Elements of Statistics and Probability

Dr. Hazar Khogeer

Department of Mathematical Sciences
Umm Al-Qura University
E-mail: hakhogeer@uqu.edu.sa

December 2, 2022

TEXTBOOK:

ELEMENTARY STATISTICS A STEP BY STEP APPROACH

By
Allan Bluman
Tenth Edition, 2018

The main purpose for this course :

- Introduce basic concepts of probability and statistics.
- Students are able to understand and perform simple data analysis.
- The emphasis will be on concepts and problem solving rather than on theory.

Topics:

1. Descriptive Statistics:
   - Chapter 1: The Nature of Probability and Statistics.
   - Chapter 2: Frequency Distributions and Graphs.
   - Chapter 3: Data Description.
   - Chapter10: Correlation and Regression.

2. Introduction to Probability:
   - Chapter 4: Probability and Counting Rules.

Assessment

Activity	Dates	weight
Class Participation (homeworks, quizzes, attendance, etc.)		20%
Major Exam	Approximately Week 7-8	30%
Final Exam (Comprehensive)	As per University Schedule	50%

Reminder :

- Only University issued excuses will be accepted.
- Attendance on time is very important as studies have shown that student that attend class regularly are more likely to complete their courses successfully.
- Homework will be assigned at the end of each chapter.

Chapter 1 (Introduction)

Definition and Basic concepts on the Nature of Probability and Statistics

التعريفات و المفاهيم الإحصاسيه حول طبيعه الإحصاء
و الإحتمال

Outline:

- Introduction.
- 1-1 <u>Descriptive</u> and <u>Inferential</u> Statistics. استدلالي وصفي
- 1-2 <u>Variables</u> and <u>Types of Data</u>. انواع البيانات متغيرات
- 1-3 <u>Data Collection</u> and <u>Sampling Techniques</u>. جمع البيانات
- 1-4 <u>Experimental</u> Design. طرق اختيار العينه
- 1-5 <u>Computers and Calculators</u>.
- Summery.

Objectives:

- Demonstrate knowledge of <u>statistical terms.</u>
- Differentiate between the <u>two branches</u> of statistics.
- Identify <u>types of data.</u> Identify the <u>measurement level</u> for each variable.
- Identify the <u>four basic</u> sampling techniques.
- Explain the difference between an <u>observational</u> and an <u>experimental</u> study.

Objectives: الاهداف

- اظهار المعرفة بالمصطلحات الإحصائية
- التمييز بين فرعي الإحصاء.
- التعرف على أنواع البيانات. تحديد مستوى القياس لكل متغير.
- التعرف على التقنيات الأربعة الأساسية لأخذ العينات.
- اشرح الفرق بين الدراسة الرصدية والتجريبية

Chapter 1

Introduction

- You may be familiar with <u>probability</u> and <u>statistics</u> through radio, television, newspapers, and magazines.

- For example, you may have read statements that are filled with <u>numerical information</u> about <u>phenomenon</u> like sports, health or social activities like the following found in newspapers.

- Such as the winning number of a football team, the most reliable cars, the daily allowance of students, the census of Saudi Arabia, and so on.

- Therefore, statistics is used in almost all fields of human endeavor, such as sports, public health and education.

- Furthermore, it is used for analysis and as a <u>tool in scientific researches</u>.

  اداه في البحث العلمي

- Other uses of it includes <u>operation research</u>, <u>quality control</u>, <u>estimation</u> and <u>prediction</u>.

  ضبط الجودة   بحوث العمليات

For more information see page 2 of Bluman's Book

<u>Introduction</u> : مقدمة

قد تكون على دراية بالاحتمالات والإحصائيات من خلال الراديو والتلفزيون والصحف والمجلات

على سبيل المثال، ربما قد تكون قرأت بيانات مليئة بالمعلومات الرقمية حول ظاهرة مثل الرياضة أو الصحة أو الأنشطة الاجتماعية مثل ما يلي الموجود في الصحف.

مثل
عدد مرات الفوز لفريق كرة القدم
السيارات الأكثر موثوقية
المصروف اليومي للطلاب
التعداد السكاني للمملكة العربية السعودية، وما إلى ذلك.

ولذلك، يتم استخدام الإحصاء في جميع مجالات النشاط البشري تقريبًا، مثل الرياضة والصحة العامة والتعليم.
علاوة على ذلك، يتم استخدامه للتحليل وكأداة في الأبحاث العلمية.
وتشمل الاستخدامات الأخرى له بحوث العمليات ومراقبة الجودة والتقدير والتنبؤ.

Chapter 1

Statistics: is the science of conducting studies

to

collect,

organize,

summarize,

analyze, and

draw conclusions from data.
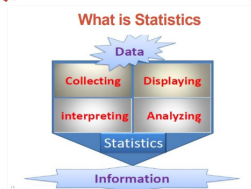
What is Statistics

Data

Collecting | Displaying

Interpreting | Analyzing

Statistics

Information

EXAMPLE

The 2010 Census of Saudi Arabia reveals the following:

- Population is 27,136,977 persons.
- Number of houses is 4,643,151.
- 69% of the population are Saudi citizens.

Students study statistics for several reasons:

قراءة وفهم الدراسات الاحصائية في مجالات

1. Like professional people, you must be able to read and understand the various statistical studies performed in your fields. To have this understanding, you must be knowledgeable about the vocabulary, symbols, concepts, and statistical procedures used in these studies.

اجراء بحث

2. You may be called on to conduct research in your field, since statistical procedures are basic to research. To accomplish this, you must be able to design experiments; collect, organize, analyze, and summarize data; and possibly make reliable predictions or forecasts for future use. You must also be able to communicate the results of the study in your own words.

3. You can also use the knowledge gained from studying statistics to become better consumers and citizens. For example, you can make intelligent decisions about what products to purchase based on consumer studies, about government spending based on utilization studies, and so on.

يدرس الطلاب الإحصائيات لعدة أسباب:

1 مثل الأشخاص المحترفين، يجب أن تكون قادرًا على قراءة وفهم الدراسات الإحصائية المختلفة التي يتم إجراؤها في مجالاتك. للحصول على هذا الفهم، يجب أن تكون على دراية بالمفردات والرموز والمفاهيم والإجراءات الإحصائية المستخدمة في هذه الدراسات.

2 قد يُطلب منك إجراء بحث في مجالك، نظرًا لأن الإجراءات الإحصائية أساسية للبحث. ولتحقيق ذلك، يجب أن تكون قادرًا على تصميم التجارب؛ جمع البيانات وتنظيمها وتحليلها وتلخيصها؛ وربما تقديم تنبؤات أو تنبؤات موثوقة للاستخدام المستقبلي. يجب أن تكون أيضًا قادرًا على توصيل نتائج الدراسة بكلماتك الخاصة.

3 يمكنك أيضًا استخدام المعرفة المكتسبة من دراسة الإحصائيات لتصبح مستهلكًا ومواطنًا أفضل. على سبيل المثال، يمكنك اتخاذ قرارات ذكية بشأن المنتجات التي سيتم شراؤها بناءً على دراسات المستهلك، وبشأن الإنفاق الحكومي بناءً على دراسات الاستخدام، وما إلى ذلك.

The purpose of this chapter

It is the purpose of this chapter to introduce the goals for studying statistics by answering questions such as the following:

- What are the branches of statistics?
- What are data?
- How are samples selected?

<div dir="rtl">

الهدف من هذا الشابتر

يهدف هذا الفصل إلى التعريف بأهداف دراسة الإحصاء من خلال الإجابة على الأسئلة التالية:

ما هي فروع الإحصاء؟

ما هي البيانات؟

كيف يتم اختيار العينات؟

</div>

المتغير   خاصية   صفة-   تتخذ قيما مختلفة

A variable: is a characteristic or attribute that can assume different values.

EXAMPLE:

درجه الحرارة

امثله على متغيرات

- temperature is a variable because it varies from day to day or from city to city or even at different time of the day,

فصيله الدم
- blood type (A,B,AB, O-,O+,.....),

- weight, height, gender, nationality, color, income, ...

الدخل   اللون   الجنسية   الجنس   الطول   الوزن

مصطلحات CH1-1

1. variable: is a characteristic or attribute that can assume different values.
2. Data: are the values that variables can assume to describe a phenomenon
3. Population: consists of all elements for the phenomenon under a study
4. Sample: is a group of elements selected randomly from a population (subset of the population).
5. Descriptive Statistic: consists of the collection, organization,summarization, and presentation of data.
6. Inferential Statistic: consists of generalizing from samples to populations by; performing estimations and hypothesis testing, determining relationships among variables, and making predictions.

1. المتغير: هو خاصية أو سمة يمكن أن تتخذ قيما مختلفة.
2. البيانات: هي القيم التي يمكن أن تتحملها المتغيرات لوصف الظاهرة
3. السكان: يتكون من جميع عناصر الظاهرة محل الدراسة   (مجتمع الدراسة)
4. العينة: هي مجموعة من العناصر يتم اختيارها عشوائيا من مجتمع ما (مجموعة فرعية من السكان).
5. الإحصاء الوصفية: وتتكون من جمع البيانات وتنظيمها وتلخيصها وعرضها.
6. الإحصاء الاستدلالي: يتكون من التعميم من العينات إلى السكان عن طريق؛ إجراء التقديرات واختبار الفرضيات، وتحديد العلاقات بين المتغيرات، ووضع التنبؤات.

عدد افراد مجتمع الدراسة (Population)

N = العدد

عدد افراد عينة الدراسة (Sample)

n = العدد

<u>Data</u>: are the values that <u>variables</u> can assume to <u>describe</u> a phenomenon.

EXAMPLE:

The values that the variable weight can take when measuring the weights of students who are studying Statistics in UQU.

<u>Population</u>: consists of <u>all elements</u> for the <u>phenomenon</u> under a study.

EXAMPLE:

- The 2010 census of Saudi Arabia reveals the following; there are 27,136,977 persons.
  Population: is 27,136,977 persons.
- Determine the leading cause of death in Saudi Arabia in 2020.
  The population: is the <u>entire people who died in Saudia Arabia in 2020</u>.
  The <u>size</u> of the population usually represented by the letter N.

Sample: is a group of elements selected randomly from a population (subset of the population).

العينة

مجموعه يتم اختيارها بشكل عشوائي من مجتمع الدراسة

EXAMPLE:

بعض سجلات الوفيات لاشخاص ماتو في مستشفيات

- Some records of the persons that died in a hospital around Saudi Arabia in 2020.

السعودية عام 2020م

- A group of students is selected randomly from all students studying in Umm AL-Qura University for a research study.

مجموعه من طلاب جامعة ام القرى لاجراء دراسة بحثيه

The size of the sample usually represented by the letter n

Population

Sample

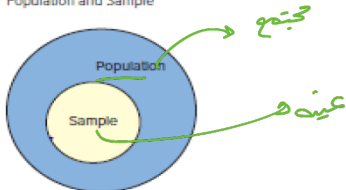مجتمع الدراسة
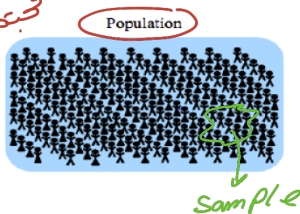
FIGURE 1–1
Population and Sample

Population

Sample

مجتمع

عينه

Population and Sample

- In most cases measuring the whole population is <u>too costly</u>, unnecessary or impossible.

- There are many possible samples when sample is a subset of the whole population.

عينة مهمه جمع ، مجتمع

- When the sample is selected it should be <u>representative</u> of the <u>entire population</u>.

- The sample could be the whole population.

① في معظم الاحيان تطبيق الدراسة على مجتمع كامل تكون (مكلفه ، غير ضروريه ، مستحيله)
② هناك خيارات عديده عند اختيار عينه من مجتمع
③ يجب اختيار العينه ان تكون عينه ممثله لكل المجتمع
④ ممكن ان تكون العينه جميع ، المجتمع

لقسم الاحصاء الى قسمين اساسين

Depending on how data are used, the subject of statistics is divided into two main branches <u>descriptive statistics</u> and <u>inferential statistics</u>.

1 Descriptive Statistic: consists of the <u>collection</u>, <u>organization</u>, <u>summarization</u>, and <u>presentation</u> of data.

2 Inferential Statistic: consists of generalizing from samples to populations by; performing estimations and hypothesis testing, determining relationships among variables, and making predictions.

Note: in any statistical study, you should start with the descriptive statistics and then the inferential statistics which is usually the goal of any study.

الى دراسة احصائيه نبدأ؟ بالاحصاد الوصفي ثم ننتقل الى
الاحصاد الاستدلالي (الهدف الاساسي لاي دراسة)

صف امثله ،

EXAMPLE (Descriptive)

العقد الـعاني هو الثـاني

- The 2010 Census of Saudi Arabia reveals the following:

- Population is 27,136,977 persons.

- Number of houses is 4,643,151.

استدلالي امثله
EXAMPLE (Inferential)

المتوسط التقديري لمتوسط طول الطلاب
- The estimated average height of UQU students is 1.72m.

- There is a relationship between the performance of students in mathematics classes and their class attendance.

- The population of Saudi Arabia is estimated to be around 37 million persons by the year 2030.

هنا له علاقة بين اداء الطالب في مواد الرياضيات وحضوره للصفوف

عدد سكان السعودية سيكون متوقع 37 مليون سنة 2030

الاحصاء الوصفي

Descriptive Statistics
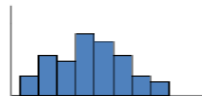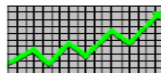
اختيار عينات تحمع البيانات

- <u>Collect data</u> (<u>Sampling</u>)

استبيان امثله

- e.g., Survey

عرض البيانات

- Present data

الجدول والرسوم البيانيه امثله

- e.g., Tables and graphs

تلخيص

- Summarize data

امثله

- e.g., Sample mean $= \dfrac{\sum X_i}{n}$ مجموع البيانات ÷ عدد القيم

موسط العينه

٦خطوات لعمليه الاستدلال

Six Step For Inference Process: ماهو السؤال عند الدراسة (الذي نهتم به)

- Step1: What is the QUESTION of interest?

- Step 2: What POPULATION is associated with the QUESTION?

ما هو المجتمع المرتبط بالدراسة

- Step 3: Take a SAMPLE from the POPULATION.

نأخذ عينه من المجتمع

- Step 4: Measure one or more VARIABLES on each unit in the sample.

قياس متغير واحد او اكثر لكل عنصر في العينة

- Step 5: Make a SUMMARY of all samples data on the VARIABLES.

عمل ملخص لكل البيانات وعينه بالعينات

- Step 6: Make an INFERENCE about the POPULATION from the sample summary.

من خلال التلخيص يمكننا الاستدلال للمجتمع من خلال العينة

POPULATION — مجتمع

Step 1

QUESTION — السؤال

Step 2

xx x x
xx x x x x x
x   xx
xx x x x x x
x   xxx

SAMPLE — عينه

xx
xxx

Step 3

Step 4

Measure VARIABLES on Sample units — هنا يعني بمتغيرات

أ) عمل رسوم بيانيه مثل مدرج تكراري ، مخطط الساق والورقه

Step 5

ب) استخدام ارقام ذات دلاله احصائيه مثل : المتوسط ، الوسيط ، التباين

Make a SUMMARY of Sample Measurements
A) Use graphs such as <u>histogram</u> and <u>stem-and-</u>leaf plot
B) Use numbers called **statistics** such as sample <u>mean</u>, <u>median</u>, <u>variance</u>

Step 6

أ) تقدير معلمه
ب) فحص فرضيه

Make an INFERENCE about the population
A) <u>Estimate</u> **parameters**
B) Test hypotheses.....etc.

EXAMPLE 1-1 Descriptive or Inferential Statistics

السؤال هل تم ا ستخدام احصاء وصفي او استدلالي

Determine whether descriptive or inferential statistics were used

a) The average price of a 30-second ad for the Academy Awards show in a recent year was 1.90 million dollars.

احصاء وصفي

b) The Department of Economic and Social Affairs predicts that the population of Mexico City, Mexico, in 2030 will be 238,647,000 people.

ا ستدلالي

c) A medical report stated that taking statins is proven to lower heart attacks, but some people are at a slightly higher risk of developing diabetes when taking statins.

استدلالي

لان عمد عينه حدده (عددها) وذكر هذا الناتج للعينه

d) A survey of 2234 people conducted by the Harris Poll found that 55% of the respondents said that excessive complaining by adults was the most annoying social media habit.

وصفي

SOLUTION EXAMPLE 1-1

a) **Descriptive statistic** (average) was used since this statement was based on data obtained in a recent year.

b) **Inferential statistics** were used since this is a prediction for a future year.

c) **Inferential statistics** were used since this conclusion was drawn from data obtained from samples and used to conclude that the results apply to a population.

d) **Descriptive statistics** were used since this is a result obtained from a sample of 2234 survey respondents.

مثال 1-1

حدد ما إذا تم استخدام الإحصاء الوصفي أو الاستدلالي

a) بلغ متوسط سعر إعلان مدته 30 ثانية لحفل توزيع جوائز الأوسكار في العام الأخير 1.90 مليون دولار.

b) تتوقع إدارة الشؤون الاقتصادية والاجتماعية أن يصل عدد سكان مدينة مكسيكو، المكسيك، في عام 2030 إلى 238,647,000 نسمة.

c) ذكر تقرير طبي أن تناول الستاتينات يخفض النوبات القلبية، ولكن بعض الأشخاص يكونون أكثر عرضة قليلاً للإصابة بمرض السكري عند تناول الستاتينات.

d) وجدت دراسة استقصائية أجريت على 2234 شخصًا بواسطة Harris Poll أن 55٪ من المشاركين قالوا إن الشكوى المفرطة من قبل البالغين كانت العادة الأكثر إزعاجًا على وسائل التواصل الاجتماعي.

الحل 1-1

a) الإحصائية الوصفية

تم استخدام (متوسط) نظرًا لأن هذا البيان يستند إلى بيانات تم الحصول عليها في العام الأخير.

b) الإحصائيات الاستدلالية

هذا توقع لسنة مقبلة.

c) الإحصائيات الاستدلالية

تم استخلاص هذا الاستنتاج من البيانات التي تم الحصول عليها من العينات واستخدامها لاستنتاج أن النتائج تنطبق على السكان.

d) الإحصائيات الوصفية

هذه هي النتيجة التي تم الحصول عليها من عينة مكونة من 2234 مشاركًا في الاستطلاع

The variable is classified into two main categories:

Quantitative and Qualitative variables.

Quantitative variables are <u>numerical in nature</u> and can be <u>counted</u> or <u>measured</u> or <u>ranked</u>, e.g., <u>number of passengers</u>, weight, age, .

Quantitative or scale variables can classified in to two groups:

- Discrete variables assume values that can be counted, e.g., number of rooms in buildings, number of students in Stat 110, number of children in families, number of hospital beds...etc.

- Continuous variables can assume all values between any two specific values. In other word, takes infinite number of possible values (a decimal fraction can be presented), e.g., heights, weights, temperature,...etc.

عين تصنيفه حسب خاصية اوصه معينة

منفيرات نوعية ؛ ٢ كمية

Qualitative variable: can be placed into <u>distinct categories</u> according to some characteristic or attribute, as <u>gender</u>,(<u>ma</u>le/ female), b<u>lood type</u> (A, B, AB, O-, O+,..) , grade (<u>A</u>+, <u>A</u>, <u>B</u>+, <u>B</u>, C+, C, D+, D, F)...etc.

امثله على بيانات نوعيه ماه جنس ، النوع،الدم ، درجة(حروف)

نوعية

Qualitative variables can be classified into two groups:

اصناف

■ <u>Nominal</u> <u>classifies</u> data into categories that <u>can not</u> be <u>ordered</u> or <u>ranked</u>, e.g., gender, color, names, departments · · ·

اسمي

■ <u>Ordinal</u> <u>classifies</u> data into <u>categories</u> that <u>can</u> be <u>ranked</u>, e.g., <u>flight classes</u>, grade letters,· · ·

ترتيبي

اسمي :- اصناف لا يمكن ترتيبها من الاصغر الى الاكبر ,مثل , اللون ,جنسية ، الديانه،

ترتيبي :- اصناف يمكن ترتيبها من الاصغر الى الاكبر ,مثل درجات (حروف) ام مستوى لتعليم

For more information see pages 6 and 7 of Bluman's Book

The classification of variables can be summarized as follows:

Exercise:

صنفصل

متصل
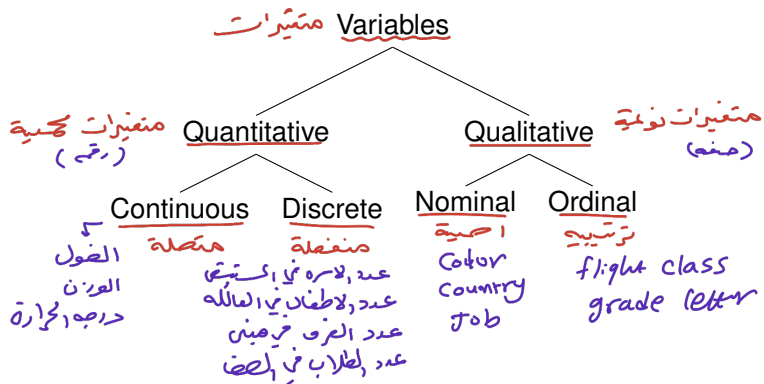
classify each variable as a discrete variable or a continuous variable.

a. The highest wind speed of a hurricane.

اعلى سرعه رياح في الاعصار

b. The weights of luggage on an airplane.

وزن الشنط في الطائرة

c. The number of pages in a statistics book.

عدد الصفحات في كتاب الاحصاء

d. The amount of money a person spends per year on online purchases.

حجم النقود التي نصرفها في الشراء

solution:

a. Continuous variable   c. Discrete variable

b. Continuous variable.   d. Discrete variable.

EXAMPLE 1-2: classify each variable as a discrete variable or a continuous variable. عدد الساعات التي يقضيها الطفل يشاهدالتلفاز

a) The number of hours during a week that children ages 12 to 15 reported that they watched television. عدد التمريرات التي يسجلها اللاعب

b) The number of touchdowns a quarterback scored each year in his college football career. منفصل

c) The amount of money a person earns per week working at a fast-food restaurant.

d) The weights of the football players on the teams that play in the NFL this year. وزن اللاعب

solution:

a) Continuous, since the variable time is measured.

b) Discrete, since the number of touchdowns is counted.

c) Discrete, since the smallest value that money can assume is in cents.

d) Continuous, since the variable weight is measured.

حدود الفئة
Class Boundaries

بيانات متصلة

When **continuous data** is measured, answers must be rounded.
Usually, answers are rounded to the nearest given unit.

في البيانات المتصله يتم ترتيب القيم الى اقرب وحدة (حدود الفئه)

TABLE 1-1 Recorded Values and Boundaries		
Variable	**Recorded value** القيمه المعله	**Boundaries** حدود
Length	15 centimeters (cm)	14.5–15.5 cm
Temperature	86 degrees Fahrenheit (°F) $\{$ −0.5 +0.5	85.5–86.5°F
Time	0.43 second (sec)	0.425–0.435 sec
Mass	1.6 grams (g)	1.55–1.65 g

0.430−0.005
0.430+0.005

1.60 − 0.05
1.60+0.05

الفضل يرجع الزمن كله

لتحديد حدود الفئه
① اذا كان عدد صحيح

② نحو صال وجود اعشار
① نضيف ونطرح لنين الرقم
② كن هذا الصف نجم ونطرح 5

0.5 → نجمع ← 15
0.5 → نطرح

12.20

12.20−
12.25

12.20−
12.15

12.15   12.25

أوجد حدود كل قياس

EXAMPLE 1-3: Class Boundaries.

Find the boundaries for each measurement.

a) 17.6 inches. $17.55 - 17.65$

b) 23 Fahrenheit. $22.5 - 23.5$

c) 154.62 mg/dl. $154.615 - 154.625$

solution:

a) 17.55-17.65 inches.

b) 22.5-23.5 Fahrenheit.

c) 154.615-154.625 mg/dl

a) 17.60

$$17.60 - \quad\quad 17.60$$
$$00.05 \quad\quad 00.05 +$$
$$\overline{17.55} \quad - \quad \overline{17.65}$$

b) $23 - 0.5 = 22.5$
$23 + 0.5 = 23.5$

c) $154.620 - \quad\quad 154.620$
$000.005 \quad\quad 100.005 +$
$\overline{154.615} \quad\quad \overline{154.625}$

In addition to classify the variables as <u>qualitative</u> or <u>quantitative</u>, variables can be classified by how they are <u>categorized</u>, <u>counted</u>, or <u>measured</u>.

Next we will describe the different (<u>measurement of scale</u> of <u>the data</u>)

مستويا ت القياس

Why do we need to Know the <u>Level of Measurement</u> of a Data?

- The level of measurement of the data **dictates the calculations** that can be done to summarize and present the data. معرفه مستوى القياس يحدد طريقه الحسابات التى نجريها على البيانات

- To determine the **statistical tests** that should be performed on the data. معرفه مستوى القياس يحدد الاختبارات الاحصائيه التى علين القائم بها

Four Levels of Measurement scale

1 **Nominal :** the variable is classified into categories in which no order or ranking can be imposed on the data. Example, race, gender, eye color.

2 **Ordinal:** the variable is classifies into categories hows values are order or scaled. Example age group(classified into, for example: less than 21 years, 22-40 years and greater than 40 years), and educational level (classified into, for example: high school, bachelor, master, and PhD).

3 **Interval** the variable is measured numerically but has no true zero, with meaningful amounts of order between data. For example temperature, IQ 'intelligence quotient' tests do not measure people who have no intelligence.

4 **Ratio** The variable is measured numerically with true and meaningful zero. Differences and ratios are meaningful for this level of measurement. Example, monthly income, distance traveled, wight, and height.

Characteristics of Levels of Measurement
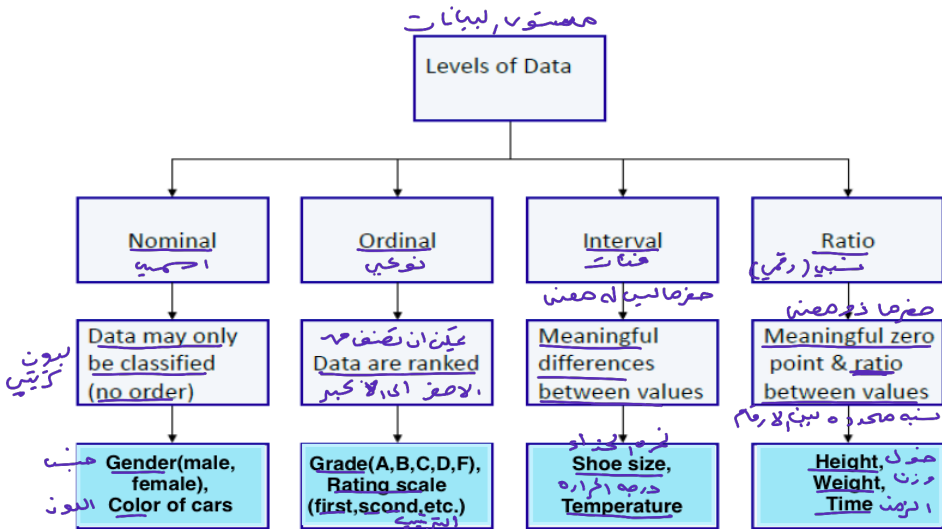
مستوى البيانات

Levels of Data

Nominal	Ordinal	Interval	Ratio
اسمية	نوعي	فئات	نسبي (رقمي)

Nominal — اسمية
Data may only be classified (no order) — بدون ترتيب

Ordinal — نوعي
عين ان تضف جه
Data are ranked — الاصغر الى الا كبر

Interval — فئات
حيث الفرق بين القيم له معنى
Meaningful differences between values

Ratio — نسبي (رقمي)
صفر ذو معنى
Meaningful zero point & ratio between values — نسبة محددة و نسبة لا رقم

Gender(male, female), Color of cars — حبة اللون

Grade(A,B,C,D,F), Rating scale (first,scond,etc.) — درجة العلامة (الترتيب)

Shoe size, Temperature — نمر الجذاء

Height, Weight, Time — طول وزن الزمن

EXAMPLE 1-4: Measurement Levels.

What level of measurement would be used to measure each variable (Nominal,Ordinal,Interval,Ratio)?

a) The ages of authors who wrote the hardback versions of the top 25 fiction books sold during a specific week.

b) The colors of baseball hats sold in a store for a specific year.

c) The highest temperature for each day of a specific month.

d)The ratings of bands that played in the homecoming parade at a college.

solution:

a) Ratio.

b) Nominal.

c) Interval

d) Ordinal

Exercise: Chose the correct answer

Number of siblings in a family is an example of:

عدد الأخوات

    (a.) Discrete variable   b. Continuous variable

    c. Qualitative variable   d. Nominal scale

رقم ، عنزل

House number is an example of:

    a. Discrete variable   b. Continuous variable

    c. Ordinal variable   d. Nominal variable

    → number of Houses

عدد البيوت خيادي

رقم ، الجوال
الرقم ، كأعمى ← بنات ، أسميه
رقم ، الهويه
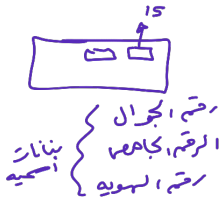
The time it takes a student to solve the exam is an example of:

    a. Ordinal variable   b. Continuous variable

    c. Discrete variable   d. Nominal variable

Temperature is measured on:

    a. Nominal scale   b. Ratio scale

    c. Interval scale   d. Ordinal scale

HOMEWORK:Chapter 1

- Exercises 1 -1: page 5 (4, 9-17).
- Exercises 1-2: page 10-11 (5-10, 11-16, 23-30)

حرف اصئ الصنيات

لجح بيان ر

Data can be collected in a variety of ways. One of the most common ways is the use of <u>surveys</u> that can be done by using a variety of methods.

استبيان

Three of the most common methods are:

ضرف جح, لاستبان

- Telephone surveys   هاتف
- Mailed q<u>uestionnai</u>re surveys   ۱ لا بعبل
- Personal interview surveys   مقابله ر شخصه

طرق اخذ العينات

Some Sampling Techniques

العلاقه المتوانه للسلفه

<u>Simple random sampling</u> where every particular sample of a specific sample size has the <mark>same chance of being selected</mark>, thus the member of the sample are chosen independently of each other.

جميع الاشخاص لديهم نفس الرزمه للاختيار

How to Choose a Random Sample:

- each subject or object in the population is assigned a number from 1 to N.
- to select (n) objects/subjects from the population a random number generator is used.

منتج للارقام العشوائيه

- the random number generator could be a <u>table</u>, a <u>calculator</u>, or a <u>computer</u>. Reject any numbers that do not correspond to any population member. (Ignore any repeated occurrence of the same number.)

حاسبه   جدول

- select a total of (n) items from the numbers from 1 to N.

عينه   جميعه

SRS

Example

عشوائي

A group of 1<u>0</u> students is selected using <u>random numbers</u> from 30 students to check the performance of a class.


عشوائي
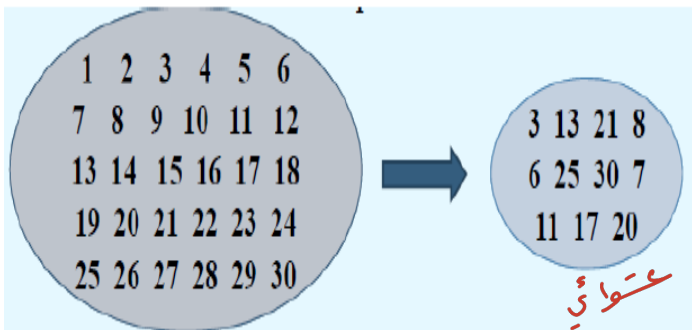
منظمة

Systematic every kth subject.

30/6

N/n هم خطوه

Example

Suppose there were 30 subjects in the population and a sample of 6 subjects were needed. Since $30/6 = 5$, then $k = 5$, the first subject (numbered between 1 and 5) would be selected at random then every 5 subject is selected. Suppose subject 3 were the first selected; then the sample would be

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
16	17	18	19	20	21	22	23	24	25	26	27	28	29	30

→

3	8
13	18
23	28
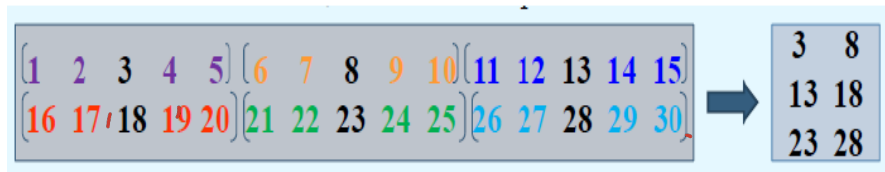
$$N/n \qquad 15/5 = 3 \; \text{خطوه}$$

2.Systemetic

Population

Sample

عينات طبقيه

علوم   حاسبا   صندقه
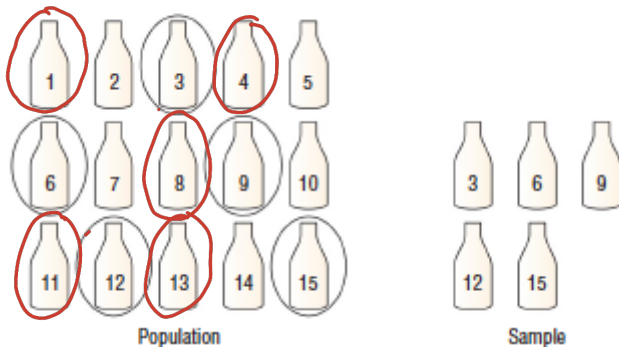
عشوائي

Stratified samples are selected by <u>dividing</u> the <u>popula</u>tion into groups (<u>strata</u>), then subjects are randomly selected from each group or strata.
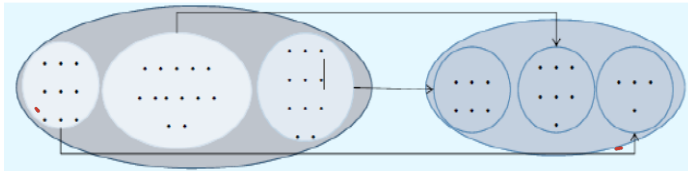
يتم تقسيم المجتمع الى مجموعات ثم
يتم الاختيار بشكل عشوائي من كل مجموعة
طبقات

Example

Suppose we want to get a sample form a high school but we need to be sure that the sample contains subjects from each level in the school.

3.Stratified

Men

4

Population

Women

4

Sample

العنقودية
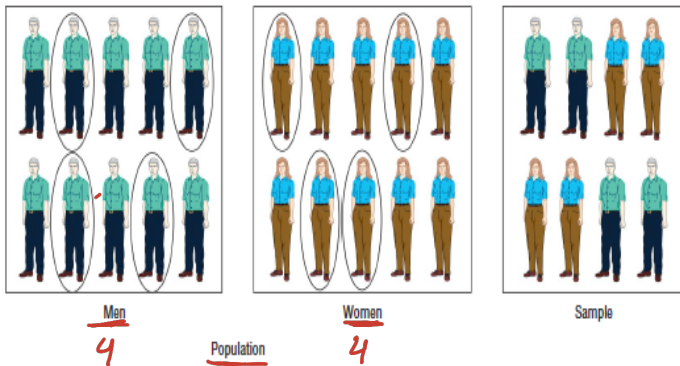
Cluster samples are selected by using intact groups called clusters and then selecting one or more clusters and using all members in the cluster(s) as the members of the sample.
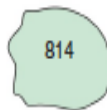
4.Cluster

اختيار مجموعات محدده ليتم دراسة جميع عناصرها
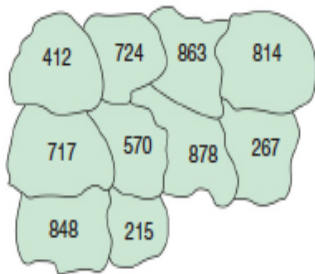
| 412 | 724 | 863 | 814 |

| 717 | 570 | 878 | 267 |

| 848 | 215 |

Area codes
Population

814

717

Sample

EXAMPLE 1-5: Measurement Levels.
State which sampling method was used?

a) Out of 10 hospitals in a municipality, a researcher selects one and collects records for a 24-hour period on the types of *Cluster* emergencies that were treated there. *M* *F*

b) A researcher divides a group of students according to gender, major field, and low, average, and high grade point average. Then she randomly selects six students from each group to answer questions in a survey. *stratified*

c) The subscribers to a magazine are numbered. Then a sample of these people is selected using random numbers. *simple random*

d) Every 10th bottle of Energized Soda is selected, and the amount of liquid in the bottle is measured. The purpose is to see if the machines that fill the bottles are working properly. *systematic*

اذكر طريقة أخذ العينات التي تم استخدامها؟

أ) من بين 10 مستشفيات في إحدى البلديات، يختار الباحث واحدًا ويجمع السجلات لمدة 24 ساعة حول أنواع حالات الطوارئ التي تم علاجها هناك.

ب) قام الباحث بتقسيم مجموعة من الطلاب حسب الجنس والمجال الرئيسي والمعدل التراكمي المنخفض والمتوسط والمرتفع. ثم تختار بشكل عشوائي ستة طلاب من كل مجموعة للإجابة على الأسئلة في الاستطلاع.

ج) يتم ترقيم المشتركين في المجلة. ثم يتم اختيار عينة من هؤلاء الأشخاص باستخدام أرقام عشوائية.

د) يتم اختيار كل زجاجة عاشرة من الصودا النشطة، ويتم قياس كمية السائل في الزجاجة. والغرض من ذلك هو معرفة ما إذا كانت الآلات التي تملأ الزجاجات تعمل بشكل صحيح

solution:
a) Cluster
b) Stratified
c) Random
d) Systematic

There are two types of studies:

Observational study: the researchers collects data from subjects by only observing and has no control on the conditions of the experiments and tries to draw conclusions based on the observations.

Example, the experimenter observed that number of babies on the airplane that receive fluid during take off and landing that do not cry.

Experimental study: the researcher manipulates one of the variables and tries to determine how that influences other variables.

Example, randomly select babies on the airplane to receive fluid during take off and landing, to see if this going to prevent them from crying.

For more information see pages 18-20 of Bluman's Book

EXAMPLE (<u>Observational Study</u>) دراسة رصدية

"Motorcycle owners are getting older and richer".

الدخل   الاعمار

Data were collected on the <u>ages</u> and <u>incomes</u> of <u>motorcycle</u> owners for the years <u>1980</u> and <u>1998</u> and then <u>compared</u>. The <u>findings</u> showed <u>considerable</u> differences in the <u>ages</u> and <u>incomes</u> of motorcycle owners for the two years. In <u>this study</u>, the researcher observed what had happened to the motorcycle owners over a period of time.

نتيجة

نتيجة لدراسة ::- و صفت الباحث كيف تغيرت الاعمار ودخل
مالكي الدراجات خلال العامين

أصحاب الدراجات النارية يكبرون وأكثر ثراءً. "
تم جمع البيانات عن أعمار ودخل أصحاب الدراجات النارية للسنوات 1980 و 1998 ثم
مقارنة. أظهرت النتائج اختلافات كبيرة
في العصور والدخل من أصحاب الدراجات النارية للاثنين
سنوات.
في هذه الدراسة ، لاحظ الباحث ما حدث
لأصحاب الدراجات النارية على مدى فترة من الزمن.

درسة تجريبة

EXAMPLE (Experimental Study)

Students are divided into two groups and had the students perform as many sit-ups as possible in 90 sec. The first group was told only to "Do your best," while the second group was told to try to increase the actual number of sit-ups done each day by 10%. After four days, the subjects in the group who were given the vague instructions to "Do your best" averaged 43 sit-ups, while the group that was given the more specific instructions to increase the number of sit-ups by 10% averaged 56 sit-ups by the last days session. The conclusion then was that athletes who were given specific goals performed better than those who were not given specific goals.

تم تقسيم الطلاب إلى مجموعتين ويقوم الطلاب بأداء أكبر عدد ممكن من تمارين البطن في 90 ثانية. طُلب من المجموعة الأولى فقط أن "تبذل قصارى جهدها"، بينما طُلب من المجموعة الثانية أن تحاول زيادة العدد الفعلي لتمارين الجلوس التي يتم إجراؤها كل يوم بنسبة 10%. بعد أربعة أيام، بلغ متوسط عدد تمرينات الجلوس في المجموعة التي أعطيت تعليمات غامضة "بذل قصارى جهدك" 43 تمرينًا، في حين أن المجموعة التي تلقت تعليمات أكثر تحديدًا لزيادة عدد تمرينات الجلوس بنسبة 10٪ بلغ متوسطها 56 تمرينًا. الاعتصامات في جلسة الأيام الأخيرة. وكان الاستنتاج إذن هو أن الرياضيين الذين أعطوا أهدافًا محددة كان أداؤهم أفضل من أولئك الذين لم يحصلوا على أهداف محددة

1-4 Independent variable Vs Dependent variable

Statistical studies usually include one or more independent variables and one dependent variable.

The independent variable: is being manipulated by the researcher. The independent variable is also called the explanatory variable.

The dependent variable: or outcome variable is the resultant variable. Thus, the outcome variable is the variable that is studied to see if it has changed significantly due to the manipulation of the independent variable.

For more information see pages 18-20 of Bluman's Book

EXAMPLE 1-6 Experimental Design

Researchers randomly assigned 10 people to each of three different groups. Group 1 was instructed to write an essay about the hassles in their lives. Group 2 was instructed to write an essay about circumstances that made them feel thankful. Group 3 was asked to write an essay about events that they felt neutral about. After the exercise, they were given a questionnaire on their outlook on life. The researchers found that those who wrote about circumstances that made them feel thankful had a more optimistic outlook on life. The conclusion is that focusing on the positive makes you more optimistic about life in general. Based on this study, answer the following questions.

a) Was this an observational or experimental study?
b) What is the independent variable? *essay type*
c) What is the dependent variable? *the score of questionnaire*
d) What may be a confounding variable in this study? *age, income, job*
e) What can you say about the sample size? *n = 30*
f) Do you agree with the conclusion? Explain your answer.

قام الباحثون بتخصيص 10 أشخاص بشكل عشوائي لكل مجموعة من ثلاث مجموعات مختلفة. طُلب من المجموعة الأولى أن تكتب مقالاً عن المتاعب التي تواجههم في حياتهم. طُلب من المجموعة الثانية أن تكتب مقالاً عن الظروف التي جعلتهم يشعرون بالامتنان. طُلب من المجموعة 3 كتابة مقال عن الأحداث التي شعروا بأنها محايدة تجاهها. وبعد التمرين، تم إعطاؤهم استبيانًا حول نظرتهم للحياة. ووجد الباحثون أن أولئك الذين كتبوا عن الظروف التي جعلتهم يشعرون بالامتنان كانت لديهم نظرة أكثر تفاؤلاً للحياة. الخلاصة أن التركيز على الإيجابيات يجعلك أكثر تفاؤلاً بالحياة بشكل عام. بناء على هذا

الدراسة، والإجابة على الأسئلة التالية.

أ) هل كانت هذه دراسة رصدية أم تجريبية؟

ب) ما هو المتغير المستقل؟

ج) ما هو المتغير التابع؟

د) ما هو المتغير المربك في هذه الدراسة؟

هـ) ماذا يمكنك أن تقول عن حجم العينة؟

و) هل توافق على الاستنتاج؟

solution:

a) This is an experimental study since the variables (types of essays written) were manipulated.

b) The independent variable was the type of essay the participants wrote.

c) The dependent variable was the score on the life outlook questionnaire.

d) Other factors, such as age, upbringing, and income, can affect the results; however, the random assignment of subjects is helpful in eliminating these factors.

e) In this study, the sample uses 30 participants total.

f) Answers will vary.

Statistical Packages

ا ضله حمل تكو لوهيا
ستـحدنا في الاحصاء

- Excel, SPSS, MINITAB, SAS and the TI-83 graphing calculator can be used to perform statistical computations.

- Students should realize that the computer and calculator merely give numerical answers and save time and effort of doing calculations by hand.

الحسوبة دوض
الوقت و الجهد

HOMEWORK:Chapter 1

- Exercises 1 -3: page 17 (5-10, 11-16).
- Exercises 1-4: page 24 (15-18, 19-22)