Data Profiling Report

Basic Statistics

Raw Counts

Name	Value
Rows	7,582
Columns	14
Discrete columns	8
Continuous columns	6
All missing columns	0
Missing observations	158
Complete Rows	7,424
Total observations	106,148
Memory allocation	834.3 Kb

Percentages

Memory Usage: 834.3 Kb

Data Structure

root (Classes 'data.table' and 'data.frame': 7582 obs. of 14 variables:)

- X (num)
- age (num)
- bmi (num)
- children (num)
- smoker (chr)
- location (chr)
- location_type (chr)
- education_level (chr)
- yearly_physical (chr)
- exercise (chr)
- married (chr)
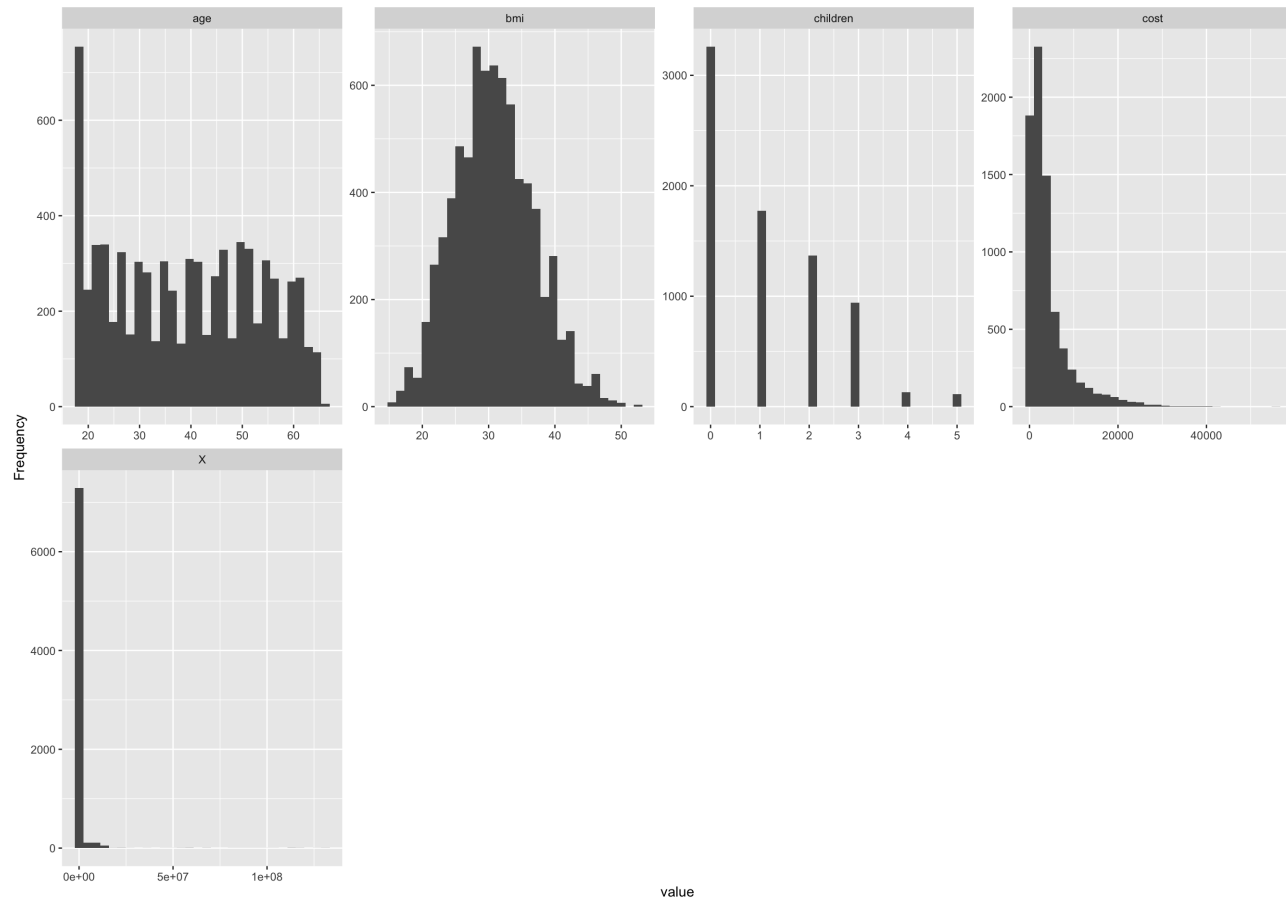- hypertension (num)
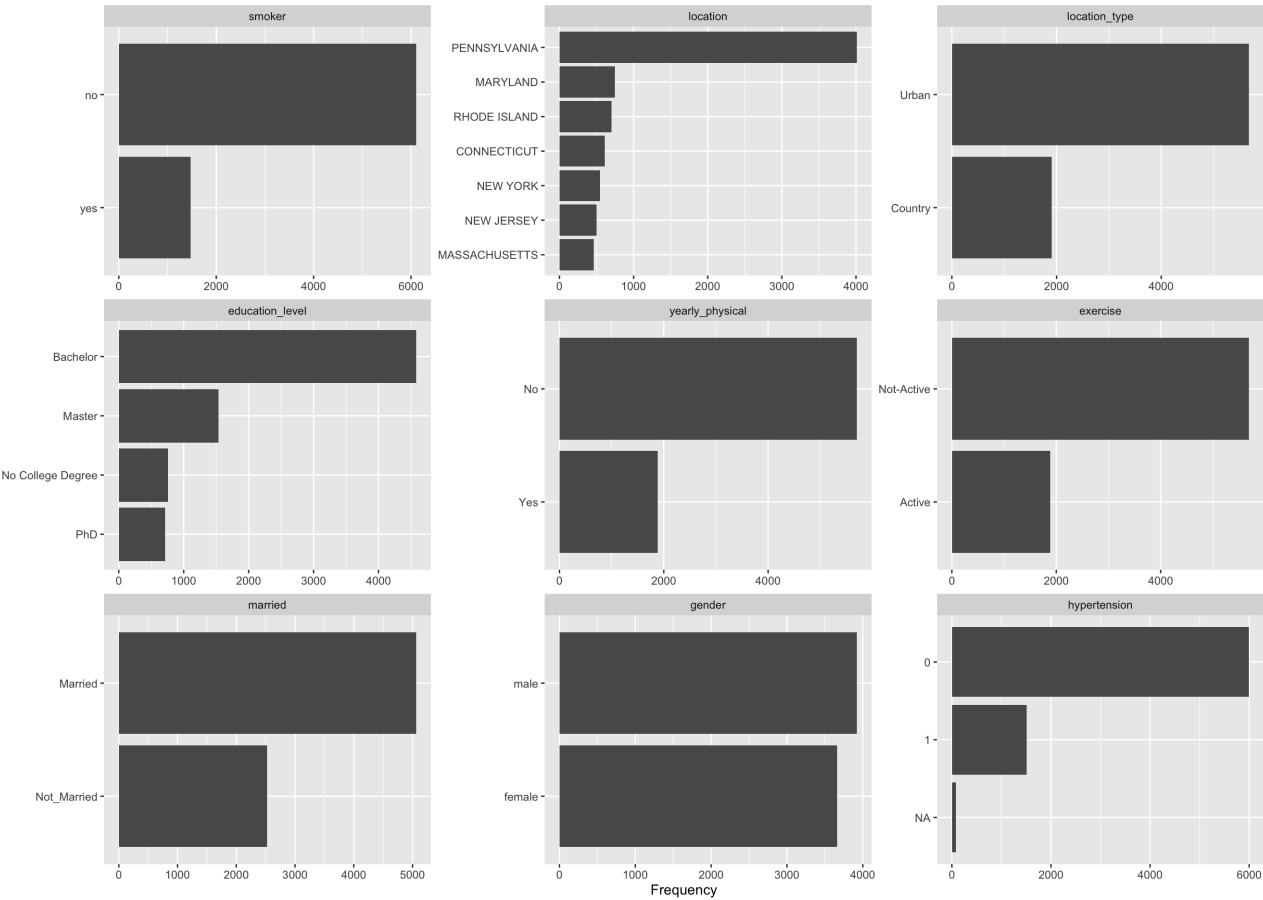- gender (chr)
- cost (num)
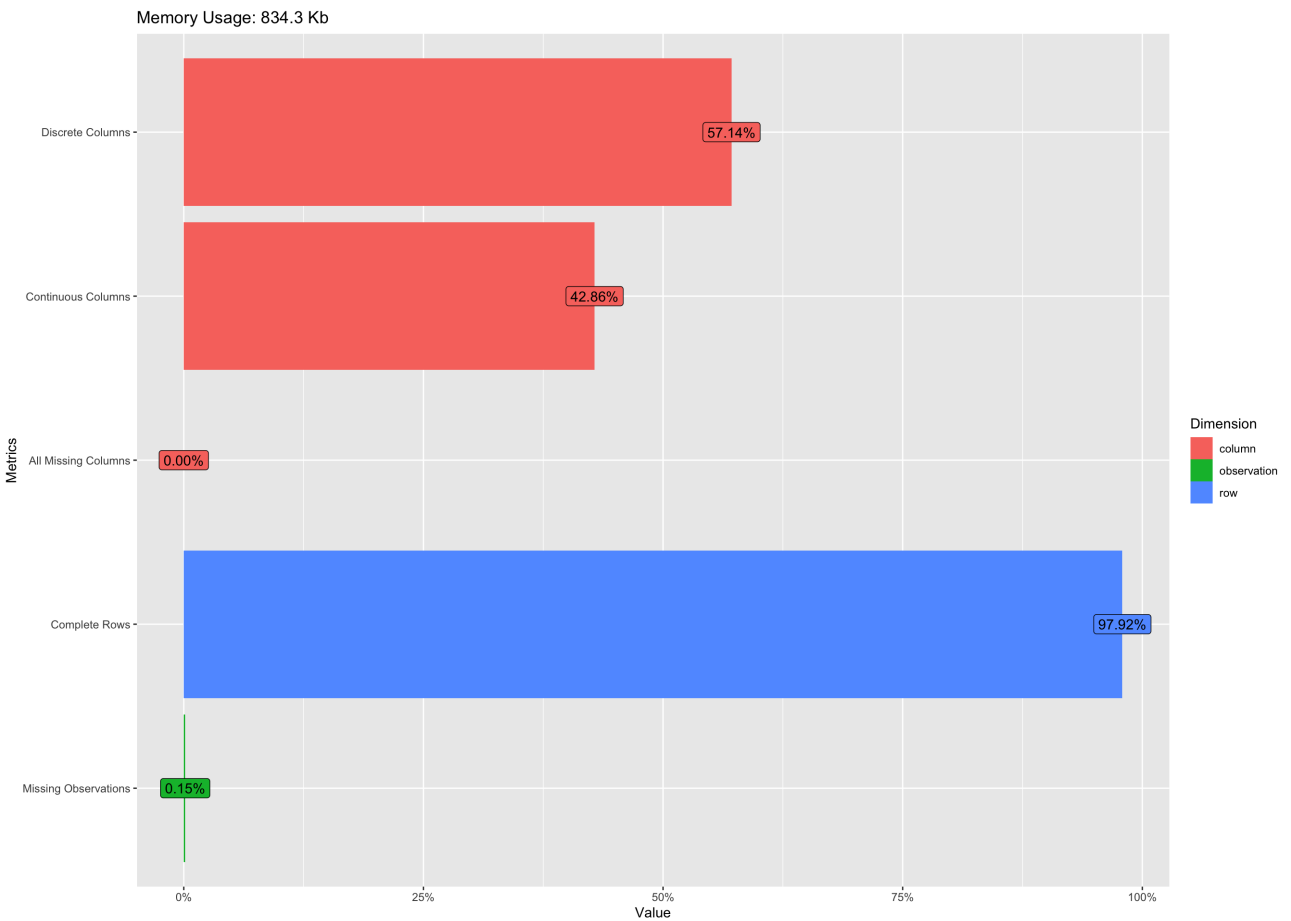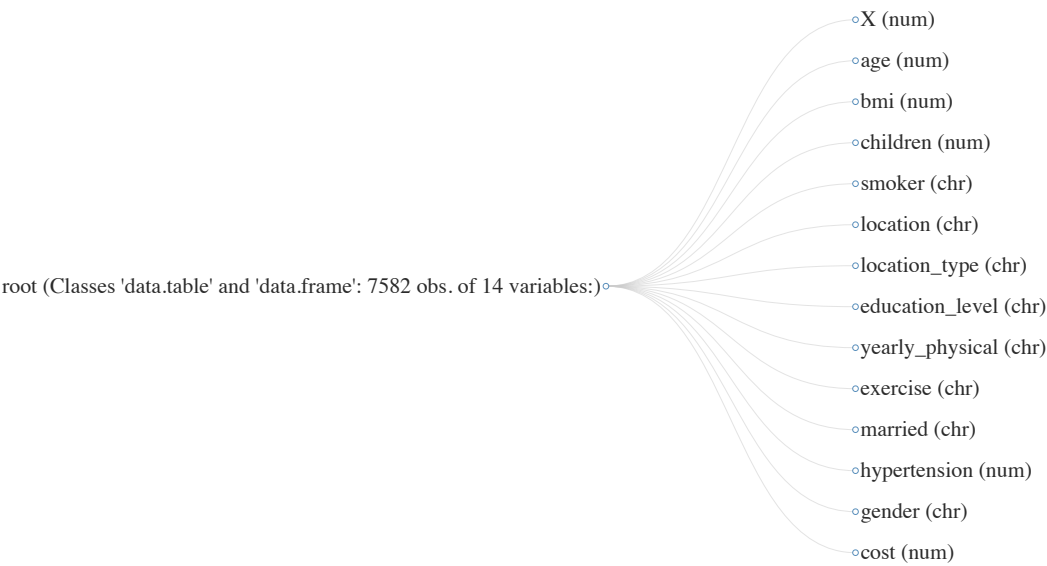
Missing Data Profile
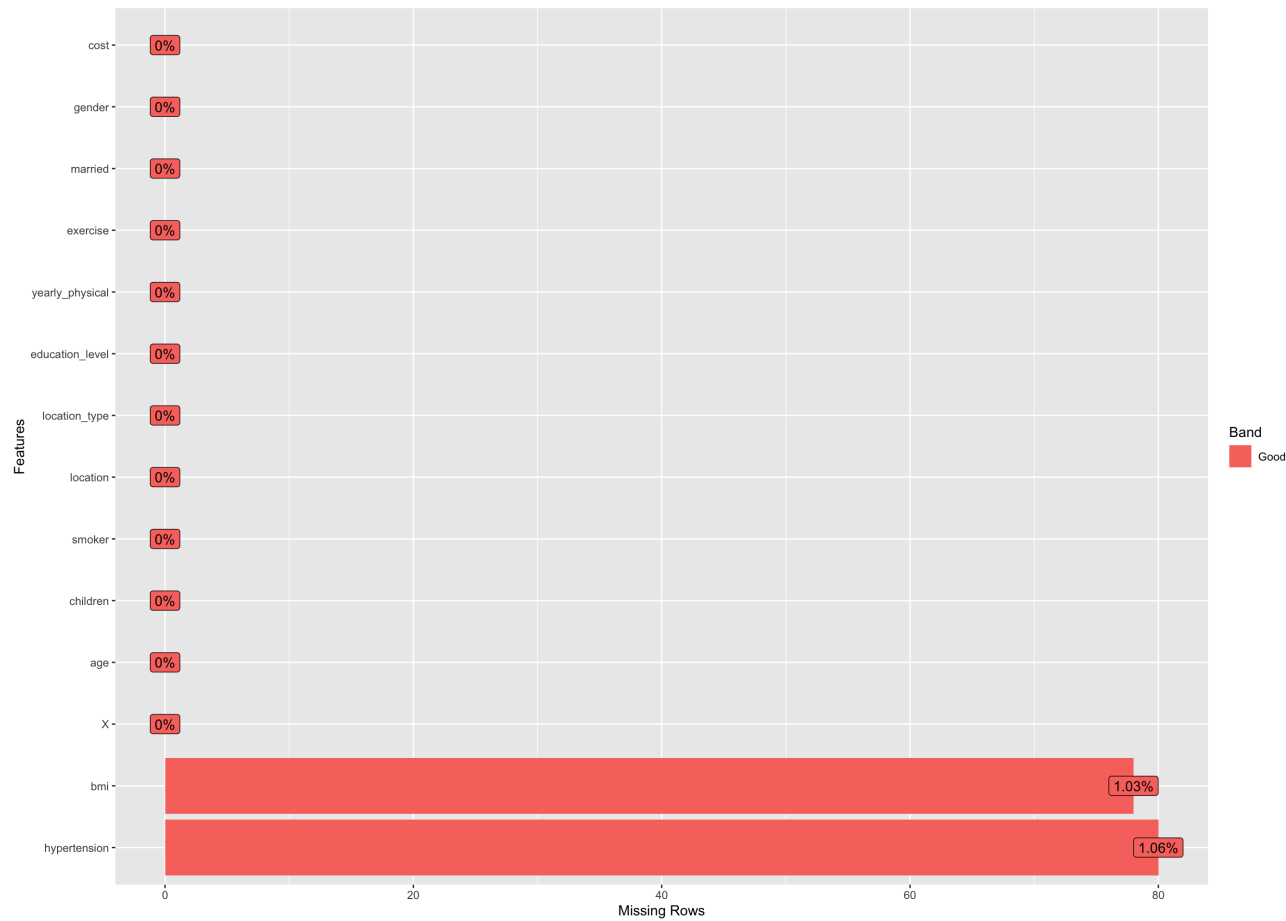
Univariate Distribution
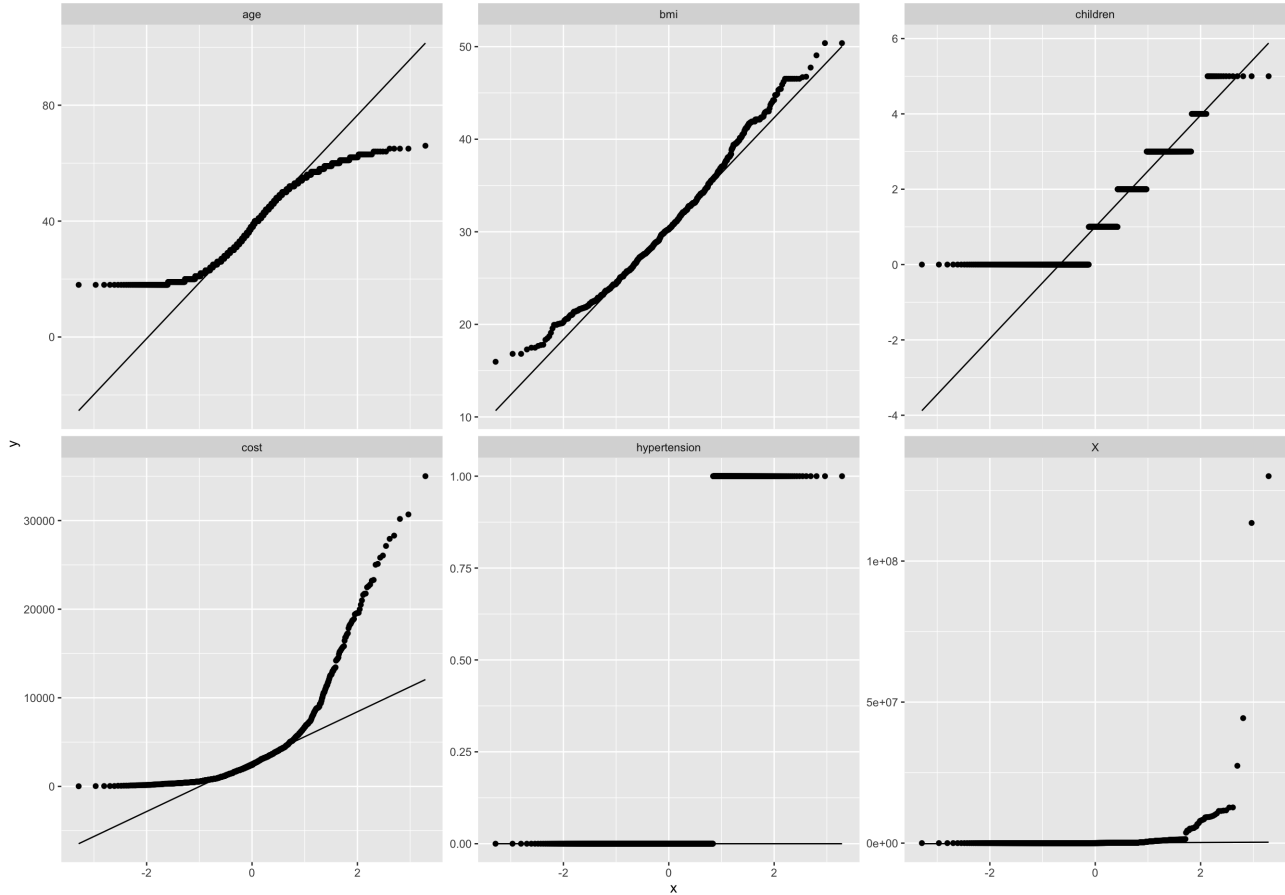
Histogram
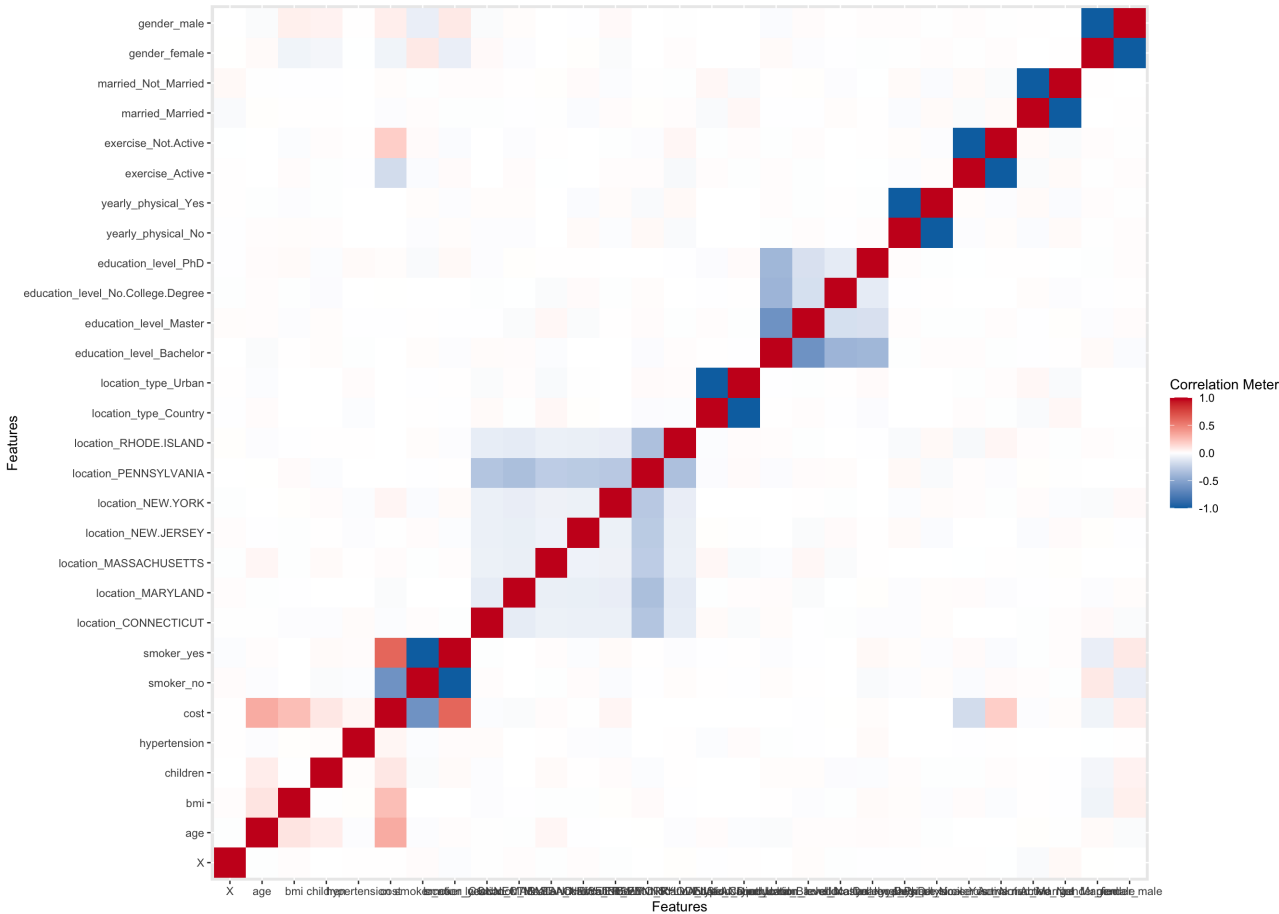

Bar Chart (with frequency)

QQ Plot

```
## Warning: Removed 21 rows containing non-finite values (stat_qq).
```

```
## Warning: Removed 21 rows containing non-finite values (stat_qq_line).
```

Correlation Analysis

Principal Component Analysis

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

Page 1

PC13 PC14

Relative Importance