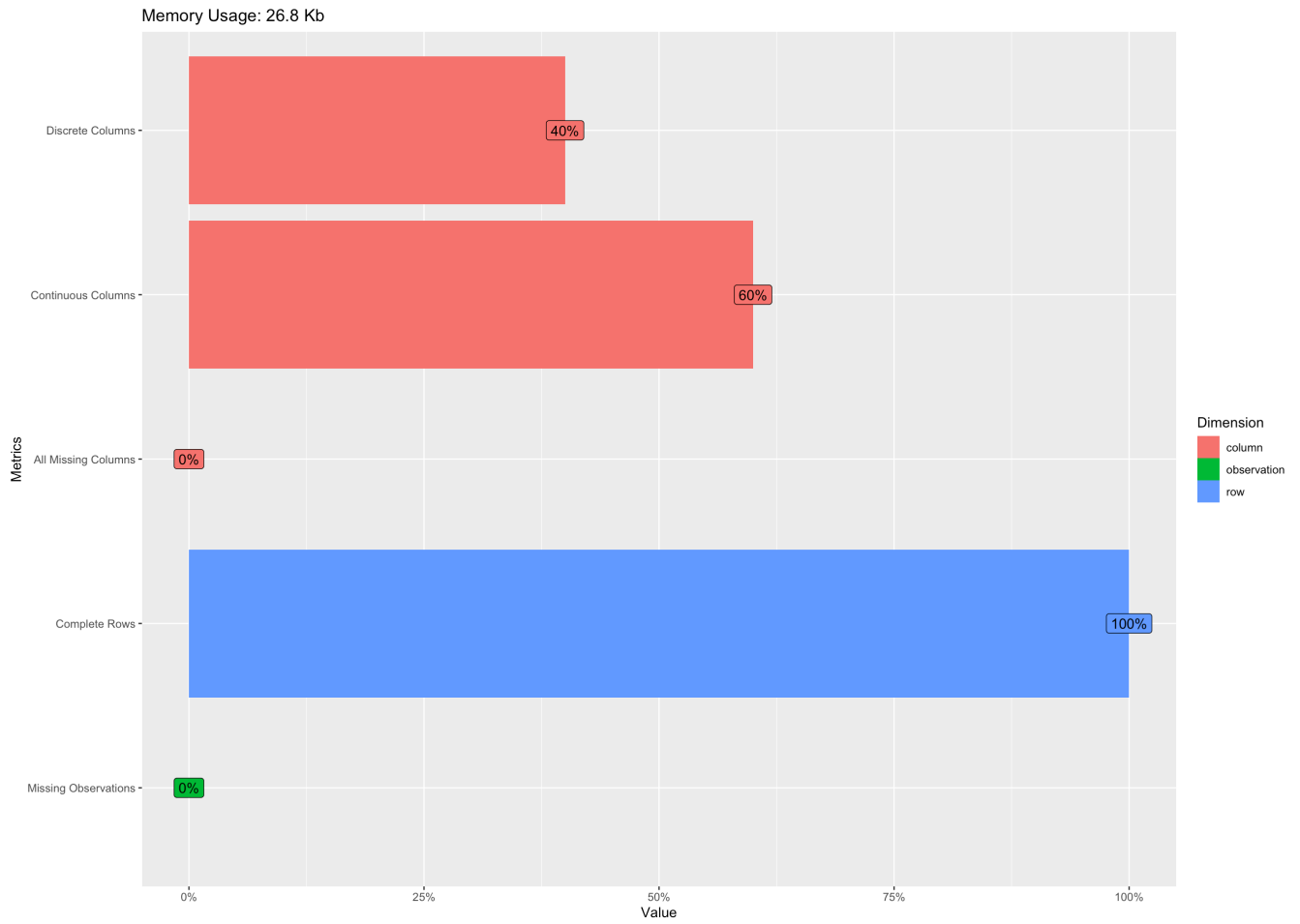# Appendix 7: Data Profiling Report

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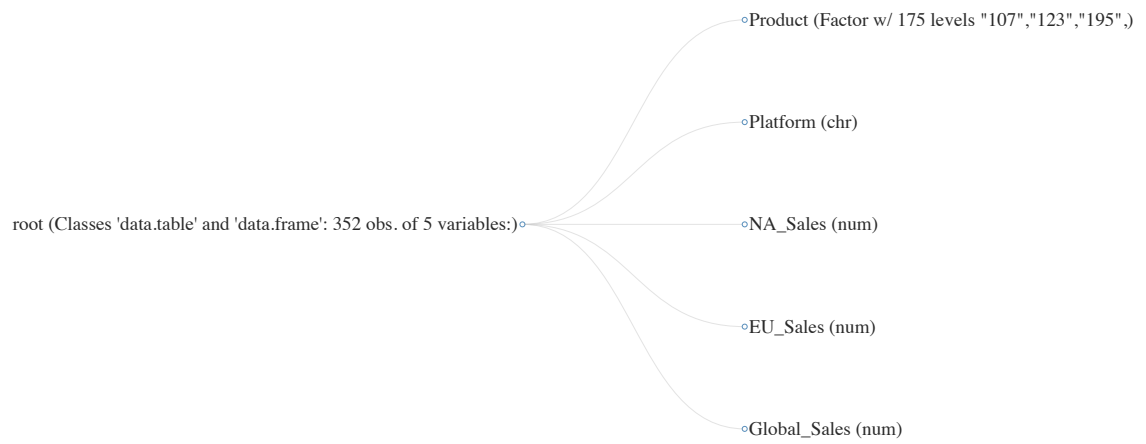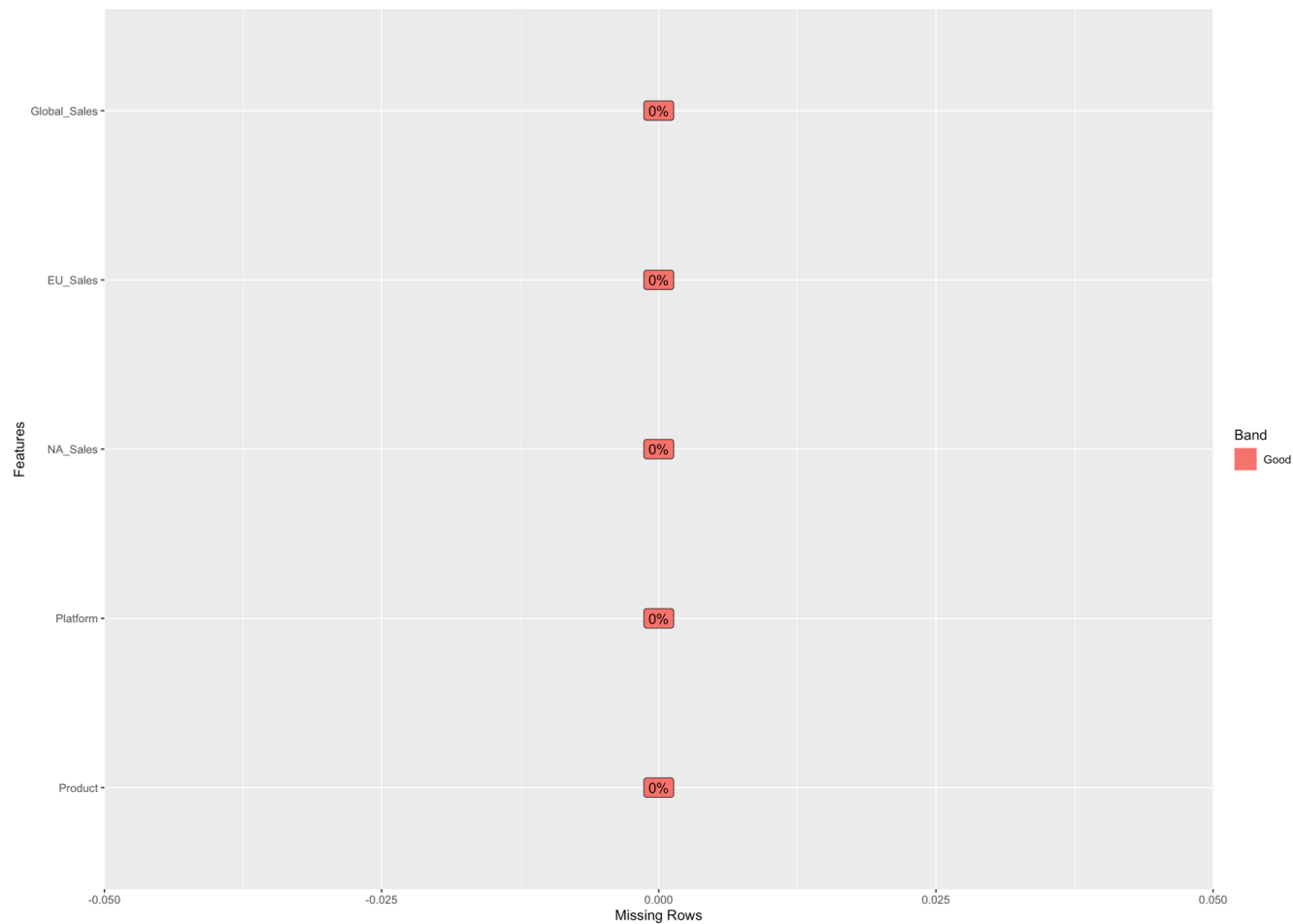- Principal Component Analysis

Basic Statistics

Raw Counts

Name	Value
Rows	352
Columns	5
Discrete columns	2
Continuous columns	3
All missing columns	0
Missing observations	0
Complete Rows	352
Total observations	1,760
Memory allocation	26.8 Kb

Percentages

Memory Usage: 26.8 Kb

Data Structure

◦Product (Factor w/ 175 levels "107","123","195",)

◦Platform (chr)

root (Classes 'data.table' and 'data.frame': 352 obs. of 5 variables:)◦
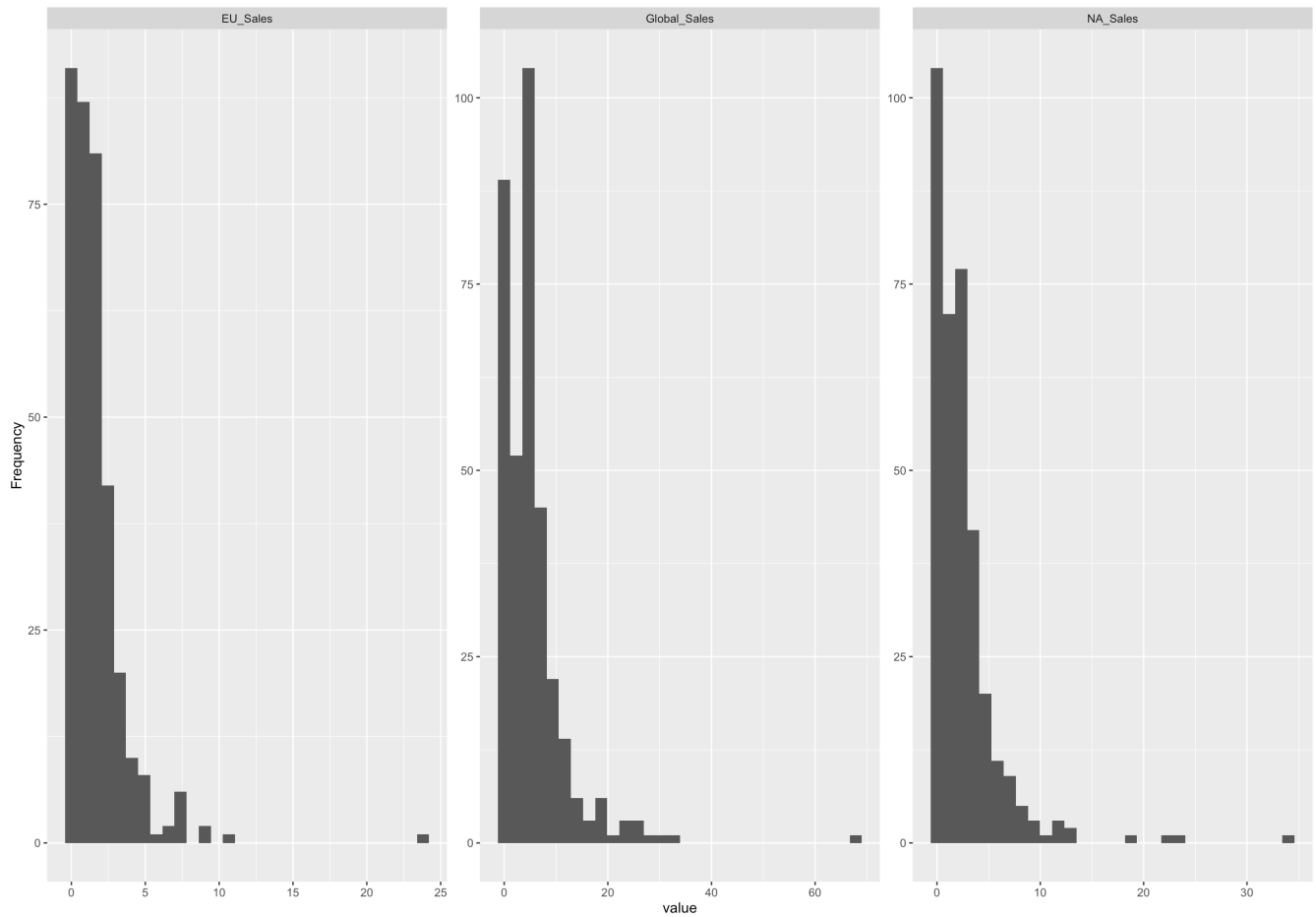
◦NA_Sales (num)

◦EU_Sales (num)

◦Global_Sales (num)
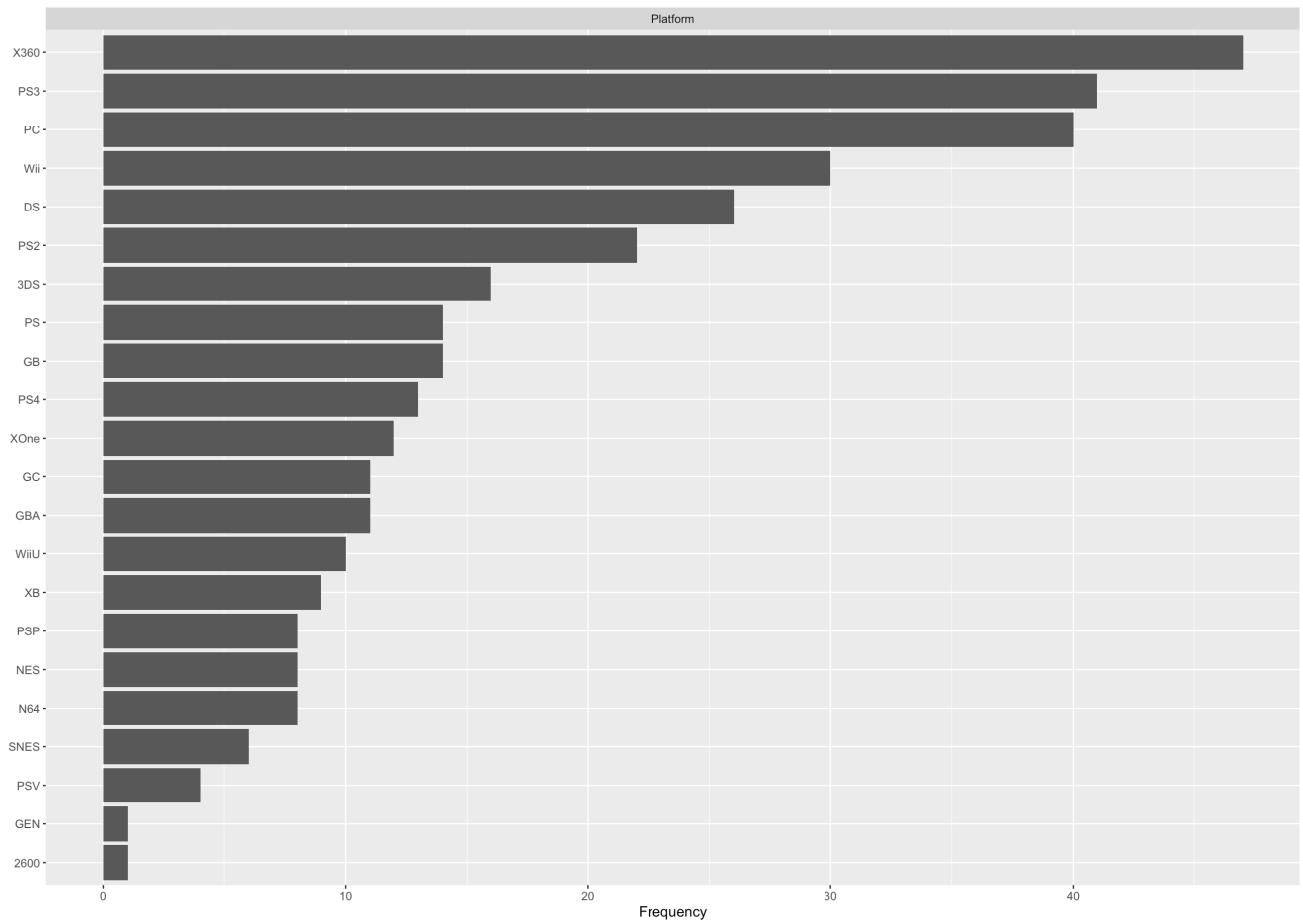
Missing Data Profile

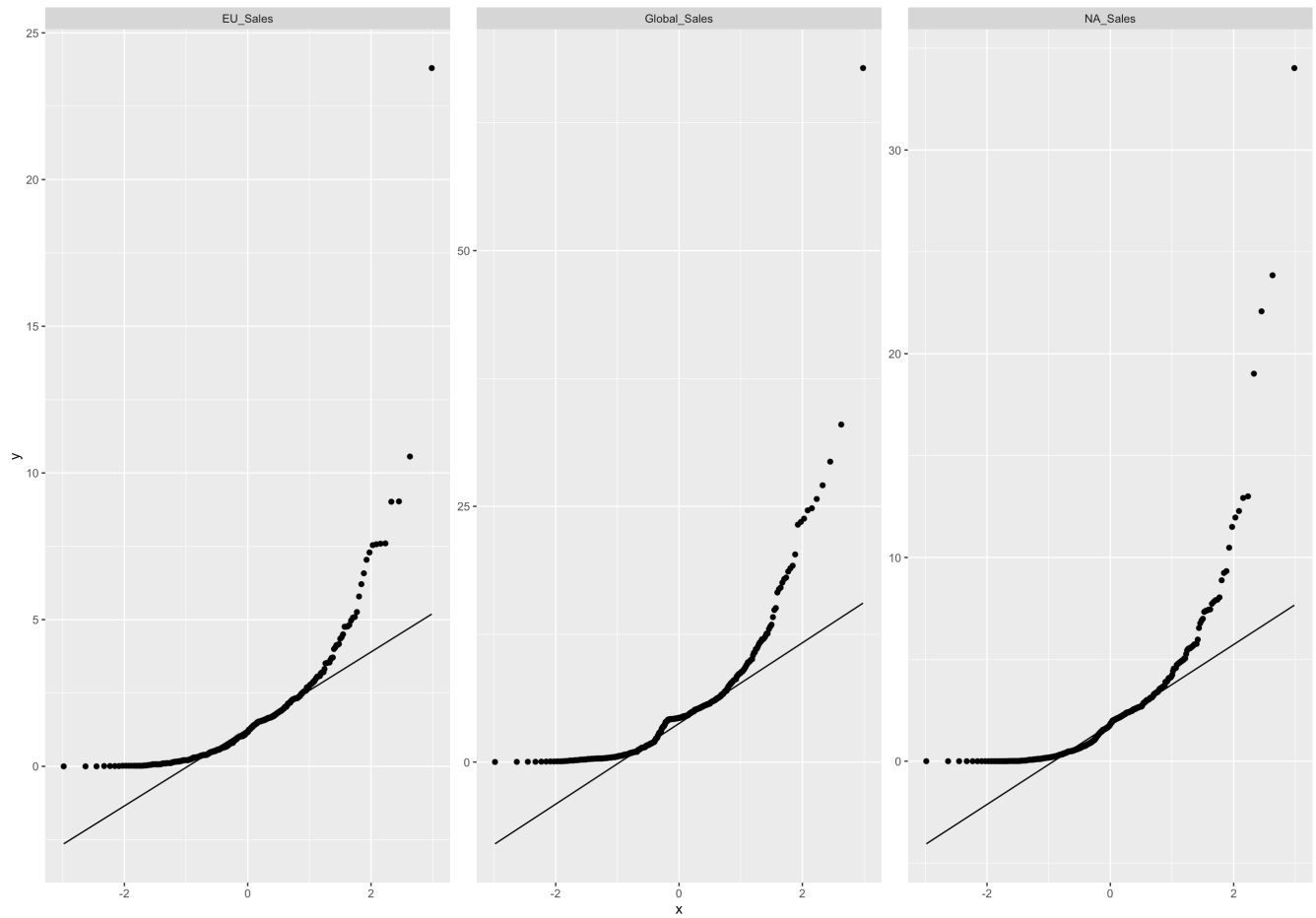## Univariate Distribution
Histogram

Bar Chart (with frequency)

```
## 1 columns ignored with more than 50 categories.
## Product: 175 categories
```
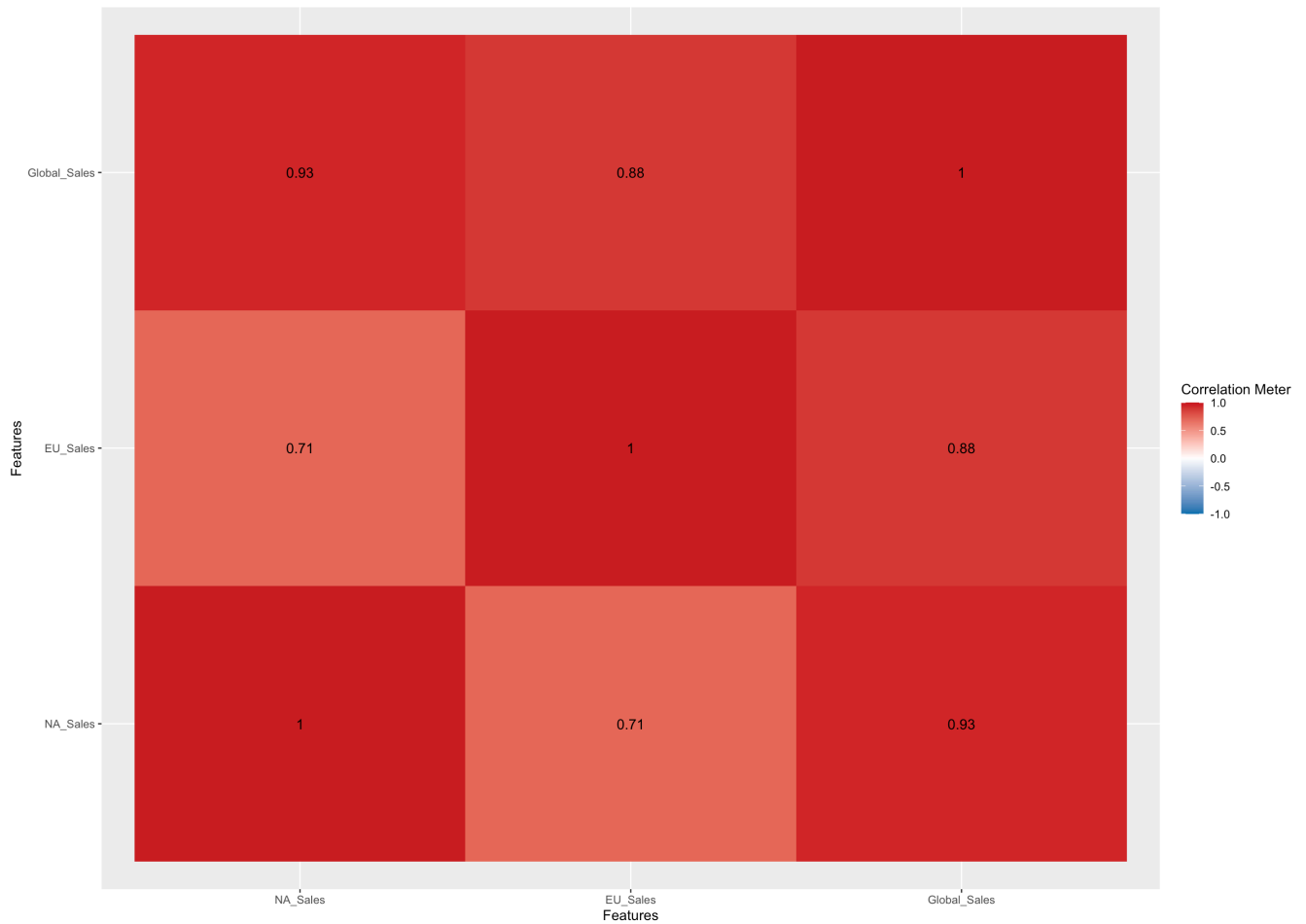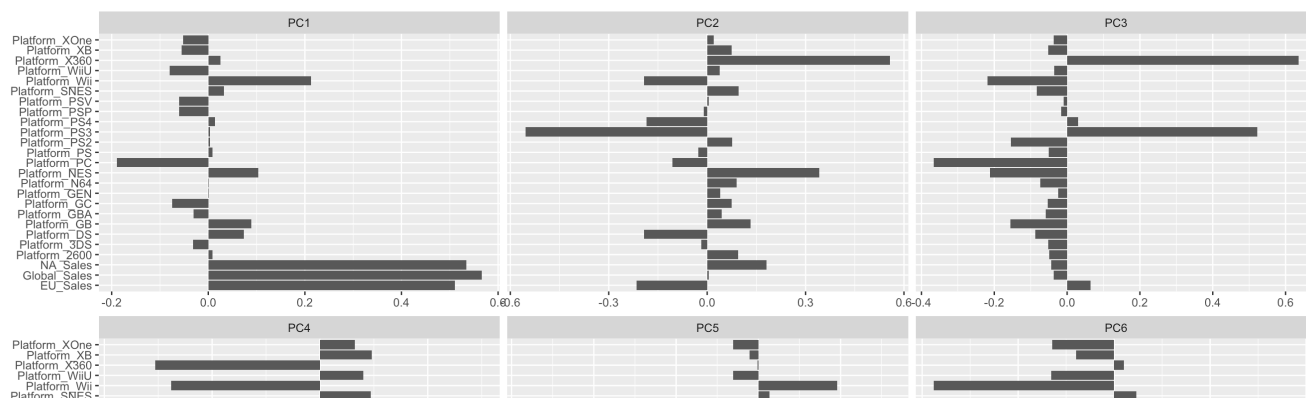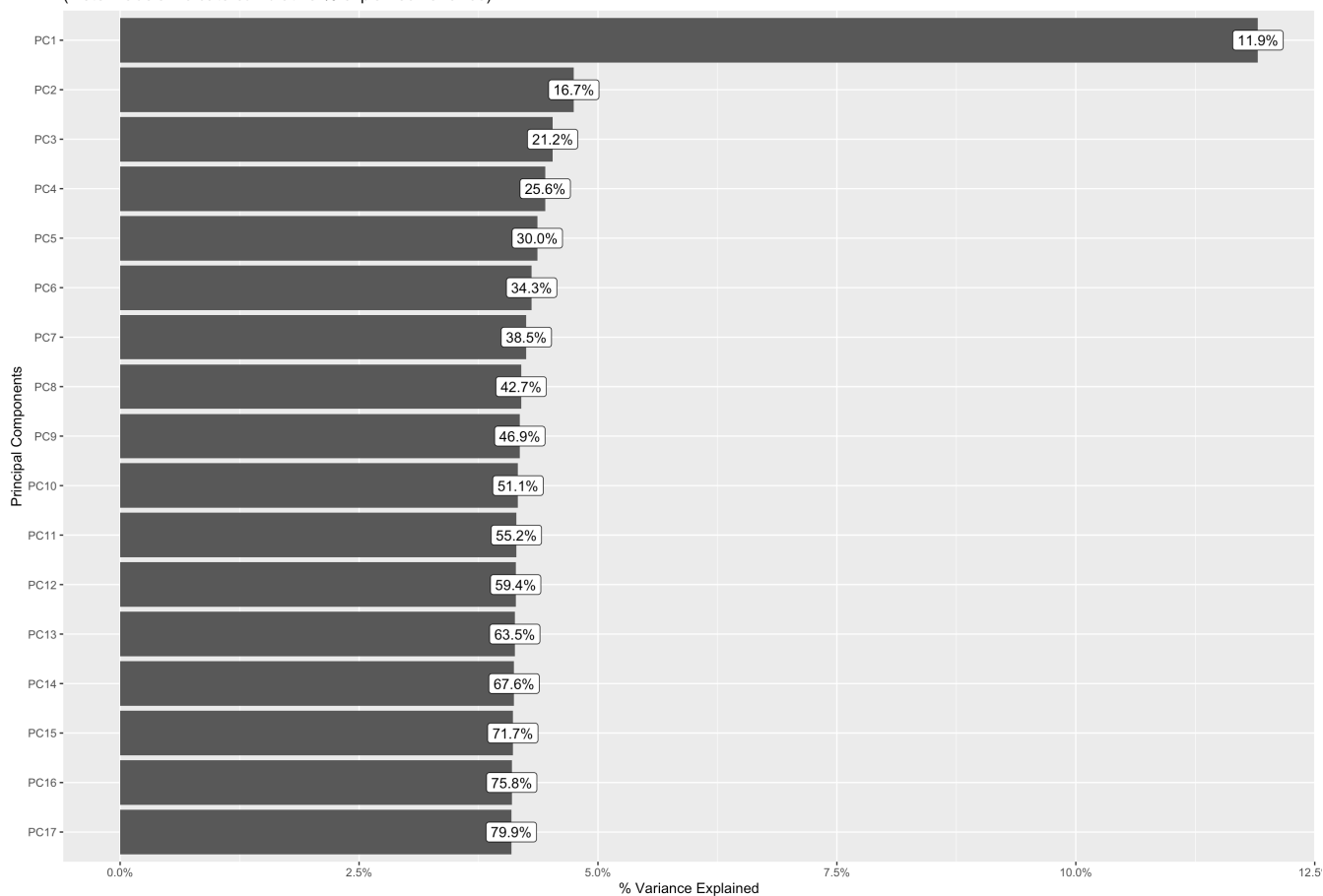
Platform

QQ Plot

Correlation Analysis

```
## Warning in dummify(data, maxcat = maxcat): Ignored all discrete features since `maxcat` set to 20 cate
gories!
```
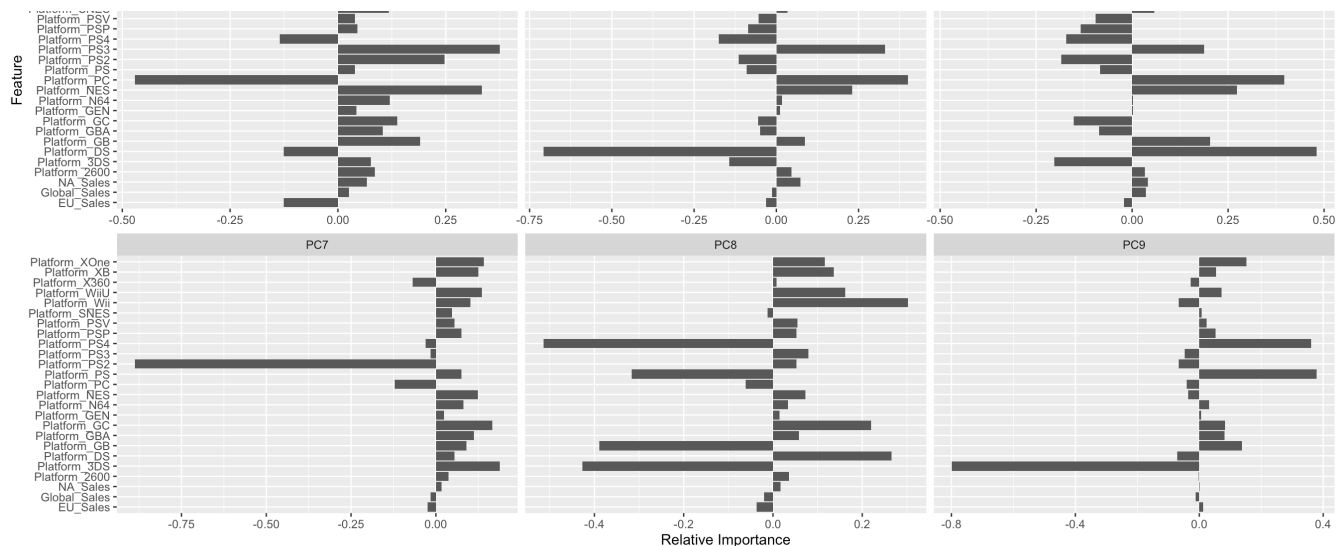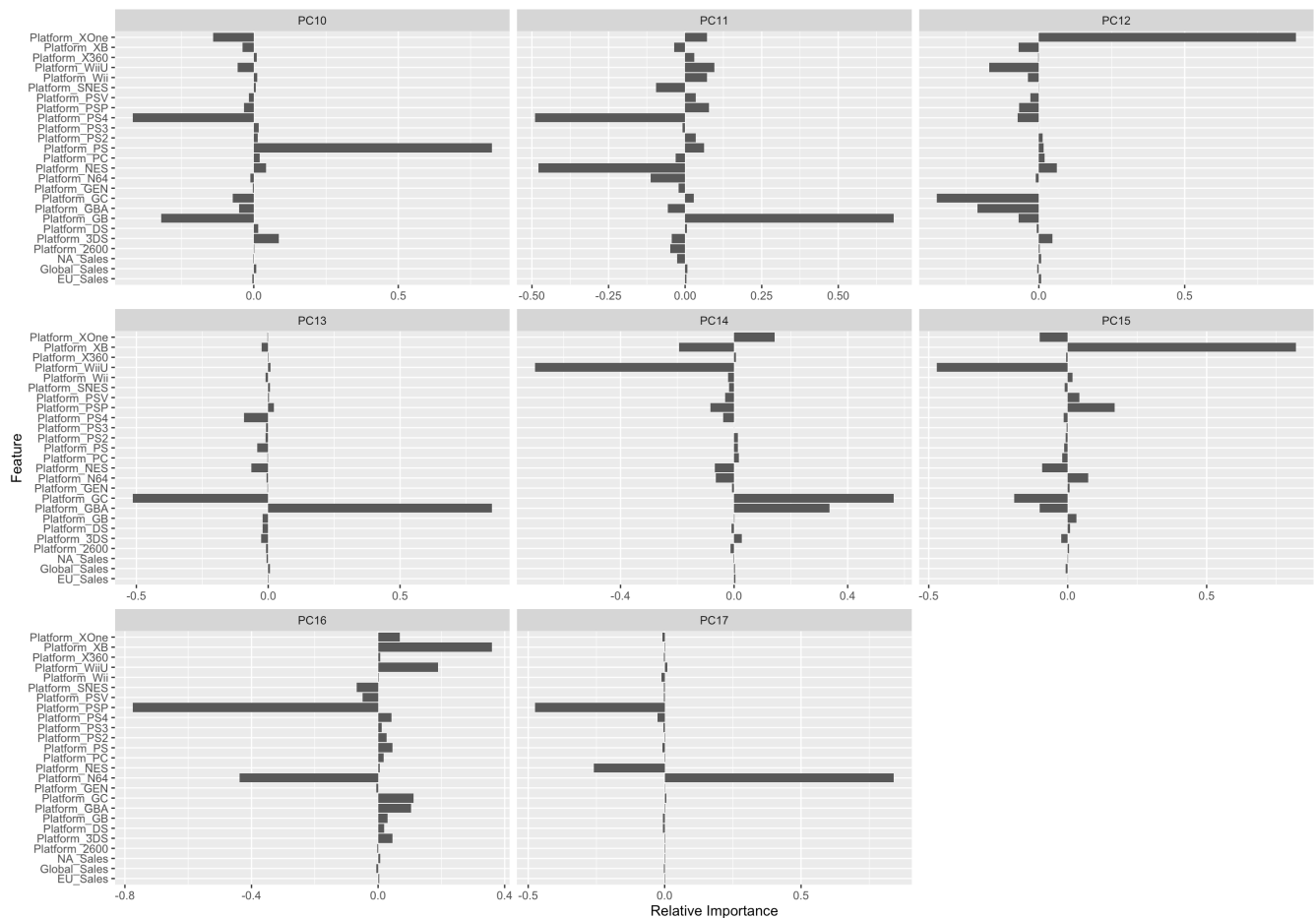
Principal Component Analysis

```
## 1 features with more than 50 categories ignored!
## Product: 175 categories
```

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

Page 1

Page 2