الفصل الأول

مقدمة عن علم الإحصاء

تعريف علم الإحصاء

ان علم الإحصاء هو فرع من فروع الرياضيات وهو يشمل النظريات والطرائق التي تهدف جمع البيانات ووصفها و معالجتها من اجل اتخاذ القرارات .

والإحصاء يُمكننا من جمع الحقائق عن الظواهر المختلفة في صور قياسية رقمية وعرضها بيانيا ووضعها في جداول تلخيصية بطريقة تسهل تحليلها بهدف معرفة اتجاهات هذه الظواهر وعلاقات بعضها ببعض .

ولقد كان الهدف الرئيس من علم الإحصاء قديما هو عد أو حصر الأشياء المراد توفير بيانات إحصائية عنها ، وكانت الجهة التي تقوم بإعداد الإحصاءات على مستوى الدولة تعرف بمصلحة التعداد ولذلك كان التعريف القديم لعلم الإحصاء أنه علم العد ، أي العلم الذي يشتمل على أساليب جمع البيانات الكمية عن المتغيرات والظواهر موضوع الدراسة . ولكن مع تطور المجتمعات ، لم يعد مجرد توفير البيانات الكمية عن المتغيرات والظواهر موضوع الدراسة يفي بحاجات متخذي القرارات إلى تكوين صورة متكاملة الجوانب عن مجتمعهم والمجتمعات المحيطة به , فقام العلماء بتحديث نظريات علم الإحصاء وأساليبه وأدواته لكي يُعين الباحثين وغيرهم على استخلاص استنتاجات معينة من البيانات الكمية التي أمكنهم جمعها عن طريق العد. ومن ذلك على سبيل المثال نظرية العينات التي ساعدت الباحثين على استخلاص استنتاجات عديدة من دراسة عدد صغير من الأفراد أو الأشياء (العينة) وتعميم تلك الاستنتاجات على المجتمع الذي سحبت منه العينة بأكمله .

ويعرف علم الإحصاء حديثاً بأنه (علم متكامل يتضمن الأسلوب العلمي لتقصي حقائق الظواهر واستخلاص النتائج عنها، كما يتضمن أيضاً النظرية اللازمة للقياس واتخاذ القرار في ميادين الحياة كافة) .

أهمية علم الإحصاء

لقد أصبح لعلم الإحصاء أهمية كبيرة في حياتنا المعاصرة فصارت الإحصاءات مألوفة لدينا وتمثل جانبا مهما من المعلومات التي نطالعها كل يوم مثل جداول النقاط التي تحرزها أندية كرة القدم والتقديرات الخاصة بالتنبؤات الجوية وانجازات الحكومة في مجال الإسكان والتعمير والتغيرات التي تطرأ على أسعار السلع والعملات .

وربما يتساءل الفرد عن أهمية الإحصاء بالنسبة للباحث في العلوم التربوية والنفسية , معتقداً أن الإحصاء موضوع يدخل في صميم تخصص الاقتصاديين والرياضيين فقط ، والواقع أن الباحث والمختص في العلوم التربوية والنفسية بوجه عام يحتاج في كثير من الأحيان إلى استخدام الأرقام لكي يلخص ويعرض بها مجموعة من البيانات التي تتعلق بظاهرة يهتم بدراستها ، فقد يطلب منه أن يقدم بحثا عن مدى التطور الذي حققه برنامج معين للتخفيف من القلق لدى متعلمي المؤسسة التي يعمل بها ، وقد يكلف بدراسة الأسباب التي تجعل الطلبة لا يرغبون في مادة دراسية معينة .

كما ان للإحصاء أهمية كبيرة في الأبحاث والدراسات العلمية والطبيعية ، إذ لا تخلو أي دراسة أو بحث من معالجة إحصائية تتعرض لأصل الظاهرة أو الظواهر المدروسة فتصور واقعها بصورة ارقام او بيانات كمية ، وتنتهي إلى اتخاذ القرارات .

أن النتائج التي تتمخض عن تطبيق الوسائل الإحصائية ليست نتائج قطعية أو غير قابلة للتمحيص والمراجعة والتعديل وبعبارة أخرى يقتصر دور الوسائل الإحصائية على توفير مؤشرات مبدئية تساعد الباحث على رفض أو قبول الفرضيات التي يقوم بدراستها في حدود درجة معينة من الثقة .

ومما يعكس أهمية الإحصاء أنه يستخدم في توجيه عملية جمع البيانات وتفسير العلاقات التي تعكسها تلك البيانات . ومن ابرز المجالات التي تستخدم فيها المعالجات الإحصائية إجراء مقارنة بين عدد من الظواهر او المتغيرات . ويمكن القول أن الحياة الإنسانية سلسلة من المواقف التي يتخذ فيها الفرد قراره بناءً على ما تسفر عنه المقارنة التي يجريها بين عديد من الاحتمالات وهذه المقارنة في جوهرها عملية إحصائية تقترن بالقياس والتقييم والتقدير . فنجاح الإنسان في حياته يتحدد وفق مقياس معين في ذهنه يقدر به هذا النجاح.

تطور علم الإحصاء

لقد مرَ علم الاحصاء في مراحل تطور عدة ، وتم ذلك بفضل جهود كثير من المتخصصين , وكان التطور في السابق بطيئا إلى أن جاء القرن العشرين ليشهد تطورا هائلا في المفاهيم الإحصائية واساليب تطبيقاتها .

يرجع الاهتمام بعلم الإحصاء إلى عصور قديمة ، وان تعداد السكان عند القدماء المصرين والصينيين أمثلة توضح اهتمام الحكومات منذ القدم بالمعلومات الاحصائية وذلك لأغراض التنظيم والتخطيط في أحوال السلم والحرب .

ظهرت كلمة إحصاء (statistics) لأول مرة في عام (1749) وهي مشتقة من الكلمة اللاتينية (status) أو الايطالية (statista) وهي تعني الدولة السياسية . اذ كانت الدولة أول

من اهتم بجمع البيانات والمعلومات الكمية وذلك لإدارة شؤون البلاد، وامتدت لتشمل إحصاءات حجم السكان والمواليد والوفيات والإنتاج والاستهلاك والثروة .

لقد تطور علم الإحصاء من مجرد فكرة الحصر والعد إلى أن أصبح في الوقت الحاضر علما له قواعده ونظرياته واساليبه ، ويرجع الفضل في ذلك إلى كثير من العلماء من أمثال برونلي وفردريك جاوس وكيتليه وجولتون وأخيرا كارل بيرسون وبولي وبول فيشر وغيرهم .

ان تطور علم الاحصاء جاء ملازما وموازيا لظهور وتطور علوم عدة مثل نظرية الاحتمالات التي نشأت على أساس رياضي في عام (1494) عن طريق العالم باسيولي ، والدراسات الفلكية لكل من كبلر (1517–1630) وجاليليو (1564–1642). غير أن التاريخ الحقيقي لنظرية الاحتمالات بدء في القرن السابع عشر اذ وضعت أسسها في عام (1654) بواسطة كلا من العالمين : الفيلسوف الفرنسي باسكان (1623 1662) عالم الرياضيات والفيزياء– و العالم فرمات (1608 – 1665) .

وفي عام (1620 – 1674) قام جرونت بنشر ملاحظاته عن معالجة البيانات المتعلقة بالحكومة خاصة في النواحي الطبيعية والسياسية والتجارية والنمو والوفيات والأمراض.

ويعد العالم البلجيكى كتيليه (1796 – 1874) أول من وضع قواعد محددة لعلم الإحصاء ، وكلمة (إحصاء) في الوقت الحاضر ذات معان عدة مثل جمع بيانات تبين الحالة في الدولة كعدد المواليد والوفيات وبيانات عن المحاصيل والتجارة الخارجية ...الخ ويسمى نشر الأجهزة الحكومية لمثل هذه المعلومات في شكل كتب وتقارير " بالإحصاء الرسمي ".

لقد تطور علم الإحصاء وتنوعت نظرياته واساليبه ، وأصبح علما مستقلا يمكن الاستعانة به في معالجة البيانات بانواعها . كما برز دور الإحصاء بما يقدمه من بيانات وإحصاءات في إجراءات التخطيط والتنمية التي تمر بها مجتمعاتنا اليوم .

بمعنى أنه للحصول على معلومات ذات قيمة من تلك البيانات الرقمية فإنها يجب أن تخضع للتحليل الإحصائي (Statistical Analysis) بمساعدة تلك الأساليب والإجراءات والأدوات التي يوفرها لنا علم الإحصاء.

ونجد أن بداية الإحصاء كان مرتبطاً في الغالب بالمجالات الاقتصادية والاجتماعية المتمثلة بتعداد السكان ومعرفة خصائصهم الاجتماعية والاقتصادية وكانت الأساليب الإحصائية المستخدمة تمتاز بالبساطة بحيث لم توفر للإحصاء الأسس والمقومات الكافية لأن يصبح علماً. واستخدم الإحصاء في عصره الأول في جمع البيانات عن السكان وحصرهم من قبل الدولة لأهداف معينة تتمثل في استخدامهم في الجيوش أو توجيههم لتنفيذ بعض المباني أو لغرض فرض الضرائب أو توزيع الأراضي الزراعية على السكان بطريقة عادلة .

وفي القرن السابع عشر والذي يمكن اعتباره العصر الإحصائي الثاني تم استخدام الطريقة الرقمية للدلالة على الظواهر موضوع البحث اذ أن هذه الطريقة أدق في التعبير عن هذه الظواهر وتركز الهدف من هذه الطريقة في معرفة عدد السكان وعدد المواليد وعدد الوفيات ومقدار الثروة والدخل ومقدار الضرائب المحصلة وكمية الناتج من المحاصيل الزراعية.

ويمكن تحديد بداية العصر الإحصائي الثالث مع تطور علوم الرياضيات في القرن الثامن عشر وظهور بعض النظريات العلمية الهامة مثل نظرية الاحتمالات التي كان لها دور كبير في تطور هذا العلم واكتسابه أهمية كبرى بحيث أصبح علماً مستقلاً وانتشر استخدامه

وبدأ الاهتمام من قبل العلماء في تطبيق النظريات والطرائق والأساليب الإحصائية في كثير من فروع العلم الحديث كالهندسة والطب والصيدلة والزراعة والصناعة والجغرافيا والفلك وعلم النفس باعتباره الطريقة الصحيحة والأسلوب الأمثل إتباعه في البحث العلمي.

أنواع الإحصاء

ان علم الإحصاء لا يختلف عن غيره من العلوم فهو يتضمن عددا من المصطلحات او المفاهيم الأساسية التي ينبغي على الباحث او المختص الإلمام بتعريفاتها لكي يعي المقصود منها ويتسنى له معرفة كيفية التعامل معها عندما تعرض له في دراساته وبحوثه ومن ثم يتفادى الخلط بين المصطلحات المختلفة عندما يحاول اختيار الأداة الإحصائية المناسبة لمعالجة البيانات التي قام بجمعها .

ان الأساليب الإحصائية تختلف فيما بينها من حيث الهدف والتدرج من البساطة إلى التعقيد , واختيار الأسلوب الملائم يتحدد وفقا لأهداف البحث ونوعية البيانات المتاحة . وبشكل عام هناك نوعين أساسيين من الإحصاء هما :

1- الإحصاء الوصفي Descriptive Statistics

وهو نوع من الإحصاء يهدف إلى تلخيص البيانات بهدف تحويلها من مجرد كم من الأرقام إلى شكل أو صورة أخرى يمكن فهمها واستيعابها بشكل ملخص ، ومن أغلب الأساليب المستخدمة في الاحصاء الوصفي مقاييس النزعة المركزية و مقاييس التشتت ومعاملات الارتباط والانحدار .

ويقتصر الاحصاء الوصفي على معالجة مجموعة بيانات بقصد استخلاص عدد من الجداول الإحصائية وعرضها في عدد من الأشكال كالرسوم البيانية ، وأن العمليات الإحصائية تدور في جملتها حول إيجاد المتوسطات ودرجات التشتت للبيانات , ولهذا يطلق على العمليات الإحصائية التي تقوم بهذه الوظيفة مصطلح الإحصاء الوصفي , وعلى هذا يستخدم الإحصاء الوصفي في تنظيم وتلخيص ووصف معلومات تخص عينة من العينات , فمن عينة محددة من الطلبة يمكن حساب متوسط التحصيل الدراسي الذي حصلوا عليه , وهذه المقاييس كلها وصفية بحتة لا تفيد في حد ذاتها في الاستنتاج أو التنبؤ وإنما تصف الكيفية التي تتوزع بها البيانات التي تم الحصول عليها من الطلبة موضوع البحث ٠

ويعد الوصف من الوظائف الأساسية لعلم الإحصاء وباستخدام أسلوب التحليل الإحصائي للبيانات أصبح من اليسير إمكانية تحديد خصائص الظاهرة المدروسة حتى عن طريق الأشكال البيانية التي تمثل بيانات الظاهرة والتي تسهل وتبسط تحديد خصائص الظاهرة واتجاهاتها العامة . وتعد عملية جمع البيانات من أقدم وظائف الإحصاء .

2- الإحصاء الاستدلالي Inferential Statistics

يهدف هذا النوع من الأساليب الإحصائية الى الوصول إلى تقديرات لمعالم وخصائص مجتمعات الدراسة من خلال ما هو متوافر من معلومات عن العينات المختارة من تلك المجتمعات فضلا عن اختبار الفرضيات الإحصائية عن مجتمع البحث على أساس البيانات المتاحة عن عينات الدراسة . ويطلق على هذا النوع من الأساليب أكثر من تسمية تؤدي جميعها إلى نفس المعنى فأحيانا يسمى بالإحصاء الاستدلالي أو الاستنباطي (Inductive) أو التعميمي (Generalizing) فهو يهدف الوصول إلى تعميمات عن مجتمع الدراسة من

خلال العينة المختارة من هذا المجتمع . ويشمل هذا النوع من الأساليب الإحصائية ، العينات ، اختبار الفرضيات ، الاستدلال من خلال عينة واحدة أو أكثر وما يتضمنه ذلك من اختبارات مختلفة .

المتغيرات Variables :

تعرف المتغيرات بانها خصائص يشترك فيها أفراد المجتمع الإحصائي ولكنها قد تختلف من فرد إلى فرد آخر، مثل التحصيل الدراسي , الذكاء ، ألطول , مستوى الدخل ، الاتجاه نحو العولمة . وتتميز هذه المتغيرات بأنها قابلة للقياس الكمي، أي بإمكانية تحديد قيمة كمية او رقمية معينة لكل منها . والمتغيرات كذلك هي ظواهر أو أحداث أو خصائص تأخذ قيما تتغير من ظرف لآخر وهي الوحدات الأساسية للتحليل الإحصائي . والمتغيرات التي تقاس كمياً تنقسم من حيث قيمتها العددية إلى نوعين أساسيين هما :-

1 – المتغير المتصل او المستمر Continuous Variable .

ان المتغير يكون متصلا او مستمرا عندما يأخذ أي قيمة متدرجة على المقياس المستخدم . مثال ذلك قياس درجات الحرارة باستخدام المحرار فالمتغير يأخذ أي قيمة بين رقمين صحيحين، بمعنى أن المتغير يمكن أن يأخذ أي قيمة بين 40 درجة و 41 درجة مثل (40,1 ، 40,2 .... الخ) . ومن امثلة المتغيرات المتصلة في العلوم التربوية والنفسية هي : التحصيل الدراسي ، الاتجاه نحو المواد الدراسية , الميول المهنية , التفكير الاستدلالي وغيرها من المتغيرات .

2 – المتغير المتقطع او المنفصل Discrete Variable

ان المتغير المتقطع او المنفصل هو الذي يحتوي مداه على عدد محدود من القيم أو يحتوي عدد لا نهائي من القيم ولكن لكل منها قيمة محددة يمكن عدها أو ترتيبها في نهاية الأمر مثل عدد الأولاد أو الأفراد في الأسرة , والذي لابد أن يكون عددا صحيحا مثل 1 ، 2 ، 3 ، 4 .... وهكذا . ويتم التعبير عن المتغيرات المنفصلة او المتقطعة بقيم عددية غير قابلة للتجزئة اذ يرمز الباحث للذكور برقم (1) وللإناث برقم (2) على سبيل المثال ، ولا توجد قيمة تتوسطهما , وكذلك الحال بالنسبة لسعة الوحدة السكنية ، فالشقة إما أن تكون غرفة واحدة أو غرفتين أو ثلاث أو أكثر وليس هناك جزء من غرفة . والبيانات التي يتم جمعها عن المتغيرات المتقطعة تكون بيانات متقطعة أيضاً أي أنها غير قابلة للتجزئة ولا نجد لها كسورا اعتيادية او عشرية . فلا يستطيع الباحث أن يدعي أن العينة تتكون من عشرة ذكور ونصف , ومن أمثال المتغيرات المتقطعة عدد الطلاب في صف معين ، عدد أيام الإنتاج في احد المصانع عدد حوادث السيارات وهكذا .

الفصل الثاني

تبويب البيانات وعرضها

الفصل الثاني

تبويب البيانات وعرضها

يقصد بتبويب البيانات هو عرض البيانات او الدرجات الخام في جداول او مخططات معينة بهدف تلخيصها واستيعابها واستنتاج النتائج منها ومقارنتها بغيرها من البيانات ، كما يسهل الرجوع إليها في صورة جداول ومخططات دون الاطلاع على الاستمارات الأصلية التي قد تحمل أسماء أصحابها مما يخل بمبدأ سرية البيانات . اذ يعد عرض وتبويب البيانات الإحصائية الخطوة الثانية بعد تجميع هذه البيانات الخام في مفهوم التحليل الإحصائي. وتتوقف طريقة عرض البيانات على نوع هذه البيانات وعلى الحقائق المطلوب إبرازها. وهناك طريقتان أساسيتان لعرض وتبويب البيانات الإحصائية وهما :

أولاً : عرض البيانات الإحصائية باستخدام الجداول:

بعد عملية تبويب وتعيين الصفات او البيانات التي تميز المفردات او افراد العينة ، ترصد النتائج في جداول مناسبة توضح الشكل النهائي للمجموعات المميزة , وتسمى هذه العملية التي يتم تجميع البيانات في مجموعات مميزة ومتجانسة بعملية التصنيف , ويمكن التمييز بين أنواع من الجداول الإحصائية نذكرها فيما يلي:-

(أ) تبويب البيانات الخام في جدول تكراري بسيط :

ان المقصود بالجدول التكراري البسيط هو ذلك الجدول الذي يتم وضع قيم الدرجات او البيانات فيه مرتبة ترتيباً تصاعدياً أو تنازليا في عموده الأول أما العمود الثاني فيسمى بعمود التكرار ويرصد فيه عدد مرات تكرار كل درجة أو حدث .

مثال :

البيانات الاتية هي درجات حصل عليها (30) تلميذا في مادة العلوم في امتحان نهاية السنة :

9	7	7	6	4	9	7	8	4	6
5	6	8	7	6	8	10	5	4	10
5	8	6	4	10	7	4	5	6	8

المطلوب تبويب هذه البيانات في جدول توزيع تكراري بسيط .

الحل :

يتم ترتيب البيانات بوضع هذه البيانات في العمود الأول من الجدول وتسمى (س) ثم وضع عدد مرات تكرارها باستخدام العلامات في العمود الثاني أما العمود الثالث فيمثل التكرار رقما ويرمز له بالرمز (ك) وكما في الجدول الأتي :-

ك	العلامات	س
5	~~////~~	4
4	////	5
6	/ ~~////~~	6
5	~~////~~	7
5	~~////~~	8
2	//	9
3	///	10
30	المجموع	

مثال :

الجدول الأتـي يمثل تقديرات (20) طالباً في مادة علم النفس التربوي , والمطلوب هو وضـع هذه البيانات في جدول تكراري بسيط ؟

جيد	مقبول	امتياز	جيد	متوسط	جيد	جيد جداً	جيد جدا	جيد	جيد جداً
امتياز	متوسط	جيد	جيد جداً	مقبول	جيد	متوسط	جيد	جيد	مقبول

الحل :

التقدير	العلامات	التكرار
مقبول	///	3
متوسط	///	3
جيد	~~////~~ ///	8
جيد جداً	////	4
ممتاز	//	2
المجموع		20

(ب) تبويب البيانات في جدول تكراري ذي فئات :

الفئـة هـي مجموعـة مـن البيانـات متشـابهة فـي الصـفات إلـى حد كبيـر جداً ، وفـي حالـة زيـادة عدد البيانـات الخـام وزيـادة انتشـارها لا يمكـن اسـتخدام الجداول البسـيطة فـي التعبيـر عن هذه الحـالات ، اذ اننـا سنحتـاج إلـى جـدول يضـم عددا كبيـرا مـن الصـفوف ، لـذلك يـتم تقسـيم البيانات إلى مجموعات متقاربة ومتشابهة في الصفات تسمى فئات . وتوجد عدة طرائق لكتابـة الفئات هي :

الطريقة الأولى :

نذكر كلا من الحد الأدنى والحد الأعلى للفئة كما في الجدول الأتي :

الفئة	التكرار
20−10	7
30−20	9
40−30	11
50−40	8

وهذه الطريقة فيها سلبية كبيرة وذلك لأن نهاية الفئة الأولى تساوي (20) هي نفسها بداية الفئة الثانية وهكذا في بقية الفئات , وفي هذه الحالة لا نعرف إلى أي فئة ينتمي الرقم (20) .

الطريقة الثانية :

في هذه الطريقة نذكر كلا من الحد الأدنى والحد الأعلى لكل فئة ولكن نقوم بترك فاصل مقداره واحد صحيح بين نهاية كل فئة وبداية الفئة التي بعدها مباشرة ، وكما في الجدول الاتي .

الفئة	التكرار
19−10	7
29−20	9
39−30	11
49−40	8

ومن سلبيات هذه الطريقة هي أنها لا تصلح في حالة البيانات التي تحتوي على كسور، فمثلا نحن لا نستطيع تمثيل البيانات (19,8 − 29,2 − 39,9) في الجدول أعلاه .

الطريقة الثالثة :

نـذكر الحـد الأدنـى فقـط للفئـة ونضـع بعـده شـارحة (−) وهـذه الطريقـة تصـلح لكافـة الظواهر. وكما في الجدول الاتي :−

التكرار	الفئة
7	−10
9	−20
11	−30
8	−40

الطريقة الرابعة :

هذه الطريقة تصلح لكافة الظواهر أيضاً ولكنها أقل شيوعاً , اذ نذكر الحد الأعلى فقط للفئة ونضع قبله شارحة وكما في الجدول الأتي :−

التكرار	الفئة
7	20 −
9	30 −
11	40 −
8	50 −

(ج) بناء جدول التوزيع التكراري ذي الفئات :−

من اجل إعداد او بناء جدول تكراري ذي فئات , يمكن إتباع الخطوات الآتية :−

1− نحسب المدى من خلال طرح أكبر قيمة من أصغر قيمة في البيانات .

2− نحسب عدد الفئات من خلال العلاقة الآتية :−

عدد الفئات = 3,3 مضروبا في لوغاريتم (عدد البيانات)

3- نحسب طول الفئة من خلال قسمة المدى على عدد الفئات .

4- نختار الحد الأدنى للفئة الأولى (أي بدايتها) والذي يساوي أقل قيمة موجودة ضمن البيانات أو أقل منها بقليل .

5- نعد الجدول ونضع العلامات التي تمثل التكرار .

مثال : قام باحث بجمع بيانات تمثل درجات (50) طالباً في مقياس الاتجاه نحو مادة الرياضيات ، وكانت درجاتهم كما في الجدول الأتي :-

42	84	30	46	55	40	23
57	39	35	63	59	36	25
	53	25	63	47	60	45
	55	48	82	39	65	33
	42	26	65	61	58	64
	55	70	45	53	52	50
	64	55	54	49	45	65
	78	51	52	41	42	75

والمطلوب هو إعداد جدول توزيع تكراري ذي فئات للبيانات أعلاه ؟

الحل :

من اجل حل هذا المثال نقوم بإتباع الخطوات الآتية :

1- نقوم بحساب المدى :

المدى = أكبر قيمة – أصغر قيمة

61 = 23 – 84 =

2- نحسب عدد الفئات = 3,3 × لو (ن)

= 3,3 × لو (50)

= 3,3 × 1,699 = 5.6

3- نقرب عدد الفئات لأقرب رقم صحيح فيكون عدد الفئات = 6

4 – طول الفئة = المدى / عدد الفئات

= 61 / 6

= 10,17

5- نقرب طول الفئة لأقرب رقم صحيح فتصبح

طول الفئة = 10

6- نختار بداية الفئة الأولى , ومن اجل تسهيل الحسابات نختار الرقم (20) كبداية للفئة الأولى .

7- نبدأ في إعداد الجدول كالتالي :

الفئات	العلامات	التكرار
20-	////	4
30-	/ ~~////~~	6
40-	// ~~////~~ ~~////~~	12
50-	//// ~~////~~ ~~////~~	14
60-	//// ~~////~~	9
70-	///	3
80-90	//	2
المجموع		50

ثانياً : العرض البياني للبيانات الإحصائية

يُعد العرض البياني للبيانات الإحصائية بمثابة تلخيص لهذه البيانات في شكل يسهل منه استيعاب خصائص موضوع الدراسة ، وتختلف طرائق عرض البيانات المبوبة عن البيانات الغير مبوبة ، وسنتعرض لكل منها بالتفصيل فيما يلي :-

أ : العرض البياني للبيانات غير مبوبة :

والمقصود بالبيانات غير المبوبة هي تلك البيانات المفردة , أي لا يوجد فيها فئات ولا تكرارات وهناك عدة طرائق لعرضها منها :-

(1) طريقة الأعمدة البيانية البسيطة :

وفي هذه الطريقة نعد شكلا بيانيا اذ يمثل محور السينات قيم المتغير أما محور الصادات فيمثل القيمة المقابلة لقيمة المتغير ويتم رسم مستطيل ارتفاعه يمثل قيمة المتغير .

مثال :

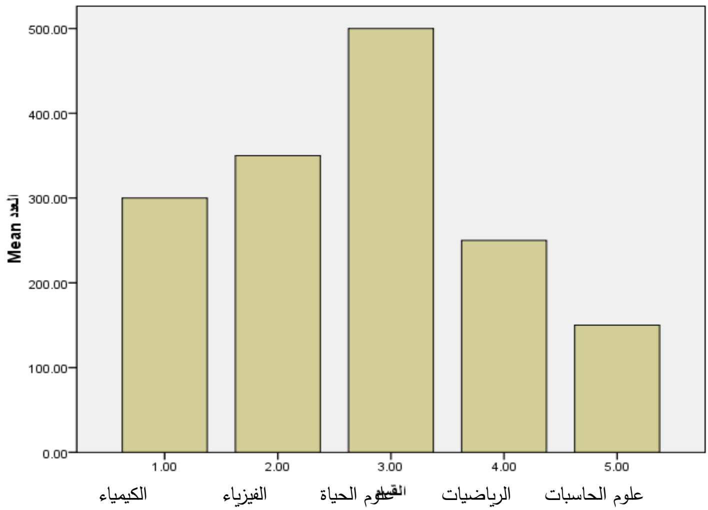
في الجدول التالي أعداد الطلبة في بعض أقسام كلية التربية (ابن الهيثم) في جامعة بغداد للعام الدراسي 2013-2014

القسم	الكيمياء	الفيزياء	علوم الحياة	الرياضيات	علوم الحاسبات
عدد الطلبة	300	350	500	250	150

والمطلوب عرض هذه البيانات باستخدام طريقة الأعمدة البيانية البسيطة ؟

الحل :

نعد شكلا بيانيا يمثل محور السينات فيه متغير القسم , اما محور الصادات فيمثل عدد الطلبة ، وكما في الشكل الأتي :-



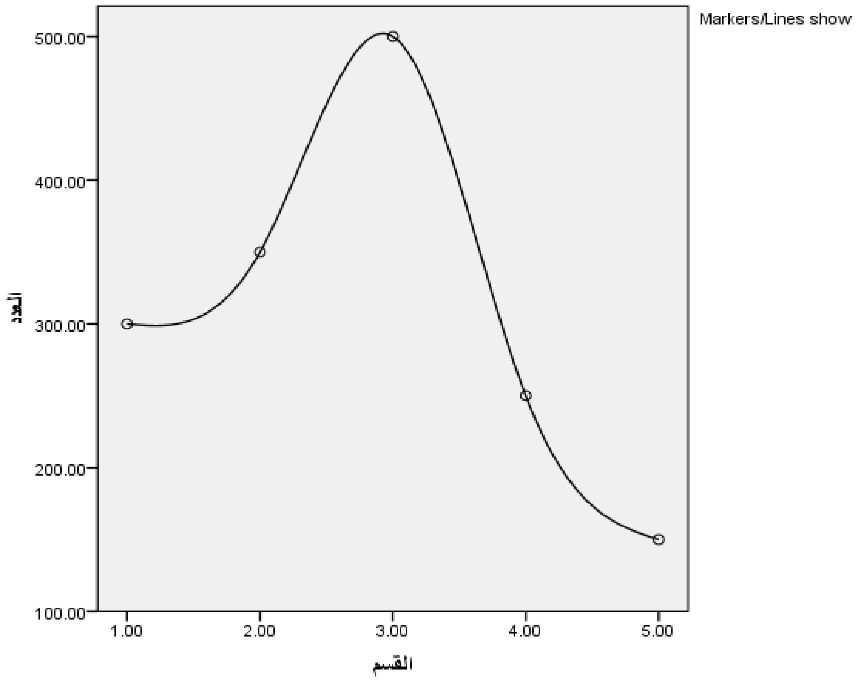
(2) طريقة المنحنى البياني البسيط :

في هذه الطريقة كما في الطريقة السابقة يمثل محور السينات المتغير أما محور الصادات يمثل قيمة المتغير ويتم وضع نقاط بين كل قيمة من قيم المتغير على محور السينات والقيمة المقابلة على محور الصادات ثم يتم توصيل تلك النقاط بخط منحني .

مثال :

كيف يتم عرض البيانات في المثال السابق بطريقة المنحني البياني البسيط ؟

الحل :-



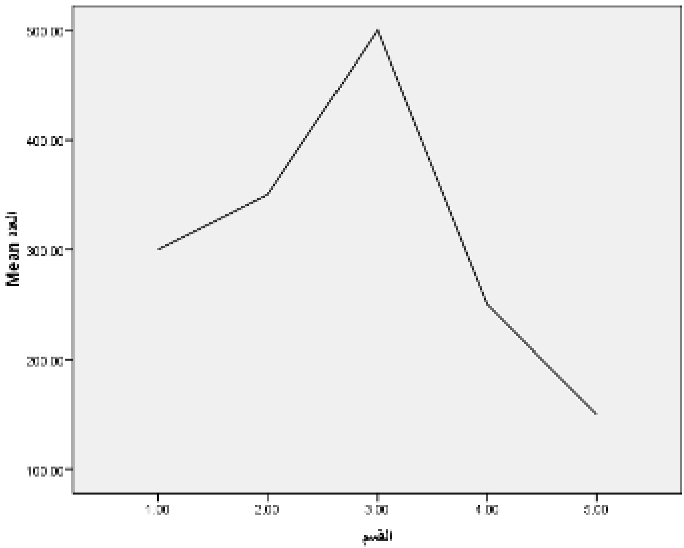
علوم الحاسبات الرياضيات علوم الحياة لفيزياء الكيمياء

(3) طريقة المضلع التكراري البسيط :

وفي هذه الطريقة يمثل محور السينات المتغير أما محور الصادات يمثل قيمة المتغير ويتم تحديد نقاط بين كل قيمة من قيم المتغير على محور السينات والقيمة المقابلة على محور الصادات ثم يتم توصيل تلك النقاط بقطع (خطوط) مستقيمة .

مثـــال : كيـف يـتم عـرض البيانــات فــي المثــال الســابق بطريقــة المضــلع التكــراري البسيط ؟

الحل :-



الكيمياء الفيزياء علوم الحياة الرياضيات علوم الحاسبات

(4) طريقة الدائرة البيانية :

هذه الطريقة تختلف عن الطرائق السابقة اذ يتم رسم دائرة ثم نحسب زاوية قطاع كل قيمة على حدة ونقوم برسم تلك الزاوية داخل الدائرة , اذ يتم حساب زاوية قطاع كل قيمة من العلاقة :

$$\text{زاوية قطاع القيمة} = \frac{\text{تكرار القيمة}}{\text{مجموع التكرارات}} \times 360$$

مثال : كيف يتم عرض البيانات في المثال السابق بطريقة الدائرة البيانية ؟

الحل : نقوم بحساب مجموع التكرارات =

$$1550 = 150 + 250 + 500 + 350 + 300$$

مقدار زاوية قطاع قسم الكيمياء = $\frac{300}{1550} \times 360 = 69,68$

مقدار زاوية قطاع قسم الفيزياء = $\mathbf{\frac{350}{1550} \times 360 = 81,29}$

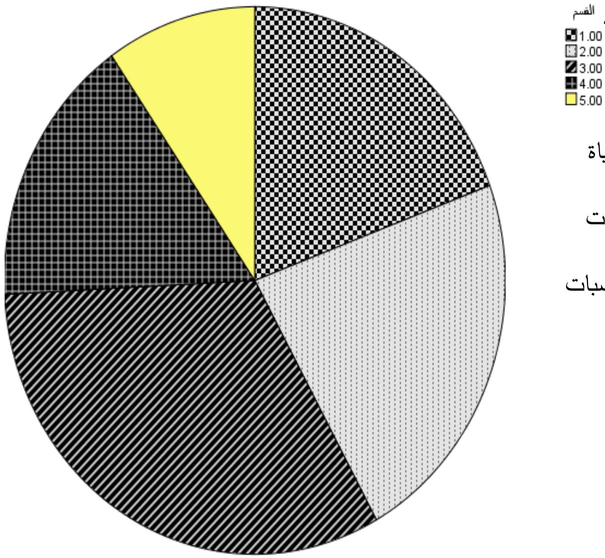
مقدار زاوية قطاع قسم علوم الحياة = $\mathbf{\frac{500}{1550} \times 360 = 116,13}$

مقدار زاوية قسم الرياضيات = $\frac{250}{1550} \times 360 = 58,06$

مقدار زاوية قسم علوم الحاسبات = $\frac{150}{1550} \times 360 = 34,84$

ملاحظة :- من اجل التأكد من ان العمليات الحسابية التي اجريناها صحيحة نقوم بجمع مقادير الزوايا التي قمنا بحسابها فاذا كان مجموعها يساوي **(360)** فهذا يدل على ان العمليات الرياضية التي اجريناها صحيحة .

وبذلك تكون الدائرة البيانية كما في الشكل الاتي :-



الكيمياء

الفيزياء

علوم الحياة

الرياضيات

علوم الحاسبات

طريقة عرض البيانات بيانيا باستخدام الحقيبة الاحصائية (SPSS) :- (استخدم الاصدار (11) في هذه الطبعة) .

قبل ان نتناول موضوع طريقة عرض البيانات بيانيا باستخدام الحقيبة الاحصائية (SPSS) سنعطي فكرة مختصرة عن هذه الحقيبة .

ان الحقيبة الاحصائية هي عبارة عن برنامج يستخدم في معالجة البيانات باحدث الطرائق الاحصائية، ويطلق عليها بالرمز (SPSS) وهي الاحرف الاولى من (Statistics Package for Social Sciences) والتي تعني (الحقيبة الاحصائية للعلوم الاجتماعية) .

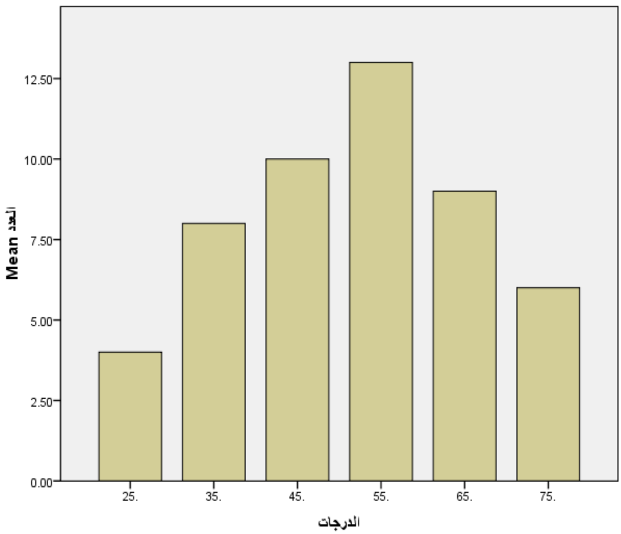
في البدء ينبغي تنصيب برنامج الحقيبة الاحصائية في الحاسبة , ومن اجل تشغيل هذا البرنامج في الحاسبة الالكترونية ننقر فوق زر " ابدأ " أو "Start" من شاشة تشغيل النوافذ

فئات الدرجات	-20	-30	-40	-50	-60	-70
عدد الطلاب	4	8	10	13	9	6

الحل : نقوم باعداد جدول وكالاتي :-

الفئة	التكرار	مركز الفئة
-20	4	25
-30	8	35
-40	10	45
-50	13	55
-60	9	65
-70	6	75

ونرسم مدرج تكراري وكالاتي :-



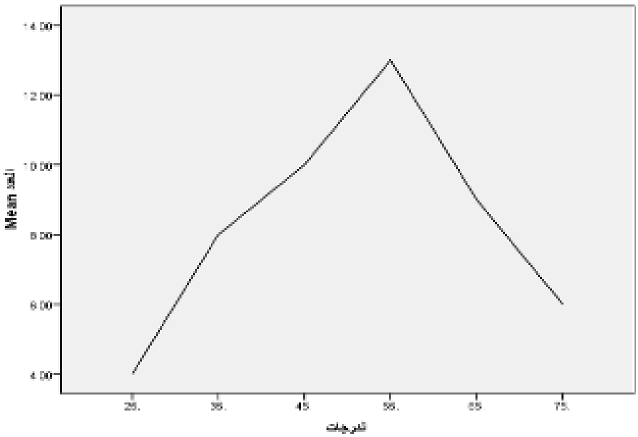
(2) طريقة المضلع التكراري :

في هذه الطريقة يكون الاحداثي السيني للشكل هو مركز الفئة بينما الاحداثي الصادي هو التكرار ، ، ثم نوصل كل نقطتين متتاليتين بقطعة مستقيمة .

مثال :

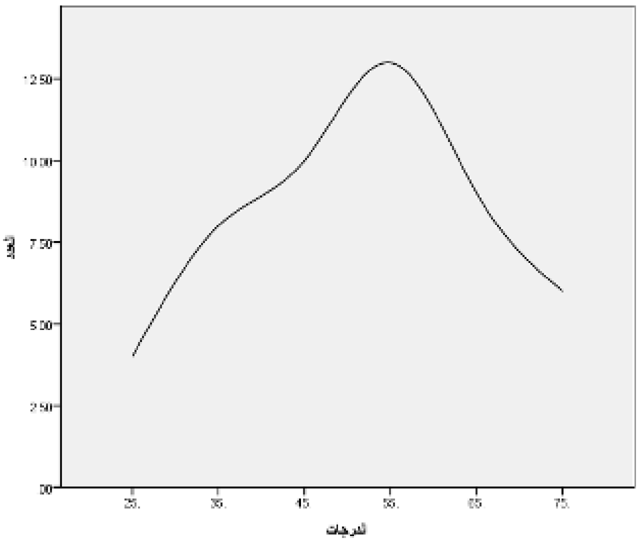
اعرض الجدول السابق بيانياً باستخدام طريقة المضلع التكراري ؟

الحل : نقوم برسم الشكل الاتي :



(3) طريقة المنحنى التكراري :

بعد رصد النقاط كما في طريقة المضلع التكراري نوصل كل نقطتين متتاليتين بمنحنٍ بسيط , وكما في الشكل الاتي :-



الفصل الثالث

مقاييس النزعة المركزية

Measures Of Central Tendency

الفصل الثالث

مقاييس النزعة المركزية

Measures Of Central Tendency

تعتمد الطرائق البيانية المستخدمة في تحليل ودراسة المتغيرات لتحديد الخصائص والاتجاهات والعلاقات في دقتها على دقة التمثيل البياني نفسه ، وبذلك من الممكن ان تختلف الخصائص من رسم إلى آخر لنفس المتغير ، كما ان هذه الطرائق لا تستخدم اذا كان هدفنا من الدراسة المقارنة بين مجموعتين او اكثر, وعليه فإنه من الأفضل اللجوء إلى طرائق جديدة تعتمد الطريقة الرياضية في القياس والتمثيل . ومن هذه الطرائق هي حساب مقاييس النزعة المركزية .

يعرف مقياس النزعة المركزية بانه قيمة مركزية قريبة من النقطة التي يتجمع عندها اكبر عدد من البيانات او الدرجات . كما يعرف بانه الدرجة التي يمكن ان تعد بانها ممثلة للدرجات او البيانات الموجودة في المجموعة .

ان الهدف الأساسي من استخدام مقاييس النزعة المركزية هو تلخيص البيانات في محاولة أخرى لوصفها عن طريق التعرف على مركزها ، ومن خلال هذا المؤشر يتمكن الباحث من فهم بعض أبعاد الظاهرة قيد الدراسة.

ومن أهم مقاييس النزعة المركزية التي سنتناولها هي :- الوسط الحسابي والوسيط والمنوال ، كما سنتعرض بالدراسة لحساب كل منهم من البيانات المفردة (الغير مبوبة) ومن البيانات المبوبة .

أولاً : الوسط الحسابي Mean

مقدمة

ويطلق عليه احيانا بـ (المتوسط الحسابي) والوسط الحسابي لمجموعة من القيم هو مجموع هذه القيم مقسوماً على عددها ويرمز له بالرمز ($\bar{س}$) أو ($\bar{X}$) ، ونحيط القارئ الكريم علما باننا سنعتمد الرموز العربية والانكليزية في كتابة القوانين لاننا سنضمن كتابنا هذا مبادئ كيفية استخدام الحقيبة الاحصائية (spss) في حساب الوسائل الاحصائية وهي تستخدم الرموز الانكليزية حصرا .

حساب الوسط الحسابي للبيانات غير المبوبة :

يحسب الوسط الحسابي من البيانات غير المبوبة من العلاقة التالية

$$\bar{س} = \frac{\text{مجـ س}}{\text{ن}} \qquad \bar{x} = \frac{\sum x}{n}$$

اذ ان :-

$\bar{س}$ = $\bar{x}$ = الوسط الحسابي

مجـ = Σ = مجموع

س = x = القيمة او الدرجة

ن = n = عدد الأفراد او عدد الدرجات

مثال :- احسب الوسط الحسابي لدرجات (7) تلاميذ في مادة اللغة العربية والتي كان درجاتهم كالأتي: 4 – 7 – 5 – 6– 10 – 8 – 9

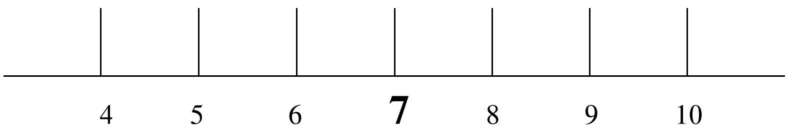
الحل :

$$\bar{\text{س}} = \frac{\text{مجـ س}}{\text{ن}}$$

$$\bar{\text{س}} = \frac{9 + 8 + 10 + 6 + 5 + 7 + 4}{7}$$

$$\bar{\text{س}} = \frac{49}{7} = 7$$

من خلال رسم خط الأعداد للبيانات السابقة



نجد ان قيمة الوسط الحسابي (7) تتوسط تقريبا القيم , لذلك فان الوسط الحسابي كما ذكرنا سابقا هو القيمة التي تتمركز حولها البيانات او الدرجات .

حساب الوسط الحسابي من البيانات المبوبة

يمكن حساب الوسط الحسابي للبيانات المبوبة عن طريق العلاقة الاتية :-

$$\bar{\text{س}} = \frac{\text{مجـ (س × ك)}}{\text{مجـ ك}} \qquad \bar{X} = \frac{\Sigma\,(x.f)}{\Sigma\, f}$$

اد ان :-

$\bar{\text{س}}$ = $\bar{x}$ = الوسط الحسابي

مجـ = Σ = مجموع

س = x = مركز الفئة

ك = f = التكرار

مثال :

قام باحث بقياس مستوى الطموح لدى عينة مكونة من (100) طالب ونظم البيانات في جدول تكراري كما في ادناه والمطلوب حساب الوسط الحسابي لدرجات العينة .

فئات درجات المقياس	40-	50-	60-	70-	80-	90-	100-110
عدد الطلاب	9	12	18	32	15	7	7

الحل :

من ملاحظة العلاقة الخاصة بحساب الوسط الحسابي للبيانات المبوبة نجد اننا بحاجة الى حساب كل من (س× ك) و (مجـ ك) ، لذا نقوم باعداد الجدول الاتي :-

ف	ك	س	س × ك
40-	9	45	405
50-	12	55	660
60-	18	65	1170
70-	32	75	2400
80-	15	85	1275
90-	7	95	665
100-110	7	105	735
المجموع	100	525	7310

$$\bar{س} = \frac{\text{مجـ (س × ك)}}{\text{مجـ ك}}$$

$$\bar{س} = \frac{7310}{100} = 73,10$$

خصائص الوسط الحسابي :

نورد فيما يأتي بعض الخصائص التي يتميز بها الوسط الحسابي :

1- المجموع الجبري لانحرافات القيم عن وسطها الحسابي يساوي صفر دائما، ففي المثال السابق في صفحة (60) اذا قمنا بحساب انحرافات القيم عن الوسط الحسابي أي طرح كل قيمة من الوسط الحسابي فاننا سنحصل على النتيجة الاتية وكما في الجدول

س	س – $\overline{س}$
4	3−
7	صفر
5	2−
6	1−
10	3+
8	1+
9	2+
المجموع	صفر

2- يتأثر الوسط الحسابي بالقيم المتطرفة (زيادة او انخفاضا) , وبهذه الحالة فان الوسط الحسابي لا يعطي صورة صحيحة للبيانات التي تضم قيما متطرفة . وكما في المثال الاتي :-

نفرض ان هناك مدرسا رغب في حساب الوسط الحسابي لدرجات طلابه الخمسة والتي كانت درجاتهم (65 , 74 , 81 , 73 , 90) ومن خلال تطبيق قانون الوسط الحسابي وجد ان الوسط الحسابي يساوي (76,6) وبعد فترة انضم احد الطلاب الجدد الى هذه المجموعة وقد كانت درجته في هذه المادة (32) وقام المدرس باعادة حساب الوسط الحسابي لطلابه الستة ووجد بانه يساوي (69,17) !!!!!.

أهمية الوسط الحسابي

إن لاستخراج قيمة الوسط الحسابي لمجموعة من البيانات او الدرجات أهمية كبيرة تتجلى في النقاط الآتية :

1- يستخدم لتلخيص مجموعة كبيرة من البيانات او الدرجات .

2- يستخدم في حساب بقية مقاييس النزعة المركزية كالمنوال او الوسيط .

3- يستخدم في حساب كثير من الوسائل الحسابية والاحصائية مثل الانحراف المتوسط والانحراف المعياري والتباين والاختبارات التائية بانواعها كما سنلاحظ في الفصول القادمة .

ثانياً : الوسيط Median

يعرف الوسيط على أنه القيمة التي تتوسط مجموعة من القيم إذا رتبت ترتيباً تصاعدياً أو تنازلياً .

حساب الوسيط من البيانات غير المبوبة (المفردة)

يعتمد حساب الوسيط من البيانات غير المبوبة على عدد تلك البيانات فهناك احتمالين :

(1) إذا كان عدد البيانات فرديا :-

ففي هذه الحالة نقوم بالخطوات الاتية :-

أ- نرتب الدرجات او البيانات تصاعديا (من اقل درجة الى اكبر درجة) او تنازليا (من اكبر درجة الى اقل درجة) .

ب- نقوم بحساب تسلسل الوسيط من خلال العلاقة :

$$\frac{(ن + 1)}{2}$$

اذ ان (ن) تمثل عدد القيم او البيانات , ونحدد قيمة الوسيط من خلال تسلسله الذي حصلنا عليه .

مثال : احسب الوسيط للبيانات الاتية :-

17 - 8 – 20 – 11 – 7 – 10 – 13 – 12 – 19

الحل :

نرتب البيانات او الدرجات ترتيبا وليكن تنازليا , فستكون بالترتيب الاتي :-

7 - 8 - 10 - 11 - 12 - 13 - 17 - 19 - 20

نحسب ترتيب الوسيط من العلاقة :

$$\frac{ن+1}{2} = \frac{9+1}{2} = 5$$

وهذا يعني ان تسلسل الوسيط في البيانات بعد ترتيبها هو الخامس , وكما في الشكل

7	**8**	**10**	**11**	**12**	**13**	**17**	**19**	**20**

أي ان الوسيط يساوي (12) .

(2) إذا كان عدد البيانات زوجيا

في هذه الحالة لا توجد قيمة تتوسط القيم كما في الحالة السابقة لان عدد البيانات زوجيا لذلك نقوم بالخطوات الاتية :-

أ- نرتب الدرجات او البيانات تصاعديا (من اقل درجة الى اكبر درجة) او تنازليا (من اكبر درجة الى اقل درجة) .

ب– نقوم بحساب تسلسل القيمتين اللتين تتوسطان القيم من خلال العلاقتين :

$$\frac{ن}{2} \quad و \quad \frac{ن}{2} + 1$$

ج – نحدد القيمتين الوسطيتين وبعد ذلك نقوم بحساب الوسط الحسابي لهما , والناتج يمثل الوسيط .

مثال : لديك البيانات الاتية :–

12	17	7	14	9	5	10	16

والمطلوب حساب الوسيط لهذه البيانات .

الحل :

لحل هذه المسالة نقوم بالخطوات الآتية :

1– نرتب البيانات ترتيبا وليكن ترتيبا تنازليا وكالاتي :

17	16	14	12	10	9	7	5

1- نحسب ترتيب القيمتين الوسطيتين من العلاقتين

$$\frac{ن}{2} \quad و \quad \frac{ن}{2} + 1 \quad \text{اذ ان } (ن = 8)$$

أي ان ترتيب القيمة الاولى = $\frac{ن}{2} = \frac{8}{2} = 4$

و ترتيب القيمة الثانية = $\frac{ن}{2} + 1 = \frac{8}{2} + 1$

$= 4+1 = 5$

أي ان ترتيبي القيمتين الوسطيتين هما الرابع والخامس وكما مؤشر في الاتي :–

17	16	14	12	10	9	7	5

2- نحسب الوسط الحسابي لهاتين القيمتين وكما ياتي :-

$$\frac{12 + 10}{2} = \frac{22}{2} = 11$$

وهذه القيمة تمثل الوسيط للبيانات أعلاه .

حساب الوسيط من البيانات المبوبة

توجد عدة طرائق في حساب قيمة الوسيط للبيانات المبوبة , وسنكتفي بطريقة واحدة هي طريقة التكرار المتجمع الصاعد , وذلك باستخدام العلاقة الاتية :-

$$\text{و} = \text{الحد الأدنى للفئة الوسيطية} + \frac{\text{ترتيب الوسيط} - \text{التكرار المتجمع الصاعد السابق}}{\text{التكرار المتجمع الصاعد اللاحق} - \text{التكرار المتجمع الصاعد السابق}} \times \text{ل}$$

اذ ان :-

و = الوسيط

$$\text{ترتيب الوسيط} = \frac{\text{مجموع التكرارات}}{2}$$

ل = طول الفئة .

مثال : البيانات في الجدول الاتي تبين درجات (200) طالبا بعد إكمالهم لاختبار في مادة الفيزياء والمطلوب حساب الوسيط لهذه الدرجات .

فئات الدرجات	40 -	50 -	60-	70 -	80-	90 -100
التكرار	25	34	53	41	30	17

الحل : من ملاحظة قانون الوسيط نجد باننا بحاجة الى حساب التكرار المتجمع الصاعد , لذا نقوم باعداد جدول وكالاتي :-

الفئات	التكرار	التكرار المتجمع الصاعد
40 -	25	صفر
50 -	34	25
60 -	53	59
70 -	41	112
80 -	30	153
90 - 100	17	183
المجموع	200	200

$$\text{ترتيب الوسيط} = \frac{\text{مجموع التكرارات}}{2} = \frac{200}{2} = 100$$

نبحث في الجدول السابق في عمود التكرار المتجمع الصاعد عن القيمتين التي يقع بينهما ترتيب الوسيط , وهاتان القيمتان هما (59 , 112) , ونؤشر على كلا القيمتين , وكما في الجدول الاتي :-

الفئات	التكرار	التكرار المتجمع الصاعد
40 -	25	صفر
50 -	34	25
60 -	53	59
70 -	41	112
80 -	30	153
90 - 100	17	183
المجموع	200	200

وبذلك فأن :-

الحد الادنى للفئة الوسيطية = 60

التكرار المتجمع الصاعد السابق = 59

التكرار المتجمع الصاعد اللاحق = 112

طول الفئة = الحد الاعلى للفئة – الحد الادنى للفئة = 50 – 40 = 10

$$\text{الوسيط} = \text{الحد الأدنى للفئة الوسيطية} + \frac{\text{ترتيب الوسيط} - \text{التكرار المتجمع الصاعد السابق}}{\text{التكرار المتجمع الصاعد اللاحق} - \text{التكرار المتجمع الصاعد السابق}} \times \text{ل}$$

$$= 60 + \frac{59 - 100}{59 - 112} \times 10$$

$$= 67{,}73$$

أهمية الوسيط

ان لاستخراج قيمة الوسيط لمجموعة من البيانات او الدرجات اهمية كبيرة تتجلى في النقاط الاتية :

1- يستخدم لتلخيص مجموعة كبيرة من البيانات او الدرجات .

2- يستخدم في حساب بقية مقاييس النزعة المركزية كالوسط الحسابي او المنوال .

3- يستخدم في حساب بعض من الوسائل الحسابية مثل الالتواء كما سنلاحظ في الفصل القادم .

ثالثاً : المنوال Mode

يعرف المنوال لمجموعة من الدرجات او البيانات بانه القيمة الأكثر شيوعاً أو تكراراً في تلك الدرجات او البيانات .

حساب المنوال من البيانات الغير مبوبة

اذا كان لدينا مجموعة من الدرجات او البيانات , ففي حالة تكرار درجة او قيمة واحدة فيتم اختيارها كمنوال , أما في حالة تكرار درجتين او رقمين بنفس عدد مرات , فيتم اختيارهما معاً كمنوال , , وفي حالة عدم تكرار أي درجة او رقم , ففي هذه الحالة نقول انه لا يوجد منوال .

مثال : احسب المنوال لكل مجموعة من البيانات الاتية :-

6	9	3	6	8	4	2	المنوال = 6
8	5	6	8	5	7	5	المنوال =5
4	7	2	4	8	9	7	المنوال = 4 و 7
6	2	3	9	5	7	1	لا يوجد منوال

حساب المنوال من البيانات المبوبة

توجد عدة طرائق لحساب المنوال من البيانات المبوبة , وسنكتفي بشرح طريقة واحدة , والتي يطلق عليها بطريقة الرافعة . اذ يمكن حساب المنوال لعدد من البيانات باستخدام العلاقة الاتية :

$$\text{المنوال} = أ + \frac{ك1}{ك1 + ك2} \times ل$$

اذ ان :-

أ = الحد ألأدنى لفئة المنوال والمقصود بدايتها .

ك1 = تكرار الفئة التي تسبق فئة المنوال

ك2 = تكرار الفئة التي تلي فئة المنوال

ل = طول الفئة.

ملاحظة : فئة المنوال هي الفئة التي يكون لها اكبر تكرار .

مثال :

الجدول الاتي يمثل درجات (100) طالب في مادة علم الاحياء .

فئات الدرجة	40-	50-	60-	70-	80-	90-100
عدد الطلاب	7	18	32	20	15	8

والمطلوب حساب المنوال لهذه الدرجات .

الحل : نعد جدولا يشمل مراكز الفئات والتكرارات وكما يأتي :-

من خلال ملاحظة الجدول اعلاه يمكن ان نستنتج ان فئة المنوال هي التي تتراوح ما بين (60-70) , لانها تضم اكبر تكرار (32) , وبذلك تكون قيمتي ك1 و ك2 والتي تساوي (18 , 20) على التوالي وكما في الجدول الاتي :-

الفئات	التكرار	
40-	7	
50-	18	ك1
60-	32	فئة المنوال
70-	20	ك2
80-	15	
90-100	8	

طول الفئة (ل) $= 50-40 = 10$

$$\text{المنوال} = \text{أ} + \frac{\text{ك}1}{\text{ك}1 + \text{ك}2} \times \text{ل}$$

$$= 60 + \frac{18}{18 + 20} \times 10$$

$$= 60 + \frac{180}{38} = 64,74$$ (بعد التقريب) .

أهمية المنوال

ان لاستخراج قيمة المنوال لمجموعة من البيانات او الدرجات اهمية تتجلى في النقاط الاتية :

1- يستخدم لتلخيص مجموعة كبيرة من البيانات او الدرجات .

2- يستخدم في حساب بقية مقاييس النزعة المركزية كالوسط الحسابي او الوسيط .

العلاقة بين مقاييس النزعة المركزية

هناك علاقة بين مقاييس النزعة المركزية (الوسط الحسابي و الوسيط و المنوال) وكما في العلاقة الاتية :-

المنوال = 3 × الوسيط – 2 × الوسط الحسابي

أي إننا نتمكن من حساب أي مقياس من مقاييس النزعة المركزية من معرفة قيمتي المقياسين الآخرين .

الفصل الرابع

مقاييس التشتت

Measures Of Tendency

الفصل الرابع

مقاييس التشتت

Measures Of Tendency

مقدمة

لا تعد مقاييس النزعة المركزية او التمركز كافية لوصف مجموعة من البيانات وصفاً كاملاً ، أي انها لا تكفي لوصف التوزيع ومعرفة خصائصه بشيء من الدقة والتفصيل , فقد تتساوى بعض العينات في الوسط الحسابي بالرغم من اختلاف توزيع بياناتها حول مركزها (درجة تجانس البيانات). فالمجموعات الاتية ذات وسط حسابي متساو هو(6) ولكنها بلا شك تختلف عن بعضها في صفة الانتشار.

المجموعة أ	4	5	6	7	8
المجموعة ب	6	7	6	5	6
المجموعة ج	6	6	6	6	6
المجموعة د	1	11	6	5	7

ان مقياس النزعة المركزية يمثل مركز البيانات او متوسطها , لكنه لا يبين مدى انتشار أو تشتت البيانات حول هذا المقياس، ولهذا لابد من وجود مقياس آخر مع المقاييس المركزية لقياس درجة الانتشار أو التشتت في داخل هذه البيانات. وهي ما تسمى بمقاييس التشتت والتي تستخدم لمعرفة مدى انتشار او تشتت البيانات وتباينها من حيث التوزيع . ومن أهم مقاييس التشتت المعروفة هي :

1- المدى. **Range**

2- الانحراف المتوسط **Average Deviation.**

3- الانحراف المعياري. **Standard Deviation**

4- التباين. **Variance**

وسنأخذ هذه المقاييس بشيء من التفصيل لاهميتها في بحوثنا :-

أولاً : المدى Range

يعد المدى من ابسط مقاييس التشتت ويعرف بانه الفرق بين أكبر قيمة او درجة وأصغر قيمة او درجة في مجموعة البيانات .

ويمكن حساب المدى من البيانات غير المبوبة عن طريق العلاقة الاتية :-

المدى = أكبر قيمة – أصغر قيمة

مثال :

احسب المدى للبيانات التالية :

12 – 42 – 23 – 39 – 49 – 10 – 40 – 34 – 19

الحل :

أعلى قيمة هي : 49

اقل قيمة هي : 10

اذن المدى = 49 – 10 = 39

اما في حالة البيانات المبوبة فنستخدم العلاقة الآتية لحساب المدى :

المدى = الحد الأعلى للفئة الأخيرة – الحد الأدنى للفئة الأولى

مثال:

احسب المدى للدرجات في الجدول الاتي :

الفئات	5-	10-	15-	20-	25-30
التكرارات	10	15	40	20	15

الحل :

الحد الأعلى للفئة الأخيرة = 30

الحد الأدنى للفئة الأولى = 5

المدى = الحد الأعلى للفئة الأخيرة – الحد الأدنى للفئة الأولى

المدى = 30 – 5 = 25

ثانياً : الانحراف المتوسط Average Deviation

يعرف الانحراف المتوسط بأنه متوسط الانحرافات المطلقة للدرجات او البيانات عن الوسط الحسابي لهذه الدرجات او البيانات .

ونقصد بالانحراف المطلق بانه الفرق بين الدرجة والوسط الحسابي بغض النظر عن الإشارة (نعد الفرق موجبا دائما) .

يمكن حساب الانحراف المتوسط للبيانات غير المبوبة من خلال العلاقة الاتية :-

$$\text{الانحراف المتوسط} = \frac{\text{مجـ} \,|\, \bar{\text{س}} - \text{س} \,|}{\text{ن}}$$

اذ ان :

س = القيمة او الدرجة

$\bar{\text{س}}$ = الوسط الحسابي للقيم او الدرجات

ن = عدد القيم او الدرجات

مثال :

احسب الانحراف المتوسط للبيانات الاتية:-

4 – 9 – 6 – 5 – 8 – 4

الحل :

$$\text{نحسب } \bar{س} = \frac{4 + 8 + 5 + 6 + 9 + 4}{6} = \frac{36}{6} = 6$$

نعد الجدول الأتي :

س	| س – $\bar{س}$ |
4	2
8	2
5	1
6	صفر
9	3
4	2
المجموع	10

$$\text{الانحراف المتوسط} = \frac{10}{6} = 1,67 \text{ (بعد التقريب) .}$$

ويمكن حساب الانحراف المتوسط للبيانات المبوبة من خلال العلاقة الاتية :–

$$\text{الانحراف المتوسط} = \frac{\text{مجـ } (\,|\, سَ - \bar{س} \,|\, \times ك\,)}{\text{مجـ ك}}$$

مثال : احسب الانحراف المتوسط للبيانات في الجدول ادناه :–

الفئات	16–	20–	24–	28–	32–36
التكرار	10	15	40	20	15

الحل :

في البداية نستخرج قيمة الوسط الحسابي لهذه البيانات (المبوبة) كما مر بنا سابقا وكما يأتي :-

الفئة	التكرار	س	س . ك
16-	10	18	180
20-	15	22	330
24-	40	26	1040
28-	20	30	600
32-36	15	34	510
المجموع	100		2150

$$\bar{س} = \frac{\text{مج س . ك}}{\text{مج ك}}$$

$$\bar{س} = \frac{2150}{100} = 21,5$$

ثم نستخرج قيمة الانحراف المتوسط , وكما في الجدول الاتي :-

ف	ك	س	| س ـ $\bar{س}$ |	| س − $\bar{س}$ | × ك
16-	10	18	3,5	35
20-	15	22	0,5	7,5
24-	40	26	4,5	180
28-	20	30	8,5	170
32-36	15	34	12,5	187,5
المجموع	100			580

$$\text{الانحراف المتوسط} = \frac{\text{مجـ} (| \bar{س} - س | \times ك)}{\text{مجـ ك}}$$

$$= \frac{580}{100} = 5.8$$

ثالثا : التباين والانحراف المعياري Variance& Standard Deviation

يرمز للتباين بالرمز $ع^2$ او S^2 , بينما يرمز للانحراف المعياري بالرمز ع او S .

ونرى في كل كتب وأدبيات الإحصاء ان الباحثين يجمعون بين الانحراف المعياري والتباين , ان سبب ذلك الجمع هو العلاقة الوطيدة بين المفهومين , اذ ان :-

الانحراف المعياري = الجذر ألتربيعي للتباين .

أو التباين = مربع الانحراف المعياري

ويعد التباين من اهم مقاييس التشتت وذلك لان انحرافات القيم او الدرجات عن الوسط الحسابي قد تكون قيما سالبة او موجبة او تكون قيمتها مساوية للصفر ، وان المجموع الجبري لهذه القيم يساوي صفر كما ذكرنا سابقا في موضوع الوسط الحسابي . اما في حالة التباين (او الانحراف المعياري) فاننا نعتمد على مربعات هذه الانحرافات , اذ تكون قيمها موجبة او تساوي صفر .

هناك طريقتان معروفتان لحساب التباين وهما :-

1- طريقة الانحرافات :-

اذ نستخدم العلاقة الاتية لحساب التباين :

$$ع^2 = \frac{\text{مج}\,(\overline{س} - س)^2}{ن - 1}$$

ويمكن ان نستخدم هذه العلاقة اذا كان لدينا عدد قليل من البيانات او الدرجات وتكون هذه البيانات صغيرة ، وكذلك اذا كانت قيمة الوسط الحسابي عددا صحيحا او يحوي كسرا عشريا بسيطا .

مثال :-

احسب التباين والانحراف المعياري لمجموعتي الدرجات الآتيتين :-

المجموعة	الدرجات				
أ	8	2	6	4	5
ب	4	3	5	7	6

الحل :-

- بالنسبة للمجموعة أ :

نستخرج الوسط الحسابي لدرجات المجموعة :

$$8+2+6+4+5 = 25$$

$$\text{الوسط الحسابي لدرجات المجموعة أ} = \frac{25}{5} = 5$$

من اجل حساب التباين لدرجات المجموعة (أ) نعد الجدول الاتي:

$(س - \bar{س})^2$	$س - \bar{س}$	س
9	3	8
9	3−	2
1	1	6
1	1−	4
صفر	صفر	5
20		المجموع

$$ع^2 = \frac{مج\ (\bar{س} - س)^2}{ن - 1}$$

$$= \frac{20}{4} = 5$$

اذن الانحراف المعياري لدرجات المجموعة أ =

$$\sqrt{5} = 2,236$$

- بالنسبة للمجموعة ب :

$$\text{الوسط الحسابي لدرجات المجموعة ب} = \frac{6+7+5+3+4}{5} = 5$$

من اجل حساب التباين لدرجات المجموعة (ب) نعد الجدول الاتي:

$(س - \bar{س})^2$	$س - \bar{س}$	س
1	1−	4
4	2−	3
صفر	صفر	5
4	2	7
1	1	6
10		المجموع

الفصل الخامس

معاملات الارتباط

Correlation Coefficients

الفصل الخامس

معاملات الارتباط

Correlation Coefficients

مقدمة

تهدف عدد من البحوث التربوية والنفسية تحليل العلاقة بين متغيرين او اكثر , اذ يهتم الباحث بتحديد كيف وإلى أي مدى يرتبط متغيران أو أكثر، والإحصاءات المستخدمة في التحليلات ثنائية المتغير تشابه متعددة المتغيرات الى حد كبير , فالمنطق متشابه وإن كانت الإحصاءات المستخدمة في دراسة العلاقات متعددة المتغير تتسم بدرجة اكبر من التعقيد.

وعند تحليل العلاقة بين متغيرين يهتم الباحث بالإجابة عن ثلاثة تساؤلات هي :-

1- هل ترتبط هذان المتغيران ؟

2- ما هو اتجاه وشكل الارتباط الموجود ؟

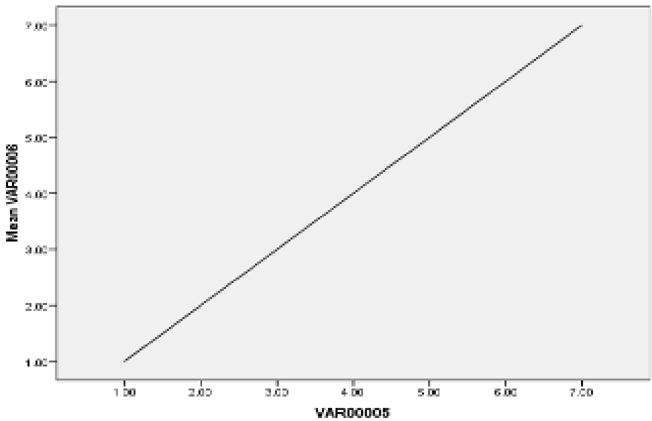
2- هل هناك احتمال أن يكون الارتباط الذي تمت ملاحظته بين حالات العينة هو أحد خصائص مجتمع البحث أم أن هذا الارتباط هو نتاج لصغر حجم العينة التي قد تكون غير ممثلة لمجتمع البحث ؟

يمكن تحديد الارتباط بين متغيرين من خلال استخدام مجموعة من الوسائل الإحصائية تعرف باسم معاملات الارتباط .

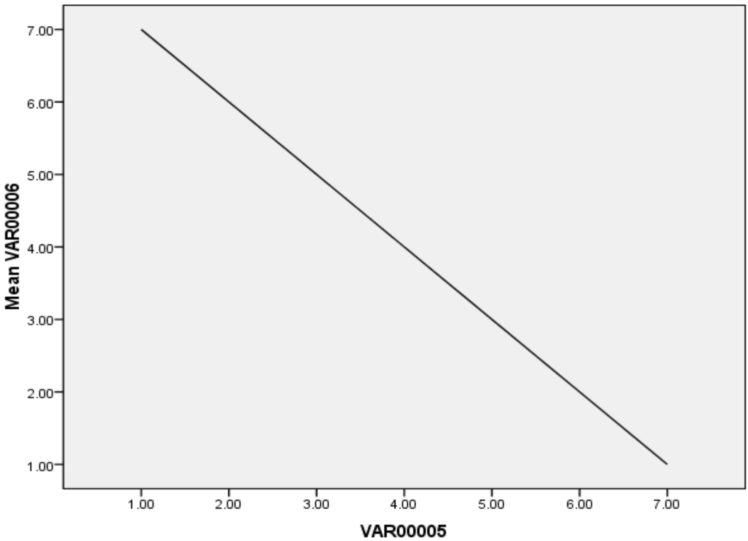
ان معامل الارتباط هو رقم يلخص التحسن في تخمين القيم على متغير واحد لأي حالة على أساس معرفة قيم المتغير الثاني، فكلما ارتفع المعامل قوي الارتباط ،

ومن ثم تحسنت قدرتنا التنبؤية أو التفسيرية. تتراوح قيم معاملات الارتباط بين (+1) و (-1) وكما ياتي :-

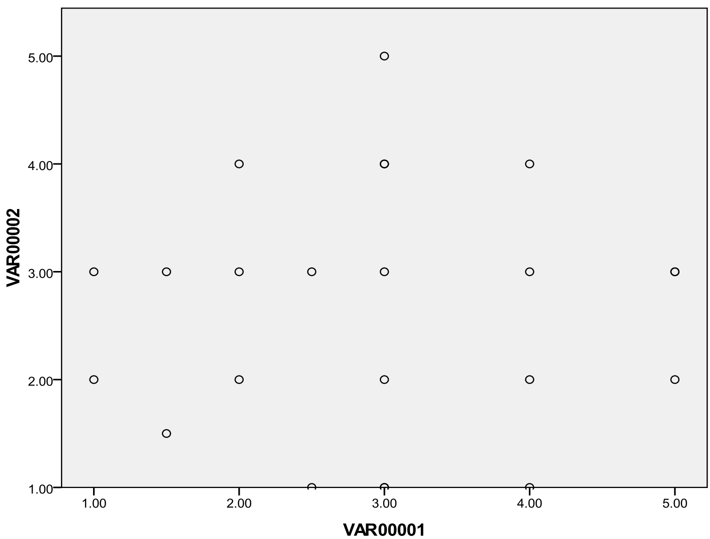
اذا كانت قيمة معامل الارتباط اكبر من الصفر واقل او تساوي (+1) فهذا يدل على وجود علاقة ايجابية او طردية بين المتغيرين , أي ان زيادة قيمة احد المتغيرين ترافقه زيادة في قيمة المتغير الثاني وبالعكس , ويكون شكل العلاقة كما في الشكل الاتي :-



اذا كانت قيمة معامل الارتباط اقل من الصفر واكبر او تساوي (-1) فهذا يدل على وجود علاقة سالبة او عكسية بين المتغيرين , أي ان زيادة قيمة احد المتغيرين ترافقه انخفاض في قيمة المتغير الثاني وبالعكس ، ويكون شكل العلاقة كما في الشكل الاتي :-



اذا كانت قيمة معامل الارتباط تساوي الصفر فهذا يدل على عدم وجود علاقة بين المتغيرين ، ويكون شكل العلاقة كما في الشكل الأتي :-



تفسير قيمة معامل الارتباط

قامت مجموعة من المختصين في مجال الإحصاء بوضع معايير نسبية يمكن ان تستخدم في تفسير قيم معاملات الارتباط ، وكما في الجدول الاتي :-

التفسير	قيمة معامل الارتباط
علاقة طردية تامة	1+
ارتباط طردي قوي	من 0.7 إلى أقل من +1
ارتباط طردي متوسط	من 0.4 إلى أقل من 0.7
ارتباط طردي ضعيف	من صفر إلى أقل من 0.4
الارتباط منعدم	صفر
ارتباط عكسي تام	1-
ارتباط عكسي قوى	من -0.7 إلى أقل من -1
ارتباط عكسي متوسط	من -0.04 إلى أقل من -0.7
ارتباط عكسي ضعيف	من صفر إلى أقل من -.04

انواع معاملات الارتباط :-

هناك أنواع عدة من معاملات الارتباط , وذلك تبعا لنوعي المتغيرين اللذين نهدف الى الكشف عن قيمة واتجاه الارتباط بينهما، اذ ان اختلاف نوع البيانات او الدرجات يستوجب اختلاف الطريقة او العلاقة المستخدمة في حساب معامل الارتباط وسنأخذ اهم انواع معاملات الارتباط وكما ياتي :-

1- معامل ارتباط بيرسون Coefficient Pearson Correlation

يستخدم معامل ارتباط بيرسون لحساب قيمة معامل الارتباط عندما يكون المتغيران المراد قياس الارتباط بينهما متغيرات متصلة او مستمرة , ويشترط تساوي عدد حالات كلاً من المتغيرين .

لحساب قيمة معامل ارتباط بيرسون نستخدم القانون الأتي:

$$ر = \frac{ن\ مجـ (س \times ص) - مجـ\ س \times مجـ\ ص}{\sqrt{[ن\ مجـ\ س^2 - (مجـ\ س)^2] \times [ن\ مجـ\ ص^2 - (مجـ\ ص)^2]}}$$

$$R = \frac{N \Sigma x.y - \Sigma x \cdot \Sigma y}{\sqrt{(N \Sigma x2 - (\Sigma x)2) \cdot (N \Sigma y2 - (\Sigma y)2)}}$$

اذ ان :-

ر : R قيمة معامل ارتباط بيرسون .

س : x قيم المتغير الاول.

ص : y قيم المتغير الثاني .

ن : N عدد قيم احد المتغيرين (س او ص) .

مثال :-

قام معلم بقياس درجات (5) تلاميذ في مادتي الرياضيات والعلوم ، بَين قيمة واتجاه العلاقة بين درجات التلاميذ في مادة الرياضيات ودرجاتهم في مادة العلوم .

درجة مادة العلوم	3	5	9	8	2
درجة مادة الرياضيات	4	6	7	4	3

الحل :-

نرمز لدرجات مادة الرياضيات بـ "س" ودرجات مادة العلوم بـ "ص" (ويجوز العكس) ، ثم نعد الجدول الأتي :

س	ص	س × ص	س^2	ص^2
3	4	12	9	16
5	6	30	25	36
9	7	63	81	49
8	4	32	64	16
2	3	6	4	9
27	24	143	183	126

$$ر = \frac{\text{ن مجـ (س×ص) – مجـ س × مجـ ص}}{\sqrt{[\text{ن مجـ س}^2 - (\text{مجـ س})^2] \times [\text{ن مجـ ص}^2 - (\text{مجـ ص})^2]}}$$

$$ر = \frac{5 \times 143 - 27 \times 24}{\sqrt{[5 \times 183 - (27)^2] \times [5 \times 126 - (24)^2]}}$$

ر = 0,668

من خلال ملاحظة قيمة معامل الارتباط يمكن ان نستنتج ان العلاقة متوسطة . ومن خلال ملاحظة اشارة قيمة معامل الارتباط ، نجد انها اشارة موجبة ، وهذا يدل على ان العلاقة موجبة او طردية .

حساب قيمة معامل ارتباط بيرسون باستخدام الحقيبة الإحصائية :

من اجل حساب قيمة معامل ارتباط بيرسون بين متغيرين باستخدام الحقيبة الإحصائية فإننا نتبع الخطوات الاتية :-

ا- نفتح واجهة الحقيبة .

ب- ندون بيانات او درجات المتغير الأول في العمود الأول .

ج- ندون بيانات او درجات المتغير الثاني في العمود الثاني .

مثال :-

قام أحد الباحثين بتطبيق بحث للكشف عن العلاقة بين الجنس ونتيجة الامتحان الوزاري ، فاخذ عينة من 10 طلاب وطالبات , وكانت نتائجهم كما ياتي :-

الاسم	**احمد**	**حسين**	**رافد**	**رانية**	**سهى**	**علاء**	**جلال**	**سناء**	**محمد**	**سرى**
النتيجة	**راسب**	**ناجح**	**راسب**	**ناجحة**	**ناجحة**	**راسب**	**ناجح**	**راسبة**	**راسب**	**ناجحة**

في البدء ننظم البيانات في مصفوفة , تحوي متغيرين فقط هما الجنس والنتيجة ، اذ نحسب عدد الطلاب (الذكور) الناجحين وعددهم (2) وندون عددهم في الخلية الاولى (أ) , ونحسب عدد الطلاب (الذكور) الراسبين وعددهم (4) وندون عددهم في الخلية الثانية (ب) ، ونحسب عدد الطالبات (الاناث) الناجحات وعددهن (3) وندون عددهن في الخلية الثالثة (ج) ، ونحسب عدد الطالبات (الإناث) الراسبات وعددهن (1) وندون عددهن في الخلية الرابعة (د) ونحسب مجاميع الصفوف والأعمدة وكما في الجدول الأتي :-

النتيجة / الجنس	نجاح	رسوب	المجموع
ذكور	2	4	6
اناث	3	1	4
المجموع	5	5	10

نطبق قانون معامل فاي :-

$$\text{معامل فاي} = \frac{\text{أ} \times \text{د} - \text{ب} \times \text{ج}}{\sqrt{\text{هـ} \times \text{و} \times \text{ز} \times \text{ح}}}$$

$$= \frac{3 * 4 - 1 * 2}{\sqrt{6 * 4 * 5 * 5}}$$

$$= \frac{10 - 2}{24.49} = - 0,41$$

ان الاشارة السالبة لقيمة معامل الارتباط تدل على وجود علاقة عكسية بين الجنس والنتيجة .

اهمية معامل ارتباط فاي في البحوث التربوية والنفسية :

على الرغم من الفائدة المحدودة لهذا المعامل في البحوث التربوية والنفسية ، الا ان له استخداما عندما يرغب الباحث في الكشف عن العلاقة بين متغيرين ثنائيين فقط ، مثل الكشف عن العلاقة بين النوع (ذكور , اناث) والقلق (عالي , واطئ) .

3- معامل التوافق Coefficient Of Contingency

يستخدم معامل التوافق لحساب قيمة معامل الارتباط عندما يكون المتغيران المراد قياس الارتباط بينهم متغيران متقطعان احدهما ثنائي والأخر رباعي فأكثر .

لحساب قيمة معامل التوافق نستخدم القانون التالي:

$$\text{معامل التوافق} = \sqrt{\frac{\text{جـ} - ١}{\text{جـ}}}$$

اذ ان

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

مثال :

قام أحد الباحثين بإجراء بحث ارتباطي عن علاقة السلوك العدواني بمشاهدة أفلام العنف ، وقد حصل على النتائج الآتية :-

وهذا يدل على ان العلاقة طردية متوسطة .

أهمية معامل التوافق في البحوث التربوية والنفسية :

ان لهذا المعامل أيضا أهمية محدودة جدا في البحوث التربوية والنفسية وسبب ذلك هو قلة المتغيرات المتقطعة فيها، اذ ان معظم المتغيرات في البحوث التربوية والنفسية هي من نوع المتغيرات المتصلة او المستمرة .

4– معامل ارتباط الرتب لسبيرمان Spearman Rank Correlation Coefficient

يستخدم معامل ارتباط الرتب لسبيرمان لحساب قيمة معامل الارتباط عندما يكون المتغيران المراد قياس الارتباط بينهما متغيرين رتبيين , ويشترط تساوي عدد حالات كل من المتغيرين أيضاً ونستخدم القانون التالي لحساب قيمة معامل ارتباط الرتب لسبيرمان:

$$\text{ر} = 1 - \frac{6 \text{ مجـ ف}^2}{\text{ن (ن}^2 - 1)}$$

$$r = 1 - \frac{6 * \Sigma d^2}{n(n^2 - 1)}$$

اذ ان :–

ر : r معامل ارتباط الرتب لسبيرمان

ف : d رتب المتغير الأول – رتب المتغير الثاني

ن : n عدد الحالات

مثال :- الجدول الاتي يوضح درجات مجموعة من الطلاب في اختبار معين تم إجراؤه على نفس الطلاب مرتين متتاليتين , والمطلوب حساب قيمة معامل ارتباط الرتب لسبيرمان بين درجات الاختبارين ؟

درجة الاختبار الأول	3	5	9	8	2
درجة الاختبار الثاني	4	6	7	4	3

الحل :-

نفترض أن درجات الاختبار الأول هى "س" ودرجات الاختبار الثاني هي "ص" و نقوم بترتيب قيم س تصاعديا , ويتم تحويل الدرجات الى رتب متسلسلة لان المطلوب استخراج معامل ارتباط سبيرمان , مع ملاحظة أنه إذا تساوى عددان أو أكثر في القيمة يأخذ كل منهم متوسط رتبهم . وكما في الجدول الاتي :

س	ص	رتب س	رتب ص	ف	$ف^2$
2	3	1	1	0	0
3	4	2	2.5	0.5−	0.25
5	6	3	4	1−	1
8	4	4	2.5	1.5	2.25
9	7	5	5	0	0
المجموع					3.5

$$ر = 1 - \frac{6 \text{ مجـ } ف^2}{ن\,(ن^2 - 1)}$$

$$ر = 1 - \frac{6 \times 3.5}{5\,(25 - 1)}$$

$$ر = 1 - \frac{21}{} = 1 - 0.175 = 0,825$$

24×5

ويمكن ان نستنتج ان الارتباط هو طردي وقوي .

اهمية معامل ارتباط الرتب في البحوث التربوية والنفسية :

ان لهذا المعامل ايضا استخدامات محدودة جدا في البحوث التربوية والنفسية وسبب ذلك هو قلة المتغيرات الرتبية فيها ونادرا ما يقوم الباحث بتحديد رتب لعينة البحث حسب المتغير المدروس اذ ان معظم المتغيرات في البحوث التربوية والنفسية هي من نوع المتغيرات المتصلة او المستمرة .

5- معامل الارتباط الثنائي النقطي Correlation Coefficient Point Biserial

يستخدم معامل الارتباط هذا اذا كان لدينا متغيرين احدهما متصل او مستمر، والاخر متقطع ثنائي بشكل طبيعي مثل الجنس

ويحسب معامل الارتباط الثنائي النقطي من العلاقة :-

$$ر = \sqrt{\frac{\bar{س}1 - \bar{س}2}{ع ك} * نس1 * نس2}$$

$$r = \sqrt{\frac{X1-X2}{S} * P . Q}$$

اذ ان :-

ر = r معامل الارتباط الثنائي النقطي

$\bar{س}1$ = X1 الوسط الحسابي لدرجات المجموعة الاولى