

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

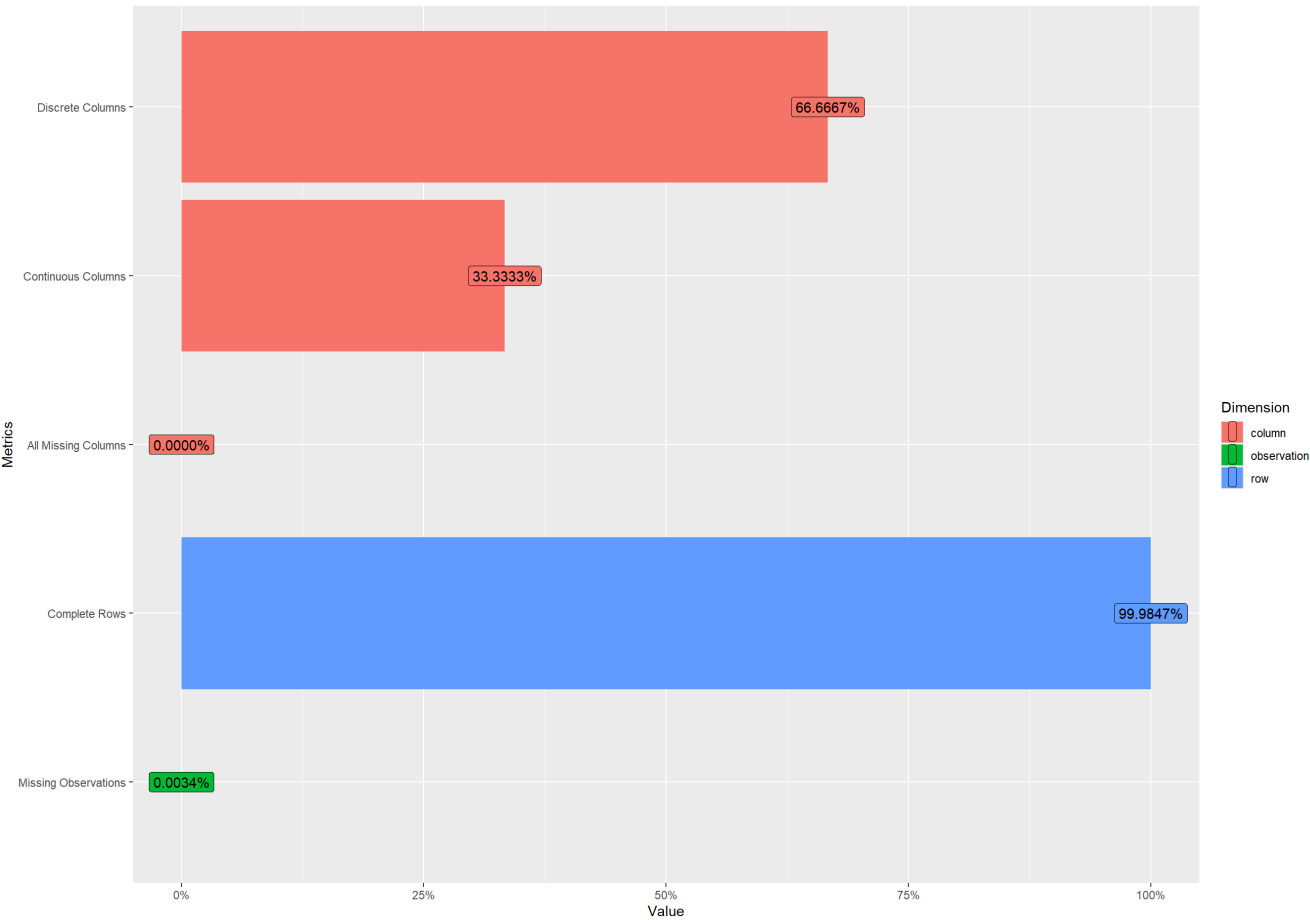
Basic Statistics

Raw Counts

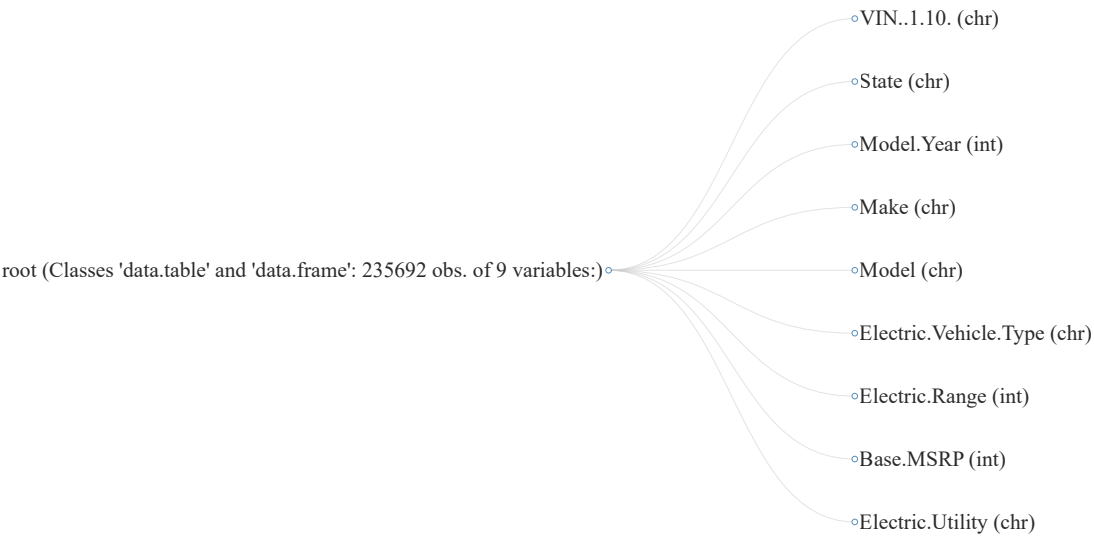
| Name | Value |
|----------------------|-----------|
| Rows | 235,692 |
| Columns | 9 |
| Discrete columns | 6 |
| Continuous columns | 3 |
| All missing columns | 0 |
| Missing observations | 72 |
| Complete Rows | 235,656 |
| Total observations | 2,121,228 |
| Memory allocation | 14.4 Mb |

Percentages

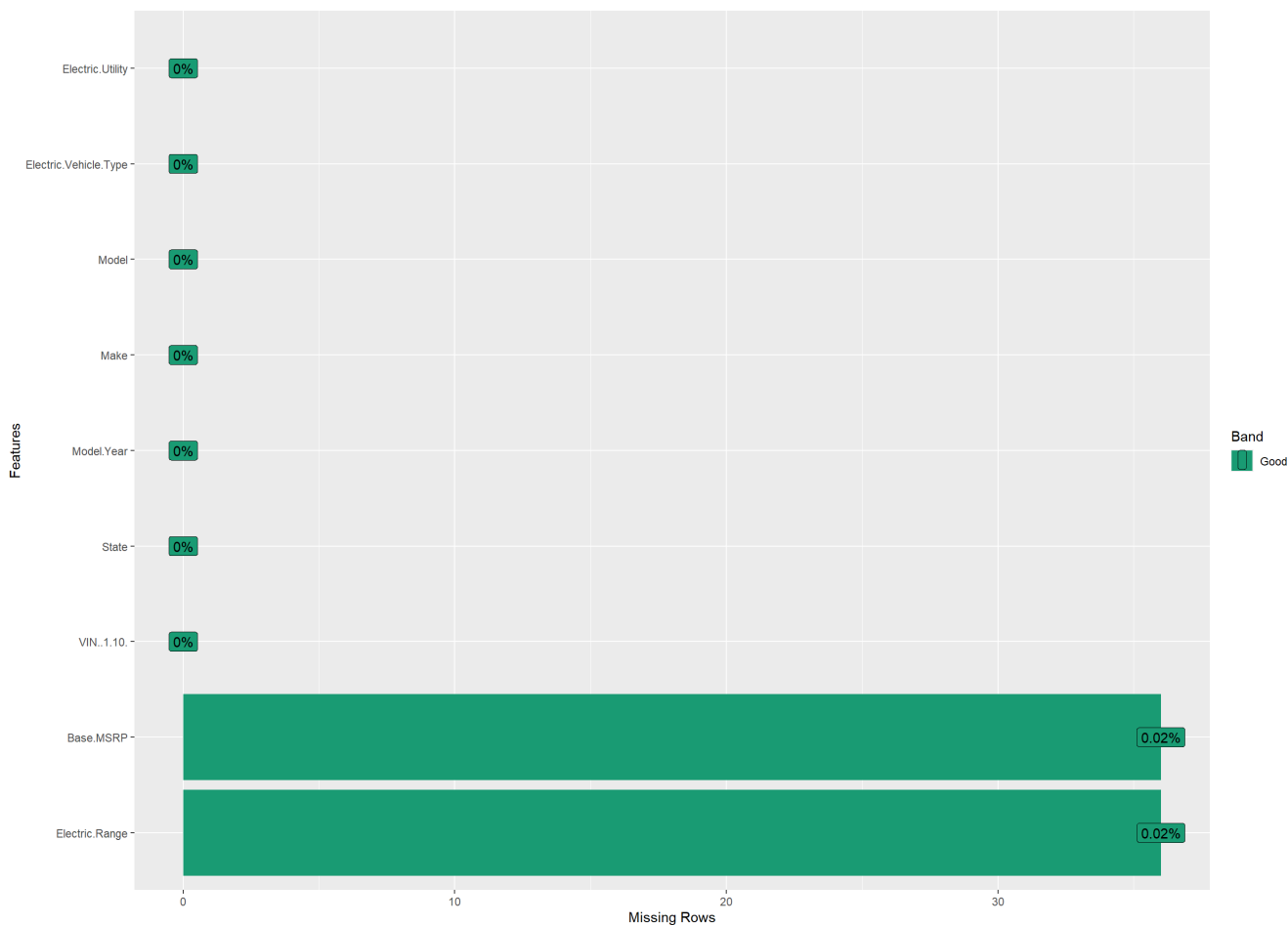
Memory Usage: 14.4 Mb



Data Structure

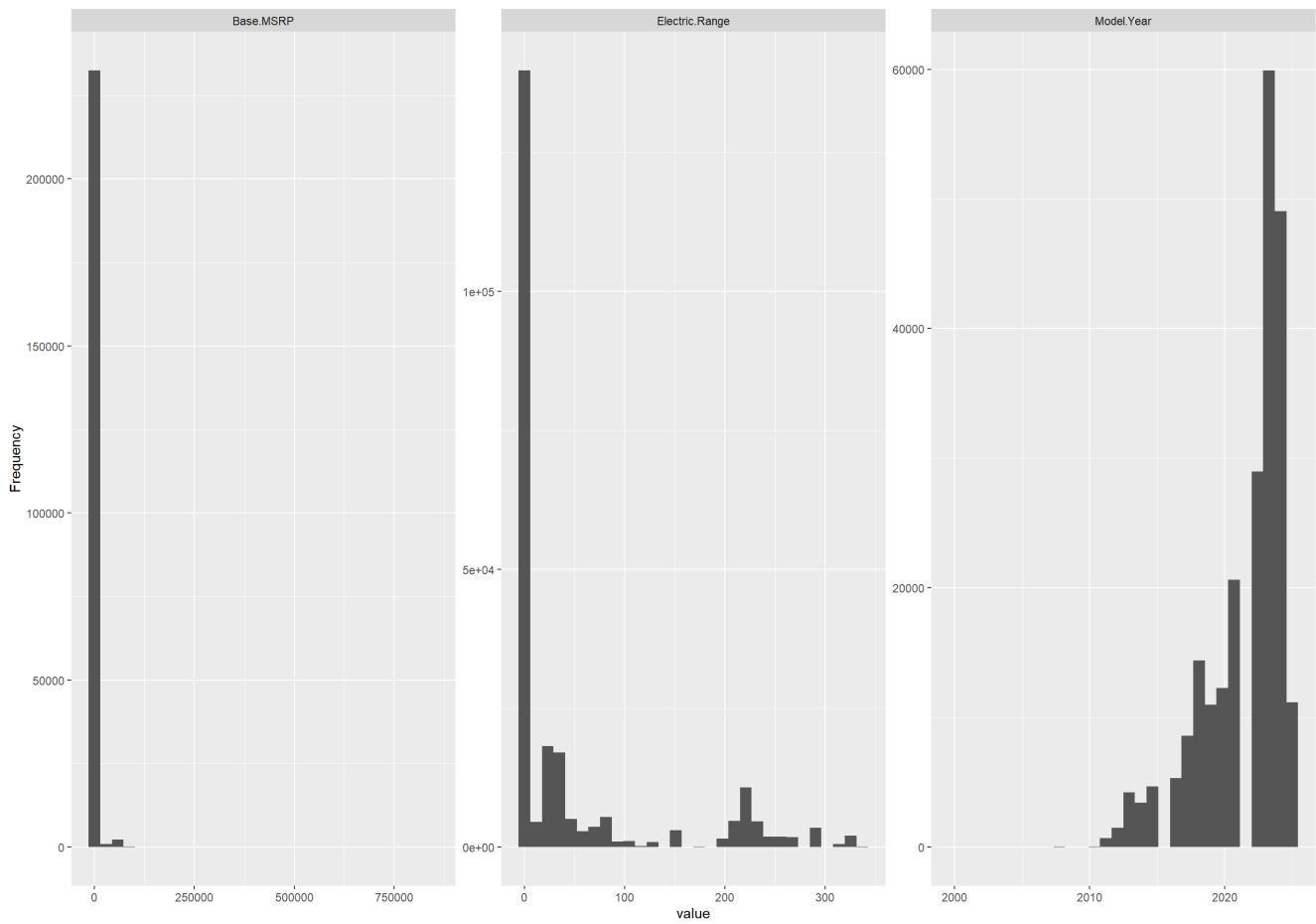


Missing Data Profile



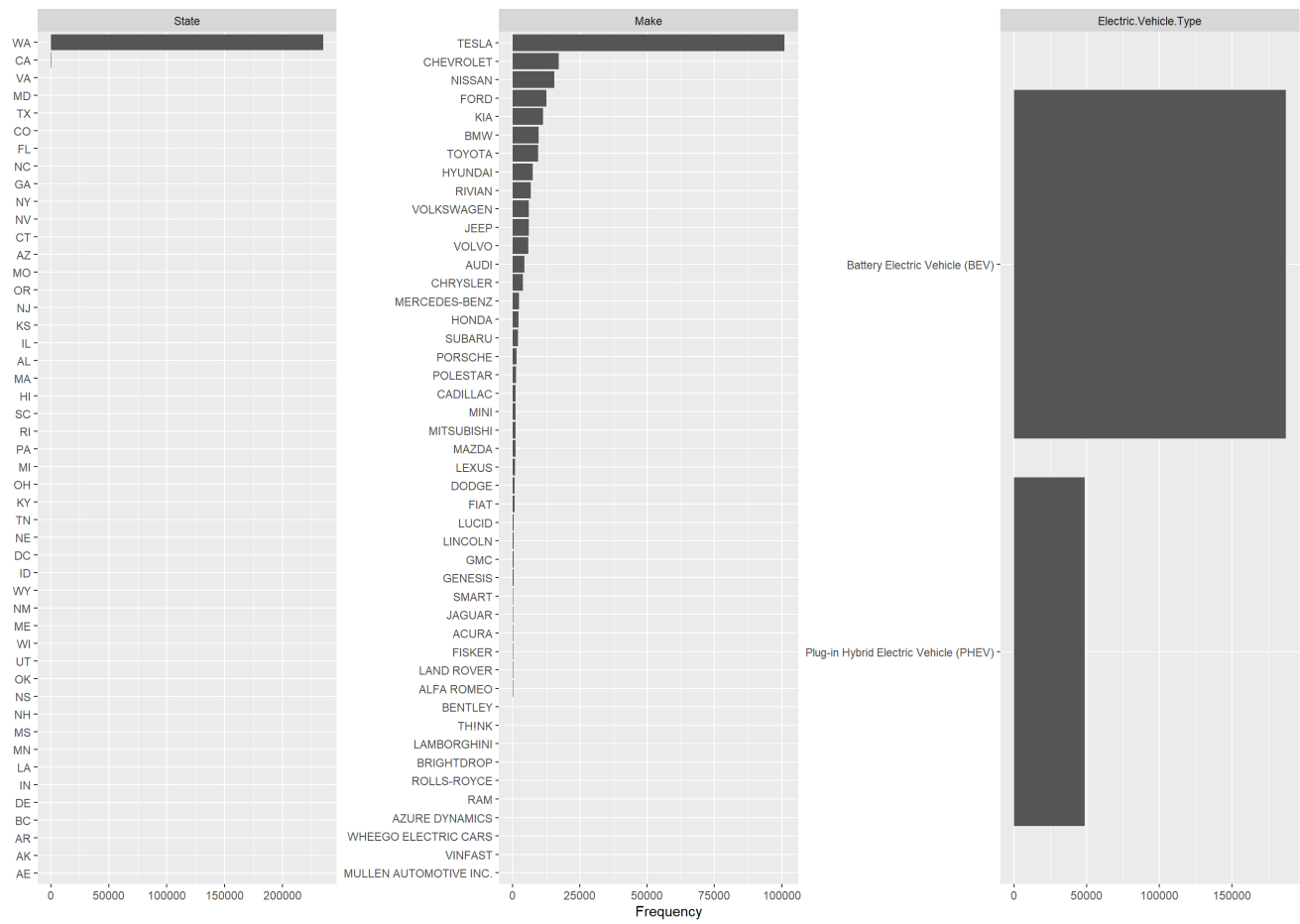
Univariate Distribution

Histogram

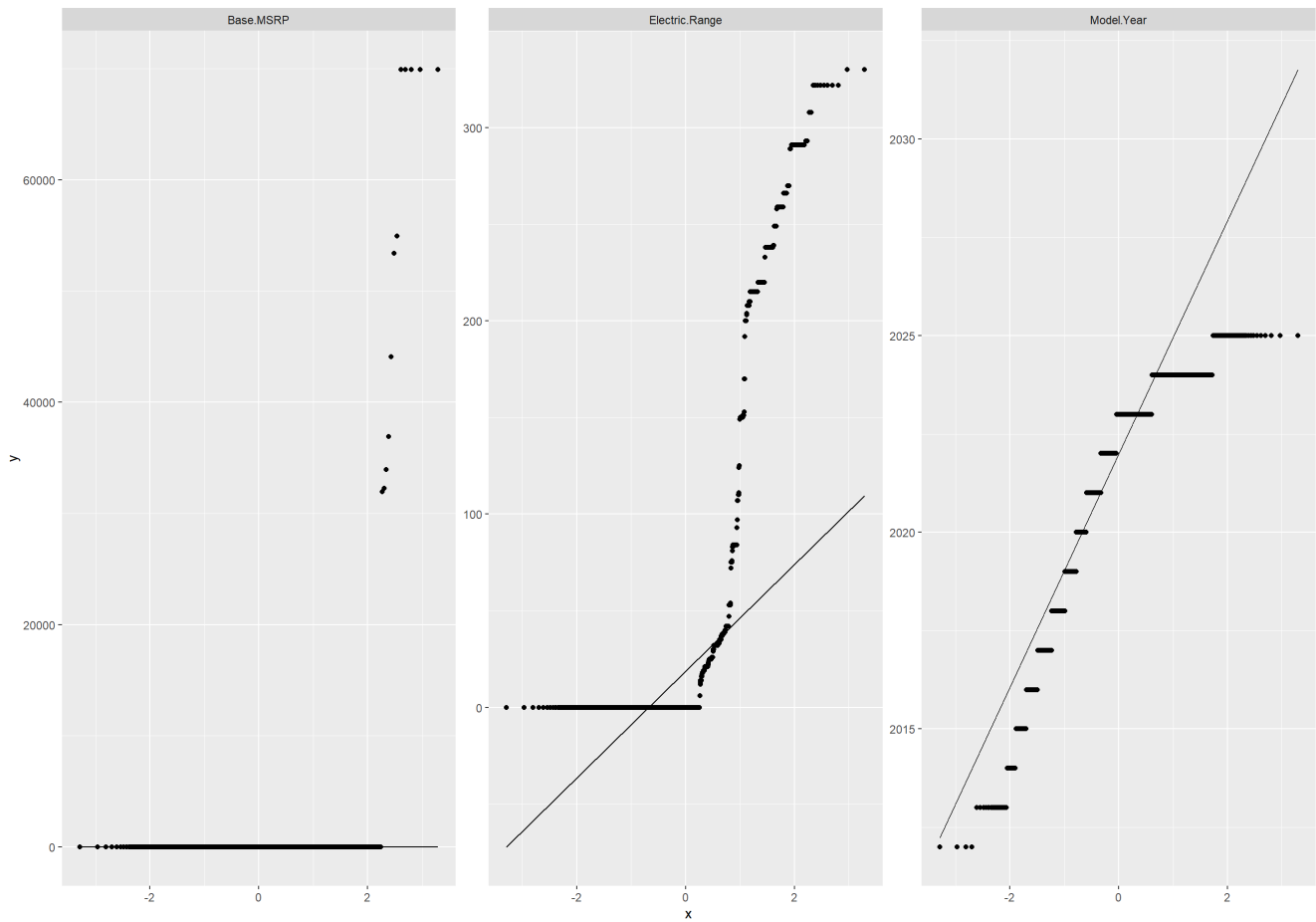


Bar Chart (with frequency)

```
## 3 columns ignored with more than 50 categories.  
## VIN..1.10.: 13763 categories  
## Model: 171 categories  
## Electric.Utility: 77 categories
```



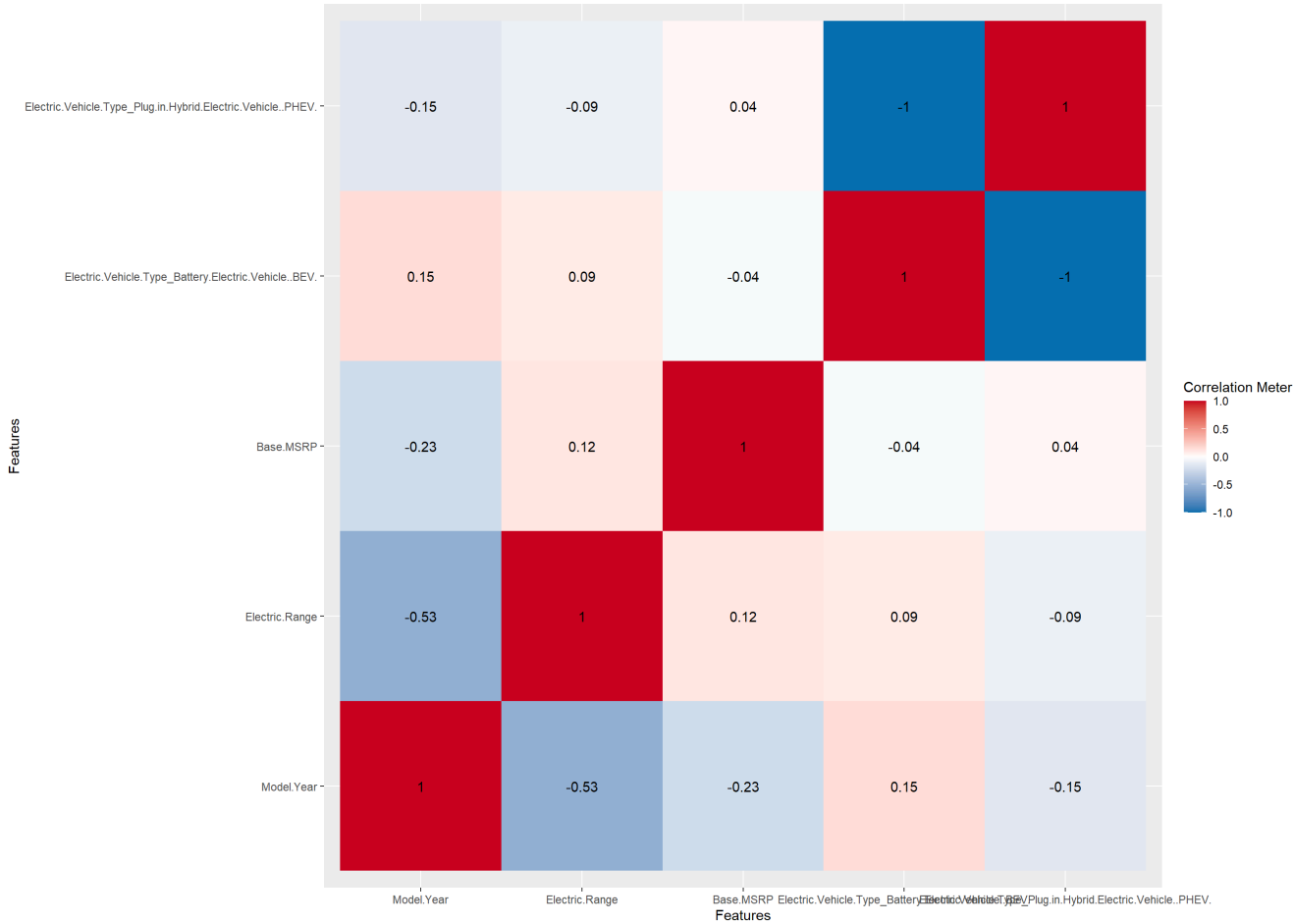
QQ Plot



Correlation Analysis

```
## 5 features with more than 20 categories ignored!  
## VIN..1.10.: 13732 categories  
## State: 48 categories  
## Make: 46 categories  
## Model: 170 categories  
## Electric.Utility: 77 categories
```

Processed 201352 groups out of 235656. 85% done. Time elapsed: 3s. ETA: 0s.Processed 235656 groups out of 235656. 100% done. Time elapsed: 3s. ETA: 0s.

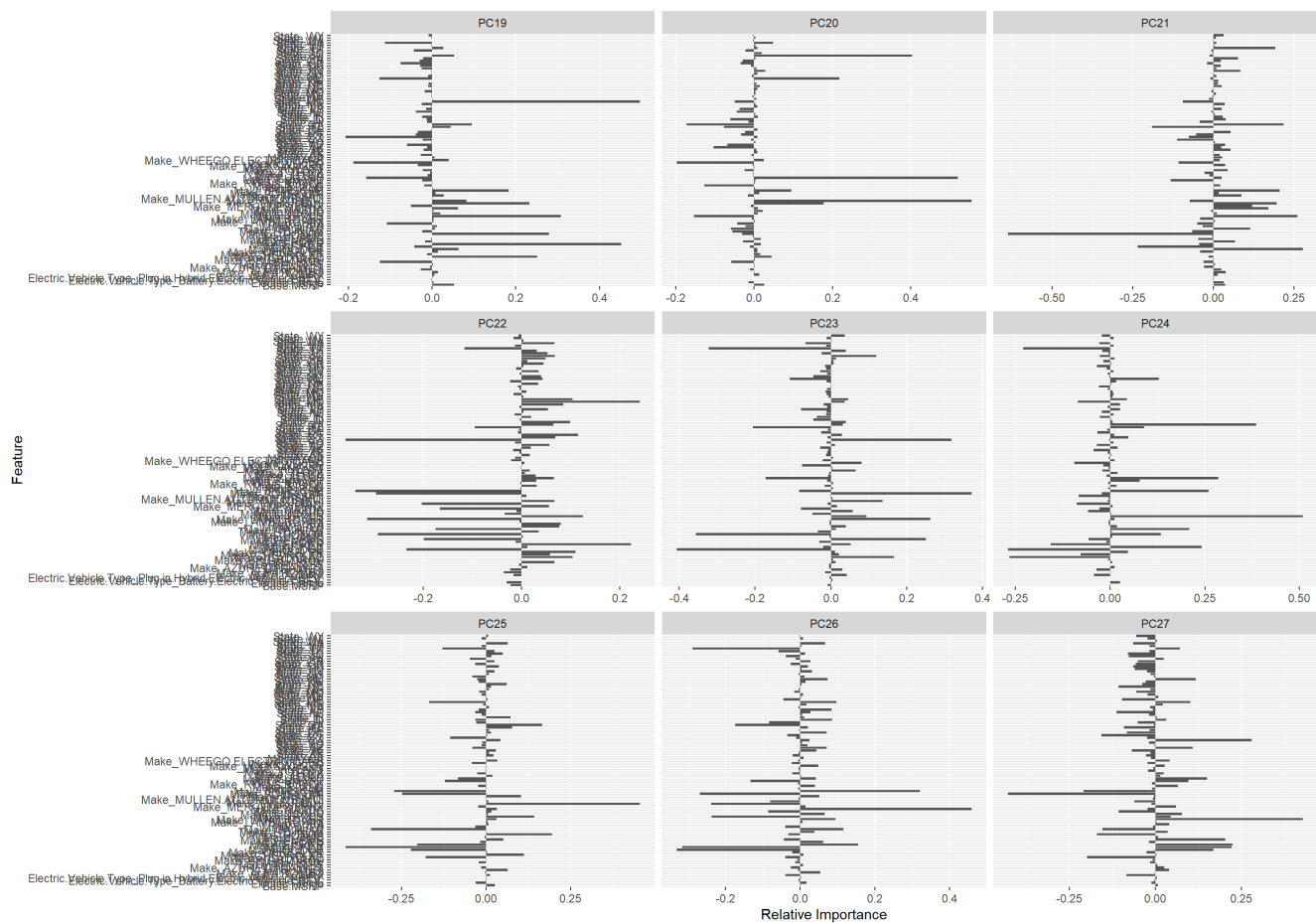
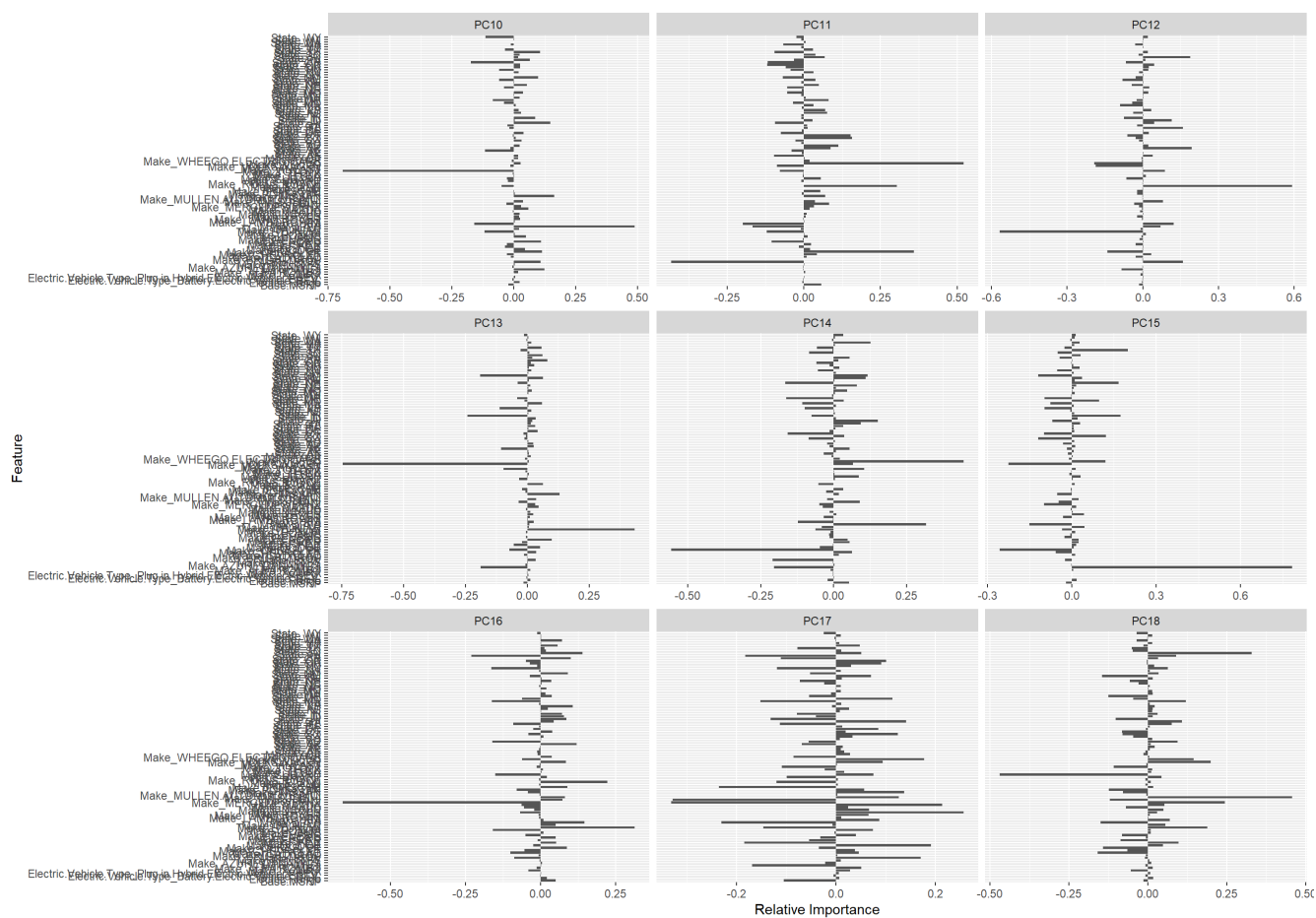


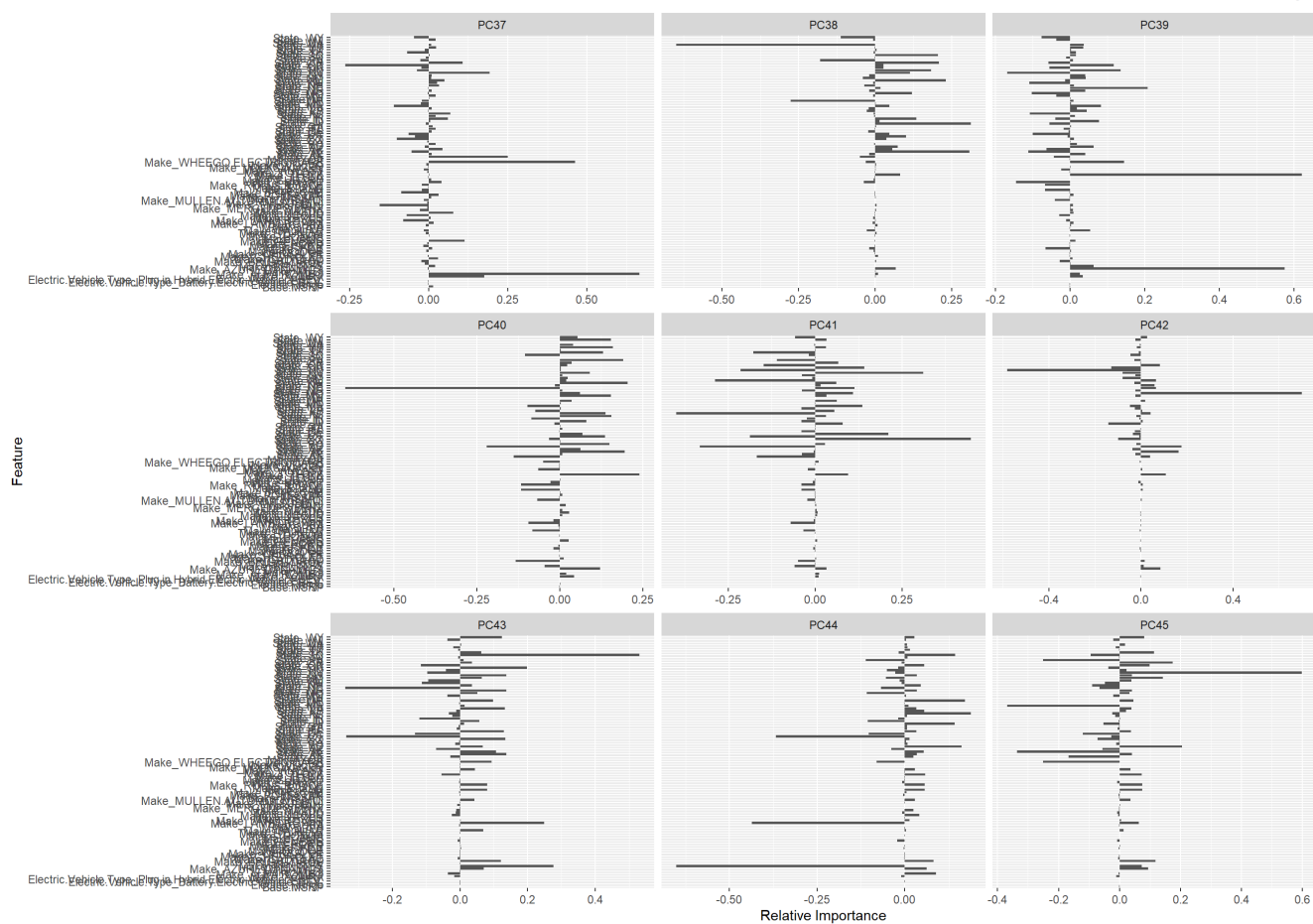
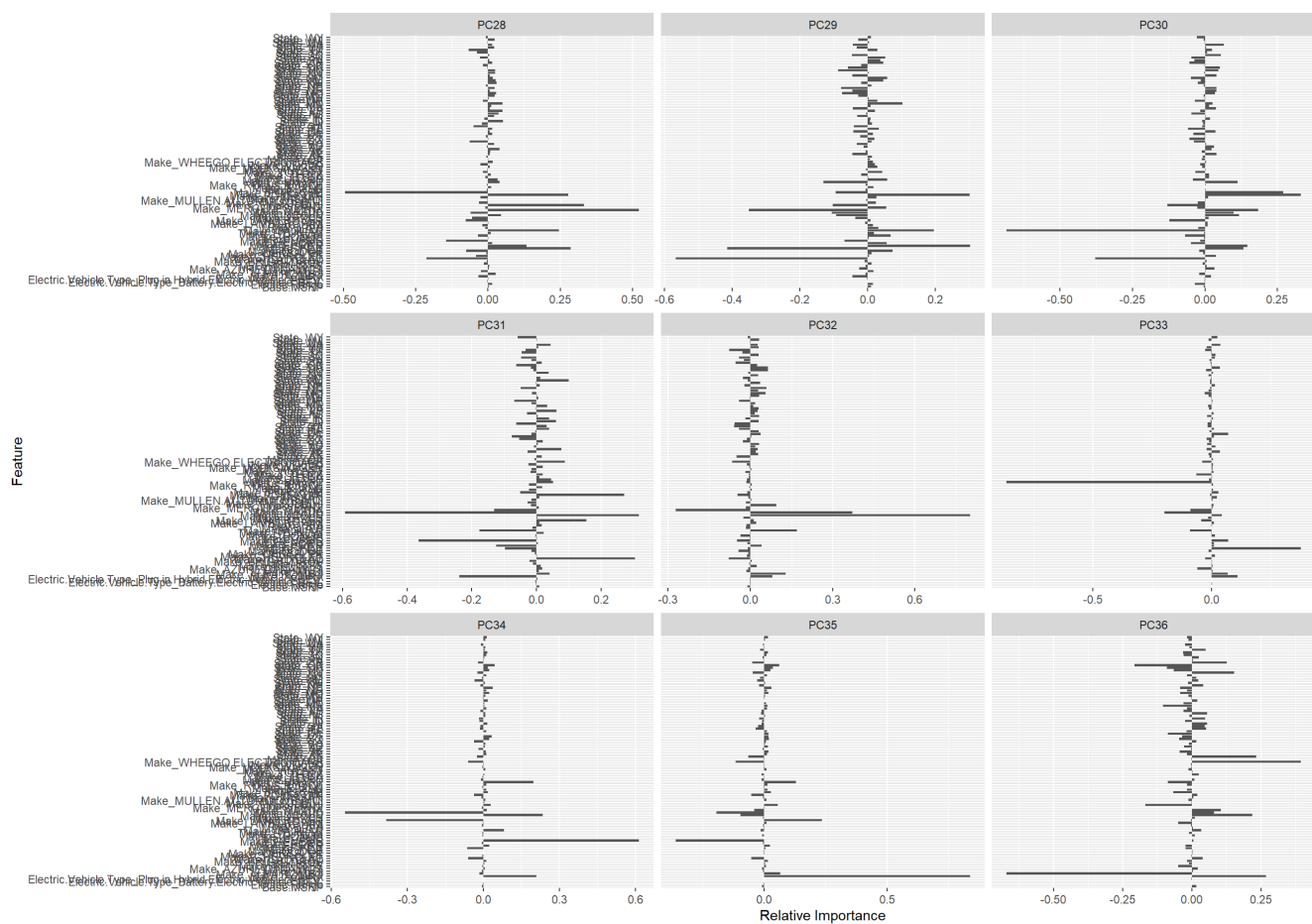
Principal Component Analysis

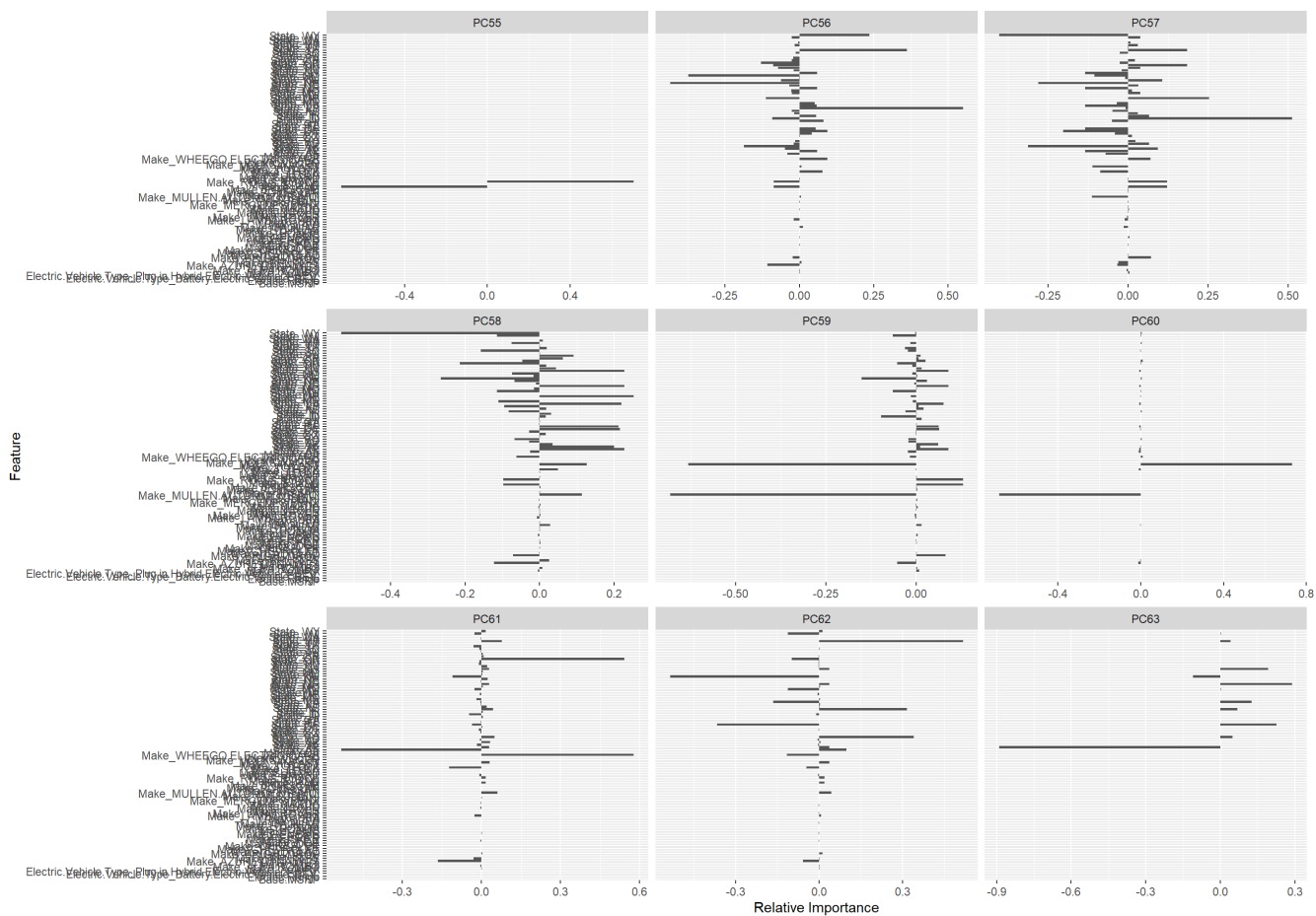
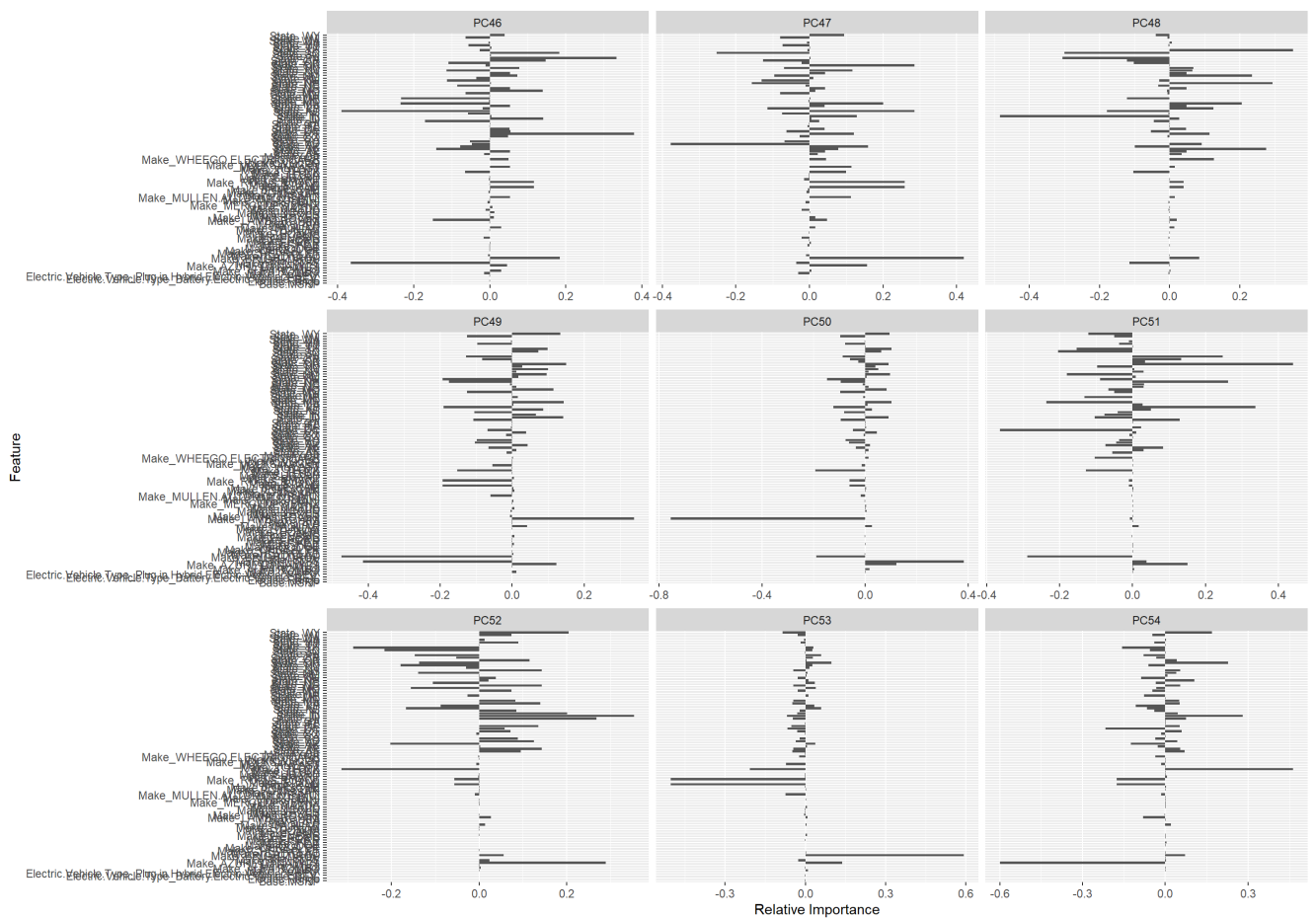
```
## 3 features with more than 50 categories ignored!  
## VIN..1.10.: 13732 categories  
## Model: 170 categories  
## Electric.Utility: 77 categories
```

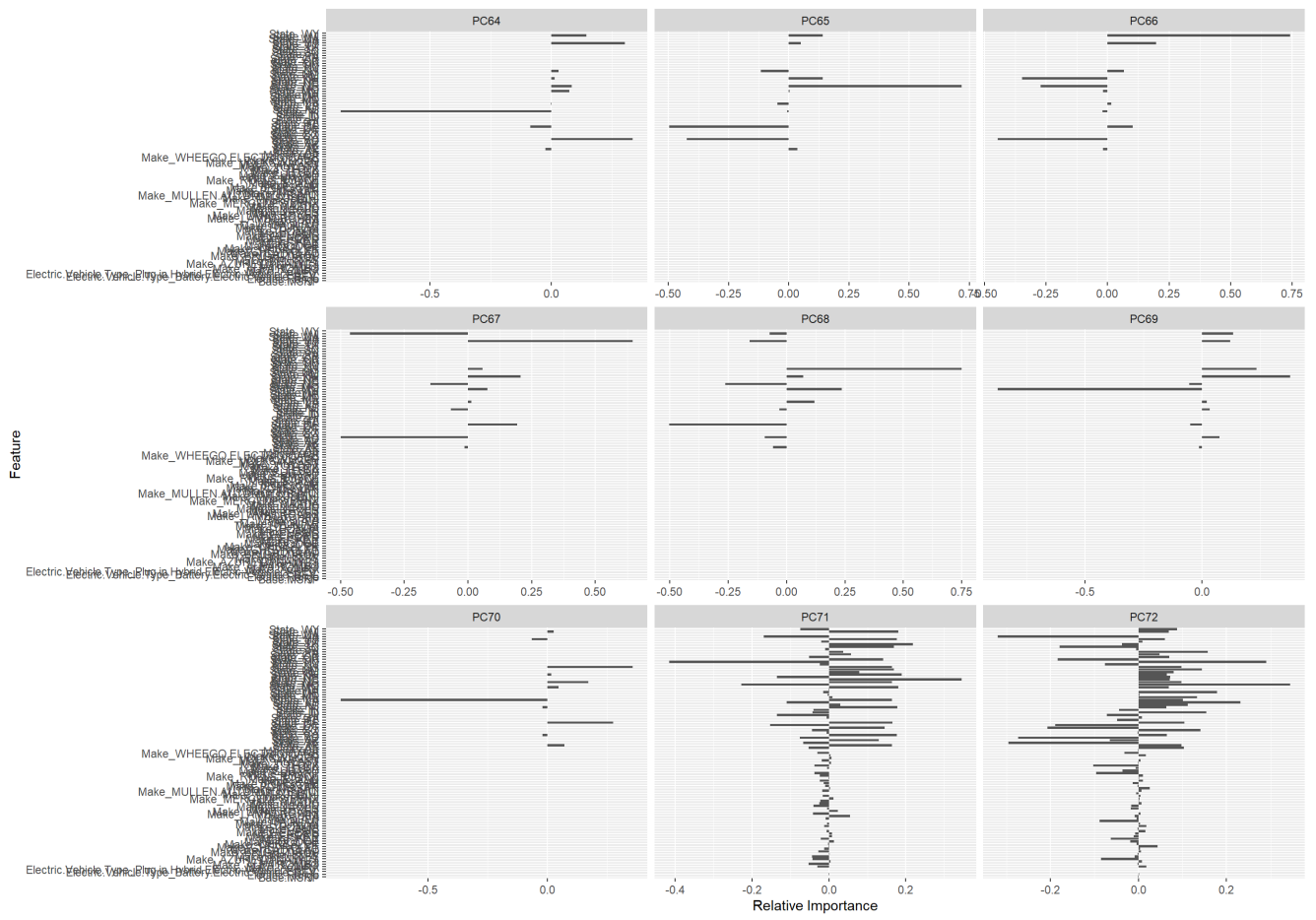
Processed 174710 groups out of 235656. 74% done. Time elapsed: 3s. ETA: 1s.Processed 230218 groups out of 235656. 98% done. Time elapsed: 4s. ETA: 0s.Processed 235656 groups out of 235656. 100% done. Time elapsed: 4s. ETA: 0s. Processed 174417 groups out of 235656. 74% done. Time elapsed: 3s. ETA: 1s.Processed 235656 groups out of 235656. 100% done. Time elapsed: 3s. ETA: 0s. Processed 168326 groups out of 235656. 71% done. Time elapsed: 3s. ETA: 1s.Processed 228883 groups out of 235656. 97% done. Time elapsed: 4s. ETA: 0s.Processed 235656 groups out of 235656. 100% done. Time elapsed: 4s. ETA: 0s.

2.9%

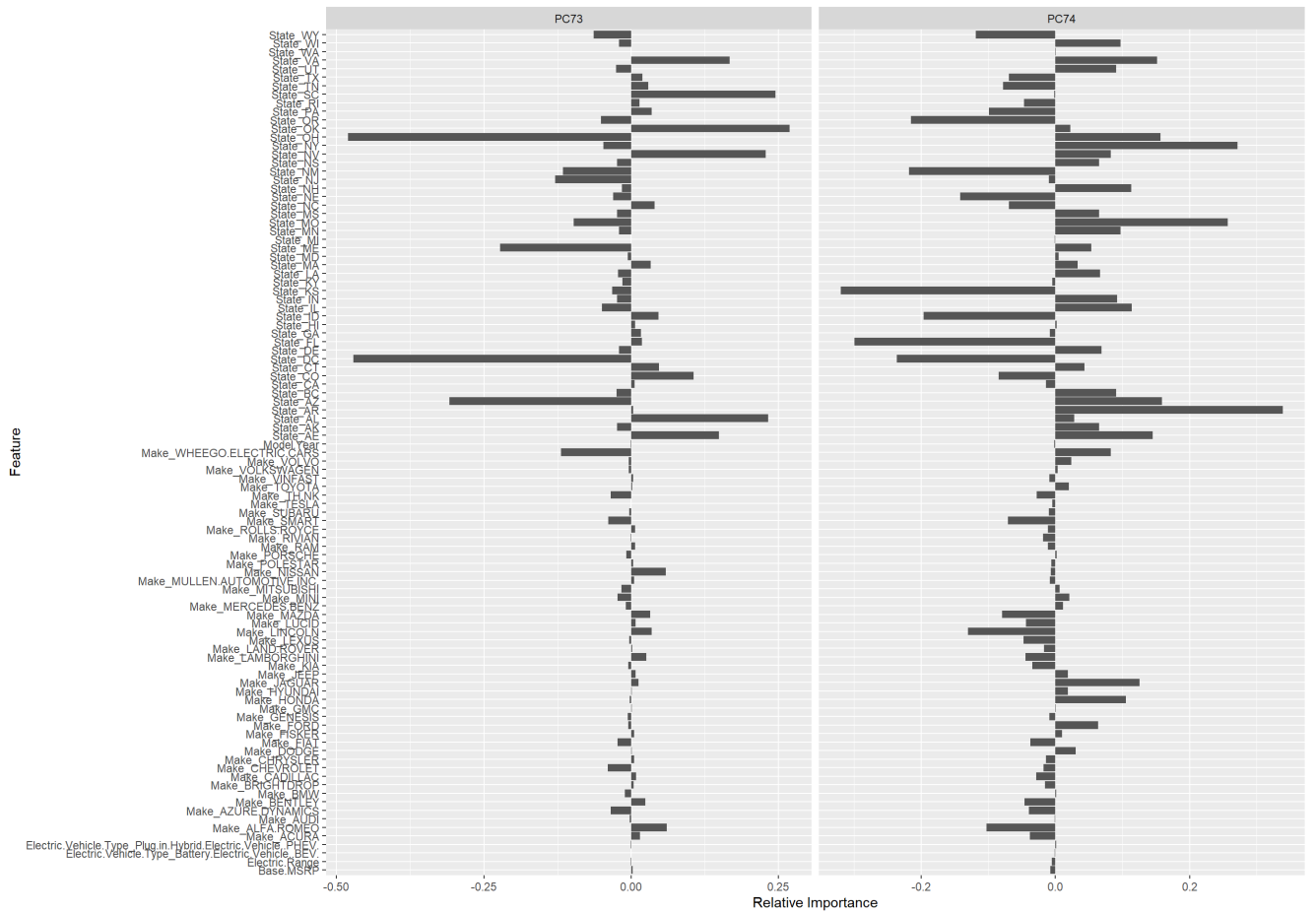








Page 8



Page 9