

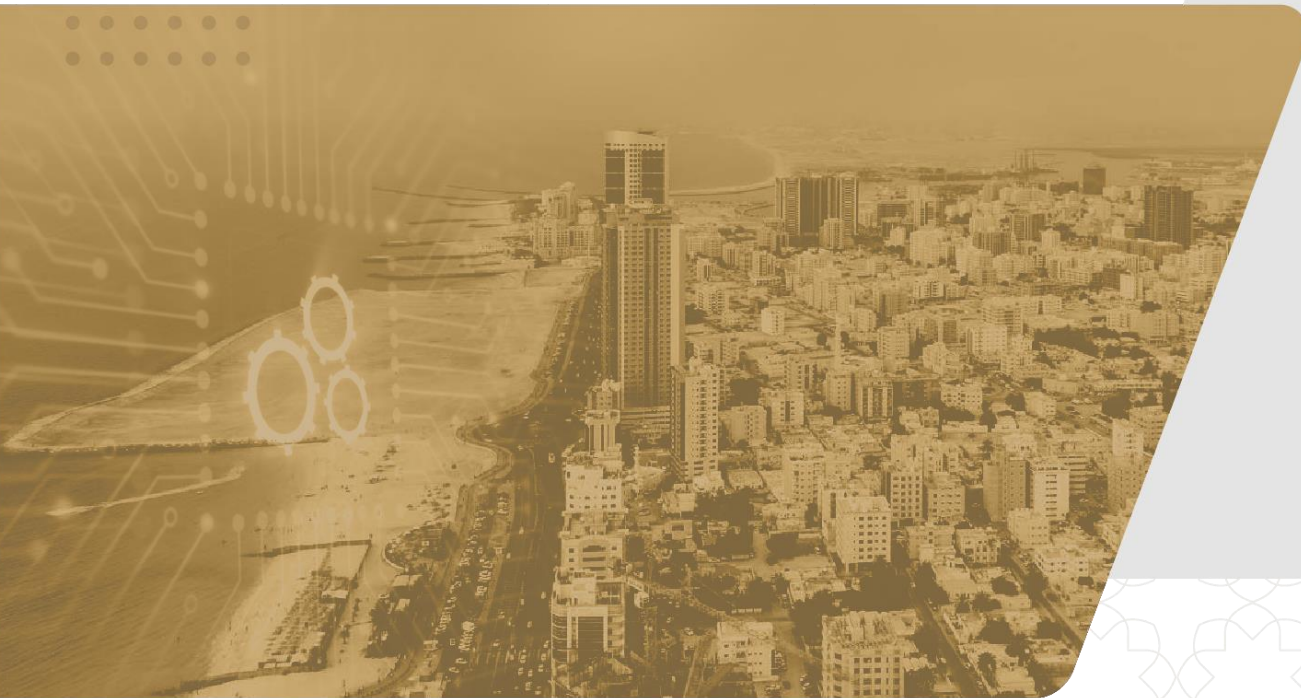
حُكُومَةُ عَجمَان

Government of Ajman

مَرْكَزُ الإحصَاء

Statistics Center

دليل مبادئ التحليل الإحصائي



إصدار ديسمبر 2023

www.scc.ajman.ae

جميع الحقوق محفوظة © مركز الإحصاء

حكومة عجمان - الإمارات العربية المتحدة @ 2023

يمنع نسخ أو استعمال أي جزء من هذا الكتاب من قبل أي شخص أو شركة أو جهة
بأية وسيلة تصويرية أو إلكترونية أو ميكانيكية بما في ذلك التسجيل الفوتوغرافي
والتسجيل على أقراص مقروءة أو بأية وسيلة نشر أخرى بما فيها حفظ المعلومات
و استرجاعها دون الحصول على موافقة مسبقة صادرة من مركز عجمان للإحصاء،
حكومة عجمان، دولة الإمارات العربية المتحدة.

في حالة الاقتباس يرجى الإشارة إلى المطبوعة كالتالي:

مركز الإحصاء – حكومة عجمان
دليل مبادئ التحليل الإحصائي
الإصدار الأول - ديسمبر 2023

للتواصل وطلب البيانات الإحصائية يرجى التواصل:

مركز عجمان للإحصاء

البريد الإلكتروني: info.scc@ajman.ae

رقم الهاتف: +971 6 701 6770

الموقع الإلكتروني: scc.ajman.ae

ص.ب: 6556، عجمان - دولة الإمارات العربية المتحدة

@sccajman

التعريف بمركز عجمان للإحصاء

تم إنشاء "مركز عجمان للإحصاء" استناداً للمرسوم الأميري رقم (8) لسنة 2022. ويعتبر المركز هو الجهة المختصة محلياً في إمارة عجمان والمصدر الرئيس والمرجع الوحيد فيها في الشؤون الإحصائية المنصوص عليها في هذا المرسوم. يهدف المركز إلى تحقيق الغايات التالية:

1. تنظيم وتطوير العمل الإحصائي بما يحقق مصالح الدولة والإمارة.
2. بناء نظام إحصائي محلي متكامل.
3. دعم منظومة اتخاذ القرار في الحكومة ببيانات ومعلومات دقيقة وحديثة.

الرؤية



بالمعرفة نعزز مستقبل عجمان.

الرسالة



الإرتقاء بالعمل الإحصائي من خلال تطبيق أفضل الممارسات بإتباع المنهجيات العلمية الإحصائية والمعايير الموصى بها دولياً لتلبي إحتياجات مستخدمي البيانات ومتخذي القرار في الإمارة.

القيم



الجودة / الحيادية / الإحترافية / الموثوقية / الإبداع والابتكار / السرية / الشفافية

دليل مبادئ التحليل الإحصائي

المحتويات

| | |
|----|------------------------------------|
| 7 | المقدمة |
| 8 | الفصل الأول |
| 8 | المنهجية |
| 8 | 1.1 أهداف الدليل |
| 8 | 2.1 أهمية الدليل |
| 8 | 3.1 مستخدمي الدليل |
| 9 | 4.1 نطاق التطبيق: |
| 9 | 5.1 محتويات الدليل |
| 9 | 6.1 مفاهيم أساسية |
| 13 | الفصل الثاني |
| 13 | جمع البيانات |
| 13 | 1.2 أساليب جمع البيانات |
| 13 | 1.1.2 الحصر الشامل |
| 13 | 2.1.2 جمع البيانات بأسلوب العينة |
| 13 | 2.2 طرق جمع البيانات |
| 13 | 1.2.2 الاستبيانات |
| 15 | 2.2.2 طريقة المشاهدة (الملاحظة) |
| 16 | الفصل الثالث |
| 16 | عرض البيانات الإحصائية |
| 16 | 1.3 العرض الجدولي للبيانات |
| 16 | 1.1.3 الجداول البسيطة |
| 17 | 2.1.3 الجداول المزدوجة |
| 17 | 3.1.3 الجداول المركبة |
| 18 | 2.3 الأشكال البيانية |
| 18 | 1.2.3 الأشرطة البيانية المفردة |
| 19 | 2.2.3 الأشرطة البيانية المركبة |
| 20 | 3.2.3 الدائرة البيانية |
| 22 | 4.2.3 الخط البياني |
| 23 | 3.3 العرض الهندسي للبيانات المبوبة |
| 23 | 1.3.3 المدرج التكراري |

| | |
|----|---|
| 24 | 2.3.3 المضلع التكراري |
| 25 | 3.3.3 المنحنى التكراري |
| 26 | الفصل الرابع |
| 26 | المقاييس الإحصائية |
| 26 | 1.4 مقاييس النزعة المركزية |
| 26 | 1.1.4 الوسط الحسابي |
| 28 | 2.1.4 الوسيط |
| 28 | 3.1.4 المنوال |
| 29 | 2.4 مقاييس التشتت |
| 29 | 1.2.4 المدى |
| 30 | 2.2.4 الانحراف المعياري : |
| 31 | 3.4 مقاييس الالتواء |
| 31 | 4.4 مقاييس التفرطح (التفلطح) |
| 33 | الفصل الخامس |
| 33 | التقدير الإحصائي واختبار الفرضيات |
| 33 | 1.5 التقدير: |
| 33 | 1.1.5 التقدير بنقطة |
| 34 | 2.1.5 التقدير بفترة الثقة |
| 35 | 2.5 اختبار الفروض |
| 35 | 1.2.5 الفرض والفرض الإحصائي |
| 35 | 2.2.5 فرض العدم و الفرض البديل |
| 36 | الفصل السادس |
| 36 | الارتباط والانحدار الخطي البسيط |
| 36 | 1.6 الارتباط الخطي |
| 37 | 1.1.6 رسم الانتشار |
| 38 | 2.1.6 معامل الارتباط |
| 39 | 2.6 الإنحدار الخطي البسيط |
| 40 | 1.2.6 طريقة المربعات الصغرى العادية لتقدير معالم النموذج الخطي البسيط |
| 42 | المراجع |

المقدمة

التحليل الإحصائي هو عبارة عن مجموعة من الأساليب والتقنيات المستخدمة لتلخيص وتحليل وفهم البيانات المترابطة، سواء كانت كبيرة أو صغيرة، حيث يهدف إلى تنظيم وتفسير البيانات للكشف عن الأنماط والعلاقات والاتجاهات التي قد تكون مخفية أو غير واضحة للوصول إلى استنتاجات دقيقة واتخاذ القرارات الأفضل. أيضاً يساعد في فهم العلاقات بين مجموعات البيانات والمتغيرات المختلفة، مما يمكن من اكتشاف العلاقات السببية والتأثيرات بين الظواهر. يُساعد التحليل الإحصائي أيضاً في اتخاذ القرارات السليمة بناءً على البيانات والأدلة، سواء كان ذلك في المجال الأكاديمي أو الصناعي أو الحكومي. كما أنه يساهم في تحليل البيانات التاريخية للتنبؤ بالاتجاهات المستقبلية ووضع الخطط والاستراتيجيات الفعالة. ويتم تحضير وتجهيز البيانات ومن ثم دراستها واستخراج النتائج منها من خلال التحليل الإحصائي، وتتم عملية تحليل البيانات من خلال اتباع عدد من الطرق الرياضية المنطقية، حيث يتم ربط العلاقات بالمضمون، وبالتالي تشكل معنى جديد ذو أهمية من علاقات لا يوجد لها أي معنى في حال وجدت بشكل منفرد، كما يعمل التحليل الإحصائي على شرح صفات ومعالم مجتمع ما، وتوضيح الصفات التي تميزه عن باقي المجتمعات، ويصل الباحث إلى هذه الصفات من خلال أخذ عينة من المجتمع ليتم دراستها، وتعميم نتائجها على مجتمع الدراسة، كما يساعد الباحث على دراسة العينات الإحصائية الكبيرة، يعد التحليل الإحصائي الحل المثالي والمناسب للعلوم كافة، ولذلك فإنها تستخدمه لتحليل بياناتها، وإصدار النتائج، كما تعد النتائج التي يظهرها التحليل الإحصائي نتائج دقيقة حيث يلعب دوراً كبيراً في مساعدة الباحث على ضبط البحث والابتعاد عن التشتت في حال كانت عينة الدراسة كبيرة، كما تستخدم الشركات التجارية التحليل الإحصائي بكثرة وذلك من أجل دراسة المجتمع الذي تنوي الاستثمار فيه، حيث يوفر لها التحليل الإحصائي معلومات حول هذا المجتمع فتوفر عليها الوقت والجهد، ويساعدها على اتخاذ القرارات الحاسمة.

الفصل الأول

المنهجية

يهدف دليل مبادئ التحليل الإحصائي إلى توفير منهج واضح وشامل للمهتمين بالعمل الإحصائي ، حيث يعتبر الدليل أداة ضرورية للتأكد من استخدام أساليب وإجراءات موحدة ودقيقة في تحليل البيانات وتقديم النتائج بشكل يضمن الشفافية والموثوقية، ويقوم مركز الإحصاء في إمارة عجمان بإعداد الكثير من الأدلة والمنهجيات لكي تساعد في الوصول إلى أفضل مستويات الجودة والموثوقية للبيانات الإحصائية، ولتحقيق هذه الأهداف تم إعداد الدليل والذي يتضمن مفهوم وأهمية التحليل الإحصائي، والخطوات الرئيسية لجمع البيانات الإحصائية، التعرف على الطرق المختلفة لعرض البيانات الإحصائية، المقاييس الإحصائية، التقديرات واختبار الفرضيات، كما يحتوي هذا الدليل على العديد من التعاريف والمصطلحات المستخدمة في العمل الإحصائي.

1.1 أهداف الدليل

- تقديم الشرح الكافي للمفاهيم الإحصائية الأكثر تداولاً بين مستخدمي البيانات فيما يختص بالتحليل الإحصائي.
- التعرف على طرق التحليل الإحصائي في مركز عجمان للإحصاء.
- توفير منهج واضح وشامل للمهتمين بالعمل الإحصائي.
- التعرف على طرق عرض البيانات الإحصائية.
- التعرف على المقاييس الإحصائية.

2.1 أهمية الدليل

- يعتبر الدليل منهجاً للاستدلال به والرجوع إليه في التحليل الإحصائي.
- زيادة المعرفة للأفراد عن طريق تقديم توجيهات حول التعاريف والأساليب الإحصائية المستخدمة في التحليل الإحصائي.

3.1 مستخدمي الدليل

- العاملين في مركز عجمان للإحصاء وجميع المراكز الإحصائية.
- المهتمين بالعمل الإحصائي.

4.1 نطاق التطبيق:

يعد هذا الدليل مرجعاً لتطبيقه في التحليل الإحصائي في إمارة عجمان كونه يحتوي وصفاً واضحاً للتعريف المستخدمة والمفاهيم المتعلقة بالتحليل الإحصائي.

5.1 محتويات الدليل

يتضمن الدليل الفصول التالية :

- الفصل الأول: المنهجية.
- الفصل الثاني : جمع البيانات.
- الفصل الثالث : عرض البيانات الاحصائية.
- الفصل الرابع : المقاييس الاحصائية.
- الفصل الخامس: التقدير الإحصائي واختبار الفرضيات.
- الفصل السادس: الارتباط والانحدار الخطي البسيط.

6.1 مفاهيم أساسية¹

إن عملية وضع المفاهيم والمصطلحات من شأنها توحيد اللغة المستخدمة في العمل الاحصائي عموماً والتحليل الاحصائي تحديداً ، حيث تم الاعتماد في وضع المفاهيم والمصطلحات على المفاهيم المعتمدة محلياً ودولياً، و فيما يلي المفاهيم الأساسية التي تم استخدامها:

- **التحليل الإحصائي:** يعتبر فرع من فروع علم الرياضيات الذي يشمل النظريات والطرق الموجهة نحو جمع البيانات والعمل على وصف البيانات، بالإضافة إلى القدرة على إتخاذ القرار.
- **الإحصاء الوصفي:** والذي يشمل الأساليب المستخدمة في تنظيم المعلومات لكي يتم فهم المعلومات بطريقة واضحة من خلال الاعتماد على الجداول التكرارية والرسوم البيانية وطرق حساب مقاييس النزعة المركزية ومقاييس التشتت وغيرها من المقاييس الأخرى ويعتمد على استخدام المنهج الوصفي.
- **الإحصاء الاستدلالي:** يهدف إلى استدلال معالم المجتمع من خلال الاعتماد على المعلومات المتوافرة عن العينة التي يقوم بدراستها بمعنى انه يهدف إلى التعميم من خلال العينة التي تم التطبيق عليها وبذلك فإن الإحصاء الاستدلالي يلعب دور بارز باختبار الفرضيات المتعلقة بالفروق بين المتوسطات او النسب المئوية المتعلقة بالعينة ويمكن للباحث اللجوء إليه عوضاً عن استخدام الاسلوب الوصفي.
- **البيانات:** هي مجموعة من المعطيات، أو الحقائق الأولية، أو التعليمات، أو الأرقام والتي تكون مواد خام ولا تكون ذات معنى ولا تحمل أي غرض محدد.

¹ <https://cutt.us/Jxq99>

- **المعلومات:** هي البيانات التي خضعت للمعالجة والتنقيح وموضوعة بطريقة ذات فائدة وأهمية ويمكن أن يتم استخدامها لاتخاذ القرارات والتنبؤات وما إلى ذلك وتعتمد المعلومات على البيانات.
- **البيانات النوعية:** وهي التي تصف البيانات التي تندرج لتصنيفات وغالباً لا يتم تصنيفها كأرقام، بل تصنف بالنصوص.
- **البيانات الاسمية:** هي القيم الاسمية الممثلة كوحداث منفصلة وتستخدم لتسمية المتغيرات التي ليس لها قيمة كمية.
- **البيانات الترتيبية:** هي القيم الترتيبية الممثلة كوحداث منفصلة والمرتبة وهي تشبه البيانات الاسمية، باستثناء أن الأمر يتعلق بالترتيب ويمكن أن يتم ترتيبها تصاعدياً أو تنازلياً ويمكن أن يكون لها تدرج وتصنيف.
- **البيانات الكمية:** هي البيانات التي تهتم بصور خاصة بالأعداد.
- **البيانات المنفصلة:** هي البيانات التي تكون قيمها مميزة ومنفصلة ويمكن أن تأخذ فقط قيماً معينة.
- **البيانات المتصلة:** هي البيانات التي يمكن قياسها وهي تتسم بأنها غير محدودة وغير معدودة.
- **البيانات الفئوية:** هي بيانات بصورة عددية مرتبطة بمتغير ما ويتم تمثيلها على شكل فترات وعلى الفترات أن تكون مقسمة بشكل متساوي.
- **البيانات النسبية:** هي بيانات مرتبطة بمتغير ما، يمكن القيام بكل العمليات الحسابية حتى الضرب والقسمة لأنها تنتج معلومات مفيدة لهذا النوع من البيانات.
- **المتغيرات:** أي خاصية أو صفة يمكن رصدها أو قياسها بشكل كمي، مثل الطول، والوزن، والعمر حيث تأتي بقيم رقمية، او لا يمكن قياسها بشكل كمي، ولكن يمكن تصنيفها في فئات، مثل اللون، والجنس، والتصنيف الاجتماعي.
- **التعداد:** هو عبارة عن إجراء يقوم بدراسة المجتمع ككل ويقوم هذا المصدر بتحليل بيانات كافة المجتمع لذلك يعتبر دقيق ويكلف الكثير من الوقت والجهد والمال.
- **المسح بالعينة:** هو أسلوب علمي وعملي منظم لجمع المشاهدات الإحصائية من بعض المفردات التي تنتمي إلى مجموعة كبيرة من المفردات بفرض تقدير الإحصاءات التي يمكن الاستدلال بها على ملامح ومعالم المجموعة الأم (الكبيرة) التي يتعذر دراستها بأسلوب الحصر الشامل.
- **استطلاعات الرأي:** هو من أنواع المسح بالعينة وهو عن آراء أو أفكار تخص مجموعة من المجتمع تجاه موضوع معين أو الفكرة المعروفة والمنتشرة لدى فئة من السكان.
- **المصادر الإدارية:** هي مجموعة من البيانات يتم وضعها في سجلات ويتم استخدامها في الوزارات والمؤسسات والمراكز الحكومية.
- **النوع:** متغير وصفي تقاس بياناته بمعياري إسمي " ذكر - أنثى ".

- **الحالة الاجتماعية:** متغير وصفي تقاس بياناته بمعيار اسمي " متزوج - أعزب - أرمل - مطلق " .
- **الجنسية:** متغير وصفي يقاس بياناته بمعيار اسمي (مواطن غير مواطن) وهذا النوع من البيانات يمكن تكويد مجموعاته بأرقام فمثلا الجنسية يمكن إعطاء الجنسية مواطن الرقم (1)، غير مواطن الرقم (2).
- **المستوى التعليمي:** متغير وصفي تقاس بياناته بمعيار ترتيبي "أمي - يقرأ ويكتب - ابتدائية - متوسطة - ثانوية - جامعية - أعلى من جامعية" .
- **البيانات الكمية:** وهي بيانات يعبر عنها بشكل رقمي (أرقام عددية) عن ظاهرة محددة، وتسمى بالبيانات المقاسة، كقياس الأوزان والمسافة والأطوال والسرعة والتعبير عن الزمن وغيرها والتي قد تكون بيانات متصلة أو منفصلة.
- **الاستبيان:** هو أحد الطرق التي تستخدم لجمع بيانات عن موضوع معين عن طريق شريحة كبيرة جداً من الناس وعن طريق هذه البيانات يمكنك عمل دراسة تحليلية لذلك الموضوع.
- **مجتمع الدراسة :** المجموعة الكلية من العناصر التي يتم تعميم النتائج عليها ذات العلاقة بالظاهرة المدروسة.
- **الإطار:** يتكون من جميع وحدات العينة أو مفرداتها والتي يرى المراجع أنها في متناول يده وقد يتفق كل من الإطار والمجتمع بحسب ما إذا كانت المفردات في متناول المراجع أم لا، ويعتبر الإطار هو القائمة المكونة من عدد من وحدات المعاينة التي يتكون منها المجتمع.
- **المجتمع :** هو كافة وحدات الظاهرة التي نحن بحاجة لدراستها، وقد يكون المجتمع محدد وقد يكون غير محدد.
- **العينة:** هي جزء من المجتمع يتم إختيارها عشوائيا أو قد تكون غير عشوائية، وعلى اساس تمثيلها لخصائص المجتمع المسحوبة منه العينة.
- **العد الشامل:** هو أسلوب من اساليب جمع البيانات الإحصائية يتم من خلاله شمول جميع وحدات المجتمع وعادة ما يستخدم أسلوب العد الشامل في التعدادات السكانية والزراعية والصناعية.
- **المعاينة :** عملية اختيار العينة من المجتمع بحيث تمثل مفردات المجتمع المراد دراسته.
- **خطط المعاينة :** هي مجموعة الإجراءات الخاصة باستخدام العينات الإحصائية لتحقيق هدف محدد من أهداف المراجعة.
- **خطأ المعاينة :** هو خطأ إحصائي يحدث عندما لا يحدد المحلل عينة تمثل مجتمع البيانات بالكامل. ونتيجة لذلك، فإن النتائج الموجودة في العينة لا تمثل النتائج التي يمكن الحصول عليها من المجتمع الإحصائي بأكمله.

- **العينة العشوائية:** هي عينة يكون فيها لكل مجموعة ممكنة من العناصر في المجتمع احتمال متساوي للدخول في العينة.
- **الباحث:** الشخص الذي يقوم باستطلاع الآراء وجمع المعلومات عند تنفيذ الدراسة الاستطلاعية.
- **البيانات التعريفية:** يقصد بها الخصائص الرئيسية والديمغرافية لمجتمع الدراسة (الاسم – الهاتف - النوع - مكان الإقامة - الحالة الاجتماعية ...).
- **الاستبيان الإلكتروني:** هو استبيان يكتب على أجهزة الحاسب الآلي ويتم إدراج الاستبيان الإلكتروني على صفحات ومواقع الشبكة العنكبوتية، أو عن طريق إرسال روابط الكترونية للمستهدفين ويتم تعبئته بشكل الكتروني.
- **الإحصاءات:** مثل احصاءات وزارة الزراعة، السكان، المواليد، الوفيات، الانتاج، الاستهلاك، الصادرات، الواردات والإحصائيات اللازمة للعلوم الاجتماعية.
- **المؤشرات:** هي التي يحصل عليها الباحث من العينة التي تشكل مجتمع الدراسة ويقصد بالعينة انها المجموعة الجزئية من الوحدات محل الدراسة.
- **المعلمة:** هي قيم حقيقية لمؤشرات المجتمع تحسب ببيانات المجتمع ككل.
- **الإحصاء:** هي مؤشرات العينة وتحسب باستخدام بيانات العينة وتستخدم كتقدير لقيمة المعلمة في المجتمع.
- **المقدر:** هو إحصاء أو معادلة للحصول على تقدير لمعلمة المجتمع.
- **التقدير:** هو القيمة العددية للمقدر.
- **المشاهدة/الملاحظة:** عملية جمع المعلومات أو البيانات عن طريق مراقبة سلوك أو أحداث معينة دون التدخل فيها مباشرة.
- **الارتباط:** قياس درجة العلاقة بين متغيرين. يستخدم معامل الارتباط لتقدير قوة واتجاه هذه العلاقة. يتراوح معامل الارتباط بين -1 و +1.

الفصل الثاني

جمع البيانات

جمع البيانات يُعتبر أحد الخطوات الأساسية في التحليل الإحصائي، حيث يمثل حجر الأساس للحصول على معلومات دقيقة وموثوقة لإجراء التحليلات اللاحقة التي ستستخدم لصنع القرارات والذي بدوره يمثل الإطار النظري للاقتراحات والتوصيات المطلوبة فيمكن اكتشاف الظواهر ومعرفة الصلات التي تربط بين جميع البيانات. يتم جمع البيانات من مصادر مختلفة وبأساليب متنوعة سيتم التطرق لها في هذا الفصل.

1.2 أساليب جمع البيانات

يتم جمع البيانات بأسلوبين هما أسلوب الحصر الشامل وأسلوب المعاينة:

1.1.2 الحصر الشامل

يتم بمقتضاه جمع بيانات عن كل مفردة من مفردات المجتمع، والتطبيق الشائع لهذا الأسلوب يتم في التعدادات مثل تعداد السكان وتعداد المنشآت والتي يتم من خلالها حصر شامل لكافة مفردات المجتمع. ولا شك أن أسلوب الحصر الشامل يتطلب تعبئة لجهود بشرية وموارد مالية وإمكانات فنية ضخمة تحول دون إمكانية تكراره إلا على فترات متباعدة ونظراً لضخامة عملية التعداد وارتفاع تكلفتها فإن البيانات التي يتم جمعها من خلال التعداد تقتصر على البيانات الأساسية دون الدخول في تفاصيل قد تكون ذات أهمية قصوى للمخطط وملتخذ القرار.

2.1.2 جمع البيانات بأسلوب العينة

يتم اختيار مجموعة جزئية من مجتمع الدراسة بطريقة مناسبة وإجراء الدراسة عليها ومن ثم استخدام تلك النتائج، وتعميمها على كامل مجتمع الدراسة الأصلي فمثلاً لدراسة القوى العاملة في إمارة عجمان فإنه يتم تقسيم الإمارة إلى أحياء يتم أخذ عينة من أفرادها بطريقة عشوائية، فالعينة تمثل جزءاً من مجتمع الدراسة من حيث الخصائص والصفات ويتم اللجوء إليها عندما يستغنى الباحث عن دراسة كافة وحدات المجتمع.

2.2 طرق جمع البيانات

بسبب تعدد طبيعة المجتمعات الإحصائية واختلاف المعطيات المطلوب جمعها والإمكانات المالية المتاحة للدراسة، فقد تعددت طرق جمع المعطيات تبعاً لذلك كما موضح في الطرق أدناه:

1.2.2 الاستبيانات

يتم جمع البيانات باستخدام الاستبيانات بتوزيعها على المجتمع المراد دراسته (إلكترونياً أو يدوياً)، بحيث يقوم المستجيبون بالإجابة على أسئلة الاستبيان، وتعتبر هذه الطريقة من الطرق الفعالة في الحالات التي يكون فيها موضوع المسح وأسئلة الاستمارة تخص المبحوثين مباشرة، كالاستفسار عن طبيعة السكن الذي يرغبون فيه أو لغرض شمولهم بإعفاءات ضريبية أو تقديم خدمات مجانية أو مخفضة لهم أو ما شابه، وتأخذ الطريقة عند تنفيذها أحد الأساليب الآتية:

الأسلوب الأول : يمكن أن يتم جمع البيانات عن طريق المقابلة الشخصية، وينقسم هذا الأسلوب الى نوعين:

1. التواصل مع المبحوثين شخصياً للحصول على إجابات منهم، وتعد الطريقة ملائمة للحالات التالية:
أولاً: في حال كان عدد وحدات المشمولين صغيراً.
ثانياً: في حال كان معظم الأشخاص المشمولين أميين.
ثالثاً: إذا كانت طبيعة الاستبيان معقدة وتحتاج إلى شرح وتوضيح.
تؤثر شخصية الباحث في هذا النوع من أساليب جمع البيانات تأثيراً كبيراً على دقة المعطيات، من خلال طريقة وأسلوب تعامله مع المبحوثين أثناء مقابلتهم، لذا فمن الضروري أن تتوفر في الباحث الذي يجمع البيانات بهذه الطريقة المواصفات التالية:

- أن يكون الباحث حسن السير والسلوك.
- أن يتمتع الباحث بالمرونة والمهارة في الحديث والقدرة على الإقناع.
- أن يتمتع برحابة الصدر والصبر والقدرة على المجاملة.
- أن يحترم العادات والتقاليد الخاصة بالمبحوثين.
- أن يقوم باحترام الأسماء والألقاب الخاصة بالمبحوثين.

2. يقوم الباحثون بزيارة أفراد العينة المستهدفة وشرح هدف المسح الإحصائي وأهميته، ثم يتم تحديد فترة زمنية معينة لتعبئة إستمارة المسح من قبل المبحوث وذلك لاستعادتها عند إنتهاء الفترة الزمنية المتفق عليها من قبل الطرفين.

الأسلوب الثاني: يمكن إجراء المقابلة باستخدام الهاتف حيث يتم من خلاله جمع المعلومات عن طريق إجراء محادثة هاتفية بين الباحث والعينة المشمولة في الدراسة.

الأسلوب الثالث : ترسل إستمارة الإستبيان عبر البريد لأفراد العينة ومن ثم يقوم أفراد العينة بتعبئة الإستمارة وإرسالها إلى مؤسسة المسح، تعتبر هذه الطريقة أقل تكلفة مقارنة بأسلوب المقابلة الشخصية والمقابلة الهاتفية وذلك لعدم الحاجة لوجود المسحيين، التكلفة الوحيدة هنا هي تكلفة طباعة إستمارات الإستبيان و الرموز البريدية. ويصلح استخدام هذه الطريقة في المجتمعات التي تقل نسبة الأمية فيها وترتفع فيها درجة الاعتماد على البريد واستخدامه.

الأسلوب الرابع : يعتبر الأسلوب الشائع والمستخدم بكثرة حالياً، حيث تُرسل إستمارة الإستبيان عبر البريد الإلكتروني أو يتم نشرها على منصات الويب أو وسائل التواصل الاجتماعي، تسمح هذه الطريقة بسهولة الوصول إليها وتوفير الوقت والمال والتكلفة.

2.2.2 طريقة الملاحظة (الملاحظة)

تمثل الملاحظة عاملاً مهماً في فحص الظواهر وتسجيل الأحداث بدقة. وتستخدم طريقة الملاحظة في حالتين تمثل الأولى مراقبة الظواهر مع استخدام المنطق في تفسير ما يحدث، وتستخدم في الغالب في بعض الحقول العلمية في دراسات اجتماعية أو تربوية وأخرى نفسية، أما الحالة الثانية مراقبة الظواهر لغرض التدوين (التسجيل) فقط، وفيها يقوم الباحث بمراقبة الظاهرة وتدوين الحقائق كما هي، وكما يحصل عند مراقبة تغيرات المناخ فيقوم بتدوين درجات الحرارة وأنواع الاثرية وغيرها من الظواهر. أحد طرق الملاحظة الحديثة هي الملاحظة الفعالة حيث يتم تحديد مؤشرات محددة للمراقبة مثل تقييم تفاعل المستخدمين مع منتجات معينة عبر وسائل التواصل الاجتماعي.

الفصل الثالث

عرض البيانات الإحصائية

عرض البيانات الإحصائية يُعتبر جزءاً حيوياً من التحليل الإحصائي؛ حيث يشمل تقديم وترتيب البيانات بطريقة تسهل فهمها وتحليلها والتي لها دور كبير في توضيح الظاهرة محل الدراسة، كما تساعد في إبراز العلاقات بين المتغيرات وإجراء مقارنات مع الظواهر الأخرى المتشابهة من أجل مقارنته بالدراسات السابقة، وذلك إما باستخدام الجداول و الرسوم البيانية وذلك لجعلها أكثر وضوحاً، لأن عرض البيانات بطريقة خاطئة سيؤدي إلى زيادة التعقيد وعدم الفهم لهذه البيانات لذا عرضها بالأسلوب الصحيح خطوة مهمة والتي ستجعل عملية الفهم واستخلاص النتائج أكثر كفاءة، وهذا الفصل يوضح طرق عرض البيانات الإحصائية :

1.3 العرض الجدولي للبيانات

الجداول تمثل طريقة لعرض البيانات الإحصائية بشكل مبسط ومفهوم، حيث يتم عمل ترتيب للبيانات الإحصائية في شكل جداول لها أعمدة وصفوف تساعد في جمع البيانات ومقارنتها. وتناسب هذه الطريقة المعلومات الفردية وتمتاز بتلخيصها للمعلومات بشكل مبسط وتمكن من مقارنة المعلومات الكمية كدرجات الطلاب في المواد الدراسية. ويمكن أن يتم العرض الجدولي باحد الأشكال الآتية:

1.1.3 الجداول البسيطة

هذا النوع من الجداول يصف قيم متغير واحد فقط ويسمى جدول تكراري بسيط، عندما يتم تحليل الجدول البسيط فإن ذلك يحدث عن طريق مقارنة التكرارات أو النسب المئوية للمتغيرات. هناك عدة أمور يجب التأكد من وجودها في أي جدول وهي:

- يجب أن يكون للجدول رقم لكي يتم تمييزه بسهولة في حالة وجود أكثر من جدول واحد.
- كتابة عنوان للجدول لإعطاء القارئ فكرة عن مضمون الجدول ومحتواه.
- هيكل الجدول والذي يحتوي على أعمدة وصفوف ويعتبر من أهم الأجزاء في الجدول.
- كل متغير له عمود يحتوي على عنوان المتغير وقيمه.
- وجود أعمدة تحتوي على قيم المتغيرات وعدد تكرارها بالنسبة للبيانات الكمية.
- من المهم وجود عمود إضافي للنسب المئوية التي ستساعد القارئ على المقارنة وتمييز القيم بسهولة أكبر.

يوضح الجدول (1.1.3) البيانات الإحصائية للقوى العاملة حسب النوع، حيث أن نسبة الذكور أكبر من نسبة الإناث، وباستخدام الجدول البسيط تتم معرفة الفرق بين المتغيرات.

جدول (1.1.3)

التوزيع النسبي للقوى العاملة حسب النوع

| النوع | النسبة المئوية |
|---------|----------------|
| ذكور | 92% |
| إناث | 8% |
| المجموع | 100% |

2.1.3 الجداول المزدوجة

الجداول المزدوجة تصف البيانات التي تتعلق بمتغيرين إثنين فقط، في حالة أن المتغيرين يحتويان على قيمتين فقط فإن هذا النوع من الجداول يسمى بالجدول التوافقي، كما أن للجداول المزدوجة اسم آخر وهو الجداول التقاطعية وذلك بسبب تقاطع قيم المتغيرين ببعضهما البعض.

في الجدول (2.1.3) يوجد متغيرين وكل متغير يحتوي على قيمتين لذا يسمى هذا الجدول بالجدول التوافقي، تم عمل إحصائية في مركز لرعاية أصحاب الهمم لمعرفة عدد أصحاب الهمم الذين لديهم إعاقة عقلية إذا كانوا مصابين بنوع آخر من الإعاقة أم لا، لذا تم رسم هذا الجدول لتوضيح ذلك حيث نلاحظ أن المصابين بأمراض عقلية ويعانون من إعاقة حركية في نفس الوقت يمثل عددهم 100 شخص أما السليمين عقلياً ولكن يعانون من إعاقة حركية بلغ عددهم 145 شخص، وعدد المصابين بأمراض عقلية ويعانون من إعاقة بصرية في نفس الوقت بلغ عددهم 39 شخص أما السليمين عقلياً ولكن يعانون من إعاقة بصرية بلغ عددهم 20 شخص.

جدول (2.1.3)

أصحاب الهمم حسب نوع الإعاقة

| نوع الإعاقة | إعاقة عقلية | سليم عقلياً |
|-------------|-------------|-------------|
| حركية | 100 | 145 |
| بصرية | 39 | 20 |

3.1.3 الجداول المركبة

تحتوي الجداول المركبة على بيانات تصنف حسب 3 صفات (متغيرات) أو أكثر في نفس الوقت وكل خلية من خلايا الجدول تمثل التكرار المشترك بين ظاهرتين.

مثلا في الجدول المركب (3.1.3) تم ترتيب بيانات الموظفين المهمة وهي إسم الموظف، والمسعى الوظيفي، والتخصص وأخيراً الرواتب، تم وضع هذه البيانات على شكل جدول لتسهيل قراءة البيانات.

جدول (3.1.3)

رواتب الموظفين حسب المسمى الوظيفي والتخصص

| الرواتب | التخصص | الوظيفة | اسم الموظف |
|---------|-------------|----------------|------------|
| 20,000 | إدارة أعمال | المدير | محمد |
| 15,000 | إدارة أعمال | نائب المدير | احمد |
| 10,000 | محاسبة | المحاسب | علي |
| 5,000 | دبلوم ترجمة | موظف الاستقبال | عمر |

2.3 الأشكال البيانية

الطريقة الأخرى التي تساعد على عرض البيانات بشكل ملفت وأكثر وضوحاً هي عن طريق الأشكال البيانية، هذا النوع من طرق العرض يوفر على القارئ ويعطيه الفكرة بسرعة ويسهل الفهم. وهناك عدة أنواع للأشكال البيانية منها:

1.2.3 الأشرطة البيانية المفردة

عبارة عن أعمدة بيانية تخص قيم متغير واحد مثل عدد إنتاجات القطن أو عدد طلاب الجامعة وتقسيم الطلاب على الكليات المتاحة، مثلاً يتم رسم الأعمدة وذلك بوضع أسماء أقسام المستشفى الموضحة في الجدول (1.2.3):

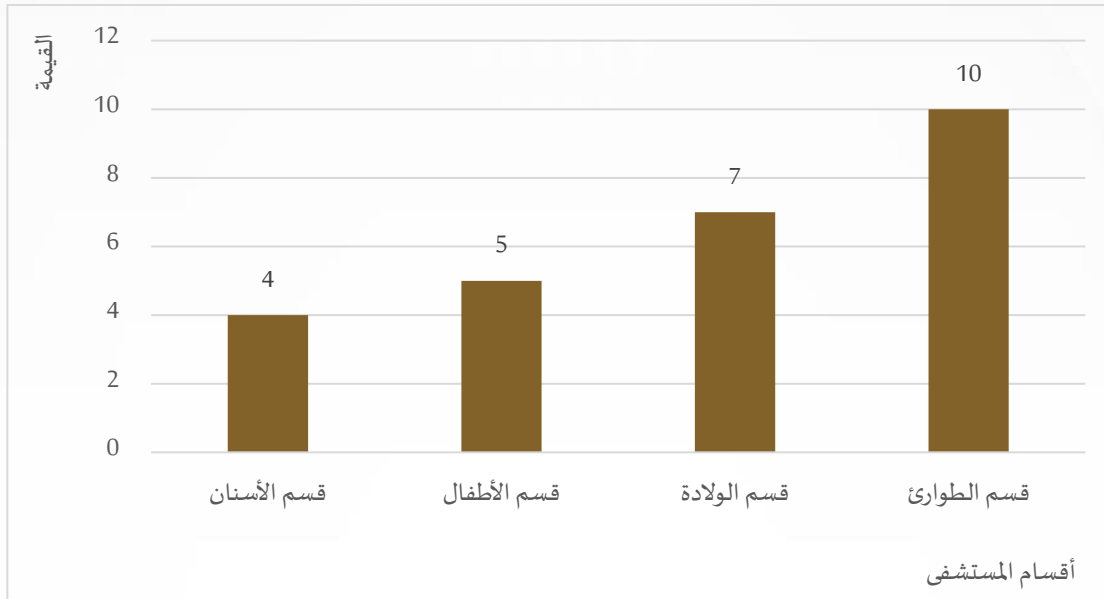
جدول (1.2.3)

أقسام المستشفى وعدد الأطباء في كل قسم

| عدد الاطباء | أقسام المستشفى |
|-------------|----------------|
| 4 | قسم الأسنان |
| 5 | قسم الأطفال |
| 7 | قسم الولادة |
| 10 | قسم الطوارئ |

ويتم برسم الأقسام في محور السينات وعدد الأطباء في محور الصادات كما موضح في الشكل (1.2.3)، وكما هو ملاحظ أن المقارنة سهلة في هذا النوع من الرسومات.

شكل (1.2.3)
أقسام المستشفى وعدد الأطباء



2.2.3 الأشرطة البيانية المركبة

هو مخطط متعدد الأعمدة يتم فيه تمثيل متغيرين أو أكثر من البيانات المترابطة بطول يتناسب مع حجم الخصائص، كما أنه يبسط المقارنة بين أكثر من حدث واحد.
مثال (2.2.2): في هذا المثال تم جمع عدد الرحلات التي ذهب لها طلاب السنوات الثلاث الأولى لكلية تكنولوجيا المعلومات كما يوضح الجدول (2.2.3).

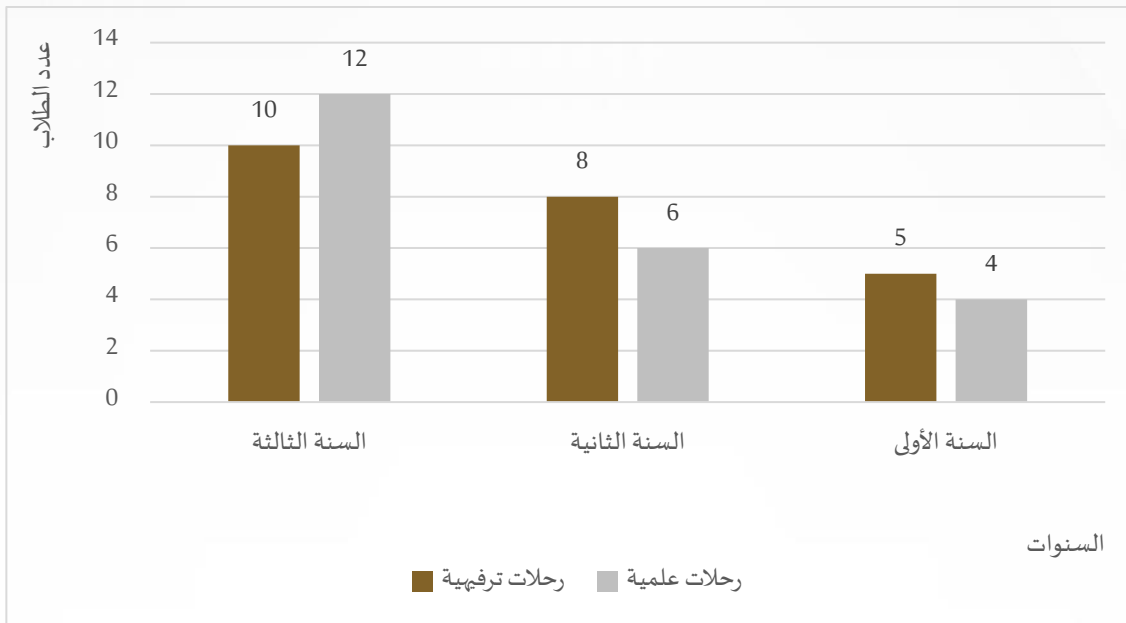
جدول (2.2.3)
أنواع الرحلات وعدد الطلاب الحاضرين

| السنوات | السنة الأولى | السنة الثانية | السنة الثالثة |
|---------------|--------------|---------------|---------------|
| رحلات ترفيهية | 5 | 8 | 10 |
| رحلات علمية | 4 | 6 | 12 |
| المجموع | 9 | 14 | 22 |

في هذا المثال يمثل محور السينات السنوات الثلاثة أما محور الصادات فيمثل المجموع وعدد الرحلات الترفيهية والعلمية، موضح في الشكل (2.2.3)

شكل (2.2.3)

أنواع الرحلات وعدد الطلاب الحاضرين



3.2.3 الدائرة البيانية

عبارة عن رسم بياني على شكل دائري ومقسم إلى أجزاء أو قطاعات وكل جزء يمثل قيمة المتغير الموجود ، ولتبسيط الموضوع فإن هذا النوع من الرسوميات يمثل الجزء من الكل كما أنه يسهل عملية المقارنة لإختلاف أحجام القطاعات في الرسم الواحد، يتم استخدام هذا النوع في مجالات عديدة كالتجارة والإنتاج وغيرها.

ولمعرفة كيفية رسم كل جزء من الدائرة فإن العملية بسيطة للغاية وهي تقسيم الجزء المطلوب على العدد الكلي للقيم الموجودة كما أن المثال التالي يوضح آلية العملية.

يتم استخدام هذه المعادلة الحسابية لمعرفة حجم الجزء المطلوب: $100 \times \frac{\text{الجزء}}{\text{الكل}}$ ولمعرفة الزوايا الخاصة بكل جزء يتم احتساب هذه المعادلة الحسابية:

$$360^\circ \times \frac{\text{الجزء}}{\text{الكل}}$$

على سبيل المثال في حالة تم عمل مسح يخص الحالة الاجتماعية للأفراد في مدينة ما والذين يبلغ عددهم 1000 فرد، فإن الدائرة البيانية تعتبر من أفضل الأساليب التي يمكنها تمثيل النتائج بصورة واضحة ومفهومة وتساعد في ملاحظة النسبة الأكبر بطريقة أسرع.

عدد الأفراد العزاب: 470 فرد

$$47\% = 100 \times \frac{470}{1000} = 100 \times \frac{\text{الجزء}}{\text{الكل}}$$

عدد الأفراد المتزوجين: 360 فرد

$$36\% = 100 \times \frac{360}{1000} = 100 \times \frac{\text{الجزء}}{\text{الكل}}$$

عدد الأفراد المطلقين: 150 فرد

$$15\% = 100 \times \frac{150}{1000} = 100 \times \frac{\text{الجزء}}{\text{الكل}}$$

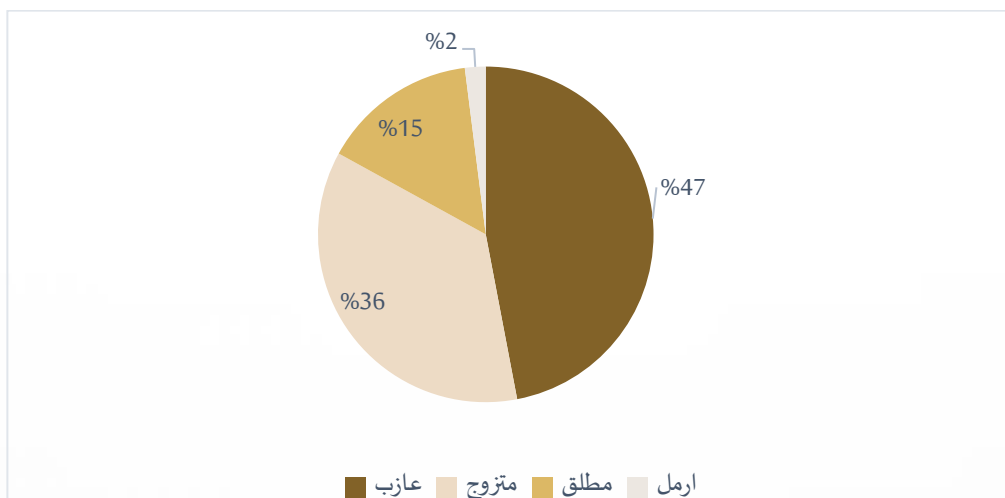
عدد الأرامل: 20 فرد

$$2\% = 100 \times \frac{20}{1000} = 100 \times \frac{\text{الجزء}}{\text{الكل}}$$

ويتم تمثيله بيانياً كما في الشكل (3.2.3)

شكل (3.2.3)

الحالة الاجتماعية



4.2.3 الخط البياني

تُستخدم الرسوم البيانية الخطية لمتابعة التغيرات على مدى فترات زمنية قصيرة وطويلة وعرض الأشياء التي تتغير مع الوقت، يقوم هذا النوع من الرسوم بربط نقاط متصلة من البيانات ببعضها البعض بخط مستقيم ويتكون الرسم البياني الخطي من محورين وهما محور السينات ومحور الصادات، كما أنها تسهل تعلم المهارات الأساسية لقراءة الرسم البياني ولكن من الأفضل عدم استخدامه مع الكثير من المتغيرات، لأن استخدام المخطط الشريطي سيكون أفضل لقدرته على توضيح الفروق لمقارنتها ولتحديد الاتجاهات، من الأمثلة المناسبة لاستخدامه كمثال لمعرفة سعر الأسهم، أو درجات هطول الأمطار أو الرطوبة. مثلاً لتمثيل البيانات في الجدول (4.2.3) كالتالي:

جدول (4.2.3)

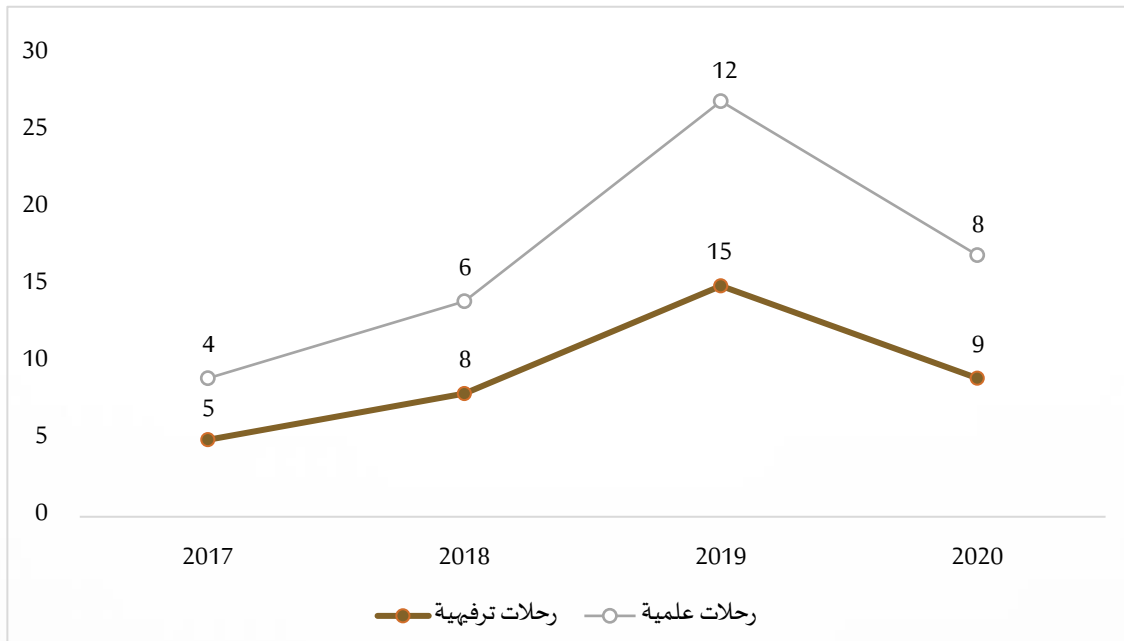
أنواع الرحلات وعدد الطلاب الحاضرين في كل سنة

| السنوات | 2017 | 2018 | 2019 | 2020 |
|---------------|------|------|------|------|
| رحلات ترفيهية | 5 | 8 | 15 | 9 |
| رحلات علمية | 4 | 6 | 12 | 8 |

يمثل الخط البياني في الشكل (4.2.3) هذه المتغيرات.

شكل (4.2.3)

أنواع الرحلات وعدد الطلاب الحاضرين في كل سنة دراسية



3.3 العرض الهندسي للبيانات المبوبة

للبيانات المبوبة رسوم بيانية معينة ومناسبة لهذا النوع من البيانات المرتبة والتي يتم وضعها في جدول تكراري وضمن فئات محددة، وكمثال عندما يتم تقسيم الأعمار على حسب الأحجام في ملابس الأطفال فمثلاً يتم وضع الأطفال من عمر 1-2 سنة في فئة والفئة الثانية الأعمار من 3-4 والفئة الثالثة من الأعمار 5-6 والفئة الرابعة من الأعمار 7-8 وسيتم توضيح الرسوم الهندسية واعطاء أمثلة.

1.3.3 المدرج التكراري

يتم تمثيله برسم بياني من أعمدة متصلة ببعضها البعض وتحتوي على توزيع تكراري لمتغير معين، تمثل قاعدة العمود (محور السينات) الفئات، أما ارتفاع العمود (محور الصادات) يمثل قيمة تكرار كل فئة. المدرج التكراري يوفر شكل رائع يسهل من الفهم والملاحظة مما يساهم إتخاذ القرارات. مثال (1.3.2): يمثل هذا المدرج التكراري في الشكل (1.3.3) الفئات العمرية وعدد التكرارات لكل فئة عند جمع عدد زوار متحف عجمان في أحد الأيام وإجراء إحصائية لأعمارهم كما يوضح الجدول التكراري (1.3.3) وكان عدد الزوار 39 زائر.

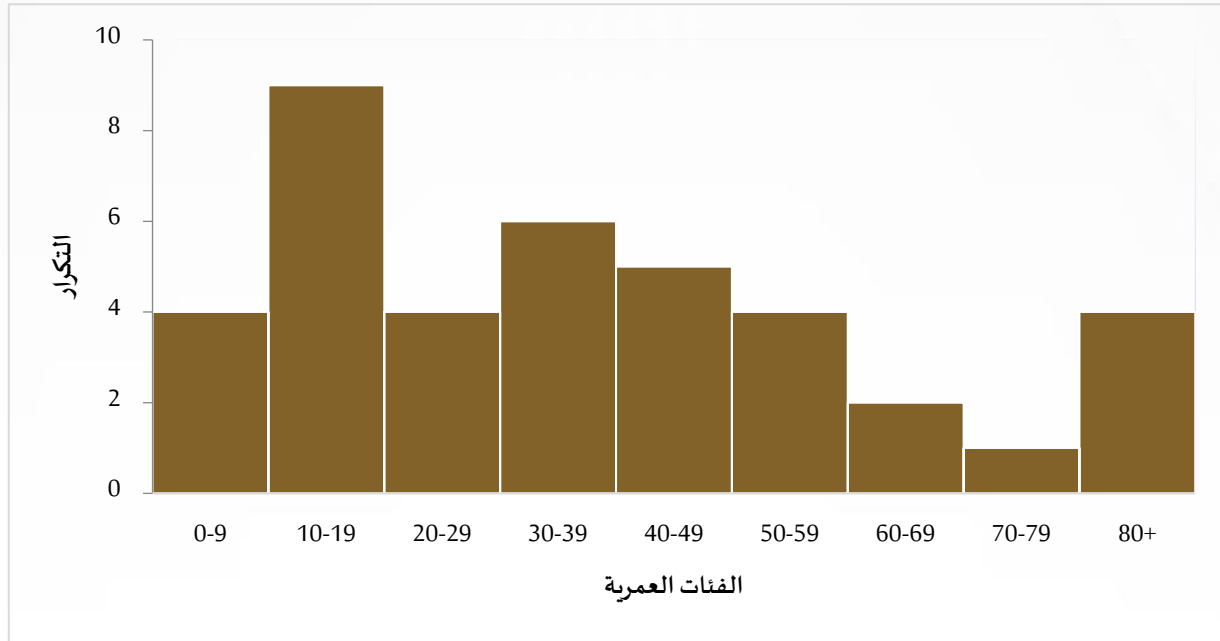
جدول (1.3.3)

عدد زوار متحف عجمان وفئاتهم العمرية

| التكرار | الفئات العمرية |
|---------|----------------|
| 4 | 0-9 |
| 9 | 10-19 |
| 4 | 20-29 |
| 6 | 30-39 |
| 5 | 40-49 |
| 4 | 50-59 |
| 2 | 60-69 |
| 1 | 70-79 |
| 4 | 80+ |
| 39 | المجموع |

شكل (1.3.3)

عدد زوار متحف عجمان وفئاتهم العمرية



2.3.3 المضلع التكراري

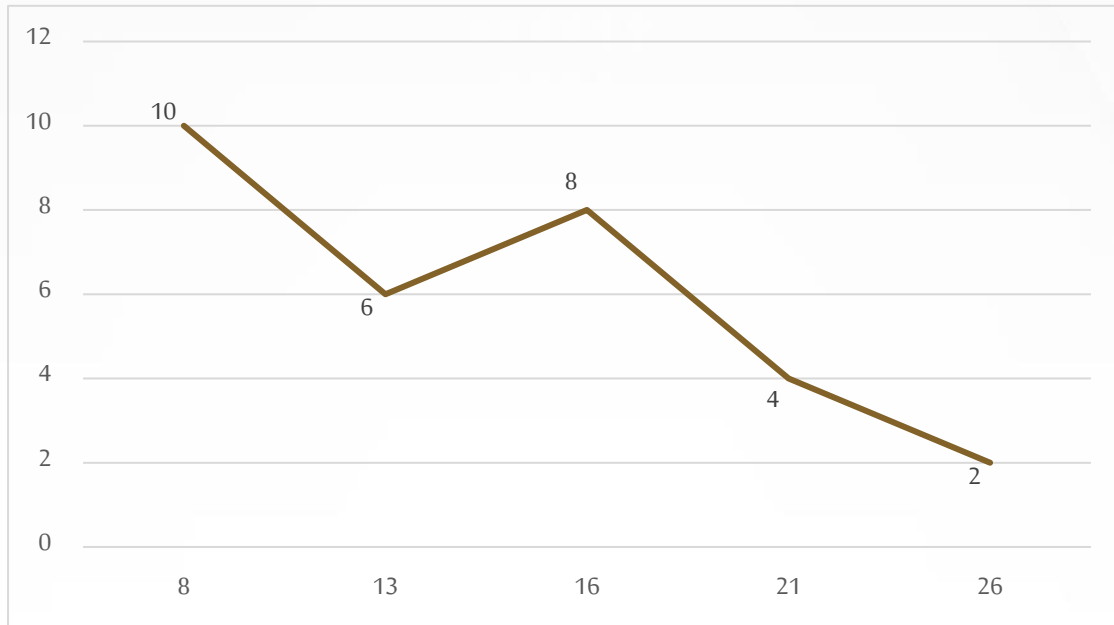
يُمثل برسم بياني يتكون من نقاط تسمى مراكز الفئات تتصل هذه النقاط ببعضها البعض مكونة خط، يتواجد في المحور السيني مراكز الفئات والتي يمكن إيجادها عن طريقة حساب المتوسط الحسابي لكل حد أدنى وحد أعلى لكل فئة عن طريق المعادلة الحسابية الآتية (الحد الأدنى + الحد الأعلى) / 2 وفي المحور الصادي يتواجد به التكرار. مثلاً لتمثيل المضلع التكراري لأعمار الطلاب في الجدول (2.3.3) يمكن رسمه عن طريق حساب عدد مرات غياب طلاب المدارس والدراسات العليا نسبة إلى أعمارهم، لذا محور السينات سيحوي على مراكز الفئات كما هو موضح في الشكل (2.3.3) أما التكرار فسيكون في محور الصادات.

جدول (2.3.3)

أعمار طلاب المدارس والدراسات العليا بحسب الفئات العمرية

| مراكز الفئات | التكرار | الفئات العمرية |
|--------------|---------|----------------|
| 8 | 10 | 10-6 |
| 13 | 6 | 15-11 |
| 16 | 8 | 18-14 |
| 21 | 4 | 23-19 |
| 26 | 2 | 28-24 |

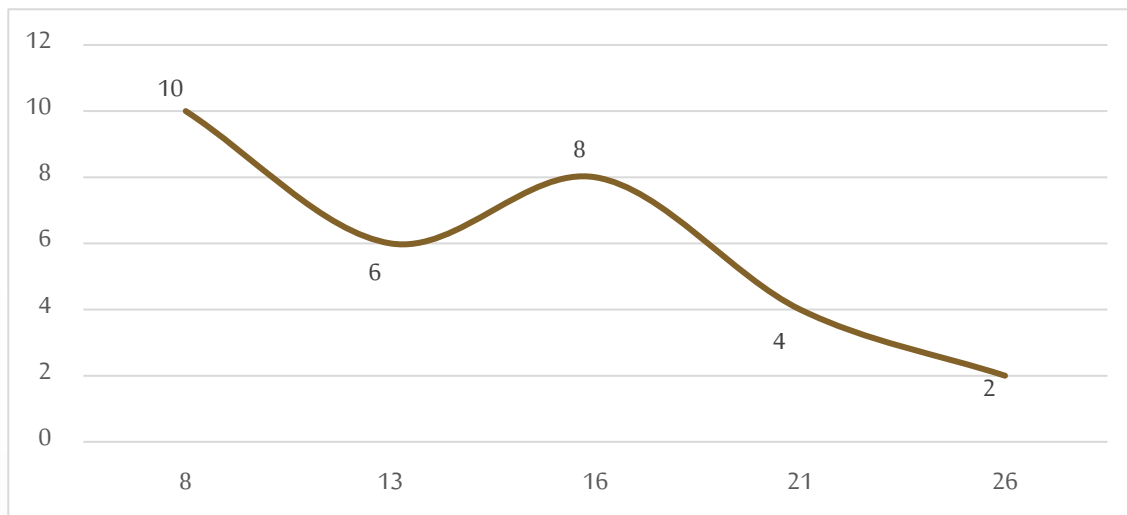
شكل (2.3.3)
المضلع التكراري



3.3.3 المنحنى التكراري

يُمثل برسم بياني يتكون من نقاط وهو مشابه للمضلع التكراري ولكن يتم رسم الخط على شكل منحنى بدلاً من رسمه كخط مستقيم، كما في الشكل (3.3.3) كما في الشكل أدناه.

شكل (3.3.3)
المنحنى التكراري



الفصل الرابع

المقاييس الإحصائية

تعتبر المقاييس الإحصائية من أهم الأدوات الفعالة والمفيدة للبحث العلمي حيث تستخدم كقواعد أساسية لمعرفة مدى تلائم النتائج مع الأهداف كما تستخدم المقاييس الإحصائية في عملية التقييم الكمي للبيانات. وتستخدم بشكل أساسي في مقارنة الأداء الخاص بالبيانات في البحث، كما تتعامل مع البيانات الكمية، وتقوم بعملية تحليل ووصف مجموعة من البيانات الإحصائية، وتقسم إلى أربعة أنواع: مقاييس النزعة المركزية، ومقاييس التشتت، ومقاييس الارتباط، ومقاييس التوزيع.

1.4 مقاييس النزعة المركزية

تُستخدم للحصول على قيمة واحدة لوصف البيانات، عن طريق تحديد المركز لتلك البيانات، ولهذا تسمى أحياناً مقاييس الموقع المركزي، وفيما يلي أهم ثلاثة مقاييس للنزعة المركزية:

1.1.4 الوسط الحسابي

يحسب بأخذ مجموع قيم المتغير وقسمته على العدد الإجمالي للقيم. سنتطرق هنا لنوعين منه:

• الوسط الحسابي البسيط

يعتبر الوسط أو المتوسط الحسابي من أهم المقاييس الإحصائية التي يتم استخدامها في الحياة العملية والنواحي التطبيقية لأنه سهل الحساب، يصف الوسط الحسابي جميع البيانات الموجودة ويلخصها في رقم واحد، كما في المثال التالي:

حصل الطلاب على درجات متفاوتة وتم تجميعها (19، 18، 14، 20، 15، 20، 17)، لمعرفة المتوسط الحسابي يجب استخدام هذا القانون: (جمع الأرقام وتقسيمها على عددها).

$$\bar{x} = \sum_{i=1}^n \frac{x_i}{n} = \frac{19 + 18 + 14 + 20 + 15 + 20 + 17}{7} = 17.6$$

لذا الوسط الحسابي هو: 17.6

ومن الضروري إن كان هناك قيمة متطرفة في مجموعة من البيانات أن يتم تعديلها أو حذفها لأنها ستؤدي إلى ظهور مؤشر مضلل ولن يكون استخدام المتوسط الحسابي مناسب في هذه الحالة والقيمة المتطرفة هي القيمة الغير متناسقة مع بقية البيانات في المجموعة وتظهر اختلاف شديداً، إما أن تكون صغيرة جداً أو كبيرة جداً.

• الوسط الحسابي المرجح أو الموزون

يتم حساب الوسط الحسابي البسيط بعملية بسيطة وهي جمع القيم وتقسيمها على عددها، ولكن، في حالة أن هناك متغير مهم آخر ويجب أن يتم حسابه في الوسط الحسابي فإن هذه الطريقة التقليدية لن تجدي نفعاً، ولهذا يمكن استخدام الوسط الحسابي المرجح وهو يوفر نتائج أكثر دقة. يعمل الوسط الحسابي المرجح في قانون رياضي بسيط وهو:

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

w_i = الوزن

x_i = القيمة

ويفسر القانون الرياضي بأنه إجمالي حاصل ضرب القيمة في وزنها مقسوماً على إجمالي الأوزان، فمثلاً تم جمع البيانات عن مجموعة من خمسة طلاب ودرجاتهم وعدد ساعات الدراسة في امتحان الإحصاء كما في الجدول (1.1.4)، وأول خطوة يتم ضرب كل قيمة في وزنها وبعد ذلك يتم التطبيق على القانون.

لتكون قيمة الوسط الحسابي المرجح تساوي

$$\bar{x} = \frac{\sum_{i=1}^n w_i x_i}{\sum_{i=1}^n w_i} = \frac{339}{20} = 16.95$$

جدول (1.1.4)

الطلاب وعدد ساعات

| الطلاب | القيمة (الدرجة في الامتحان) | الوزن (عدد ساعات الدراسة) | القيمة x الوزن |
|---------|-----------------------------|---------------------------|----------------|
| 1 | 20 | 6 | 120 |
| 2 | 18 | 3 | 54 |
| 3 | 15 | 4 | 60 |
| 4 | 17 | 5 | 85 |
| 5 | 10 | 2 | 20 |
| المجموع | - | 20 | 339 |

2.1.4 الوسيط

يعتبر الوسيط كذلك أحد المقاييس الإحصائية التي تندرج تحت فئة مقاييس النزعة المركزية، ويمكن تعريفه بأنه القيمة التي توجد في منتصف القيم، التي من خلالها يتم تقسيم كافة القيم إلى نصفين، وكل من النصفين متساويين، إن قياس الوسيط في غاية السهولة والبساطة فهو عملية حسابية لا تتأثر بالقيم وخاصة القيم المتطرفة، يقوم الوسيط بوصف البيانات التي تتوسط مجموعة من البيانات فهي القيمة المتوسطة التي تقع في منتصف البيانات بعد ترتيبها بشكل تصاعدي أو تنازلي، ولفهم هذه العملية الحسابية فأن أول خطوة هي ترتيب 55 القيم الموجودة في مجموعة البيانات بترتيب تصاعدي أو تنازلي كما ذكر سابقاً، بعد ذلك يتم عد الب58888+ يانات ومعرفة العدد الكلي، إن كان العدد الكلي للبيانات عدد فردي يتم اختيار العدد الذي يقع في المنتصف كوسيط، أما في حالة إن كان العدد الكلي عدد زوجي، يتم جمع العددين اللذان يقعان في المنتصف ومن ثم تقسيمهما على اثنين وبهذه الطريقة يمكن إيجاد الوسيط لكل الأعداد.

وكمثال عند ترتيب هذه المجموعة التي تحتوي على بيانات تخص عدد الكتب التي يتم استعارتها من المكتبة من قبل كل طالب {2,3,3,4,5,6,1,5} يتم ترتيبها بشكل تصاعدي لتصبح: 1,2,3,3,4,5,6 ولأن عدد البيانات هو 7 قيم لذا الوسيط هو 3.

وفي حالة اضافة رقم متطرف فأن العملية لن تتأثر به على الاطلاق. {2,3,3,4,5,6,1,5,40} وعند ترتيبها بشكل تصاعدي تصبح: 1,2,3,3,4,5,6,40، وهنا، عدد القيم في المجموعة بعد إضافة الرقم المتطرف أصبح 8 وهو عدد زوجي لذا سيكون الوسيط من رقمين هما: 4 و3 ولكن يجب القيام بالعملية لكي يكون الوسيط متكون من رقم واحد فقط:

$$\text{Median} = \frac{4 + 3}{2} = \frac{7}{2} = 3.5$$

والوسيط لهذه المجموعة من البيانات هو العدد 3.5.

3.1.4 المنوال

يعتبر المنوال مقياساً إحصائياً يستخدم لمعرفة القيمة الأكثر تكراراً في مجموعة القيم، ولكن وجودها ليس دائماً وفي حالة عدم وجود قيم متكررة للبيانات يسمى عديم المنوال، وفي حالة وجود قيم متكررة يكون المنوال متعدد.

يمكن استنتاج المنوال في المثال التالي حيث تتوفر البيانات التالية عن عدد أكواب الماء المشروب من قبل مجموعة من الأشخاص :

(7,8,1,4,8,2) من الملاحظ أن الرقم 8 الأكثر تكراراً وقد تكرر ثلاثة مرات لذا يمكن اعتباره المنوال لهذه المجموعة من البيانات.

2.4 مقاييس التشتت

تساعد هذه المقاييس في تفسير التباين في البيانات، أو بمعنى آخر معرفة مقدار البيانات المتجانسة أو غير المتجانسة، وتنقسم إلى:

1.2.4 المدى

يعد المدى مقياساً سهل الحساب يتم حسابه بإيجاد الفرق بين أكبر قيمة وأصغر قيمة في البيانات، ويمتاز المدى بسهولة الفهم والحساب ومستقل بذاته، أما عيوبه فإنه لا يعتبر مقياساً موثوقاً لحساب التشتت، ويعتمد اعتماداً كلياً على تغيير المقياس.

ويُفضل أن لا يتم استخدامه في حالة وجود قيم متطرفة لأنه يتأثر بشكل كبير بها، وإذا كان المدى كبير فإن ذلك يعني أن التشتت (التباعد) كبير وبين المدى والتشتت علاقة طردية.

وكمثال على المدى يمكن وضع مجموعة من البيانات التي تخص أعمار الطلاب في مادة الإحصاء بهذا الشكل

19،18،20،21،19،17

ويلاحظ أن أكبر قيمة هي 21 وأصغر قيمة هي 17 لذا يتم عملية الطرح بينهما لقياس المدى وتحديد درجة التباين أو التشتت لمجموعة البيانات،

$$21 - 17 = 4$$

لذلك التشتت قليل في هذا المثال.

ومع مثال آخر يحتوي هذه المرة على أعمار المدرسين في إحدى الجامعات والأعمار مرتبة كالآتي:

25،55،53،60،63،58

وفي هذا المثال كانت أكبر قيمة هي 63 والأصغر هي 25، من الواضح أن الفرق بين الرقمين كبير ويمكن إثبات ذلك عن طريق حساب المدى

$$63 - 25 = 38$$

وتم الذكر سابقاً أن الفرق كبير بسبب وجود قيمة متطرفة وهو الرقم 25 والذي تسبب في زيادة التشتت وفي حالة تم حذف هذه القيمة فإن المدى سيكون

$$63 - 53 = 10$$

وهو أقل تشتت من وجود القيمة المتطرفة.

2.2.4 الانحراف المعياري :

يعتبر الانحراف المعياري القيمة الأكثر استخداماً وشيوعاً بين مقاييس التشتت الإحصائي لقياس مدى التبعثر الإحصائي، أي أنه يدل على مدى امتداد مجالات القيم ضمن مجموعة البيانات الإحصائية ونظراً لكون الانحراف المعياري هو الجذر التربيعي للتباين والتباين هو معدل مربعات انحرافات العلامات في التوزيع عن الوسط الحسابي. تؤثر القيم المتطرفة على التباين أو الانحراف المعياري بينما لا تؤثر كثيراً تغيرات العينة، ويرتبطان بالوسط الحسابي للتوزيع، أي أن التشتت في البيانات ينسب إلى الوسط الحسابي.

في الوهلة الأولى سيبدو أن التباين يشبه المدى ولكن التباين أفضل للاستخدام في حالة وجود قيم متطرفة لأنه يوفر نتيجة أكثر دقة من مؤشر المدى، ويقوم التباين بقياس مدى التشتت بين البيانات بين بعضها البعض وبين المتوسط الحسابي، فإن كانت قيمة التباين كبيرة فإن ذلك يعني أن القيم متباعدة عن بعضها وعن الوسط الحسابي والعكس صحيح، ومن المهم معرفة أن للتباين عدد إيجابي دائماً وذلك لأنه يمثل مربع الانحراف المعياري. ويتم التعبير عنه رياضياً بهذه المعادلة:

$$\frac{\sum_{i=1}^N (X_i - \mu)^2}{N - 1}$$

$$X_i = \text{القيم}$$

$$\mu = \text{المتوسط الحسابي}$$

$$N = \text{عدد القيم}$$

مثلاً ما هو التباين للعينة الآتية التي تمثل أطوال الأشجار بالمتر في أم القيوين: 20، 31، 44، 51، 69؟
وبين الجدول (2.2.4) أطوال الأشجار بالمتر في أم القيوين حيث الوسط الحسابي لهذا المثال هو:

$$\mu = \frac{17 + 20 + 31 + 44 + 51 + 69}{6} = 38.9$$

ومن ثم نعوض بالبيانات ونستخدمها في هذا القانون:

$$\begin{aligned} \frac{\sum_{i=1}^N (X_i - \mu)^2}{N - 1} &= \frac{470.9 + 349.7 + 59.3 + 28.1 + 151.3 + 918.1}{6 - 1} \\ &= \frac{1977.4}{5} = 395.48 \end{aligned}$$

وتكون قيمة التباين هي 395.48

جدول (2.2.4) أطوال الأشجار في أم القيوين

| قيم المشاهدات | الانحراف عن الوسط الحسابي (القيمة- الوسط الحسابي) | مربع الانحراف عن الوسط الحسابي (القيمة- الوسط الحسابي) ² |
|---------------|--|--|
| 17 | -21.7 | 470.9 |
| 20 | -18.7 | 349.7 |
| 31 | -7.7 | 59.3 |
| 44 | 5.3 | 28.1 |
| 51 | 12.3 | 151.3 |
| 69 | 30.3 | 918.1 |
| المجموع | - | 1977.4 |

3.4 مقاييس الإلتواء

تدل خاصية الإلتواء على درجة ابتعاد المنحنى التكراري عن التماثل، فإذا كانت معظم القيم في الطرف الأدنى من التوزيع، وامتد الطرف الأيسر للمنحنى ناحية اليسار بحيث يكون الوسط الحسابي أكبر من الوسيط، فغن البيانات تتوزع توزيعاً ملتوياً ناحية اليسار، أما إذا كان العكس فيقال بأن المنحنى ملتوي التواء موجب.

$$sk = \frac{3(\bar{x} - M)}{s}$$

$$sk = \frac{\sum_{i=1}^n (x_i - \bar{x})^3}{s^3(N - 1)}$$

حيث SK هو الإلتواء و \bar{x} هي الوسط الحسابي و M هي الوسيط و S الانحراف المعياري و x_i هي قيم المتغير بينما N يمثل حجم المجتمع.

4.4 مقاييس التفطح (التفطح)

تشير خاصية التفطح إلى درجة تركيز التكرارات في منطقة الوسط للبيانات، بالنسبة للتركيز في الطرفين، مقارنة بالتوزيع الطبيعي القياسي. وهو مقياس يقيس درجة علو أو انخفاض أي منحنى توزيع تكراري بالنسبة للمنحنى الطبيعي للبيانات، وهو منحنى متماثل حول الرأس يمر بالمتوسط، فإذا كان للتوزيع قمة مرتفعة (أكبر من التوزيع الإعتدالي) يقال أنه مدبب. وإذا كان التوزيع ذو قمة مسطحة يقال أنه

مفلطح، وإذا كانت قمة التوزيع متوسطة (ليست مدببة وليست مفلطحة) يسمى متوسط التفلطح .
وصفة التفلطح ليس لها علاقة بالمتوسط الحسابي للتوزيع فقد يكون هناك أكثر من توزيع لهم نفس
المتوسط الحسابي ولكن يختلف شكل المنحنى من مدبب أو مسطح.
فإذا كان ارتفاع قمة التوزيع الاعتدالي تساوي 3 تقريباً، فإن التوزيع يكون مفلطحاً عندما يكون معامل
التفلطح أقل من 3، ويكون التوزيع مدبباً عندما يكون معامل التفلطح أكبر من 3 يحسب معامل التفلطح
من الصيغة الرياضية التالية:

$$sk = \frac{\sum_{i=1}^n (x_i - \bar{x})^4}{s^4(N - 1)}$$

حيث SK هو التفرطح و \bar{x} هي الوسط الحسابي و S الانحراف المعياري و x_i هي قيم المتغير بينما N يمثل
حجم المجتمع.

الفصل الخامس

التقدير الإحصائي واختبار الفرضيات

دراسة المجتمع وخصائصه تتم من خلال اخذ عينة من مفرداته، فتتم إجراء عملية تقدير إحصائي للمعلومات من معطيات العينة المختارة، وتسمى كل من هذه المقدرات بالإحصاءة. حيث أن قرار قبول المقدرات واعتمادها في دراسة خصائص المجتمع الذي سحبت منه العينة، مرتبط بعملية تقييم لتلك المقدرات، لأن تقدير المؤشر الإحصائي المسحوب من مفردات العينة قد لا يساوي معلمة المجتمع. وأن وجود فروق بين هذين المؤشرين قد يكون بالإمكان تجاهلها إذا كانت بسيطة. وليس لها خطورة على قيمة المقدّر، ولكن قد يكون الفرق واضحاً وجوهرياً.

تهدف عملية التقدير الإحصائي إلى الوصول إلى أفضل مقدر أو أكثر لمعلومات المجتمع. أما اختبار الفرضيات الإحصائية فينطوي على بناء أساليب تعتمد على البيانات قيد الدراسة لإتخاذ قرار بشأن فرضية تصاغ قبل التعامل مع بيانات العينة. سنسرد في هذا الفصل التقدير وأنواعه وخطواته واختبارات الفروض.

1.5 التقدير:

تتم عملية التقدير إما بالسعي للحصول على قيمة مقدرة محددة، مشتقة من بيانات عينة من المجتمع، بحيث جعلها أقرب ما يمكن إلى قيمة المعلمة الحقيقية، أو بحساب حدود يتوقع أن تقع القيمة الحقيقية للمعلمة ضمنها باحتمال معين، وكلما كان ذلك الإحتمال عالياً كلما كانت هناك موثوقية أكبر في الحصول على القيمة الحقيقية للمعلمة ضمن مدى الثقة. وعليه فإن التقدير ينقسم إلى:

1.1.5 التقدير بنقطة

يتم من خلاله حساب قيمة مقدرة لمعلمة المجتمع من العينة، وتقدير النقطة هو حساب قيمة معينة (واحدة) من العينة على أنها تمثل قيمة معلمة المجتمع وذلك بإحتمال خطأ تقدير معين (α).

إن تقدير النقطة هو إجراء يتقرر بموجبه اعتماد المقدّر $\bar{\theta}$ لمعلمة المجتمع θ ، ولا يعتمد الفرق الذي يعبر عنه بالمقدار $(\theta - \bar{\theta})$ للحكم على دقة عملية التقدير، ويستعاض عنه أحياناً بمربع الفرق $(\theta - \bar{\theta})^2$ وذلك للتخلص من أثر الإشارة في الفروق، أو أحياناً يستعاض عنه بالقيمة المطلقة، حيث يتم اختيار المقدّر الذي يجعل القيمة المتوقعة لمربعات الفروق بين قيمة المعلمة وقيمة المقدّر أقل ما يمكن، ويسمى المقدّر حينها بالمقدّر ذو أقل متوسط مربعات خطأ.

مثلاً إذا كان المطلوب هو التعرف على مؤشر متوسط إنفاق الفرد في الإمارة. فإن ذلك يتطلب بيانات شاملة لكافة أفراد سكان إمارة عجمان ويتم سؤالهم عن مقدار دخل كل فرد، مما يعني إجراء عملية

مسح شامل لكافة الأفراد الأمر الذي سيؤدي إلى كلفة مالية عالية وإلى وقت زمني طويل ونتيجة المؤشر وهو متوسط إنفاق الفرد لن تكون بالمستوى المطلوب من الدقة وذلك لكبر حجم عملية جمع البيانات وكثرة الفرق الميدانية العاملة واحتمالات الأخطاء المختلفة التي ستعكس بالنهاية على قيمة معلمة المجتمع.

إن البديل المناسب لعملية المسح الشامل هو إجراء مسح بالعيننة من خلال اختيار عيننة من أفراد الإمارة واستيفاء بيانات الإنفاق لكل فرد من أفراد العيننة وحساب متوسط إنفاق الفرد من بيانات العيننة أي مقدار $\bar{\theta}$ واعتباره تقديراً لمعلمة متوسط إنفاق مجتمع أفراد الإمارة أي مقدار θ .

لقياس دقة القيمة المقدرة للمؤشر فإن المقدّر يتضمن نسبة خطأ محددة ولكنها غير معروفة وهذا الخطأ من مصدرين رئيسيين الأول من العيننة ويدعى بالأخطاء العيننة أو أخطاء العيننة، والثاني من إجراءات وعمليات جمع البيانات ويدعى بالأخطاء غير العيننة والتي لا يمكن قياسها وإنما يمكن الحد منها لتكون في أقل مستوى من خلال ضبط إجراءات وعمليات جمع ومعالجة البيانات.

أما خطأ العيننة فيمكن قياسه بالإعتماد على مقدار الانحراف المعياري S لبيانات أفراد العيننة العشوائية البسيطة التي تشتمل على n من الوحدات، من خلال حساب ما يعرف بخطأ المعاينة وهو:

$$\frac{s}{\sqrt{N}} \times \left(\frac{N-n}{N} \right)$$

حيث N هو حجم المجتمع الكلي.

مثلاً إذا كان المطلوب هو تقدير متوسط حجم الأسرة في إمارة عجمان فإن معلمة المجتمع المبحوث هنا هو متوسط حجم الأسرة، اخذت عيننة حجمها 5000 أسرة، وتم استيفاء بيان عدد الأفراد لكل أسرة، بعد ذلك تم حساب مقدار متوسط حجم الأسرة في الإمارة اعتماداً على بيانات العيننة وكان المقدّر هو 6.4. وعند حساب مقدار الإنحراف المعياري لبيانات العيننة وجد أن $S=2.121$ ، من هنا إن مقدار خطأ المعاينة يحسب من خلال قسمة الإنحراف المعياري على الجذر التربيعي لحجم العيننة، ويتم تجاهل

المعامل $\left(\frac{N-n}{N} \right)$ باعتبار أنه سيكون مقدار صغير جداً عندما يكون حجم العيننة كبير.

بناء على ما سبق ان خطأ المعاينة هو:

$$\frac{2.121}{\sqrt{5000}} = \%3.0$$

2.1.5 التقدير بفترة الثقة

هي طريقة تعطي نطاقاً تقديرياً للقيم التي من المحتمل أن تتضمن معلمة المجتمع غير المعروف، حيث يتم حساب النطاق المقدّر من مجموعة معينة من بيانات العيننة، وبلغه أخرى فإن فترة الثقة هي النسبة المئوية للإحتمال أن معلمة المجتمع تم اختيارها مثلاً (بلغت نسبة التأكد بأن المعلمة تم اختيارها من المجتمع بصورة صحيحة بنسبة 90% أو 95%) وكدلالة إحصائية فإن هذه النسبة تعني أن المعلمة وبحسب الاختيار العشوائي تكون أقل من أو أكبر من الوسط الحسابي للمجتمع حسب التوزيع الطبيعي والتي تقابلها قيم

قياسية تسمى بقيم التوزيع الطبيعي القياسي مثلاً نسبة الثقة 90% تقابلها القيمة $Z=1.645$ مستوى الثقة 95% تقابلها $Z=1.96$.

2.5 اختبار الفروض

غالباً ما يتم اتخاذ قرارات في الحياة العملية حول المجتمعات بالاعتماد على معلومات من العينة. وتسمى بالقرارات الإحصائية ومن أمثلة ذلك تقرير فعالية دواء معين في علاج مرض ما بناء على بيانات العينة، أو ميزة أسلوب تدريس معين عن أسلوب آخر، أو أن تكون عملة معينة غير متوازنة.

1.2.5 الفرض والفرض الإحصائي

يعتبر الفرض إدعاء أو تخمين لقيمة شيء معين، بينما الفرض الإحصائي هو إدعاء أو تخمين بأن المتغير العشوائي يتبع توزيع معين (تخمين لقيمة معالم توزيع معين).

2.2.5 فرض العدم والفرض البديل

في محاولة للوصول إلى قرار فإنه من المفيد وضع افتراضات أو تخمينات حول المجتمعات ومن هذه الافتراضات التي قد تكون صحيحة أو غير صحيحة ما يسمى بالفروض الإحصائية وهي بصفة عامة عبارة عن قرارات حول التوزيعات الاحتمالية للمجتمعات، وكمثال لذلك لمعرفة ما إذا كانت قطعة عملة معينة غير متوازنة فإنه يتم وضع فرض أن قطعة العملة متوازنة بمعنى أن احتمال الحصول على الصورة هو $P = 0.5$ ، وبالمثل لمعرفة أن أسلوب ما أفضل من أسلوب آخر، فأنتنا نقوم بوضع الفرض أنه لا يوجد اختلاف N بين الأسلوبين (أي أن الاختلاف يرجع إلى التقلبات من نفس المجتمع) ومثل هذا الفرض يطلق عليه عادة بفرض العدم ويرمز له بالرمز H_0 . وأي فرض يختلف عن فرض العدم يسمى بالفرض البديل. ويعرف فرض العدم بأنه الفرض موضع الاختبار أي الذي يراد إتخاذ قرار بشأن رفضه أو قبوله (غالباً يكون الهدف رفضه)، ويرمز له بالرمز (H_0) ، أما الفرض البديل هو الذي يتم اللجوء إلى قبوله في حالة رفض فرض العدم، ويرمز له بالرمز (H_1) .

الفصل السادس

الارتباط والانحدار الخطي البسيط

يحدد الارتباط الدرجة والاتجاه الذي يرتبط به متغيران أما تحليل الانحدار فهو من التحليلات الإحصائية المهمة المستخدمة من قبل الباحثون الذي يصف ويوضح العلاقة بين المتغيرات في شكل معادلات، وفي الوقت الذي يمكن تحليل الانحدار بمتغير تابع واحد وعدة متغيرات مستقلة يمكن تحليل الارتباط بمجموعة من المتغيرات التابعة باستخدام مجموعة من المتغيرات المستقلة على حسب نوع بيانات البحث. من المهم التمييز بين النظرية الرياضية الكامنة وراء تحليل البيانات الإحصائية، والقرارات المتخذة بعد إجراء التحليل. إن فهم المخاطر وراء النتائج الإحصائية أمر بالغ الأهمية في عملية صنع القرار لذلك يجب دراسة العلاقة بين الظواهر من خلال تحليل البيانات الإحصائية عن طريق الارتباط والانحدار، والتي تغطي على وجه الخصوص كيفية اتخاذ القرارات المناسبة من خلال تطبيق تحليل البيانات الإحصائية. سنتطرق في هذا الفصل لكيفية معرفة الارتباط وطريقة رسمه وحساب معامل الارتباط ونموذج الانحدار الخطي البسيط وطريقة تقدير معامل.

1.6 الارتباط الخطي

أن الهدف من دراسة الارتباط هو الكشف عن قوة أو درجة العلاقة بين متغيرين أو أكثر، وتتراوح درجة العلاقة بين أي متغيرين يعبر عنها بمعامل الارتباط بين +1 و -1 ، فكلما كانت درجة الارتباط قريبة من 1 فأن ذلك يعني أن الارتباط قوياً بين المتغيرين، وكلما قلت درجة الارتباط كلما ضعفت العلاقة بين المتغيرين. وقد تتخذ العلاقة الارتباطية بين المتغيرين أحد شكلين (علاقة طردية، علاقة عكسية) عموماً يمكن اعتبار أن العلاقة ضعيفة اذا كانت قيمة معامل الارتباط أقل من 0.30، ويمكن اعتبارها متوسطة اذا تراوحت قيمة معامل الارتباط بين 0.30 إلى 0.70 أما اذا كانت قيمة معامل الارتباط أكثر من 0.70 فتعتبر العلاقة قوية بين المتغيرين. ومن الجدير بالذكر أن الارتباط يدل على وجود علاقة ما بين متغير و آخر، إلا أنه يجب أن ندرك بأن هذه العلاقة لا تدل على السببية، فهي لا تدل على وجود أثر لمتغير على آخر، فقد تكون هنالك علاقة طردية بين شرب القهوة ومعدلات الوفيات إلا أن شرب القهوة لا يعتبر سبباً في زيادة معدلات الوفيات بين الناس، فقد يكون هنالك عامل آخر كالتدخين مثلاً ينتج عن زيادة معدلات شرب القهوة ويؤثر في معدلات الوفيات فزيادة معدلات شرب القهوة تؤدي إلى الزيادة في التدخين مما يؤثر في زيادة معدلات الوفيات.

بطريقة أخرى يقال بأنه يوجد ارتباط بين متغيرين عندما ترتبط قيم متغير واحد بطريقة ما بقيم المتغير الآخر يحدد الارتباط إلى أي مدى يرتبط متغيران كميان X و Y ، فإذا كان هناك نمطاً في البيانات فيقال أن هناك ارتباطاً في البيانات. لرؤية هذا النمط، عندما ترتبط القيم العالية ل X بالقيم العالية ل Y ، يوجد

ارتباط إيجابي. عندما ترتبط القيم العالية ل X بالقيم المنخفضة ل Y ، يوجد ارتباط سلبي وتستخدم الكلمات "ضعيف" و "معتدل" و "قوي" لوصف قوة العلاقة بين المتغيرين. لقياس الارتباط بين المتغيرات يتم استخدام طريقة رسم الإنتشار (Scatter plot) و معامل الارتباط الخطي.

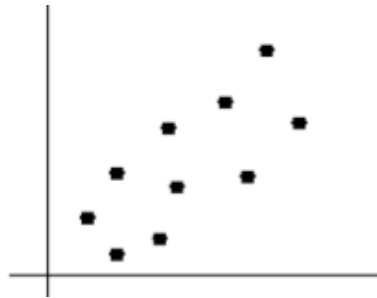
1.1.6 رسم الإنتشار

إن الخطوة الأولى لقياس الارتباط هي إنشاء مخطط للبيانات (Scatter plot). متوقع أن يكشف المخطط عن أنواع الارتباط التالية:

- ارتباط إيجابي ويعني أن قيم عالية ل X مرتبطة بقيم عالية ل Y .
 - ارتباط سلبي يقصد به قيم عالية ل X مرتبطة بقيم منخفضة ل Y .
 - لا يوجد ارتباط يقصد به أن قيم X ليست تنبؤية على الإطلاق بقيم Y .
- يتم توضيح هذه الأنماط في الأشكال التالية:

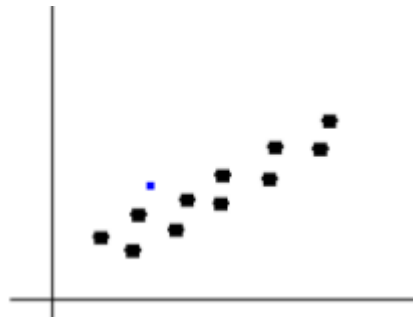
الشكل (1)

علاقة ارتباط موجب قوي



الشكل (2)

علاقة ارتباط موجب قوي



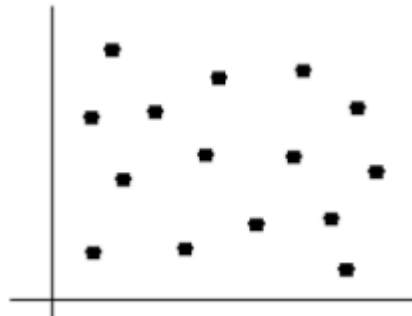
الشكل (3)

علاقة ارتباط سلبي



الشكل (4)

لا وجود علاقة ارتباط بين المتغيرات



2.1.6 معامل الارتباط

إن معامل الارتباط الخطي يصف قوة العلاقة الخطية بين المتغيرين. ويسمى أيضاً معامل ارتباط بيرسون وتكون المعادلة العامة كالتالي:

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

حيث يمثل n حجم العينة و x المتغير المستقل بينما y يمثل المتغير التابع.

ويرمز لمعامل الارتباط الخطي للعينة بالرمز r . أما رمز معامل ارتباط المجتمع الإحصائي هو ρ وتتراوح قيمته بين $(1-)$ و $(1+)$ حيث تدل علامة الموجب على العلاقة الطردية وتدل علامة السالب على العلاقة العكسية كما يوضح الجدول (1.2.1.6).

جدول رقم (1.2.1.6)
أنواع الارتباط واتجاه العلاقة

| تفسير القيمة | قيمة معامل الارتباط |
|-------------------|---------------------|
| ارتباط طردي تام | 1+ |
| ارتباط طردي قوي | 0.99 - 0.70 |
| ارتباط طردي متوسط | 0.69 - 0.50 |
| ارتباط طردي ضعيف | 0.49 - 0.01 |
| لا يوجد ارتباط | 0 |
| ارتباط عكسي تام | 1- |
| ارتباط عكسي قوي | 0.99- - 0.70- |
| ارتباط عكسي متوسط | 0.69- - 0.50- |
| ارتباط عكسي ضعيف | 0.49- - 0.01- |

هناك عدة أنواع لمعاملات الارتباط تستخدم باختلاف الحالة وحجم العينة وغيرها، كما موضح في الجدول (2.2.1.6).

جدول رقم (2.2.1.6)

| حالة الاستخدام | اسم معامل الارتباط |
|--|--------------------|
| العينات الكبيرة والبيانات الكمية | معامل بيرسون |
| عينات صغيرة وبيانات وصفية | معامل سبيرمان |
| قياس العلاقة بين ظاهرتين لكل منهما صفتين | معامل الاقتران |
| قياس العلاقة بين ظاهرتين لكل منهما أكثر من صفتين | معامل التوافق |
| يعني كم يفسر متغير من متغير آخر | معامل التحديد |

2.6 الإنحدار الخطي البسيط

يتعلق تحليل الإنحدار بدراسة العلاقة بين ظاهرة رئيسية وظاهرة أخرى أو مجموعة من الظواهر يعتقد أنها تفسر الظاهرة الرئيسية محل الدراسة، حيث تكمن أهمية الإنحدار والهدف الأساسي منه في التالي:

- تقدير الصورة الرياضية للعلاقة بين متغير مستقل ومتغير تابع.
- دراسة مدى تأثير متغير مستقل واحد أو أكثر على متغير تابع محدد.
- التنبؤ بقيم المتغير التابع اذا كانت قيم المتغير المستقل أو المتغيرات المستقلة معلومة.

ويجب أن تتوفر شروط أساسية لإجراء الإنحدار حتى تكون النتائج دقيقة ويمكن الوثوق بها، حيث ينبغي أن يكون توزيع المتغيرين المستقل والتابع توزيعاً طبيعياً، كما ينبغي أن تكون العينة مختارة بشكل عشوائي.

يتم صياغة العلاقة بين المتغير التابع والمتغيرات المستقلة في شكل نموذج رياضي إحصائي ويطلق على هذا النموذج نموذج الانحدار، كما في المعادلة أدناه:

$$Y_i = \alpha + \beta X_i$$

حيث أن Y_i هو المتغير التابع و X_i هو المتغير المستقل و α و β هي ثوابت تحدد العلاقة بين X و Y ، أي أن التغير في X يتبعه تغير في Y .

لفهم وتفسير الظاهر محل الدراسة وعلاقتها بالتغير في المتغير المستقل لابد من إيجاد قيم حسابيه للثوابت الموجوده في هذا النموذج وهي α و β .

يسمى هذا النموذج بالانحدار الخطي لأنه يعرض شكل العلاقة بين X و Y على أنها علاقة خطية. تتميز العلاقة الخطية بين المتغيرات الإحصائية بسهولة فهمها و دراستها وبالتالي سهولة تقدير معالم هذه العلاقة (α و β).

لتقدير معالم النموذج (α و β) لا بد من توفير عينة مشاهدات لقيم X و Y لأنه في الغالب لا يمكن حصر جميع المشاهدات المتعلقة ب X و Y في المجتمع.

فإذا تم تقدير هذه المعالم سيكون من الممكن ليس فقط في فهم العلاقة بين X و Y وإنما يمكن أيضاً استخدام النموذج في التنبؤ بقيم مستقبلية للمتغير التابع ما يساعد في التخطيط واتخاذ القرار بشكل علمي.

إذا تم تقدير قيم α و β من بيانات العينة فإنه يتوجب أيضاً فحص ما إذا كانت هذه القيم موجودة فعلاً في العلاقة الفعلية في المجتمع بمعنى قد لا تكون هناك علاقة أصلاً بين X و Y (بمعنى أن $\beta=0$) لإجراء هذا الفحص فإنه يلزم القيام بإجراء اختبار للمعنوية بالنسبة ل α و β للاستدلال على أن هذه المعالم غير مساوية للصفر وبالتالي يمكن الاستدلال على وجود علاقته بين X و Y .

1.2.6 طريقة المربعات الصغرى العادية لتقدير معالم النموذج الخطي البسيط

تمتاز طريقة المربعات الصغرى بسهولة الفهم و بالخصائص الجيده لمقدراتها، ويتم إيجاد قيمة الميل β وقيمة القاطع α حسب المعادلات التالية :

$$\hat{\beta} = \frac{n \sum xy - (\sum x)(\sum y)}{n (\sum x^2) - (\sum x)^2}$$

$$\hat{\alpha} = \frac{\sum y - \hat{\beta} \sum x}{n}$$

فمثلاً إذا تم تقدير معادلة خط الانحدار للوفيات خلال الأعوام كالآتي:

$$\hat{Y} = \hat{\alpha} - \hat{\beta}(x)$$

$$\hat{Y} = 34.78 - 1.7(x)$$

تقرأ المعادلة أن 34.78 هي قيمة الوفيات المقدرة عندما يظل متغير الزمن (الأعوام) ثابت، وأن معدل التغير في عدد الوفيات بالنسبة لمتغير الزمن (الأعوام) هو (-1.7)، وهذا يعني أنه كلما تغيرت قيمة الأعوام بوحدة واحدة تتناقص قيمة أعداد الوفيات من الحوادث المرورية بمقدار 1.7.

المراجع

المراجع العربية

- أساسيات التحليل الإحصائي (3 أساسيات مهمة) - هدير عبد الحميد-سندك للاستشارات الأكاديمية والترجمة -2022
<https://www.sanadkk.com/blog/post/529/%D8%A3%D8%B3%D8%A7%D8%B3%D9%8A>
- دليل مبادئ التحليل الإحصائي- أدلة المنهجية والجودة-دليل رقم (10)- مركز الإحصاء أبو ظبي
<https://scad.gov.ae/documents/20122/0/1692026477992>
- دليل توجيهات رفع المعرفة -مركز عجمان للإحصاء-2021
<https://scc.ajman.ae/sites/default/files/publications/%D8%AF%D9%84%D9%8A.pdf>
- التقدير واختبارات الفروض – موقع SCRIBD
<https://www.scribd.com/document/559802993/%D8%B4%D8%B1%D8%AD-%D8%A7%>
- التحليل الإحصائي-قسم الاساليب الكمية – كلية إدارة الأعمال-جامعة الملك فيصل
<https://www.scribd.com/presentation/537839254/%D8%AA%D9%88%D8%B2%D9%8A>
- مبادئ الإحصاء والإحتمالات -صلاح علي مبخوت-جامعة ذمار
https://www.researchgate.net/publication/326327243_mbady_alahsa_walahtmalat
- مبادئ الطرق الإحصائية – الدكتور جلال الدين الصياد – الدكتور عبد الحميد محمد ربيع
<https://books-library.net/free-88676369-download>

المراجع الإنجليزية

- Coconino Community College- Kate Kozak
https://www.coconino.edu/resources/files/pdfs/academics/sabbatical-reports/kate-kozak/chapter_10.pdf
- San Jose State University- Stat Primer (Version 7.0)- B. Burt Gerstman 2003, 2006, 2016
<https://www.sjsu.edu/faculty/gerstman/StatPrimer/correlation.pdf>
- Stanford University - Introduction to Statistical Inference-2016
<https://web.stanford.edu/class/archive/stats/stats200/stats200.1172/Lecture18.pdf>