

Causes & Impact of Missing Values

Important Concepts

1. Handling Missing Values.
2. Handling Outliers.
3. Learning some Important Data Manipulation Functions useful for Cleaning the Data.
4. How to Deal with Real World Problems.

Missing Value

- Missing values occur when there is no data or value stored for the variable in an observation.
- Missing data are a common occurrence and can have a significant effect on the conclusions drawn from the data.
- There can be several causes for the occurrence of missing values in a data set.

Example

- The data collected from a survey can have missing values because some participants refuse to answer some or all of the questions or they do not know the answer to any question or the answer has been skipped.
- **Conclusion:** Data can be missing due to Multiple Reasons such as accidental or data entry error.

Problems with Missing Data

1. Most statistical procedures require a value for each variable.
2. The missing data can cause a bias in the estimation of parameters and the accuracy of our machine learning model can be affected.