# Exploratory Data Analysis Report

*2019-03-15*

- Exploratory Data analysis (EDA)
    - 1. Overview of the data
    - 2. Summary of numerical variables
    - 3. Distributions of Numerical variables
        - Quantile-quantile plot for Numerical variables - Univariate
        - Density plots for Numerical variables - Univariate
        - Box plots for all numeric features vs categorical dependent variable - Bivariate comparision only with categories
    - 4. Summary of categorical variables
    - 5. Distributions of categorical variables

# Exploratory Data analysis (EDA)

Analyzing the data sets to summarize their main characteristics of variables, often with visual graphs, without using a statistical model.

## 1. Overview of the data

Understanding the dimensions of the dataset, variable names, overall missing summary and data types of each variables

```
# Overiview of the data
ExpData(data=data,type=1)
# Structure of the data
ExpData(data=data,type=2)
```

**Overview of the data**

| Descriptions<br><fctr> | Obs<br><fctr> |
|---|---|
| Sample size (Nrow) | 1000 |
| No. of Variables (Ncol) | 9 |
| No. of Numeric Variables | 2 |
| No. of Factor Variables | 3 |
| No. of Text Variables | 2 |
| No. of Logical Variables | 2 |
| No. of Date Variables | 0 |
| No. of Zero variance Variables (Uniform) | 0 |
| %. of Variables having complete cases | 55.56% (5) |
| %. of Variables having <50% missing cases | 44.44% (4) |
| 1-10 of 12 rows | Previous **1** 2 Next |

**Structure of the data**