

# Data Profiling Report

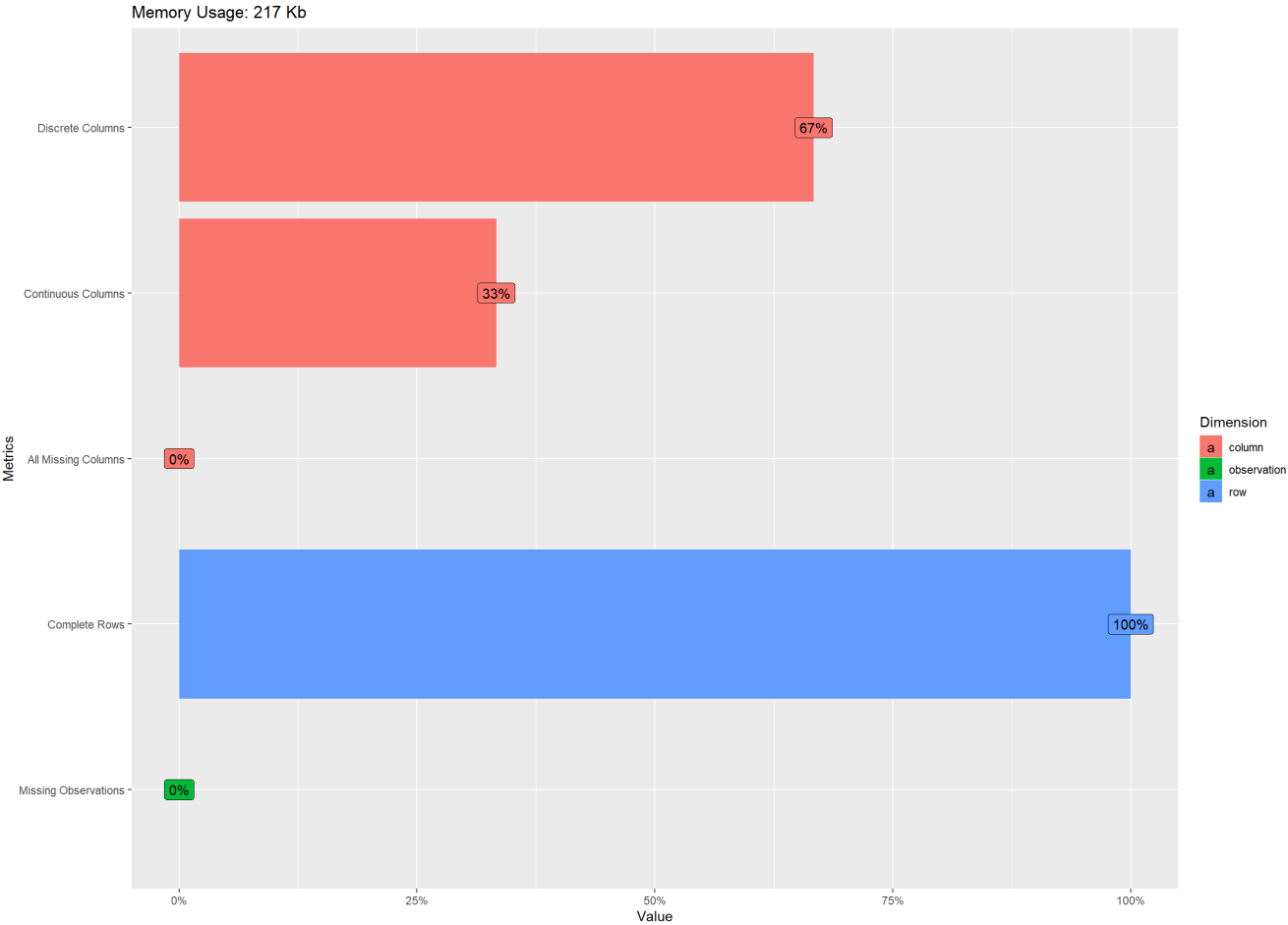
- Basic Statistics
  - Raw Counts
  - Percentages
- Data Structure
- Missing Data Profile
- Univariate Distribution
  - Histogram
  - Bar Chart (by frequency)
  - QQ Plot
- Correlation Analysis
- Principal Component Analysis

## Basic Statistics

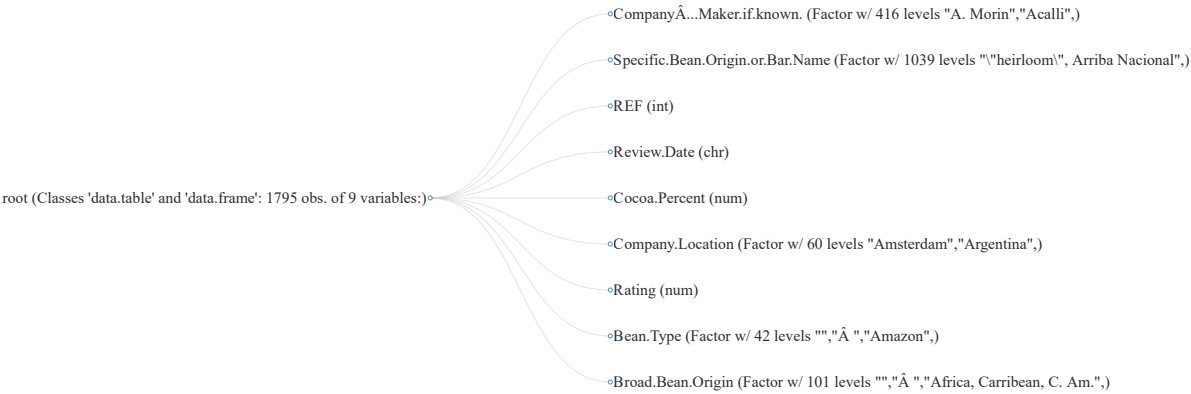
### Raw Counts

Name	Value
Rows	1,795
Columns	9
Discrete columns	6
Continuous columns	3
All missing columns	0
Missing observations	0
Complete Rows	1,795
Total observations	16,155
Memory allocation	217 Kb

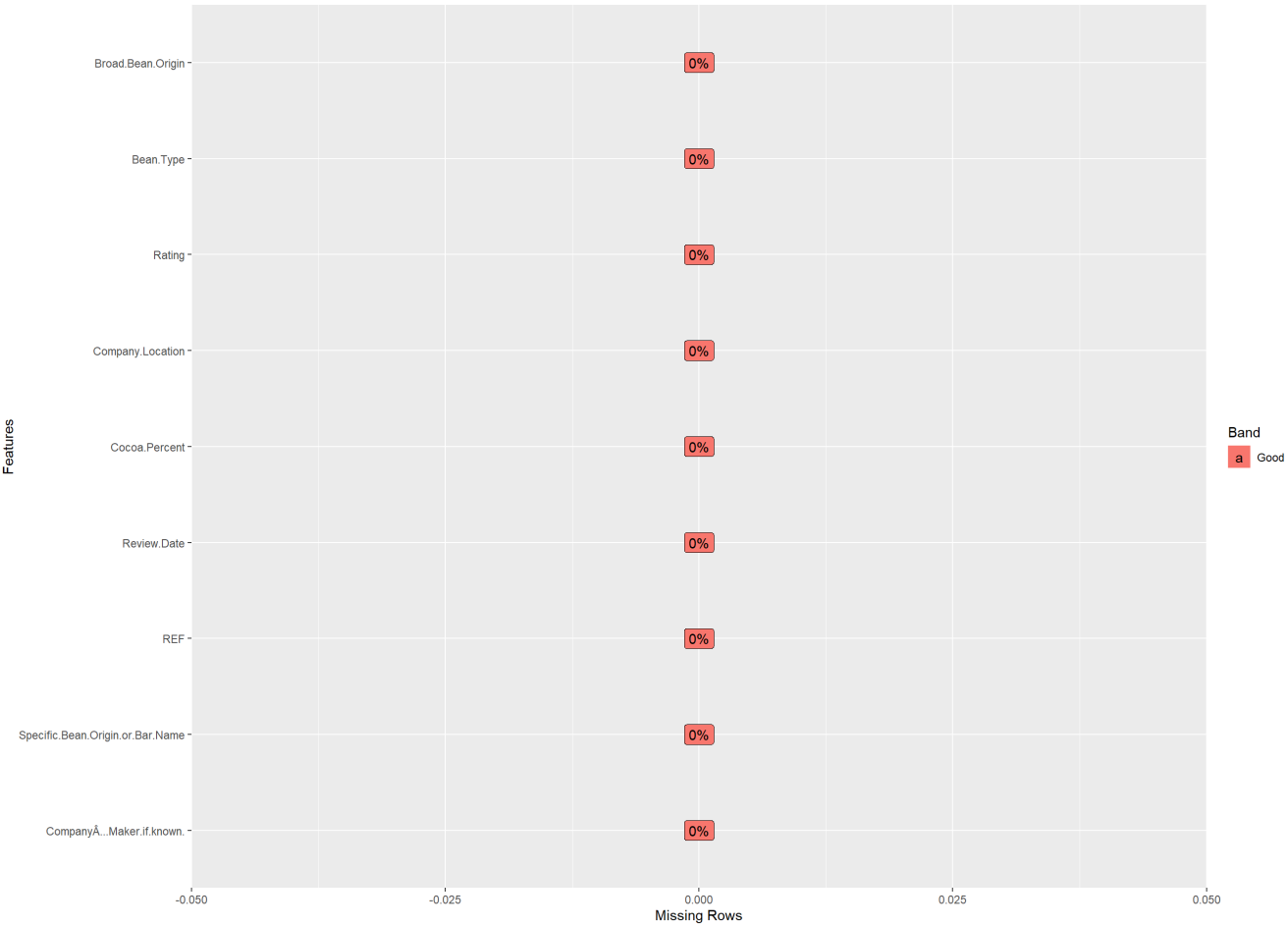
### Percentages



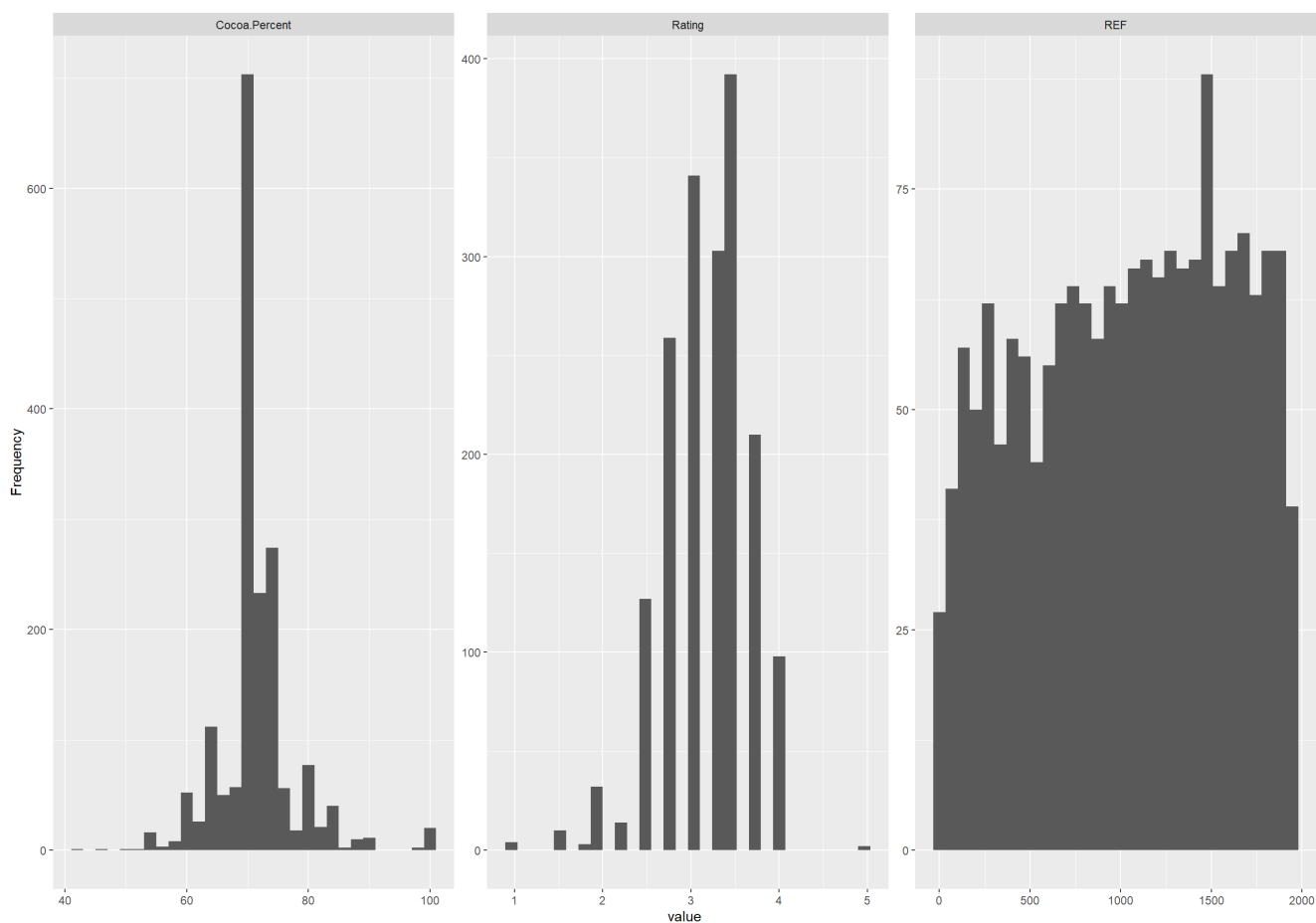
Data Structure



Missing Data Profile

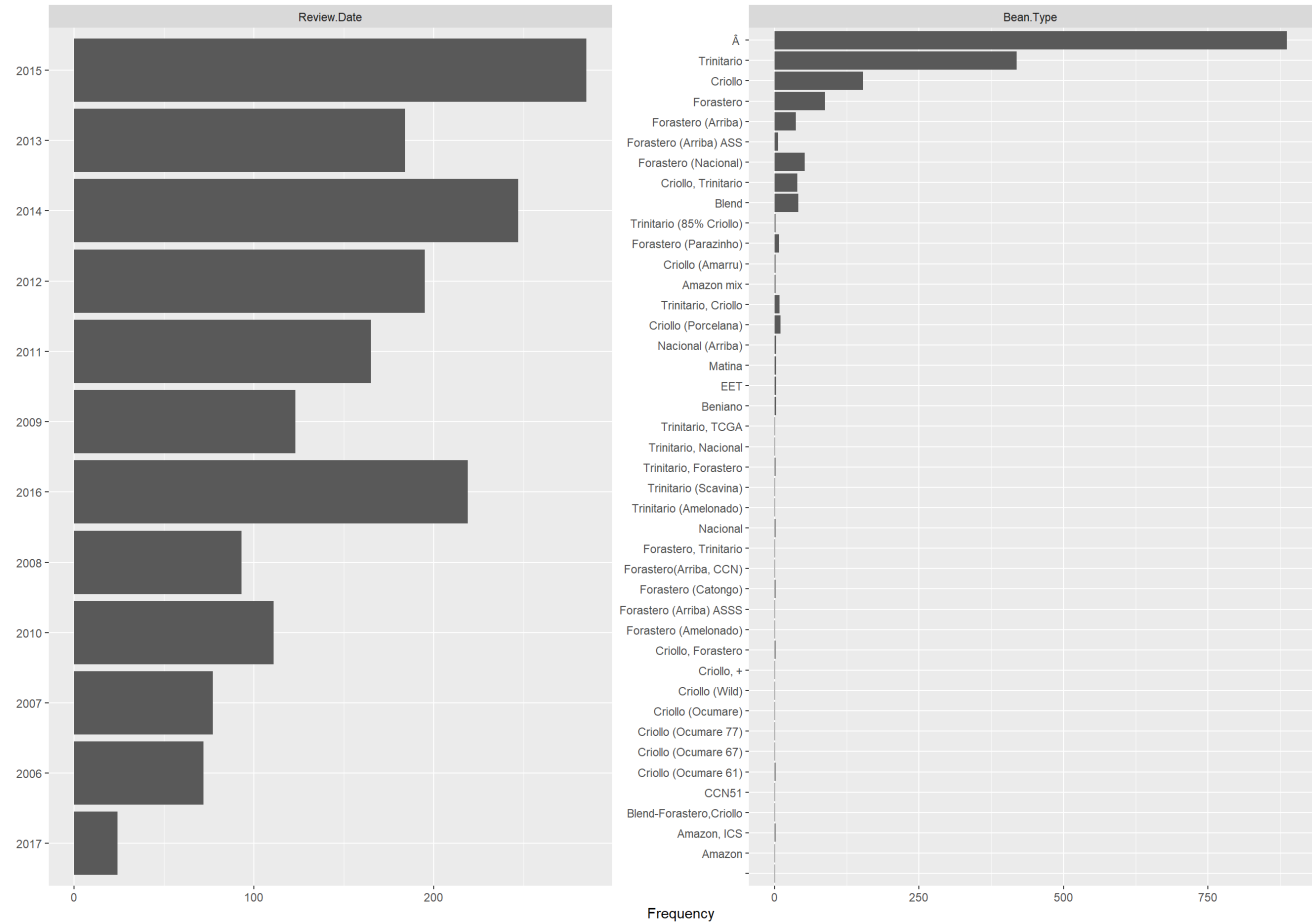


Univariate Distribution  
Histogram

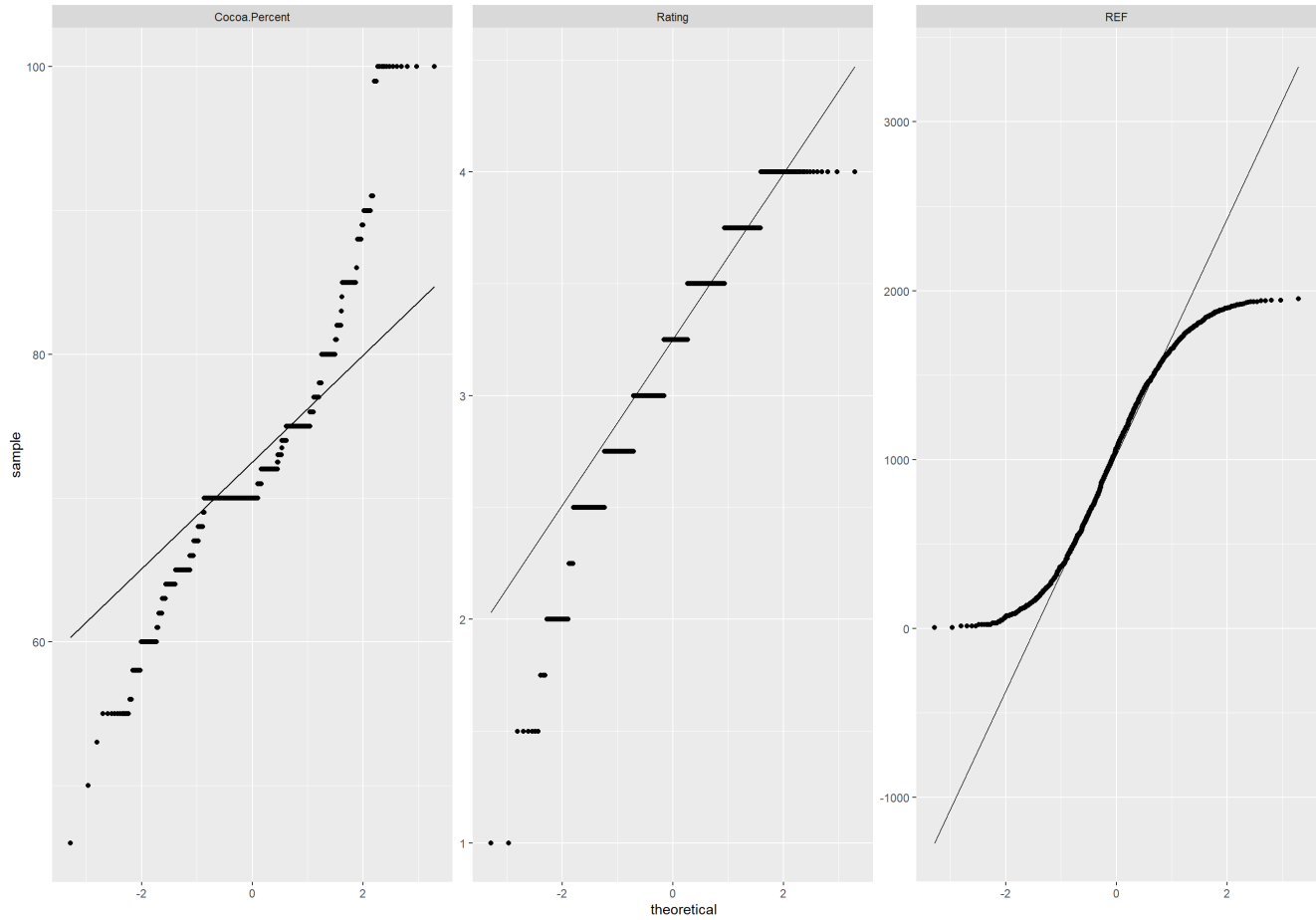


### Bar Chart (by frequency)

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

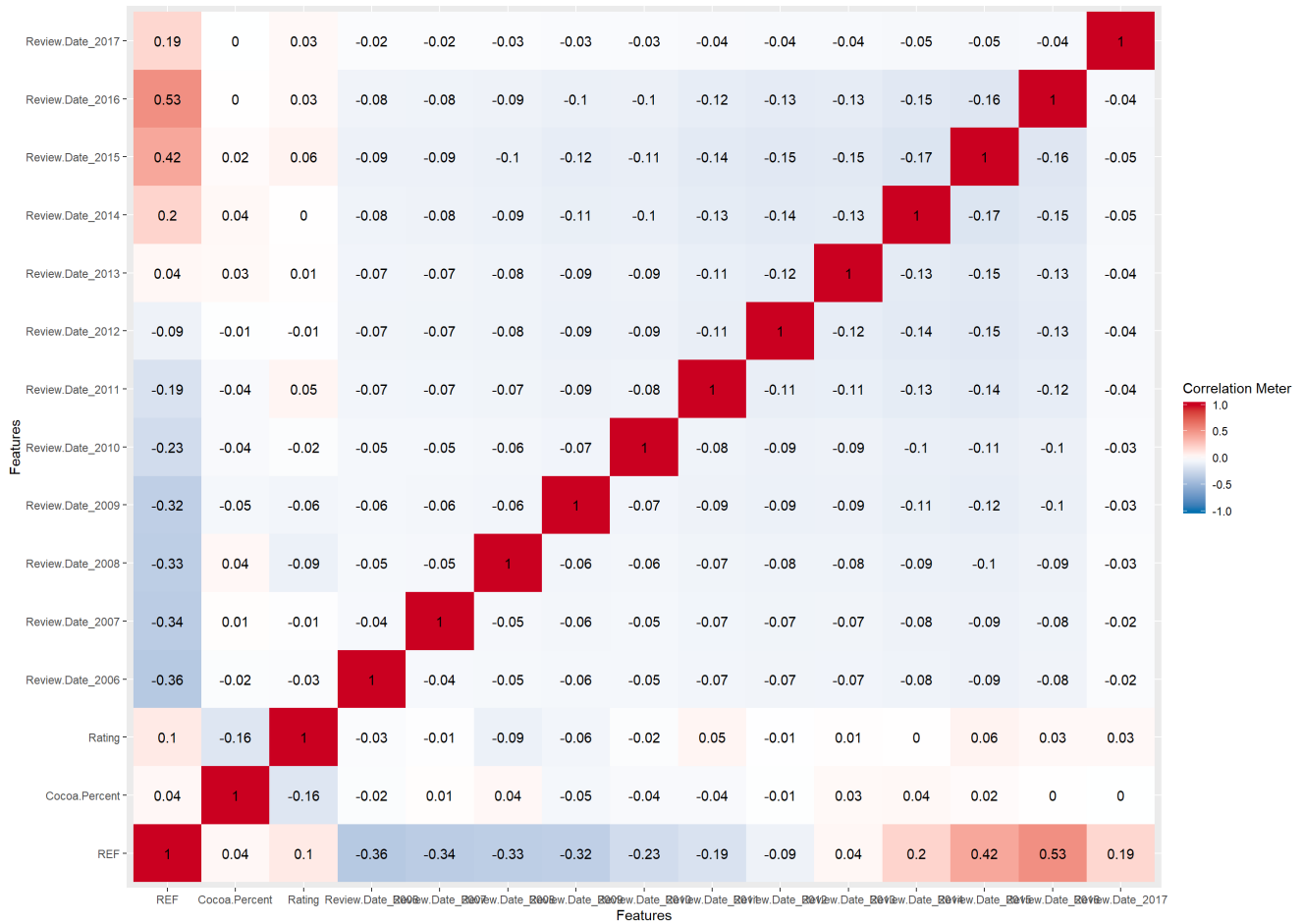


QQ Plot



Correlation Analysis

```
## 5 features with more than 20 categories ignored!  
## CompanyA...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
```



Principal Component Analysis

```
## 4 features with more than 50 categories ignored!  
## CompanyA...Maker.if.known.: 416 categories  
## Specific.Bean.Origin.or.Bar.Name: 1039 categories  
## Company.Location: 60 categories  
## Broad.Bean.Origin: 101 categories
```





% Variance Explained By Principal Components  
(Note: Labels indicate cumulative % explained variance)

