Data Exploration

Stages of Data Exploration

Reading the Data

Univariate Analysis

Missing Value Treatment

Variable Transformation

Variable Identification

Bi-variate Analysis

Outlier Treatment

Analytics Vidhya
Learn everything about analytics

Reading the Data

Two common types of file formats.

1. Csv format - .csv

2. Excel format - .xlsx/.xls

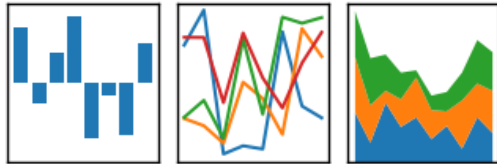# Reading the Data

Reading csv file – *read_csv()*

Size of the data - *shape*

Name of the variables in the dataset - *columns*

Look at top 5 observations in the dataset – *head()*