Basic Statistics

DigitalSkola
Uncover The World Of Digital Skills With Us

Profile

linkedin.com/in/romansyasetyo/

DigitalSkola
Uncover The World Of Digital Skills With Us

A Data scientist is one who knows more statistics than a programmer and more programming than a statistician.

- Josh Wills

DigitalSkola
Uncover The World Of Digital Skills With Us

Table of Content

What will We Learn Today?

1. Introduction to Statistics
2. Data Types in Statistics
3. Descriptive Statistics

DigitalSkola
Uncover The World Of Digital Skills With Us

Introduction to Statistics

DigitalSkola
Uncover The World Of Digital Skills With Us

Berapa persen pengguna mobil di Jakarta?

DigitalSkola
Uncover The World Of Digital Skills With Us

Bagaimana tingkat kepuasan customer?

DigitalSkola
Uncover The World Of Digital Skills With Us

Bagaimana revenue company dalam enam bulan terakhir?

Statistics

- Statistics is the science of **collecting, analyzing, presenting, and interpreting data**.

- Menarik kesimpulan :

  - Descriptive Statistics

    summarize the characteristics of a data set

  - Statistical Inference

    test a hypothesis or assess whether the data is generalizable to the broader population

Statistics

Descriptive Statistics

- **Describe the features** of populations and/or samples.
- Organize and present data in a purely **factual way**.
- Present final results visually, using **tables**, **charts**, or **graphs**.
- Draw conclusions based on **known data**.
- Use measures like **central tendency**, **distribution**, and **variance**.

Statistical Inference

- Use samples to **make generalizations** about larger populations.
- Help us to **make estimates** and predict future outcomes.
- Present final results in the form of **probabilities**.
- Draw conclusions that go **beyond the available data**.
- Use techniques like **hypothesis testing**, **confidence intervals**, and **regression and correlation analysis**.

Statistics : Population & Sample

- **Population**

  The entire group that you wish to draw data from

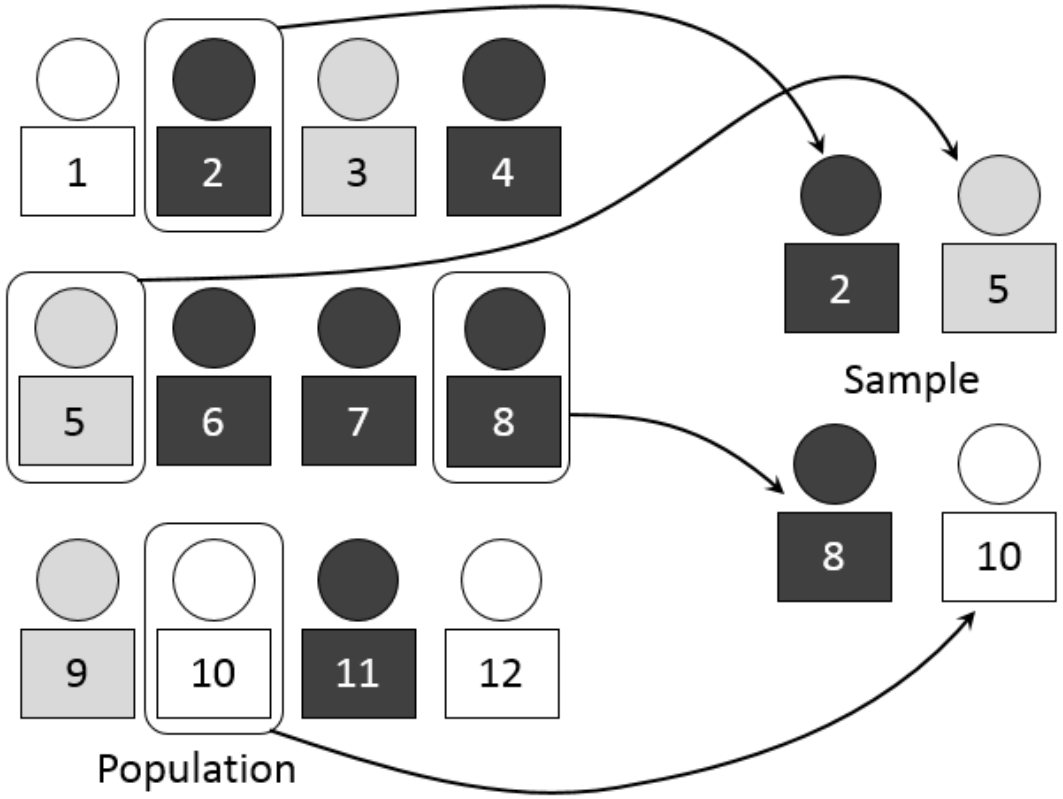
  (Contoh : Seluruh Masyarakat Jakarta)

- **Sample**

  A representative group of a larger population

  (Contoh : beberapa orang yang terpilih pada masing-masing kota di Jakarta)

Statistics : Population & Sample



Statistics : Population & Sample

Raka pergi membeli buah lengkeng di toko buah. Kumpulan buah lengkeng ditumpuk pada sebuah rak. Untuk memastikan rasa dari buah lengkeng tersebut, **Raka** kemudian mengambil beberapa buah lengkeng dan dicoba langsung. Karena buah lengkeng yang dicobanya manis, maka Raka pun membeli buah di toko tersebut.

Statistics : Population & Sample

Populasi :

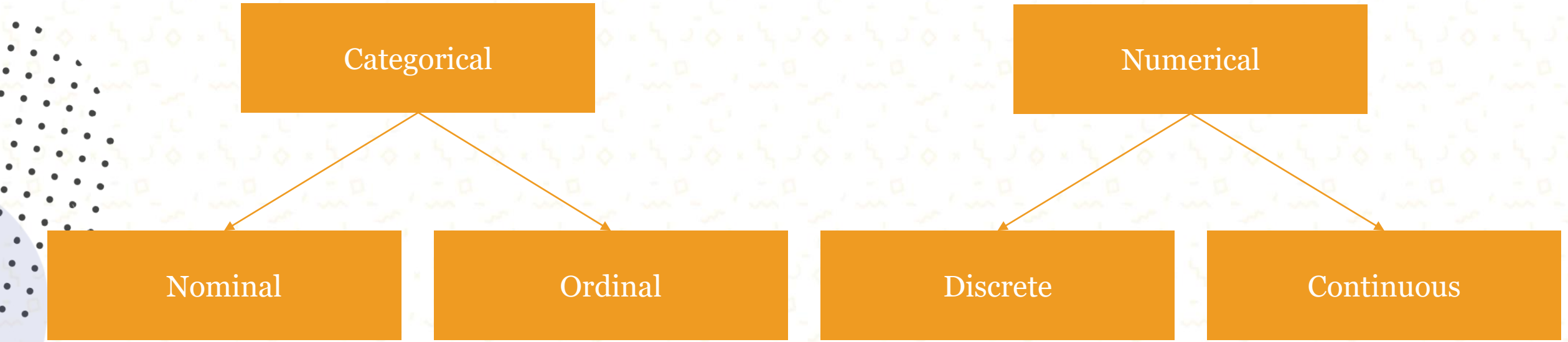
Populasi pada ilustrasi di atas adalah **seluruh buah lengkeng** yang ada pada rak di toko buah tadi.

Sampel :

Buah yang Raka ambil untuk dicicipi dengan tujuan untuk menentukan rasa manisnya.

DigitalSkola
Uncover The World Of Digital Skills With Us

Data Types in Statistics



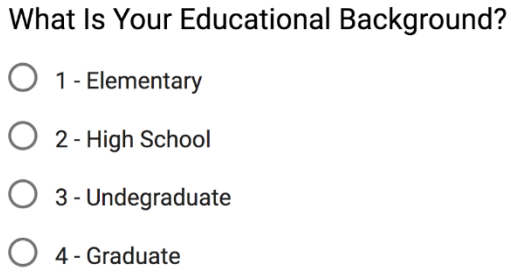
Categorical

- Nominal

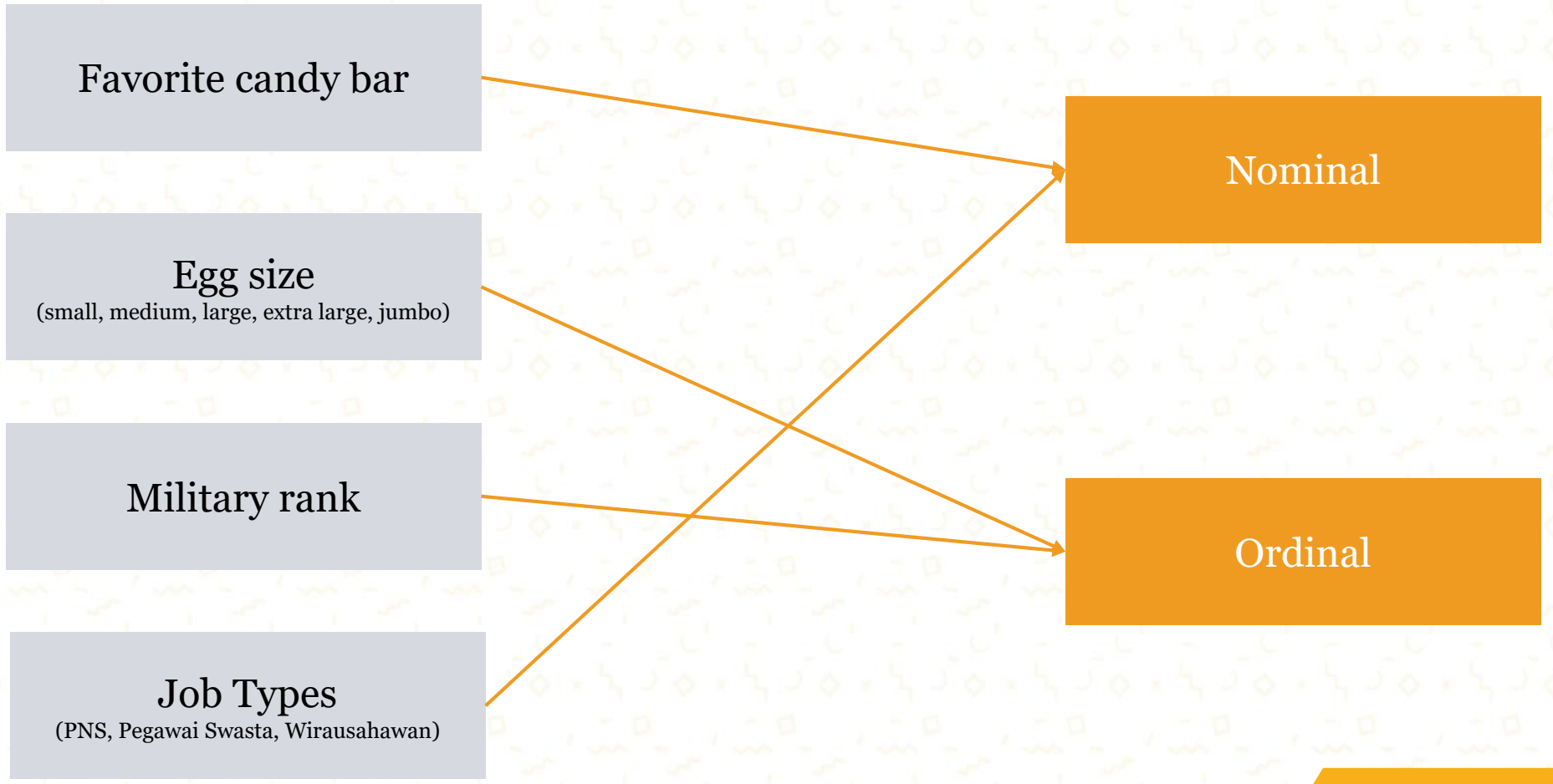
  Variable with categories that **do not have** a natural order or ranking

- Ordinal

  Variable with categories that **have** a natural order or ranking



Categorical



Numerical

- Discrete

  This type of data **can't be measured** but **it can be counted** (1,2,3,...)

  Jumlah murid di kelas

  Jumlah kelengkeng dalam keranjang

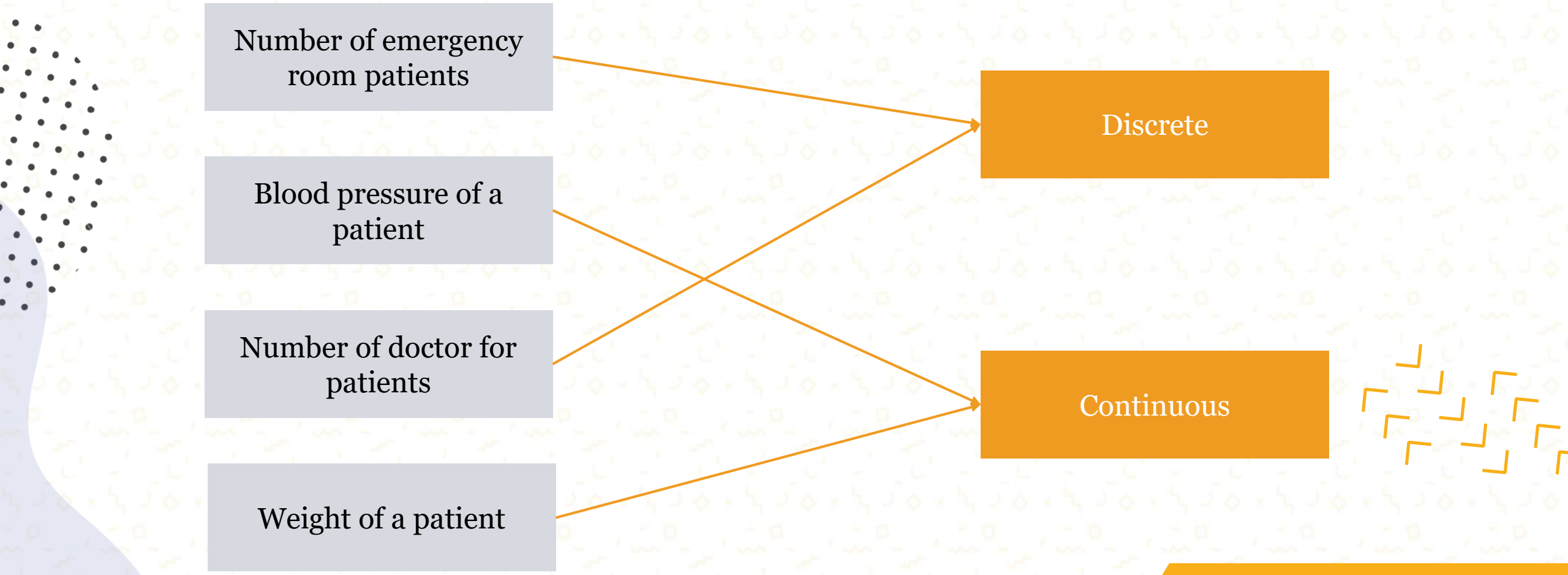
- Continuous

  Represents measurements and therefore the values **can't be counted** but they **can be measured**

  Temperatur ruangan

  Kecepatan angin

Numerical



DigitalSkola
Uncover The World Of Digital Skills With Us

Descriptive Statistics

Descriptive Statistics

Descriptive statistics are used to **describe the characteristics** or features of a dataset.

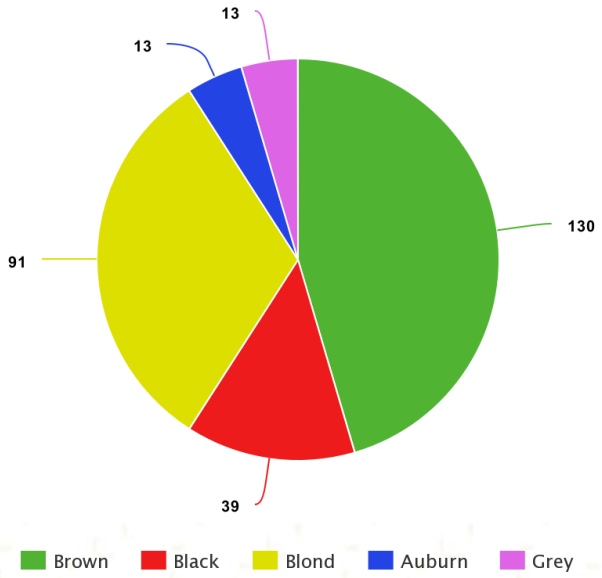
- Distribution
- Central Tendency
- Variability

Descriptive Statistics : Distribution

Distribution shows us the **frequency** of different outcomes (or data points) in a population or sample.

As a basic example, the following list shows the number of those with different hair colors in a dataset of 286 people.

- Brown hair: 130
- Black hair: 39
- Blond hair: 91
- Auburn hair: 13
- Gray hair: 13



Descriptive Statistics : Central Tendency

Central tendency is the name for measurements that look at the typical central values within a dataset.

- The **mean**: The average value of all the data points.
- The **median**: The central or middle value in the dataset.
- The **mode**: The value that appears most often in the dataset.

DigitalSkola
Uncover The World Of Digital Skills With Us

Descriptive Statistics : Central Tendency

Mean

The average value of all the data points

$$\mu = \frac{x_1 + \cdots + x_n}{n}$$

3,3,4,4,4,4,4,4,4,5,5,5,6,7,7,7,7,9,9,10

Mean = (3+3+4+4+4+4+4+4+4+5
+5+5+6+7+7+7+7+9+9+10)/20
= 111 / 20
= **5.55**

Descriptive Statistics : Central Tendency

Median

The central or middle value in the dataset

Median

n is odd,

$$\text{Median} = \left(\frac{n+1}{2}\right)^{th} \text{ observation}$$

n is even,

$$\text{Median} = \frac{\left(\frac{n}{2}\right)^{th} + \left(\frac{n}{2}+1\right)^{th} \text{ observation}}{2}$$

Letak ke- 20/2 = letak ke- 10 = 5
Letak ke- 20/2 + 1 = letak ke- 11 = 5

Median = (5+5)/2 = **5**

3,3,4,4,4,4,4,4,4,5,5,5,6,7,7,7,7,9,9,10

Descriptive Statistics : Central Tendency

Mode

The value that appears most often in the dataset

3,3,4,4,4,4,4,4,4,5,5,5,6,7,7,7,7,9,9,10

Data	3	4	5	6	7	9	10
Freq	2	7	3	1	4	2	1

Mode = 4

Descriptive Statistics : Variability

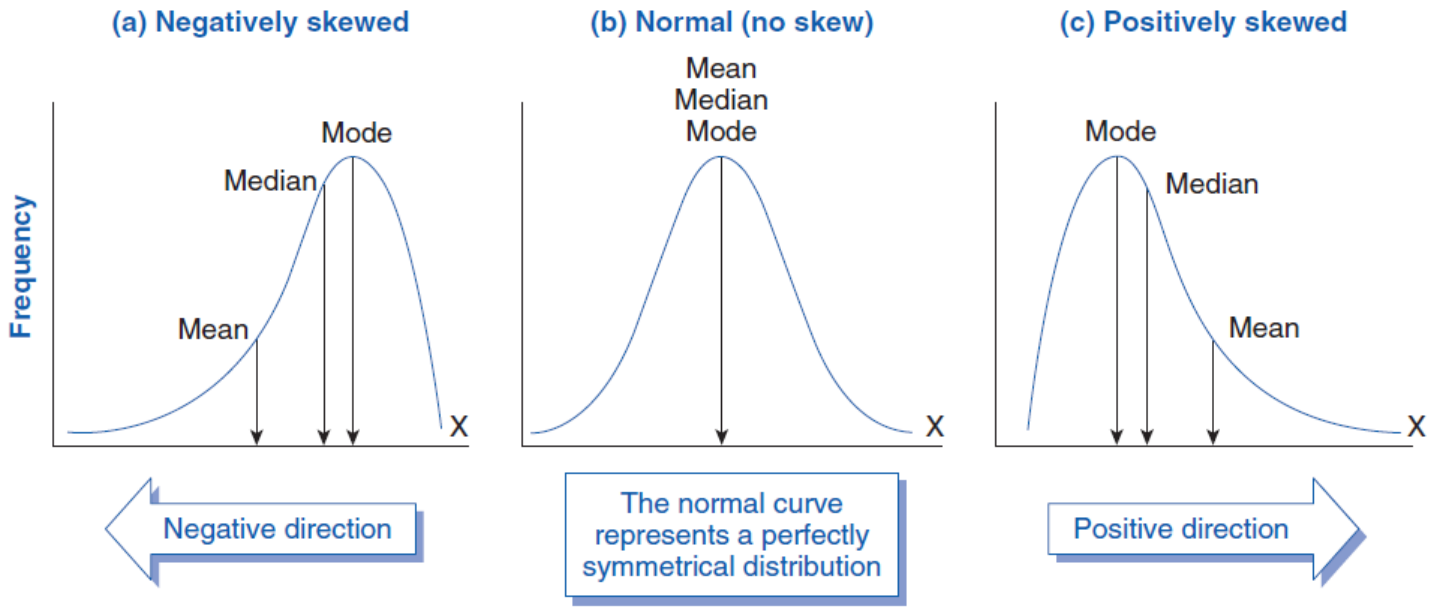
Variability describes how values are distributed or spread out.

- **Standard deviation**: This shows us the amount of variation or dispersion. Low standard deviation implies that most values are close to the mean. High standard deviation suggests that the values are more broadly spread out.
- **Minimum and maximum values**: These are the highest and lowest values in a dataset or quartile
- **Range**: This measures the size of the distribution of values.
- **Skewness**: This is a measure of a dataset's symmetry.

Descriptive Statistics : Variability

Skewness

This is a measure of a dataset's symmetry



DigitalSkola

Uncover The World Of Digital Skills With Us

Thank You