

# Chapter 1 Data and Statistics

**Statistics** → The science of collecting, analyzing, presenting and interpreting data  
إحصاء علم جمع - تحليل - عرض وتفسير البيانات

**Data** → Fact and Figures (measurement) that are collected  
معلومات

**Data set** → all the data collected in a particular study  
مجموعة بيانات

← أي معلومة (Data) لها قيمة  
تتكون من بيانات هي معلومات (Information)

**Elements** → the entities on which the data are collected  
عناصر ex → People, mobiles, firms  
العنصر أي شيء له قيمة

**variables** → characteristic of interest of the element  
متغير element  
العلاقة المتغيرة بالعنصر والتي نريد تحديدها

**Observation** → The set of measurements for a particular element  
ملاحظة  
المجموعة لقياس لعنصر معين / قياس العنصر  
المعنى مجموعة بيانات هي التي تم تسجيلها على  
صفتها

**Population** → The collection of all element in a study  
مجتمع  
المجموعة جمع العناصر التي يتم دراستها

Sample → a subset of the population  
عينة له قسمه من الكل بالقطع

Census → Collecting data on the entire population  
مع شامل له يعني تجميع البيانات من كل عناصر المجتمع بأكمله

Survey → Collecting data on a sample  
استبيان له يعني تجميع البيانات من عينة من المجتمع

Scale of measurement مستوى القياس  
له وهو ينشأ بالنتيجة من مراحله

Nominal

Data are just labels or names  
البيانات مجرد رموز أو أسماء

Order or rank is not meaningful  
له الترتيب ليس له أي معنى

يعني أنا ما يقدر أرتب البيانات بعضها  
بغير أسماء

Data are either numeric or nonnumeric  
له البيانات إما أن تكون رقمية أو غير رقمية

وهذا النوع الأرقام تكون غير أرقام  
فمثلا عندما نكتب Female مرقم 1

يعني لو الرقم غير أصلي وإنما بدل على اسم معين

example → Gender (male, female)  
 Blood type (A, B, AB, O)  
 Phone numbers &  
 student numbers

لأنه رقمها النسوي على سبيل المثال  
 لو اتا معنا رقم الهاتف الخاص بي مع رقم  
 الهاتف الخاص بـ صديقي  
 له الاتع يكون بلا معنى

في جميع عدم الرفض بين الاستبار نظيرة الموضع

13

### 12) Ordinal

Data are labels or names  
لأن البيانات هي رموز أو أرقام

Order or rank is **meaningful**  
لأن الترتيب له معنى في بعض الحالات  
الاسم أو الرقم يمكن وضعها ترتيب ذو معنى  
مثال: A/A+/C ← ترتيب

Data either numeric or nonnumeric

لأن البيانات تكون رقمية أو غير رقمية

لكن الأرقام لا يمكن ترتيبها وترتيبها هي لرمز تكون أمثلة

- example →
- service rating ( good, poor )
  - Clothes size ( large, xlarge, small )
  - Satisfaction ( approve, Disapprove )

### 13) Interval

Data are numeric

لأن البيانات عبارة عن أرقام يمكن جمعها وطرحها  
وتكون ذات معنى

Order or rank is **meaningful**  
لأن الترتيب له معنى

The interval (**difference**) between value exists  
لأن الفرق بين رقمين يعني  $20 - 10 = 10$

١٥٥٥  
 No true or meaningful zero

لها الصفر لا يعني انه لا يوجد شيء  
 فالصفر هناك يعتبر نقطة توقف  
 فيكون ما هو اقل من الصفر

مثلا درجة الحرارة عند صفر لا تتوقف  
 فيكون مثلا درجة حرارة 1 تحت الصفر

Example → Temperature  
 Date of birth  
 SAT scores  
 IQ test      الاختبار الذكاء

## 14) Ratio

□ Data are numeric

له البيانات رقمية اقلية ذات معنى

□ Order or rank is meaningful

له الترتيب لها ذو معنى

□ Ratios are meaningful

له النسب ذات معنى والنسبة هي

نتج لعدة ابي رقميين

## مثلاً

□ True or meaningful zero

له هنا الصفر يكون بعض الصفر الحقيقي يعني

لا يوجد شيء

فالصفر لها صيغة توقف

مثلاً عندما يكون الشئها لا يوجد معه

نقود ويقول هو في شكل

هنا يكون الصفر (Ordinal)

Example → Profit

Age

Distance

Salary

Family size

بعض اقلية رتبة فيما زالجها الى كتور

12 Type of variables (Data)

13 Qualitative (Categorical) التسمية  
Data are nominal or Ordinal

14 Quantitative الرقمية  
Data are Interval or Ratio

15 Type of quantitative data أنواع البيانات الرقمية

16 Discrete متقطع  
Data values can be counted and gaps between value exist. usually, no decimals  
(How many) → كمية أو عدد

لا يمكن حساب قيمه بلانا، وتكون  
فجوات بين القيم. عادة لا توجد كسور عشرية  
بعض لا تقبل للشار  
فقط لا يمكن ان نقول عدد افراد الاسرة 4 افراد، ونقول

Example → Family size, number of chairs in rooms

17 Continuous متصل  
Infinite number of value, no gap, include decimals  
(How much)

لا يوجد لنهاية من القيم. لا يوجد فجوة  
وهي تقبل للشار  
فقط يمكن ان نقول عدد الوزن 45.5

Example → Revenues, weight, Amount of Calories in Cake

□ Type of Statistics

1) Descriptive statistics الوصفية

① Tabular → Frequency and Cumulative Frequency Distribution

الجدول التكراري والتكرار التراكمي

② Graphical → Histogram, Pie Chart, Ogive  
المخطط التكراري، منطوق دائري

③ Numerical → proportion, Percentage, Average  
النسبة، النسبة المئوية، متوسط

2) Statistical Inference →

using data from sample to make estimates about the population

استخدام بيانات العينة من المجتمع لتقدير المجتمع ككل

مثلاً: عند زيارتي لجامعة فإحدى فصولها

لقد التقيت 20% من المجتمع

فأخذت عينة من هذا المجتمع فإحدى 100 طالب sample

وكان ذلك 20% من العينة فإحدى 20  
منهم ذكور، نسبة على المجتمع الجامعة ككل



### Data Sources: مصادر البيانات

① Existing sources مصادر موجودة / موجودة مسبقاً  
↳ Data that already exist  
ex → Customer profiles, Bank records.

② Statistical studies دراسة إحصائية

Ⓐ Experimental تجريبية  
ex ↳ studying how a new drug affect the Blood pressure

لدراسة تأثير دواء جديد على ضغط الدم  
من خلال تجارب معية ومراقبة وتغييره  
على، مثل نتائج مثل نسبة السكر في الدم

Ⓑ Observational مراقبة / قائمة بيانات  
↳ survey and questionnaires  
مثل استبيان

□ Cross-sectional data بيانات تم جمعها على فترة واحدة  
↳ Data collected at the same, or approximately the same time  
له بيانات تم جمعها على فترة واحدة

□ Time serious data البيانات التي تم جمعها على فترات  
↳ Data collected over several period of time  
له بيانات تم جمعها على أكثر من فترة  
مثل المعدل السنوي

### □ Parameter

↳ numerical measure based on the population  
قياس رقمي على أساس المجتمع ككل

### □ Statistic

↳ a numerical measure based on the sample  
قياس رقمي على أساس عينة المجتمع  
Pānā qānā

### Proportion, Percentage, Average

Example Blood type → A, AB, O, A, B, A, B, AB, AB, A, AB

① proportion

$$\text{proportion of AB} = \frac{5}{12} = 0.42$$

② percentage

$$\text{percentage of Blood O} = \frac{1}{12} \times 100 = 8.33\%$$

③ Average

Agas, 12, 15, 20, 57, 16, 19

$$\text{Average} = \frac{12+15+20+57+16+19}{6}$$

$$= \frac{139}{6}$$

$$= 23.17$$