الفصل العاشر

مبادئ وأهداف وطرق الدراسات الإحصائية
طبيعة الإحصاء Nature of Statistics

10.1. مقدمة Introduction

يلعب علم الإحصاء دورا هاما فى مختلف نواحي العلوم وبخاصة العلوم الهندسية والصناعية وعلوم الإقتصاد والإدارة والتسويق، و يطبق هذا العلم فى جميع أوجه النشاط البشرى مثل تحديد القصور فى مستوى جودة المنتجات وتحسين مستوى خدمة العملاء وتحديد مستوى وجودة المنتجات والخدمات والأسعار طبقاً لآليات ومتطلبات السوق، ويمتد استخدام تطبيقات علم الإحصاء إلى كثير من العلوم مثل الأحياء والزراعة والطب والفيزياء والكيمياء والفضاء والعلوم السياسية وعلم النفس وعلم الاجتماع والتربية والعديد من المجالات الحيوية الأخرى.

والأصل اللغوي فى اسم هذا العلم (Statistics) مشتق من كلمة (State) أى الدولة ومعناه "البيانات والحقائق المتعلقة بشئون الدولة"، وعلم الإحصاء (Statistics) هو ذلك العلم المعنى بقواعد تجميع ووصف وتحليل وتفسير و استقراء ومعالجة وترجمة البيانات الرقمية.

Statistics is the science compromising rules and procedures for collecting، describing، analyzing، and interpreting numerical data.

و يخطئ من يظن أن الإحصائي (Stasitician) هو ذلك الشخص الذي يقوم بعد وحساب نقاط مباراة كرة سلة، أو الشخص الذي يقوم بقراءة النشرة الجوية ويقوم بسرد مجموعة من الأرقام، إن الأمر أبعد وأعمق من ذلك بكثير، فالإحصائى هو هذا الشخص الذي يجمع و يصنف و يصف ويحلل ويفسر ويترجم ويتحاور مع الأرقام للوصول إلى نتائج محددة و موثوق بها.

ولا يكاد يخلو يوم إلا ونسمع فيه نتيجة إحصائية مثل :

- تشير النتائج المالية إلى أن أرباح الشركة هذا العام ستزيد بنسبة 14% عن العام السابق.
- من المتوقع فى ضوء البيانات المتاحة تقدم الحزب المعارض في الانتخابات المحلية الحالية.
- لقد تضاءلت فرصة إدارة المشروع فى إنهاء مشروعهم فى التاريخ المحدد سلفا.
- من المتوقع أن تكون درجات الحرارة حول معدلها الطبيعي فى مثل هذا الوقت من كل عام.

فكل هذه الأمثلة ما هي إلا نتائج لبيانات كثيرة جداً، تم تجميعها وتحليلها واستخلاص هذه الاستنتاجات والنتائج منها.

10.2. لمحة تاريخية عن الإحصاء Historical Tip

يرجع استخدام الإحصاء – بمعنى الحصر والعد- للقدماء المصريين الذين قاموا بتعداد وحصر لسكان مصر وثرواتها، واستخدموا هذه النتائج فى تنظيم مشروعهم لبناء الأهرامات، كما استخدمته سلطات الدولة الرومانية فى القرن الأول الميلادي في حصر السكان والأراضي لأغراض جمع الضرائب والمكوس، ثم تطور ليشمل المواليد والوفيات والزواج...

ومع ازدياد كم البيانات المتداولة بصورة كبيرة - وخاصة في القرن التاسع عشر- ومع ازدياد صعوبة التعامل معها، كانت أولى خطوات الاعتماد على أخذ عينات Samples فى محاولة أولية لكشف وإدراك خصائص البيانات الكلية أو المجتمع Population، ثم إجراء المعالجات اللازمة على هذه العينة، ثم إسقاط نتائج هذه المعالجة علي المجتمع مرة أخرى.

10.3. استخدامات الإحصاء في الأعمال Uses of Statistics in Business

في الماضي، ولما كانت الأعمال صغيرة Small Business كان يمكن للمديرين والقائمين علي هذه الأعمال حل مشاكلهم بسهولة ويسر من خلال معالجة البيانات المتاحة لديهم، ولكن مع إزدياد حجم الأعمال وازدياد تعقدها، كان علي هؤلاء المدراء تجميع البيانات وتصنيفها ووصفها وتحليلها واستقرائها للمساهمة في اتخاذ القرارات السليمة، وهذا ما أدي لازدهار علم الإحصاء.

ولما كان هذا العلم مفيدا و قادرا علي تحليل البيانات واستنباط الأحكام واستخلاص القرارات، فقد زاد الاهتمام به و الاعتماد عليه.

ومن المجالات التي تستخدم علم الإحصاء ما يلي :-

- التحكم ومراقبة الجودةQuality Control: فباستخدام الطرق الإحصائية يمكن زيادة الجودة وتحسين الإنتاجية.
- تخطيط الإنتاج Product Planning فمن خلال تحليل العوامل الإقتصادية والاحتياجات يمكن تخطيط الإنتاج والمخزون والمبيعات.
- التنبؤ Forecasting وذلك للتنبؤ بحجم الإنتاج المستقبلي وتوقع قيمة المبيعات ونوعيتها.
- التقارير السنوية Yearly Reports فمعظم الذين نقدم لهم التقارير السنوية لا يتقنون فهم الأرقام البحتة، إنما يجيدون فهم واستقراء البيانات المعالجة كالنسب والخرائط والإتجاه العام.
- إدارة أداء العاملين Personnel Performance Management : فالإحصاء يساهم في تقييم الأداء وقياس الإنتاجية والفعالية والمشاركة.

- دراسة السوق Market Research وفيه يتم دراسة الأسواق واحتياجاتها وسلوك المشترين وعاداتهم ومتطلباتهم.
- الدراسات الاجتماعية Social Studies وبخاصة فى تحليل القضايا السياسية وبحث الظواهر الاجتماعية.
- التحليل المالى Financial Analysis لتوظيف رؤوس الأموال وتوجيهها والبحث عن آفاق جديدة للاستثمار.
- تصميم التجارب العلمية Design of Experiments لإنتاج سلالات محسنة ومتميزة من المحاصيل الزراعية و كذلك لإنتاج أدوية جديدة لمكافحة الأمراض والأوبئة.

10.4. فروع علم الإحصاء Branches of Statistics

وينقسم علم الإحصاء إلى فرعين أساسيين هما:

أ- الإحصاء الوصفي Descriptive Statistics

وهو فرع يعنى بجمع البيانات وتنظيمها وتصنيفها وعرضها عن طريق الجداول أو الرسوم البيانية بهدف توضيح وعرض الخصائص الحالية للبيانات المتاحة سواء كانت عينات Samples أو مجتمعاً Population.

ب- الإحصاء الاستنتاجي/ الإستدلالى Inferential Statistics

وهو فرع يعنى بتحليل البيانات واستنتاج واستقراء واستنباط الخصائص المستقبلية للمجتمع Population من خلال دراسة خصائص العينة Samples وهدفه اتخاذ القرارات وبناء التقديرات المبنية على هذه البيانات وهو ما نطلق عليه التقديرات والقرارات المرتكزة والمبنية على البياناتData Driven Estimations and Decisions ، وهذا ما نطمح إليه جميعاً وهو أحد الأركان القوية التي ترتكز عليها منهجية التطوير المستمر Six Sigma Methodology.

و تجدر الإشارة إلى أنه من المحتمل أن تكون الاستنتاجات التي نتوصل إليها غير دقيقة، لأننا نعتمد فى التوصل إلى الاستنتاجات – وبالتالي اتخاذ قرارات معينة فى ضوئها – على فروض واعتبارات Hypothesis تحت ظروف غير مؤكدة، وفى عبارة أخرى، فإن الظروف التى تأخذ فيها البيانات ديناميكية وتتغير بإستمرار، ولكن بالرغم من عدم التأكد من دقة الاستنتاجات والقرارات، فإن استخدام الإحصاء الاستنتاجى فى التوصل إليها يعتبر أفضل من عملية التخمين المطلق، ذلك لأن الإحصاء بطرقه الاستقرائية المتنوعة والمتعددة، قادرٌعلى تحديد وتقدير مدى ثبات وصحة هذه الاستنتاجات واحتمال حدوثها عن طريق عوامل الصدفة، ولذلك فالاستنتاج الاستدلالي يزودنا بالأساسيات اللازمة للتعرف على مدى

الخطأ فى استنتاجات حول مجموعة بيانات عندما نعتمد فى توصلنا إلى هذه الاستنتاجات عن طريق الاعتماد على جزء صغير من هذه البيانات، وهو بذلك يسهم فى تقييم النتائج والاستنتاجات التي نحصل عليها من التجارب والدراسات(مجدى عبد الكريم 1996).

إذن فالإحصاء علم يهتم بطرق جمع البيانات وتمثيلها وعرضها (الإحصاء الوصفي) ومن ثم تحليلها و تفسيرها والتوصل إلى الاستنتاجات (الإحصاء الاستنتاجى).

أى أننا عندما نرغب فى القيام بدراسة ما فإننا نقوم بجمع بيانات لاستخدامها فى الإجابة عن أسئلة معينة أو إختبار فروض محددة. وعندئذ نستخدم عينة من المجتمع ونجمع بيانات عنها ثم نحاول تعميم نتائج العينة إلى المجتمع وهذا ما يطلق عليه بالإستدلال الإحصائى Statistical inferenceومعنى هذا أن الهدف الاسمى كما قلنا هو تعميم نتائجها على المجتمع. والاساليب الاحصائية التى تستخدم لاختبار صحة الفروض والاستنتاج عن خصائص المجتمع من بيانات العينة هى اساليب الإحصاء الاستدلالى

ج- اساليب الإحصاء الاستنتاجي/ الإستدلالى

1- اختبار المقارنة بين متوسطى مجموعتين (اختبار ت t-test).
2- اختبار الفروق بين متوسطات عدة مجموعات (تحليل التباين ANOVA).
3- الارتباط المتعدد للتنبؤ بدرجات متغير من عدة متغيرات أخرى.
4- أساليب الاحصاء اللابارامترى مثل : مربع كاي Chi square ، مان-ويتنى Mann-whitney ، ويلكوكسون Wilcoxon ، كروسكال واليس -Kruskal Wallis وغيرها

وعموما فإن الاساليب الاحصائية الاستدلالية تستخدم قيم حرجة لاختبار الفروض الصفرية (يرجى مراجعة الفصل السابع عشر من هذا الكتاب) ، فإذا كانت القيم الحرجة أعلى من القيم الجدولية فإن احتمال وجود الفروق (أو العلاقة) يكون مرتفعاً (أو دالاً) ، وعليه يتم رفض الفرض الصفرى عند مستوى دلالة معين ، ومستويات الدلالة المقبولة هى 0.05 أو أقل، ويعرف مستوي الدلالة Alpha α بأنه إحتمال الخطأ فى رفض الفرض الصفرى الصحيح (وهو ما يطلق عليه خطأ النوع الأول).

وتشترط الاساليب الاحصائية البارامترية ما يلى:

1- العشوائية فى اختيار العينة المستخدمة فى الدراسة، وإلا فإن الاستدلال يكون متحيزاً، وصعب حينئذ تعميم النتائج على المجتمع.
2- إعتدالية توزيع العينة ، وتقل أهمية هذا الشرط بزيادة حجم العينة حيث تيقترب توزيع العينات الكبيرة (100 فأكثر) من الاعتدالية، أما فى حالة العينات الصغيرة فإن الامر يتطلب تحويل العينات (يرجى مراجعى الفصل الثامن من هذا الكتاب).

10.5. مصادر البيانات Sources of Data

تعتمد النتائج المستخلصة من بيانات العينات Samples بشكل كبير جداً على دقة وجودة بيانات هذه العينات، فكلما كان جمع البيانات دقيقا زادت ثقتنا في الاعتماد على النتائج المستخلصة منها، ولذا يتحتم الإجابة على عدة أسئلة للتأكد من مدى جودة هذه البيانات قبل التسليم بالنتائج:

- هل بيانات هذه العينات دقيقة؟
- هل هذه العينات، وبالتالي البيانات المأخوذة منها تمثل خصائص المجتمع Population فعلاً؟
- هل تم تجميع بيانات العينات كما ينبغي؟

وقبل الإجابة على هذه الأسئلة يجب أولا البحث في مصدر هذه البيانات والذي ينقسم إلى مصدرين:

أ- مصادر البيانات الأولية Primary Data Source

وهي البيانات التي يتم تجميعها لأول مرة من أجل موضوع محدد، من خلال الأسئلة المباشرة في صورة إستبيان Questionnaire، أو من خلال تصميم خطة لجمع بيانات محددة عن موضوع معين، وهذه البيانات لم تكن موجودة من قبل، وهذا النوع من البيانات يحتاج لمجهود كبير وغالبا ما يصاحبه تكاليف كبيرة.

ب- مصادر البيانات الثانوية Secondary Data Sources

وهي البيانات التي تم تجميعها من قبل لغرض آخر غير غرض البحث الحالي، وغالبا ما تكون بكميات كبيرة، وفي بعض الحالات قد لا تكون مناسبة بصورتها الراهنة للتعامل معها، ومن أمثلة ذلك التقارير الحكومية والدولية والتقارير المالية وبيانات الأرشيف.

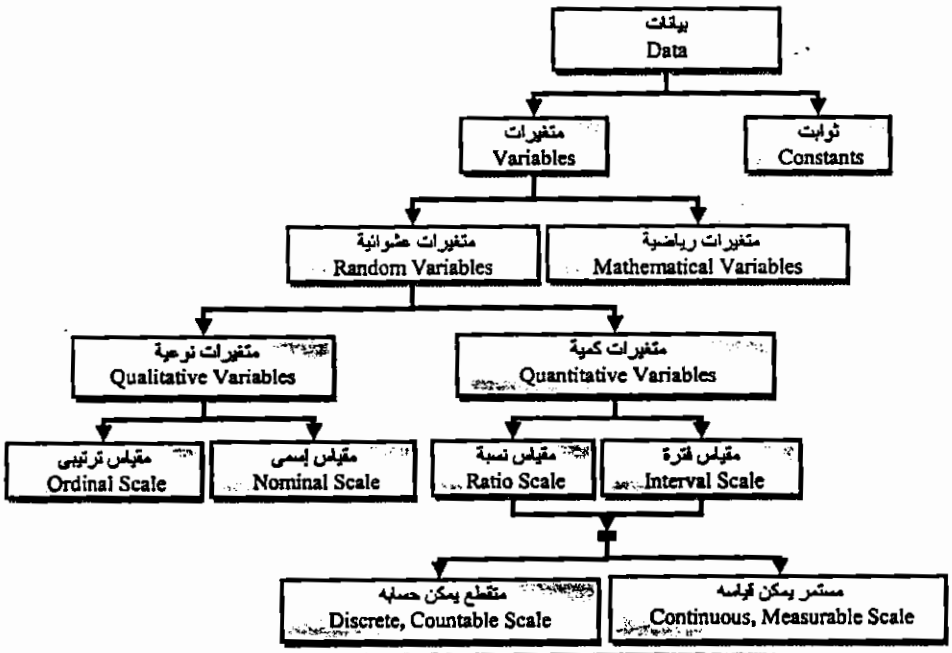
وعيوب هذه البيانات أنها قد تكون كبيرة، وقد تكون قديمة، وغالبا ما تحتاج لمعالجة قبل استخدامها، وقد لا نعرف الظروف المصاحبة لتجميعها، و يندرج تحت هذا النوع أيضا تلك البيانات التي كانت في الأصل أولية Primary ثم تمت معالجتها وتحويلها إلى صورة أخرى، وعموما فمن الأفضل التعامل مع البيانات الأولية Primary Data.

هذا و يوجد تصنيف آخر لأنواع البيانات Data (إلا أنه تصنيف غير شائع) وهو كما يلي:

- بيانات ثابتة Static Data مثل العينات التي تؤخذ مرة واحدة، ولا تتغير خصائصها بمرور الوقت مثل المتوسط السنوى لدرجات الحرارة والرطوبة ومعدل سقوط المطر فى منطقة ما.
- بيانات ديناميكية Dynamic Data مثل القراءات التي تؤخذ بصفة دورية (يومية- أسبوعية – شهرية-.....)، ويقال أنها بيانات حيه Live Data أى متجددة، ومن خصائصها أننا لو حسبنا المتوسط كل مرة فسوف نجده بقيمة تختلف عن المرة السابقة والمرة التالية.

10.6. أنواع البيانات الرقمية Numerical Data Types

يوجد نوعان رئيسيان من البيانات Data وهما الثوابت Constants و المتغيرات Variables، والشكل 10-1 التالي يعطينا فكرة عامة عن أنواع البيانات الرقمية، وسنتناول هذه الأنواع والأقسام بشئ من التفصيل فى الصفحات القليلة القادمة.



شكل رقم 10- 1 الأنواع المختلفة للبيانات

أ- الثوابت Constants

هي السمات والخواص التي لا تتغير بمرور الوقت وهى تصف ماهية المواد فى ظروف معينة مثل الكثافة النوعية ومعامل التمدد ومعامل الاحتكاك لسطح ما، وقوة الجاذبية و العدد الذرى، ومثل باقي الثوابت الهندسية والفيزيائية والكيمائية.

ب- المتغيرات Variables

هي السمات والخواص أو الصفات التي تتغير قيمتها من عنصر لآخر بمرور الوقت مثل: ضغط الدم و مستوى السكر فى الدم و كمية نزول المطر و حجم المبيعات والمخزون ودرجة الحرارة وزمن الانتظار وزمن التشغيل وعدد الموظفين ونسبة الإنجاز وأسعار المواد والأجور.

والمتغير هو أى ظاهرة أو حدث أو خاصية تأخذ قيما تتغير من ظرف لآخر،أما المتغير العشوائي Random Variable هو متغير حقيقي يخضع لمؤثرات غير منظورة مثل درجة الحرارة والضغط والتآكل والإجهاد ولذلك لا نستطيع التحكم فيه تجريبيا، وهذا النوع من المتغيرات يكون له توزيع احتمال، أى أن حدوثه أو ظهور أى عنصر من عناصره يكون مصحوبا باحتمال ما مثل معدل درجات الحرارة ومعدل الوفيات.

أما المتغير غير العشوائي أو الرياضي Mathematical Variable فليس له توزيع احتمال، ولا يخضع لتأثير عوامل عشوائية، وإنما يخضع لقانون رياضي، ويمكن التحكم فيه تجريبيا، أو تحديد قيمته مقدما، أو قياسها بدقة، مثل الزمن اللازم لبلوغ مسافة معينة عند السير بسرعة منتظمة.

و تنقسم المتغيرات العشوائية إلى قسمين رئيسيين:

أ- المتغيرات الوصفية (النوعية) Qualitative Variables

وهى متغيرات لا يمكن قياس مفرداتها عدديا، ولكن يمكن تقسيم هذه المفردات بحسب اشتراكها فى صفة أو خاصية ما، ويمكن تقسيم هذا النوع إلى نوعين فرعيين كما يلي:

1. المقياس الاسمى أو التدريج الاسمى Nominal Scale

يعد أقل وأضعف مستويات القياس، وهو مجرد تقسيم أو تصنيف بالاسم فقط مثل تقسيم الأشخاص حسب الجنس (ذكور – إناث)، وحسب الجنسية (مصري - سوداني- أردنى- فرنسي - برازيلي..)، وتقسيم المشروعات إلى (مدني – كهرباء- ميكانيكا – صحي...)، وتقسيم الكتب بالمكتبة حسب الموضوع إلى (دين- فلسفة – اجتماع- علوم هندسية...) وهكذا، إذن فهو مقياس يستخدم لتحديد هوية العناصر المقاسة اسميا، وبالتالي فهو مقياس يصنف العناصر التي تختلف فى النوعية لا فى الكمية، وفى هذا النوع نجد أنه من الصعب ترتيب المفردات حسب الأهمية أو الأفضلية، وقد تستعمل الأرقام والأعداد للمساعدة فى تحديد هوية العناصر، ومثال ذلك ذكور= 1& إناث =2 أوأحمر=2 & أخضر = 3 & أزرق = 5، وفى هذه الحالة لايكون للعدد ذلك المدلول الكمي، مثل أعمار الذكور والإناث

فى المثال السابق، الذي يفهم منه عادة التصنيف ، ولا يجرى عليها عمليات حسابية مثل الجمع والطرح.

2. المقياس الترتيبى Ordinal Scale

وهو أفضل من التدريج السابق حيث يتم التصنيف و التقسيم فيه على أساس الرتبة أو الأهمية النسبية، مثل درجات الطلاب على أساس ممتاز – جيد جداً – جيد مقبول، أو توزيع السكان حسب الحالة التعليمية: أمى – ابتدائي – اعدادى – ثانوي، وفى هذا القياس يمكن ترتيب القيم وإجراء المقارنات، حيث يمكن القول بأن الحاصل على تقدير جيد يكون مستوى تحصيله أفضل من الحاصل على تقدير مقبول، وبالرغم من أننا نستطيع الترتيب والمقارنة إلا أننا لا نستطيع تحديد مقدار الفروق بين القيم.

إذن فهو مقياس يستخدم للتصنيف والمفاضلة مثل نتائج الامتحانات / B=Second A=First أو 1=First & 2=Second أو كما فى الرتب العسكرية، وفي هذا النوع من البيانات الذي يهم هو ترتيب الأرقام أي1 سيكون أفضل من 2 بينما الفرق بينهما أو مجموعهما لا يعني شيئا علي الإطلاق.

والنوعان الاسمى Nominal و الترتيبى Ordinal هما نوعان يتضح من اسمهما أنهما في الأصل ليسا قيما، ولكن فى بعض الأحيان نحولهم إلي قيم لتسهيل التعامل فقط، ففي المقياس الاسمى Nominal مثلا نحولهم إلى قيم كما يلي Yes=1 & No=2. و في المقياس الترتيبى Ordinal نحولهم إلي قيم مثل Accepted= 1 & Good=2 & Best = 3، ولكن فى كل الحالات يجب الانتباه إلى أنه لا يمكن إجراء عمليات حسابية مثل الجمع والطرح، و لا نستطيع تحديد مقدار الفروق بين هذه القيم.

ب- المتغيرات العددية (الكمية) Quantitative Variables

وهى متغيرات يمكن التعبير عنها بالأعداد الحقيقية إما بالعد أو بالقياس، ويكون التباين فيها هو تباين من حيث المقدار، أى يمكن تقسيم مفرداتها بحسب الأصغر والأكبر، ويمكن تقسيم هذا النوع إلى نوعين فرعيين كما يلي:

1. المقياس الفترى Interval Scale

وهذا المقياس يعد أقوى من السابقين، حيث نستطيع تحديد الفروق بين القيم مثل: درجات الحرارة، ودرجات الاختبارات الرقمية، وعدد ساعات العمل الإضافى، ووحدة القياس فيه واحدة، أى أن المسافة على المقياس أو التدريج أو الميزان بين أى درجتين هي نفسها

المسافة بين أى درجتين أخريين مثل الفرق بين درجتى الحرارة "30" و "31" درجة مئوية هي نفس الفرق بين درجتى الحرارة "39" و "40" درجة مئوية، ولكن يؤخذ على هذا القياس عدم وجود نقطة الصفر المطلق، بمعنى أن الصفر هنا لا يعنى انعدام الخاصية، وبالتالي لا نستطيع إجراء النسبة بين القيم المقاسة، فمثلاً لا نستطيع القول بأن كمية الحرارة عند درجة الحرارة 20 تساوى ضعف كمية الحرارة عند درجة الحرارة 10، أو أن الطالب الحاصل على 10 درجات يكون مستواه فى التحصيل مساويا لخمسة أمثال الحاصل على درجتين فقط.

2. المقياس النسبي Ratio Scale

ويعد أقوى المقاييس الرقمية حيث يسمح بإيجاد النسب بين قيم المتغيرات، مثال ذلك الأوزان والأطوال ودرجات الحرارة بالكلفن Kelvin والسرعة، ويمكن التفريق بينها وبين الفترة Interval في أنها تحتوي علي قيمة صفرية مطلقة، أى أن الصفر على هذا التدريج يعنى العدم (للدكتور مصطفى 1989)، بمعنى إنعدام الظاهرة أو الشيئ أو عدم وجوده وللتفريق بين النسبة Ratio و الفترة Interval اسأل نفسك هل لو ضاعفت البيانات Data هل سيؤدي إلي مضاعفة تأثير القيمة أم لا ؟ فإذا كانت الإجابة نعم تكون بيانات نسبة Ratio Data وإذا كانت الإجابة لا تكون بيانات فترة Interval Data، ففي درجات الحرارة، هل 100 درجة تساوى ضعف الخمسين درجة والإجابة لا فيكون نوع البيانات Data هي بيانات فترة Interval .

هذا ويلاحظ أننا قد عرضنا المقاييس الرقمية الأربعة السابقة بالترتيب حسب قوتها، بحيث يحمل كل مقياس مزايا المقاييس التي تسبقه بالإضافة إلى مزايا أخرى، وفى هذا الصدد يجب الإشارة إلى نقطتين هامتين:

- كلما زاد مستوى القياس للمتغيرات، كلما زادت الدقة فى القياس، وكلما أمكن استخدام أساليب إحصائية على درجة أفضل.
- المتغيرات بمستوى قياس معين يكون التعامل معها بالأساليب الإحصائية المخصصة لهذا المستوى من القياس، كما أنه يمكن أيضا استخدام الأساليب الإحصائية المخصصة لمستويات القياس الأقل.

والجدول 10-1 يعرض ملخصا لما سبق ذكره

جدول رقم 10 - 1 ملخص لمقاييس البيانات

Feature	إسمى	ترتيبى	فترى	نسبة	الخاصية
Order data is meaning full	N	Y	Y	Y	هل الترتيب يعنى شيئا؟
Differences between data values is meaningful	N	N	Y	Y	هل الفرق بين البيانات يعنى شيئا؟
Zero point represent total absence	N	N	N	Y	هل الصفر يعنى العدم ؟

هذا وتنقسم المتغيرات الكمية أو العددية التي يمكن التعبير عنها بالأعداد الحقيقية إما بالعد أو بالقياس، إلى قسمين كما يلي:

أ- المتغيرات المتصلة Continuous Variables

وهى المتغيرات التي نحصل عليها عن طريق القياس Measurement وبأى درجة من الدقة، مثل: الطول والوزن والعمر ودرجة الحرارة و شدة الزلزال و الزمن الذي تنتظره سيارة عند إشارة ضوئية و مثل الأرقام 1.5 & 2 & 4.25 & 1.75 & 1.82 & 5.253 & 3.21 أي إذا كنت أستطيع قياسها كل لحظة فهي Continuous، وفى هذه المتغيرات نستطيع أن نتصور أن المفردات المقاسة يمكن أن تختلف عن بعضها بمقادير صغيرة صغرا لا نهائيا من الناحية النظرية، وإذا تم توقيع وتمثيل قيم متغير متصل Continuous Variable بين قيمتين مثل X و Y فإنها تمثل بيانيا بجميع نقط قطعة مستقيمة محدودة بهاتين القيمتين، ويمكن تشبيه البيانات المتصلة Continuous Data بخرطوم المياه الذي تتدفق منه المياه.

ب- المتغيرات الوثابة Attribute (Discrete) Variables

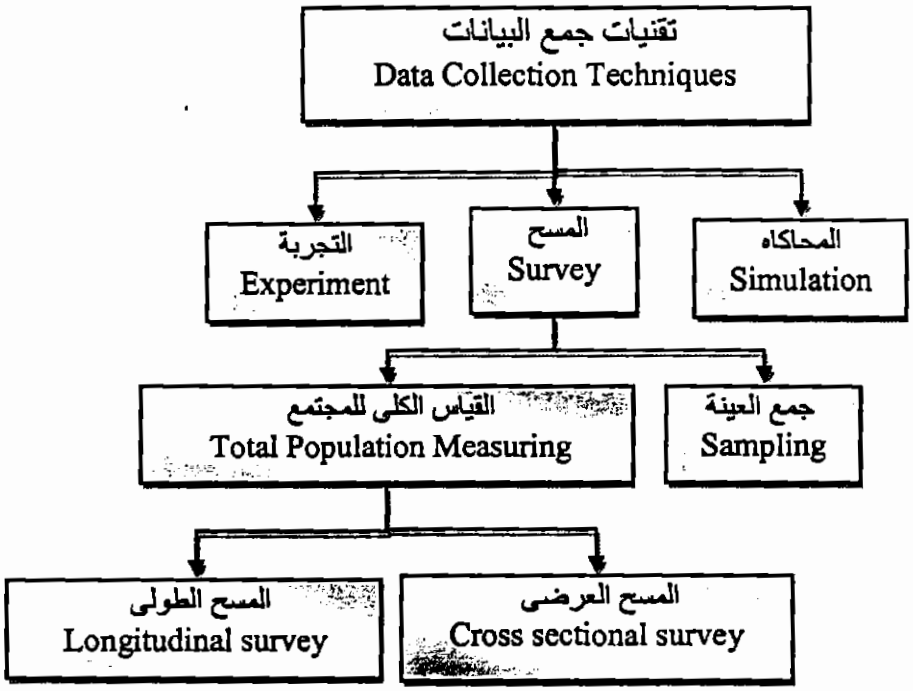
وهى المتغيرات التي نحصل عليها عن طريق العد Counting، مثل عدد أفراد الأسرة، وعدد الطلاب بالفصل، وعدد المشروعات فى السنة، وعدد ضربات القلب وعدد المنتجات المعيبة وعدد مرات أعطال المعدات، وإذا تم توقيع وتمثيل قيم متغير وثاب Attribute (Discrete) Variable بين قيمتين مثل X، Yفإنها تمثل بيانيا بنقاط منفصلة على خط مستقيم، وقيم هذه المتغيرات ليس فيها صفة الإستمرارية أي أنها متقطعة أو وثابة مثل 2.5 1.5 & 2 & 4 &قبرغم وجود عدد عشري مثل 1.5 فإن ذلك لا يعنى أنها مستمرة إذ لو كانت مستمرة لوجدنا في نفس مجموعة الأرقام السابقة مثلا 1.23 أو5.5612، ويمكن تشبيه

بيانات المتغيرات الوثابة Discreet Data بطلقات المسدس أو بحركة عقرب الثواني الذي يقفز، فهي حركة متقطعة.

10.7. تقنيات وطرق جمع البيانات Data Collection Techniques

إن من أهم مراحل بحث أي مشكلة هي مرحلة جمع المعلومات والبيانات عن هذه المشكلة، و قبل البدء في تجميع البيانات يجب تجهيز خطة واضحة المعالم تشمل ماذا سيتم تجميعه من بيانات، وكيف سيتم ذلك، ومتى ستبدأ عملية الجمع ومتى ستنتهي، ومن سيقوم بالعملية، ومن أين سيتم التجميع.

وعموما فإنه يتم جمع البيانات من خلال عدة طرق كما يتضح فى الشكل 10-2 التالى:



شكل رقم 10- 2 تقنيات جمع البيانات

أ- التجربة Experiment

تتميز التجربة فى كثير من الأحيان بدقة النتائج لما فيها من حرية وتحكم فى المتغيرات، ويلزم لإجرائها قدر من الخبرة بمجال التجربة، غير أنها قد تكون مكلفة فى بعض الأحيان، وهى وسيلة فعالة لفهم العلاقة بين المتغيرات قيد الدراسة.

ب- المحاكاة Simulation

فى بعض الحالات تكون عملية جمع بيانات معينة عملية صعبة أو غيرأخلاقية أو تتنافى مع العرف والتقاليد والموروث، وهنا يتم عمل نماذج يطلق عليها النماذج التمثيلية Prototype والتي ينبغي أن تشبه إلى حد كبير جداً العمليات التي نريد بحثها ودراستها، وهى تحتاج إلى مهارات وتقنيات وتكاليف عالية للوصول إلى نموذج يقترب فى أدائه من العملية الأصل، كذلك يحتاج إلى تهيئة الجو العام والظروف المحيطة بالنموذج وقت التجربة إلى جو وظروف يناظران الظروف الحقيقية بنسب معينة ومحسوبة، و يمكن عن طريق المحاكاة توليد البيانات اللازمة للدراسة باستخدام الكمبيوتر اصطناعيا Artificially، ومن أهم طرق المحاكاة الشائعة تلك المعروفة باسم مونت كارلو، وهى تعتمد على العينات العشوائية والتوزيعات الاحتمالية.

ج- المسح Survey

و عند تجميع البيانات بأسلوب المسح فإنه يتم التعامل مع الزمن بعدة أساليب منها:

1. المسح العرضي Cross Sectional Survey: إذا تم تجميع البيانات عند نقطة زمنية محددة أو فى تاريخ محدد.
2. المسح الطولي Longitudinal Survey: إذا تم تجميع البيانات عن فترة زمنية معينة، قد تمتد فى الماضي أو فى المستقبل، وهذا النوع ينقسم إلى عدة أنواع أخرى:

- تحليل التوجه Trend Analysis: دراسة ظاهرة محددة فى أوقات زمنية مختلفة تسمى دراسة الاتجاه أو Trend Analysis، مثل دراسة ارتفاع أسعار البترول خلال الفترة من يناير 1992 حتى فبراير 1995.
- دراسة الفوج Cohort Study: إجراء نفس الدراسة على عينات مختلفة من نفس المجتمع عبر فترات زمنية مختلفة تسمى دراسات الفوج Cohort Studies، مثل دراسة ارتفاع الإصابة بمرض السكر بين المدخنين خلال الأعوام 1994، 1996، 1999.
- دراسة الشريحة Panel Study: إجراء نفس الدراسة على نفس العينة عبر فترات زمنية مختلفة تسمى دراسات الشريحة أو Panel Studies، مثل دراسة تأثير عقار طبي على مجموعة محددة من المرضى.

وخلال إجراء كل عمليات المسح Survey السابق ذكرها يتم الاعتماد على طريقتين أساسيتين، فإما أن يتم تجميع كل البيانات من خلال الحصر الشامل للمجتمع، أو أن يتم أخذ عينات كما يتضح من الشرح التالي:

- **الحصر الشامل للمجتمع Total Population Measuring**

وتتم إذا كان من الممكن قياس كافة عناصر المجتمع Population محل الدراسة مثل قياس أوزان جميع الطلاب فى إحدى المدارس، ولا يكون ذلك إلا إذا كان المجتمع صغير نسبيا، وهذا النوع محدود التطبيق ومكلف وغالبا ما يحتوى على أخطاء بسبب كبر حجم البيانات ويستغرق وقتا طويلا.

- **أخذ العينات Sampling**

وهو الأسلوب الأكثر شيوعا، وفيه نلجأ إلى دراسة جزء من كل، أى نجمع بيانات عن بعض العينات Samples وليس كل عناصر المجتمع ، ويتم ذلك في الحالات التي يتعذر فيها قياس كل عناصر المجتمع إما لكثرتها كما في قضايا السلوك الإنساني، أو لاستحالتها كما في الاختبارات الاتلافية، وعملية أخذ العينات Sampling أقل كلفة من المسح الشامل وتستغرق جهدا ووقتا أقل، كما أن أخذ العينات وما يستتبعه من إمكانيات أكبر فى استخدام جامعي بيانات، و مراجعين أكثر تدريبا وخبرة يقلل من حجم الأخطاء المصاحبة لهذه العملية، ويجب أن تكون العينات المنتقاة معبرة عن المجتمع (كما سيأتي شرحه فى الفصل الثالث عشر تحت عنوان العينات وطرق انتقائها Samples and Sampling فى هذا الكتاب).

10.8. الدراسة الإحصائية Statistical Study

الدراسة الإحصائية هى تلك الدراسة التى تتعلق بمحاولة استقراء واستنتاج خصائص معلمات المجتمع Population Parameters من خلال حساب إحصاء العينة Samples Statistics

أ- أهداف الدراسة الإحصائية Statistical Study Goals

إن هدف أى دراسة إحصائية يكمن فى رغبتنا فى معرفة خصائص أى عنصر من عناصر المجتمع، وذلك بتجميع البيانات اللازمة عنه، ويتم ذلك - كما أشرنا سابقا- بإحدى طريقتين:

1. الحصر الكامل، وتتم إذا كان من الممكن قياس كافة عناصر المجتمع محل الدراسة وهو لا يمكن فى كثير من الأحوال.
2. أخذ العينات، وهو الشائع، وفيه نلجأ إلى دراسة جزء من كل، أى نجمع بيانات عن بعض العينات وليس كل معلمات المجتمع Population Parameters، ثم نقوم بحساب وقياس إحصاء العينة Samples Statistics، أى خصائص محددة للعينات مثل المتوسط $\bar{X}$ والانحراف المعياري S، ثم عن طريق الإحصاء الاستنتاجي

Inferential Statistics نقوم باستنتاج واستقراء واستنباط مقادير أو قياسات خصائص معلمات المجتمع Population Parameters مثل المتوسط μ، والانحراف المعياري σ، وهذا ما يطلق عليه التقدير Estimation (وسنتناوله بشئ من التفصيل فى الفصل الخامس عشر من هذا الكتاب)، فإذا عرفنا هاتين القيمتين μ & σ كان ذلك كافيا وكفيلا للإجابة عن كافة الأسئلة الاحتمالية حول عناصر هذا المجتمع ككل، ومن ثم فإن القيمة الاحتمالية للمتغير X والذي يمثل أى فرد أو أى عنصر فى ذلك المجتمع يكون قد تحدد تماما، ويصبح بامكاننا الإجابة عن الأسئلة المتعلقة بالاحتمالات حول أى متغير X.

والخلاصة أن الدراسة الإحصائية تمكننا من:

- وصف العينة، كأن نحسب متوسط العينة أو مدى البيانات فى العينة.
- استقراء واستنتاج خصائص المجتمع الذي أخذنا منه العينة وهو الأهم لنا.

ب- خطوات الدراسة الإحصائية Statistical Study Steps

لإجراء دراسة إحصائية على عملية ما، فإنه توجد عدة خطوات محددة، و يمكننا إجمال هذه الخطوات كما يلي:

1. عرف وعين المجتمع محل الدراسة.
2. حدد الاستفسارات التي تريد معرفتها فيما يتعلق بالمجتمع ، أى هل تريد تحديد المتوسط Mean مثل متوسط العدد المباع من سلعة يوميا، أو متوسط الربح الاسبوعى، أو متوسط تكاليف إصلاح معدات فى مصنع ما سنويا، أم تريد تحديد الانحراف المعياري Standard Deviation، والذى يعطى دلالة عن مدى الانحراف المحتمل عن المتوسط الذى يتم حسابه.
3. حدد المتغيرات Variables ذات الصلة والتي ستقوم بجمعها لتساعدك فى الإجابة على الاستفسارات المطلوبة أى هل سنقيس الوقت أم نعد المنتجات المعيبة؟
4. اسحب عينة بحيث تكون معبرة وممثلة للمجتمع .
5. أوجد قيم القيم الإحصائية Statistics المرغوبة مثل المتوسط والانحراف المعيارى... إلى آخره والتي تخص العينة التي تم سحبها.

6. طبق طرق الإحصاء الاستنتاجى Inferential Statistical واحسب قيم معلمات المجتمع Population Parameters.

7. طبق طرق الإحصاء الإستنتاجى لاختبار الفروض التي تم فرضها للتأكد من صحة النتائج.

ويلاحظ أنه بمعرفة قيم المعاملات Parameters وبمعرفة نوع توزيع المجتمع فإننا نستطيع الإجابة عن الأسئلة الاحتمالية حول هذا المجتمع، وبالتالي الإجابة عن الأسئلة المتعلقة بالاحتمالات حول أى متغير X.