Nature of Data

Types of data

- Based on Data type
  - Numerical
    - Discrete
      - Likert Scale
      - Quantitative (counts)
    - Continuous
  - Categorical: Nominal and Ordinal
- Based on Time
  - Cross-sectional
  - Time-series
  - Panel Data

Discrete: Likert Scale

- An ordered, one-dimensional scale from which respondents choose one option that best aligns with their view

- e.g.
  - Question: Eating pizza in the evening more preferable than eating it at the dinner.
    1. Strongly Disagree
    2. Disagree
    3. Neither agree nor disagree
    4. Agree
    5. Strongly Agree

Discrete: Quantitative

- The data with particular values.

- e.g.
  - Count of Customers coming in a particular time slot
  - Count of patients of a particular disorder

Continuous

- Data taking any real number value
- e.g.
  - Height
  - Weight
  - Volume in litres
  - Sales

Short Quiz: Classify the data

Classify each set of data as discrete or continuous

1. The number of suitcases lost on a railway platform.
2. The height of sugarcane plants.
3. The number of sugarcane sticks grown in a plot.
4. The time it takes for a car battery to die.
5. The production of sugar by weight.
6. The opinion of passengers about the airport administration.

Short Quiz : Classify the data

Classify each set of data as discrete or continuous
1. Discrete: The number of suitcases lost on a railway platform.
2. Continuous: The height of sugarcane plants.
3. Discrete: The number of sugarcane sticks grown in a plot.
4. Continuous: The time it takes for a car battery to die.
5. Continuous: The production of sugar by weight.
6. Likert Scale: The opinion of passengers about the airport administration.

Categorical: Nominal

- **Nominal** level data is made up of values that are distinguished by name only.

- There is no standard ordering scheme to this data.

- e.g.
  - Colours in a fabric produced by a company
  - Flavours of corn flakes available in the market

Categorical: Ordinal

- There is an ordering scheme in Ordinal Scale.

- e.g.
  - Movies on a certain TV show are classified as 2 thumbs up, 1 thumb up, or 0 thumbs up.
  - The dealers are classified as follows depending on the quality of service they provide

Short Quiz : Identify the data

1. Letter grades on an English essay.

2. Flavors of yogurt.

3. Instructors classified as : Easy, Difficult or Impossible.

4. Employee evaluations classified as : Excellent, Average, Poor.

5. Political parties.

6. Ice cream flavour.

7. Students classified by their reading ability : Above average, Below average, Normal.

Short Quiz : Identify the data

1. Ordinal: Letter grades on an English essay.
2. Nominal: Flavors of yogurt.
3. Ordinal: Instructors classified as : Easy, Difficult or Impossible.
4. Ordinal: Employee evaluations classified as : Excellent, Average, Poor.
5. Nominal: Political parties.
6. Nominal: Ice cream flavour.
7. Ordinal: Students classified by their reading ability : Above average, Below average, Normal.

Types of data

- Based on Data type
  - Numerical
    - Discrete
      - Likert Scale
      - Quantitative (counts)
    - Continuous
  - Categorical: Nominal and Ordinal
- Based on Time
  - Cross-sectional
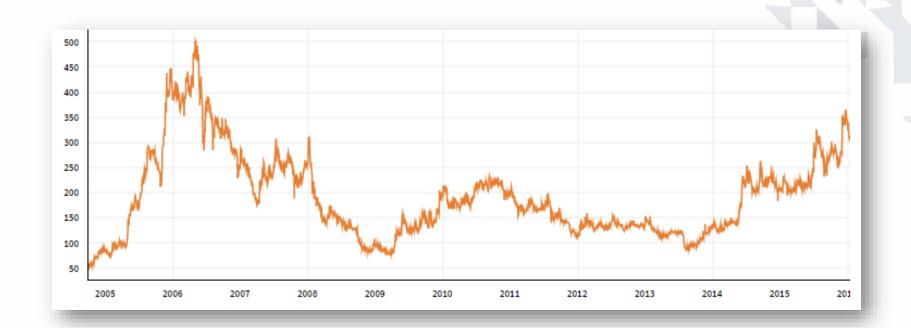 - Time-series
  - Panel Data

Cross-sectional

- Refers to data collected by observing many subjects (such as individuals, firms or countries/regions) at the same point of time or without regard to differences in time
- e.g.
  – Data of Home and Automobile insurance policies sold at different branches and their total operating cost in a specific year

Branch	Home	Automobile	Operating_Cost
B01	400	1200	124000
B02	350	360	71000
B03	600	800	136000
B04	800	1800	219000
B05	900	1600	230000
B06	200	1000	75000
B07	120	900	56000
B08	340	1100	110000
B09	490	900	120000
B10	700	800	144000

Time Series

- A sequence of data points, measured typically at successive times spaced at uniform time intervals
- e.g.
  - Month-wise Stock Price of a company

Panel Data

- Panel Data is mixture of time series analysis and cross-sectional data

StoreID	Month	Sales
1	Jan	58000
2	Jan	59000
3	Jan	36000
1	Feb	15000
2	Feb	14700
3	Feb	15400
1	Mar	18000
2	Mar	12500