

Data Profiling Report

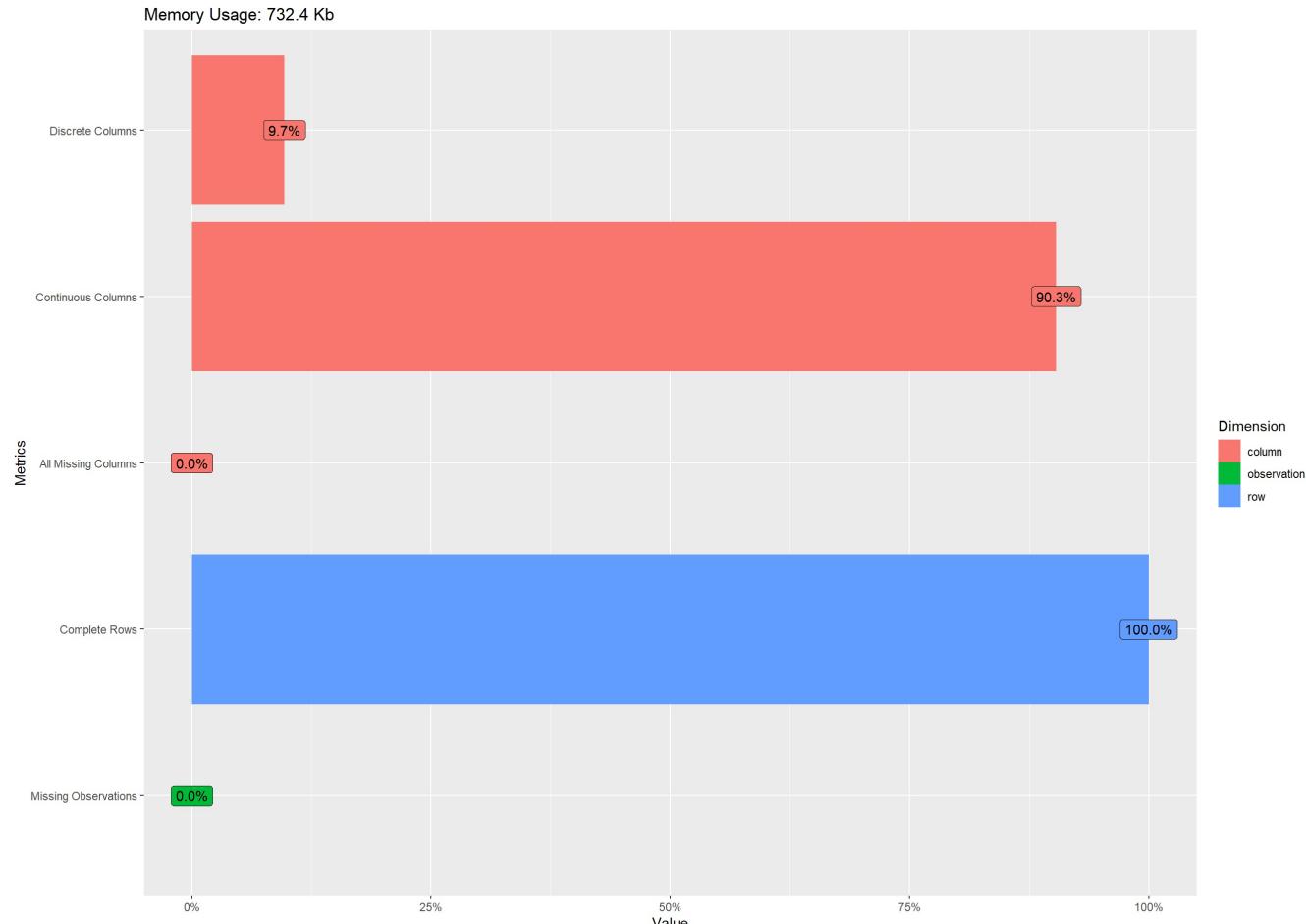
- Basic Statistics
 - Raw Counts
 - Percentages
- Data Structure
- Missing Data Profile
- Univariate Distribution
 - Histogram
 - Bar Chart (with frequency)
 - QQ Plot
- Correlation Analysis
- Principal Component Analysis

Basic Statistics

Raw Counts

Name	Value
Rows	4,971
Columns	31
Discrete columns	3
Continuous columns	28
All missing columns	0
Missing observations	0
Complete Rows	4,971
Total observations	154,101
Memory allocation	732.4 Kb

Percentages

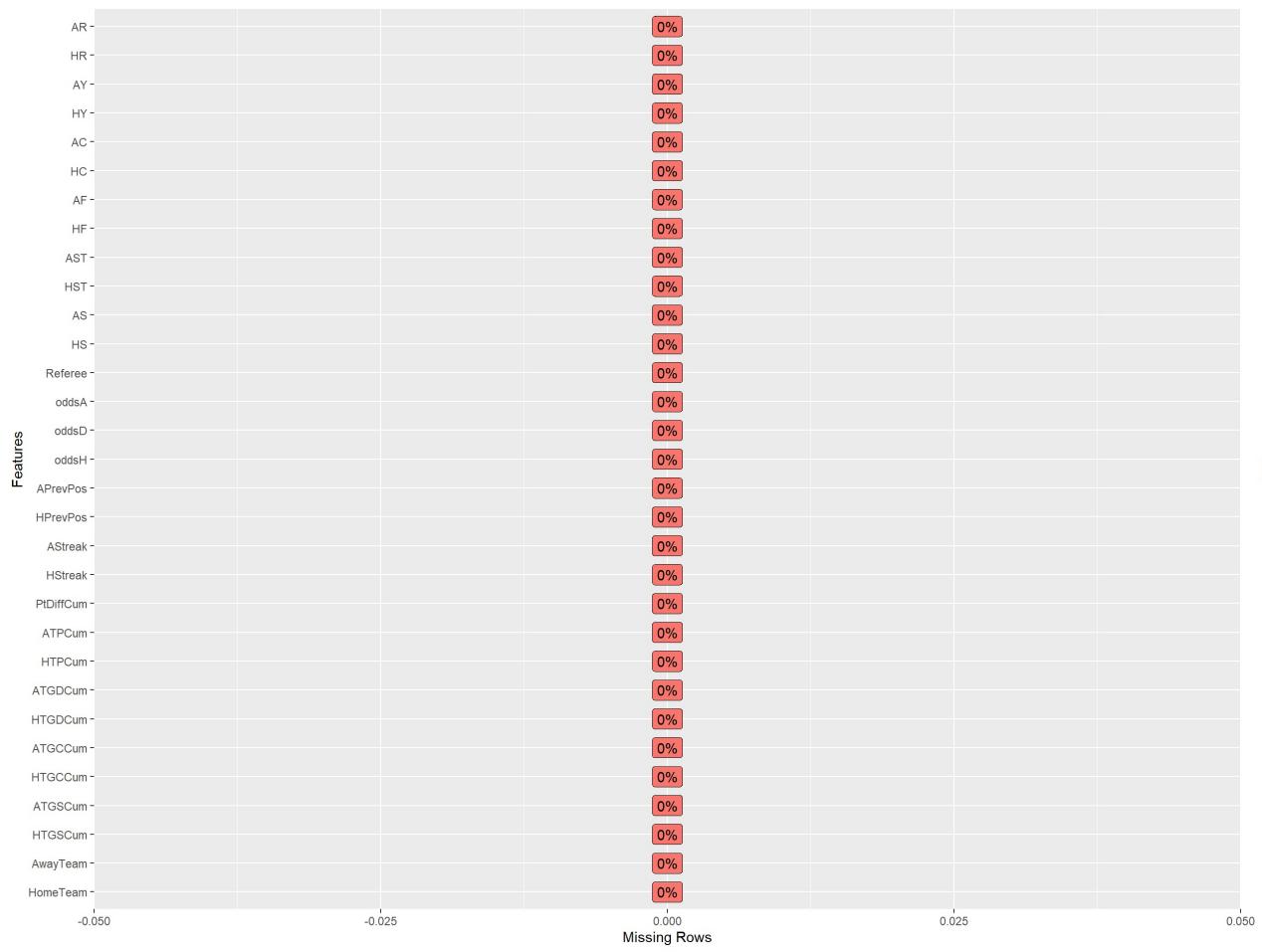


Data Structure

root (Classes 'data.table' and 'data.frame': 4971 obs. of 31 variables):=

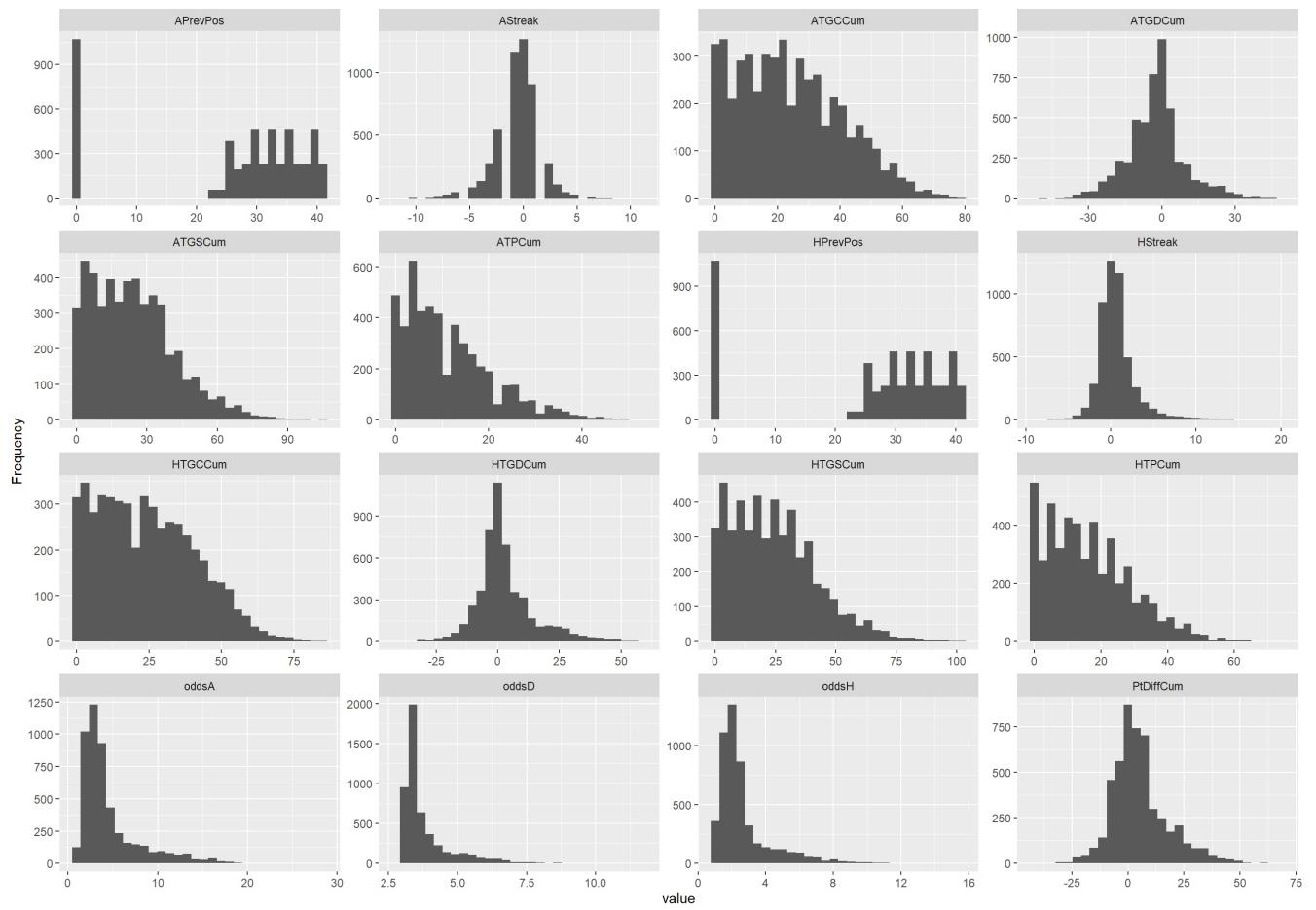
- HomeTeam (chr)
- AwayTeam (chr)
- HTGSCum (int)
- ATGSCum (int)
- HTGCCum (int)
- ATGCCum (int)
- HTGDCum (int)
- ATGDCum (int)
- HTPCum (int)
- ATPCum (int)
- PtDiffCum (int)
- HStreak (int)
- AS streak (int)
- HPrevPos (int)
- APrevPos (int)
- oddsH (num)
- oddsD (num)
- oddsA (num)
- Referee (chr)
- HS (int)
- AS (int)
- HST (int)
- AST (int)
- HF (int)
- AF (int)
- HC (int)
- AC (int)
- HY (int)
- AY (int)
- HR (int)
- AR (int)

Missing Data Profile

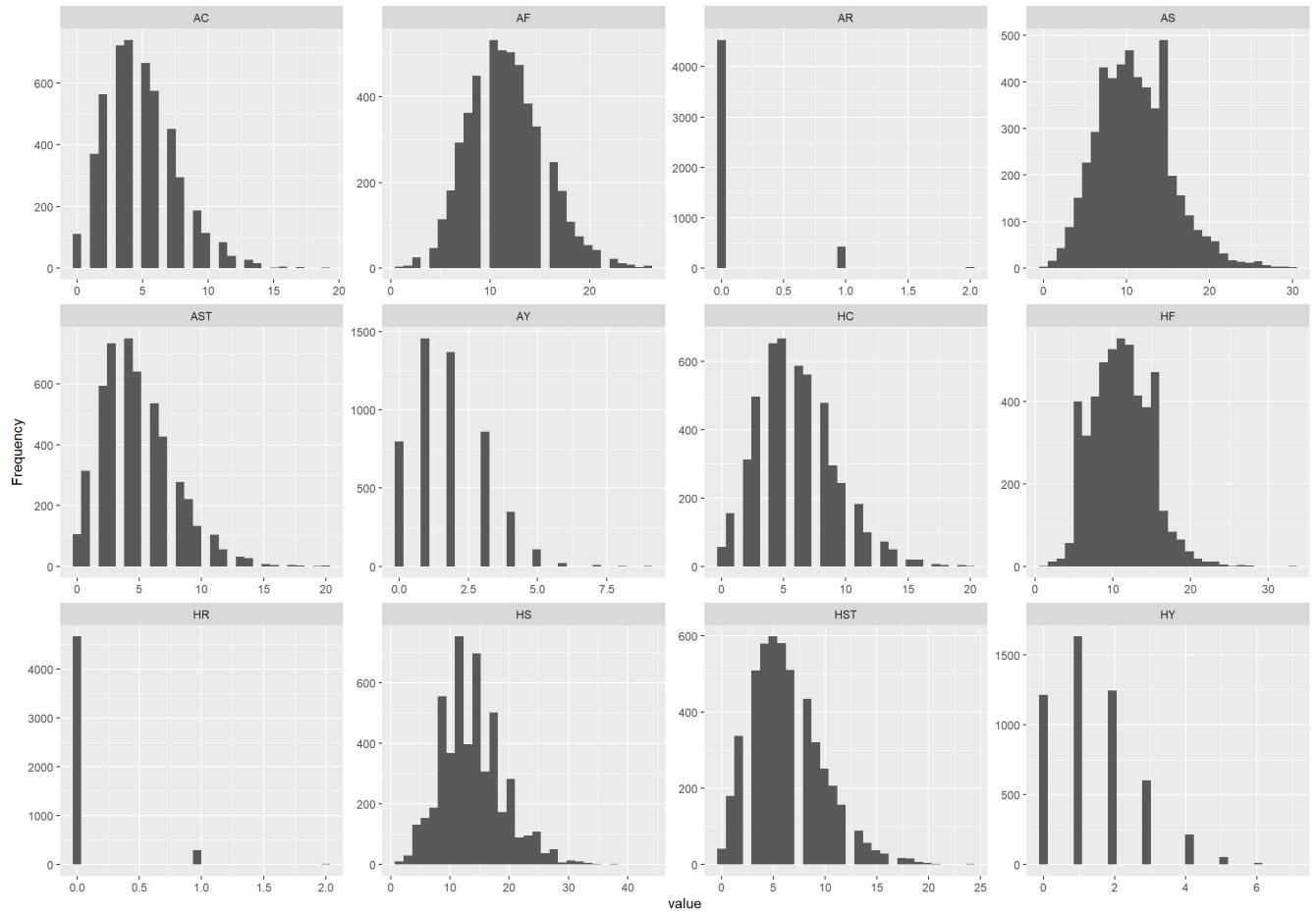


Univariate Distribution

Histogram

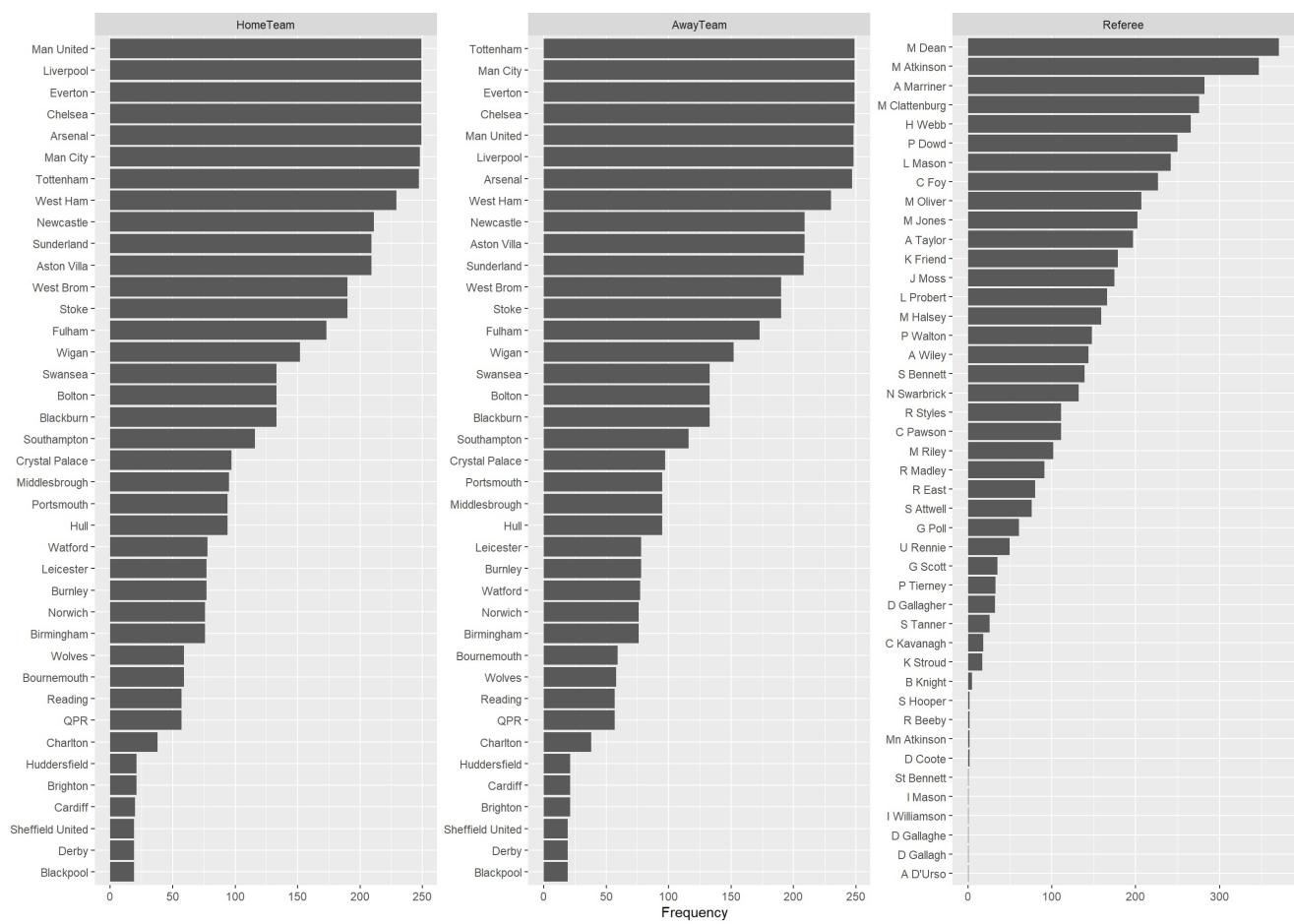


Page 1

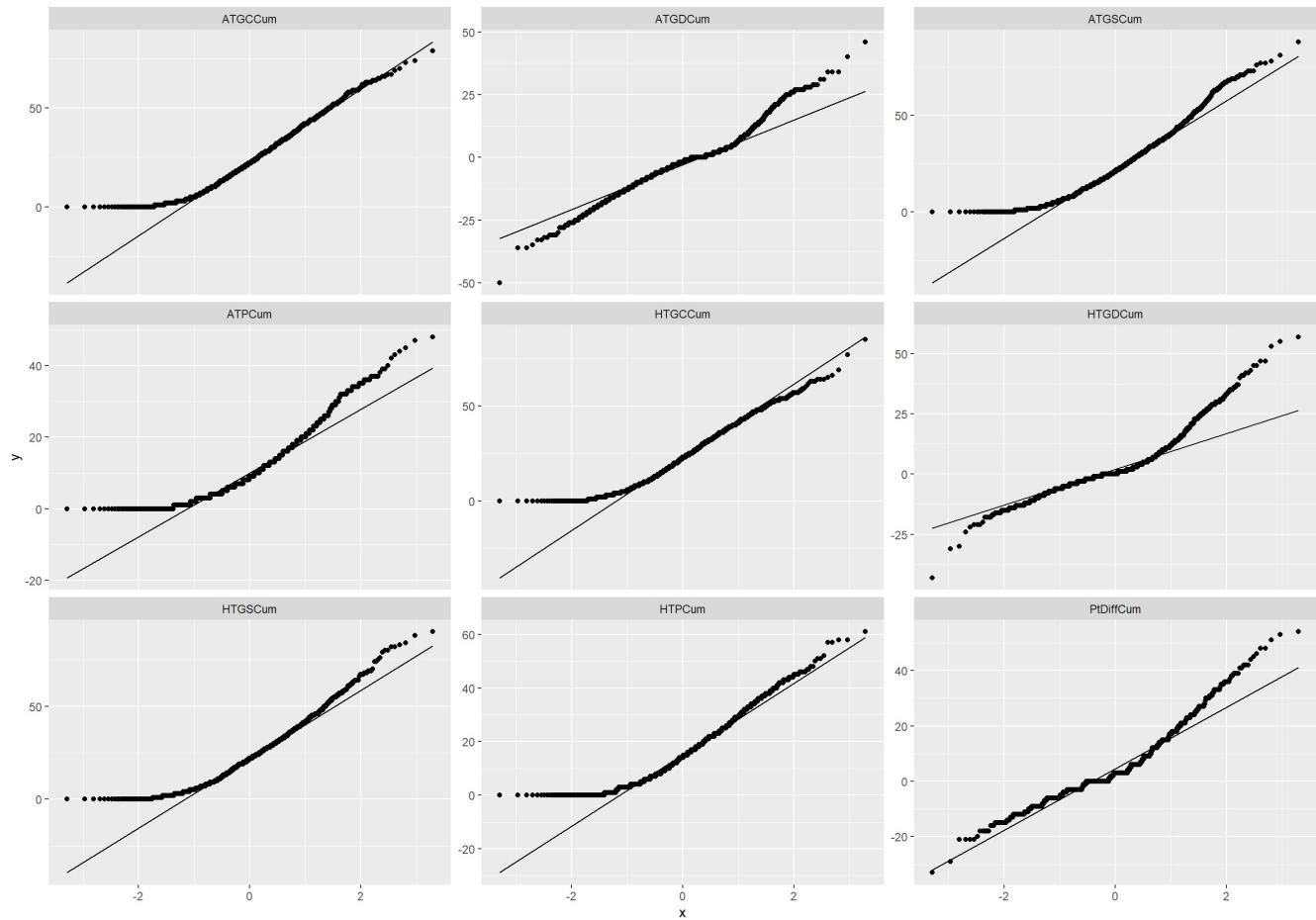


Page 2

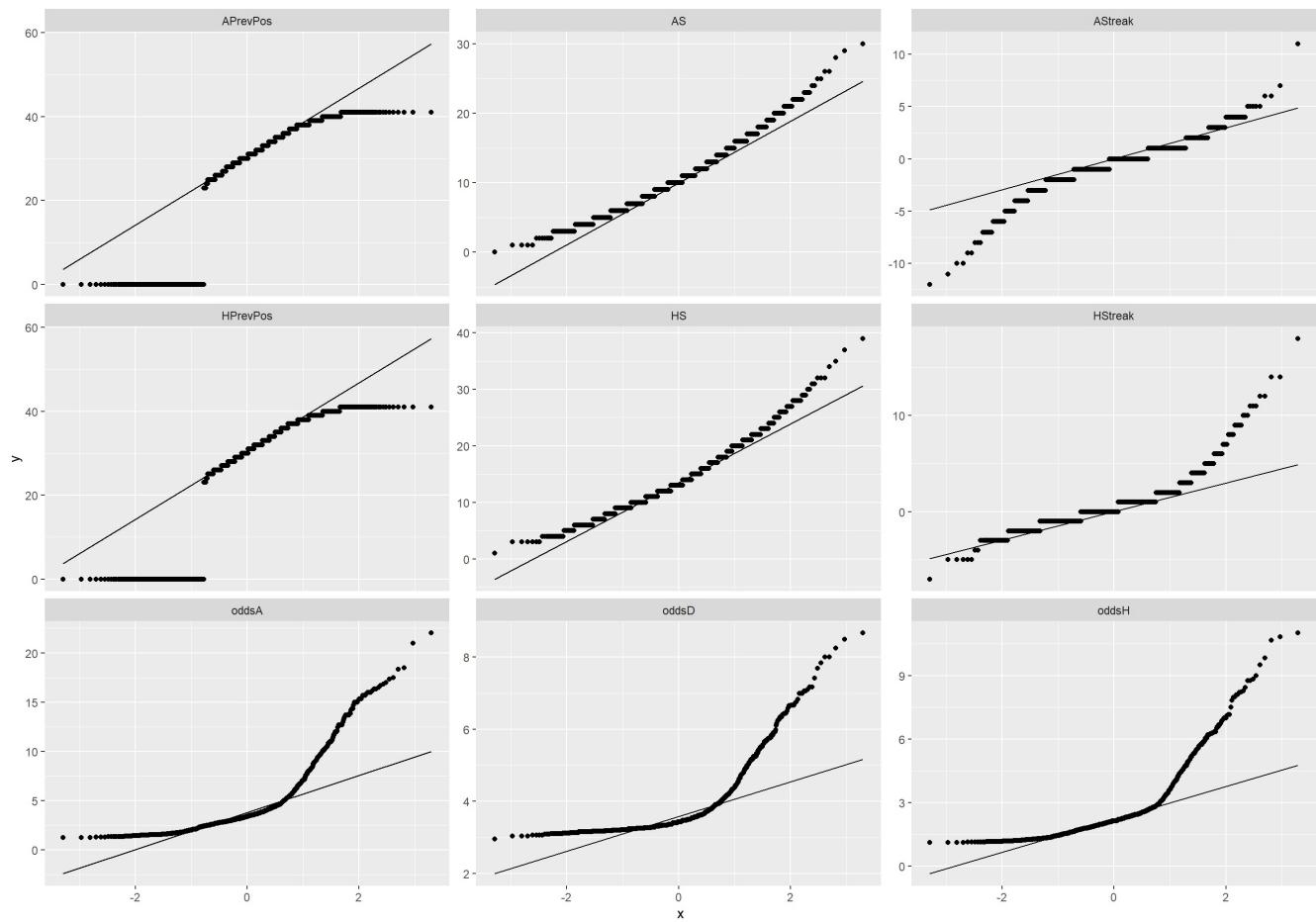
Bar Chart (with frequency)



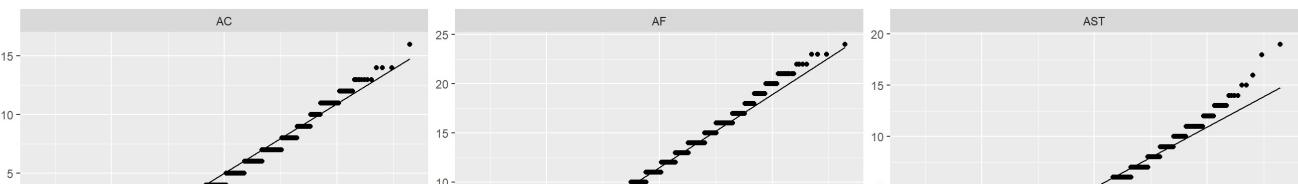
QQ Plot

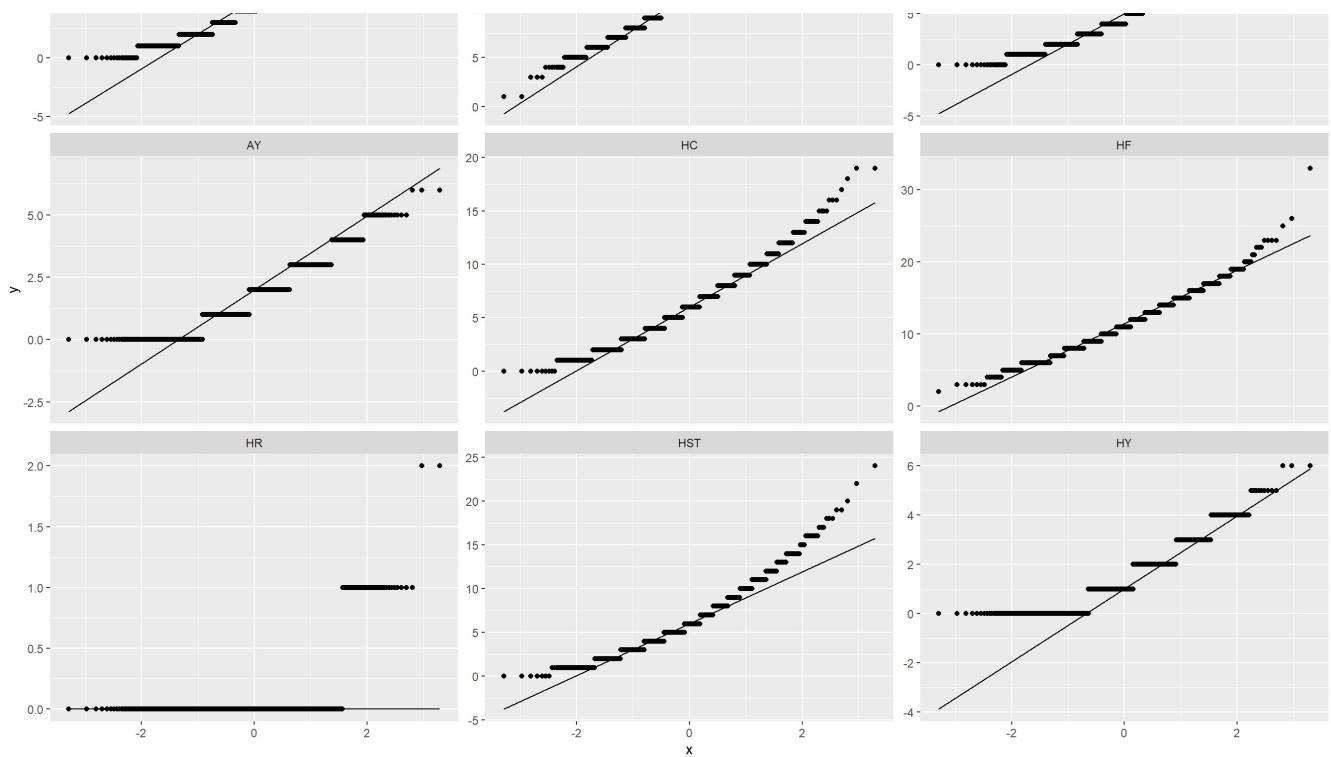


Page 1

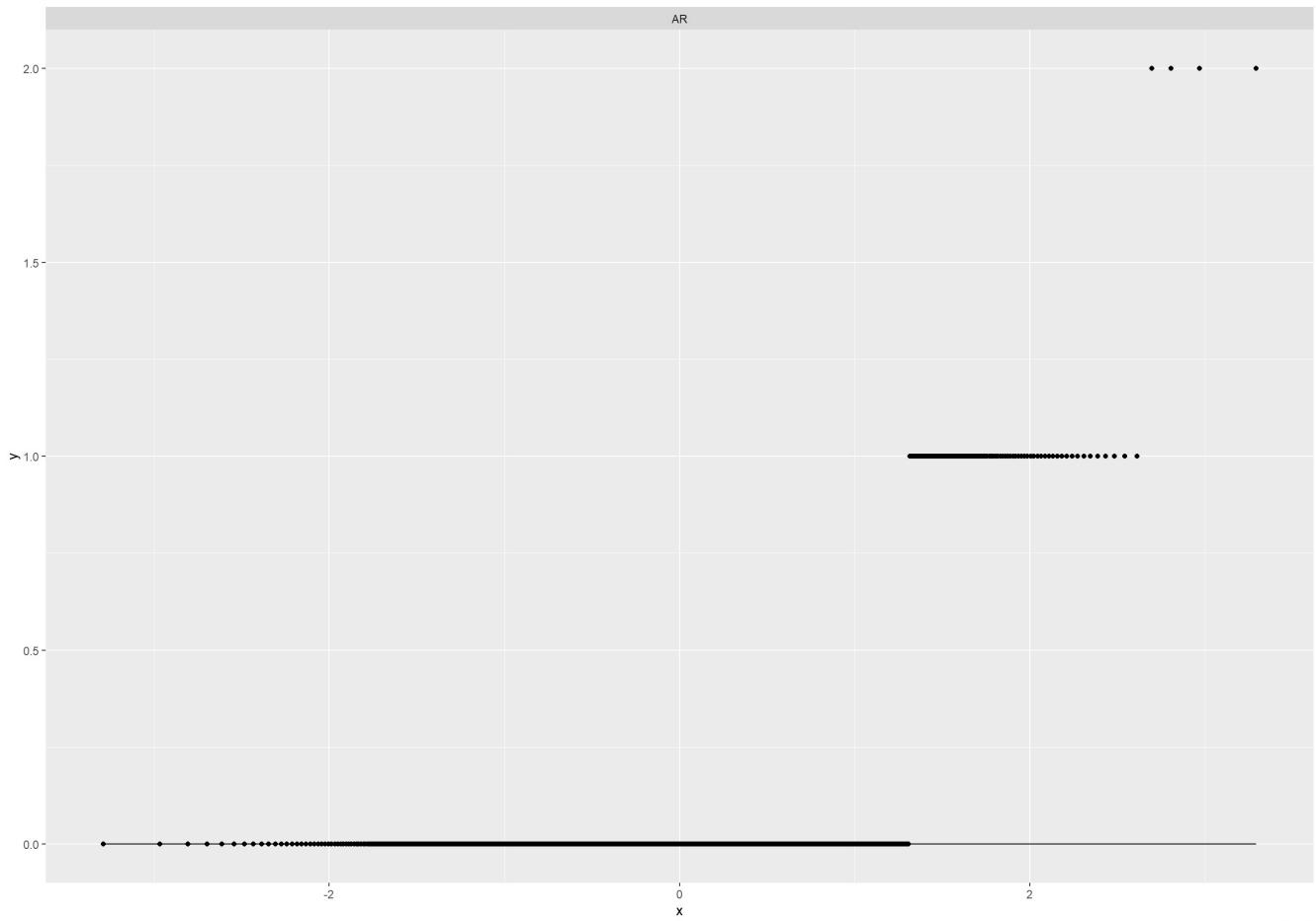


Page 2





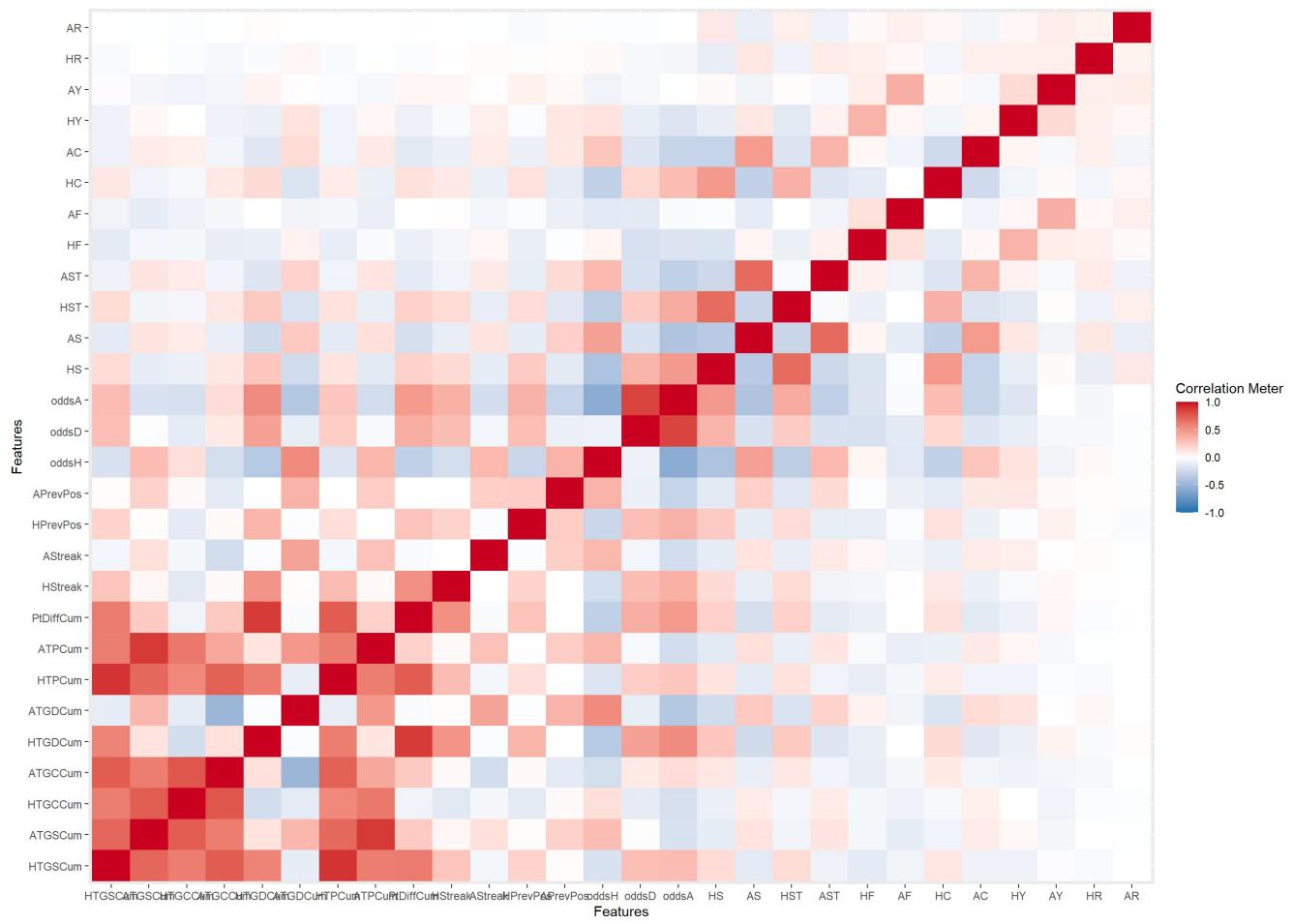
Page 3



Page 4

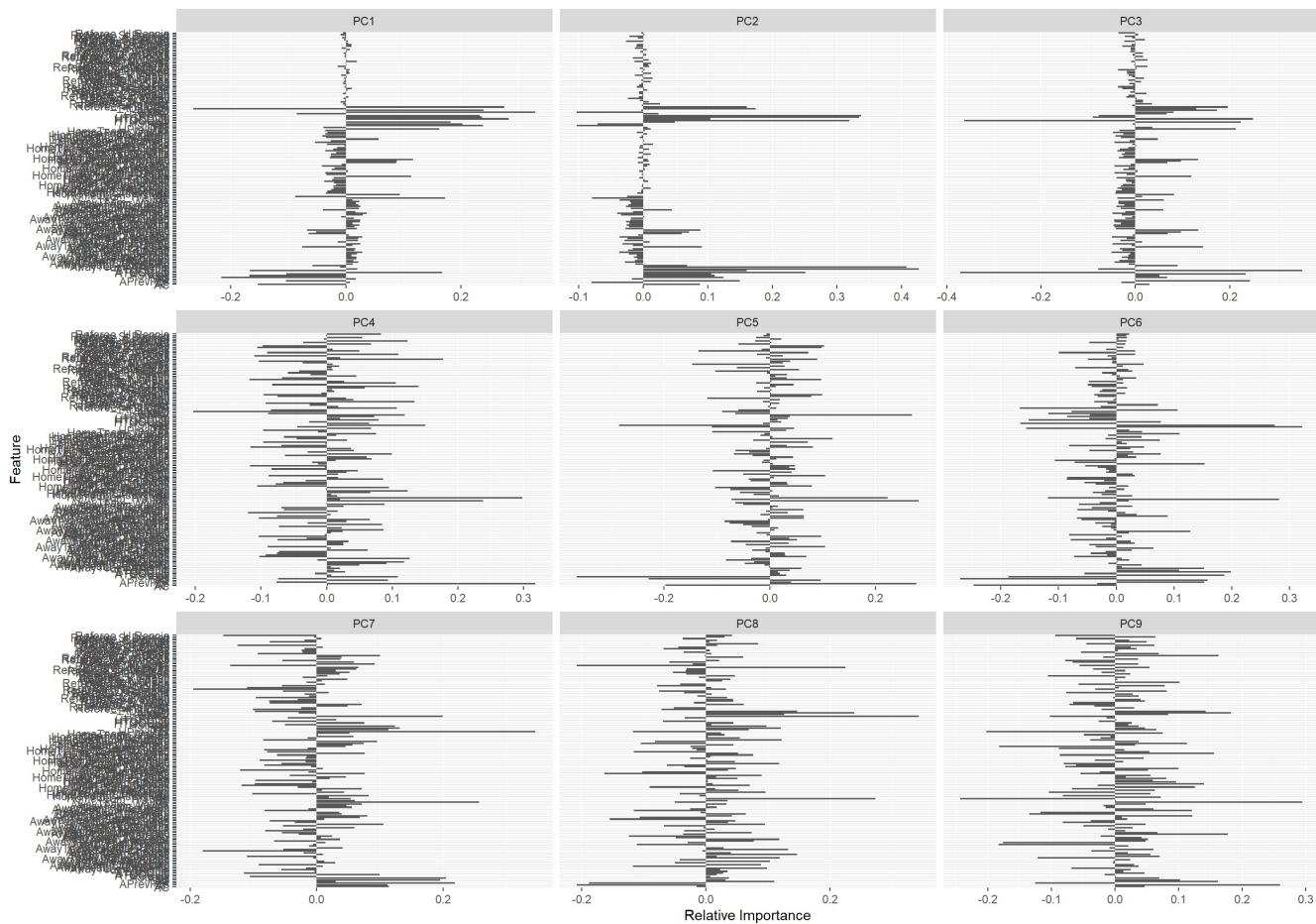
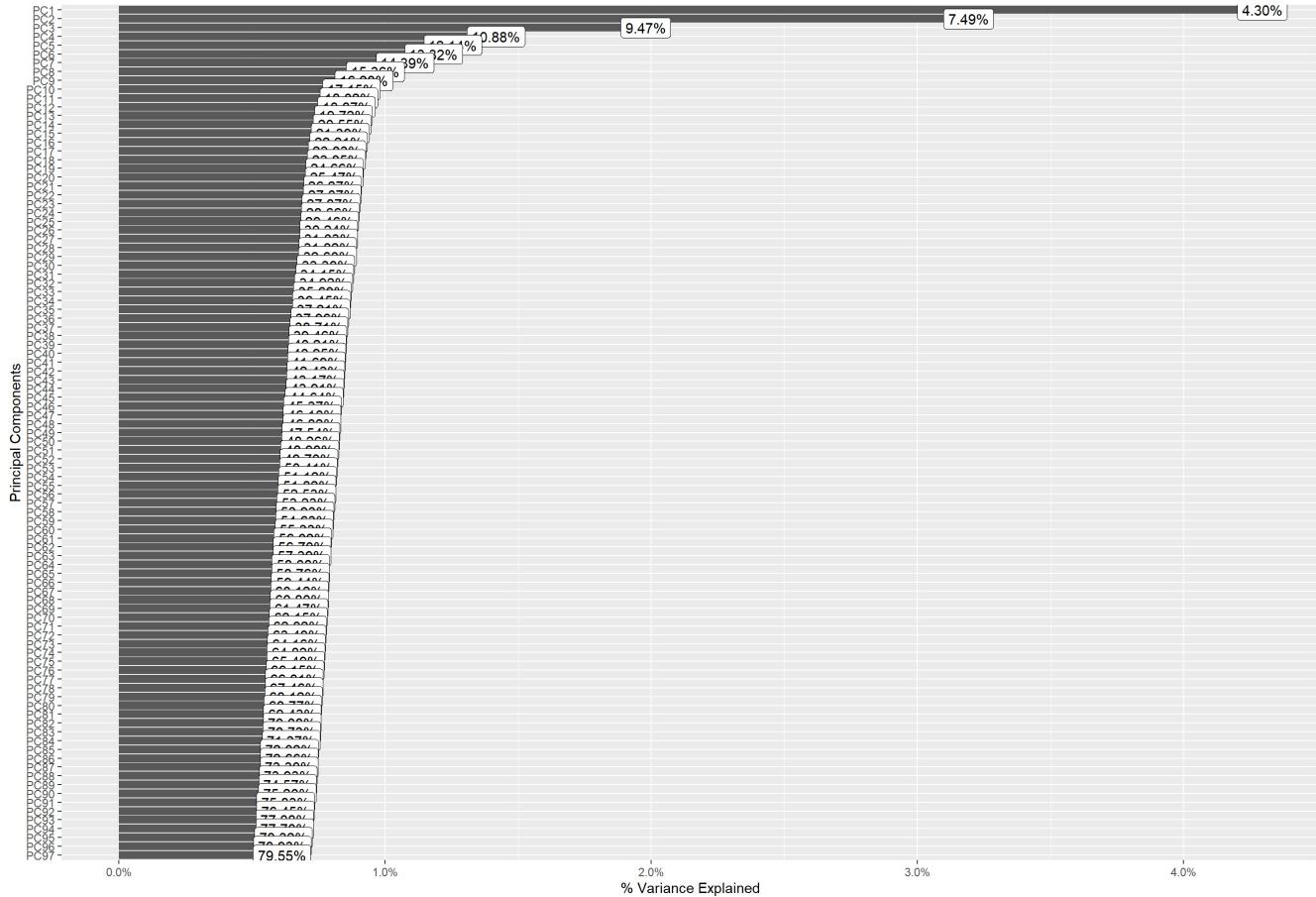
Correlation Analysis

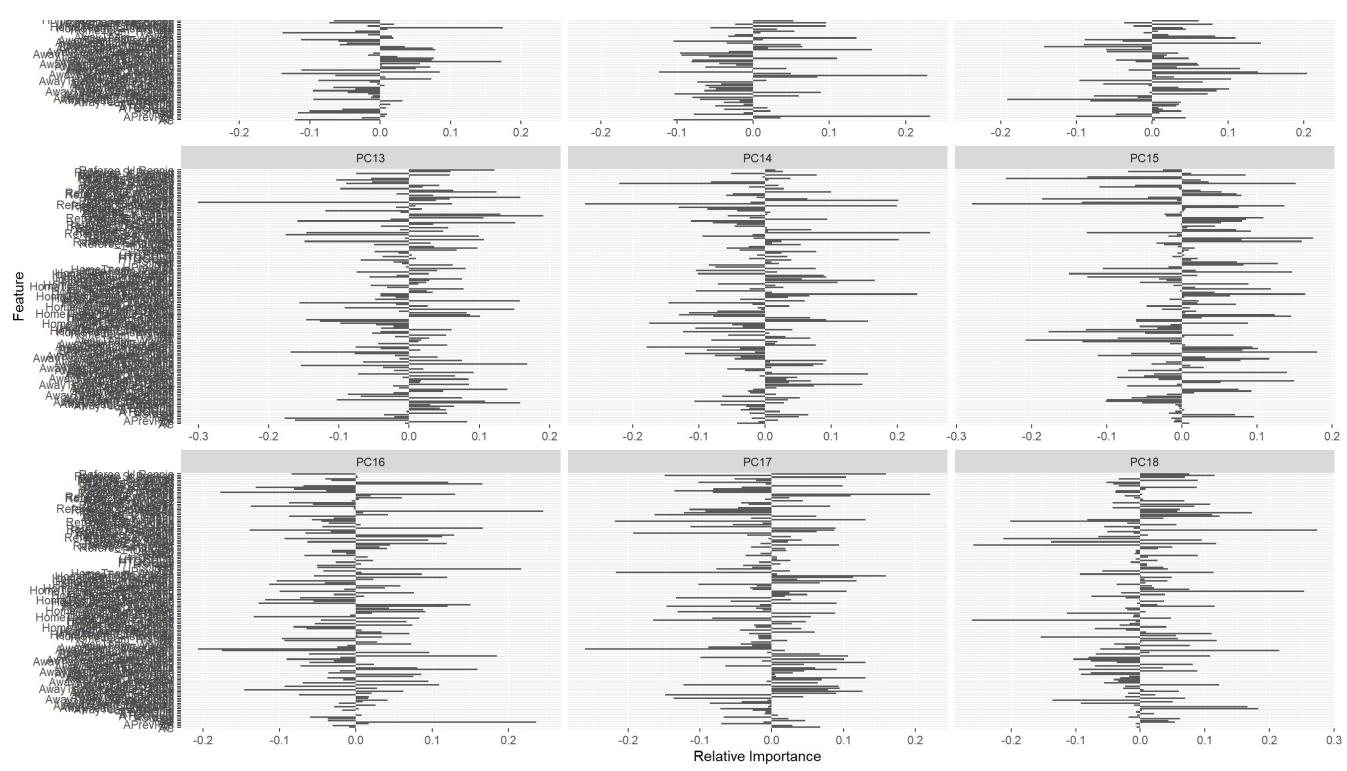
```
## Warning in dummify(data, maxcat = maxcat): Ignored
## all discrete features since `maxcat` set to 20
## categories!
```



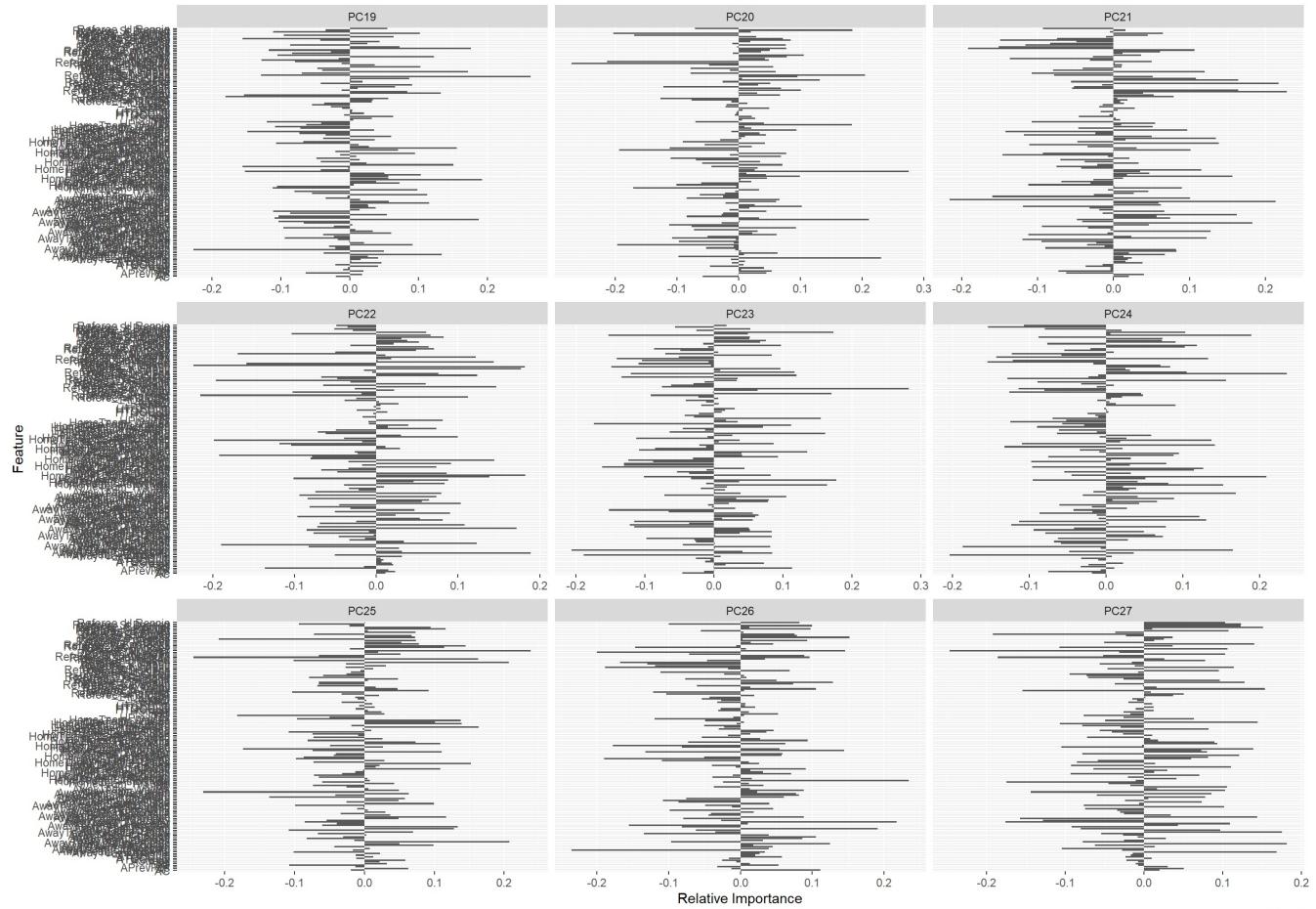
Principal Component Analysis

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

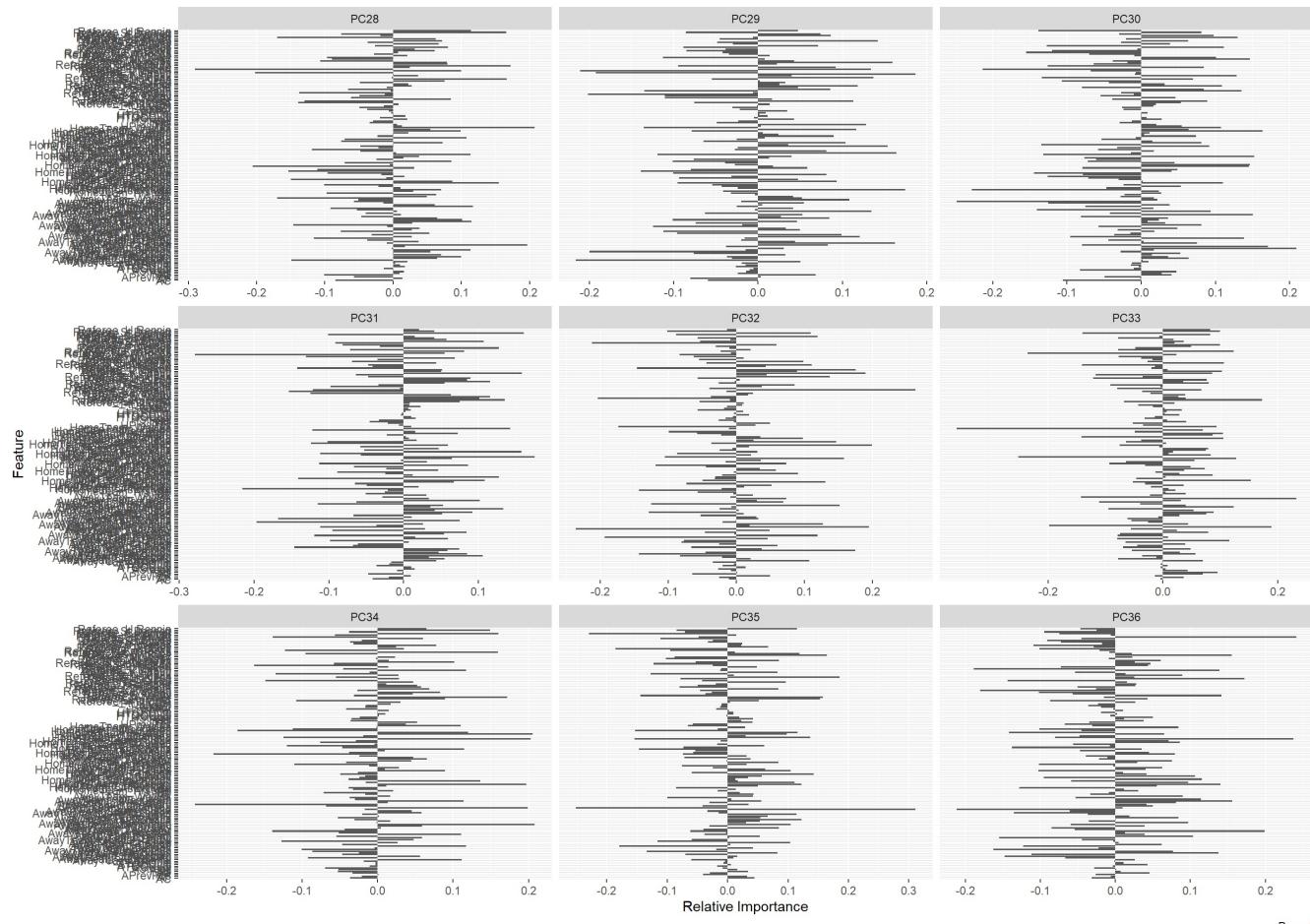




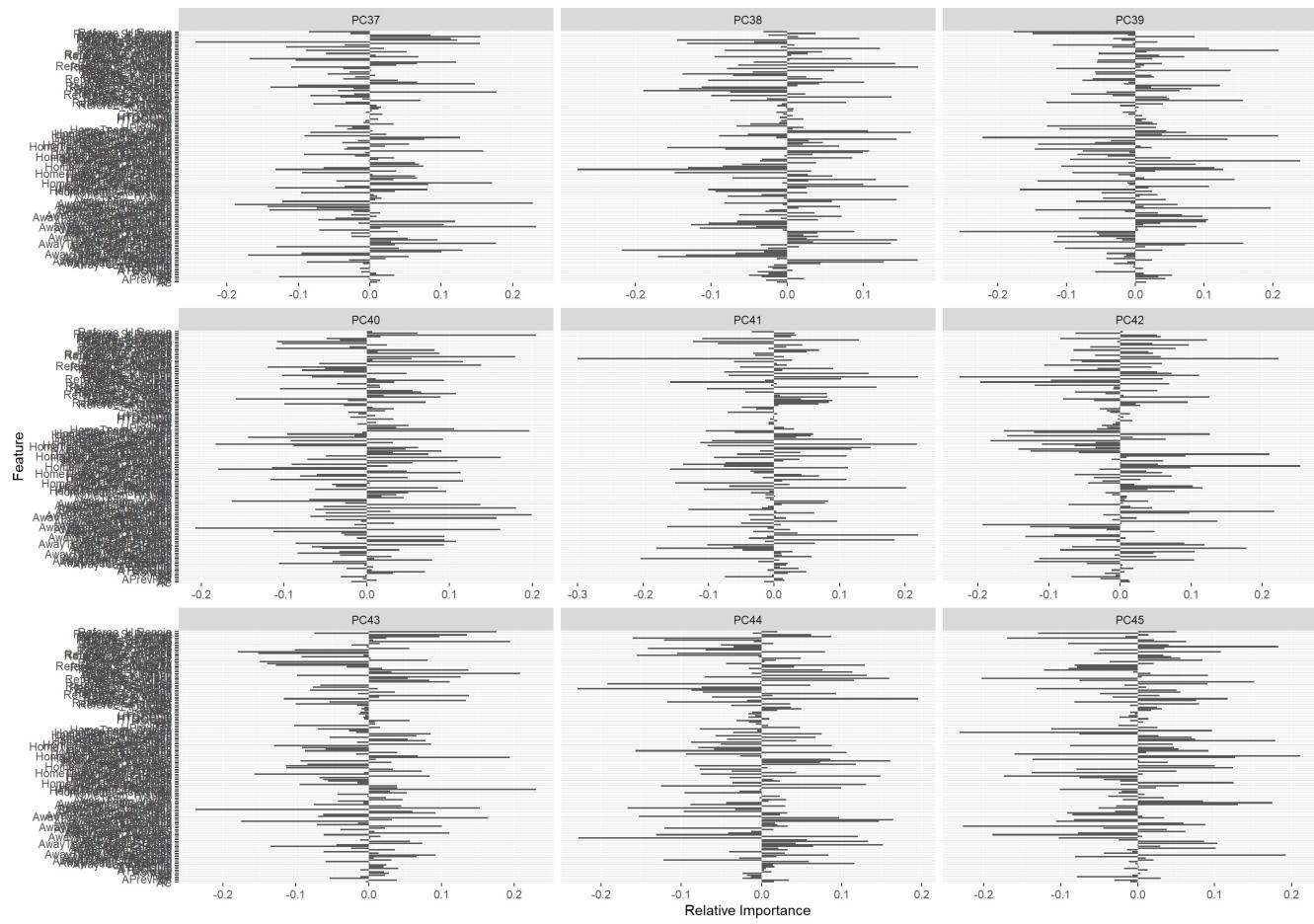
Page 2



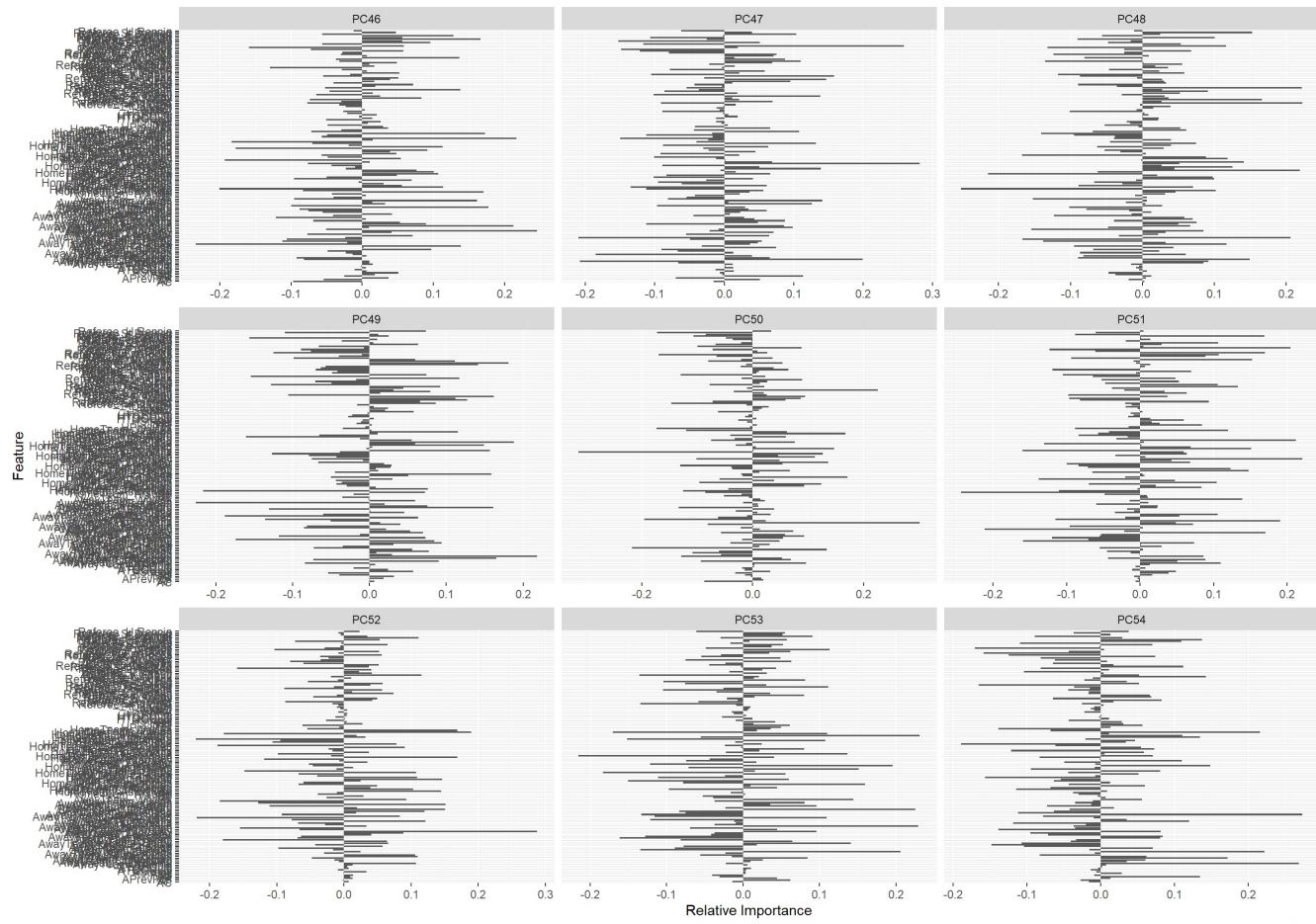
Page 3



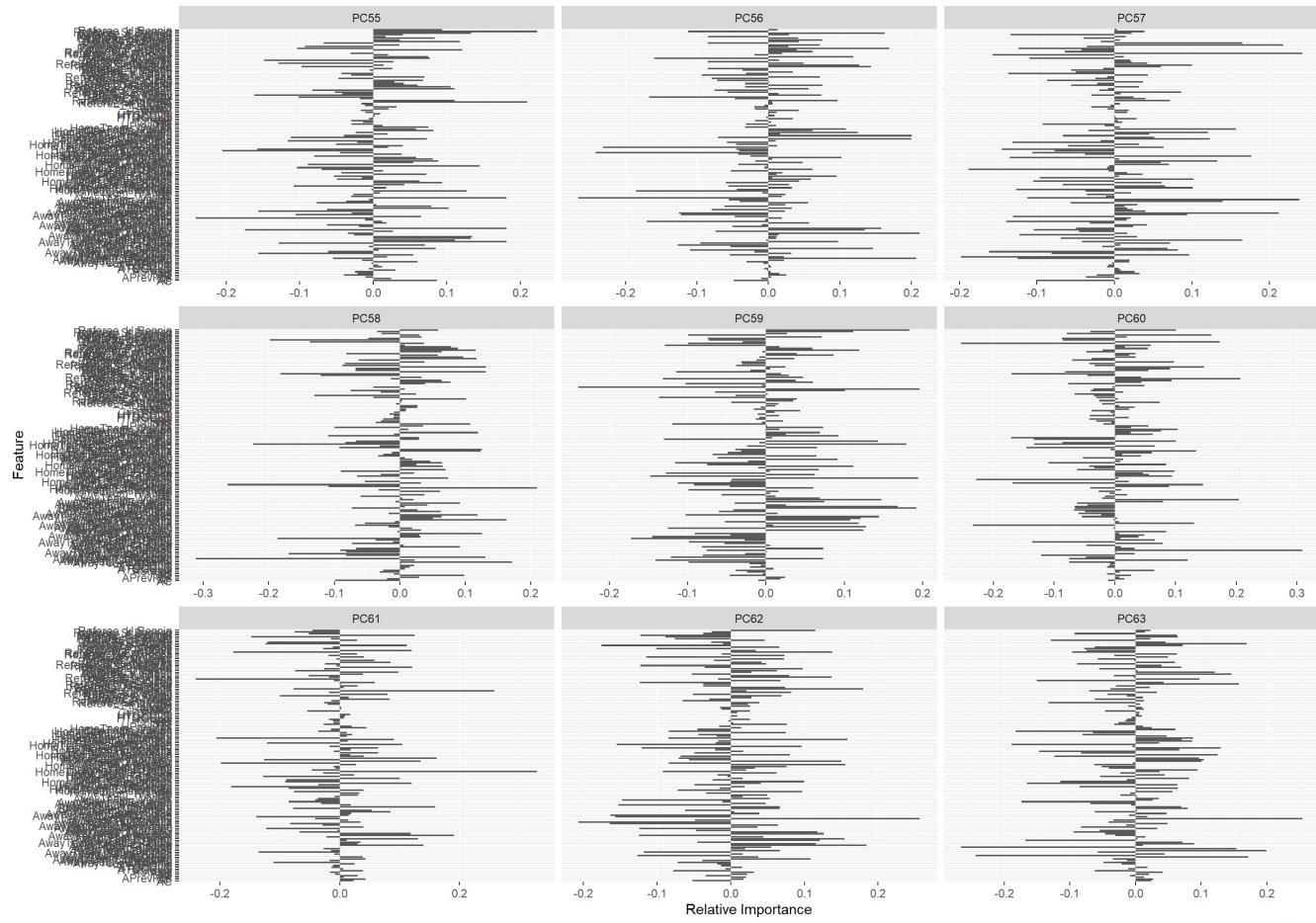
Page 4



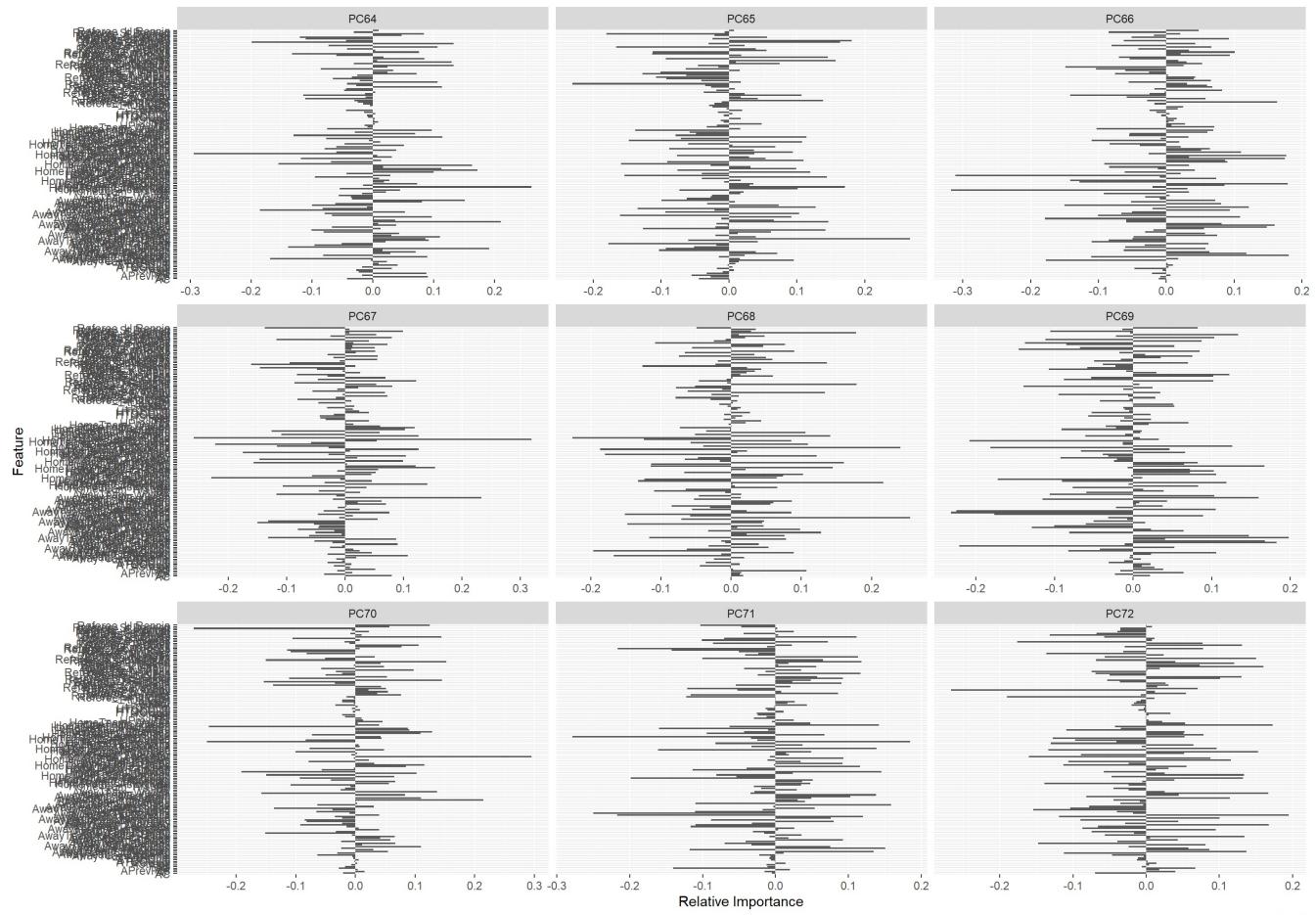
Page 5



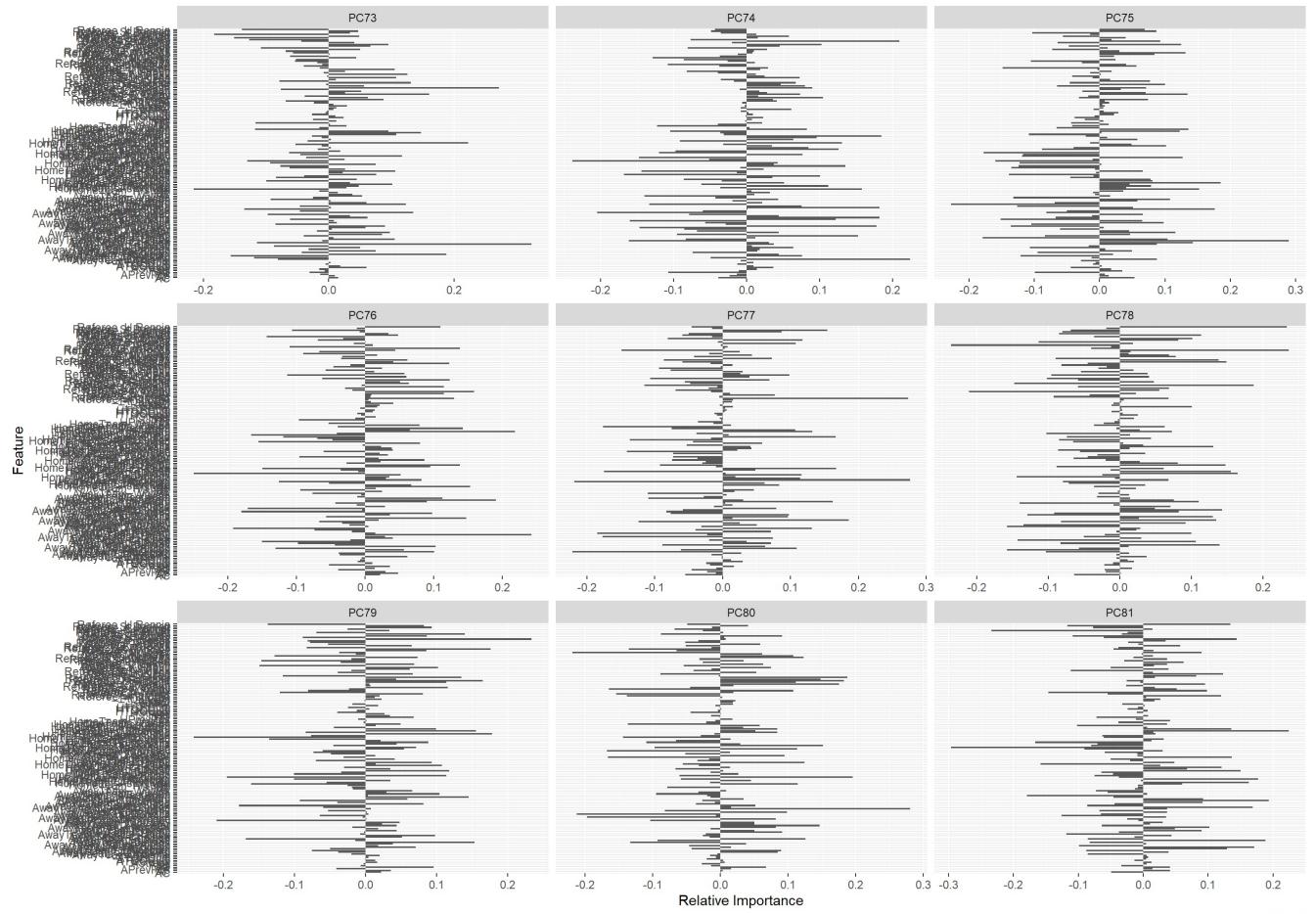
Page 6



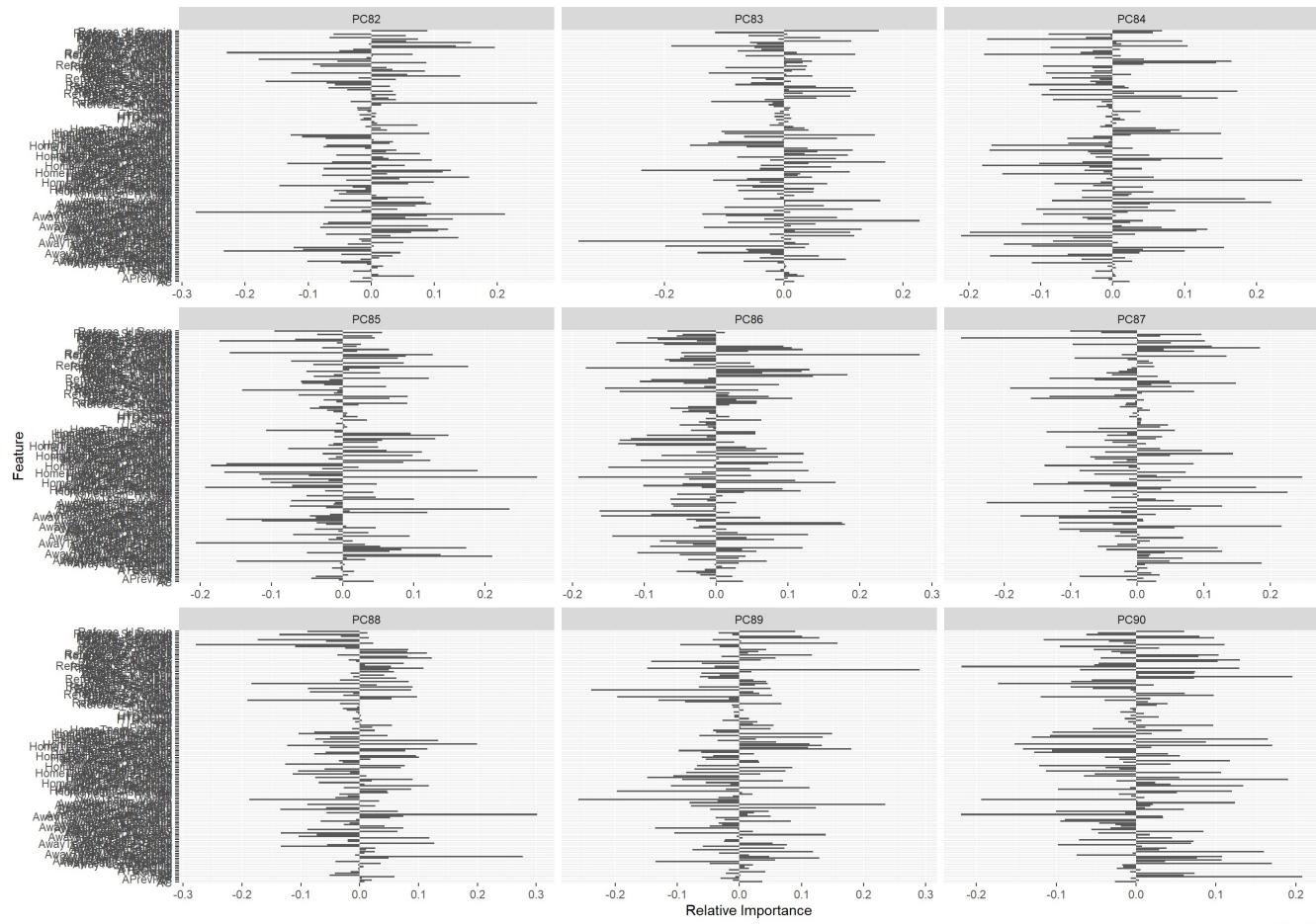
Page 7



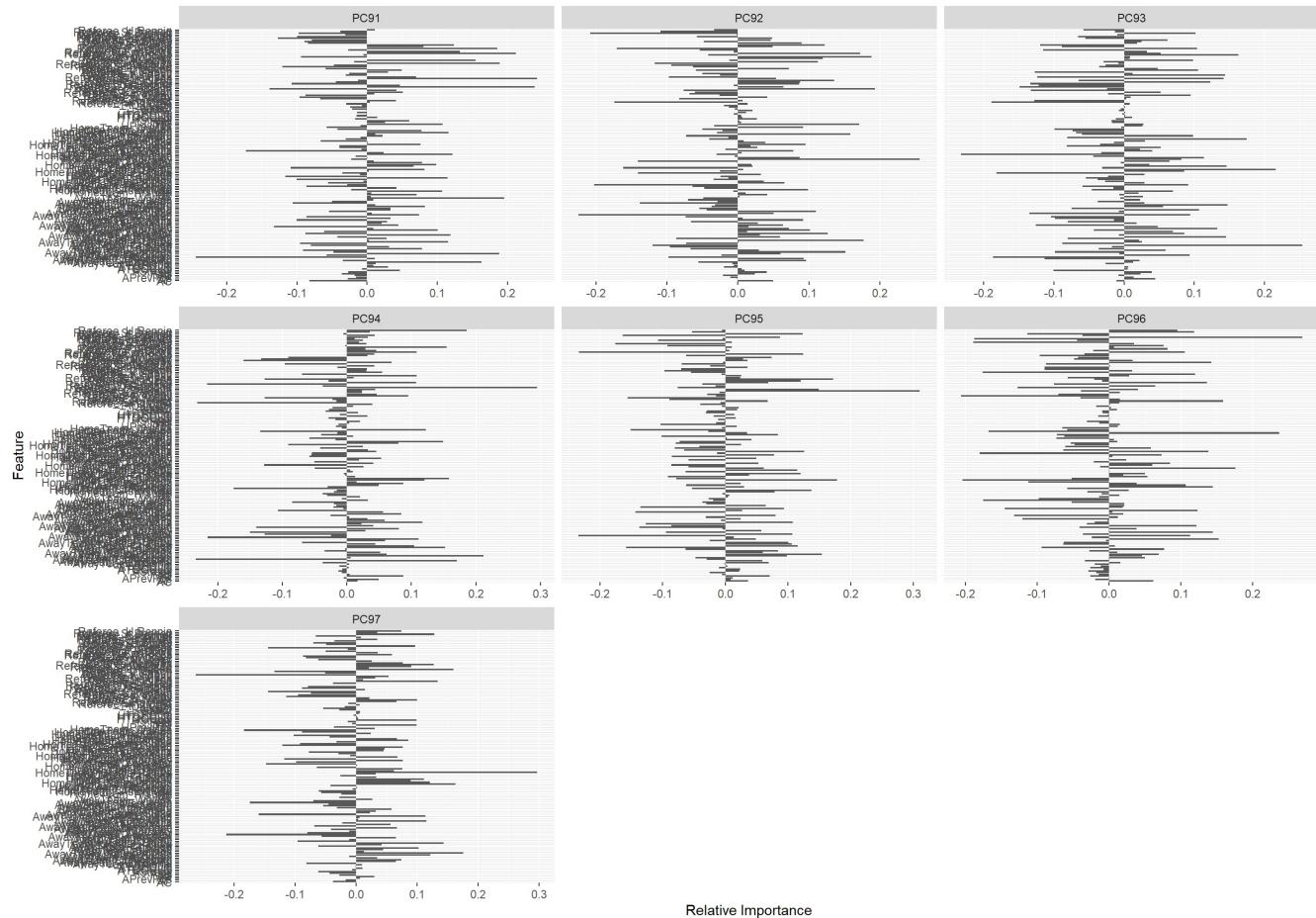
Page 8



Page 9



Page 10



Page 11