ARMS
ACADEMY

Introduction to Statistics

Lecture 1,2

Our Groups

قناة الاكاديمية

Telegram

Ch1 : Introduction to Statistics

ARMS ACADEMY

Statistics

Is the science of
planning studies and experiments, obtaining data, and then organizing, summarizing, presenting, analyzing, interpreting, and drawing conclusions based on the data

علم الإحصاء
هل علم تخطيط الدراسات والتجارب، وجمع البيانات، ثم تنظيمها وتلخيصها وعرضها وتحليلها وتفسيرها واستخلاص النتائج بناءً على البيانات

Prepare التحضير

1. Context السياق

- **What do the data represent?**
  ماذا تمثل البيانات؟
- **What is the goal of study?**
  ما هدف الدراسة؟

2. Source of the Data مصدر البيانات

- **Are the data from a source with a special interest so that there is pressure to obtain results that are favorable to the source?**
  هل البيانات مأخوذة من مصدر له مصلحة خاصة بحيث يكون هناك ضغط للحصول على نتائج تصب في مصلحة هذا المصدر؟

3. Sampling Method طريقة أخذ العينة

- **Were the data collected in a way that is unbiased, or were the data collected in a way that is biased (such as a procedure in which respondents volunteer to participate)?**
  هل جُمعت البيانات بطريقة غير متحيزة، أم جُمعت بطريقة متحيزة (مثل إجراء يسمح للمشاركين بالتطوع للمشاركة)؟

Analyze التحليل

1. Graph the Data تمثيل البيانات بيانياً

2. Explore the Data استكشاف البيانات

- **Are there any outliers (numbers very far away from almost all of the other data)?**
  هل توجد قيم متطرفة (أرقام بعيدة جدًا عن معظم البيانات الأخرى)؟
- **What important statistics summarize the data (such as the mean and standard deviation described in Chapter 3)?**
  ما الإحصاءات المهمة التي تلخص البيانات (مثل المتوسط الحسابي والانحراف المعياري المذكورين في الفصل الثالث)؟
- **How are the data distributed?**
  كيف يتوزع شكل البيانات؟
- **Are there missing data?**
  هل توجد بيانات مفقودة؟
- **Did many selected subjects refuse to respond?**
  هل رفض عدد كبير من الأشخاص المختارين الرد؟

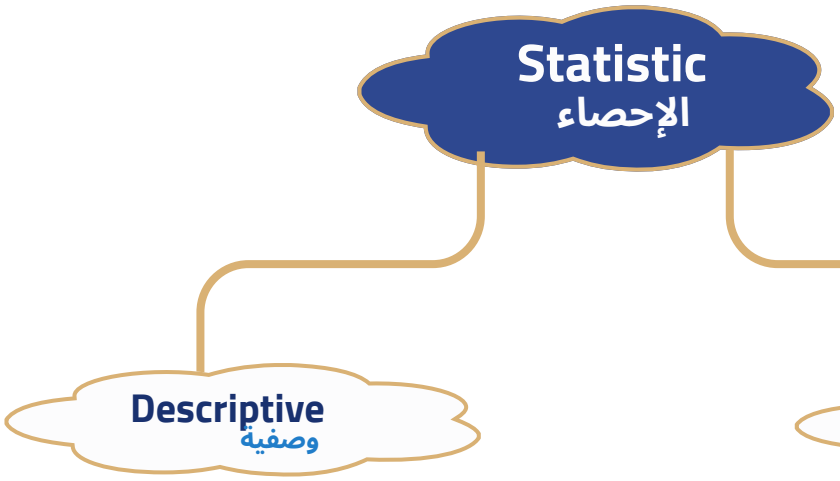
3. Apply Statistical Methods تطبيق الطرق الإحصائية

- **Use technology to obtain results**
  استخدام التكنولوجيا للحصول على النتائج

Conclude الاستنتاج

1. Significance الدلالة

- **Do the results have statistical significance?**
  هل النتائج ذات دلالة إحصائية؟
- **Do the results have practical significance?**
  هل النتائج ذات دلالة عملية (تطبيقية)؟



ARMS ACADEMY

Population المجتمع الإحصائي

The complete collection of all individuals (scores, people, measurements, and so on) to be studied ; the collection is complete in the sense that it includes all of the individuals to be studied

هو المجموعة الكاملة من جميع الأفراد أو العناصر (مثل الدرجات، الأشخاص، القياسات، وغيرها) التي يتم دراستها،
وتُعد هذه المجموعة كاملة لأنها تشمل جميع الأفراد محلّ الدراسة.

Sample عينة

Sub-collection of members selected from a population

مجموعة فرعية من الأفراد يتم اختيارها من المجتمع الإحصائي

Example :

In a survey of 410 human resource professionals, 148 of them said that job candidates were disqualified because of information found on social media postings (based on data from The Society for Human Resource Management)

في استطلاع شمل 410 من المتخصصين في الموارد البشرية، ذكر 148 منهم أن المتقدمين للوظائف تم استبعادهم بسبب معلومات تم العثور عليها في منشورات وسائل التواصل الاجتماعي (اعتمادًا على بيانات من جمعية إدارة الموارد البشرية).

In this case, the population and sample are as follows:

Population: All human resource professionals المجتمع الإحصائي : جميع المتخصصين في الموارد البشرية

Sample: The 410 human resource professionals who were surveyed

العينة: الـ 410 من المتخصصين في الموارد البشرية الذين شملهم الاستطلاع

Parameter المعامل الإحصائي

A numerical measurement describing some characteristic of a population هو قياس عددي يصف خاصية معينة في المجتمع الإحصائي

Population ↕ Parameter

متوسط المجتمع (μ)، ونسبة المجتمع (p)، والانحراف المعياري للمجتمع (σ) — وكلها تُعد معاملات إحصائية.

Statistic الإحصاء

A numerical measurement describing some characteristic of a sample

قياس عددي يصف بعض الخصائص الخاصة بـ العينة.

sample ↕ statistic

Ex. The sample mean $(\overline{x})$, proportion $(\hat{p})$, and the sample standard deviation **(s)** are statistic.

متوسط العينة $(\overline{x})$, النسبة المئوية للعينة $(\hat{p})$, والانحراف المعياري للعينة **(s)** جميعها تعتبر "إحصاءات".

Examples

1- In studying the Loggerhead turtle, scientists observe the average (mean) number of hatchlings in all 253 nests. The given value is a (Parameter/Statistic)

"عند دراسة سلحفاة (Loggerhead)، راقب العلماء متوسط عدد الصغار في جميع الأعشاش البالغ عددها 253 عشاً. القيمة المعطاة هي (Parameter/Statistic)."

2- A sample of batteries lasted an average (mean) of 9 days .The given value is a (Parameter/Statistic)

"صمدت عينة من البطاريات لمتوسط (9) أيام . القيمة المعطاة هي (Parameter/Statistic) "

Data

Collections of observations such as measurements, gender, or survey responses

ARMS ACADEMY

البيانات
"مجموعات من الملاحظات, مثل القياسات, أو النوع (الجنس), أو استجابات الاستبيان"

Data البيانات

Qualitative كيفي

consists of names or labels (representing categories)

تتكون من أسماء أو تسميات (تمثل فئات)

Examples

1-The gender (male/female) of professional athletes
جنس (ذكر/أنثى) الرياضيين المحترفين

2-Shirt numbers on professional athletes uniforms - substitutes for names
أرقام القمصان على زي الرياضيين المحترفين - بدائل للأسماء

3-Names of students in a class
أسماء الطلاب في الفصل

Quantitative كمّية

consists of numbers representing counts or measurements. (or numerical)

تتكون من أرقام تمثل أعداداً (حسابات) أو قياسات. (أو عددية)

Examples

1-The weights of supermodels
أوزان عارضات الأزياء

2-The ages of respondents
أعمار المستجيبين (المشاركين في الاستبيان)

3-The heights of students in a class
أطوال الطلاب في الفصل

Quantitative

continuous مستمرة

result from infinitely many possible values that correspond to some continuous scale that covers a range of values without gaps, interruptions, or jumps

تنتج عن عدد لا نهائي من القيم الممكنة التي تقابل مقياساً متصلاً يغطي نطاقاً من القيم دون وجود فجوات, أو انقطاعات, أو قفزات

Examples

1-The amount of milk that a cow produces; e.g. 2.343115 gallons per day
كمية الحليب التي تنتجها البقرة؛ مثال: 2.343115 جالون في اليوم

2-The ages of respondents
أعمار المستجيبين (المشاركين في الاستبيان)

3-The lengths of distances from 0 cm to 12 cm
أطوال المسافات من 0 سنتيمتر إلى 12 سنتيمتراً

Discrete متقطعة

result when the data values are quantitative and the number of values is either a finite number or 'countable'
(i.e. the number of possible values is 0, 1, 2, 3, . . .)

تنتج عندما تكون قيم البيانات كمية, ويكون عدد القيم إما عدداً محدوداً أو 'قابلاً للعد'. (أي أن عدد القيم الممكنة هو 0, 1", 2, 3, ...)

Examples

1- The number of eggs that a hen lays
عدد البيض الذي تضعه الدجاجة

2-The number of tosses of a coin before getting tails
'عدد مرات رمي العملة المعدنية قبل ظهور 'الكتابة

Examples

1- A sample of Dall sheep is measured to have an average (mean) horn length of 35.7 inches. The given value is

تم قياس عينة من أغنام الدال ووُجد أن متوسط طول قرونها هو 35.7 بوصة. القيمة المعطاة هي

a) Continuous

b) Discrete

2-The number of hatchlings from a sample of 45 bluebird nests is 135. The given value is

عدد الفرخ الحديثة (الفقس) من عينة مكونة من 45 عشاً لطائر السعادة (Bluebird) هو 135. القيمة المعطاة هي

a) Continuous

b) Discrete

ARMS ACADEMY

Levels of Measurement مستويات القياس

Another way of classifying data

1 Nominal level of measurement مستوى القياس الاسمي

categories only

data that consist of names, labels, or categories only, and the data cannot be arranged in an ordering scheme (such as low to high)

بيانات تتكون من أسماء, أو تسميات, أو فئات فقط, ولا يمكن ترتيب هذه البيانات وفق نظام ترتيب معين (مثل من المنخفض إلى المرتفع)

Examples

- Professions المهن
- Survey responses- yes, no, undecided ردود الاستبيان - نعم, لا, (غير مقرر)
- Blood group فصائل الدم
- Colors الألوان

2 Ordinal level of measurement مستوى القياس الترتيبي

categories with some order

involves data that can be arranged in some order, but differences between data values either cannot be determined or are meaningless

بيانات تتكون من أسماء, أو تسميات, أو فئات فقط, ولا يمكن ترتيب هذه البيانات وفق نظام ترتيب معين (مثل من المنخفض إلى المرتفع)

Examples

- Course grades A, B, C, D, or F
  تقديرات المقررات الدراسية (أ, ب, ج, د, أو هـ/و)
- social class: high- middle- low
  الطبقة الاجتماعية: عالية - متوسطة - منخفضة

differences but no natural zero point

3 Interval level of measurement

مستوى قياس الفترة

like the ordinal level, with the additional property that the difference between any two data values is meaningful, however, there is no natural zero starting point (where none of the quantity is present)

مثل المستوى الترتيبي، مع خاصية إضافية وهي أن الفرق بين أي قيمتين من البيانات له معنى، ومع ذلك، لا يوجد نقطة بداية صفرية طبيعية (حيث لا توجد كمية على الإطلاق)

Examples

- Years 1000, 2000, 1776, and 1492

differences and a natural zero point

4 Ratio level of measurement

مستوى القياس النسبي

The interval level with the additional property that there is also a natural zero starting point (where zero indicates that none of the quantity is present); for values at this level, differences and ratios are meaningful

هو المستوى الفتري مع خاصية إضافية وهي وجود نقطة بداية صفرية طبيعية (حيث يشير الصفر إلى عدم وجود أي قدر من الكمية)؛ وبالنسبة للقيم في هذا المستوى، تكون الفروق والنسب ذات معنى

Examples

- Prices of college textbooks ($0 represents no cost, a $100 book costs twice as much as a $50 book)
  أسعار الكتب الجامعية (يمثل الرقم 0 عدم وجود تكلفة، والكتاب الذي سعره 100 دولار يكلف ضعف الكتاب الذي سعره 50 دولاراً)
- Class times of 50 minutes and 100 minutes
  أوقات المحاضرات الممتدة لـ 50 دقيقة و100 دقيقة

Examples : specify the level of measurements

1- Doctors measure the weights (in pounds) of preterm babies. What type of data is collected?

2-Scientists classify the four major deserts in North America using different regions: the Great Basin, the Mojave, the Sonoran, and the Chihuahuan. What type of data is collected?

3-The income range of attendees at a research conference is gathered as upper, middle, and low levels. What type of data is collected?

Basics of Collecting Data
أساسيات جمع البيانات

ARMS ACADEMY

observational studies
دراسة رصدية

Observing and measuring specific characteristics without attempting to modify the subjects being studied

ملاحظة وقياس خصائص محددة دون محاولة تعديل الأفراد (الموضوعات) الذين تتم دراستهم

Experiment
دراسة تجريبية

Apply some treatment and then observe its effects on the subjects; (subjects in experiments are called experimental units)

تطبيق معالجة معينة ثم ملاحظة آثارها على الأفراد؛ (الأفراد في التجارب يُسمون الوحدات التجريبية)

Example: Ice Cream and Drownings

Observe past data to conclude that ice cream causes drownings (based on data showing that increases in ice cream sales are associated with increases in drownings). The mistake is to miss the lurking variable of temperature and the failure to see that as the temperature increases, ice cream sales increase and drownings increase because more people swim

ملاحظة بيانات سابقة للاستنتاج بأن الآيس كريم يسبب الغرق (بناءً على بيانات تظهر أن زيادة مبيعات الآيس كريم ترتبط بزيادة حالات الغرق). الخطأ هو إغفال المتغير الخفي وهو 'درجة الحرارة'، وعدم إدراك أنه مع ارتفاع درجة الحرارة، تزداد مبيعات الآيس كريم وتزداد حالات الغرق لأن المزيد من الناس يذهبون للسباحة

Conduct an experiment with one group treated with ice cream while another group gets no ice cream. We would see that the rate of drowning victims is about the same in both groups, so ice cream consumption has no effect on drownings

إجراء تجربة بحيث يتم إعطاء 'معالجة' لمجموعة واحدة متمثلة في الآيس كريم، بينما لا تحصل المجموعة الأخرى على آيس كريم. سنلاحظ حينها أن معدل ضحايا الغرق هو نفسه تقريباً في كلتا المجموعتين، مما يعني أن استهلاك الآيس كريم ليس له أي تأثير على الغرق

Here, the experiment is clearly better than the observational study

Types of Observational Studies

1-Cross sectional study
الدراسة المستعرضة

Data are observed, measured, and collected at one point in time

بيانات يتم ملاحظتها، وقياسها، وجمعها في نقطة زمنية واحدة

observational studies
دراسة رصدية

2-Retrospective (or case control) study
الدراسة الاسترجاعية (أو دراسة الحالة والشواهد)

Data are collected from the past by going back in time (past period of time)

يتم جمع البيانات من الماضي من خلال العودة بالزمن (عن طريق البحث في فترة زمنية ماضية)

3-Prospective (or longitudinal or cohort) study
الدراسة المستقبلية (أو الطولية، أو الدراسة الجماعية)

Data are collected in the future from groups sharing common factors (called cohorts) (Forward in time)

يتم جمع البيانات في المستقبل من مجموعات تشترك في عوامل مشتركة (تسمى أترابًا/ مجموعات ملاحقة)؛ (بالتقدم في الزمن)

Examples :

1- A researcher from the department of defense is studying the psychology of trauma. He plans to follow the children of service members who died in Afghanistan for the next 20 years. This observational study is ----------

يتم جمع البيانات في المستقبل من مجموعات تشترك في عوامل مشتركة (تسمى أتراباً/مجموعات ملاحقة)؛ (بالتقدم في الزمن)

2- Scientists studying the migration habits of Aleutian geese collect data from reports measuring the number of geese during "fly-offs" over the past 15 years. This observational study is

----

على "fly-offs" يقوم علماء يدرسون عادات الهجرة لإوز جزر ألوتيان بجمع البيانات من التقارير التي تقيس عدد الإوز خلال رحلات الطيران مدار الـ 15 سنة الماضية. هذه الدراسة الرصدية هي

3- In a study of cold symptoms, the body temperatures of study participants were taken. This is an ---------- study

في دراسة لأعراض نزلات البرد, تم قياس درجات حرارة أجساد المشاركين في الدراسة. هذه الدراسة هي

4-Children with attention deficit hyperactivity disorder are randomly divided into four treatment groups (the Tomatis Method, the Dore Method, the Brain Gym, and a control group) to compare differences in focus and concentration. This is an ---------- study

يتم تقسيم الأطفال المصابين بمرض نقص الانتباه وفرط الحركة عشوائياً إلى أربع مجموعات علاجية (طريقة توماتيس, طريقة دور, رياضة الدماغ, ومجموعة ضابطة) لمقارنة الاختلافات في التركيز والإنتباه. هذه الدراسة هي

Methods of Sampling
طرق اختيار العينات

Random عشوائية

A sample of n subjects is selected in such a way that individual member has an equal chance of being selected

من الأفراد بطريقة تضمن أن كل (n) يتم اختيار عينة مكونة من عدد فرد لديه فرصة متساوية في الاختيار

Examples :

1-Computer generated random numbers

أرقام عشوائية مولدة بواسطة الحاسوب

2-Drawing few balls from the box

سحب كرات قليلة من الصندوق

3-Drawing lucky winner

سحب الفائز المحظوظ

Cluster عنقودية

divide the population area into sections (or clusters); randomly select some of those clusters; choose all members from selected clusters

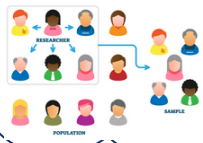
تقسيم منطقة المجتمع إلى قطاعات (أو عناقيد)؛ ثم اختيار بعض تلك العناقيد عشوائياً؛ واختيار جميع الأعضاء من العناقيد التي وقع عليها الاختيار

Convenience ميسرة

use results that are easy to get

استخدام النتائج التي يسهل الحصول عليها

CONVENIENCE SAMPLING



Systematic منتظمة

Select some starting point and then select every kth element in the population

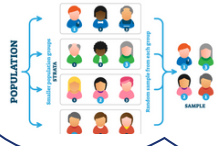
اختيار نقطة بداية معينة, ثم اختيار كل عنصر رقم في المجتمع

Stratified طبقية

Subdivide the population into at least two different subgroups so that the subjects within the same subgroup share the same characteristics. Then draw a sample from each subgroup (or stratum)

تقسيم المجتمع إلى مجموعتين فرعيتين مختلفتين على الأقل بحيث يتشارك الأفراد داخل نفس المجموعة الفرعية في نفس الخصائص. ثم يتم سحب عينة من كل مجموعة فرعية (أو طبقة)

STRATIFIED SAMPLING



Multistage متعددة المراحل

Collect data by using some combination of the basic sampling methods In a multistage sample design, pollsters select a sample in different stages, and each stage might use different methods of sampling

جمع البيانات باستخدام مزيج من طرق المعاينة الأساسية. في تصميم العينة متعددة المراحل, يختار القائمون على الاستطلاع عينة عبر مراحل مختلفة, وقد تستخدم كل مرحلة طريقة معاينة مختلفة

ARMS
ACADEMY

Examples : write the name of sampling method

1-A county water quality control officer obtains a list of all residential addresses in the county and constructs a sample of homes to monitor by selecting every 200th home on the list. The type of sampling is ----------

قام مسؤول مراقبة جودة المياه في المقاطعة بالحصول على قائمة بجميع العناوين السكنية، وقام بتكوين عينة من المنازل لمراقبتها عن طريق اختيار كل منزل رقم 200 في القائمة. نوع هذه المعاينة هو

2-To conduct a study of Johnson & Johnson shareholder attitudes, a list of shareholder zipcodes is compiled. All shareholders from 20 different zipcodes are selected to be surveyed. The type of sampling is ----------

للمساهمين. ثم تم (Zip codes) لإجراء دراسة حول آراء المساهمين في شركة جونسون آند جونسون، تم تجميع قائمة بالرموز البريدية اختيار جميع المساهمين من 20 رمزاً بريدياً مختلفاً للمشاركة في الاستطلاع. نوع هذه المعاينة هو

3-To obtain a sample of pregnant women, a researcher contacts her son's preschool teacher for a list of names. The type of sampling is ----------

للحصول على عينة من النساء الحوامل، قامت باحثة بالتواصل مع معلمة ابنها في الحضانة للحصول على قائمة بالأسماء. نوع هذه المعاينة هو

4-Biologists divide the regions where Red foxes are found into the Nearctic, Palearctic, Oriental, Ethiopian, and the Australian region. They plan to follow 100 female foxes from each region to find the average (mean) number of their offspring. The type of sampling is ----------

يقوم علماء الأحياء بتقسيم المناطق التي توجد فيها الثعالب الحمراء إلى (المنطقة القطبية الشمالية، المنطقة القطبية القديمة، المنطقة الشرقية، المنطقة الإثيوبية، والمنطقة الأسترالية). ويخططون لمتابعة 100 أنثى ثعلب من كل منطقة لإيجاد متوسط عدد صغارها. نوع هذه المعاينة هو

ARMS
ACADEMY

Introduction to Statistics

Lecture 1,2

استثمر في تعليمك اليوم ،لتحصد
ثمار ذلك غداً

Our Groups

قناة الاكاديمية

Telegram