# Data Profiling Report

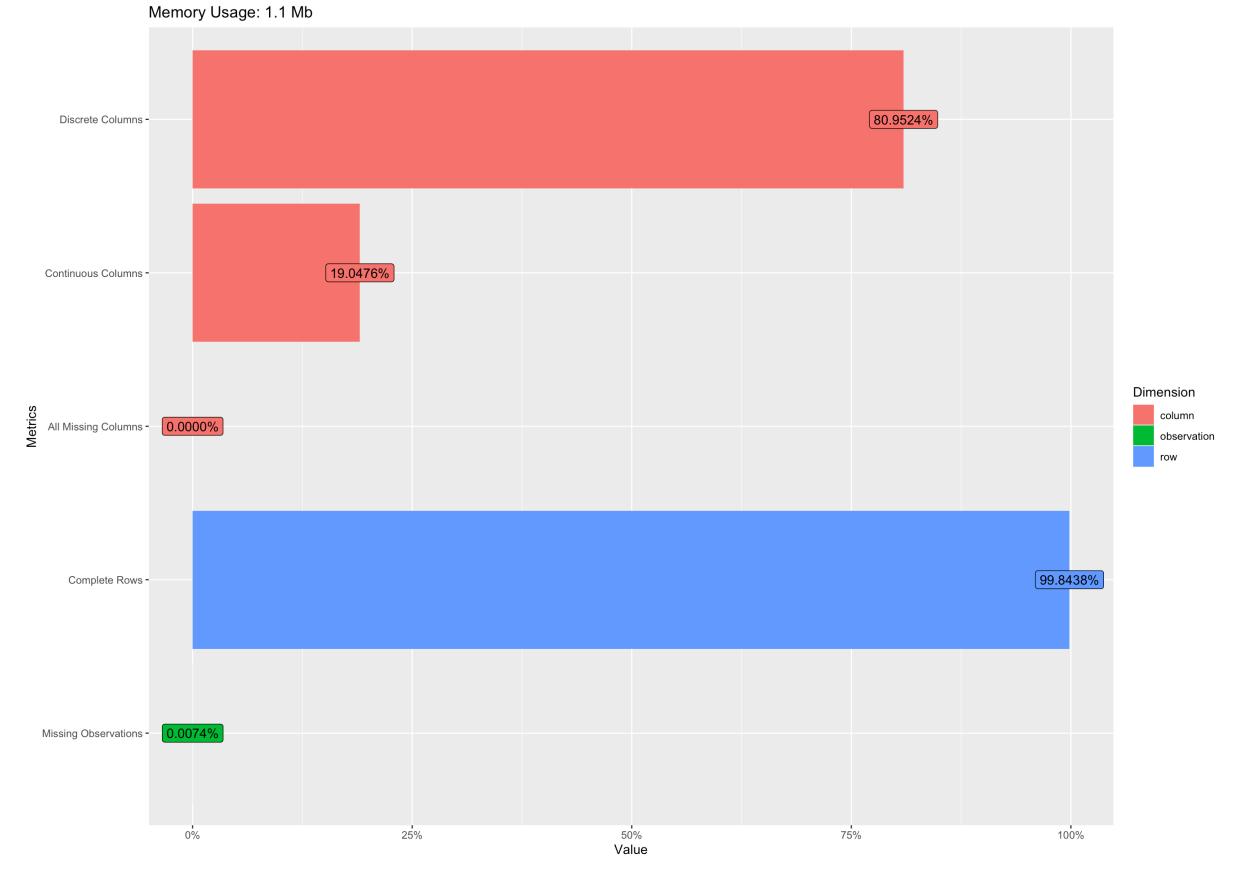
- Basic Statistics
  - Raw Counts
  - Percentages
- Data Structure
- Missing Data Profile
- Univariate Distribution
  - Histogram
  - Bar Chart (by frequency)
  - QQ Plot
- Correlation Analysis
- Principal Component Analysis

#### **Basic Statistics**

#### Raw Counts

| Name                 | Value   |
|----------------------|---------|
| Rows                 | 7,043   |
| Columns              | 21      |
| Discrete columns     | 17      |
| Continuous columns   | 4       |
| All missing columns  | 0       |
| Missing observations | 11      |
| Complete Rows        | 7,032   |
| Total observations   | 147,903 |
| Memory allocation    | 1.1 Mb  |

#### Percentages



Data Structure

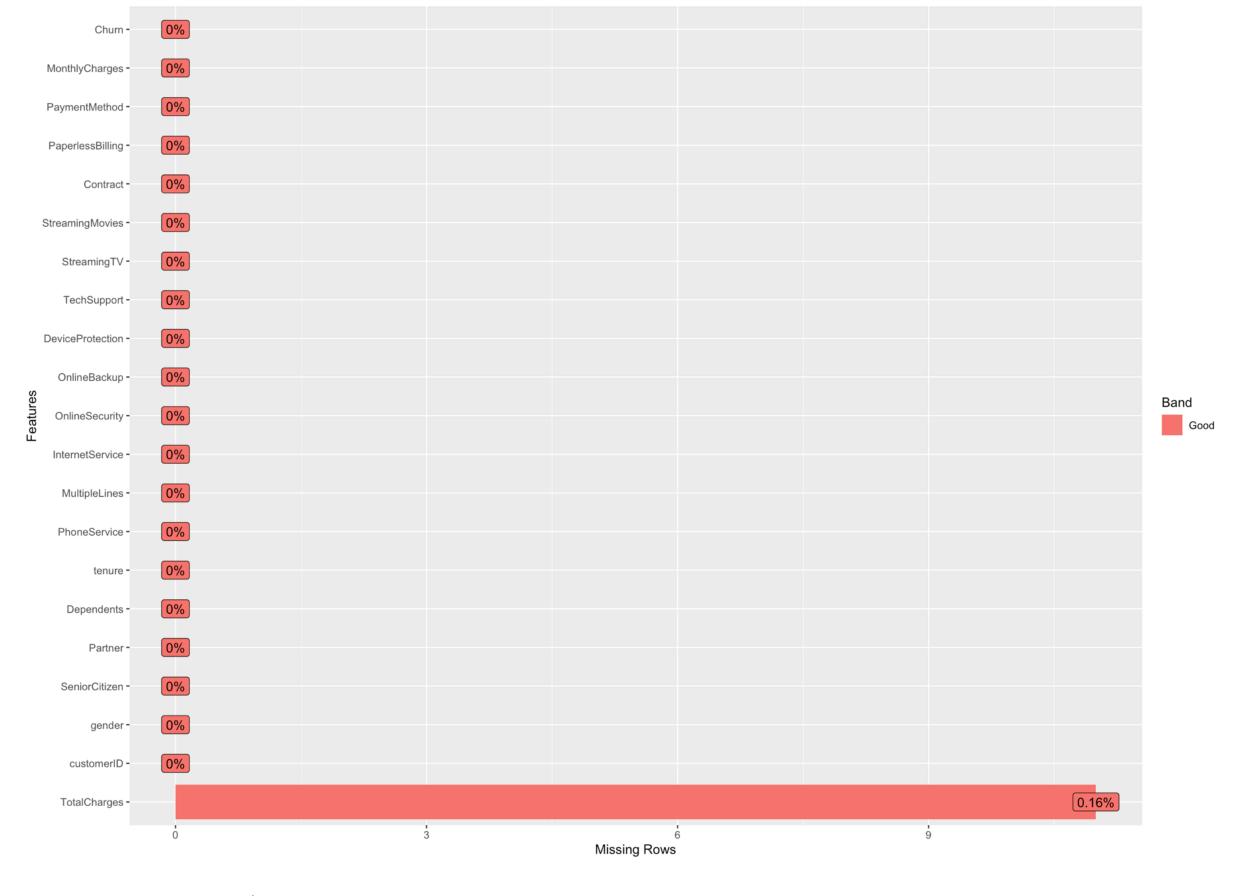
```
root (Classes 'data.table' and 'data.frame': 7043 obs. of 21 variables:) o-
```

```
ocustomerID (Factor w/ 7043 levels "0002-ORFBO", "0003-MKNFE",)
• gender (Factor w/ 2 levels "Female", "Male")
SeniorCitizen (int)
•Partner (Factor w/ 2 levels "No", "Yes")
• Dependents (Factor w/ 2 levels "No", "Yes")
otenure (int)
• PhoneService (Factor w/ 2 levels "No", "Yes")
• MultipleLines (Factor w/ 3 levels "No", "No phone service",)
oInternetService (Factor w/ 3 levels "DSL", "Fiber optic",)
OnlineSecurity (Factor w/ 3 levels "No", "No internet service",)
OnlineBackup (Factor w/ 3 levels "No", "No internet service",)
• DeviceProtection (Factor w/ 3 levels "No", "No internet service",)
• TechSupport (Factor w/ 3 levels "No", "No internet service",)
StreamingTV (Factor w/ 3 levels "No", "No internet service",)
•StreamingMovies (Factor w/ 3 levels "No","No internet service",)
°Contract (Factor w/ 3 levels "Month-to-month",)
•PaperlessBilling (Factor w/ 2 levels "No", "Yes")
• PaymentMethod (Factor w/ 4 levels "Bank transfer (automatic)",)
MonthlyCharges (num)
```

TotalCharges (num)

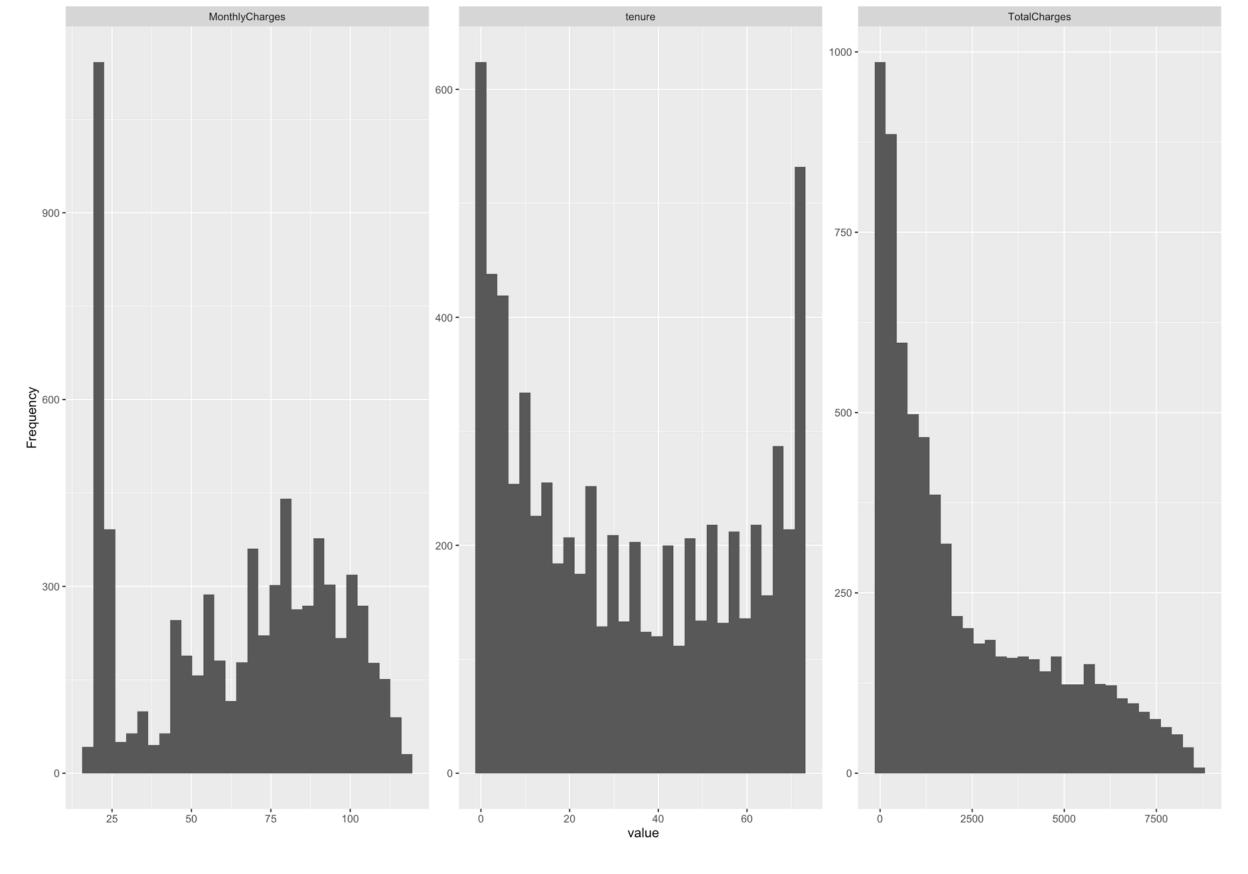
∘Churn (Factor w/ 2 levels "No","Yes")

Missing Data Profile



# Univariate Distribution

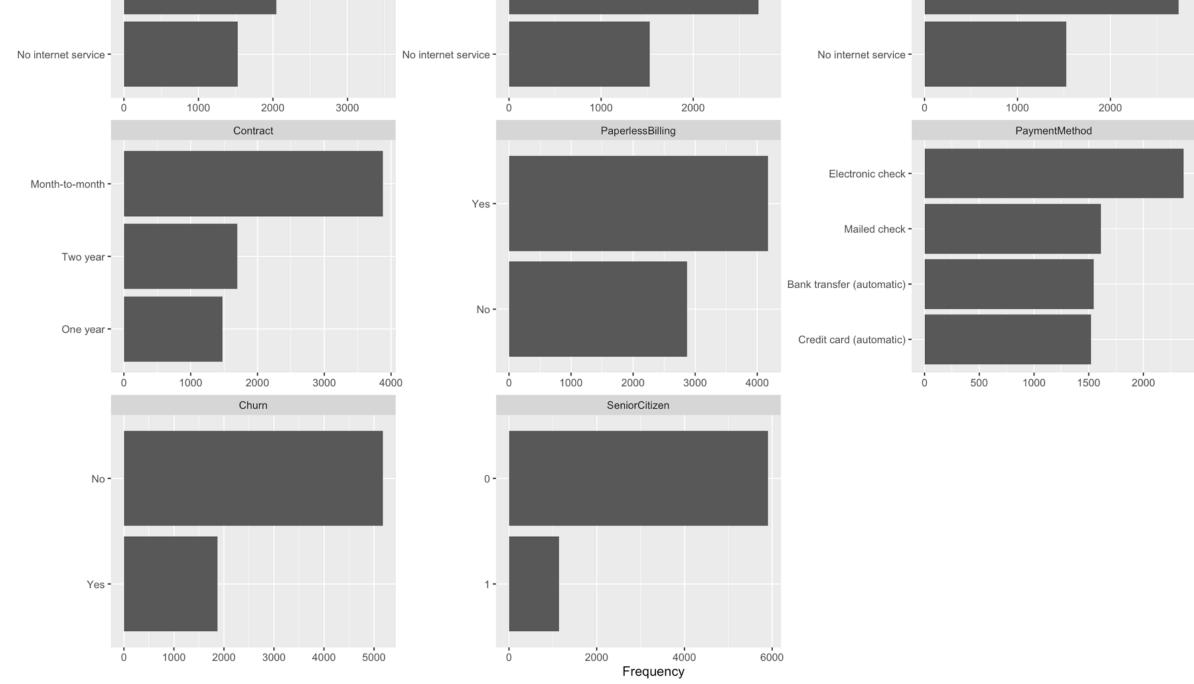
Histogram



#### Bar Chart (by frequency)

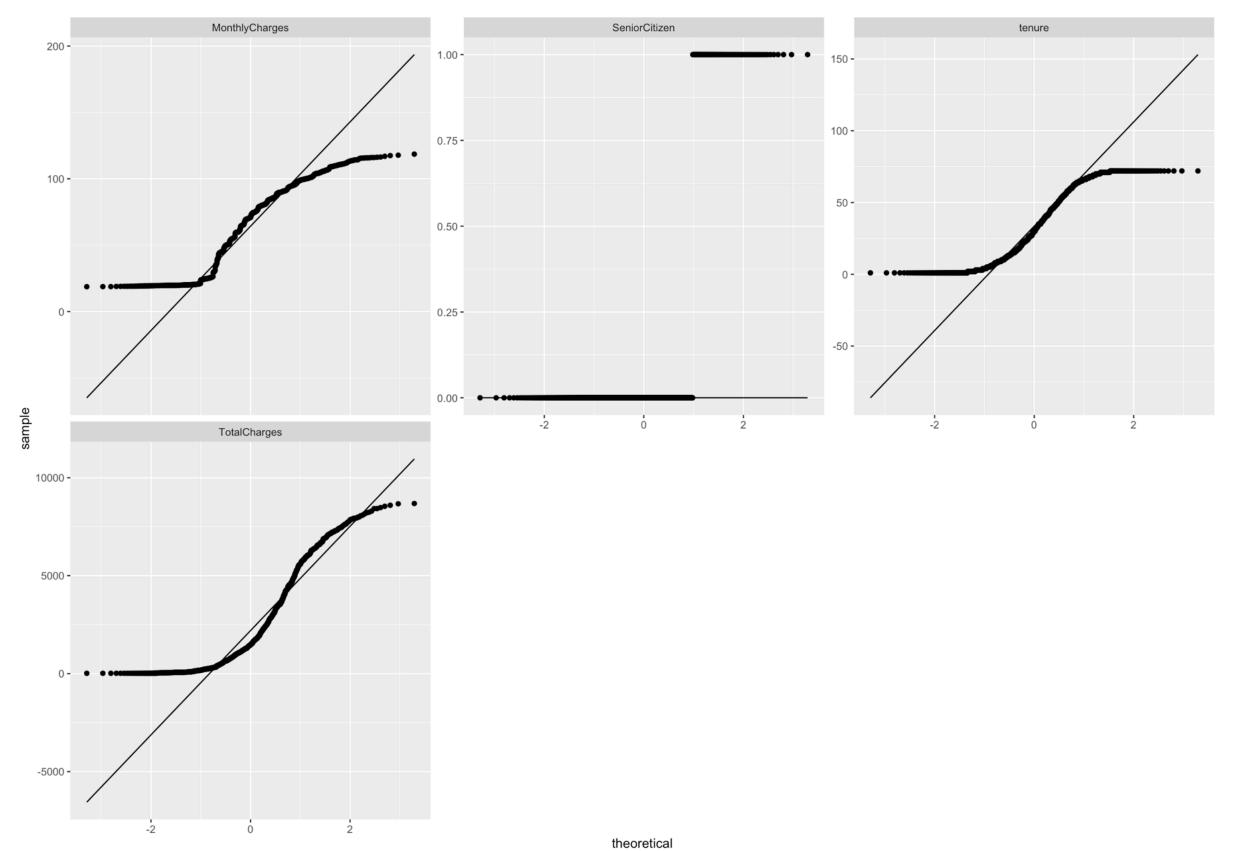
## 1 columns ignored with more than 50 categories.
## customerID: 7043 categories





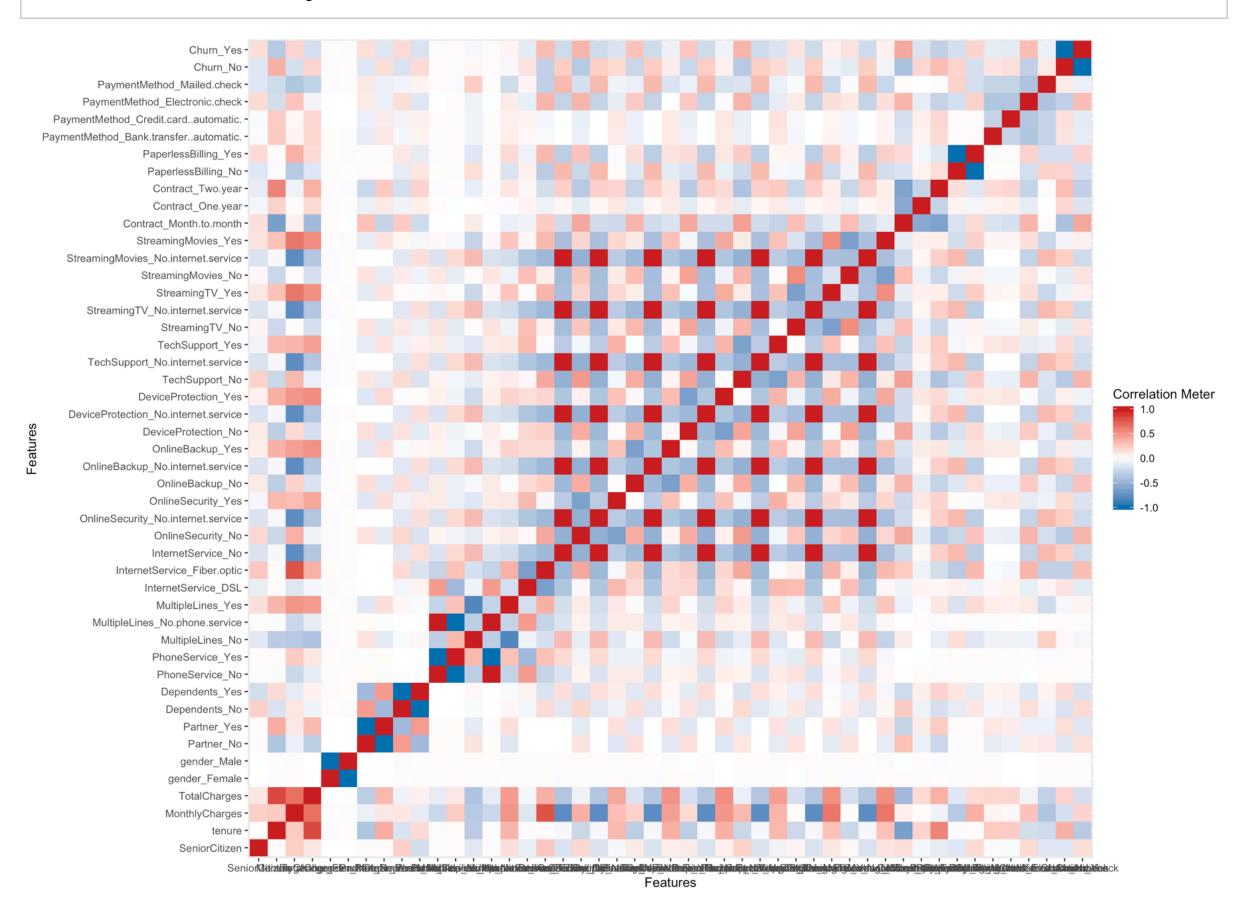
#### Page 2

# QQ Plot



# Correlation Analysis

## 1 features with more than 20 categories ignored!
## customerID: 7032 categories



# Principal Component Analysis

## 1 features with more than 50 categories ignored!
## customerID: 7032 categories

