

Data Profiling Report

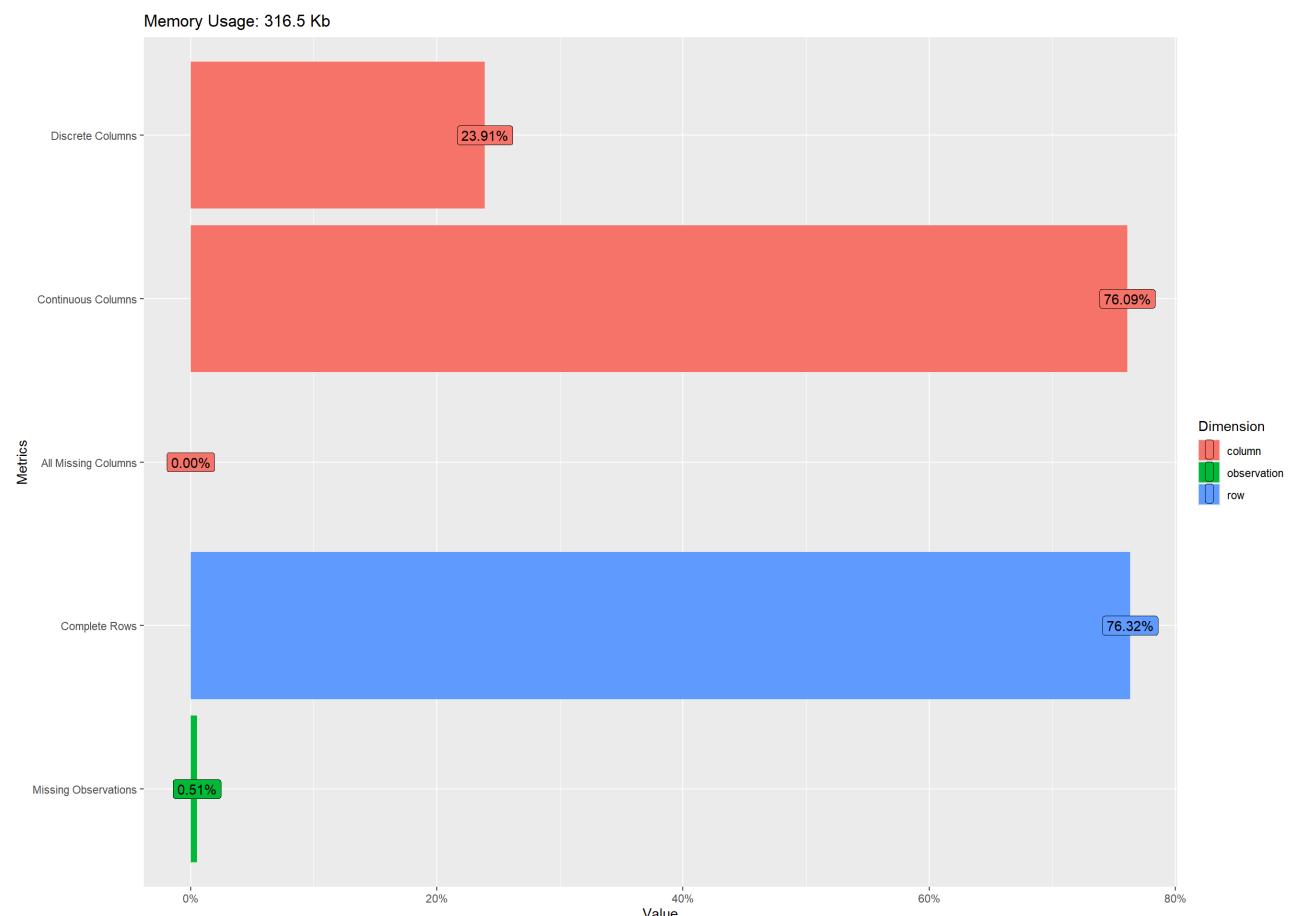
- Basic Statistics
 - Raw Counts
 - Percentages
- Data Structure
- Missing Data Profile
- Univariate Distribution
 - Histogram
 - Bar Chart (with frequency)
 - QQ Plot
- Correlation Analysis
- Principal Component Analysis

Basic Statistics

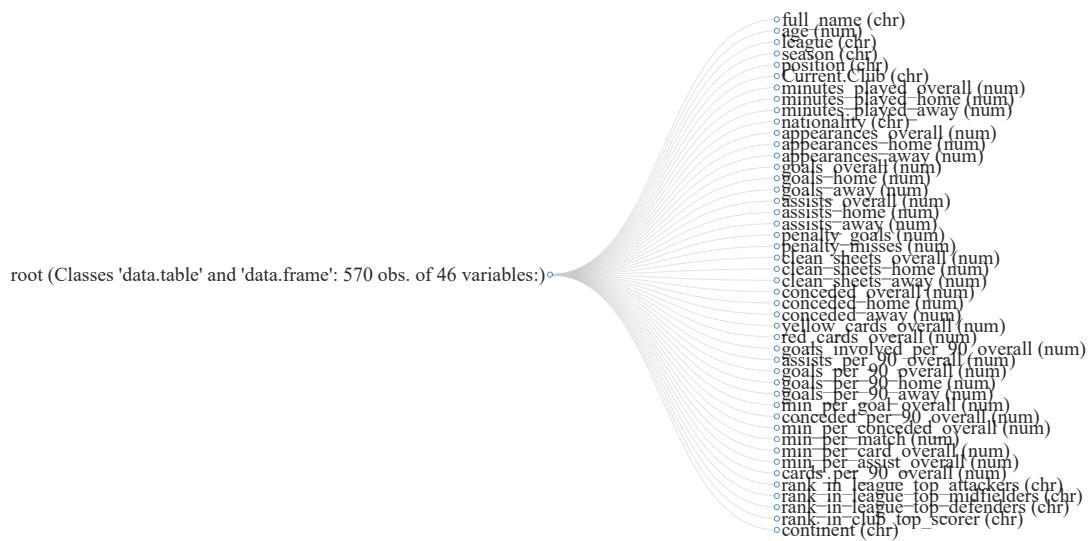
Raw Counts

Name	Value
Rows	570
Columns	46
Discrete columns	11
Continuous columns	35
All missing columns	0
Missing observations	135
Complete Rows	435
Total observations	26,220
Memory allocation	316.5 Kb

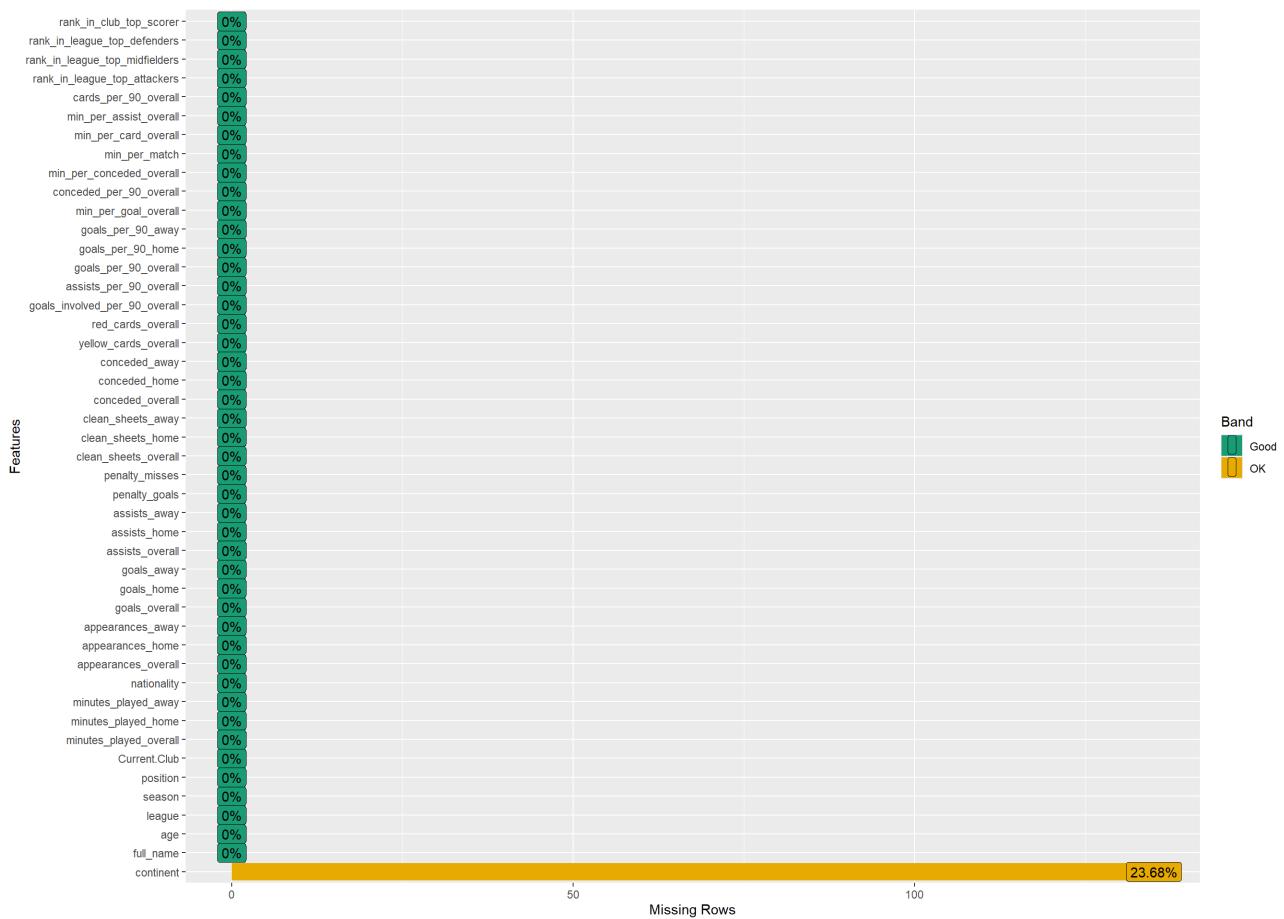
Percentages



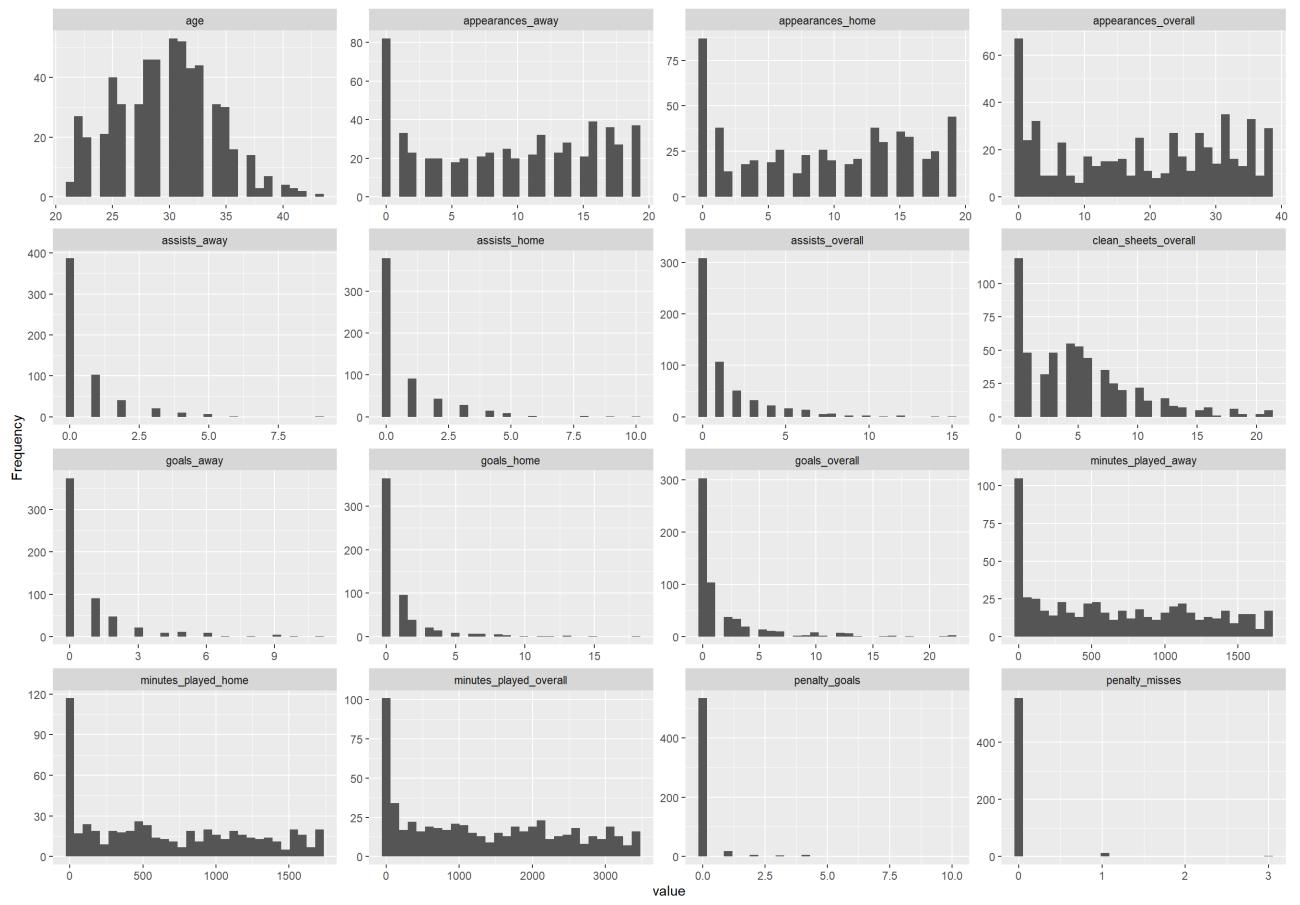
Data Structure



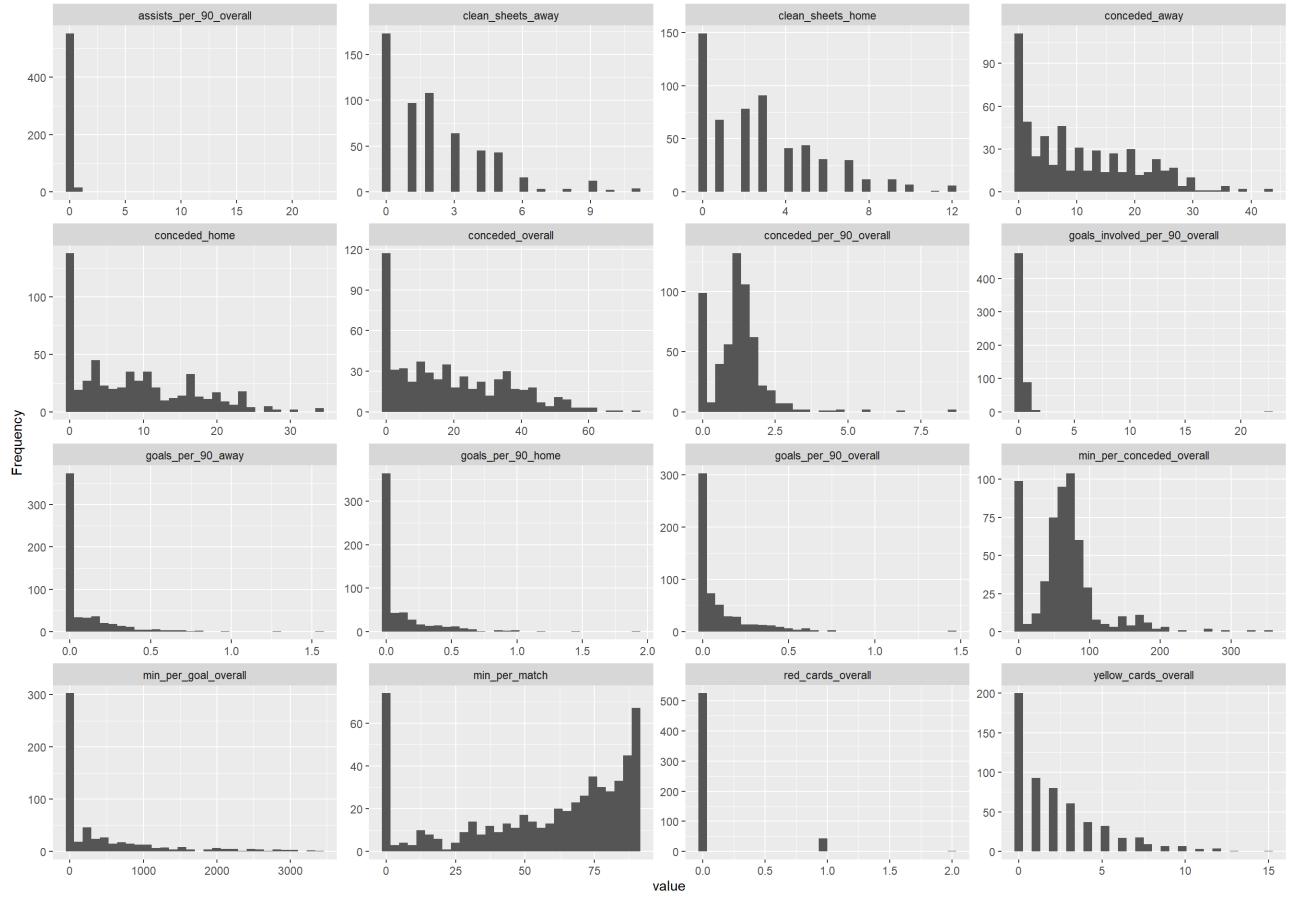
Missing Data Profile



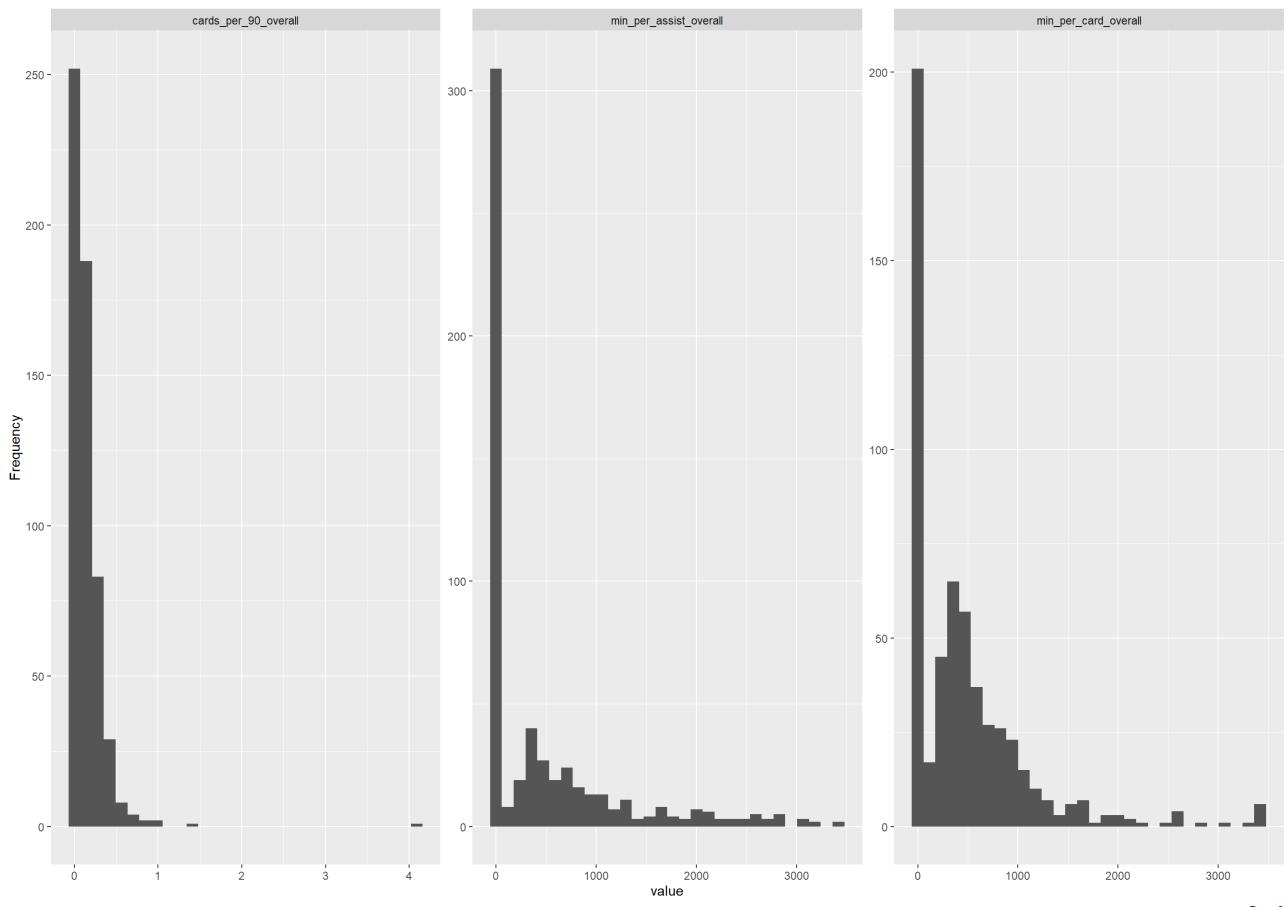
Univariate Distribution Histogram



Page 1



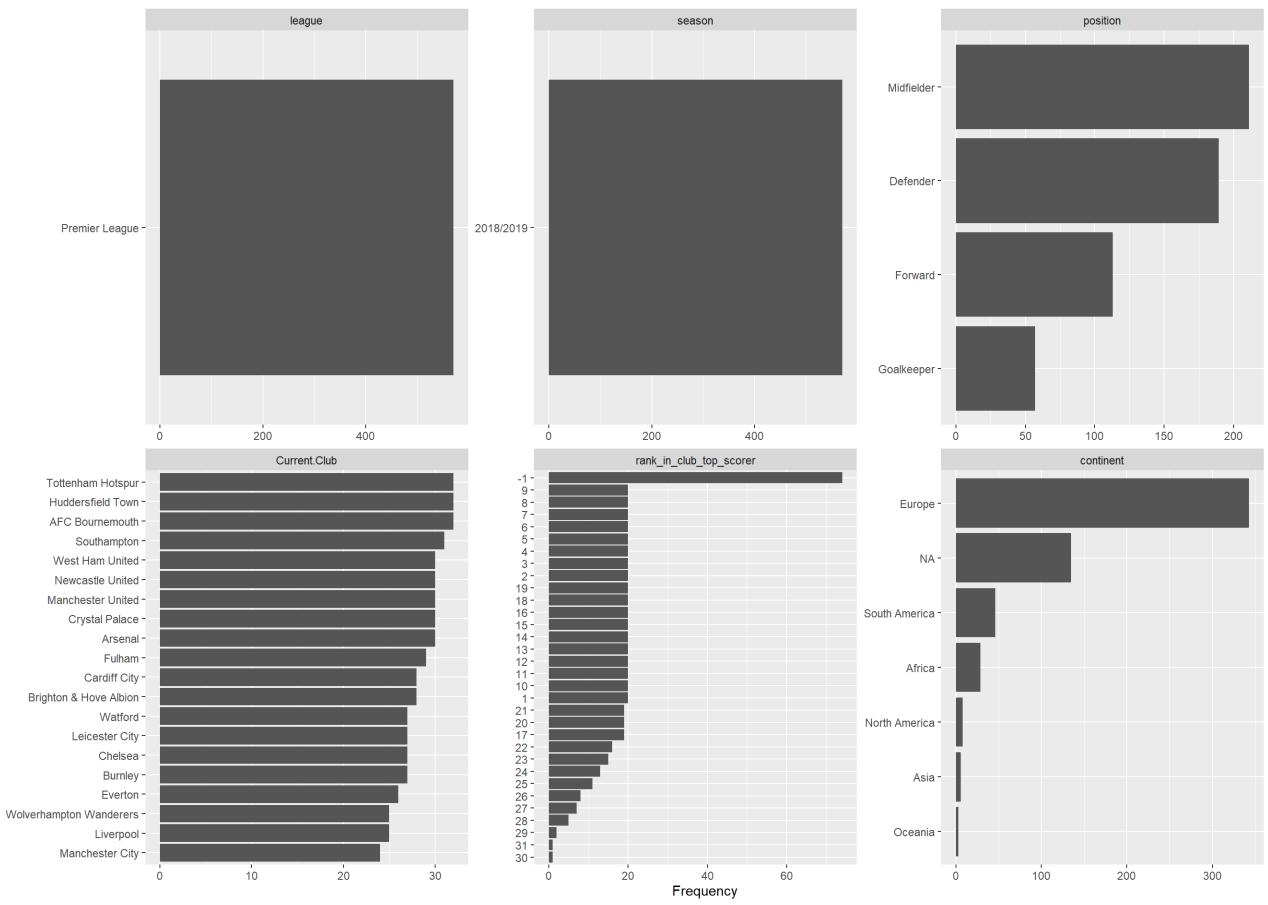
Page 2



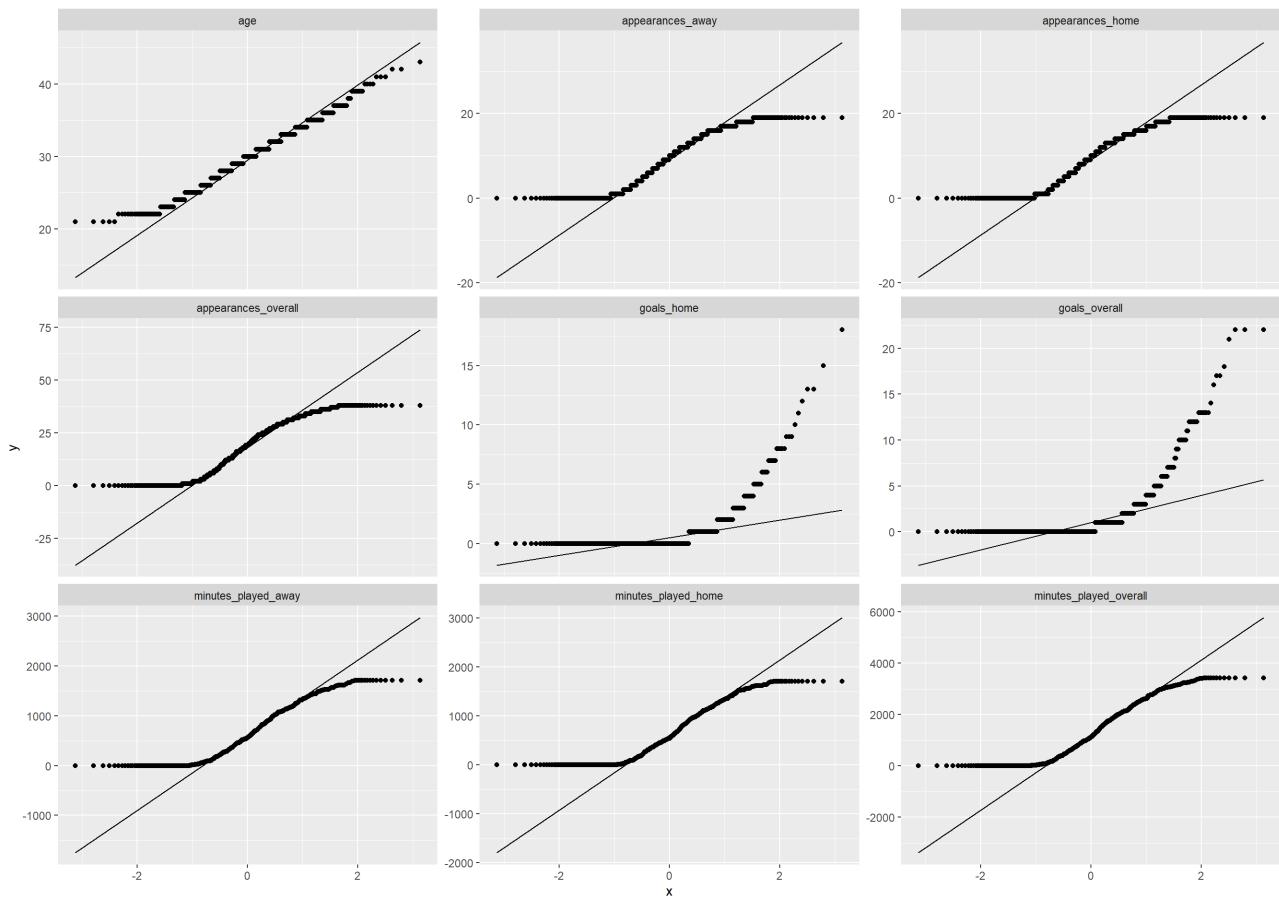
Page 3

Bar Chart (with frequency)

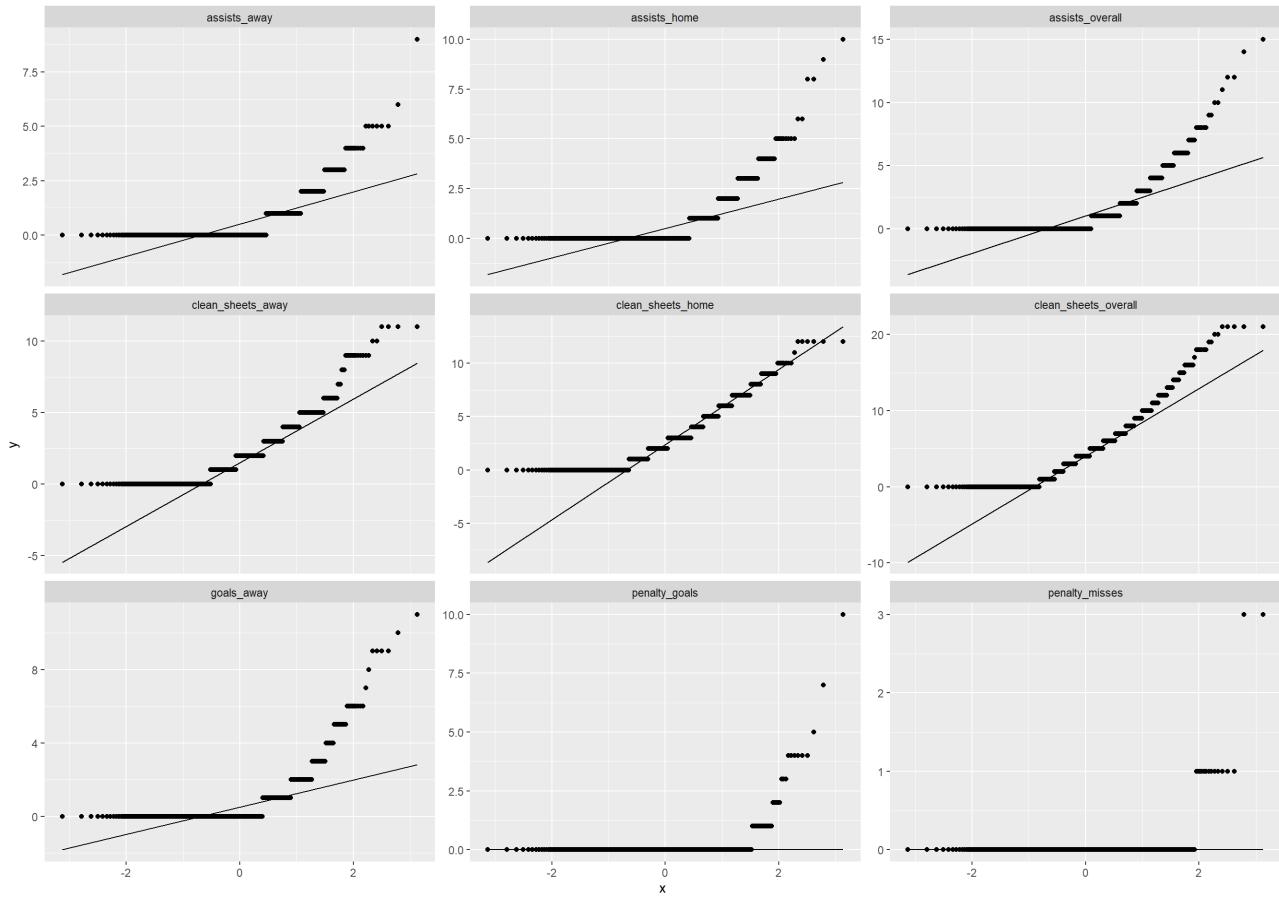
```
## 5 columns ignored with more than 50 categories.
## full_name: 569 categories
## nationality: 67 categories
## rank_in_league_top_attackers: 420 categories
## rank_in_league_top_midfielders: 420 categories
## rank_in_league_top_defenders: 176 categories
```



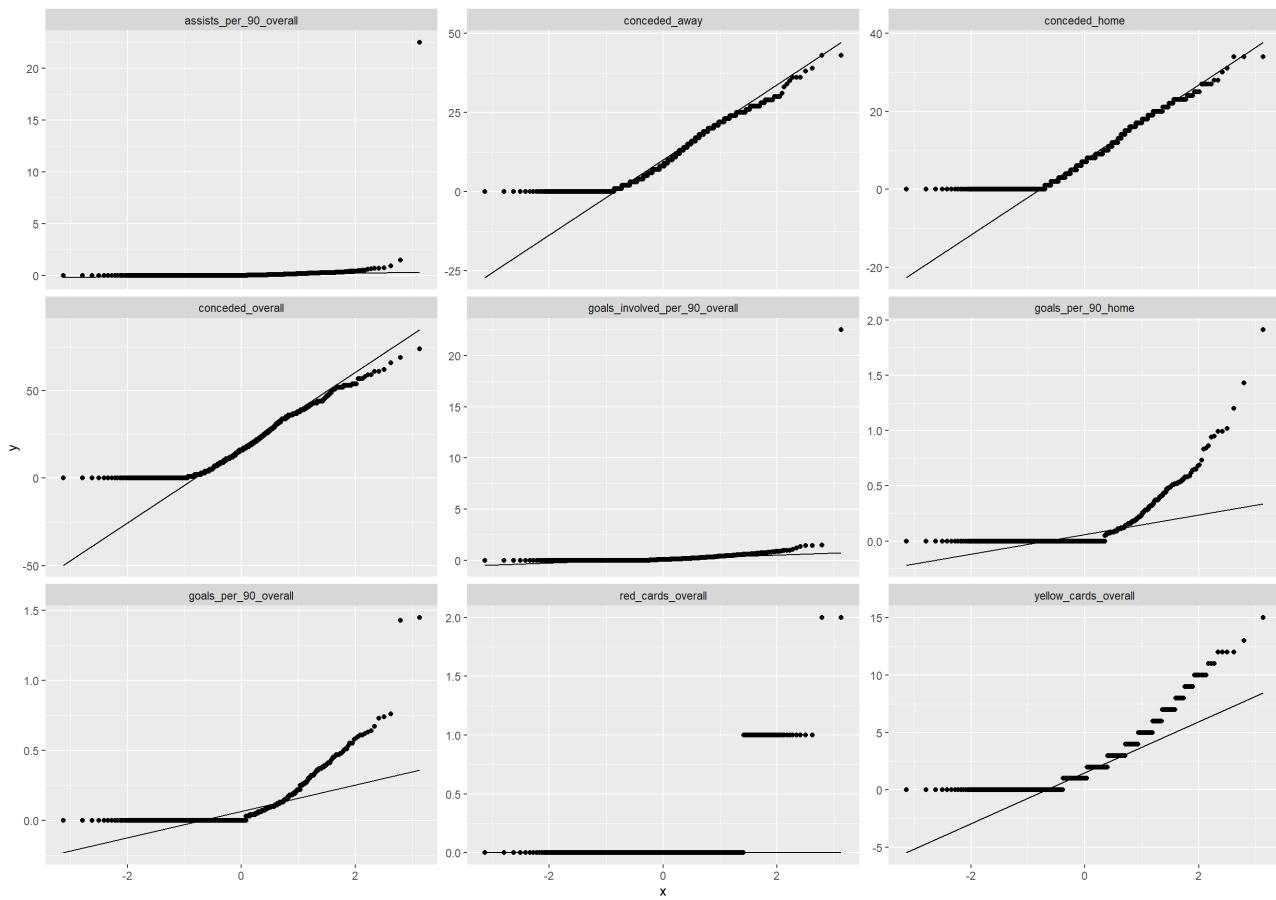
QQ Plot



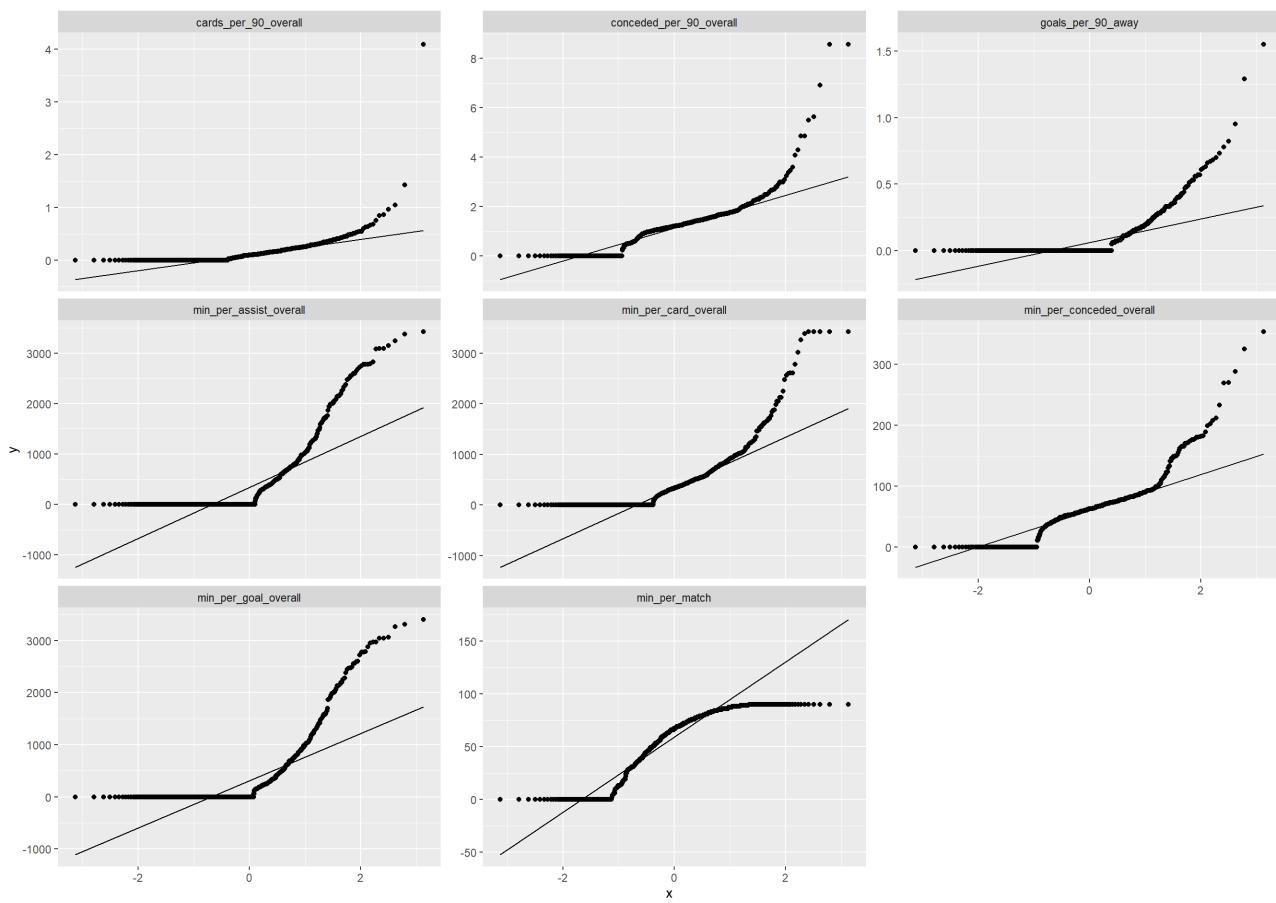
Page 1



Page 2



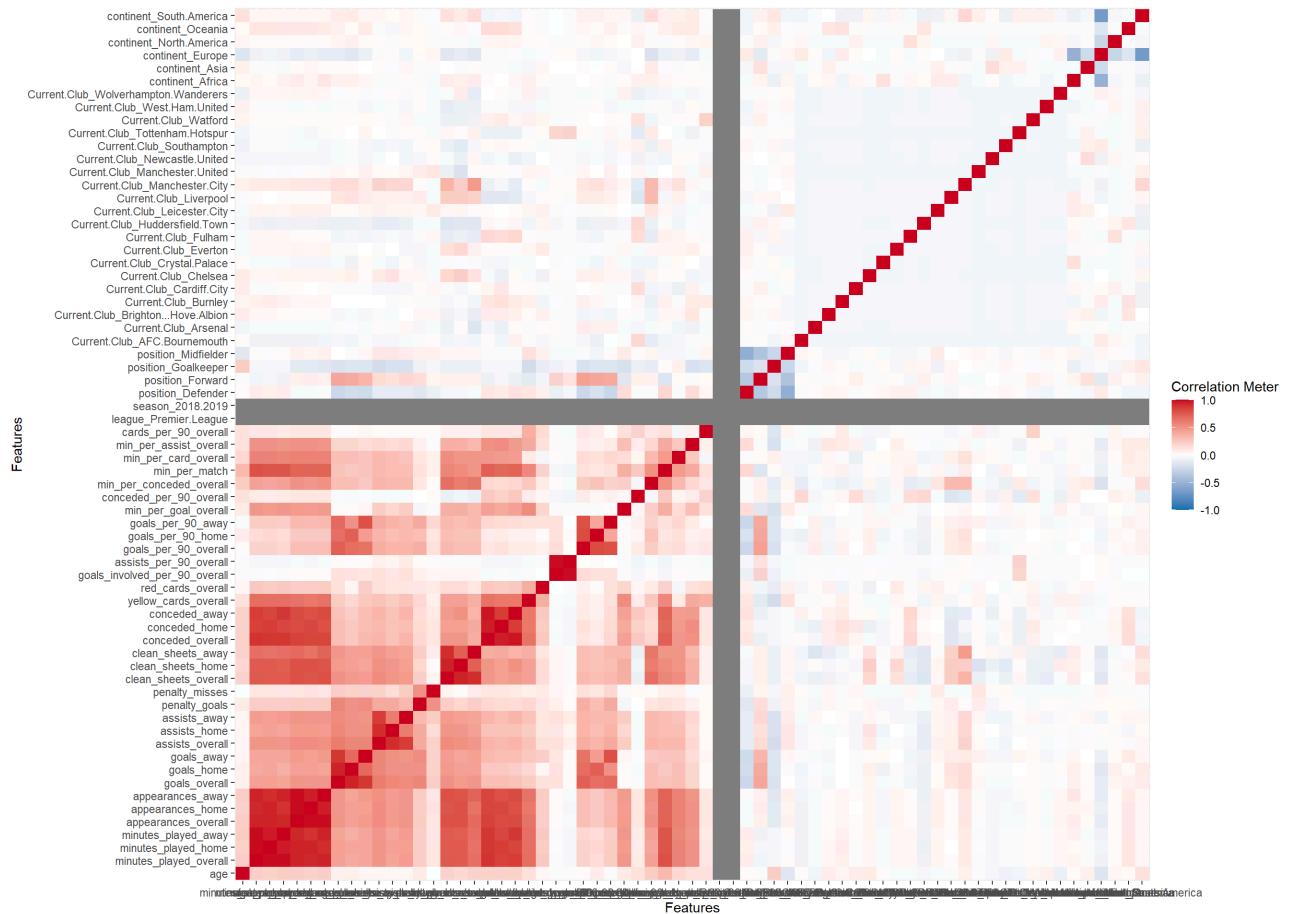
Page 3



Page 4

Correlation Analysis

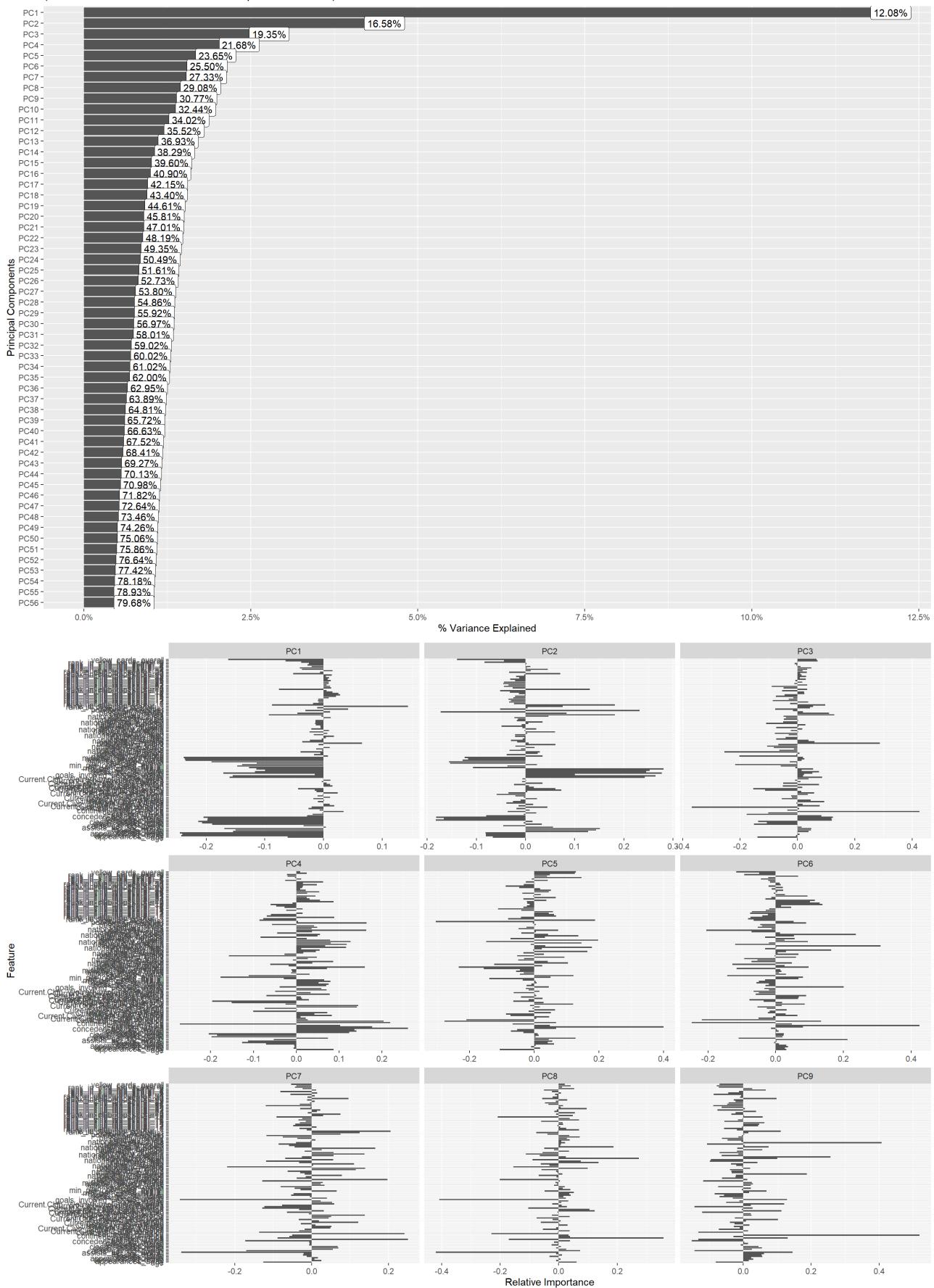
```
## 6 features with more than 20 categories ignored!
## full_name: 435 categories
## nationality: 29 categories
## rank_in_league_top_attackers: 314 categories
## rank_in_league_top_midfielders: 314 categories
## rank_in_league_top_defenders: 129 categories
## rank_in_club_top_scorer: 32 categories
```

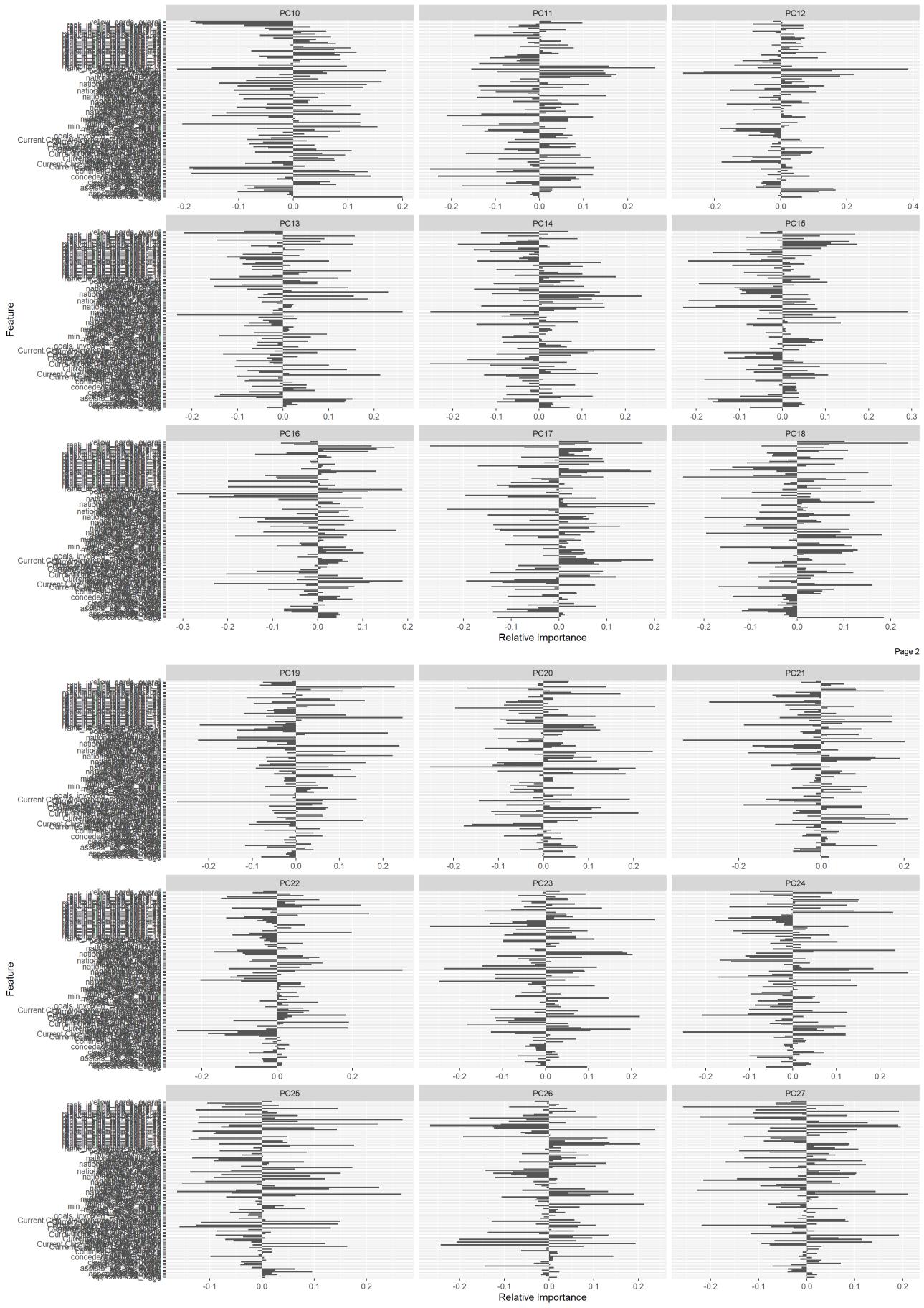


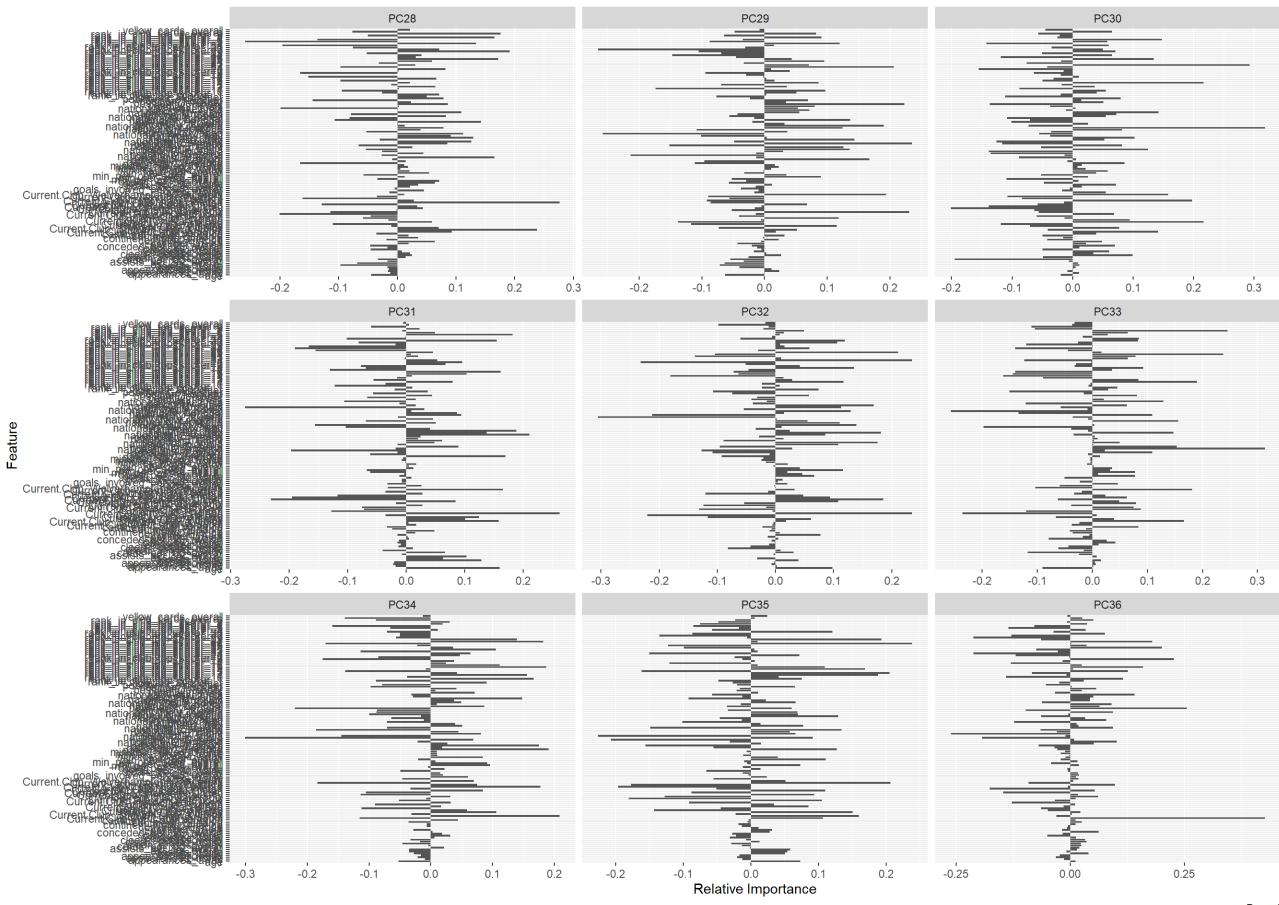
Principal Component Analysis

```
## 4 features with more than 50 categories ignored!
## full_name: 435 categories
## rank_in_league_top_attackers: 314 categories
## rank_in_league_top_midfielders: 314 categories
## rank_in_league_top_defenders: 129 categories
```

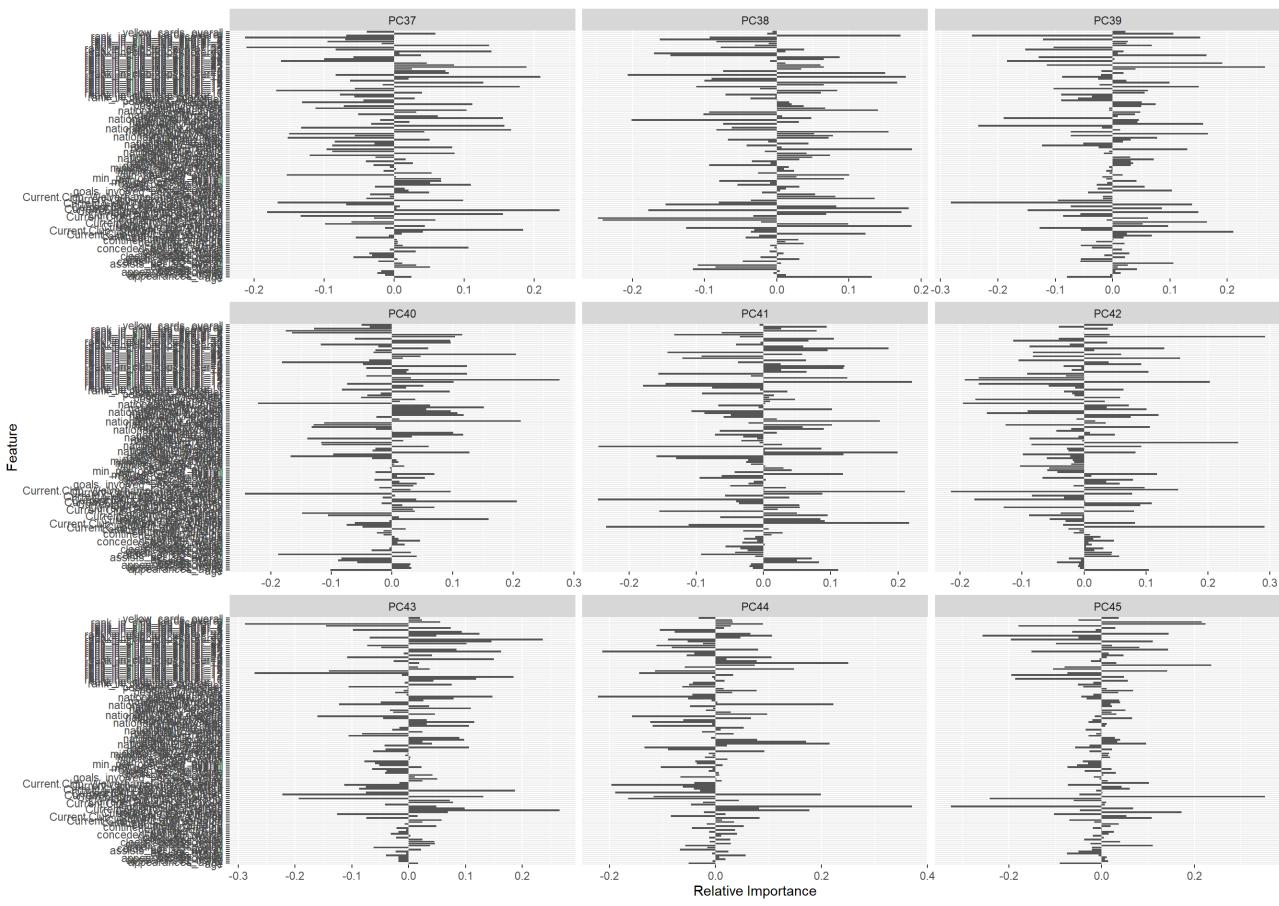
% Variance Explained By Principal Components (Note: Labels indicate cumulative % explained variance)



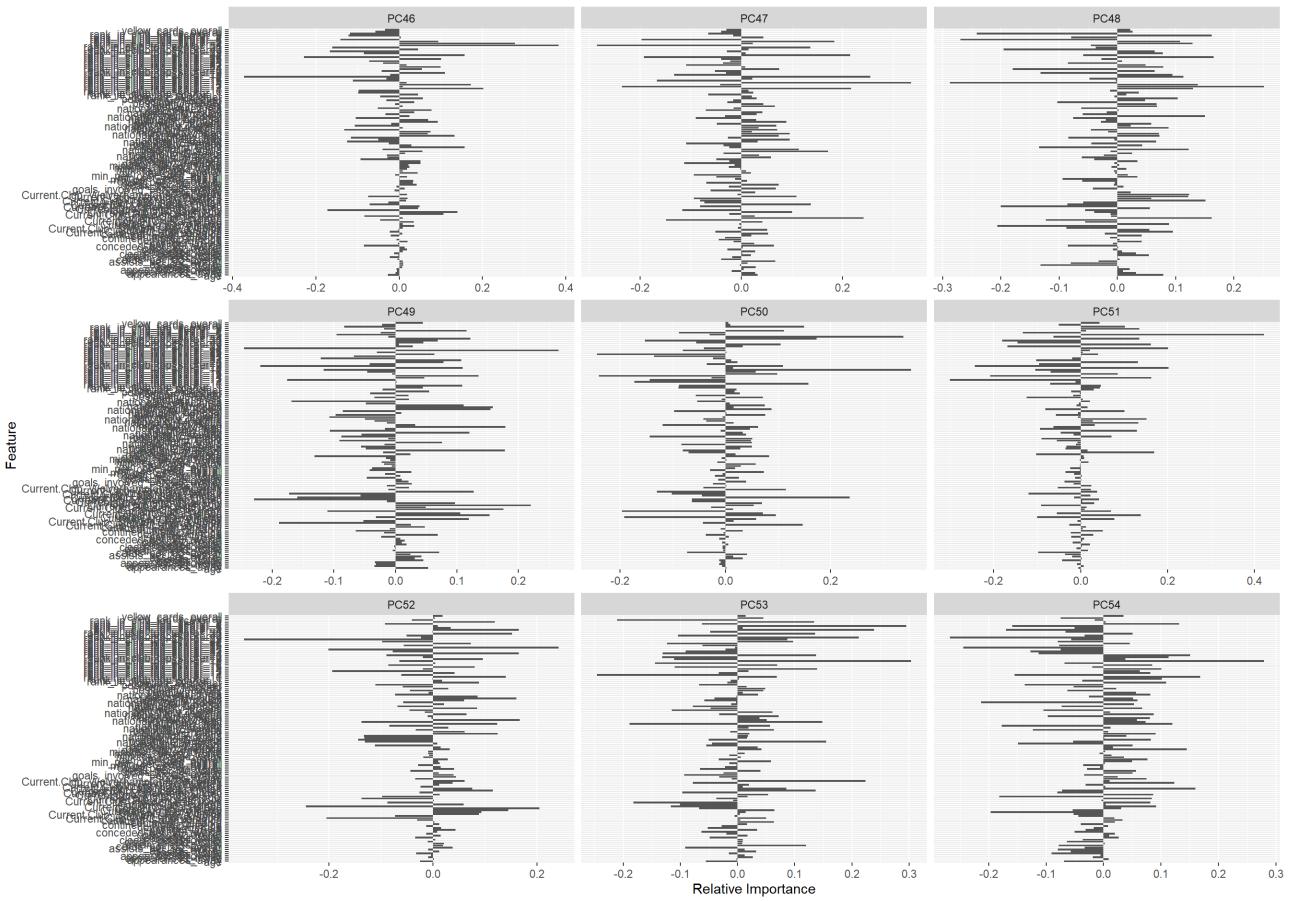




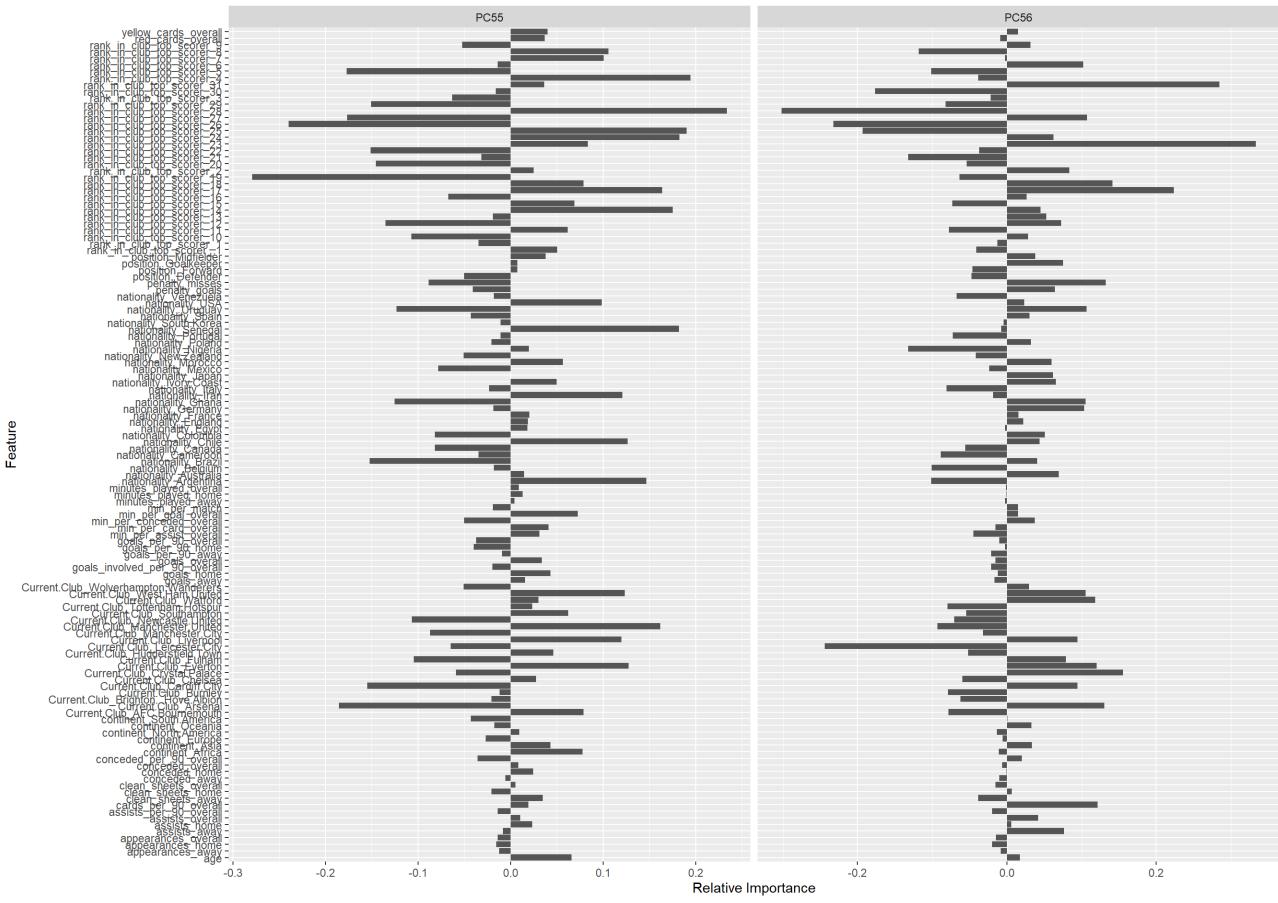
Page 4



Page 5



Page 6



Page 7