

# Data Profiling Report

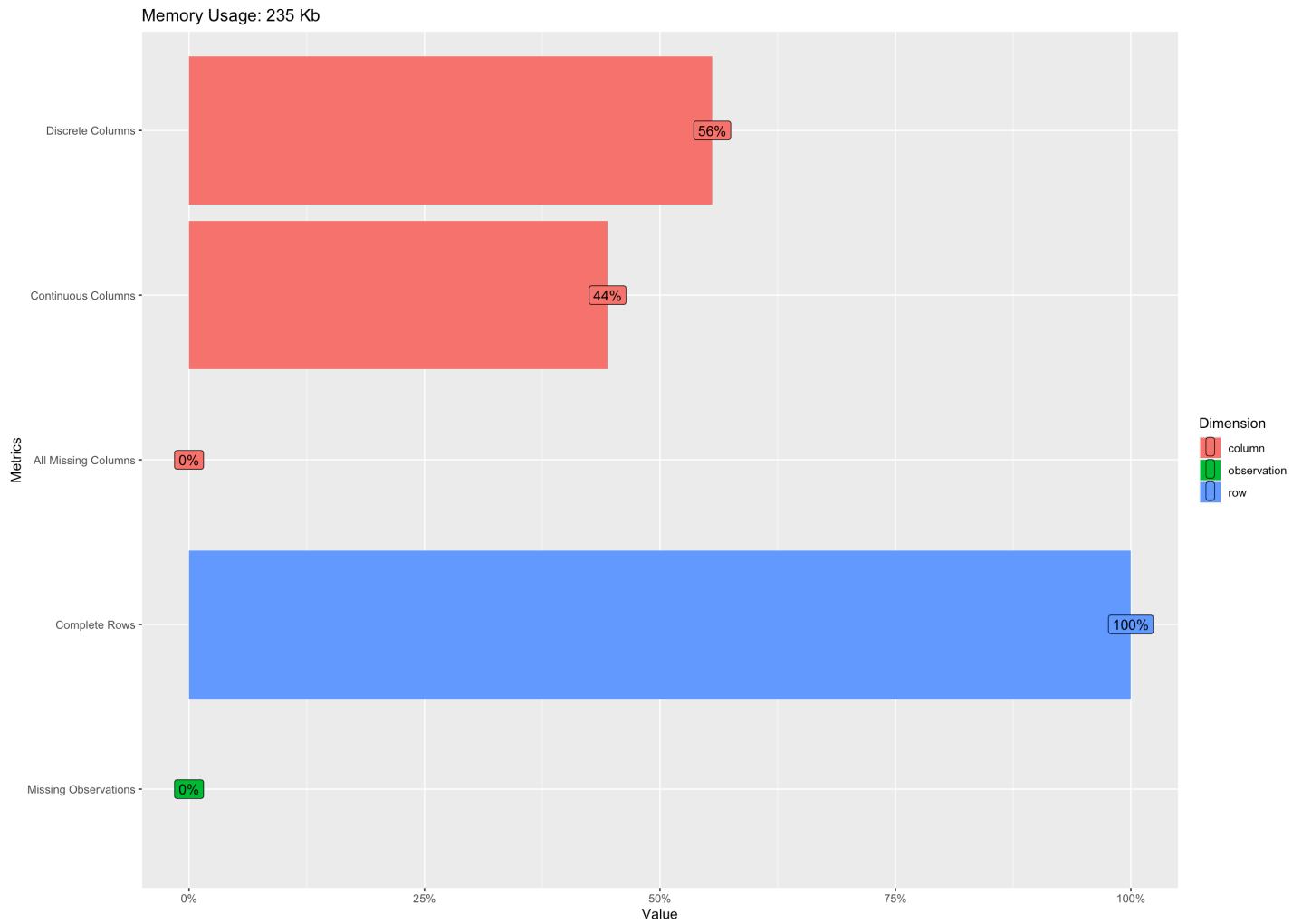
- Basic Statistics
  - Raw Counts
  - Percentages
- Data Structure
- Missing Data Profile
- Univariate Distribution
  - Histogram
  - Bar Chart (with frequency)
  - QQ Plot
- Correlation Analysis
- Principal Component Analysis

## Basic Statistics

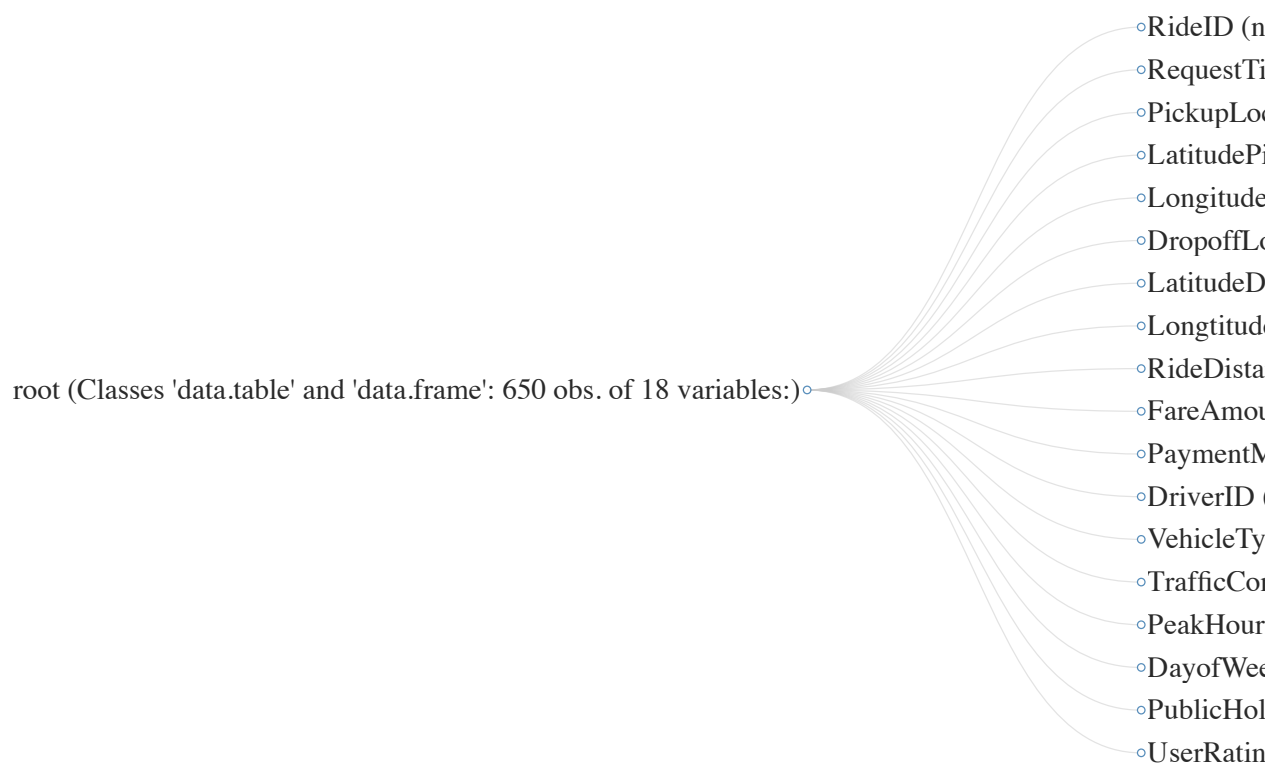
### Raw Counts

Name	Value
Rows	650
Columns	18
Discrete columns	10
Continuous columns	8
All missing columns	0
Missing observations	0
Complete Rows	650
Total observations	11,700
Memory allocation	235 Kb

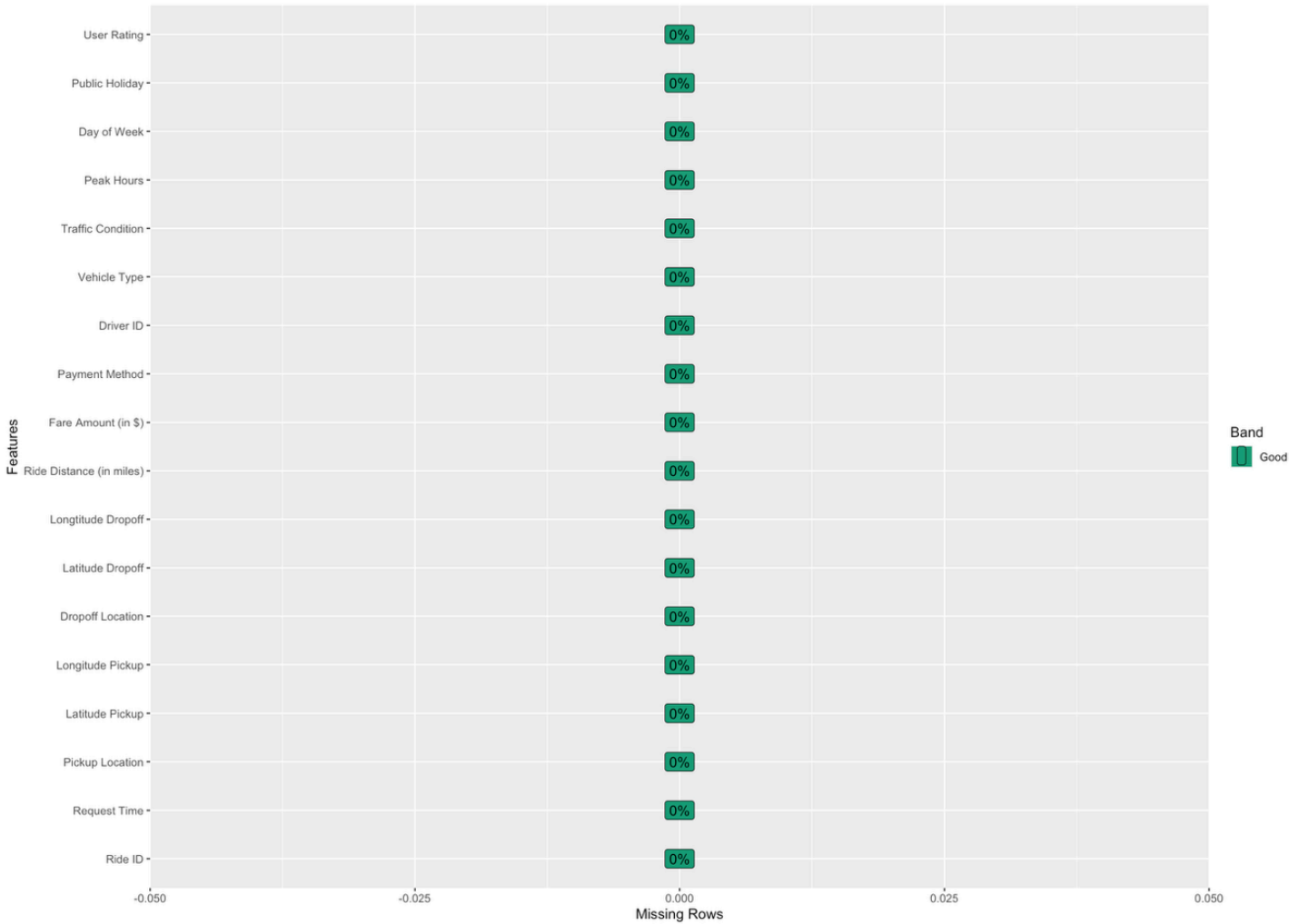
### Percentages



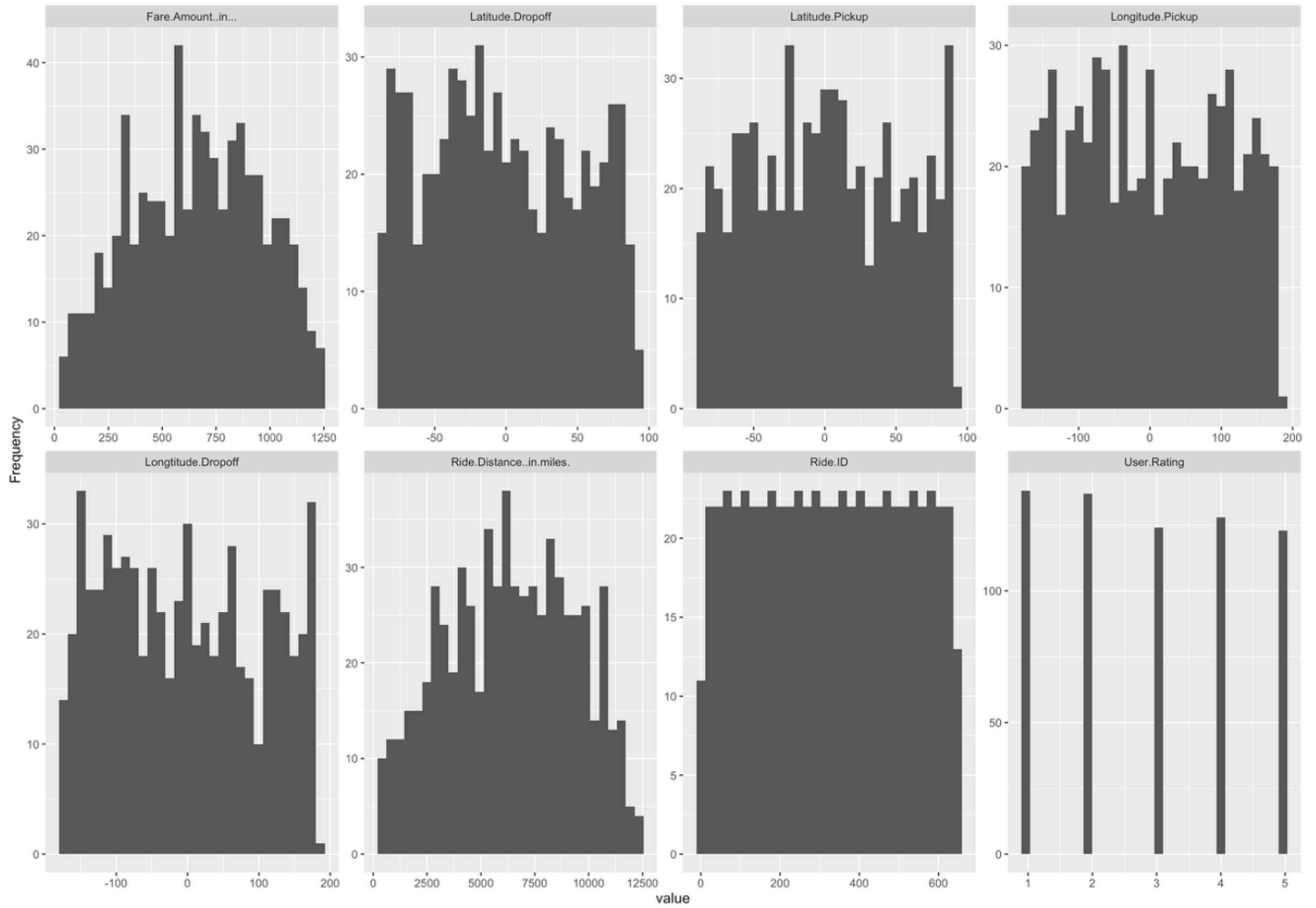
## Data Structure



# Missing Data Profile

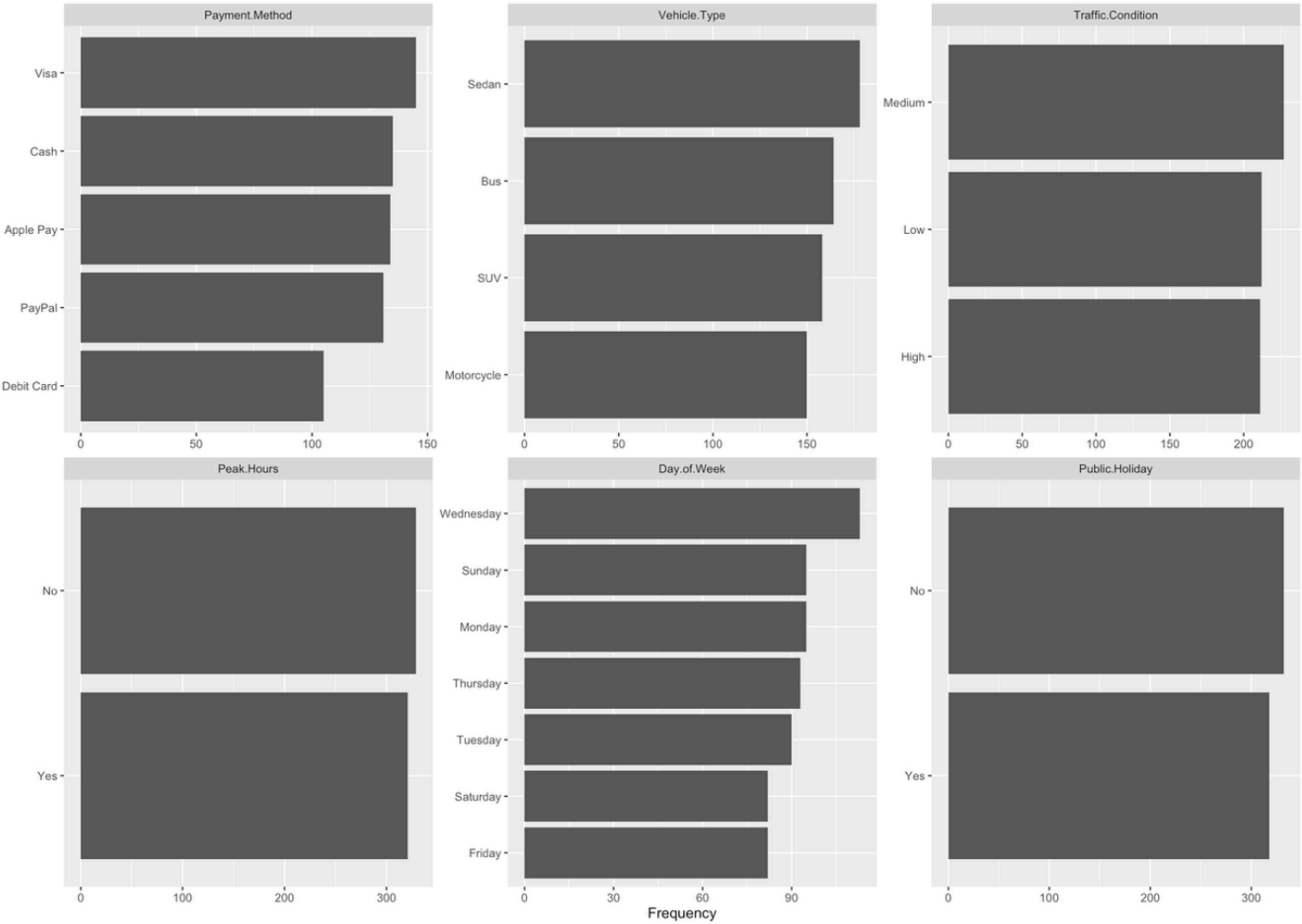


## Univariate Distribution Histogram

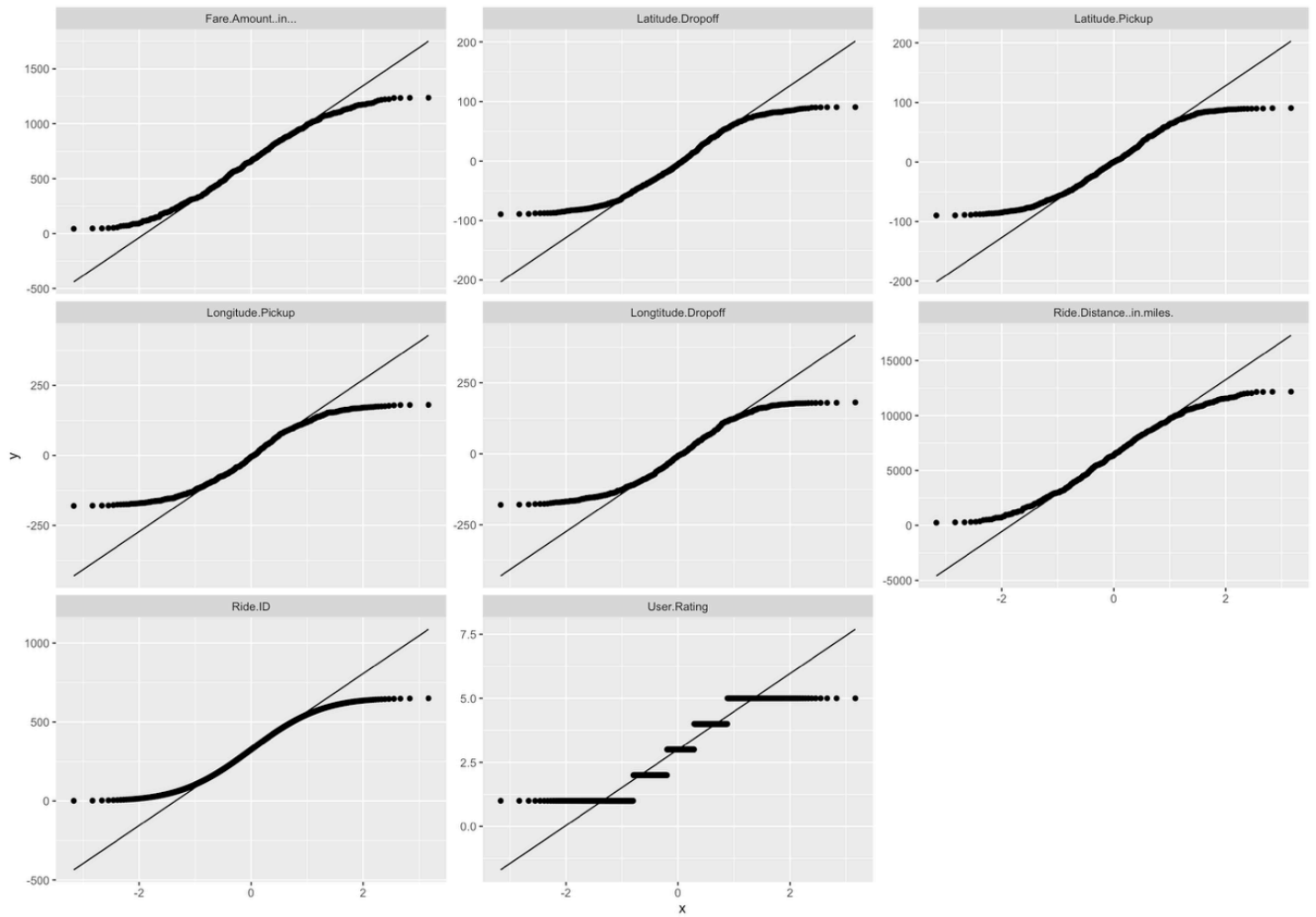


### Bar Chart (with frequency)

```
## 4 columns ignored with more than 50 categories.  
## Request.Time: 650 categories  
## Pickup.Location: 650 categories  
## Dropoff.Location: 650 categories  
## Driver.ID: 650 categories
```

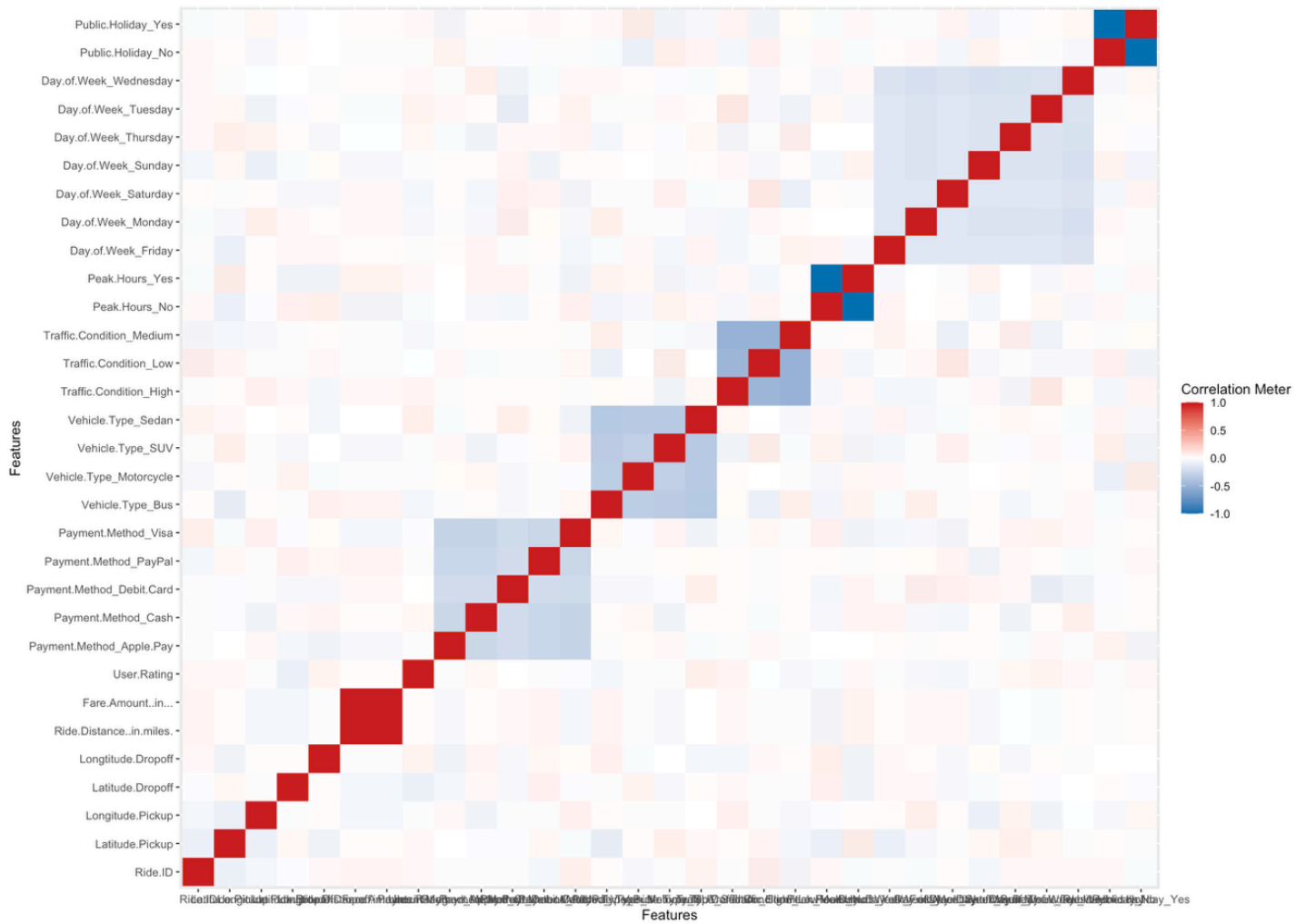


QQ Plot



## Correlation Analysis

```
## 4 features with more than 20 categories ignored!  
## Request.Time: 650 categories  
## Pickup.Location: 650 categories  
## Dropoff.Location: 650 categories  
## Driver.ID: 650 categories
```



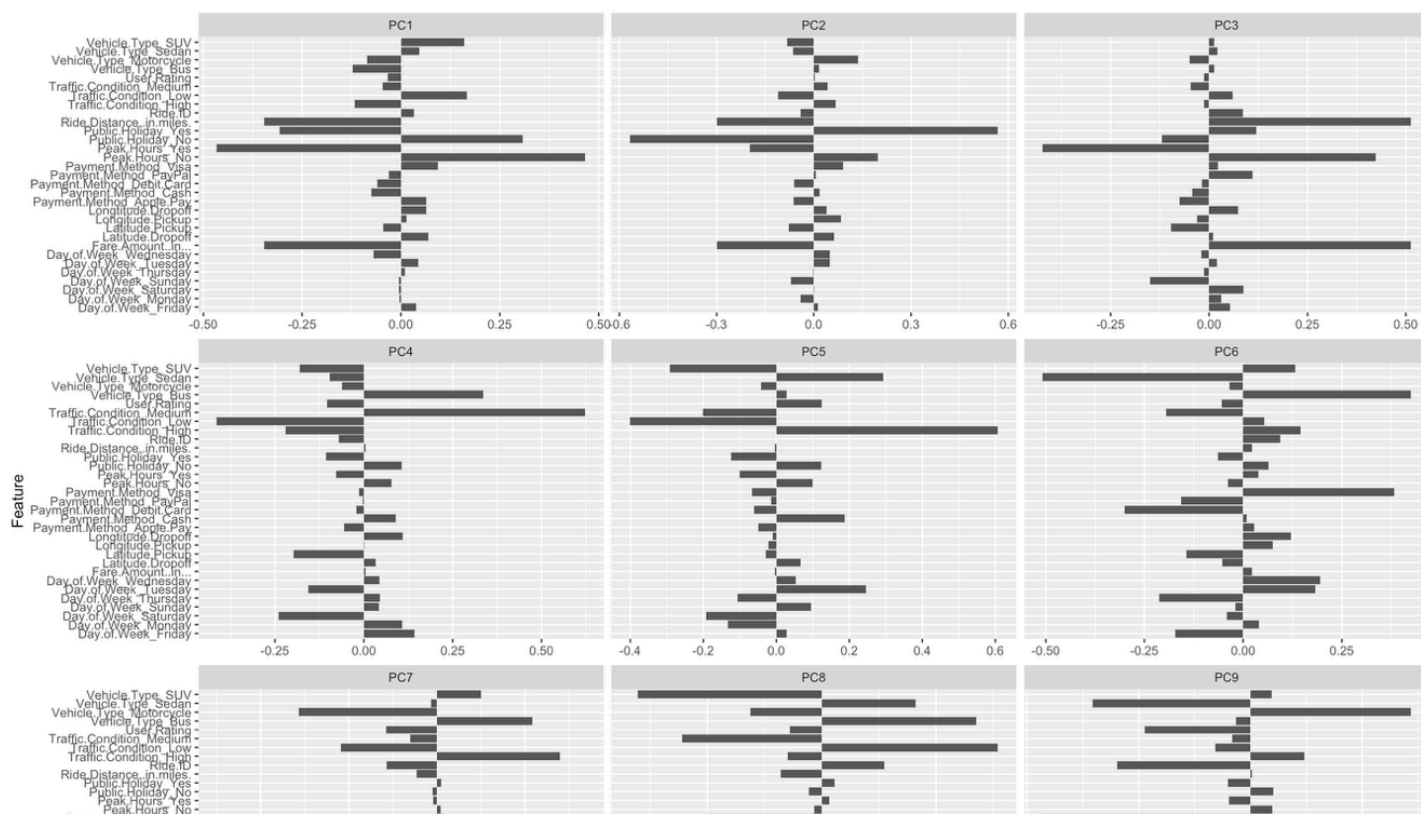
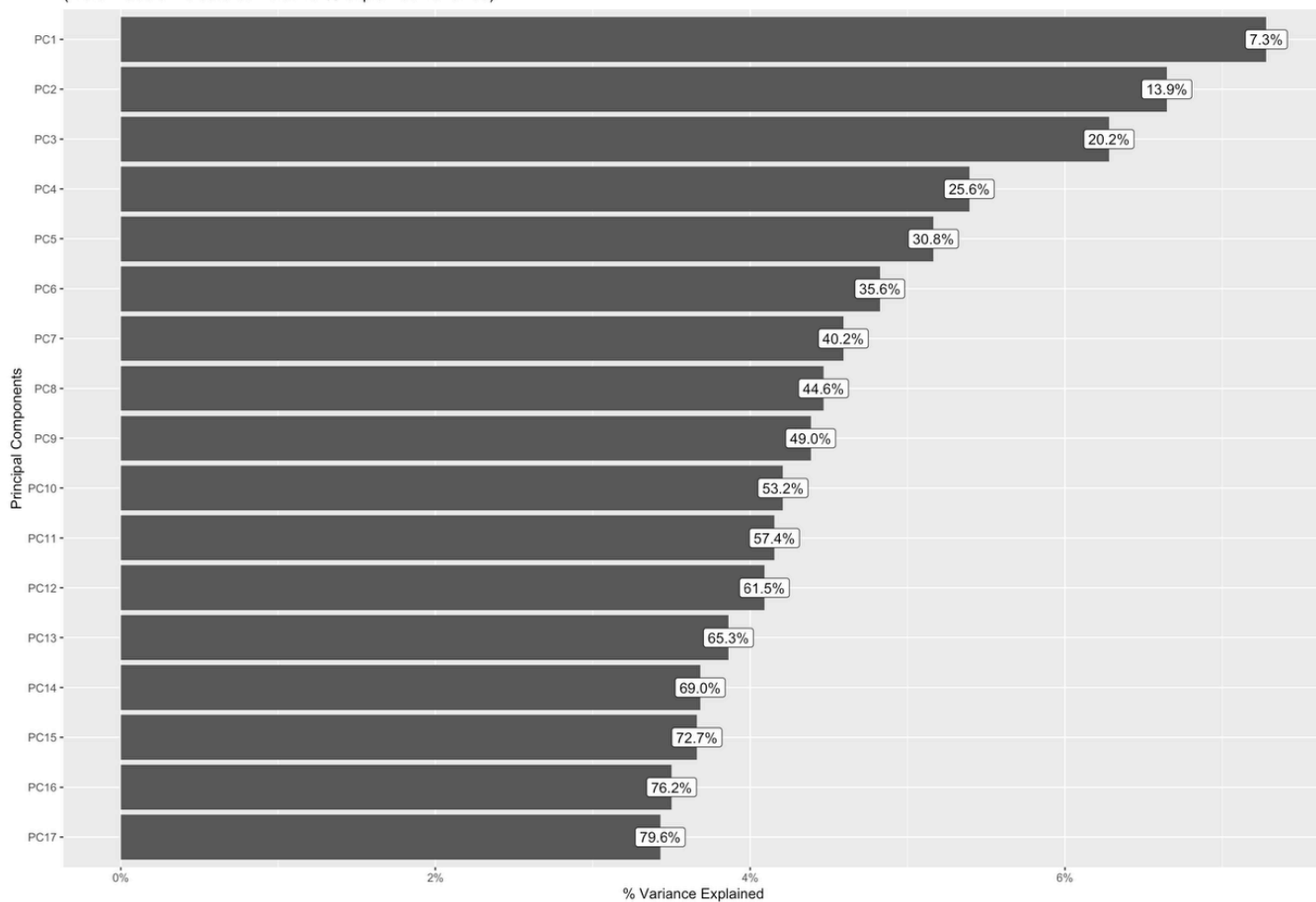
## Principal Component Analysis

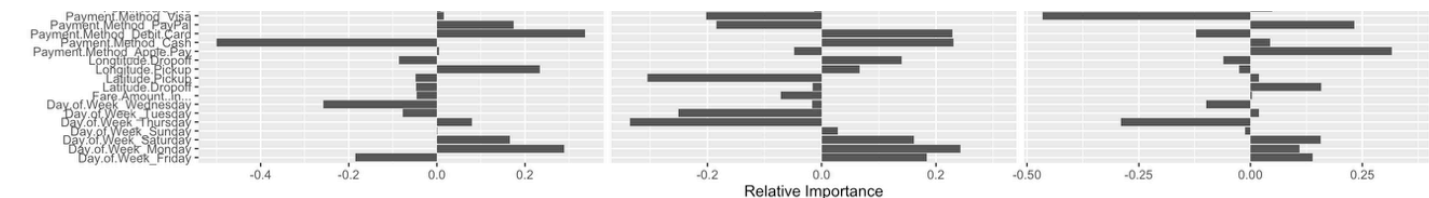
```
## 4 features with more than 50 categories ignored!  
## Request.Time: 650 categories  
## Pickup.Location: 650 categories  
## Dropoff.Location: 650 categories  
## Driver.ID: 650 categories
```



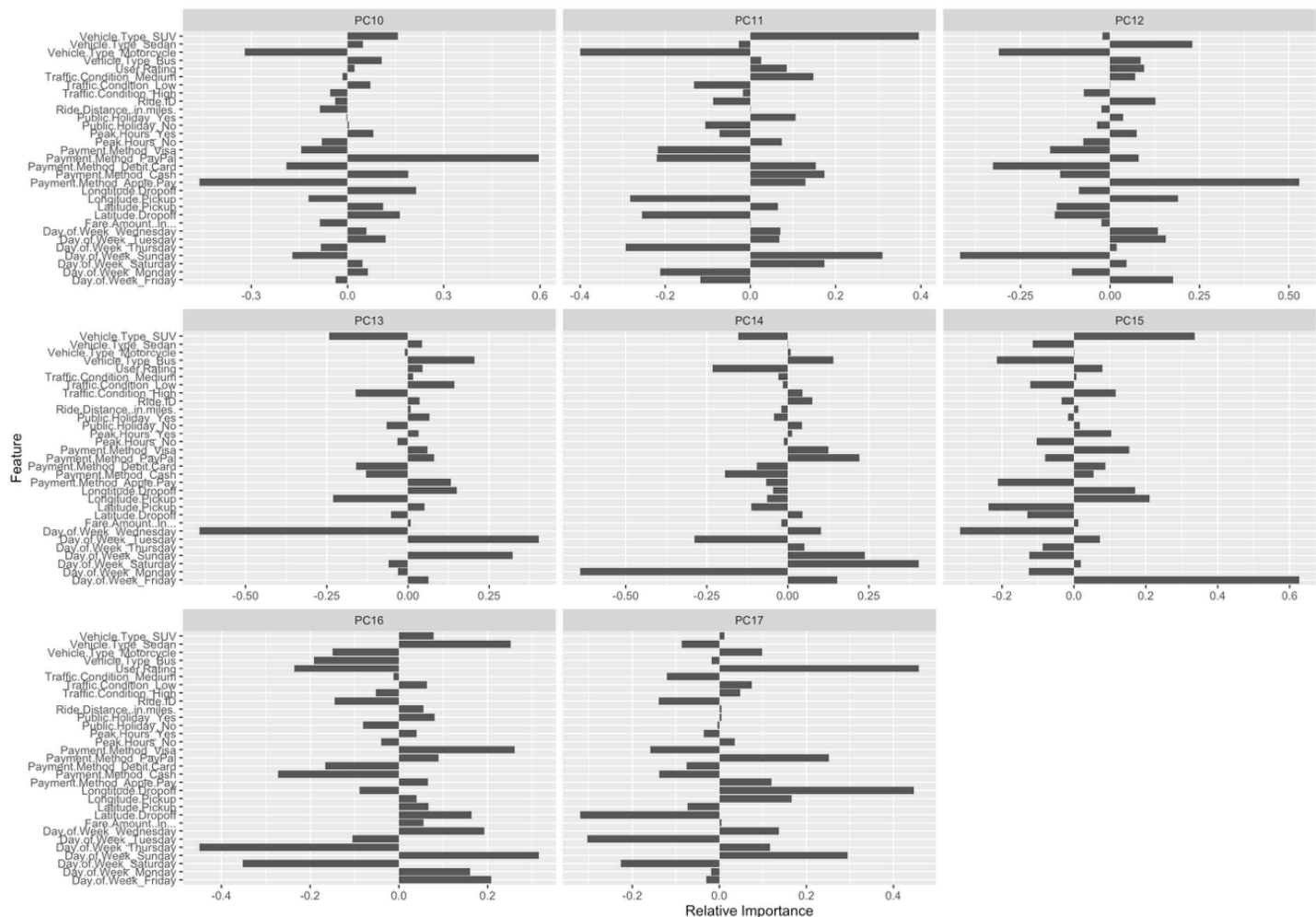


% Variance Explained By Principal Components  
(Note: Labels indicate cumulative % explained variance)





Page 1



Page 2