المحاضرة (1)– 2 ساعة

المصطلحات والمقاييس الإحصائية المستخدمة في تصميم التجارب

المحتويات

1- أنواع البيانات، وطرق جمعها.
2- الرموز المستخدمة في تصميم التجارب
3- مقاييس النزعة المركزية والتشتت

الأهداف

1- اكساب الطالب معرفة الأسس والمفاهيم الأساسية للتجربة.
2- اكساب الطالب مهارة تحديد أنواع البيانات الاحصائية الزراعية
3- حساب المقاييس الاحصائية الوصفية.

آلية التنفيذ

1- محاضرة
2- نشاط منزلي

1- أنواع البيانات ومعايير قياسها

علم الإحصاء هو العلم الذي يهتم بطرق جمع البيانات، وتبويبها، وتلخيصها بشكل يمكن الاستفادة منها في وصف البيانات وتحليلها للوصول إلى قرارات سليمة في ظل ظروف عدم التأكد.من هذا التعريف يلاحظ أن أحد وظائف هذا العلم الاهتمام بجمع البيانات **Data**. ويعتبر تحديد نوع البيانات، ومعايير قياسها من أهم الأسس التي تحدد نوع التحليل الإحصائي المستخدم، و هناك نوعين للبيانات التي تواجه الباحث في كثير من المجالات التطبيقية هما:

1- البيانات الوصفية Qualitative Data
2- البيانات الكمية Quantitative Data

أولا: البيانات الوصفيةQualitative Data

هي بيانات غير رقمية تعبر عن مجموعاتمتنافية تمثل مجال الظاهرة محل الدراسة، وقد تأخذ في بعض التطبيقات بيانات رقمية بينية ومحددة تمثل مستويات الظاهرة أو بيانات منظمة في شكل فئات رقمية تعبر عن مستويات الظاهرة ، ومن ثم تقاس البيانات الوصفية بمعيارين هما:

أ- **بيانات وصفية مقاسة بمعيار اسمي Nominal Scale:** وهي بيانات غير رقمية تتكون من مجموعات متنافية تمثل مجال الظاهرة، كل مجموعة لها خصائص تميزها عن المجموعة الأخرى، كما أن هذه المجموعات لا يمكن المفاضلة بينها، ومن الأمثلة على ذلك:

- النوع (ذكر، أنثى): متغير وصفي يأخذ مجموعتين.
- طرق التسميد (فوسفات، نترات، عضوي): متغير وصفي مكون من ثلاث مجموعات.
- الجنسية (سعودي، غير سعودي) متغير وصفي مكون من مجموعتين وتقاس بياناته بمعيار اسمي.

وعند معالجة هذه البيانات إحصائيا، يتم معالجتها من خلال تكويد مجموعات المتغير بأرقام، فمثلا الجنسية يمكن إعطاء الجنسية "سعودي" الكود (1)، والجنسية "غير سعودي" الكود (0)

ب- **بيانات وصفية مقاسة بمعيار ترتيبي Ordinal Scales:** وتتكون من مستويات، أو أرقام بينية محددة تمثل مستويات المتغير، أو بيانات منظمة في شكل فئات رقمية تعبر عن مستويا المتغير. وفي جميع الأحوال يمكن تميز هذه المستويات عن بعضها البعض بترتيبها تصاعديا أو تنازليا. ومن الأمثلة على ذلك:

- تقدير الطالب($D, D^+, C, C^+, B, B^+, A , A^+$)**:** متغير وصفي تقاس بياناته بمعيار ترتيبي وله ثمان مستويات**.**
- المستوى التعليمي (أمي ، يقرأ ويكتب ، ابتدائية ، متوسطة ، ثانوية ، جامعية أو أعلى) متغير وصفي تقاس بياناته بمعيار ترتيبي وله ستة مستويات**.**
- تحديد تركيز خلات الصوديوم(*0%، 5%، 10%*)لحفظ صدور الدجاج من البكتريا: متغير وصفي يقاس بياناته بمعيار ترتيبي وله ثلاث مستويات.
- فئات الدخل العائلي بالألف ريال في الشهر (*5> ، 5-10 ، 10-15 ، 15-20، 20<*)متغير وصفي يقاس بياناته بمعيار ترتيبي وله خمس مستويات.

ثانيا: البيانات الكمية

هي بيانات يعبر عنها بأرقام عددية تمثل القيمة الفعلية للظاهرة، وتنقسم إلى

قسمين هما:

بيانات فترة Interval Data: وهي بيانات رقمية،تقاس القيمة بمقدار بعدها عن الصفر، أي أن للصفر دلالة على وجود الظاهرة، ومن أمثلة ذلك:

- درجة الحرارة: متغير كمي تقاس بياناته بمعيار بعدي، حيث أن درجة الحرارة "***0***" ليس معناه انعدام الظاهرة، ولكنه يدل على وجود الظاهرة.
- درجة الطالب في الاختبار: متغير كمي يقاس بياناته بمعيار بعدي، حيث حصول الطالب على الدرجة "**0**" لا يعني انعدم مستوى الطالب.

بيانات نسبية Ratio Data: هي متغيرات كمية، تدل القيمة "**0**" على عدم وجود الظاهرة ومن الأمثلة على ذلك:

- إنتاجية الفدان بالطن/هكتار.
- المساحة المنزرعة بالأعلاف بالدونم.
- كمية الألبان التي تنتجها البقرة في اليوم.
- عدد مرات استخدام المزارع لنوع معين من السماد.

ويلاحظ أن بيانات الفترة لا يمكن إخضاعها للعمليات الحسابية مثل عملياتالضرب والقسمة، بينما يمكن فعل ذلك مع البيانات النسبية.

2- أنواع المتغيرات الكمية

وبصرف النظر عن معيار قياس البيانات الكمية، فإن المتغيرات الكمية يمكن، تقسيمها إلى قسمين هما:

- **متغيرات كمية منفصلة**Discrete variable: وهي التي تأخذ قيما متباينة، ومتباعدة، كما تخضع لعملية العد، ومن أمثلة ذلك : عدد المرضي المصابين بمرض ارتفاع الضغط ، عدد أفراد الأسرة ، عدد مرات زيارة المريض للطبيب المختص كل شهر، وهكذا.....
- **متغيرات كمية مستمرة**Continuous Variable: وهي التي تأخذ قيما متصلةداخل المدى أو المجال المعرف لها ، ومن ثم يوجد عدد لانهائي من القيم داخل مدي ينتمي لذا المجال. ومن الأمثلة على ذلك : الأوزان، والأطوال، ضغط الدم، كمية الإنتاج من

الحليب اليومي، السعرات الحرارية،.... وهكذا.

3- مصادر ووسائل جمع البيانات

يعتبر جمع البيانات أحد المراحل الأساسية في تصميم البحث، وهذه المرحلة من المراحل الهامة جدا التي يقوم بها الباحث، إذ يجب على الباحث بعد تحديد متغيراته تحديد مصادر بياناته، وطريقة جمعها وكذلك الوسائل المستخدمة لذلك.

أولا: مصادر جمع البيانات

وتتحدد في مصدرين هما :

- المصادر الأولية.
- المصادر الثانوية.

- **المصادر الأولية:**

وهي المصادر التي نحصل منها على البيانات بشكل مباشر، حيث يقوم الباحث نفسه بجمع البيانات من المفردة محل البحث مباشرة، فعندما يهتم الباحث بجمع بيانات عن أسرة المزارع ، يقوم بإجراء مقابلة مع رب الأسرة، ويتم الحصول منه مباشرة على بيانات خاصة بأسرته، مثل بيانات المنطقة التابع لها، والحي الذي يسكن فيه، وعدد أفراد الأسرة، ومساحة الحيازة الزراعية، والدخل المزرعي الذي تحققه المزرعة، والمستوى التعليمي للمزارع، ... وهكذا.

ويتميز هذا النوع من المصادر بالدقة والثقة في البيانات، لأن الباحث هو الذي يقوم بنفسه بجمع البيانات من المفردة محل البحث مباشرة، ولكن يعاب عليها أنها ذات كلفة عالية من الناحية الماديةوتحتاج إلى وقت ومجهود كبير.

- **المصادر الثانوية:**

وهي المصادر التي نحصل منها على البيانات بشكل غير مباشر، بمعنى آخر يتم الحصول عليها بواسطة أشخاص آخرين، أو أجهزة، وهيئات رسمية متخصصة، مثل نشرات وزارة الزراعة، ونشرات مصلحة الإحصاء، ونشرات منظمة الأغذية (الفاو)....وهكذا.

ومن مزايا هذا النوع من المصادر، توفير الوقت والجهد والمال، إلا أن درجة ثقة الباحث فيها ليست بنفس الدرجة في حالة المصادر الأولية.

ثانيا: الأدوات المستخدمة في جمع البيانات (وسائل جمع البيانات)

يمكن تصنيف وسائل جمع البيانات الإحصائية ضمن الأدوات التالية:

1) المقابلة.
2) المراسلة.
3) استخدام وسائل الاتصالات الحديثة.
4) الملاحظة.

1) المقابلة (أو الاتصال المباشر):

تعتبر المقابلة من أهم الوسائل الشائعة لجمع البيانات وهي عبارة عن محادثة تتم بين الباحث والمبحوثين بغرض تحقيق هدف الدراسة ، حيث يقوم الباحث بطرح الأسئلة المكتوبة في استمارة على المبحوث ومن ثم تدوين إجابة المبحوث على تلك الاستمارة. ولا بد من إعداد جيد للمقابلة عن طريق تحديد أهداف المقابلة بشكل واضح وتحديد الأفراد الذين سيقابلهم الباحث بالإضافة إلى تحديد الأسئلة والترتيب المسبق للمقابلة والظهور بمظهر مناسب وتهيئة الجو الملائم مما يدعو إلى ارتياح المبحوث وإزالة أي توتر لديه. ويجب تنفيذ المقابلة وفق الخطة المحددة من حيث الوصول في الوقت المحدد لإجراء المقابلة واللباقة في الدخول إلى المبحوث وتدوين الإجابات بخط واضح والانصراف بلباقة مع تقديم الشكر على تعاون المبحوث.

ومن أهم مزايا المقابلة:

- الحصول على بيانات دقيقة.
- ضمان الحصول على إجابات عن كل الأسئلة.
- إمكانية توضيح الأسئلة للمبحوثين في حالة وجود أسئلة غير مفهومة.

أهم عيوب المقابلة:

- تحتاج إلى وقت ونفقات مالية وإمكانات بشرية ضخمة.
- تتأثر بالحالة النفسية لكل من الباحث والمبحوث.
- تسبب في بعض الأحيان حرجاً للمبحوثين خاصة إذا كانت الأسئلة شخصية.

مجالات استخدام المقابلة:

تستخدم كثيراً في البحوث الميدانية في كثير من الدول النامية بسبب تدني المستوى الثقافي والوعي الإحصائي ،وبالأخص في حالة عدم إلمام المبحوثين بالقراءة والكتابة واحتياجهم إلى تفسير وتوضيح لأسئلة الباحث.

2) المراسلة (أو البريد)

يتم إرسال الاستبانات إلى المبحوثين إما بالبريد أو تسلم لهم باليد حيث يقومون بقراءة الأسئلة والإجابة عليها بأنفسهم دون وجود الباحث ومن ثم إعادتها إلى الباحث مرة أخرى بعد تعبئتها.

أهم مزايا المراسلة:

- توفير الوقت إذا كان عدد المبحوثين كبيراً ومن مناطق متباعدة.
- تتطلب إمكانات مالية وبشرية اقل من طريقة المقابلة.
- لا تسبب حرجاً للمبحوث حيث تتم الإجابة على الأسئلة بكل موضوعية.

أهم عيوب المراسلة:

- عدم الإجابة على بعض الأسئلة لعدم وضوحها أو احتياجها إلى تفسير.
- تأخر وصول بعض الإجابات.
- انخفاض نسبة المردود بسبب عدم وضوح عنوان المبحوث أو نسيان المراسلة أو إهمالها.

مجالات استخدام المراسلة:

- تستخدم كثيراً في بعض البحوث التي تنفذها المؤسسات الحكومية.
- تستخدم عندما يكون المستوى الثقافي والوعي الإحصائي مرتفعاً.

3) استخدام وسائل الاتصالات الحديثة

في هذه الطريقة يتم استخدام وسائل الاتصالات (هاتف وفاكس وتلكس وإنترنت) وذلك للحصول على إجابات سريعة مثل استطلاعات الرأي العام، وتعتبر هذه الطريقة من أسرع وسائل جمع البيانات.

أهم مزايا استخدام وسائل الاتصالات الحديثة:

- تعد أسرع الطرق وأسهلها.
- انخفاض تكاليفها مقارنة مع غيرها من الوسائل.
- يبتعد الباحث عن المخاطر خاصة إذا كان المبحوثين من الفئات الخطرة مثل

المجرمين ومتعاطي المخدرات .

أهم عيوب استخدام وسائل الاتصالات الحديثة:

- تعتمد على مدى توافر هذه الوسائل لدى المبحوثين.
- عدم إمكانية التعرف على ملامح المبحوث أثناء إجابته على الأسئلة.

مجالات استخدام وسائل الاتصالات الحديثة:

تستخدم هذه الطريقة في بعض البحوث التي تقوم بها المؤسسات الحكومية و المؤسسات الكبيرة في القطاع الخاص التي يتوافر لديها هذه الأجهزة.

4) الملاحظة أو المشاهدة

هي عملية مشاهدة ومراقبة سلوك ظاهرة ما أو مشكلة ما وذلك عن طريق اتصاله مباشرة بالأشخاص أو الأشياء التي يدرسها أو من خلال اتصاله بالسجلات التي أعدها الآخرون، مثلاً تدوين نوع مادة بناء المسكن دون الحاجة إلى طرح الأسئلة، تدوين كميات سقوط الأمطار ،ودراسة تصرفات الأفراد أثناء مشاهدتهم لمباراة لكرة القدم.

مزايا وسيلة الملاحظة:

- عدم إحراج المدلي بالبيانات .
- إمكانية استخدامها في حالات معينة لا يستطيع فيها المدلي بالبيانات إعطاء بيانات دقيقة .
- لا تتطلب عدداً كبيراً من الأفراد ليقوموا بملاحظة الظاهرة.

عيوب وسيلة الملاحظة:

- عدم الدقة في بعض الأحيان نتيجة التخمين الخاطئ للباحث.
- قد يضطر الباحث إلى الانتظار فترة طويلة لملاحظة وقوع ظاهرة معينة ، مما يترتب عليه إنفاق وقت وجهد وكلفة مرتفعة من الباحث وخاصة في ملاحظة الظواهر الكونية كالزلازل وسلوك الحيوانات .

مجالات استخدام وسيلة الملاحظة:

تستخدم هذه الطريقة في المجالات العلمية كملاحظة الظواهر الطبيعية بالإضافة إلي المجالات الاجتماعية والإدارية التي تستخدم فيها الملاحظة لدراسة سلوك الناس حول مواقف معينة.

4- تعريف الاستبيان

الاستبيان هي وسيلة لجمع البيانات قوامها الاعتماد على مجموعة من الاسئلة ترسلامابطريق البريد لمجموعة من الافراد، أو تنشر على صفحات الجرائد والمجلاتأو علىالتلفزيون او عن طريق الاذاعة او عن طريق النت. ليجيب عليها الأفرادويقوموابارسالها إلى الهيئة المشرفة على البحث أو تسلم للمبحوثين ليقوموابملئها ثم يتولىالباحث جمعها.

مزايا الاستبيان

- يفيد الباحث في حالأنأفراد البحث منتشرين في أماكن متفرقة ويصعب الاتصال بهمشخصياً.
- يتميز بأنهقليل التكاليفوالجهد.
- يعطيالاستبيان للمبحوثين فرصة كافية للاجابة علىالاسئلةبدقة.
- يوفر الاستبيان التقنين أكثر من اي وسيلةأخرى.
- يسمحبالحصول على بيانات حساسةومحرجة.

5- الرموز المستخدمة في تصميم وتحليل التجارب

مجموع قيم المشاهدات $(x_1, x_2, ..., x_n)$ يكتب:

$$\sum_{i=1}^{n} x_i = x_1 + x_2 + ... + x_n$$

مجموع مربعات القيم: $(x_1^2, x_2^2, ..., x_n^2)$ يكتب:

$$\sum_{i=1}^{n} x_i^2 = x_1^2 + x_2^2 + ... + x_n^2$$

مع ملاحظة أن: $(\sum_{i-1}^{n} x_i)^2 \neq \sum_{i=1}^{n} x_i^2$

مجموع قيمالمشاهدات $(y_{i1}, y_{i2}, ..., y_{ir})$ يكتب:

$$\sum_{j-1}^{r} y_{ij} = (y_{i1} + y_{i2} + ... + y_{ir}) = y_{i.}$$

التعريب	الإسم	الرمز
ألفا	Alpha	α
بيتا	Beta	β
إبسلون	Epsilon	ε
جاما	Gamma	γ
لامدا	Lambdaμ	λ
ميو	Mu	μ
رو	Rho	ρ
سيجما	Sigma	σ
مربع سيجما	Sigma	σ^2
الجذر التربيعي	Square	√
مجموع	Summation	∑
تاو	Tau	τ
ثيتا	Theta	θ

6- مقاييس النزعة المركزية والتشتت

أولا: مقاييس النزعة المركزية:

من أهم هذه المقاييس المتوسط الحسابي الوسيط والمتوسط الهندسي.

- **المتوسط الحسابي** Arithmetic Mean

وهو من أهم مقاييس النزعة المركزية ، وأكثرها استخداما في النواحي التطبيقية ، ويعرف 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 المتوسط الحسابي لهذه القيم ، ونرمز له بالرمز $\bar{x}$ يحسب بالمعادلة التالية :

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

ويدل الرمز Σ Summation على المجموع .

تطبيق (1-1)

فيما يلي بيانات قيمة المبيعات بالألف ريال التي حققها 8من مندوبي المبيعات خلال شهر في إحدى الشركات.

المندوب	1	2	3	4	5	6	7	8
قيمة المبيعات	40	36	40	35	37	42	32	34

والمطلوبحساب متوسط قيمة المبيعات التي يحققها المندوب في هذه الشركة خلال الشهر.

مناقشة التطبيق

لإيجاد المتوسط الحسابيلقيمة المبيعات التي يحققها المندوب في هذه الشركة خلال الشهر، تطبق معادلة المتوسط الحسابي أعلاه آنفا.

$$\bar{x} = \frac{x_1 + x_2 + \ldots + x_n}{n} = \frac{34 + 32 + 42 + 37 + 35 + 40 + 36 + 40}{8} = \frac{296}{8} = 37$$

أي أنالمتوسط الحسابي لقيمة المبيعات التي يحققها المندوب 37ألف ريال شهريا.

- **مزايا الوسط الحسابي:** أنه سهل الحساب، ويأخذ في الاعتبار كل القيم، وأنه أكثر مقاييس النزعة المركزية استخداما وفهما .
- **عيوب الوسط الحسابي:** أنه يتأثر بالقيم الشاذة والمتطرفة، ويصعب حسابه في حالة

البيانات الوصفية.

- **الوسيط Median**

هو أحد مقاييس النزعة المركزية، والذي يأخذ في الاعتبار رتب القيم ، ويعرف الوسيط بأنه القيمة التي يقل عنها نصف عدد القيم $(n/2)$ ، ويزيد عنها النصف الآخر $(n/2)$، أي أن 50% من القيم أقل منه، 50% من القيم أعلى منه. وفيما يلي خطوات حساب الوسيط .

- ترتب القيم تصاعديا .
- تحديد رتبة الوسيط، وهي : رتبة الوسيط $= \left(\frac{n+1}{2}\right)$
- إذا كان عدد القيم (n) فردي فإن الوسيط هو:

$$\text{الوسيط} = \text{القيمة رقم } \left(\frac{n+1}{2}\right)$$

- إذا كان عدد القيم (n)زوجي، فإن الوسيط يقع بين القيمة رقم $(n/2)$ ، والقيمة رقم $((n/2)+1)$ ، ومن ثم يحسب الوسيط بتطبيق المعادلة التالي:

$$\text{الوسيط} = \{\text{القيمة رقم } \left(\frac{n}{2}\right) + \text{القيمة رقم } \left(\frac{n}{2}+1\right)\} \div 2$$

تطبيق (3-3)

تم تقسيم قطعة أرض زراعية إلى17وحدة تجريبية متشابهة ، وتم زراعتها بمحصول القمح ، وتم استخدام نوعين من التسميد هما : النوع(A) وجرب على7وحدات تجريبية ، والنوع(B)وجرب على10وحدات تجريبية، وبعد انتهاء الموسم الزراعي ، تم تسجيل إنتاجية الوحدة بالطن / هكتار ، وكانت على النحو التالي :

النوع(A)	1.2	2.75	3.25	2	3	2.3	1.5			
النوع (B)	4.5	1.8	3.5	3.75	2	2.5	1.5	4	2.5	3

والمطلوب حساب وسيط الإنتاج لكل نوع من السماد المستخدم، ثم قارن بينها.

مناقشة التطبيق

أولا : حساب وسيط الإنتاج للنوع الأول(A)

- ترتيب القيم تصاعديا:

قيمة الوسيط

الإنتاج	1.2	1.5	2	2.3	2.75	3	3.25
الرتبة	1	2	2	4	5	6	7

رتبة الوسيط

- عدد القيم فردى $(n = 7)$
- إذا رتبة الوسيط هي: $((n+1)/2 = (7+1)/2 = 4)$.
- ويكون الوسيط هو القيمة رقم 4 ، أي أن وسيط الإنتاج للنوع(A) هو:

$Med_{a} = 2.3$ طن / هكتار

ثانيا : حساب وسيط الإنتاج للنوع الثاني(B):

- ترتيب القيم تصاعديا .

قيمة الوسيط $= \frac{2.5+3}{2} = 2.75$

الإنتاج	1.5	1.8	2	2.5	2.5	3	3.5	3.75	4	4.5
الرتبة	1	2	3	4	5	6	7	8	9	10

5.5

رتبة الوسيط

- عدد القيم زوجي $(n = 10)$ إذا
- رتبة الوسيط هي : $((n+1)/2 = (10+1)/2 = 5.5)$.
- الوسيط = الوسط الحسابي للقيمتين الواقعتين في المنتصف (رقم 5 ، 6).

$Med_{b} = \frac{2.5+3}{2} = 2.75$ طن / هكتار

وبمقارنة النوعين من السماد ، نجد أن وسيط إنتاجية النوع(A)أقل من وسيط إنتاجية النوع(B)، أي أن : $Med_{B} > Med_{A}$.

- **مزايا الوسيط:** لا يتأثر بالقيم الشاذة أو المتطرفة ، كما أنه سهل في الحساب، .

مجموع قيم الانحرافات المطلقة عن الوسيط أقل من مجموع الانحرافات المطلقة عن أي قيم أخرى . أي أن : $\sum |x - Med| \le \sum |x - a| \ , \ a \neq Med$

- **عيوب الوسيط:** أنه لا يأخذ في الإعتبار كل القيم عند حسابه فهو يعتمد على قيمة أو قيمتين فقط، كما يصعب حسابه في حالة البيانات الوصفية المقاسة بمعيار اسميnominal

ثانيا: مقاييس التشتت

تستخدم مقاييس التشتت في التعرف على مدى تجانس القيم التي أمكن جمعها حول ظاهرة معينة، ويعتمد عليها في مقارنة بيانات ظاهرتين لهما نفس القياس خاصة إذا كانت متوسطي الظاهرتين متساوي، كما تستخدم مقاييس الانتشار في وصف مدى انتشار البيانات حول مقياس المتوسط ويمكن الاعتماد عليها ليس فقط في مقارنة ظاهرتين لهما نفس القياس، يمكن استخدامها في مقارنة ظاهرتين مختلفتين من حيث القياس.

ومن أهممقاييس التشتت التي تصف مدى تجانس البيانات المدى Rang، ومتوسط الانحرافات Mean of Deviations، والتباين Variance والانحراف المعياري StandardDeviation، وفيما يلي عرض بعض هذه المقاييس:

- **المدى** Range

هو أبسط مقاييس التشتت، ويحسب بتطبيق المعادلة التالية.

المدى = أكبر قيمة – أدنى قيمة

- **ثانيا: متوسط الانحرافات**Mean of Deviations

وهو عبارة عن متوسط الانحرافات المطلقة عن المتوسط الحسابي، ومن ثم يحسب بتطبيق المعادلة التالية:

$$MD = \sum \left| X - \overline{X} \right| / n$$

- **التباين** Variance

من أهم مقاييس التشتت استخداما، وهو عبارة متوسط مربعات انحرافات القيم عن وسطها الحسابي، ويحسب بتطبيق المعادلة التالية:

$$S^2 = \sum (X \quad \overline{X})^2 / (n \quad 1)$$

- **الانحراف المعياري** Standard Deviation

من أكثر مقاييس التشتت استخداما، وهو عبارة عن الجذر التربيعي الموجب للتباين، وهو أوقع استخداما من التباين، ويحسب بالمعادلة التالية:

$$S = \sqrt{S^2} = \sqrt{\frac{\sum (X - \overline{X})^2}{n - 1}}$$

تطبيق (3-1)

فيما يلي بيانات قيمة صادرات المملكة إلى دول مجلس التعاون الخليجى بالمليون ريال خلا الفترة من 1999 إلى 2003م

السنوات	1999	2000	2001	2002	2003
قيمة الصادرات بالمليون	13310	15993	14761	16734	23310

والمطلوب حساب المقاييس المختلفة للتشتت المذكورة أعلاه.

مناقشة التطبيق

يتم تطبيق المعادلات وفقا للجدول التالي.

السنوات	قيمة الصادرات بالمليون X	$(X - \bar{X})$ $(X - 16821.6)$	$\lvert X - \bar{X} \rvert$ $\lvert X - 16821.6 \rvert$	$(X - \bar{X})^2$ $(X - 16821.6)^2$
1999	13310	-3511.6	3511.6	12331334.56
2000	15993	-828.6	828.6	686577.96
2001	14761	-2060.6	2060.6	4246072.36
2002	16734	-87.6	87.6	7673.76
2003	23310	6488.4	6488.4	42099334.56
الإجمالي	**84108**	**0.00**	**12976.80**	**59370993.20**

المدى Range

$$\text{المدى} = \text{أكبر قيمة} - \text{أدنى قيمة} = 23310 - 13310 = 10000$$

متوسط الانحرافات Mean of Deviations

$$MD = \frac{\sum \left| X - \bar{X} \right|}{n}$$

$$\bar{X} = \frac{84108}{5} = 16821.6$$

$$\sum \left| X - \bar{X} \right| = 12976.80$$

$$MD = \frac{\sum \left| X - \bar{X} \right|}{n} = \frac{12976.80}{5} = 2595.36$$

التباين Variance

$$S^2 = \frac{\sum (X - \overline{X})^2}{n - 1}$$

$$\overline{X} = 16821.6$$

$$\sum (X - \overline{X})^2 = 59370993.2 \quad 0$$

$$S^2 = \frac{\sum (X - \overline{X})^2}{n - 1} = \frac{59370993.2 \quad 0}{4} = 14842748.3$$

الانحراف المعياري Standard Deviation

$$S = \sqrt{S^2} = \sqrt{14842748.3} = 3852.62875 \quad 2$$

ثالثا: مقاييس الانتشار

من أهم هذه المقاييس معامل الاختلاف النسبي، والدرجة المعيارية، وفيما يلي عرض لهذه المقاييس.

- **معامل الاختلاف النسبي** Coefficient of Variation (CV)

وهو عبارة عن الانحراف المعياري كنسبة مئوية من المتوسط الحسابي، وكلما كان هذا المعامل صغيرا كلما دل ذلك على انتشار البيانات في مدى ضيق ويستدل منه على أن البيانات أكثر تجانسا، ويحسب هذا المعامل بتطبيق المعادلة التالية.

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

- **الدرجة المعيارية** Standardized Value (Z)

ويقيس مقدار انخفاض أو زيادة القيم عن وسطها الحسابي بعدد من وحدات انحراف معياري، ويطبق باستخدام التحويلة التالية.

$$Z = \frac{Value \quad x - \overline{X}}{S}$$

تطبيق (4-1)

في التطبيق (3-1) السابق احسب معامل الاختلاف النسبي، وكذلك الدرجات المعيارية مع التفسير.

مناقشة التطبيق

- **معامل الاختلاف النسبي** Coefficient of Variation (CV)

$$CV = \frac{S}{\overline{X}} \times 100 = \frac{3852.62875^{\ 2}}{16821.6} \times 100 = 22.9$$

ويدل ذلك على أن القيم متجانسة إلى حد ما.

- **الدرجة المعيارية** Standardized Value (Z)

$$Z = \frac{Value\ \ x - \overline{X}}{S} = \frac{Value\ \ x - 16821.6}{3852.62875^{\ 2}}$$

ويبين الجدول التالي الدرجات المعيارية لقيم الصادرات

السنوات	قيمة الصادرات بالمليون	الدرجة المعيارية
	X	$Z = (X - 16821.6) / 3852.628752$
1999	13310	-0.91148
2000	15993	-0.21507
2001	14761	-0.53486
2002	16734	-0.02274
2003	23310	1.684149

يلاحظ أن قيمة صادرات عام 2003م (1.68) زاد عن قيمة متوسط صادرات الفترة بـ (1.68) انحراف معياري.