

Data Profiling Report

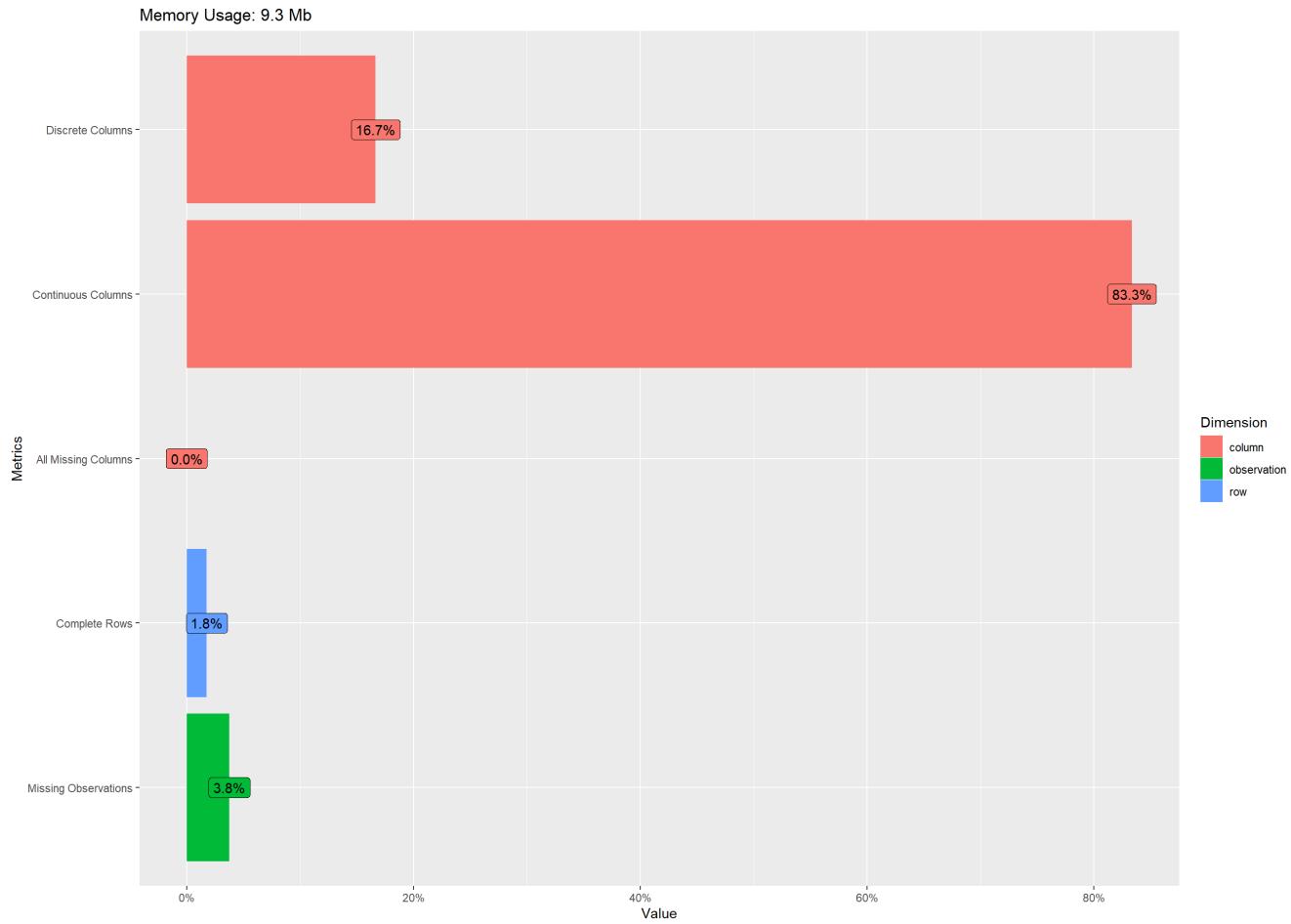
- Basic Statistics
 - Raw Counts
 - Percentages
- Data Structure
- Missing Data Profile
- Univariate Distribution
 - Histogram
 - Bar Chart (by frequency)
 - QQ Plot
- Correlation Analysis
- Principal Component Analysis

Basic Statistics

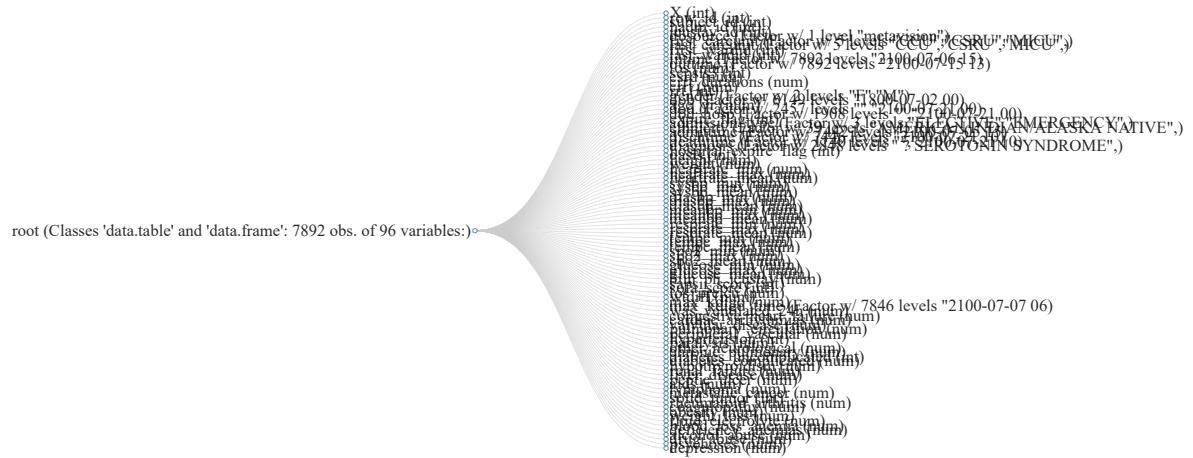
Raw Counts

Name	Value
Rows	7,892
Columns	96
Discrete columns	16
Continuous columns	80
All missing columns	0
Missing observations	28,516
Complete Rows	140
Total observations	757,632
Memory allocation	9.3 Mb

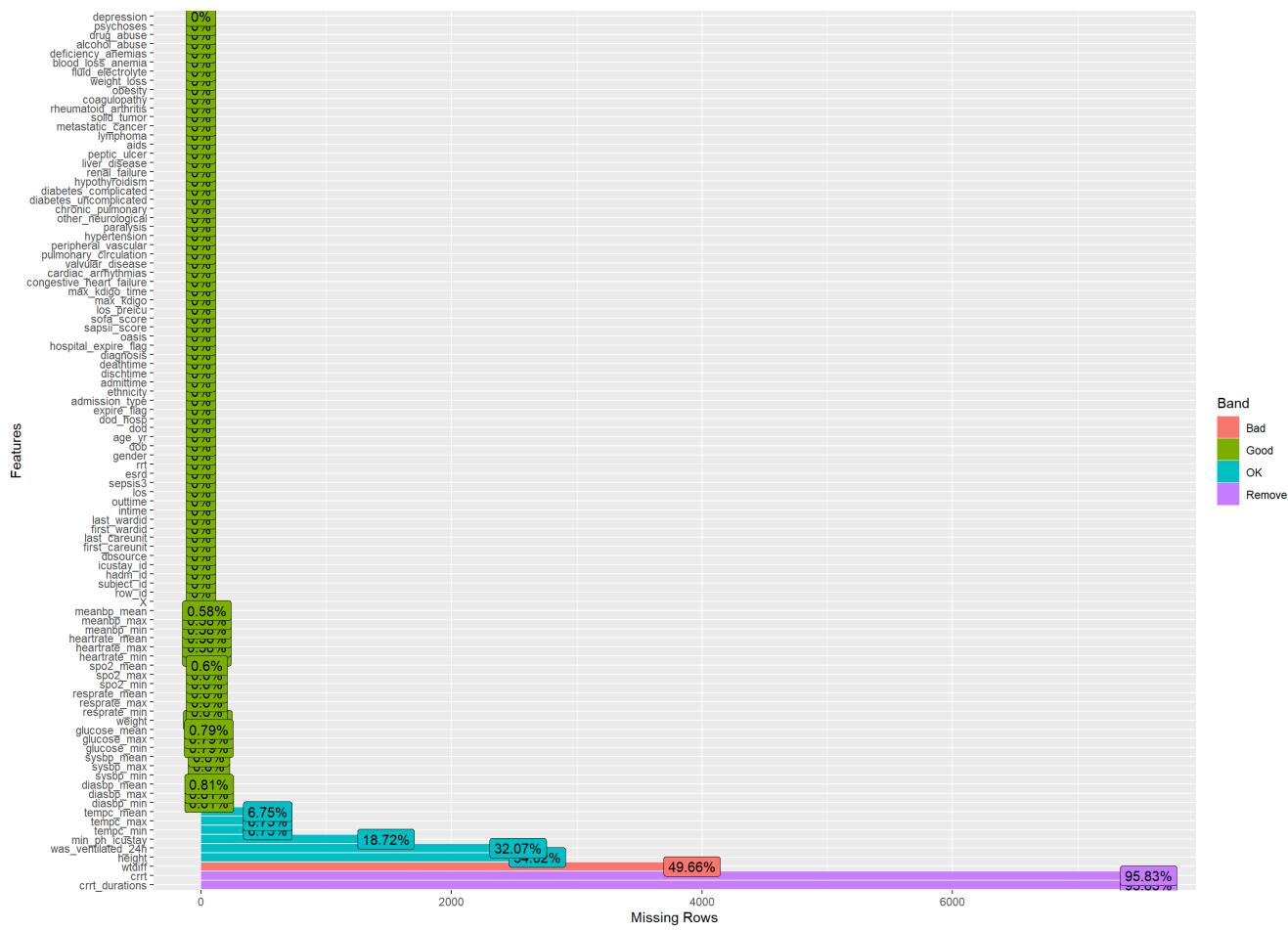
Percentages



Data Structure



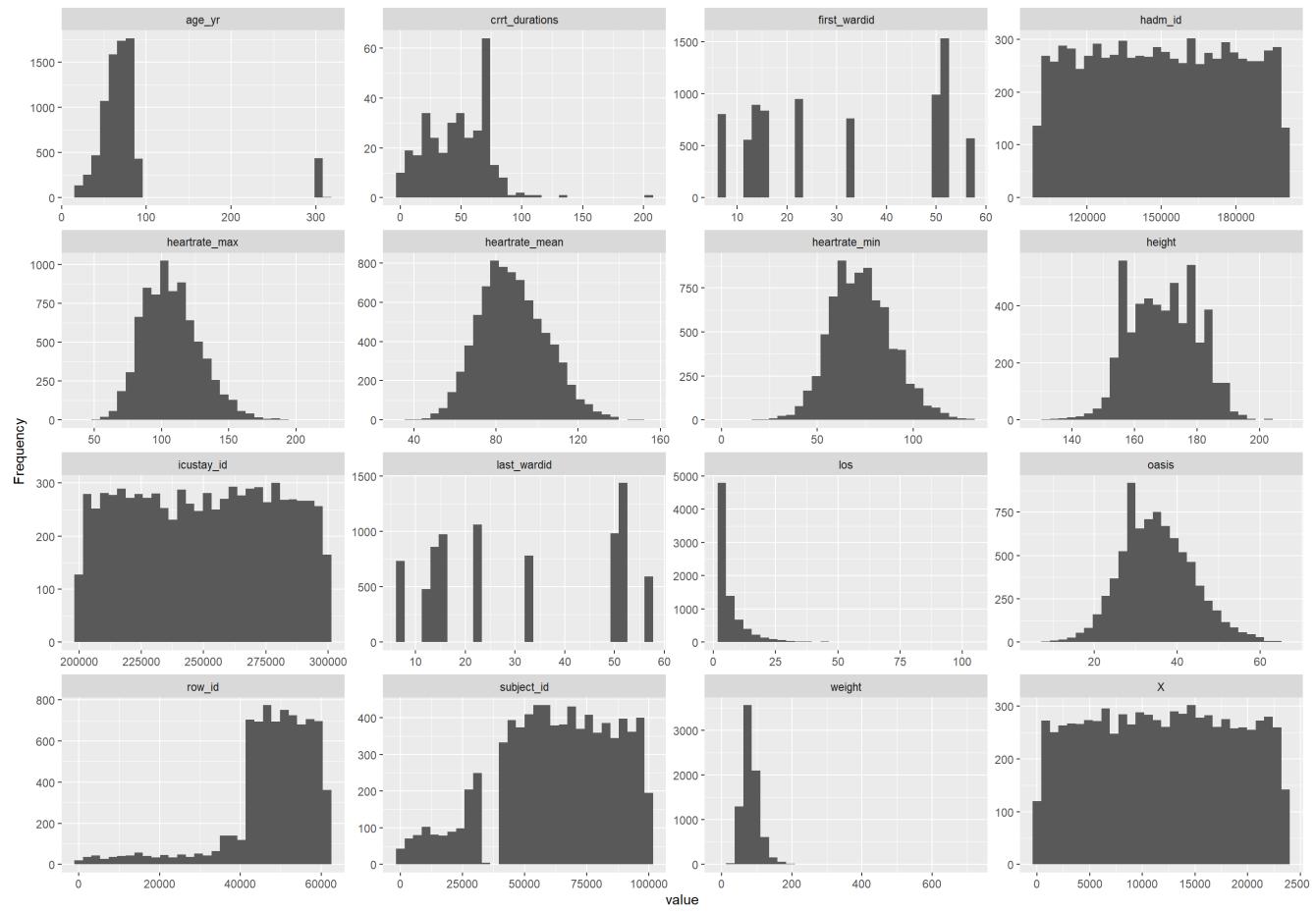
Missing Data Profile



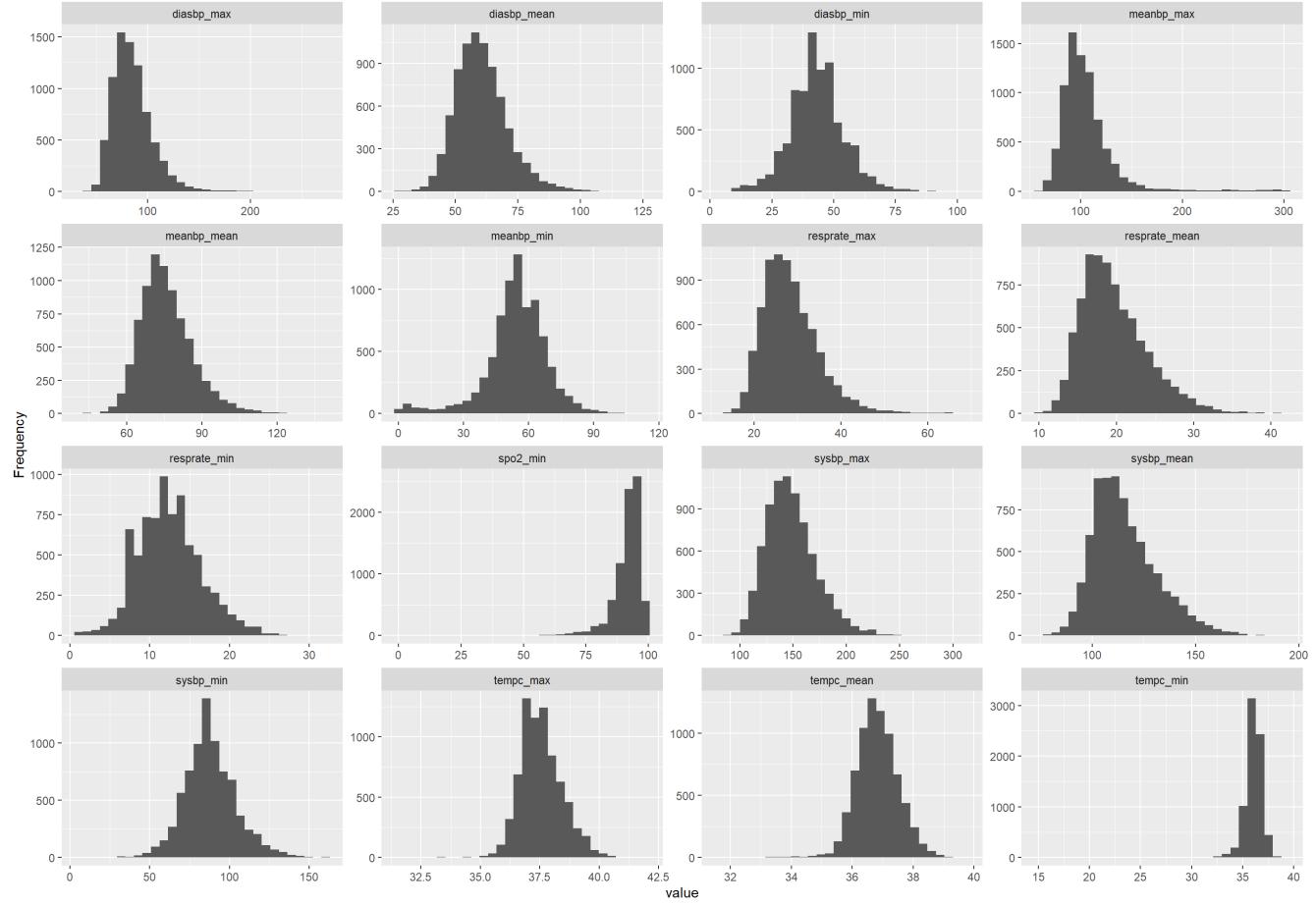
Univariate Distribution

Histogram

Data Profiling Report

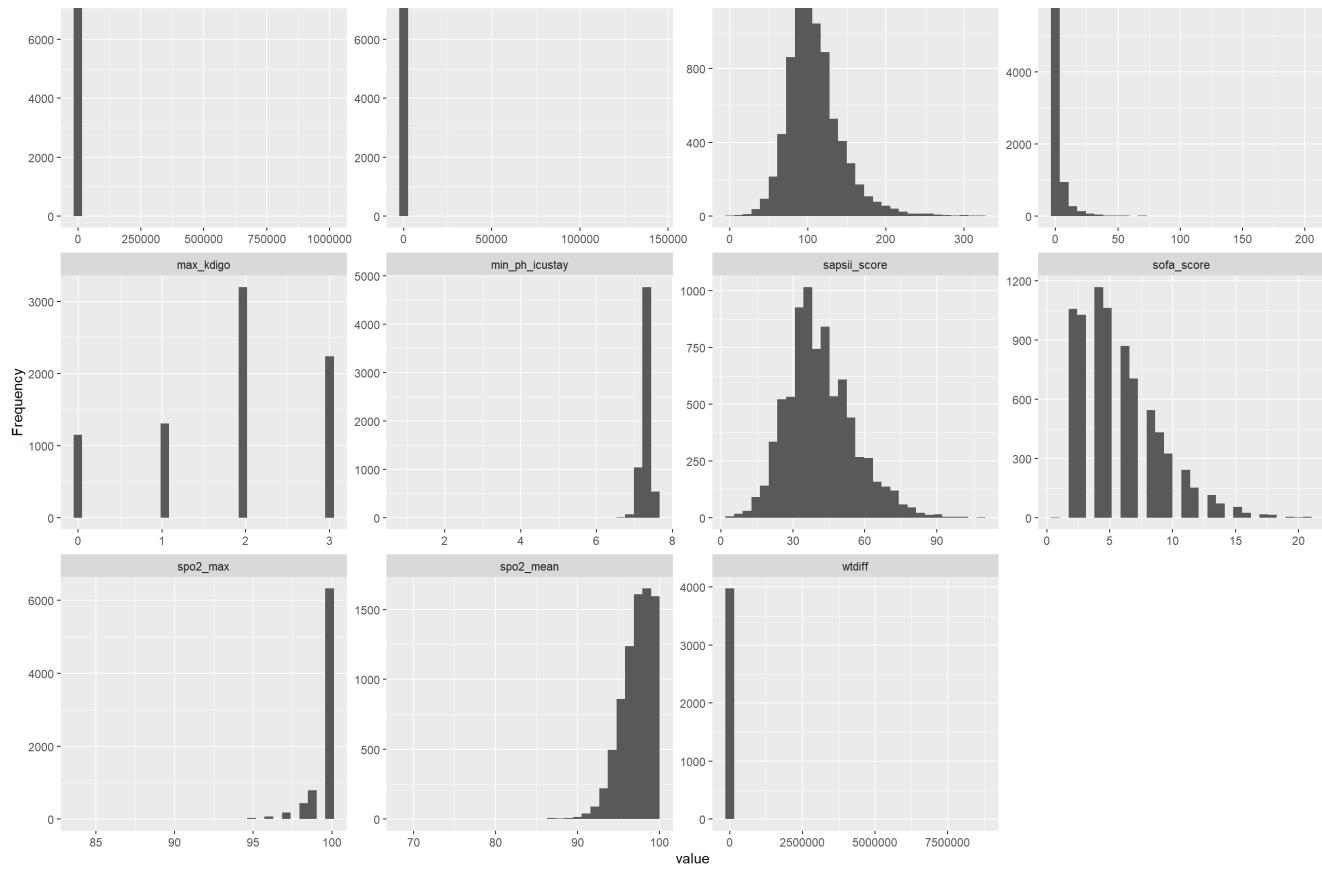


Page 1



Page 2

Data Profiling Report

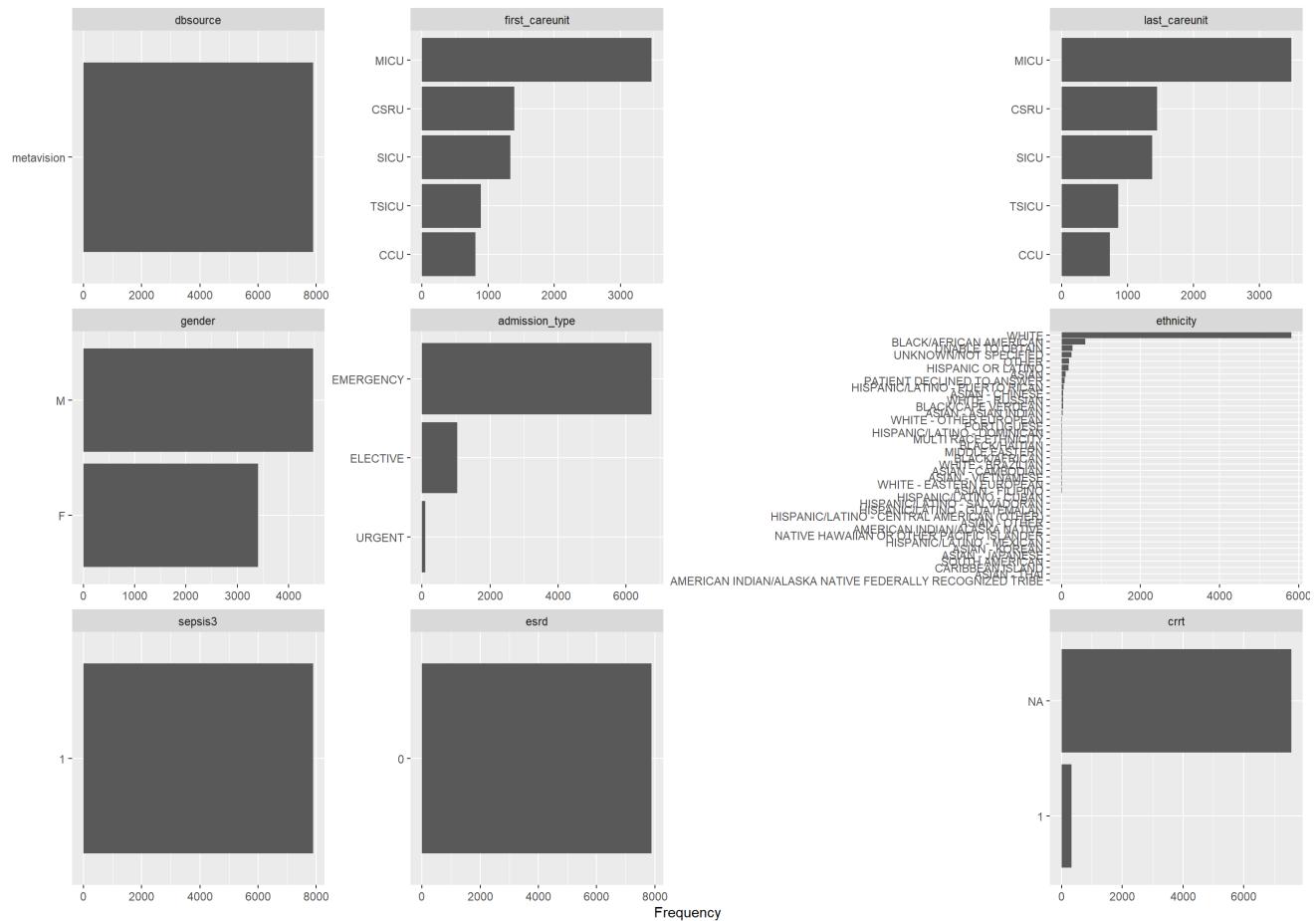


Page 3

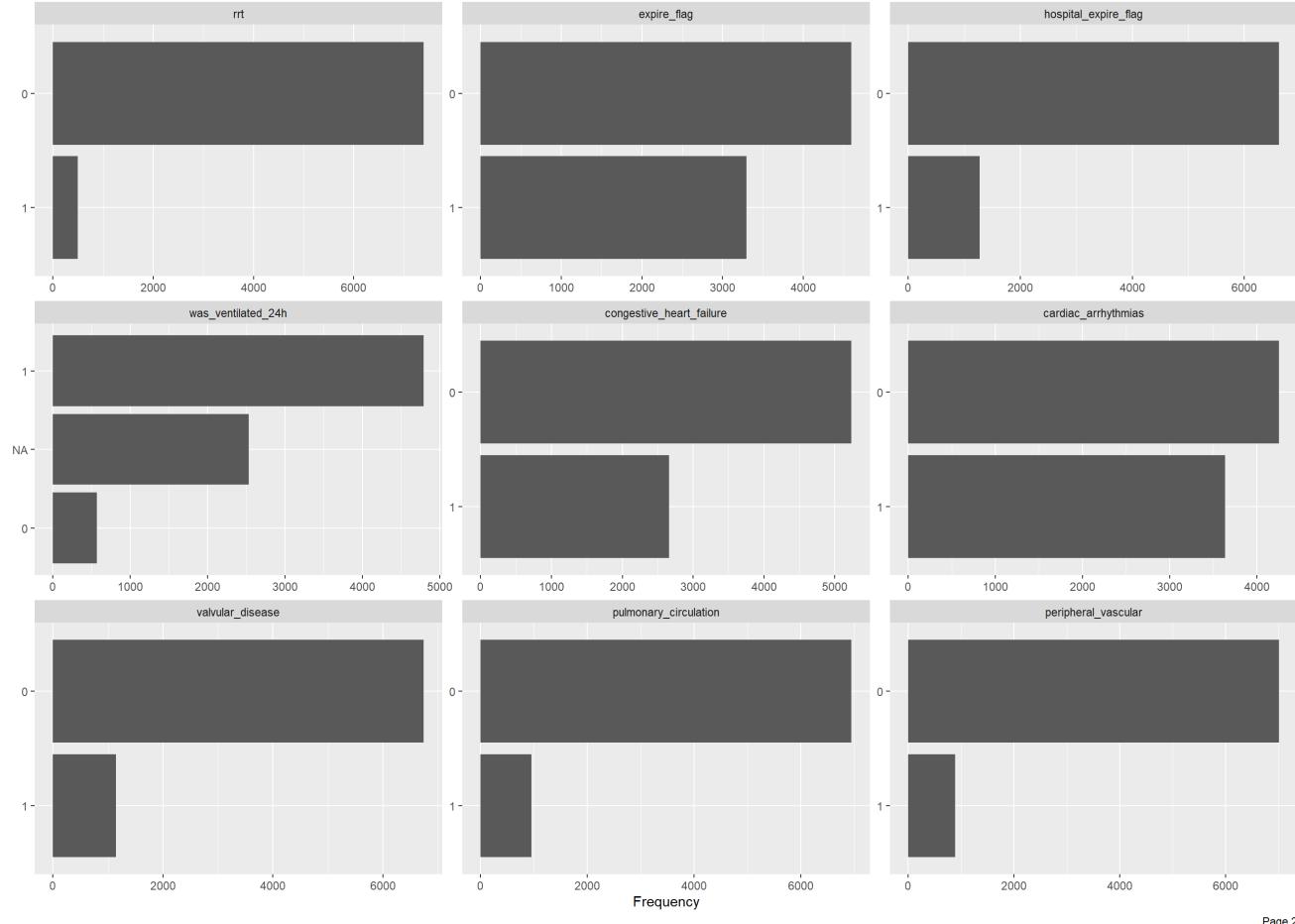
Bar Chart (by frequency)

```
## 10 columns ignored with more than 50 categories.
## intime: 7892 categories
## outtime: 7892 categories
## dob: 6149 categories
## dod: 2457 categories
## dod_hosp: 1908 categories
## admittime: 7442 categories
## dischtime: 7446 categories
## deathtime: 1140 categories
## diagnosis: 2848 categories
## max_kdigo_time: 7846 categories
```


Data Profiling Report

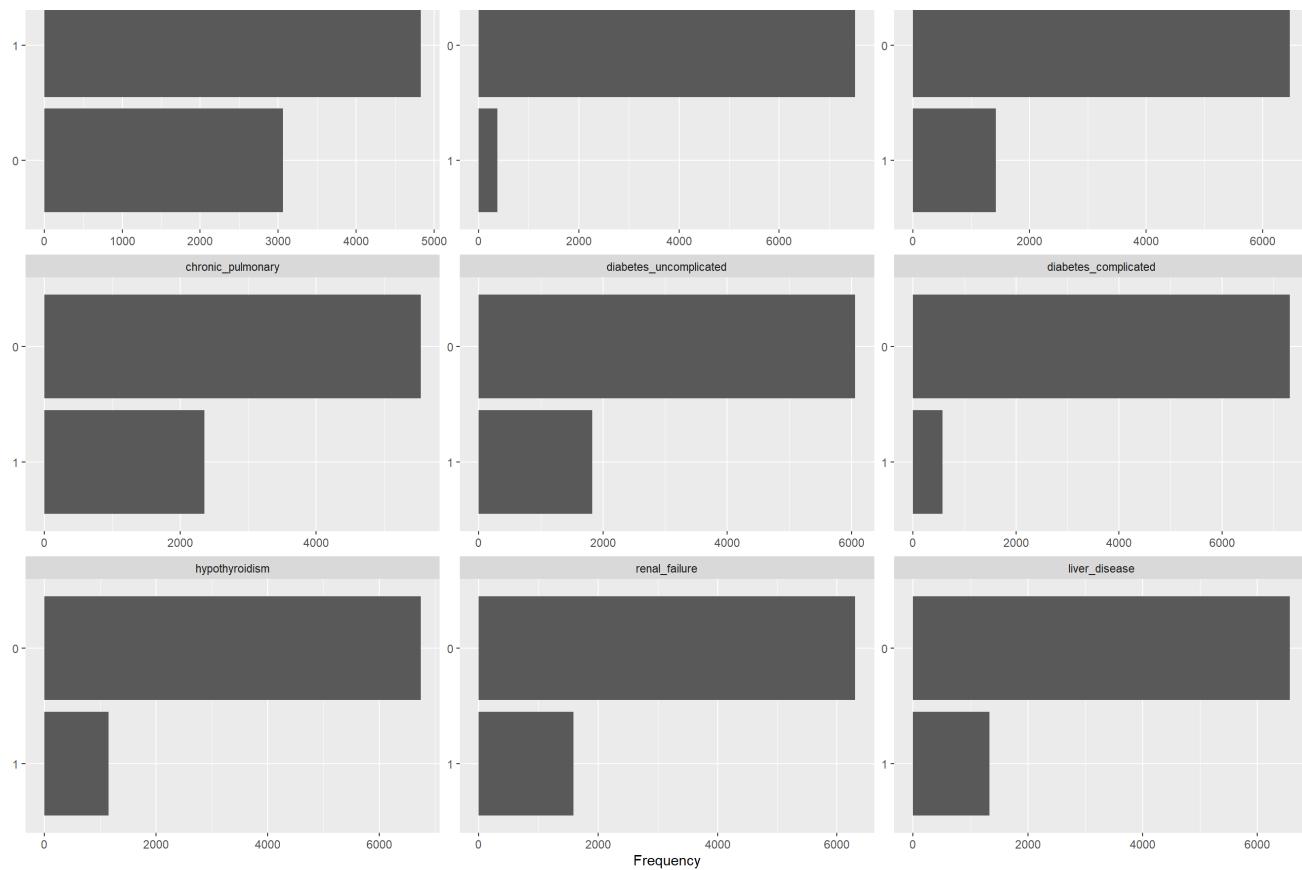


Page 1



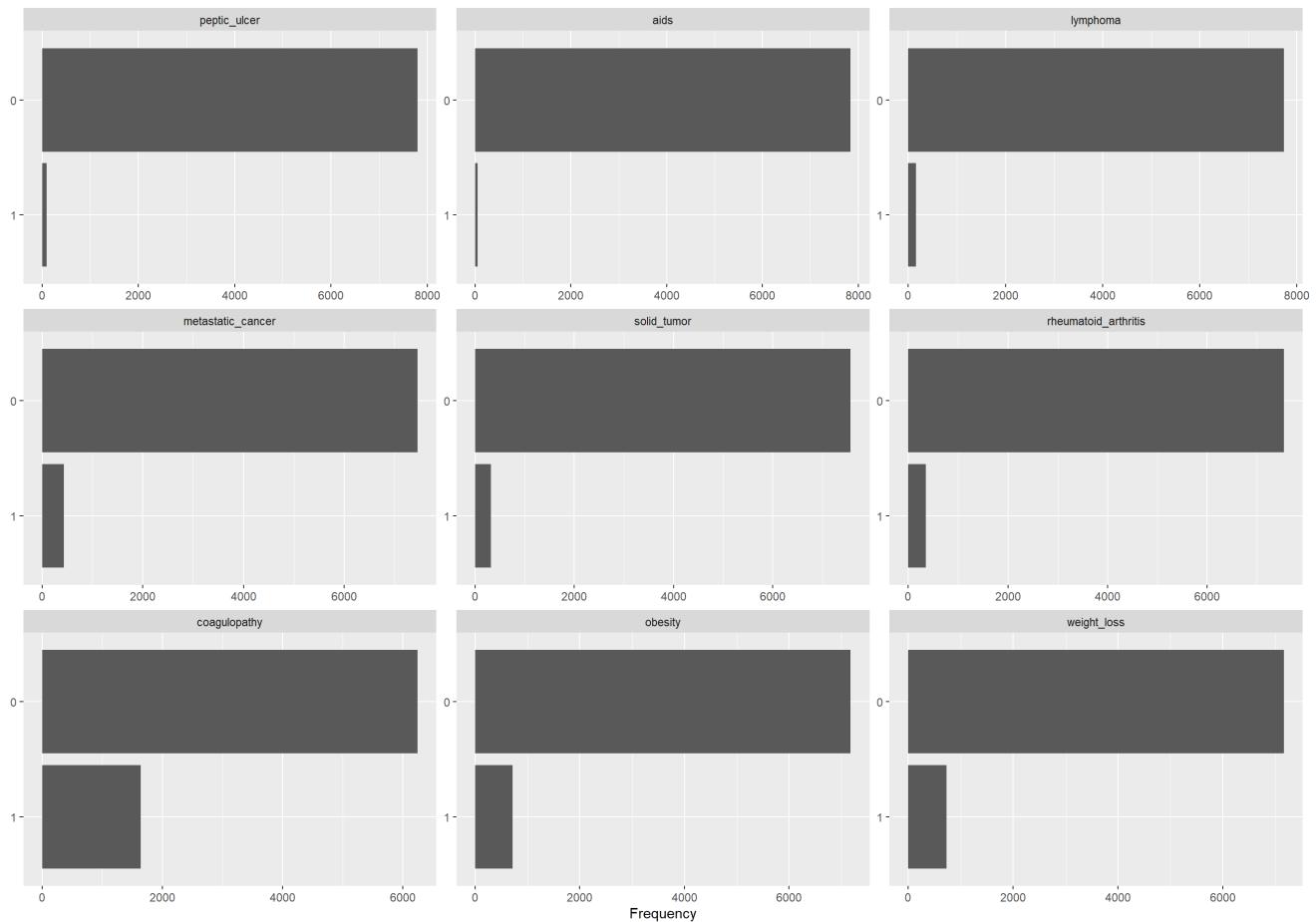
Page 2

Data Profiling Report

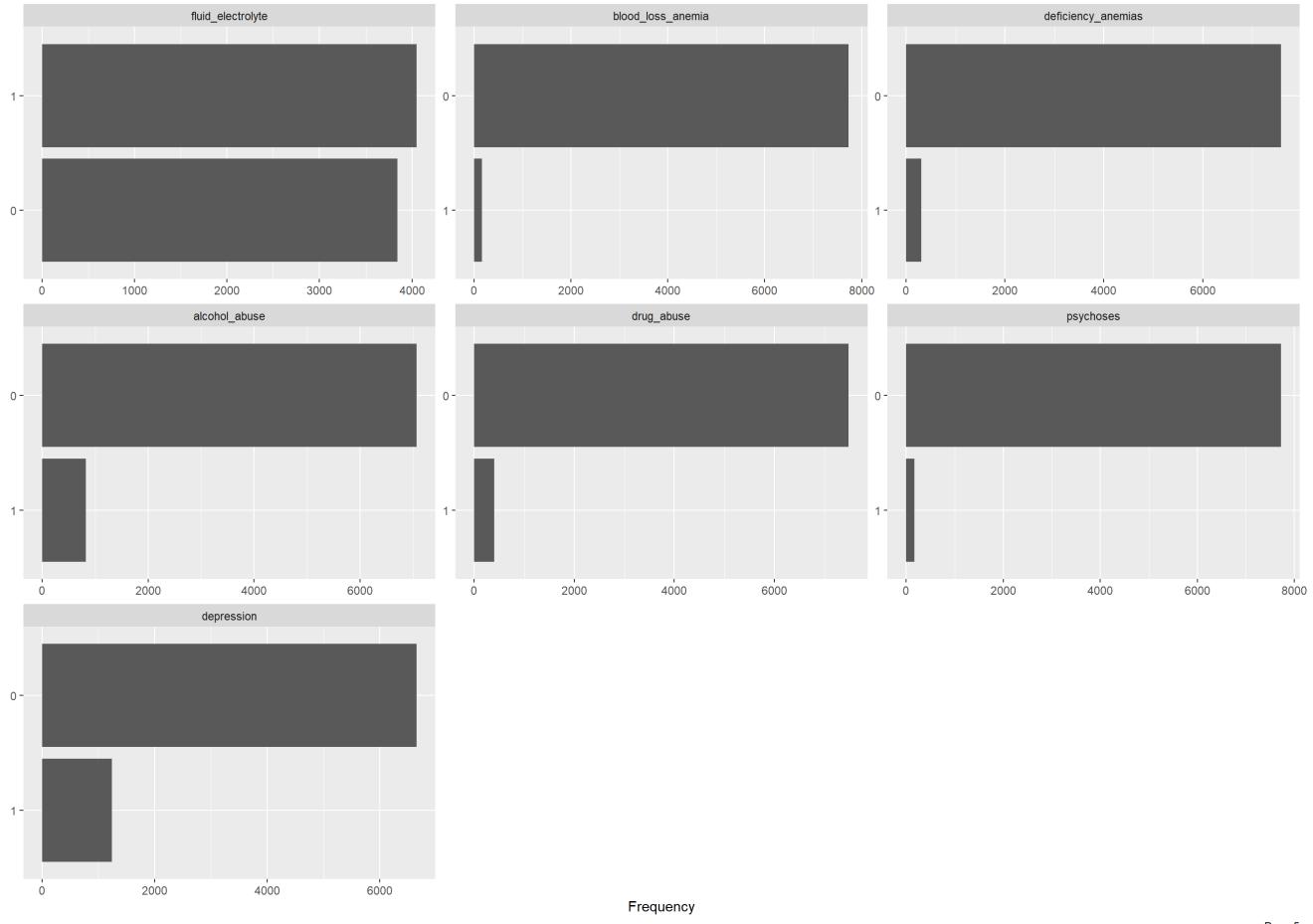


Page 3

Data Profiling Report



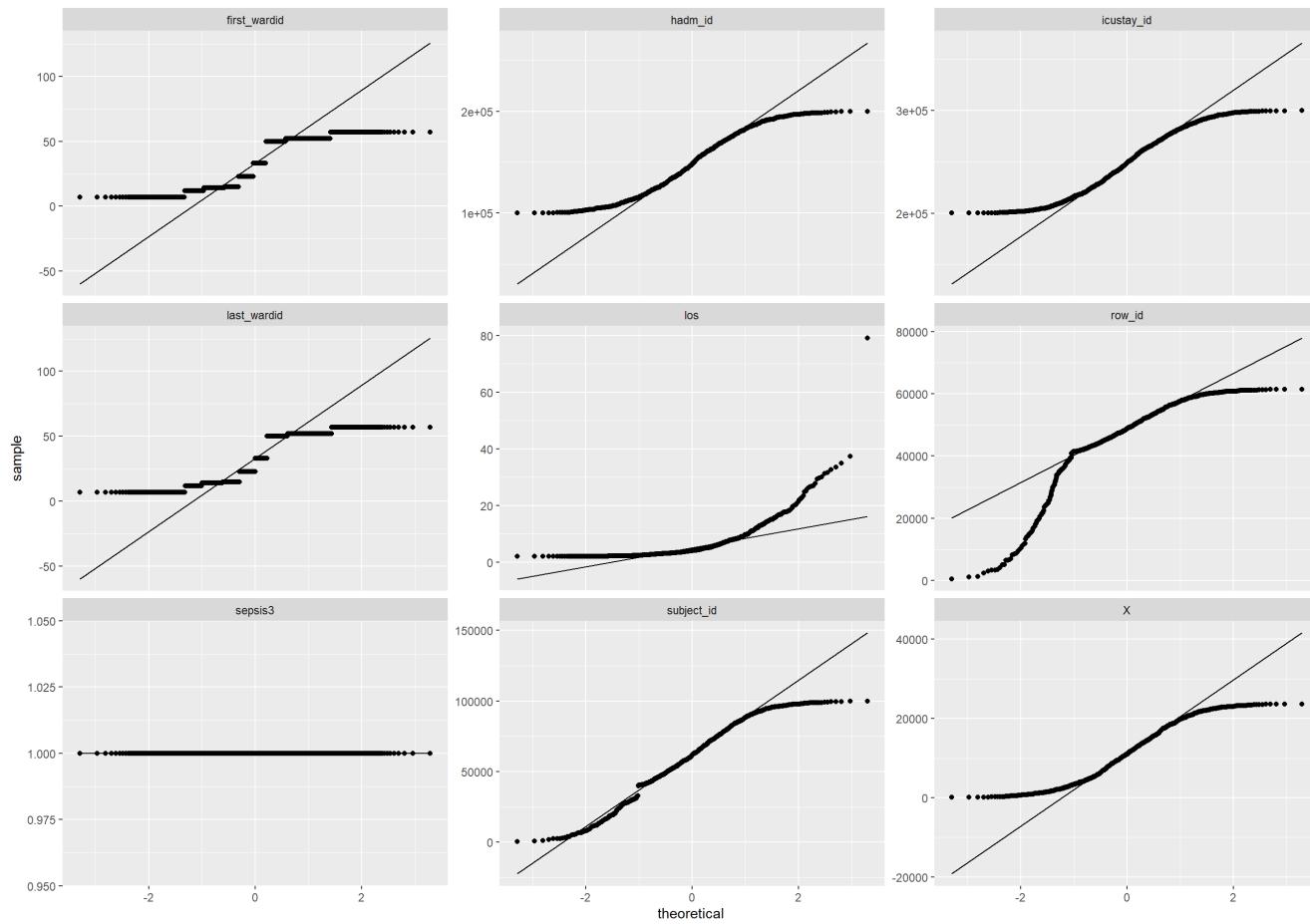
Page 4



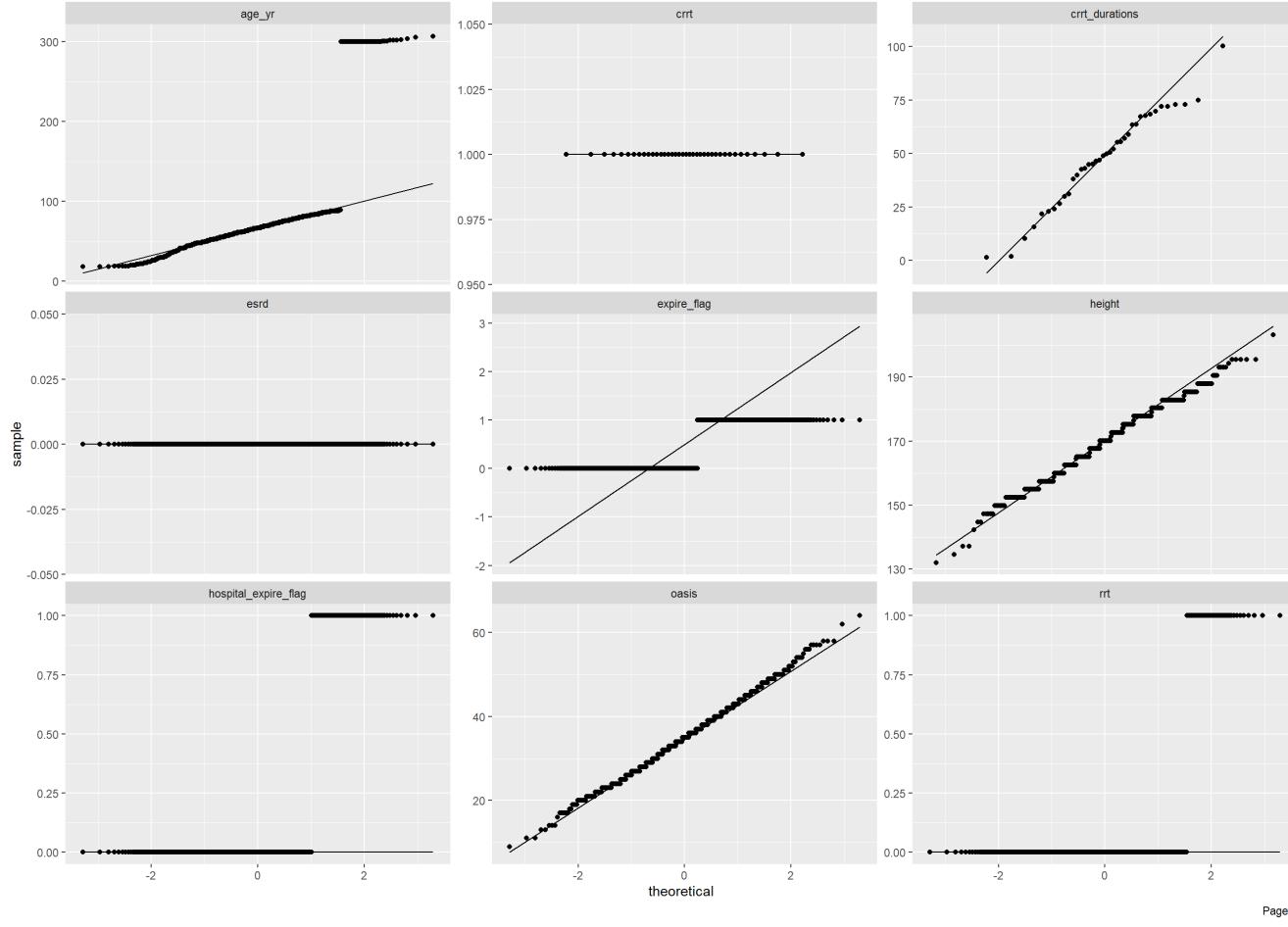
Page 5

QQ Plot

Data Profiling Report

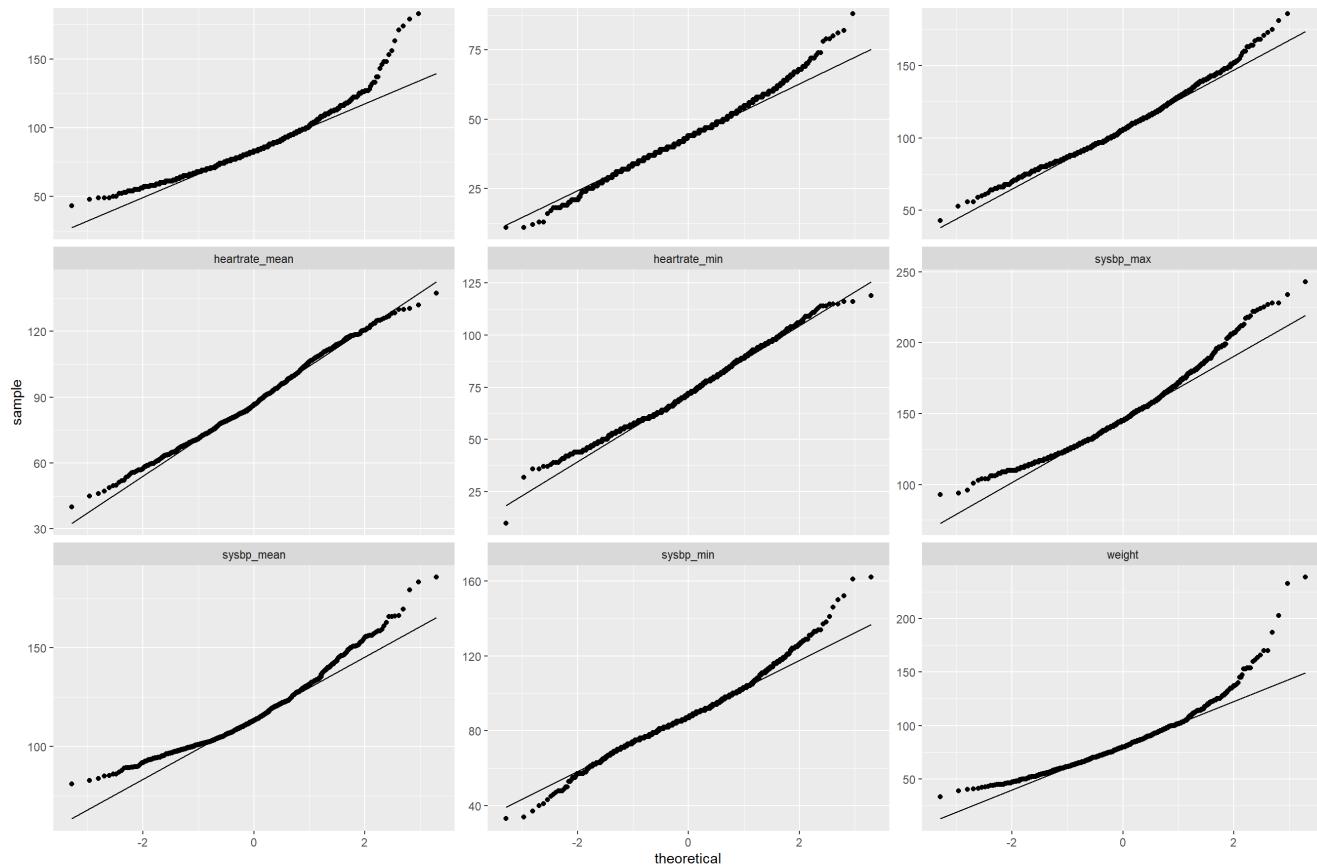


Page 1

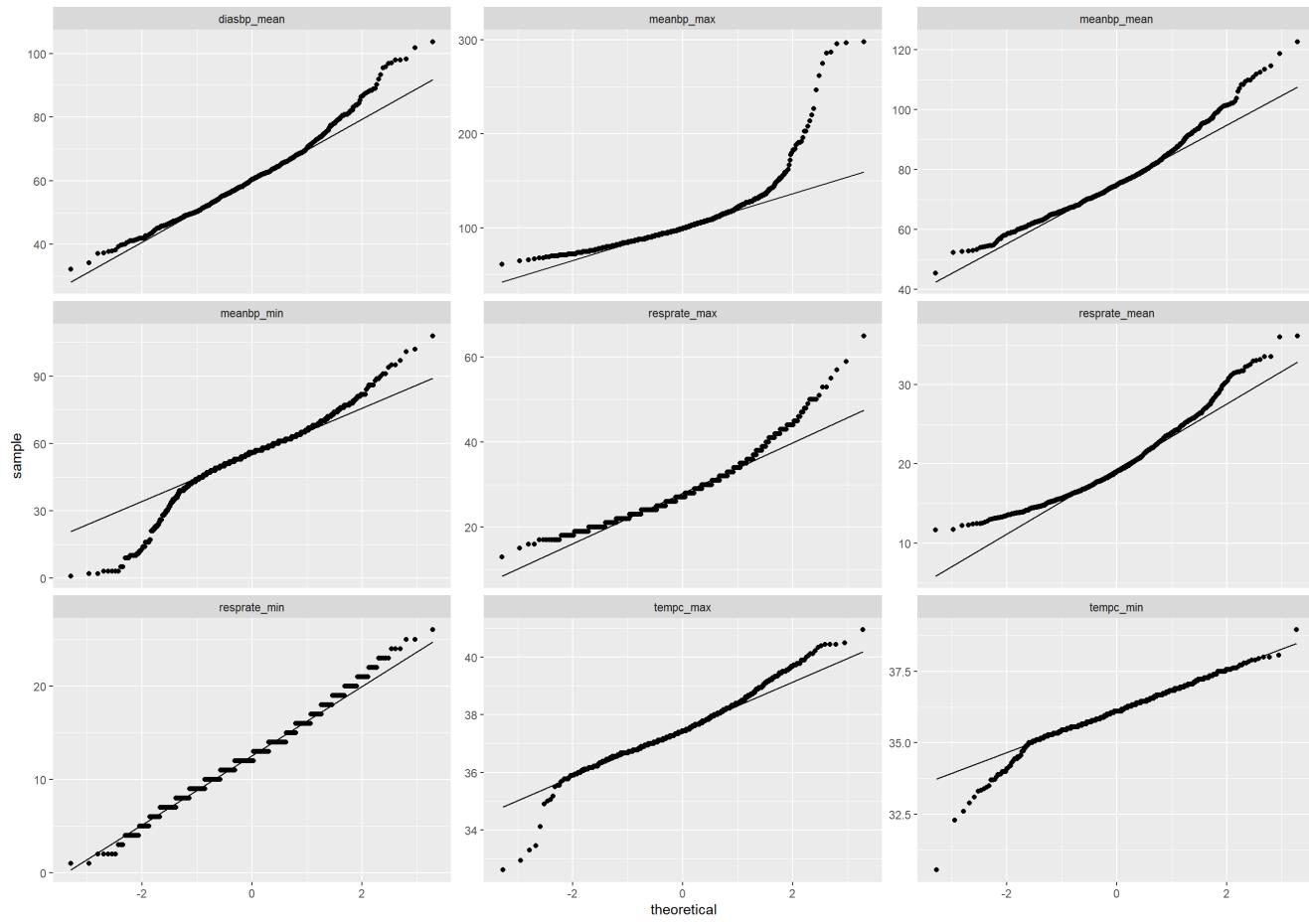


Page 2

Data Profiling Report

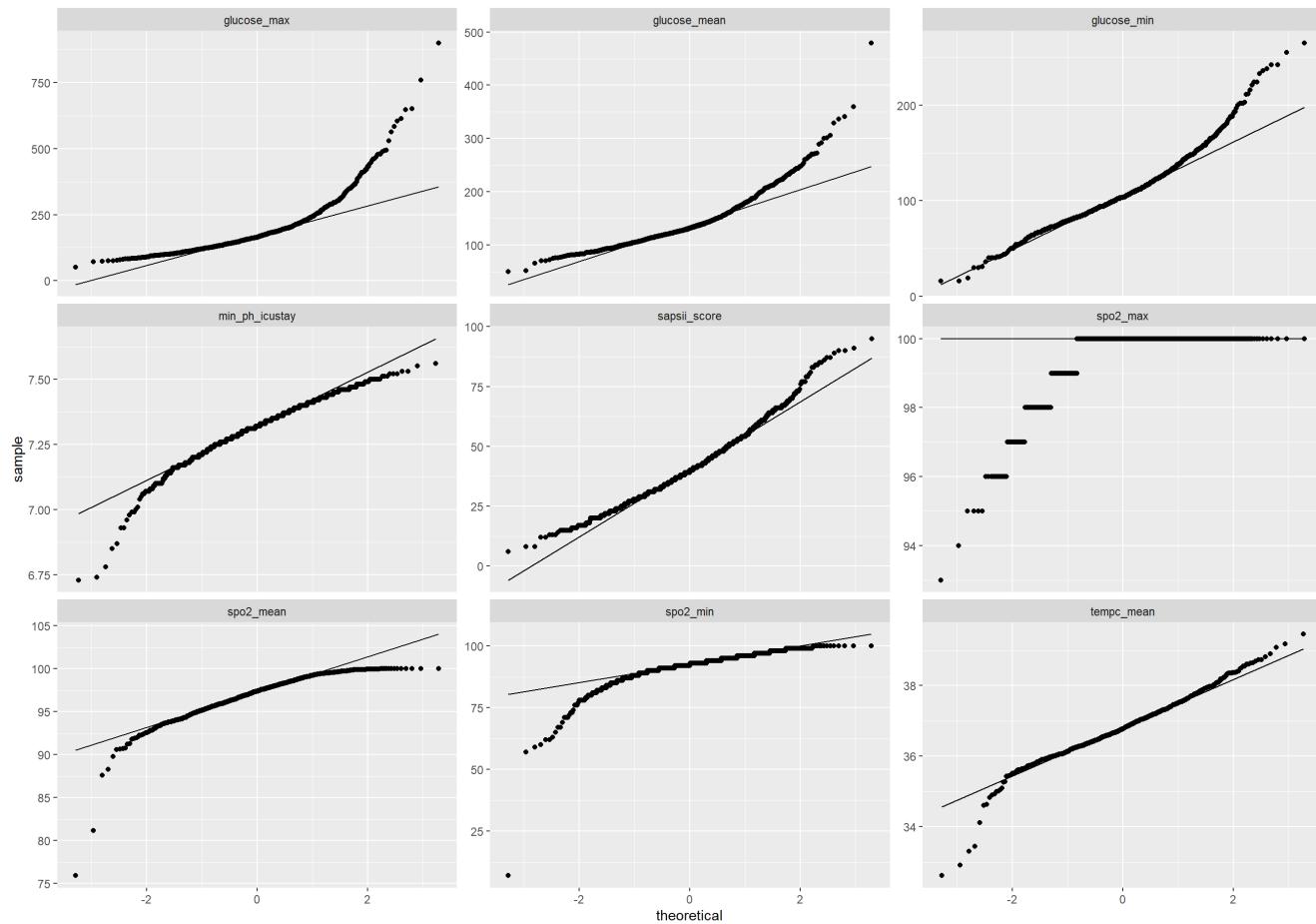


Page 3

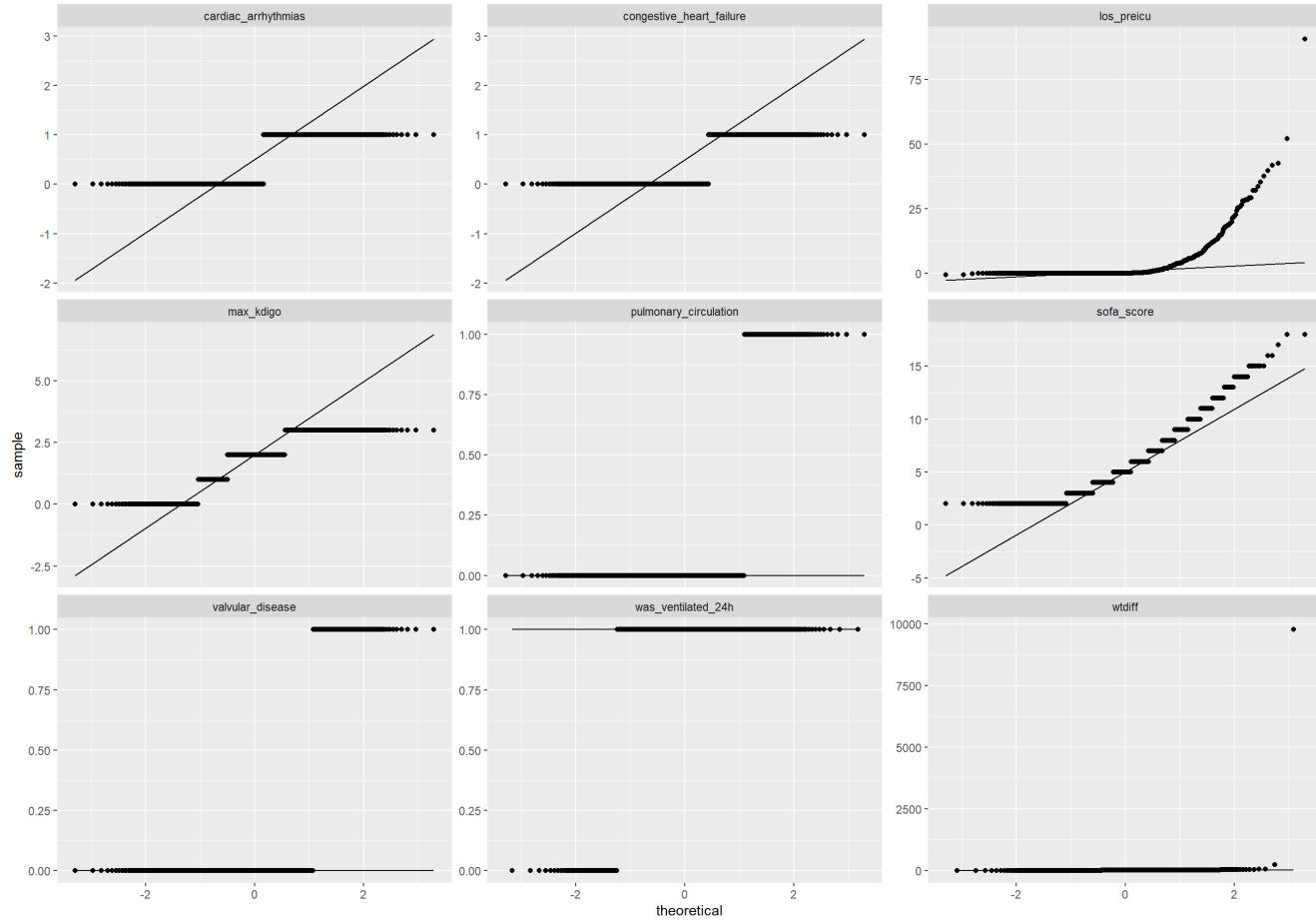


Page 4

Data Profiling Report

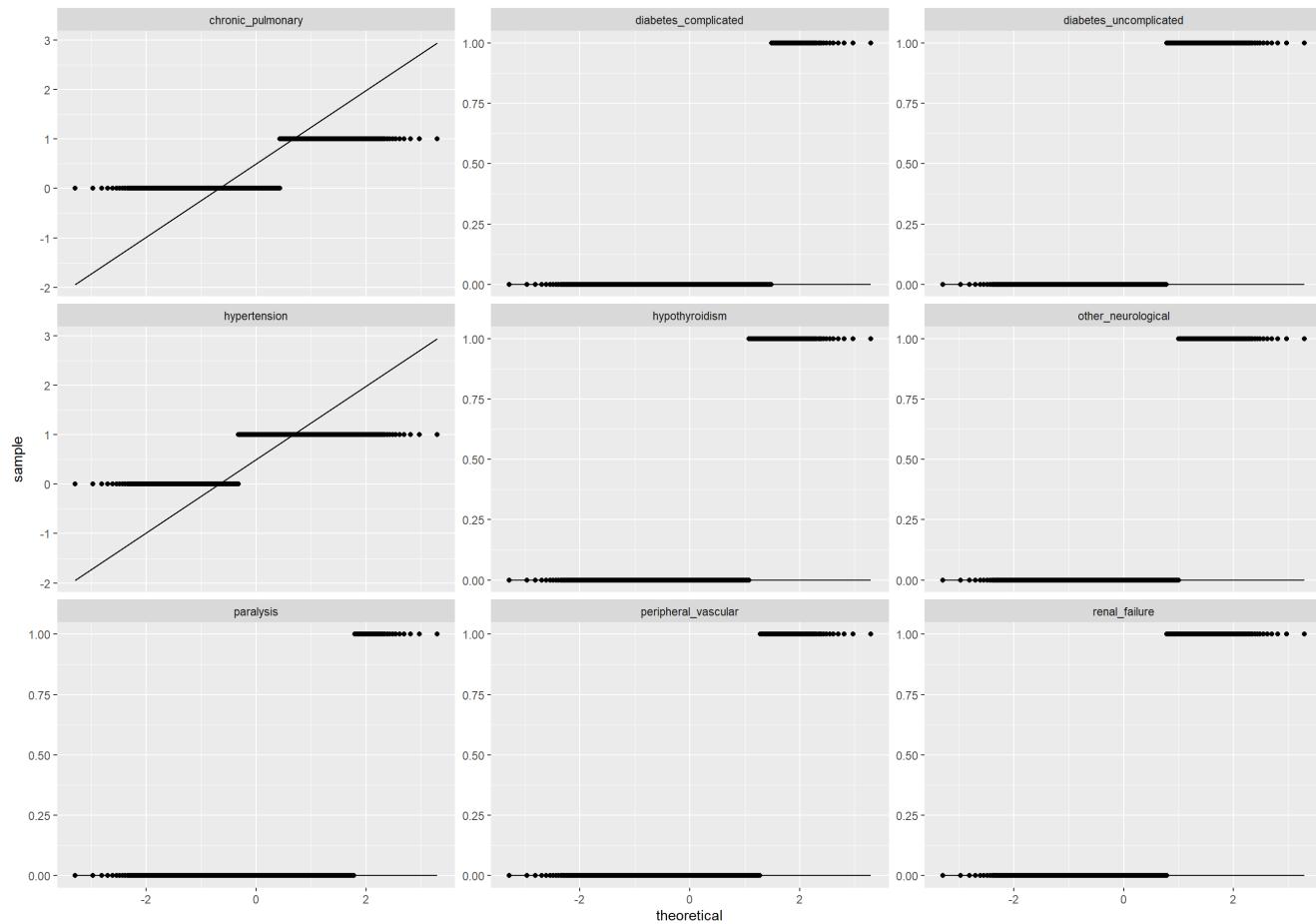


Page 5

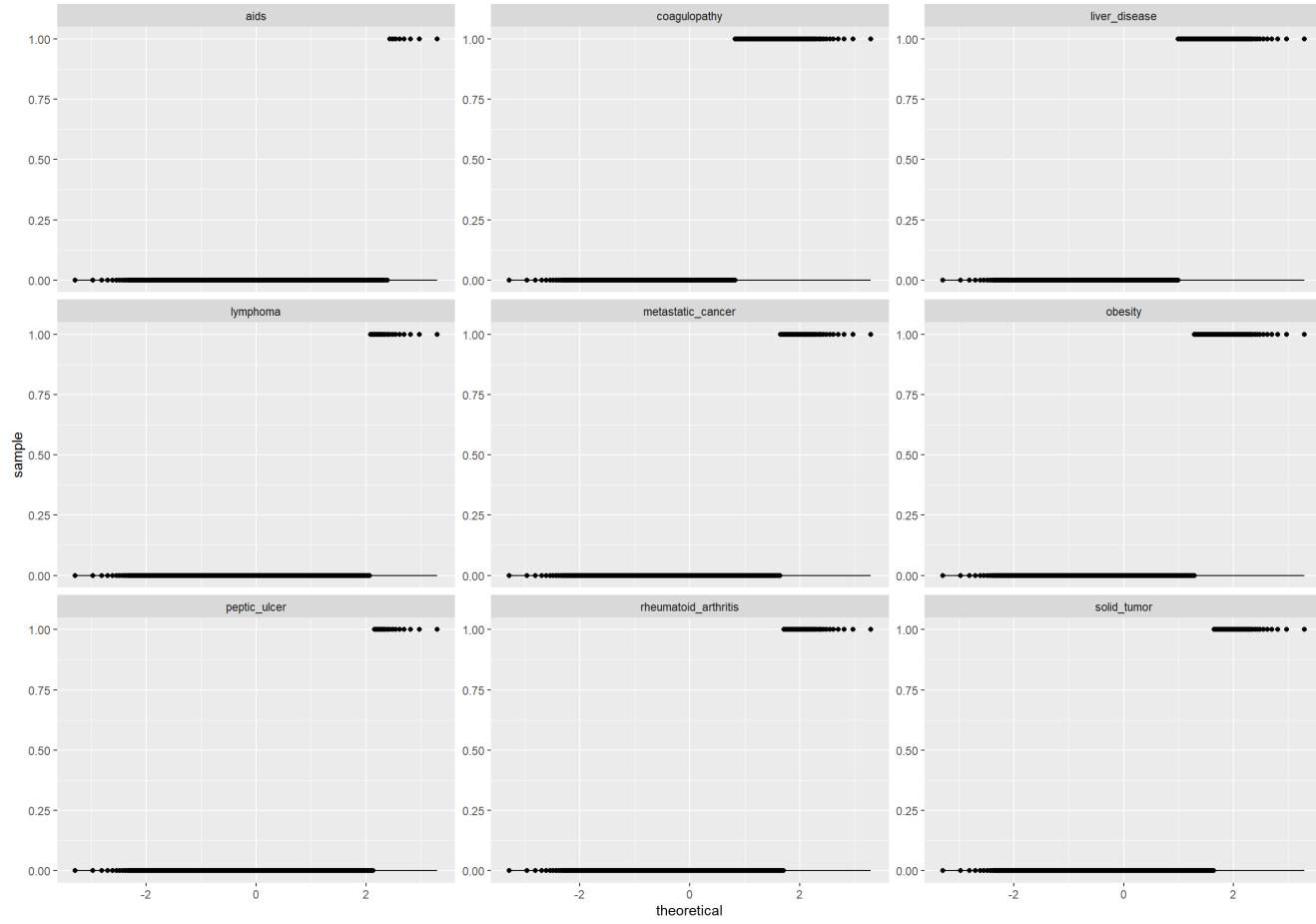


Page 6

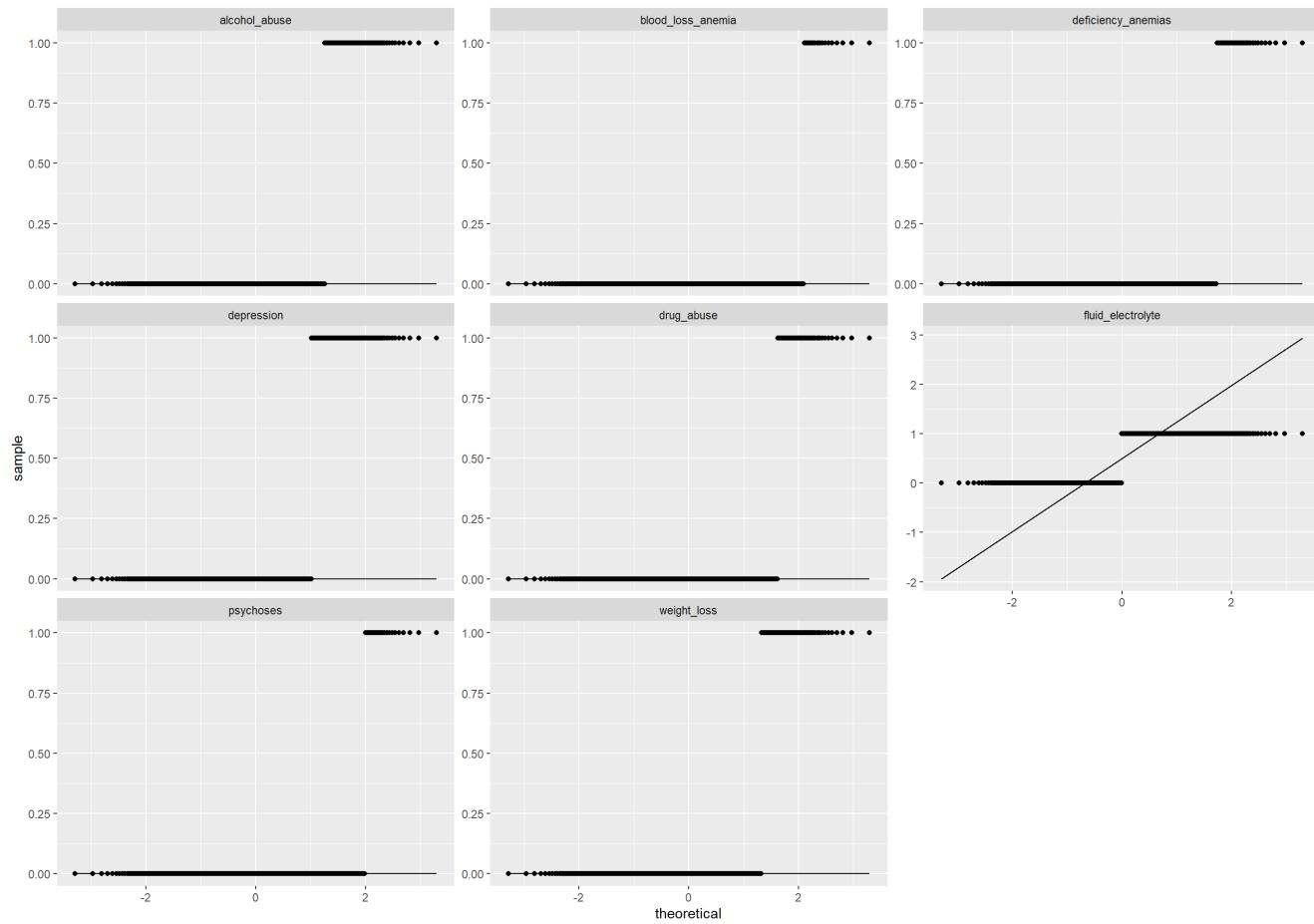
Data Profiling Report



Page 7



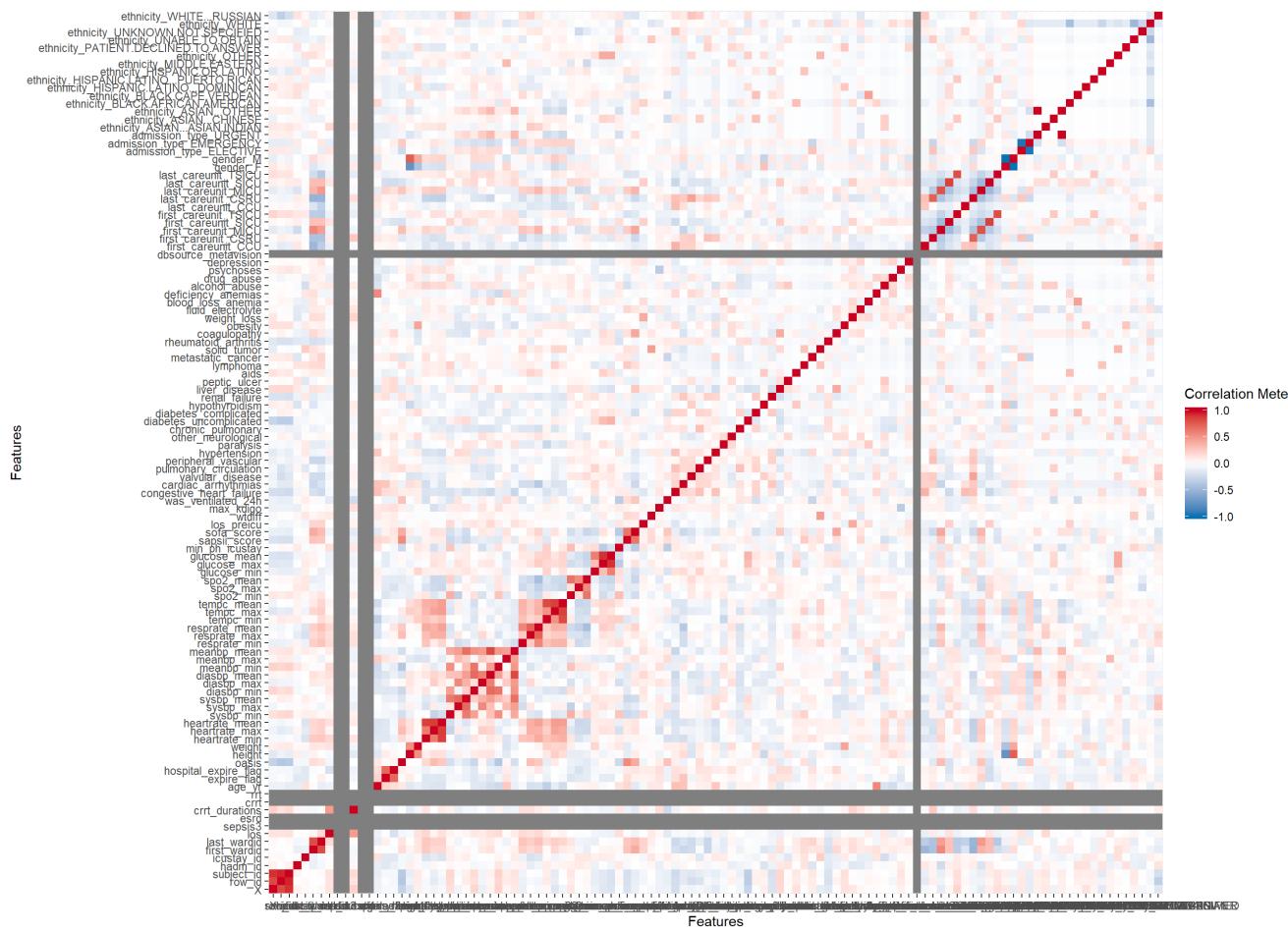
Page 8



Page 9

Correlation Analysis

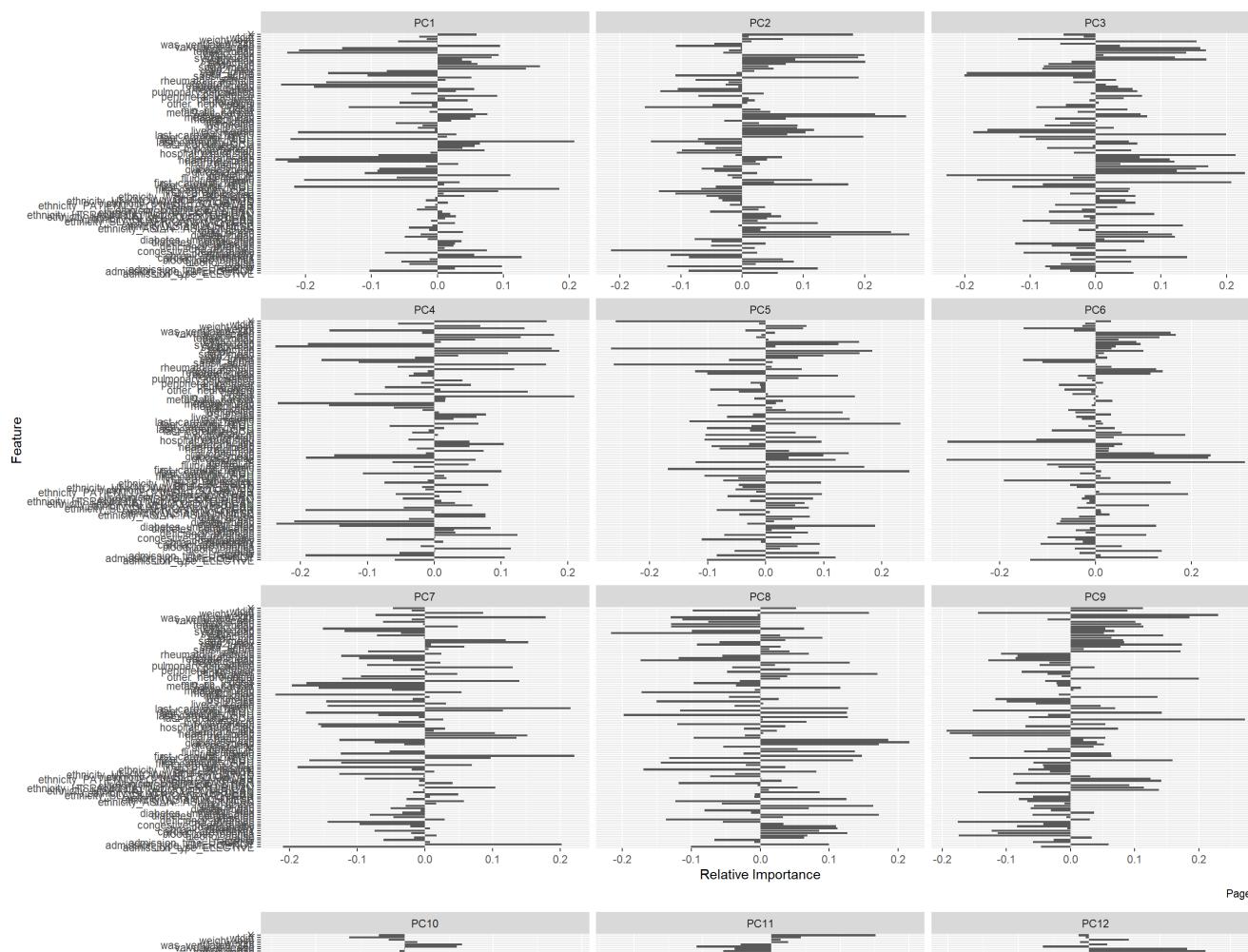
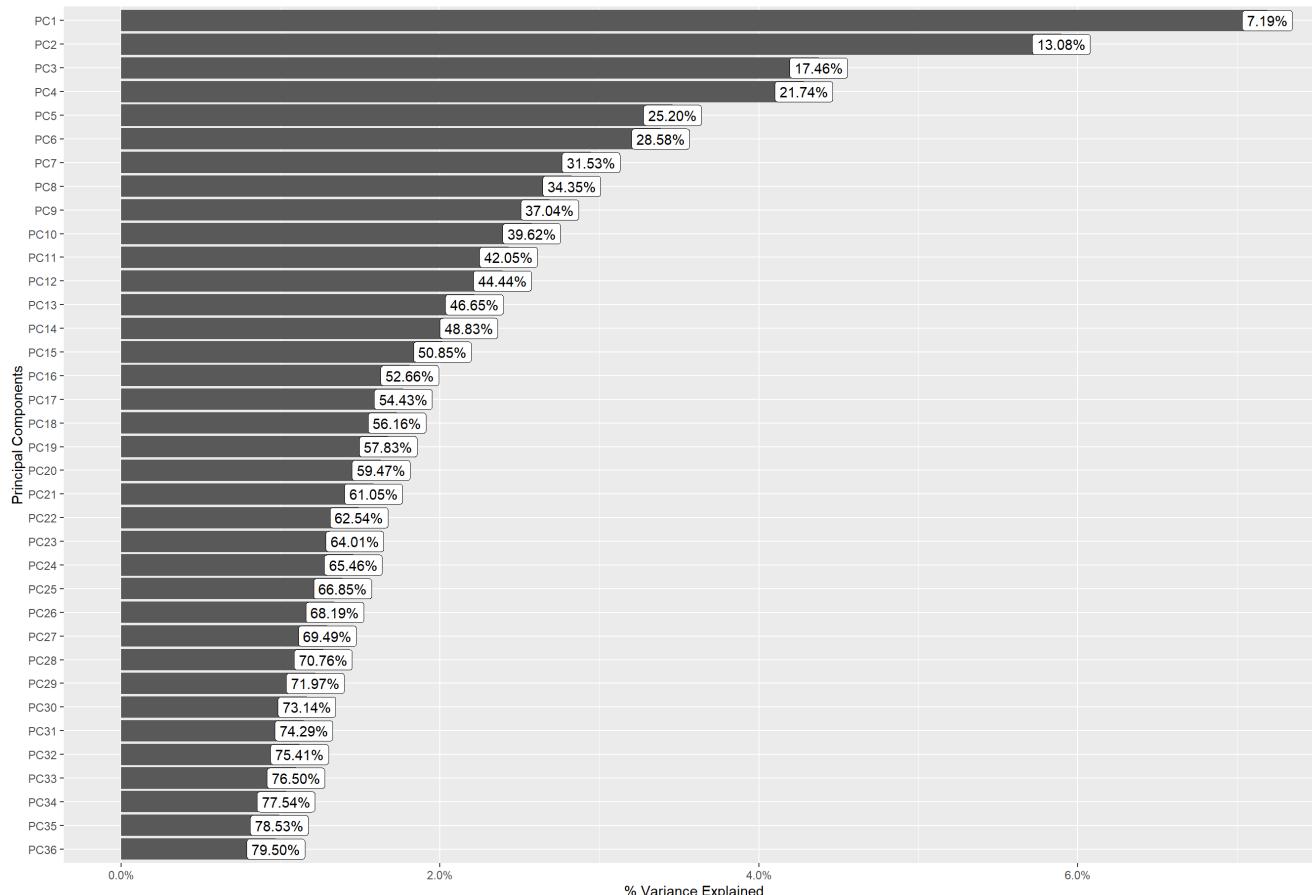
```
## 10 features with more than 20 categories ignored!
## intime: 140 categories
## outtime: 140 categories
## dob: 140 categories
## dod: 87 categories
## dod_hosp: 78 categories
## admittime: 140 categories
## dischtime: 140 categories
## deathtime: 63 categories
## diagnosis: 111 categories
## max_kdigo_time: 140 categories
```



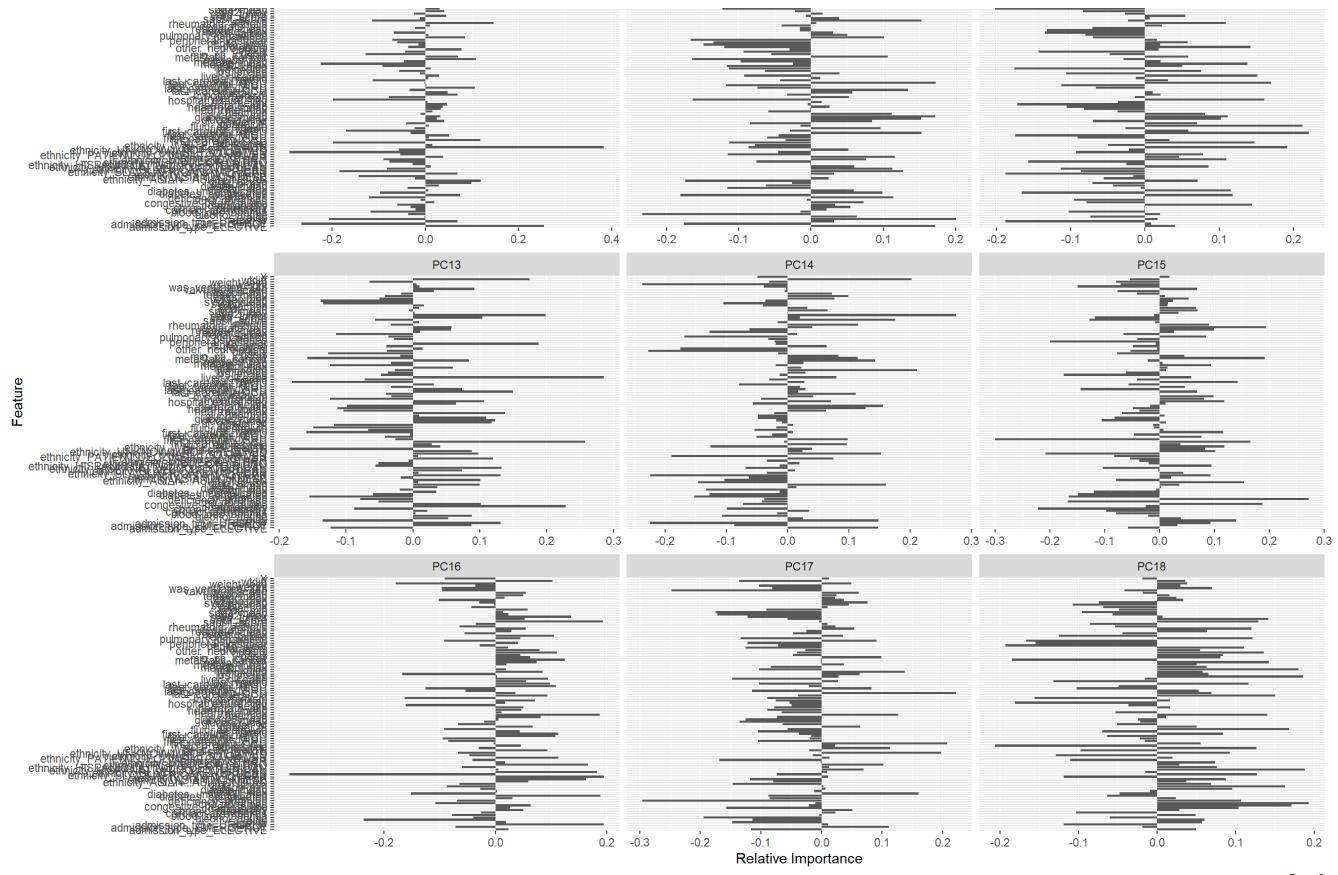
Principal Component Analysis

```
## 10 features with more than 50 categories ignored!
## intime: 140 categories
## outtime: 140 categories
## dob: 140 categories
## dod: 87 categories
## dod_hosp: 78 categories
## admittime: 140 categories
## dischtime: 140 categories
## deathtime: 63 categories
## diagnosis: 111 categories
## max_kdigo_time: 140 categories
```

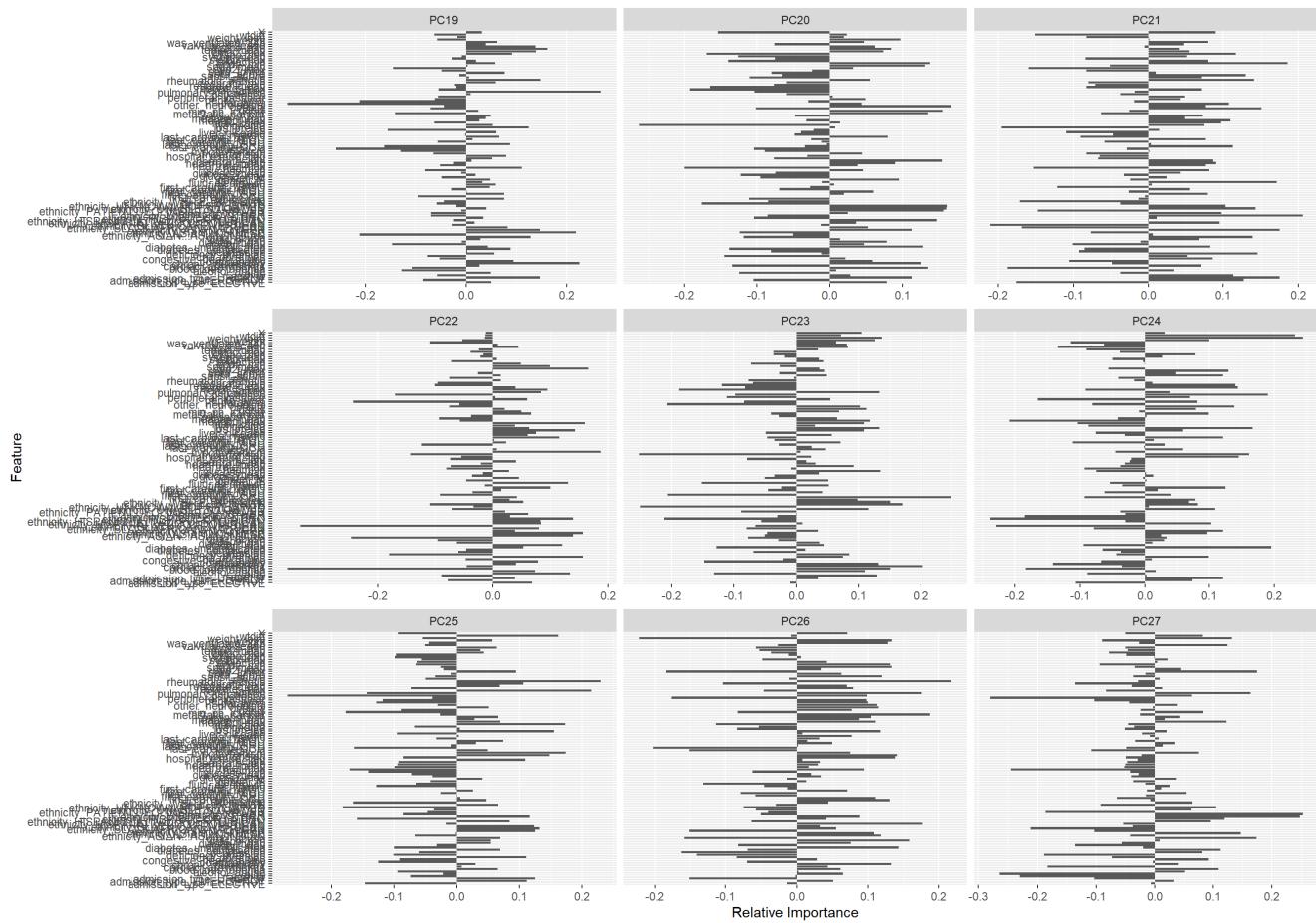

% Variance Explained By Principal Components
 (Note: Labels indicate cumulative % explained variance)



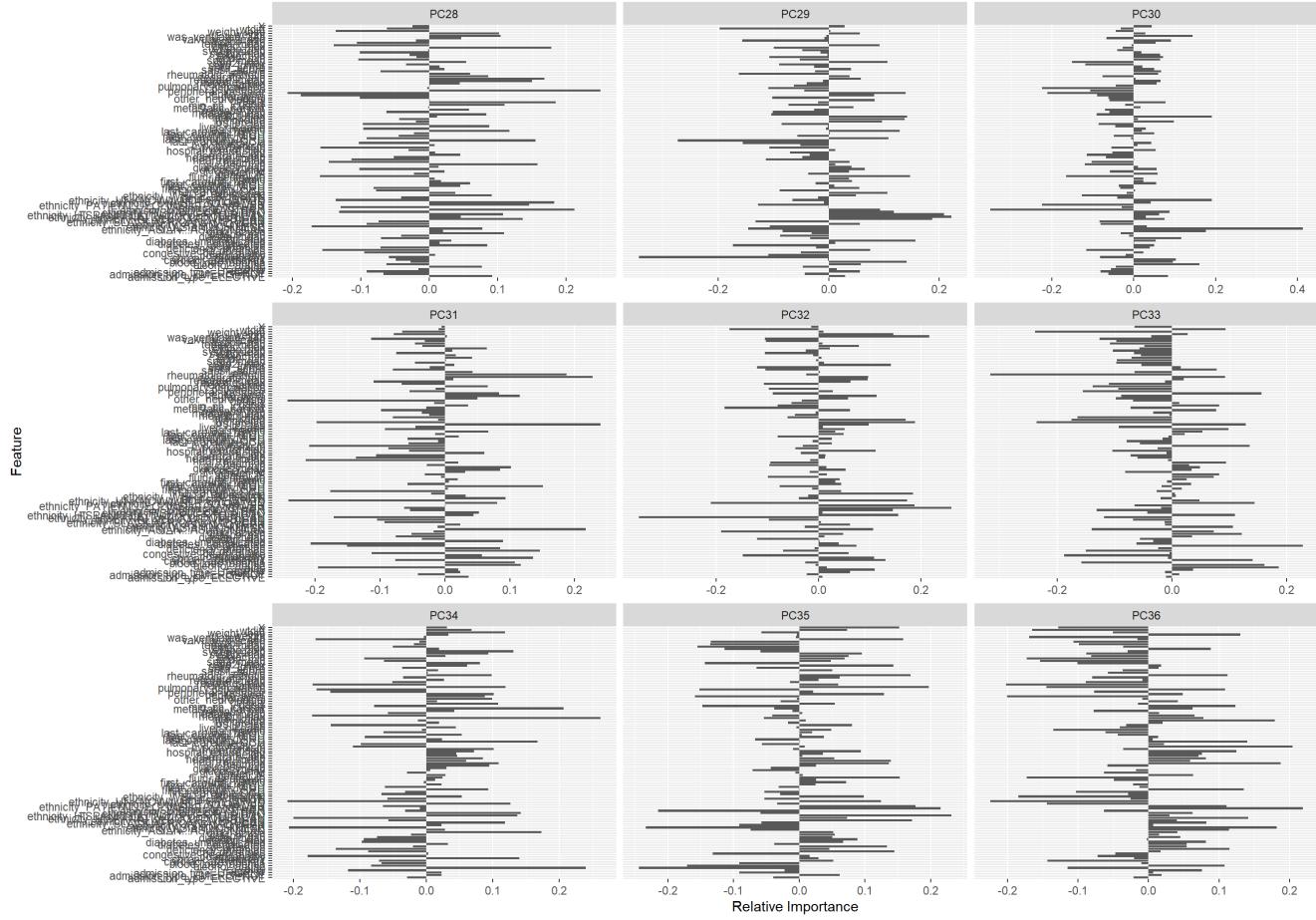
Data Profiling Report



Page 2



Page 3



Page 4