيا أو ترتيبيا أو فئويا أو نسبة ... فالمقاييس الوصفية لا تفيد في حدّ ذاتها في الاستنتاج أو التنبؤ، وإنما تصف الكيفية التي تتوزع بها البيانات التي تمّ الحصول عليها فقط.

فالوصف في الإحصاء يعتمد على مجموعة من المقاييس والاختبارات التي لا تتجاوز في مستواها كشف الحقائق وإظهار الارتباطات وتحديد الخصائص العامة للمتغيرات والظواهر. لكنها لا تصل إلى حد النتائج والاستدلالات الخاصة بالمجموعات الأساسية للظاهرة.

ب – الإحصاء الاستدلالي: يستند إلى مجموعة من النظريات الإحصائية أهمها نظرية الاحتمالات ونظرية العينات، اللتان تمثلان حلقة الوصل بين الإحصاء الوصفي والإحصاء الاستدلالي ... ويهدف هذا الأخير إلى الوصول إلى تعميمات عن مجتمع الدراسة من خلال العينة المسحوبة من هذا المجتمع، ويشمل هذا النوع من الأساليب الإحصائية، الاحتمالات، العينات، اختبار الفروض مثل اختبار ك2 .

مقارنة بين الاحصاء الوصفي و الاحصاء اللاستدلالي	
الوصفي Descriptive	**الاستدلالي Inferential**
– طرق تنظيم وتلخيص ووصف البيانات وصفا كميا. **– مجموعة من المفاهيم والأساليب الإحصائية التى تستخدم فى تنظيم وتلخيص وعرض البيانات بهدف إعطاء فكرة عامة عنها.** **– ملخص جيد لمجموعة كبيرة من المعلومات والبيانات** **– أهم صور التصنيف جداول التوزيع التكراري والرسوم البيانية التي تعبر عن هذا التوزيع** **– أما التخليص فيتخذ ثلاثة صور هي :** *** النزعة المركزية " المتوسط – الوسيط – المنوال "** *** التشتت " المدى – الانحراف المعياري – نصف المدى الربيعي "** **العلاقة أو الارتباط والانحدار**	**– مجموعة من الأساليب الإحصائية المستخدمة للتوصل إلى استنتاجات من بيانات العينة إلى المجتمع الأكبر** **– يشير إلى طرق الاستدلال عن المجتمع من بيانات العينة** **– عملية اتخاذ قرار منطقي باستخدام بيانات العينة وأسلوب إحصائي مناسب** **– يعتمد على افتراضين أساسيين هما :** **– العشوائية فى اختيار العينة المستخدمة فى الدراسة** **– التوزيع الاعتدالي للمتوسطات** **– ومنه : اختبار"ت" – تحليل التباين – اختبار مان ويتنى – النسبة الحرجة – فريدمان – كروسكال واليز –ولكوكسون –كا2**

المحاضرة 06: أساليب التحليل الاجتماعي.

علم الإحصاء أحد أهم العلوم المستخدمة في كافة المجالات العلمية والتربوية والاجتماعية وغيرها من أجل تفسير الظواهر الغامضة، والمتضاربة، والشائكة، وحل المشكلات والتنبؤ العلمي بالمستقبل وكيفية الاستفادة منه وتجنب الأخطاء والأضرار، عن طريق اتباع العديد من الأساليب الإحصائيّة التي تقوم على جمع البيانات، وتنفيذها، وفحصها، وتقييمها، واستخلاص النتائج الأولية والثانوية منها، خلال مدة زمنية معينة وبالاستعانة بأفراد وأدوات تعيين على تحقيق الغاية من اتباع الأسلوب الإحصائي والذي يجب أن يتناسب مع نوع البحث العلمي وطبيعته، وسنقدم في هذا المقال أهم المعلومات عن الأساليب الإحصائية في البحث العلمي بالتفصيل.

1– أنواع الأساليب الإحصائية المستخدمة في البحث العلمي: يستطيع الباحث أو الباحثون في أي بحثٍ علمي كان حول أي ظاهرة أو مشكلة علمية يحاولون تفسيرها أو إيجاد الحلول العلمية والمنطقيّة لها اتباع أحد الأساليب الإحصائيّة أو أكثر من أسلوبٍ بحسب طبيعة المعلومات، ونوع المشكلة وغيرها من العوامل المتحكمة في البحث العلمي، ومن تلك الأساليب ما يلي:

أ– الأسلوب الإحصائي الوصفي: هو الأسلوب الذي يقوم على جمع البيانات والمعلومات حول ظاهرةٍ ما ثم تنظيمها وترتيبها وتبويبها، عن طريق استخدام الجداول والرسومات، ويمكن الاعتماد على قوانين الإحصاء في ذلك ومنها: قانون الوسط والوسيط والمنوال والنزعة المركزيّة، ويتم استخدامه في البحوث العلميّة من أجل التقييم واستخلاص النتائج والاستدلال على الظاهرة المدروسة ومدى انتشارها أو أهميتها.

ب– الأسلوب الإحصائي التحليلي أو الاستنتاجي :هو الأسلوب الذي يقوم على جمع البيانات والمعلومات عن طريق الاستقصاء والاختبار والاستفتاء والاستبيانات وتحديد عينة عشوائية أو انتقائية لإجراء الدراسة البحثية عليها حول ظاهرةٍ ما من أجل الوصول إلى كافة النتائج الممكنة مع تحييد رأي الباحث الشخصيّ حول تلك الظاهرة، ومن ثم تحليل النتائج بحسب المعطيات الموجودة بين يديّ الباحث، ثم الخروج بالاستنتاج أو مجموعة استنتاجات منطقيّة وعقلانيّة يُمكن تطبيقها والاستفادة منها في الواقع، وكذلك توصيات للمستقبل.

2- **مراحل الأسلوب الإحصائي الاجتماعي** : أساليب الإحصاء و الاستبيان تتكون من عدد من المراحل حيث أن هذه الأساليب قد تطورت حتى وصلت إلى هذا الشكل الذي نعرفه في يومنا هذا حيث تقوم هذه الأساليب بجمع بيانات إحصائية دقيقة حول موضوع و خطوات أسلوب البحث الإحصائي العلمي وتنقسم هذه المراحل إلى ما يلي :

أيًا كان نوع الأسلوب الإحصائيّ المستخدم فلا بد لها جميعًا من مراحل وهي:

أ- **مرحلة تحديد مشكلة البحث:** إن أول خطوة يجب أن يقوم بها الباحث في بحثه عن التحليل الإحصائي هي تحديد المشكلة أو الموضوع الذي يريد أن يدرسه.

ب- **مرحلة جمع البيانات:** تتضمن هذه المرحلة جمع البيانات المتعلقة بموضوع بحث ما من مصادر مختلفة، قد تكون المصادر المعتمد عليها إما :

مصادر بيانات أولية : وهي البيانات التي تم تجميعها عبر الباحثين أو من وكالات بحثية معينة أو عن طريق أحدى مكاتب البحث العلمي، وتكون بيانات هذا المصدر غير مرتبة.

مصادر بيانات ثانوية: هي مصادر بيانات تم جمعها من الباحثين أو من وكالات بحثية معينة أو عن طريق أحدى مكاتب البحث العلمي وقد استُخدمت بالفعل في بحوث ودراسات أخرى، وتتميز هذه البيانات بأنها مرتبة ومنظمه.

ج- **مرحلة عرض وتنظيم البيانات** : حيث يقوم الباحث بجمع البيانات وتدوينها من مصادر ثانوية بشكل منظم، وتنظيمها يكون عبر الإجراءات التالية :

دراسة البيانات وتنظيمها عبر حذف المتناقضات فيها وحذف البيانات الغير دقيقة منها.

تصنيف البيانات بحسب الخصائص المشتركة فيما بينها .

وضع البيانات في اطار جداول .

العرض والذي يكون عبر مخططات أو رسوم بيانية .

د- **مرحلة تحليل البيانات الرقمية:** بعد تدوين وتنظيم البيانات يبدأ الباحث بمرحلة تحليل البيانات العدديّة لفهم الموضوع بشكل أكبر، كما أنه لتحليل البيانات الرقمية يتوجب على الباحث أن يستخدم قوانين الإحصاء الرياضي ومنها : المتوسط الحسابي – مقاييس التشتت – الارتباط – مقاييس النزعة المركزية – تحليل الالتواء – تحليل المدى والمنطق والعديد من قوانين الإحصاء و الاستبيان الأخرى.

هـ- **مرحلة تفسير البيانات التي توصل إليها** : تتضمن هذه المرحلة تفسير البيانات الرقمية والحصول منها على الاستنتاجات ، حيث يتوجب على الباحث في هذه المرحلة أن يكون يمتلك درجة كبيرة من الخبرات والمهارات ، و أن يتمتع بقدر كافي من الصدق والأمانة والمصداقية العلمية.

3- أهمية استخدام أسلوب البحث الإحصائي الاجتماعي

أهمية خطوات أسلوب البحث الإحصائي العلمي تظهر بشكل واضح في عدد من النقاط والتي سنوردها فيما يلي:

أ ـ فهم المعلومات والبيانات العلمية: ويتم فهم المعلومات والبيانات عبر تطبيقها في عمليات التحليل الإحصائي، حيث يكون بإمكان الباحث أن يقوم بفهم كافة المعلومات والبيانات الموجودة لديه كبيانات رقمية .

وقد قدمت عملية التحليل الإحصائي تبسيط لمجموعة من المعوقات والمشكلات التي واجهت الباحثين، ومنها ما يلي:

– فهم وتحويل الكميات الضخمة من المعلومات والبيانات إلى معطيات ذات معنى ويمكن استثمارها والتعامل بها.

– تساعد تلك المعطيات والمعلومات الباحث في إيصال النتائج التي يحاول الوصول إليها، بحيث تكون نتائج دقيقة ومنطقية ومفهومة بشكل واضح وعلى درجة عالية من الموثوقية والاعتمادية .

فلو لم يكن موجود التحليل الإحصائي ، لأصاب الباحثون حيرة حول ذلك الكم الهائل من المعلومات والبيانات والتي لم يستطيع الباحثون استنباط المعلومات الدقيقة منها اللازمة لإتمام بحوثهم العلمية.

ب ـ الوصول إلى الأدلة والبراهين: حيث يكون بإمكان الباحث عبر استخدام التحليل الإحصائي أن يتوصل إلى عدد من الأدلة والإثباتات بشكل بسيط ودون استنزاف كبير للوقت والجهد ، حيث أن هذه البيانات والمعلومات تساعد الباحث في تأكيد صحة بحثة العلمي. أن الجداول والرسوم البيانية الخاصة بالتحليل الإحصائي تقوم بتوضيح معظم المعلومات التي تتعلق بالموضوع الذي يقوم الباحث بالعمل عليه بدقة عالية وجودة ممتازة .

ج ـ دقة النتائج التي يصل البحث إليها: أهم شروط عملية التحليل الإحصائي العلمي تتلخص في جودة ودقة النتائج والبيانات التي يصل إليها الباحثون عن طريق تلك العملية.

د ـ الربط بين البيانات والنتائج المترتبة عليها: التحليل الإحصائي دائمًا ما يرتكز على جمع العلاقات والروابط والمتغيرات بين كافة المعلومات والبيانات التي يقوم الباحث بجمعها وفحصها وفهمها وربطها بالنتائج التي تم التوصل إليها في نهاية عملية التحليل الإحصائي العلمي الذي قام به.

المحاضرة 07: المقاييس والمعاملات الاحصائية.

تعد العمليات التي يعتمد عليها الباحث في تحليل المعطيات الاجتماعية والتي تستند إلى مختلف الأساليب الإحصائية، طريقا منهجيا ضروريا لتحويل الظواهر الاجتماعية إلى شكل تكون فيه قابلة للقياس. ولذلك يلجأ الباحثون إلى تصميم بحوثهم وفق تصورات متكاملة بين ما تقتضيه المنهجية العلمية للبحث، وما تتطلبه الأساليب الإحصائية التي تقيس الفروض التي تُقترح في تلك الدراسات. ولذلك تقتضي الضرورة منا أن نشير في هذا المقياس إلى بعض الأساليب التي تعالج بها الدراسات الأمبريقية دون التعمق فيها، تاركين ذلك لتخصص الإحصاء كمقياس مستقل للتفصيل فيها.

أولا– مقاييس النزعة المركزية ومقاييس التشتت:

تستخدم مقاييس النزعة المركزية بكثرة في العلوم الاجتماعية والإنسانية، فهي تقوم بتلخيص البيانات ذات الصفة المشتركة في قيمة واحدة . فمقاييس النزعة المركزية تساعد في:

– ايجاد ذلك الرقم المتوسط الذي يدل على خصائص أرقام مجموعة من المجموعات فيكفي أن ننظر الى ذلك الرقم المتوسط لنعرف الكثير عن خصائص هذه المجموعة من الأرقام.

– المقارنة بين عدة مجموعات في وقت واحد، فنقول مثلا أن هذه المجموعة أقوى من تلك، وذلك اعتمادا على مقارنة هذه المتوسطات بعضها ببعض. و من اهم مقاييس النزعة المركزية نجد:

1– المتوسط أو الوسط الحسابي Moyenne:

أ- **الوسط الحسابي للبيانات غير المبوبة** : يعرف الوسط الحسابي بشكل عام على أنه مجموع القيم مقسوما على عددها . فإذا كان لدينا n من القيم ، ويرمز لها بالرمز $x_1, x_2, \ldots, x_n$ فإن الوسط الحسابي لهذه القيم ، ونرمز له بالرمز $\bar{x}$ يحسب بالمعادلة التالية:

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

$$\bar{x} = \frac{x_1 + x_2 + \ldots + x_n}{n} = \frac{\sum_{i=1}^{n} x_i}{n}$$

حيث يدل الرمز Σ على المجموع .

مثال1: فيما يلي درجات 8 طلاب في مقرر مقياس الإحصاء : 40، 36، 40، 35، 37، 42، 32، 34 .

والمطلوب إيجاد الوسط الحسابي لدرجة الطالب في الامتحان .

الحل لإيجاد الوسط الحسابي للدرجات تطبق المعادلة السابقة كما يلي:

$$\bar{x} = \frac{x_1 + x_2 + \ldots + x_n}{n}$$

$$= \frac{34+32+42+37+35+40+36+40}{8} = \frac{296}{8} = 37$$

أي أن الوسط الحسابي لدرجة الطالب في اختبار مقرر الإحصاء يساوي 37 درجة.

ب- الوسط الحسابي للبيانات المبوبة: من المعلوم أن القيم الأصلية ، لا يمكن معرفتها من جدول التوزيع التكراري ، حيث أن هذه القيم موضوعة في شكل فئات ، ولذا يتم التعبير عن كل قيمة من القيم التي تقع داخل حدود الفئة بمركز هذه الفئة، ومن ثم يؤخذ في الاعتبار أن مركز الفئة هو القيمة التقديرية لكل مفردة تقع في هذه الفئة.

فإذا كانت k هي عدد الفئات ، وكانت $x_1, x_2, \ldots, x_k$ هي مراكز هذه الفئات، و $f_1, f_2, \ldots, f_k$ هي التكرارات ، فإن الوسط الحسابي يحسب بالمعادلة التالية:

$$\bar{x} = \frac{x_1 f_1 + x_2 f_2 + \ldots + x_k f_k}{f_1 + f_2 + \ldots + f_k} = \frac{\sum_{i=1}^{k} x_i f_i}{\sum_{i=1}^{k} f_i}$$

xi هي القيم التي يأخذها المتغير الإحصائيx

K هي عدد القيم xi التي يأخذها المتغير الإحصائيx

ni هي التكرارات المطلقة المناسبة للقيمxi

وهذا هو المتوسط الاكثر استعمالا في العلوم الاجتماعية لأنه ينطبق على كل انواع العينات كما أن هناك عدة أنوع اخرى من المتوسطات كمتوسط الهندسي ،المتوسط المرجح ، المتوسط التوافقي و المتوسط الهندسي ...الخ.

مثال2: الجدول التالي يعرض توزيع 40 تلميذ حسب أوزانهم.

فئات الوزن	32-34	34-36	36-38	38-40	40-42	42-44
عدد التلاميذ	4	7	13	10	5	1

والمطلوب إيجاد الوسط الحسابي.

الحل: لحساب الوسط الحسابي باستخدام المعادلة السابقة يتم إتباع الخطوات التالية :

1- إيجاد مجموع التكرارات $\sum f$

2- حساب مراكز الفئات x

3- ضرب مركز الفئة في التكرار المناظر له $\sum f$ وحساب المجموع $\sum xf$

4- حساب الوسط الحسابي بتطبيق المعادلة.

فئات الوزن (C)	التكرارات f	مراكز الفئات x	$x\ f$
32-34	4	2=33÷(32+34)	33=132×4
34-36	7	35	35=245×7
36-38	13	37	37=481×13
38-40	10	39	39=390×10
40-42	5	41	41=205×5
42-44	1	43	43=43×1
المجموع	40		1496

إذا الوسط الحسابي لوزن التلميذ هو :

$$\bar{x} = \frac{\sum_{i=1}^{6} x_i f_i}{\sum_{i=1}^{6} f_i} = \frac{1496}{40} = 37.4\ k.g$$

أي أن متوسط وزن التلميذ يساوي 37.4kg

2- الوسيط:Mediane :

يعد الوسيط احد مقاييس النزعة المركزية المهمة في التطبيقات الإحصائية. و يعرف بأنه تلك القيمة التي يقع ترتيبها وسط المجموعة عند ترتيب هذه القيم ترتيبا تصاعديا أو تنازليا، أي تقسم المجموعة إلى قسمين متساويين بحيث تتساوى عدد الحدود التي أصغر من الوسيط مع عدد الحدود الأكبر منه.

بمعنى أخر هو القيمة التي يقل عنها نصف عدد القيم ، و يزيد عنها النصف الآخر، أي أن 50% من القيم أقل منه و 50% من القيم أعلى منه. و عادة ما يرمز له بالرمز .(Me)

أ- **حساب الوسيط في حالة القيم غير مبوبة (لا توجد فئات)**

* إذا كان عدد البيانات (n) عدد فردي فإن الوسيط هو القيمة التي رتبتها n+1/ 2

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

* إذا كان عدد البيانات(n)عدد زوجي فإن الوسيط هو متوسط القيمة التي رتبتها n/2 و القيمة التي رتبتها n/2 +1

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

مثال3: احسب الوسيط للبيانات التالية:

- (9 ،3، 1 ، 4 ،7 ، 7 ،6 ، 3 ، 4 ،4 ، 2 ، 1)
- (7.5 ، 6 ،1 ، 5.5 ،5 ، 3 ، 2.5 ، 0 ،1)

الحل: المثال الأول : 1، 1، 2، 3، 3، 4، 4 ، 4، 6، 7،7، 9 عدد البيانات زوجي (12) ومنه:

$$Me = \frac{X_{\frac{n}{2}} + X_{\frac{n}{2}+1}}{2} = \frac{X_6 + X_7}{2} = \frac{4+4}{2} = 4$$

المثال الثاني: 0،1، 1، 2.5، 3، 5،5.5، 6، 7.5 عدد البيانات فردي (9) ومنه:

$$Me = X_{\frac{n+1}{2}} = X_5 = 3$$

ب - **حساب الوسيط في حالة الفئات (بيانات مبوبة)**:إذا كان لدينا جدول توزيع تكراري على شكل فئات فإننا نتبع الطرق الآتية لحساب الوسيط:

- تحديد الفئة الوسيطية : و هي أول فئة تكرارها المتجمع الصاعد أكبر او يساوي

أي : $N^{\uparrow}_{me} \geq \frac{n}{2}$

- حساب الوسيط بطريقة المد الداخلي :

$$Me = Li_{me} + \left[\frac{\frac{n}{2} - N^{\uparrow}_{me-1}}{n_{me}}\right] \times A_{me}$$

حيث : Li_{me} تمثل الحد الأدنى للفئة الوسيطية

$N^{\uparrow}_{me-1}$ يمثل التكرار المتجمع الصاعد المطلق للفئة قبل الفئة الوسيطية

A_{me} يمثل طول الفئة الوسيطية

n_{me} تمثل التكرار المطلق للفئة الوسيطية

مثال4: احسب الوسيط و اشرح النتيجة في حالة كان توزيع الأجور اليومية للعمال على النحو التالي:

الأجور اليومية	C_i	n_i	$N_i^{\uparrow}$
]550-400]	475	17	17
700-550	625	13	30
850-700	775	10	40
1000-850	925	5	45
1150-1000	1075	2	47
1300-1150	1225	2	49
1450-1300	1375	1	50

الحل:

- تحديد الفئة الوسيطية : و هي أول فئة تكرارها المتجمع الصاعد أكبر او يساوي $\frac{n}{2}$

أي : $N_{me}^{\uparrow} \geq (\frac{n}{2} = \frac{50}{2} = 25)$

و منه الفئة الوسيطية هي]700-550]

- حساب الوسيط بطريقة المد الداخلي :

$$Me = Li_{me} + \left[\frac{\frac{n}{2} - N_{me-1}^{\uparrow}}{n_{me}}\right] \times A_{me}$$

$$Me = 550 + (\frac{25-17}{13})150 = 642.30$$

الشرح : هناك 50%من العمال أجورهم اليومية أقل من 642.30 دج و 50% من العمال أجورهم اليومية أكبر من 642.30 دج .

ملاحظة: الوسيط بيانيا هو نقطة التقاطع بين المنحنى المتجمع الصاعد و النازل

وهناك مقاييس أخرى محسوبة بنفس أسلوب الوسيط و منها الرباعيات و العشاريات و المائينات تستعمل في جداول خصائص العينة.

3- المنوال:le mode :المنوال هو ثالث مقاييس النزعة المركزية و يعرف بأنه القيمة الأكثر تكراراً أو شيوعاً بين القيم، و قد يكون للمجموعة أكثر من منوال أين تتساوى تكرارات القيم أكثر من مرة.

وتوجد عدة طرق لحساب المنوال من البيانات المبوبة، و كل طريقة تؤدى إلى نتيجة مختلفة، و لذلك فإن قيمة المنوال قيمة تقريبية، و أقل دقة عن المتوسط الحسابي أو الوسيط.

أ- **حساب المنوال في حالة القيم المنفردة:** إذا كانت لدينا قيم منفردة فإن المنوال هو القيمة الاكثر تكرار.

لدينا البيانات التالية التي تمثل عدد الولادات التي تضعها مجموعة من الأرانب في السنة:

13،12،15،10،20، 08، 10، 16،15، 17 ، 15

Mo=15 ، عدد الولادات التي تضعها الأرنب في السنة الأكثر شيوعا هو 15.

ب- **حساب المنوال في حالة الفئات(بيانات مبوبة):** تحديد الفئة المنوالية: و هي الفئة التي تقابل أكبر تكرار

-حساب المنوال بدقة : بحيث

$$Mo = Li_{mo} + \left[\frac{\Delta_1}{\Delta_1 + \Delta_2}\right] \times A_{mo}$$

ص

حيث : Li_{mo} تمثل الحد الأدنى للفئة المنوالية

$$\Delta_1 = n_{mo} - n_{mo-1}$$

$$\Delta_2 = n_{mo} - n_{mo+1}$$

A_{mo} يمثل طول الفئة المنوالية

مثال5: احسب المنوال و اشرح النتيجة في حالة توزيع الأجور اليومية للعمال كان على النحو التالي:

الأجور اليومية	C_i	n_i	$N_i^{\uparrow}$
]550-400]	475	17	17
700-550	625	13	30
850-700	775	10	40
1000-850	925	5	45
1150-1000	1075	2	47
1300-1150	1225	2	49
1450-1300	1375	1	50

الحل:

- تحديد الفئة المنوالية : هي الفئة التي تقابل الأكبر تكرار(17) و هي : 400–550

$$Mo = Li_{mo} + \left[\frac{\Delta_1}{\Delta_1 + \Delta_2}\right] \times A_{mo} = 400 + \left[\frac{(17-0)}{(17-0)+(17-13)}\right] \times 150$$

$$Mo = 521.42$$

- حساب المنوال:

- **الشرح : الأجور اليومية الأكثر شيوعا هي الأجور التي تتراوح بين 400و أقل من 550 دج .**

ثانيا– مقاييس التشتت: إن مقاييس النزعة المركزية لا تكفي لوحدها لوصف البيانات و إجراء المقارنات بين التوزيعات التكرارية ، لأنها لا تعطينا فكرة عن مدى تجانس أو عدم تجانس البيانات ، فعند إجراء مقارنة بين ظاهرتين يمكن أن يتساوى متوسطهما الحسابي ، و رغم ذلك نجد أن انتشار البيانات في الظاهرتين مختلف كثيرا لأن البيانات غير متجانسة. لهذا وجدت مقاييس اخرى تعطينا فكرة عن مدى تباعد البيانات عن بعضها البعض ، تسمى هذه المقاييس بمقاييس التشتت.

ويقصد بالتشتت في مجموعة من القيم التباعد بين مفرداتها. و يكون التشتت كبير اذا كان التفاوت بينهما كبيرا، ويكون التشتت صغيرا اذا كان التفاوت بينهما صغيرا. و من أهم مقاييس التشتت نجد: المدى ، الانحراف المتوسط ، الانحراف المعياري ، الانحراف الربيعي و معامل الاختلاف. الا أن اهمها في العلوم الاجتماعية و الانسانية هو التباين و الانحراف المعياري.

إذا كان لدينا مجموعتين من الموظفين، وكل مجموعة تتكون من خمسة أشخاص حيث الأجور الشهرية (دج) لكل فرد في المجموعتين هي كالآتي:

المجموعة الأولى: 30000 ، 31000 ، 32000 ، 33000 ، 34000.

المجموعة الثانية: 16000 ، 20000 ، 24000 ، 44000 ، 56000.

لو قمنا بحساب المتوسط الحسابي لكل مجموعة، نجد أنه متساوي في كل منهما ويساوي 32000، ومع ذلك أجور المجموعة الأولى أكثر تجانسا من أجور المجموعة الثانية، من أجل ذلك لجأ الإحصائيون إلى استخدام مقاييس أخرى لقياس مدى تجانس البيانات أو مدى انتشار البيانات حول مقياس النزعة المركزية، ويمكن استخدامها في المقارنة بين مجموعتين أو أكثر، ومن هذه المقاييس مقاييس التشتت.

1– مقاييس التشتت المطلقة:

أولا– مقاييس التشتت المطلقة:

المقياس	تعريفه	القانون في حالة سلسلة إحصائية	القانون في حالة توزيع تكراري
المدى E	المدى: هو الفرق بين أكبر قيمة وأقل قيمة في البيانات، فهو من أبسط مقاييس التشتت وأسهلها في الحساب، ومن عيوبه أنه يعتمد على القيمتين المتطرفتين فقط.	$E = Max(X_i) - Min(X_i)$	$E = Max(X_i) - Min(X_i)$
المدى الربيعي IQ	هو الفرق بين الربيع الثالث والربيع الأول، ويعطينا فكرة عن المجال الذي تنتشر فيه نصف عدد البيانات متوسطة القيمة فقط.	$IQ = Q_3 - Q_1$	$IQ = Q_3 - Q_1$
الانحراف المتوسط عن المتوسط الحسابي $EM_{\bar{X}}$	هو البعد المتوسط بالقيمة المطلقة لقيم المتغير الاحصائي عن المتوسط الحسابي وهو الأكثر استعمالا، من عيوبه أنه لا يفرق بين القيم التي تكون أكبر وأقل من المتوسط الحسابي.	$EM_{\bar{X}} = \frac{\sum_{i=1}^{n} \lvert x_i - \bar{X} \rvert}{n}$	$EM_{\bar{X}} = \frac{\sum_{i=1}^{k} n_i \lvert x_i - \bar{X} \rvert}{n}$

الانحراف المتوسط عن الوسيط EM_{M_e}	هو البعد المتوسط بالقيمة المطلقة لقيم المتغير الاحصائي عن الوسيط، من عيوبه أنه لا يفرق بين القيم التي تكون أكبر وأقل من الوسيط	$EM_{M_e} = \frac{\sum_{i=1}^{n} \lvert X_i - M_e \rvert}{n}$	$EM_{M_e} = \frac{\sum_{i=1}^{k} n_i \lvert X_i - M_e \rvert}{n}$
التباين $V(X)$	يعتبر من أهم مقاييس التشتت وهو كثير الاستعمال في مجال الاحصاء $V(X) = (\delta(X))^2$	$V(X) = \frac{\sum_{i=1}^{n} (X_i - \bar{X})^2}{n}$	$V(X) = \frac{\sum_{i=1}^{k} n_i (X_i - \bar{X})^2}{n}$
الانحراف المعياري $\delta(X)$	يعتبر من أهم مقاييس التشتت وهو كثير الاستعمال في مجال الاحصاء $\delta(X) = \sqrt{V(X)}$	$\delta(X) = \sqrt{\frac{\sum_{i=1}^{n} (X_i - \bar{X})^2}{n}}$	$\delta(X) = \sqrt{\frac{\sum_{i=1}^{k} n_i (X_i - \bar{X})^2}{n}}$

ملاحظة: إذا كان المتغير الإحصائي مستمر فإننا نعوض X_i بمراكز الفئات C_i في كل المعادلات السابقة.

2- مقاييس التشتت النسبية:

إذا كنا بصدد إجراء مقارنة بين توزيعات تكرارية ليست لها نفس وحدة القياس أو ليس لهما نفس المتوسط، فمن الضروري هنا استخدام مقاييس التشتت النسبية والتي من أهمها:

المقياس	القانون
المدى النسبي $E\%$	$E\% = \frac{E}{\bar{X}} \times 100$
المدى الربيعي النسبي $IQ\%$	$IQ\% = \frac{IQ}{M_e} \times 100$
الانحراف المتوسط عن المتوسط الحسابي النسبي $EM_{\bar{X}}\%$	$EM_{\bar{X}}\% = \frac{EM_{\bar{X}}}{\bar{X}} \times 100$
الانحراف المتوسط عن الوسيط النسبي $EM_{M_e}\%$	$EM_{M_e}\% = \frac{EM_{M_e}}{M_e} \times 100$
الانحراف المعياري النسبي (معامل الاختلاف) CV	$CV = \frac{\delta(X)}{\bar{X}} \times 100$

مثال: حساب مختلف مقاييس التشتت

البيانات التالية تمثل أوزان 60 طالبا بالكيلوغرام في أحد أقسام الـ: LMD بكلية العلوم الاقتصادية وعلوم التسيير بجامعة سطيف:

أوزان الطلبة X_i	عدد الطلبة n_i
$]55-50]$	2
$]60-55]$	5
$]65-60]$	12
$]70-65]$	16
$]75-70]$	14
$]80-75]$	8
$]85-80]$	3
المجموع $\sum n_i$	60

المطلوب:

1- حساب المدى المطلق والنسبي؟

2- حساب المدى الربيعي المطلق والنسبي؟

3- الانحراف المتوسط عن المتوسط الحسابي المطلق والنسبي؟

4- الانحراف المتوسط عن الوسيط المطلق والنسبي؟

5- التباين، الانحراف المعياري ؟

6- في دراسة مماثلة عن أوزان الطلبة بجامعة أخرى تحصلنا على النتائج التالية:

$\bar{X} = 75$ ، $\delta(X) = 12$

- قارن بين تشتت الأوزان في الدراستين؟

الحل:

أوزان الطلبة X_i	n_i	C_i	$N_i^{\uparrow}$	$n_i \times C_i$	$n_i\lvert C_i - \bar{X}\rvert$	$n_i\lvert C_i - M_e\rvert$	$n_i(C_i - \bar{X})^2$
$]55-50]$	2	52,5	2	105	31.84	31.88	506,8928
$]60-55]$	5	57,5	7	287.5	54.6	54.7	596,232
$]65-60]$	12	62,5	19	750	71,04	71.28	420,5568
$]70-65]$	16	67,5	35	1080	14,72	15.04	13,5424
$]75-70]$	14	72,5	49	1015	57,12	56.84	233,0496
$]80-75]$	8	77,5	57	620	72,64	72.48	659,5712
$]85-80]$	3	82,5	60	247,5	42,24	42.18	594,7392
المجموع $\sum n_i$	60	/	/	4105	344,2	344.4	3024,584

1- حساب المدى:

أ- المدى المطلق:

$$\bar{X} = \frac{\sum_{i=1}^{7} n_i c_i}{n} = \frac{4105}{60} = 68{,}42$$

$$E = Max(X_i) - Min(X_i) = 85 - 50 = 35$$

طول المجال الذي تنتشر فيه أوزان الطلبة هو: 35 كلغ.

ب- المدى النسبي:

$$E\% = \frac{E}{\bar{x}} \times 100 = \frac{35}{68{,}42} \times 100 = 51{,}15\%$$

2- حساب المدى الربيعي:

أ- المدى الربيعي المطلق:

$$IQ = Q_3 - Q_1$$

نقوم بحساب Q_1 كمايلي:

تحديد الفئة الربيعية الأولى: وهي أول فئة تكرارها المتجمع الصاعد أكبر أو يساوي $\frac{n}{4}$ ، أي:

$$N_{Q_1}^{\uparrow} \geq \left(\frac{n}{4} = \frac{60}{4} = 15\right)$$

ومنه الفئة الربيعية الأولى هي: $]65-60]$

ج- حساب الربيع الأول بطريقة المد الداخلي:

$$Q_1 = Lim_{Q_1} + \left[\frac{\frac{n}{4} - N^{\downarrow}_{Q_1-1}}{n_{Q_1}}\right] \times A_{Q_1} = 60 + \left[\frac{15-7}{12}\right] \times 5 = 63{,}33$$

نقوم بحساب Q_3 كمايلي:

تحديد الفئة الربيعية الثالثة: وهي أول فئة تكرارها المتجمع الصاعد أكبر أو يساوي $\frac{3}{4}$ ، أي:

$$\geq \left(\frac{3n}{4} = \frac{3(60)}{4} = 45\right)$$

ومنه الفئة الربيعية الثالثة هي: $]75 - 70]$

– حساب الربيع الثالث بطريقة المد الداخلي:

$$Q_3 = Lim_{Q_3} + \left[\frac{\frac{3n}{4} - N^{\uparrow}_{Q_3-1}}{n_{Q_3}}\right] \times A_{Q_3} = 70 + \left[\frac{45-35}{14}\right] \times 5 = 73{,}57$$

ومنه: $$\boldsymbol{IQ = 73,57 - 63,33 = 10,24}$$

الشرح: طول المجال الذي تنتشر فيه الأوزان المتوسطة لنصف الطلبة هو : 10,24 كلغ.

ب – المدى الربيعي النسبي :

نقوم بحساب الوسيط في حالة البيانات المبوبة فنجده يساوي: 68,44 كلغ، أي: $M_e = 68.44$

$$\boldsymbol{IQ\% = \frac{IQ}{M_e} \times 100 = \frac{10.24}{68,44} \times 100 = 14,96\%}$$

3 – الانحراف المتوسط عن المتوسط الحسابي:

$$\bar{X} = \frac{\sum_{i=1}^{i} n_i c_i}{n} = \frac{4105}{60} = 68{,}42$$

أ – المطلق: $$EM_{\bar{x}} = \frac{\sum_{i=1}^{k} n_i |x_i - \bar{x}|}{n} = \frac{344{,}2}{60} = 5{,}74$$

يقدر الانحراف المتوسط عن المتوسط الحسابي لأوزان الطلبة بـ: 5,74 كلغ.

ب – النسبي: $$\boldsymbol{EM_{\bar{x}}\% = \frac{EM_{\bar{x}}}{\bar{x}} \times 100 = \frac{5,74}{68,42} \times 100 = 8,39\%}$$

4 – الانحراف المتوسط عن الوسيط:

$$M_e = 68{,}44$$

أ – المطلق: $$EM_{M_e} = \frac{\sum_{i=1}^{k} n_i |x_i - M_e|}{n} = \frac{344{,}4}{60} = 5{,}74$$

يقدر الانحراف المتوسط عن الوسيط لأوزان الطلبة بـ: 5,74 كلغ.

ب – النسبي: $$EM_{M_e}\% = \frac{EM_{M_e}}{M_e} \times 100 = \frac{5,74}{68,42} \times 100 = 8{,}39\%$$

5 –حساب التباين والانحراف المعياري:

أ – التباين: $$V(X) = \frac{\sum_{i=1}^{k} n_i (x_i - \bar{x})^2}{n} = \frac{3024{,}584}{60} = 50{,}41$$

ب – الانحراف المعياري: $$\delta(X) = \sqrt{\frac{\sum_{i=1}^{k} n_i (c_i - \bar{x})^2}{n}} = \sqrt{V(X)} = \sqrt{50.41} = 7.1$$

يقدر متوسط تشتت أوزان الطلبة بـ: 7,1 كلغ.

6 – المقارنة بين تشتت الأوزان في الدراستين:

أ – الدراسة الأولى:

الانحراف المعياري النسبي (معامل الاختلاف): $$CV_1 = \frac{\delta(X)}{\bar{x}} \times 100 = \frac{7,1}{68,42} \times 100 = 10{,}38\%$$

ب – الدراسة الثانية:

الانحراف المعياري النسبي (معامل الاختلاف): $$CV_2 = \frac{\delta(X)}{\bar{x}} \times 100 = \frac{12}{75} \times 100 = 16\%$$

نلاحظ أن تشتت الأوزان في الدراسة الثانية أكبر منه في الدراسة الأولى.

القراءة الإحصائية : تشتت الوزن في الدراسة الثانية أكبر منه في الدراسة الأولى ، أي أن الفوارق الموجودة في أوزان الطلبة في الجامعة الثانية أكبر ، وعلية فأوزان الطلبة في جامعة سطيف أكثر تجانس من أوزان الطلبة في الجامعة الثانية.

المحاضرة 08: المقاييس والمعاملات الاحصائية.

ثالثا – المعاملات الاحصائية:

1– معامل الاختلاف (coefficient de variation): يستخدم معامل الاختلاف لمقارنة التشتت بين مجموعتين وذلك للاختلاف الواضح في الوسط الحسابي لمجموعتين من حيث القيمة. و هو النسبة المئوية بين الانحراف المعياري والوسط الحسابي، و بالتالي يمكن استخدامه لمجموعتين مختلفتين في الوحدات، و يحسب من الصيغة الرياضية الآتية:

معامل الاختلاف = (الانحراف المعياري / المتوسط) x 100. وبالرموز تكون الصيغة: CV = (SD / xbar) x 100

2– معامل الالتواء (dissymétrie): في حالة التوزيع الغير اعتدالي / الغير طبيعي لا يجب الاعتماد على المتوسطات الحسابية و الانحراف المعياري لقياس درجة التشتت فقط، بل لا بد من استخدام مقياس يبين مدى بعد التوزيع عن الاعتدالية .فالالتواء هو بعد المنحنى عن التماثل. و يكون التوزيع التكراري متماثلا اذا كانت التكرارات موزعة توزيعاً متماثلا حول الوسط الحسابي, بمعنى أن يكون لقيم المتغير المتساوية البعد عن الوسط الحسابي نفس التكرارات.

و قد يكون الالتواء سالبا أو موجبا، و قد يتساوى توزيعان من حيث المتوسط الحسابي و الانحراف المعياري الا أنهما يختلفان من حيث الالتواء فإما أن يكونان:

* في اتجاه واحد و لكن بمقادير مختلفة.
* تتساوى درجة الالتواء و لكن تختلفان في الاشارة (موجب ، سالب)

و يحسب مقياس الالتواء وفق المعادلة التالية:

هناك طرق كثيرة لقياس الالتواء ومنها ما يلي:

1– طريقة" بيرسون Person" في قياس الالتواء: تأخذ هذه الطريقة في الاعتبار العلاقة بين الوسط والوسيط والمنوال، في حالة ما إذا كان التوزيع قريب من التماثل وليس شديد الالتواء ، وهذه العلاقة هي(5–1)

$$\text{المنوال} = \text{الوسط الحسابي} - 3(\text{الوسط الحسابي} - \text{الوسيط})$$

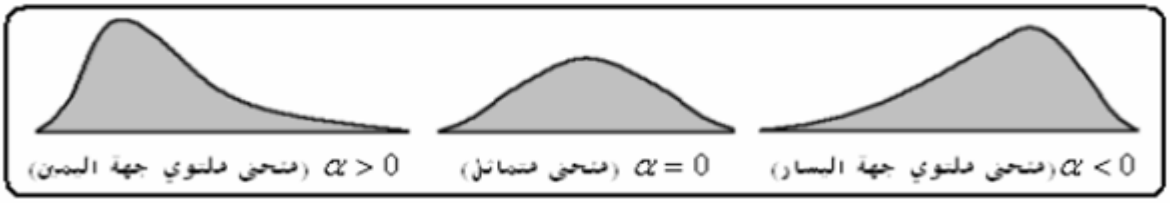
ومن ثم فإن طريقة" بيرسون "في قياس الالتواء ، تتحدد بالمعادلة التالية.(٥-٢)

$$\alpha = \frac{3(Mean - Median)}{Standard\ Deviation} = \frac{3(\bar{x} - Med)}{S}$$

حيث أن a **ألفا هو معامل الالتواء "لبيرسون** $\bar{x}$ **الوسط الحسابي** Med **هو الوسيط** S هو الانحراف المعياري، ويمكن من خلال الإشارة التي يأخذها هذا المعامل الحكم على شكل الالتواء، كما يلي :

- إذا كان الوسط الحسابي = (الوسيط)كان قيمة المعامل($a = 0$) ويدل ذلك على أن منحنى التوزيع التكراري متماثل.
- إذا كان(الوسط الحسابي > الوسيط)كان قيمة المعامل($a > 0$) ويدل ذلك على أن منحنى التوزيع التكراري ملتوي جهة اليمين.
- إذا كان(الوسط الحسابي< الوسيط)كان قيمة المعامل($a < 0$) ويدل ذلك على أن منحنى التوزيع التكراري ملتوي جهة اليسار.

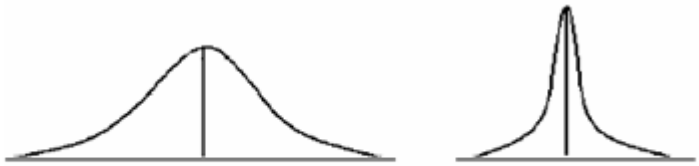
شكل(5-1). أشكال التواء البيانات



٢ـ طريقة" المئين "في قياس الالتواء: المئين ينتج من ترتيب البيانات تصاعديا، ثم تقسيمها البيانات إلى 100 جزء، يفصل بينها قيم تسمى المئين، وعلى سبيل المثال يعرف المئين 15 ويرمز له بالرمز (v_p) على أنه القيمة التي يقل عنها15% من القيم، ولحساب قيمة المئين p ونرمز له بالرمز ، (v_p) يتبع نفس الفكرة المستخدمة في حساب الربيع كما يلي

- ترتب القيم تصاعديا $x_{(1)} < x_{(2)} < \ldots < x_{(n)}$
- رتبة المئين $R = (n+1)\left(\frac{p}{100}\right)$.

3- التفرطح (kurtosis) : تقسم المنحنيات التكرارية من حيث تفرطح قمتها الى ثلاثة أنواع (معتدلة / متوسطة التفرطح، مدببة ، مفرطحة).ان وصف المنحنيات بأنها مدببة أو مفرطحة يكون بالمقارنة مع المنحنيات المعتدلة. فعندما نقول أن المنحنى مدبب، فنعنى أن عدداً كبيراً من المفردات يتراكم بالقرب من الوسط الحسابي و عند الذيلين و لا يكون في الأماكن الأخرى الا عدداً قليلاً منها, و ذلك بالمقارنة مع المنحنيات المعتدلة. كذلك حين نقول أن المنحنى مفرطح فنعنى أن عدداً قليلاً من المفردات يتراكم بالقرب من الوسط الحسابي وعند الذيلين و يكون هناك عدد كبير منها بالمواقع الأخرى, و هذا بالمقارنة مع المنحنيات المعتدلة، ويظهر ذلك من الشكل التالي :



ويمكن قياس التفرطح باستخدام عدد من الطرق، ومنها طريقة العزوم ، حيث يحسب معامل التفرطح (K) بتطبيق المعادلة التالية :

$$k = \frac{\frac{1}{n}\sum (x-\bar{x})^4}{s^4}$$

حيث أن المقدار $\sum (x-\bar{x})^4 / n$ هو العزم الرابع حول الوسط s هو الانحراف المعياري . ومعامل التفرطح في التوزيع الطبيعي يساوي 3 ، ومن ثم يمكن وصف منحنى التوزيع من حيث التفرطح ، والتدبب كما يلي: :

- إذا كان k =3 كان منحنى التوزيع معتدلا .
- إذا كان k > 3 كان منحنى التوزيع مدببا .
- إذا كان k< 3 كان منحنى التوزيع منبسطا (مفرطحا) .

4- معامل الارتباط: يشير معامل الارتباط إلى المدى الذي تقع فيه النقاط الموجودة في مخطط التشتت على خط مستقيم، وقيمته تنحصر ما بين [-1،+1] ، حيث تُستخدم معاملات الارتباط في الإحصائيات لقياس مدى قوة العلاقة بين متغيرين، وتعتبر دراسة الارتباط الإحصائي بين المتغيرات بالغ الأهمية في البحوث الاجتماعية والانسانية لأنه يعطينا معيارا نستطيع من خلاله تقدير قيمة الفرضيات التي وضعناها أثناء الدراسة بحيث يمكننا من خلال دراسة الارتباط أن نثبت هذه الفرضيات أو ننفيها نفيا باتا.

وعند دراستنا للعلاقات بين المتغيرات المختلفة نجد أن بعضها يرتبط بأكثر من متغير واحد، مثل جداول الارتباط المركبة السالفة الذكر، في هذه الحالة أما أن نقوم بدراسة علاقة المتغير بجميع المتغيرات المرتبطة به دفعة واحدة وهذا ما يسمى بالارتباط الكلي، أو ندرس علاقة المتغير المستقل بالتابع فقط وهذا ما يسمى بالارتباط الجزئي.

نوع العلاقة: و تأخذ ثلاث أنواع حسب إشارة معامل الارتباط كما يلي:

* إذا كانت إشارة معامل الارتباط سالبة ($r < 0$) توجد علاقة عكسية بين المتغيرين، بمعنى أن زيادة أحد المتغيرين يصاحبه انخفاض في المتغير الثاني، و العكس صحيح.
* إذا كانت إشارة معامل الارتباط موجبة ($r > 0$)توجد علاقة طردية بين المتغيرين، بمعنى أن زيادة أحد المتغيرين يصاحبه زيادة في المتغير الثاني، و العكس صحيح.
* إذا كان معامل الارتباط مساويا لصفرا ($r = 0$) دل ذلك على انعدام العلاقة بين المتغيرين.

و مقايس الارتباط هي الاخرى تستعمل حسب المتغير والجدول و من أهمها نجد:

أ- **معامل الارتباط الخطى البسيط " لبيرسون (Pearson)**: وهو يستعمل في حالة جمع بيانات عن متغيرين كميين ، ويحسب وفق المعادلة التالية:

$$r = \frac{n\sum xy - (\sum x)(\sum y)}{\sqrt{(n\sum x^2 - (\sum x)^2\,(n\sum y^2 - (\sum y)^2}}$$

$$r = \frac{\sum Xy - n\,\bar{x}\,\bar{y})}{\sqrt{(\sum x^2 - n\,\bar{x}^2)(\sum y^2 - n\,\bar{y}^2)}}$$

ب- معامل ارتباط الرتب اسبيرمان (Spearman): أما إذا كانت الظاهرة محل الدراسة تحتوي على متغيرين وصفيين ترتيبين أو متغيرين كميين، فإنه يمكن استخدام معامل ارتباط اسبيرمان Spearman " ، و يعبر عنه بالمعادلة التالية:

4- معامل الاقتران(coefficient d'association): نستعمل هذا المعامل في حالة وجود توزيع تكراري مزدوج بسيط التقاطع يحتوي على 4 خانات للتقاطع فقط. و هو يستعمل عادة في حساب درجة العلاقة بين المتغيرات الكيفية أي كلا من المتغير المستقل و التابع كيفيين معا أو في حالة وجود أحد المتغيرين كيفي و الثاني كمي، ويحسب وفق القانون التالي:

$$r_{rank} = 1 - \frac{6\sum D^2}{n(n^2 - 1)}$$

5- معامل التوافـق (coefficient de concordance): إذا كان كلا من المتغير المستقل والتابع كيفيين معا أو أحدهما كيفي و الثاني كمي. و كان الجدول التكراري يحتوي على 6 خانات للتقاطع فاكثر ففي هذه الحالة نستعمل معامل التوافق. الذي يحسب وفق المعادلة التالية.

$$\text{معامل التوافق} = \frac{\text{جـ} - \text{ا}}{\text{جـ}}$$

حيث تحسب (جـ) من العلاقة :

$$\text{جـ} = \text{مج} \frac{\text{مربع الخلية}}{\text{مجموع صف الخلية} \times \text{مجموع عمود الخلية}}$$

4 - اختبار الفروض الاحصائية كا^2 : تستعمل هذه الأداة بصورة رئيسية لاختبار الفرضيات التي تقوم على أساس مقارنة مجموعة من التكرارات النظرية مع مجموعة من التكرارات الفعلية لتقيم الفرق بينهما لمعرفة ما إذا كان هذا الفرق فرقا ظاهريا نتيجة قوى الحظ والصدفة أم أنه فرق حقيقي نتيجة قوى أخرى غير قوى الحظ والصدفة.

ومن أجل اختبار استقلالية الظواهر نقوم بتصنيف البيانات في جدول مزدوج ثم نقوم بالخطوات التالية:

- صياغة الفروض العلمية الأول هو الفرض الصفري و الذي نرمز له بالرمز (H_0) الذي ينفي وجود العلاقة بين متغيرين أي لا يوجد فروق ذات دلالة إحصائية، و الفرض الثاني هو الفرض البديل(H_1) والذي يقرر أن هناك علاقة بين المتغيرين .

- نحدد مستوى الدلالة : ونسبته في الغالب 1% أو 5% ، أي 0.01 إلى 0.05 في العلوم الاجتماعية.

- نحدد درجة الحرية = (dL) (عدد الصفوف -1) X (عدد الاعمدة -1)

- نحسب التكرارات المتوقعة (النظرية): وذلك بأن نعد جدول توافق نظري حيث نفترض أن المجاميع فيه متساوية ويحسب وفق القانون التالي:

$$\text{ت}_\text{م} = \frac{\text{مجموع الصف} \times \text{مجموع العمود}}{\text{المجموع الكلى}}$$

- نقوم بحساب كا^2 وفق القانون التالي:

الطريقة العامة لحساب كا^2

$$\text{كا}^2 = \text{مج} \frac{2(\text{ت}_\text{و} - \text{ت}_\text{م})}{\text{ت}_\text{و}}$$

حيث :

تو : هو التكرار الواقعى الذي يحدث بالفعل والموجود بالجدول .

تم : هو التكرار المتوقع حدوثه ويختلف حسابه باختلاف نوع الجدول المطلوب حساب كا^2 منه .

تحديد مدى دلالة كا^2 من عدمه

فى جميع الحالات نخرج من الحسابات بقيمة كا^2 المحسوبة نقارنها بقيمة كا^2 الجدولية كالتالى :

- إذا كانت كا^2 المحسوبة > كا^2 الجدولية فان كا^2 تكون دالة إحصائية .
- إذا كانت كا^2 المحسوبة < كا^2 الجدولية فان كا^2 ليست دالة إحصائية .

حالات حساب كا^2 من الجداول المختلفة :

الحالة الأولى : الطريقة العامة لحساب كا^2 من الجدول التكراري 1×2 :

يتكون الجدول 1×2 من صف واحد وعمودين دون خلايا المجموع إن وجدت بالجدول .

ولحساب قيمة كا^2 فى هذا الجدول تحسب من القانون العام :

$$\text{كا}^2 = \text{مج} \frac{(\text{ت}_\text{و} - \text{ت}_\text{م})^2}{\text{ت}_\text{و}}$$

حيث تم هنا تساوى متوسط التكرارات الواقعية الموجودة بالجدول.

مثال1 :

الجدول التالي يوضح آراء 80 شخص فى استبيان دار حول رفض أو قبول قضية الزواج العرفي .

الرأي	موافق	غير موافق	مج
التكرار	60	20	80

والمطلوب حساب قيمة كا^2 مع بيان مدى دلالتها إحصائيا عند مستوى دلالة 0.05 ؟

الحل :

حساب التكرار المتوقع (تم) :

$$\text{ت}_\text{م} = \frac{60 + 20}{2} = 40$$

حساب كا^2 المحسوبة :

نكون الجدول التالي :

ت و	ت م	ت و – ت م	(ت و – ت م)^2	(ت و – ت م)^2 / ت م
6	40	20	400	10
20	40	20-	400	10
–	–	–	مجموع	20

من الجدول مباشرة فان مجموع العمود الأخير يعطينا قيمة كا^2 :

كا^2 المحسوبة = 20 .

حساب كا^2 الجدولية :

لحسابها يلزم حساب درجة الحرية ومستوى الدلالة :

درجة الحرية = عدد الأعمدة – 1 = 2 – 1 = 1

مستوى الدلالة = 0.05 .

بالبحث فى جداول كا^2 عند درجة حرية = 1 ومستوى دلالة 0.05 نجد قيمة كا2 الجدولية = 3.841 .

تحديد مدى دلالة كا^2:

نقارن قيمة كا^2 المحسوبة بقيمة كا^2 الجدولية نجد أن :

قيمة كا^2 المحسوبة = 20 > قيمة كا^2 الجدولية = 3.841

لذا فان كا^2 دالة إحصائية عند مستوى دلالة 0.05 .

الحالة الثانية : الطريقة العامة لحساب كا^2 من الجدول التكراري 2×2 :

يتكون الجدول 2×2 من صفين وعمودين دون خلايا المجموع إن وجدت بالجدول .

ولحساب قيمة كا^2 فى هذا الجدول تحسب من القانون العام :

مثال2 : الجدول التالي يوضح العلاقة بين المتغيرين النوع وتأييد برنامج تليفزيوني معين .

النوع / الفكرة	ذكور	إناث	المجموع
مؤيد	35	37	72
معارض	14	34	48
المجموع	49	71	120

والمطلوب حساب قيمة كا^2 مع بيان مدى دلالتها إحصائيا عند مستوى دلالة 0.05 ؟

الحل :

حساب التكرار المتوقع (ت$_م$) :

$$\text{ت}_م \text{ للخلية الأولى (35)} = \frac{72 \times 49}{120} = 29.4$$

$$\text{ت}_م \text{ للخلية الثانية (37)} = \frac{72 \times 71}{120} = 42.6$$

$$\text{ت}_م \text{ للخلية الثالثة (14)} = \frac{49 \times 48}{120} = 19.6$$

$$\text{ت}_م \text{ للخلية الرابعة (34)} = \frac{71 \times 48}{120} = 28.4$$

حساب كا^2 المحسوبة :

نكون الجدول التالي :

ت$_و$	ت$_م$	ت$_و$ – ت$_م$	(ت$_و$ – ت$_م$)^2	(ت$_و$ – ت$_م$)^2 / ت$_م$
35	**29.4**	**5.6**	**31.36**	**1.06**
37	**42.6**	**5.6–**	**31.6**	**0.74**
14	**19.6**	**5.6–**	**31.36**	**1.6**
34	**28.4**	**5.6**	**31.36**	**1.1**
–	–	–	**مجموع**	**4.5**

من الجدول مباشرة فان مجموع العمود الأخير يعطينا قيمة كا^2

كا^2 المحسوبة = 4.5 .

حساب كا^2 الجدولية :

لحسابها يلزم حساب درجة الحرية ومستوى الدلالة :

درجة الحرية = (عدد الصفوف – 1) × (عدد الأعمدة – 1) = (2 – 1) × (2 – 1) = 1 × 1 = 1

مستوى الدلالة = 0.05 .

بالبحث فى جداول كا^2 عند درجة حرية = 1 ومستوى دلالة 0.05 نجد قيمة كا^2 الجدولية = 3.841

تحديد مدى دلالة كا^2 :

نقارن قيمة كا^2 المحسوبة بقيمة كا2 الجدولية نجد أن :

قيمة كا2 المحسوبة = 4.5 > قيمة كا^2 الجدولية = 3.841

لذا فان كا^2 دالة إحصائية عند مستوى دلالة 0.05 .

مثال2 :الجدول التالي يوضح العلاقة بين المتغيرين النوع وتأييد برنامج تليفزيوني معين .

الفكرة / النوع	موافق جدا	موافق نوعاً ما	لا أدرى	أرفض نوعاً ما	أرفض جداً	المجموع
ذكور	5	37	13	28	5	88
إناث	3	17	8	20	5	53
المجموع	8	54	21	48	10	141

والمطلوب حساب قيمة كا2 مع بيان مدى دلالتها إحصائيا عند مستوى دلالة 0.05 ؟

الحل :

– حساب التكرارات المتوقعة (ت$_م$) لكل خلية بنفس الطريقة السابقة فنحصل على الجدول التالي:

– حساب كا^2 المحسوبة :

نكون الجدول التالي :

ت$_و$	ت$_م$	ت$_و$ – ت$_م$	(ت$_و$ – ت$_م$)^2	(ت$_و$ – ت$_م$)^2 / ت$_م$
5	37	5	0	0
37	33.7	3.3	10.9	0.32
13	13.1	0.1–	0.01	0
28	29.95	1.59–	3.8	0.13

0.24	1.5	1.24-	6.24	5
0	0	0	3	3
0.53	10.8	3.29-	20.29	17
0	0.01	0.11	7.89	8
0.22	4	2	18	20
0.42	1.56	1.25	3.75	5
1.86	مجموع	-	-	-

من الجدول مباشرة فان مجموع العمود الأخير يعطينا قيمة كا2

كا2 المحسوبة = 1.86 .

حساب كا2 الجدولية :

لحسابها يلزم حساب درجة الحرية ومستوى الدلالة :

درجة الحرية = (عدد الصفوف – 1) × (عدد الأعمدة – 1)

= (2 – 1) × (5 – 1) = 1 × 4 = 4

مستوى الدلالة = 0.05 .

بالبحث فى جداول كا^2 عند درجة حرية = 4 ومستوى دلالة 0.05 نجد قيمة كا2 الجدولية = 9.488

تحديد مدى دلالة كا^2 :

نقارن قيمة كا^2 المحسوبة بقيمة كا^2 الجدولية نجد أن :

قيمة كا^2 المحسوبة = 1.86 < قيمة كا2 الجدولية = 9.488

لذا فان كا^2 ليست دالة إحصائية عند مستوى دلالة 0.05 .

– نقوم مقارنة قيمة كا^2 الجدولية بقيمة كا^2 المحسوبة

* فإذا كانت كا^2 الجدولية > من كا^2 المحسوبة نستنتج أن هناك فروق بين التكرارات النظرية والتكرارات الفعلية وهي فروق ظاهرية راجعة للصدفة وبالتالي نقبل فرضية العدم والاستقلال أي لا توجد علاقة بين المتغيرين.

* أما إذا كانت كا^2 الجدولية < من كا^2 المحسوبة نستنتج أن هناك فروق جوهرية ونرفض فرضية العدم أو الاستقلال ونقول أن هناك علاقة بين المتغيرين.

دليل اختيار الأسلوب الإحصائي الذي يناسب بيانات بحثك:

تمثل الدراسات المتخصصة في الإحصاء، الطريق الأوضح لجميع التطبيقات التي يمكن للباحث انتهاجها قصد التوفيق بين مرحلتي الإحصاء الوصفي و الاستدلالي. لكن هناك مقترحات مختلفة يمكنها إفادته في اختيار الأسلوب الإحصائي الذي يناسب بيانات البحث في حالة اختيار العينة، نقترح واحدا منها في الجدول التالي:

تقوم فكرة هذا الدليل على الإجابة عن الأسئلة الأربعة التالية :

س1: ما عدد العينات المستخدمة فى البحث ؟ س2: هل العينات مستقلة أم مترابطة ؟

س3: ما نوع البيانات الخاصة بمتغيرات البحث؟ س4: ما نوع التصميم التجريبى الذى يستخدمه الباحث؟

يمكن الإجابة على التساؤلات الأربعة، وكذلك الأسلوب الإحصائى المناسب فى الجدول التالى:

عدد العينات	الفرض	التصميم التجريبى	نوع البيانات	الاختبار الإحصائى
عينة واحدة	التحقق من جودة المطابقة	مجموعة واحدة ذات الاختبار الواحد	اسمية	ذى الحدين – كا2 –سمير نوف
			رتبية	سمير نوف – الإشارة
			فترية	اختبارZ – اختبار ت
عينتان مستقلتان	الفروق بين المجموعات	مجموعتان تجريبية – ضابطة	اسمية	كا2 – فشر – سمير نوف
			رتبية	الوسيط – مان ويتنى – التتابع
			فترية	اختبار ت
عينتان مترابطتان	الفروق بين القياسات	مجموعة واحدة ذات اختبارين قبلى وبعدى	اسمية	ماكنمار
			رتبية	ولكوكسن – الإشارة
			فترية	اختبار ت
عدة عينات مستقلة	الفروق بين المجموعات	المجموعات المتعددة	اسمية	كا2
			رتبية	الوسيط– كروسكال ولاس
			فترية	تحليل التباين – تحليل التغاير
عدة عينات مترابطة	الفروق بين القياسات	مجموعة واحدة ذات الاختبارات المتعددة	اسمية	كوجران
			رتبية	فريدمان
			فترية	تحليل التباين ذى القياسات المتكررة
عينة واحدة أو عينتان أو عدة عينات	الارتباط بين القياسات أو العلاقة بين المتغيرات "دراسات ارتباطية"	مجموعة واحدة ذات اختبار قبلى أو بعدى أو عدة اختبارات	اسمية	معامل ارتباط فاى– معامل التوافق – معامل الاقتران الرباعى
			رتبية	معامل ارتباط سبيرمان– معامل ارتباط كندال
			فترية	معامل ارتباط بيرسون – الارتباط القانونى – الارتباط

				المتعدد
عينة واحدة أو عينتان أو عدة عينات	"دراسات تنبؤية " للمتغيرات أو عضوية الجماعة	مجموعة واحدة أو عدة مجموعات مع عدة اختبارات	فترية	تحليل الانحدار بأنواعه المختلفة– السلاسل الزمنية التحليل التمييزى بأنواعه المختلفة
عينة واحدة أو عينتان أو عدة عينات	" دراسات عاملية" البناء العاملى	مجموعة واحدة أو عدة مجموعة مع عدة اختبارات	فترية	التحليل العاملى الاستكشافى – التحليل العاملى التوكيدى