Example 1

An all male jury of 12 people is chosen for a trial. The defence lawyer claims the jury must have been chosen in a biased manner because 50% of the city's adult residents are women and yet the jury contains no women. Assess the claim by answering the following:

If the jury were randomly chosen from the city's adult population, what is the probability that the jury would have (a) no females, (b) at least one female?

Ans. (a) 0.5^{12}; (b) $1 - 0.5^{12}$

Example 2

When testing for electrical current in a cable with five colour-coded wires, a meter is used to test two wires at a time. How many tests are required for every possible pairing of two wires? Ans. $_5C_2$

Example 3

The Biogene Research Company claims that it has developed a technique for ensuring that a baby will be a girl. In a test of that technique, 12 couples all have baby girls. Find the probability of getting 12 baby girls by chance, assuming boys and girls are equally likely and that the gender of any child is independent of the others. Does the result appear to support the company's claim? Ans. See Example 1 part (a); Small probability implies a rare event so the result supports the company's claim.

Example 4

The New York State Health Department reports a 10% rate of the HIV visrus for the "at-risk" population. Under certain conditions, a preliminary screening test for the HIV virus is correct 95% of the time. One person is randomly selected from the at-risk population.

(a) What is the probability that the person selected has the HIV virus if it is known this person has tested positive in the initial screening?

(b) What is the probability that the selected person tests positive in the initial screening if it is known that this person has the HIV virus?

(c) What is the probability that the selected person has the HIV virus if it is known that he or she has tested negative in the initial screening?

(d) What is the probability that the selected person tests negative in the initial screening if it is known that he or she has the HIV virus?

Chapter 1
Definition

- **Data** are observations (such as measurements, genders, survey responses) that have been collected.

Chapter 1
Definition

- **Data** are observations (such as measurements, genders, survey responses) that have been collected.
- **Statistics** is a collection of methods for planning experiments, obtaining data, organizing, summarizing, analyzing, interpreting, presenting, and drawing conclusions based on the data.

Chapter 1
Definition

- **Data** are observations (such as measurements, genders, survey responses) that have been collected.
- **Statistics** is a collection of methods for planning experiments, obtaining data, organizing, summarizing, analyzing, interpreting, presenting, and drawing conclusions based on the data.
- A **census** is the collection of data from *every* member of the population.
- A **sample** is a *subset* of members selected from part of a population.

Chapter 1
Section 1-2: Types of Data

- **Parameter**: a measurement describing some characteristic of a *population*. E.g. Mean mercury level of all fish in Lake Winnipeg.

Chapter 1
Section 1-2: Types of Data

- **Parameter**: a measurement describing some characteristic of a *population*. E.g. Mean mercury level of all fish in Lake Winnipeg.
- **Statistic**: a measurement describing some characteristic of a *sample*. E.g. Draw a sample of 10 fish from Lake Winnipeg and calculate their sample mean mercury level.

Chapter 1
Section 1-2: Types of Data

- **Parameter**: a measurement describing some characteristic of a *population*. E.g. Mean mercury level of all fish in Lake Winnipeg.
- **Statistic**: a measurement describing some characteristic of a *sample*. E.g. Draw a sample of 10 fish from Lake Winnipeg and calculate their sample mean mercury level.
- **Quantitative Data**: data are counts or real-valued. E.g. weights of newborn babies.

Chapter 1
Section 1-2: Types of Data

- **Parameter**: a measurement describing some characteristic of a *population*. E.g. Mean mercury level of all fish in Lake Winnipeg.
- **Statistic**: a measurement describing some characteristic of a *sample*. E.g. Draw a sample of 10 fish from Lake Winnipeg and calculate their sample mean mercury level.
- **Quantitative Data**: data are counts or real-valued. E.g. weights of newborn babies.
- **Qualitative/Categorical/Attributes Data**: data can be separated into different categories/classes. E.g. genders of newborn babies, restaurant ratings, eye colour.

Chapter 1
Section 1-2: Types of Data

- **Parameter**: a measurement describing some characteristic of a *population*. E.g. Mean mercury level of all fish in Lake Winnipeg.
- **Statistic**: a measurement describing some characteristic of a *sample*. E.g. Draw a sample of 10 fish from Lake Winnipeg and calculate their sample mean mercury level.
- **Quantitative Data**: data are counts or real-valued. E.g. weights of newborn babies.
- **Qualitative/Categorical/Attributes Data**: data can be separated into different categories/classes. E.g. genders of newborn babies, restaurant ratings, eye colour.

Parameter or Statistic?

Determine whether each number is a *parameter* or a *statistic*.

(a) In a study of cloned sheep, their average age is found to be **2.7** years.

Parameter or Statistic?

Determine whether each number is a *parameter* or a *statistic*.

(a) In a study of cloned sheep, their average age is found to be **2.7** years. Ans. parameter.

Parameter or Statistic?

Determine whether each number is a *parameter* or a *statistic*.

(a) In a study of cloned sheep, their average age is found to be **2.7** years. Ans. parameter.

(b) In a random sample of fish caught from a lake, **2%** are determined to have a high Mercury level.

Parameter or Statistic?

Determine whether each number is a *parameter* or a *statistic*.

(a) In a study of cloned sheep, their average age is found to be **2.7** years. Ans. parameter.

(b) In a random sample of fish caught from a lake, **2%** are determined to have a high Mercury level. Ans. statistic.

Continuous or Discrete?

Determine whether each level of measurement is a *continuous* or *discrete*.

(a) A sample of broad-billed hummingbirds is observed, and their mean length is found to be 3.25 inches. Lengths are

Continuous or Discrete?

Determine whether each level of measurement is a *continuous* or *discrete*.

(a) A sample of broad-billed hummingbirds is observed, and their mean length is found to be 3.25 inches. Lengths are continuous.

Continuous or Discrete?

Determine whether each level of measurement is a *continuous* or *discrete*.

(a) A sample of broad-billed hummingbirds is observed, and their mean length is found to be 3.25 inches. Lengths are continuous.

(b) In a study of birds in Alaska, 312 breeding adult red Kittiwakes are banded. The number of birds banded is

Continuous or Discrete?

Determine whether each level of measurement is a *continuous* or *discrete*.

(a) A sample of broad-billed hummingbirds is observed, and their mean length is found to be 3.25 inches. Lengths are continuous.

(b) In a study of birds in Alaska, 312 breeding adult red Kittiwakes are banded. The number of birds banded is discrete.

- **Discrete Data:** data are either finite or countable, i.e. 0, 1, 2, and so on. E.g. number of traffic accidents at an intersection.

Chapter 1
Section 1-2 (Continued)

- **Discrete Data:** data are either finite or countable, i.e. 0, 1, 2, and so on. E.g. number of traffic accidents at an intersection.
- **Continuous Data:** data are real-valued. E.g. heights and weights of newborns, daily temperature, amount of milk given by cows.

Chapter 1
Section 1-2 (Continued)

- **nominal level of measurement** is characterized by data that consists of names, labels, or categories only. The data cannot be arranged in order. E.g. Survey responses where the answer is "Yes/No/Undecided"

Chapter 1
Section 1-2 (Continued)

- **nominal level of measurement** is characterized by data that consists of names, labels, or categories only. The data cannot be arranged in order. E.g. Survey responses where the answer is "Yes/No/Undecided"
- **ordinal level of measurement** data can be arranged in some order but *differences* cannot be determined or are meaningless. E.g. course grades (A,B,C,D,F), restaurant ratings: (1,2,3,4,5).

Chapter 1
Section 1-2 (Continued)

- **nominal level of measurement** is characterized by data that consists of names, labels, or categories only. The data cannot be arranged in order. E.g. Survey responses where the answer is "Yes/No/Undecided"
- **ordinal level of measurement** data can be arranged in some order but *differences* cannot be determined or are meaningless. E.g. course grades (A,B,C,D,F), restaurant ratings: (1,2,3,4,5).

- **Interval level of measurement:**
  - (a) Like ordinal data but differences are also meaningful.
  - (b) No natural zero
  - (c) E.g. Body temperature in degrees Celcius; zero degrees is arbitrary.
  - (d) E.g. Birth year; year 0 is arbitrary.
- **Ratio level of measurement: interval data plus there is a natural zero;** *difference* **and** *ratios* **are meaningful.** E.g. heights, weights, income, ages (in days).

See Table 1-1 on p.9 of textbook for a summary.

Nominal, Ordinal, Interval, or Ratio?

Determine which of the levels of measurement (nominal, ordinal, interval, or ratio) is most appropriate?

(a) Blood types (A, B, AB, O) are

Nominal, Ordinal, Interval, or Ratio?

Determine which of the levels of measurement (nominal, ordinal, interval, or ratio) is most appropriate?

(a) Blood types (A, B, AB, O) are nominal.

Nominal, Ordinal, Interval, or Ratio?

Determine which of the levels of measurement (nominal, ordinal, interval, or ratio) is most appropriate?

(a) Blood types (A, B, AB, O) are nominal.

(b) A physician's description of a patient's drinking habit recorded as one of the following: (i) abstains from alcohol, (ii) light drinker, (iii) moderate drinker, (iv) heavy drinker is

Nominal, Ordinal, Interval, or Ratio?

Determine which of the levels of measurement (nominal, ordinal, interval, or ratio) is most appropriate?

(a) Blood types (A, B, AB, O) are nominal.

(b) A physician's description of a patient's drinking habit recorded as one of the following: (i) abstains from alcohol, (ii) light drinker, (iii) moderate drinker, (iv) heavy drinker is ordinal.

Nominal, Ordinal, Interval, or Ratio?

Determine which of the levels of measurement (nominal, ordinal, interval, or ratio) is most appropriate?

(a) Blood types (A, B, AB, O) are nominal.

(b) A physician's description of a patient's drinking habit recorded as one of the following: (i) abstains from alcohol, (ii) light drinker, (iii) moderate drinker, (iv) heavy drinker is ordinal.

(c) Birth year is

Nominal, Ordinal, Interval, or Ratio?

Determine which of the levels of measurement (nominal, ordinal, interval, or ratio) is most appropriate?

(a) Blood types (A, B, AB, O) are nominal.

(b) A physician's description of a patient's drinking habit recorded as one of the following: (i) abstains from alcohol, (ii) light drinker, (iii) moderate drinker, (iv) heavy drinker is ordinal.

(c) Birth year is interval.

Nominal, Ordinal, Interval, or Ratio?

Determine which of the levels of measurement (nominal, ordinal, interval, or ratio) is most appropriate?

(a) Blood types (A, B, AB, O) are nominal.

(b) A physician's description of a patient's drinking habit recorded as one of the following: (i) abstains from alcohol, (ii) light drinker, (iii) moderate drinker, (iv) heavy drinker is ordinal.

(c) Birth year is interval.

(d) Body mass index of a random sample of patients is

Nominal, Ordinal, Interval, or Ratio?

Determine which of the levels of measurement (nominal, ordinal, interval, or ratio) is most appropriate?

(a) Blood types (A, B, AB, O) are nominal.

(b) A physician's description of a patient's drinking habit recorded as one of the following: (i) abstains from alcohol, (ii) light drinker, (iii) moderate drinker, (iv) heavy drinker is ordinal.

(c) Birth year is interval.

(d) Body mass index of a random sample of patients is ratio.

- **Observational Study**: we observe and measure specific characteristics, but we don't attempt to *modify* the subjects being studied.

Chapter 1; Section 3: Design of Experiments
Definitions

- **Observational Study**: we observe and measure specific characteristics, but we don't attempt to *modify* the subjects being studied. E.g. studies on breast feeding (with one exception).
- **Experiment:** we apply some *treatment* and then proceed to observe it effect on the subjects.

Chapter 1; Section 3: Design of Experiments
Definitions

- **Observational Study**: we observe and measure specific characteristics, but we don't attempt to *modify* the subjects being studied. E.g. studies on breast feeding (with one exception).
- **Experiment:** we apply some *treatment* and then proceed to observe it effect on the subjects. E.g. Harvard Physicians' Health Study (aspirin)
- **Retrospective study**: data are collected from the past by going back in time (through examination of records, interviews, etc.)

Chapter 1; Section 3: Design of Experiments
Definitions

- **Observational Study**: we observe and measure specific characteristics, but we don't attempt to *modify* the subjects being studied. E.g. studies on breast feeding (with one exception).
- **Experiment:** we apply some *treatment* and then proceed to observe it effect on the subjects. E.g. Harvard Physicians' Health Study (aspirin)
- **Retrospective study**: data are collected from the past by going back in time (through examination of records, interviews, etc.)
- **Prospective/longitudinal/cohort study**: data are collected from future groups sharing common factors.

Examples of Prospective & Retrospective Studies
Pages 10-13 of the textbook

- Prospective: *The Manitoba Health Study*; See
  `http://www.mfus.ca/History.php`.
- Prospective: *The Framingham Health Study*; See `https://www.framinghamheartstudy.org/about-fhs/history.php`.

Examples of Prospective & Retrospective Studies
Pages 10-13 of the textbook

- Prospective: *The Manitoba Health Study*; See
  http://www.mfus.ca/History.php.
- Prospective: *The Framingham Health Study*; See https://www.
  framinghamheartstudy.org/about-fhs/history.php.
- Prospective: *CHILD Study*; See
  http://www.canadianchildstudy.ca/

Examples of Prospective & Retrospective Studies
Pages 10-13 of the textbook

- Prospective: *The Manitoba Health Study*; See
  `http://www.mfus.ca/History.php`.
- Prospective: *The Framingham Health Study*; See `https://www.framinghamheartstudy.org/about-fhs/history.php`.
- Prospective: *CHILD Study*; See
  `http://www.canadianchildstudy.ca/`
- Retrospective/Case Control Study: Go back in time and collect the data. E.g. `http://www.ghahramani.ca/uploads/1/7/0/4/17042208/sim_01.pdf`