

# Data Profiling Report

- Basic Statistics
  - Data Structure (Text)
  - Data Structure (Network Graph)
- Missing Values
- Data Distribution
  - Continuous Features (Histogram)
  - Discrete Features (Bar Chart)
- Correlation Analysis

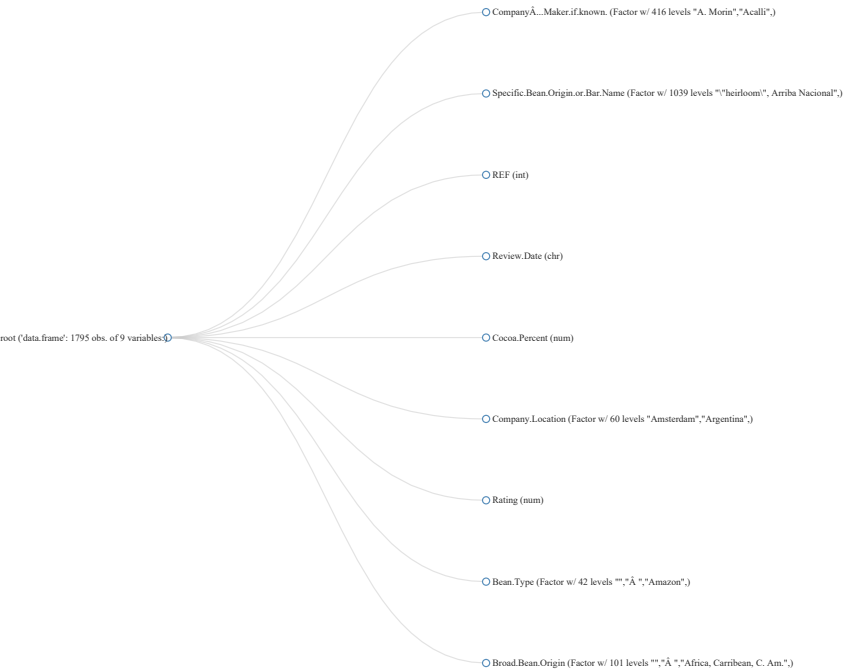
## Basic Statistics

The data is **202.5 Kb** in size. There are **1,795** rows and **9** columns (features). Of all **9** columns, **6** are discrete, **3** are continuous, and **0** are all missing. There are **0** missing values out of **16,155** data points.

### Data Structure (Text)

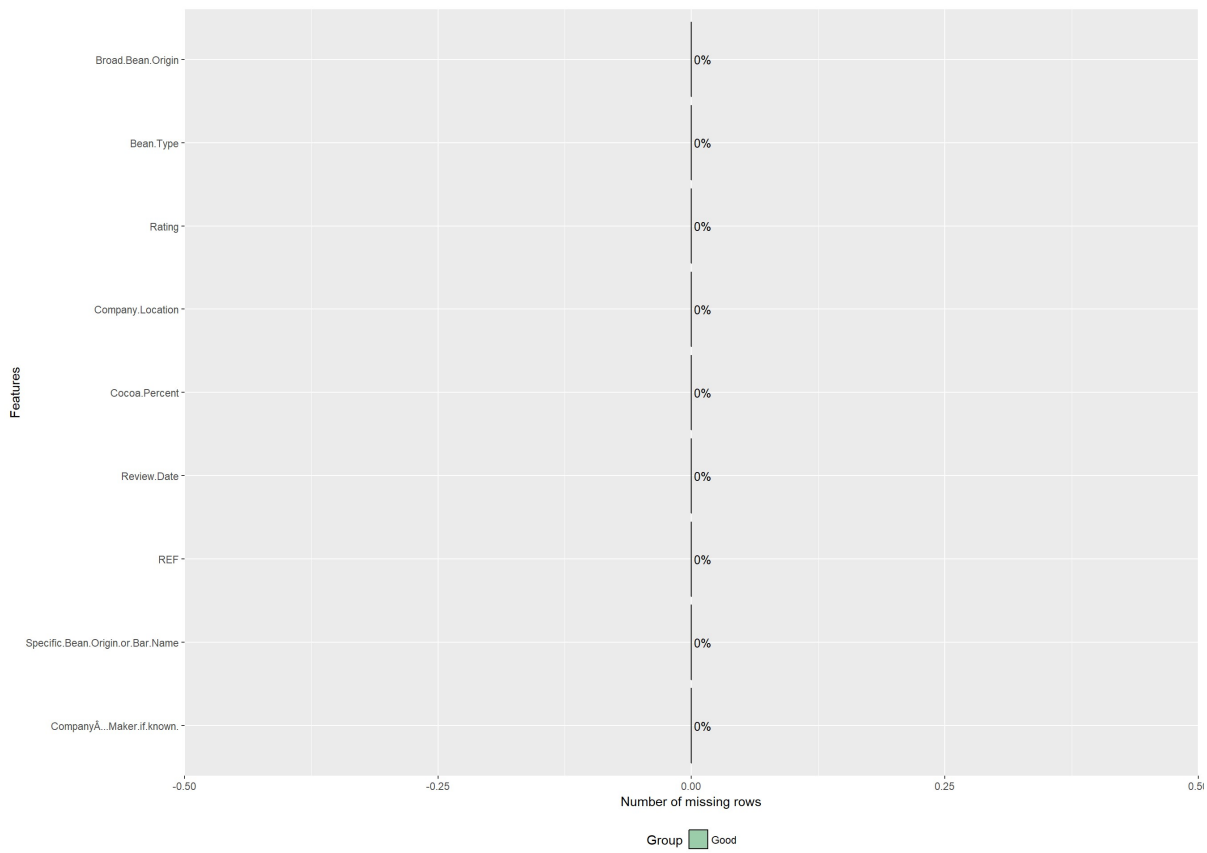
```
## 'data.frame': 1795 obs. of 9 variables:
## $ Company...Maker.if.known. : Factor w/ 416 levels "A. Morin","Acalli",...: 1 1 1 1 1 1 1 1 1 ...
## $ Specific.Bean.Origin.or.Bar.Name: Factor w/ 1039 levels "\"heirloom\"", Arriba Nacional",...: 15 494 68 16 813 175 288 9
23 805 731 ...
## $ REF : int 1876 1676 1676 1680 1704 1315 1315 1315 1319 1319 ...
## $ Review.Date : chr "2016" "2015" "2015" "2015" ...
## $ Cocoa.Percent : num 63 70 70 70 70 70 70 70 70 ...
## $ Company.Location : Factor w/ 60 levels "Amsterdam","Argentina",...: 19 19 19 19 19 19 19 19 19 ...
## $ Rating : num 3.75 2.75 3 3.5 3.5 2.75 3.5 3.5 3.75 4 ...
## $ Bean.Type : Factor w/ 42 levels "\"", "Ã ", "Amazon",...: 2 2 2 2 2 10 2 10 10 2 ...
## $ Broad.Bean.Origin : Factor w/ 101 levels "\"", "Ã ", "Africa, Carribean, C. Am.",...: 70 80 80 80 57 93 18 9
3 93 57 ...
```

Data Structure (Network Graph)



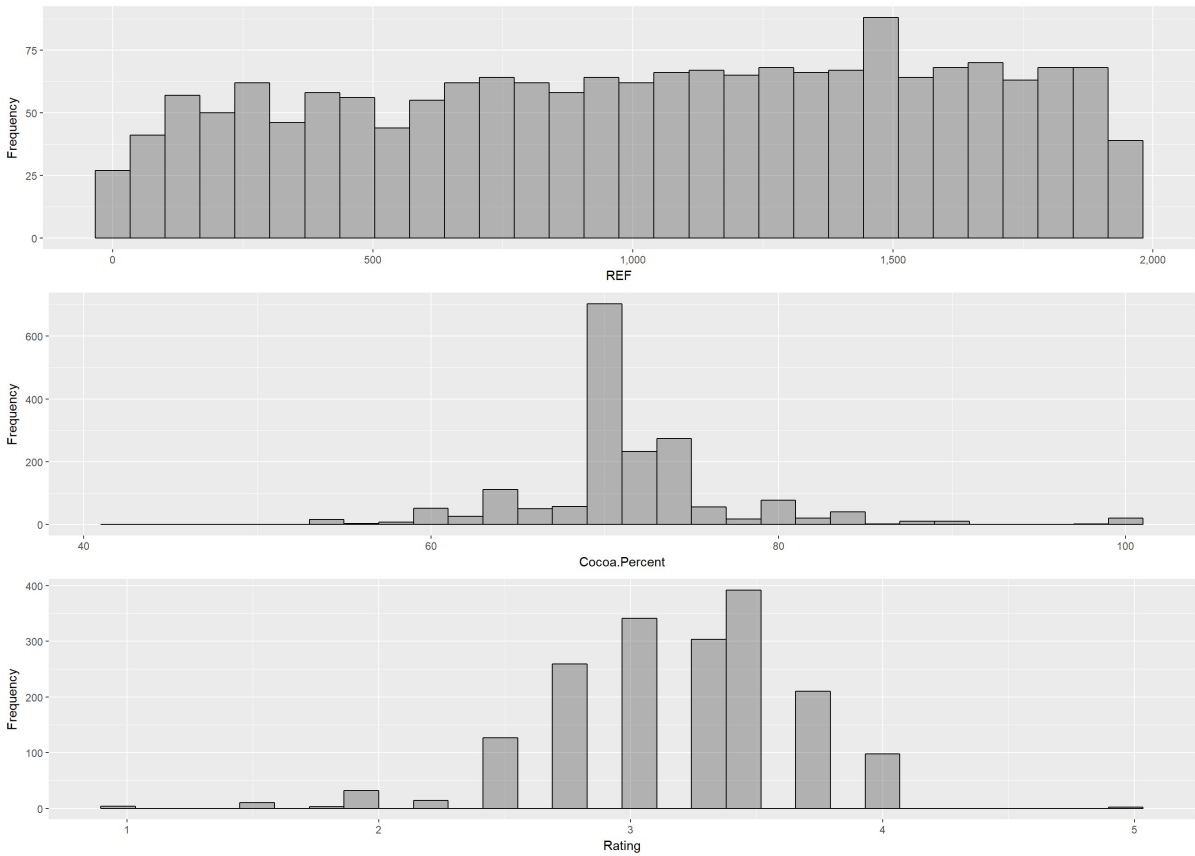
Missing Values

The following graph shows the distribution of missing values.



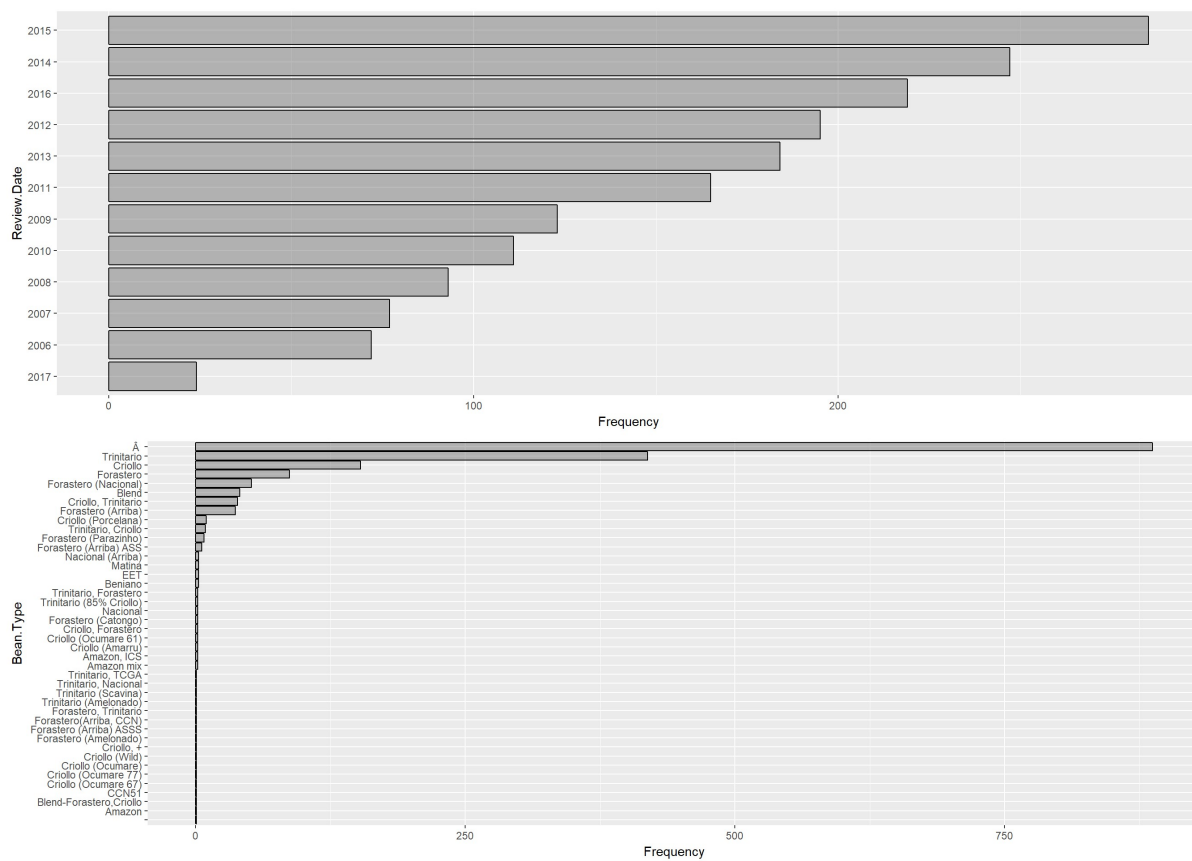
# Data Distribution

## Continuous Features (Histogram)



## Discrete Features (Bar Chart)

```
## 4 columns ignored with more than 50 categories.  
## Company's Maker,if.known.: 416 categories  
## Specific.Bean.Origin.or.Bar.Name: 1039 categories  
## Company.Location: 60 categories  
## Broad.Bean.Origin: 101 categories
```



## Correlation Analysis

```
## 5 features with more than 20 categories ignored!
## CompanyÀ...Maker.if.known.: 416 categories
## Specific.Bean.Origin.or.Bar.Name: 1039 categories
## Company.Location: 60 categories
## Bean.Type: 42 categories
## Broad.Bean.Origin: 101 categories
```

