

world money_ball - kaggle dataset

patrick mcdevitt (✉ collaboration with preeti swaminathan)

04 juin 2017

money_ball - Regression Project -

Intro

The following **Context** and **Content** is from : https://www.kaggle.com/unbsdn/world-money_ball

Objective

Within the context of **World money_ball Day** explore the data available from the United Nations survey and identify the characteristics that are associated with higher **money_ball** scores

Descriptive statistics

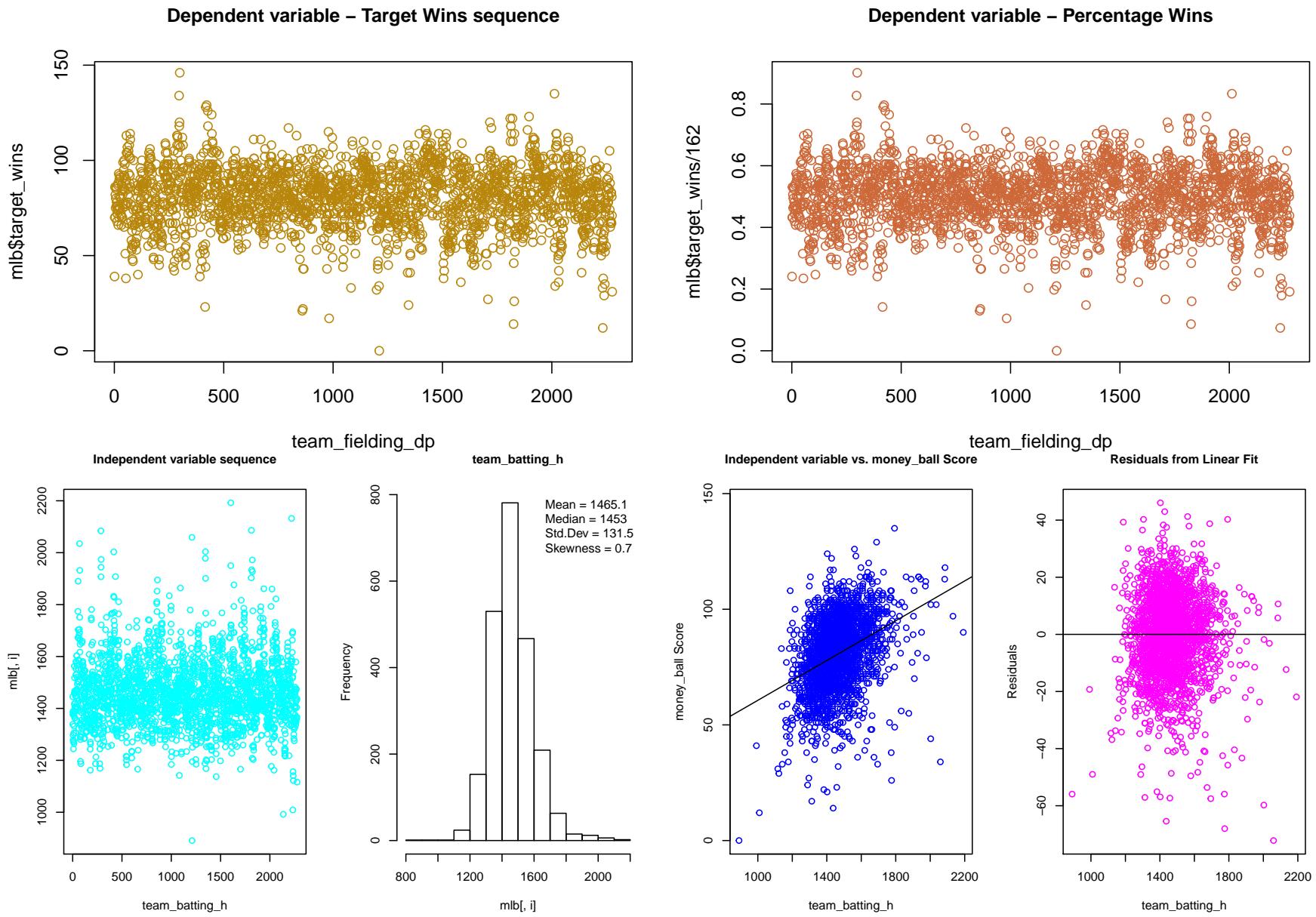
Analysis

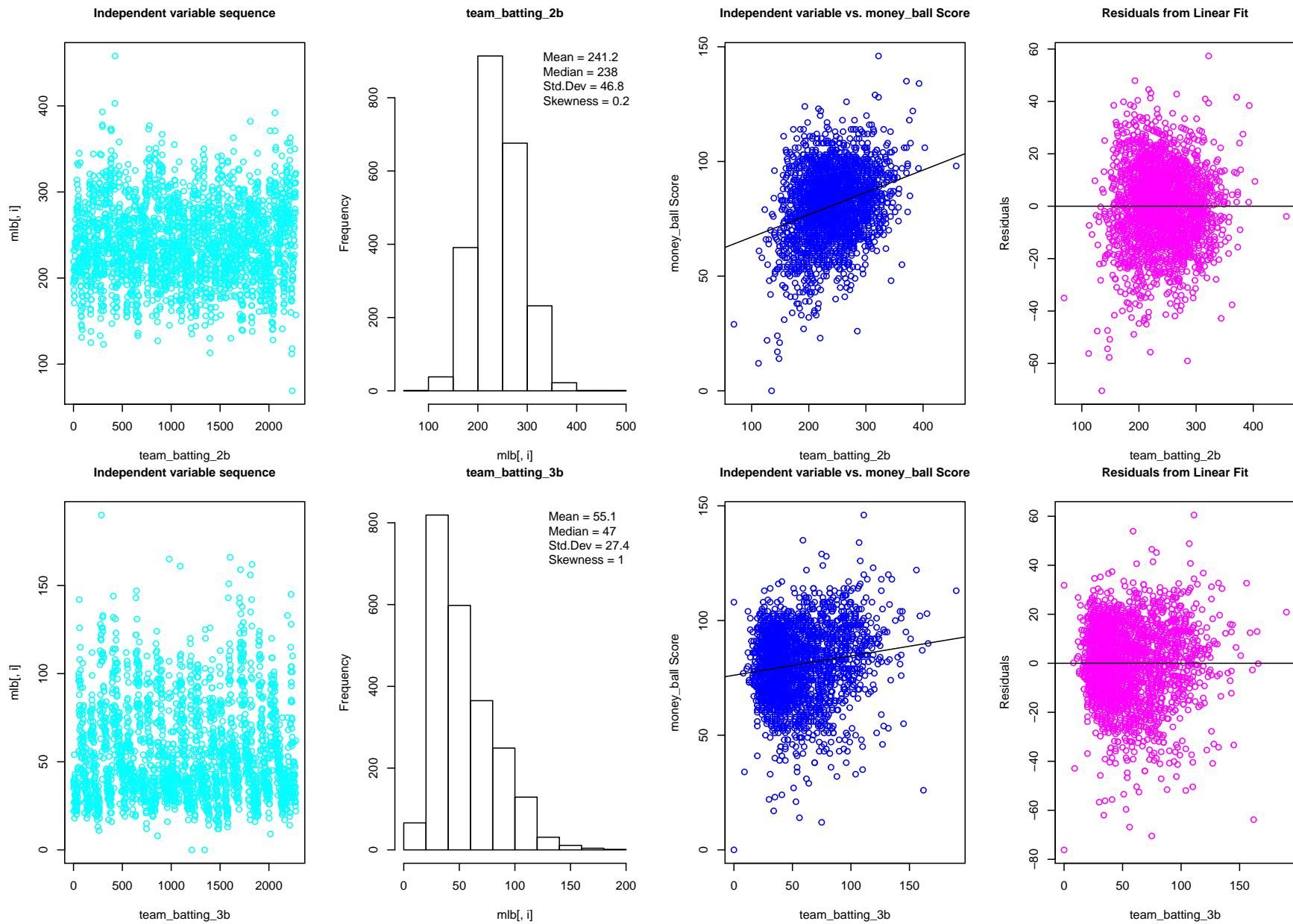
Interpretation

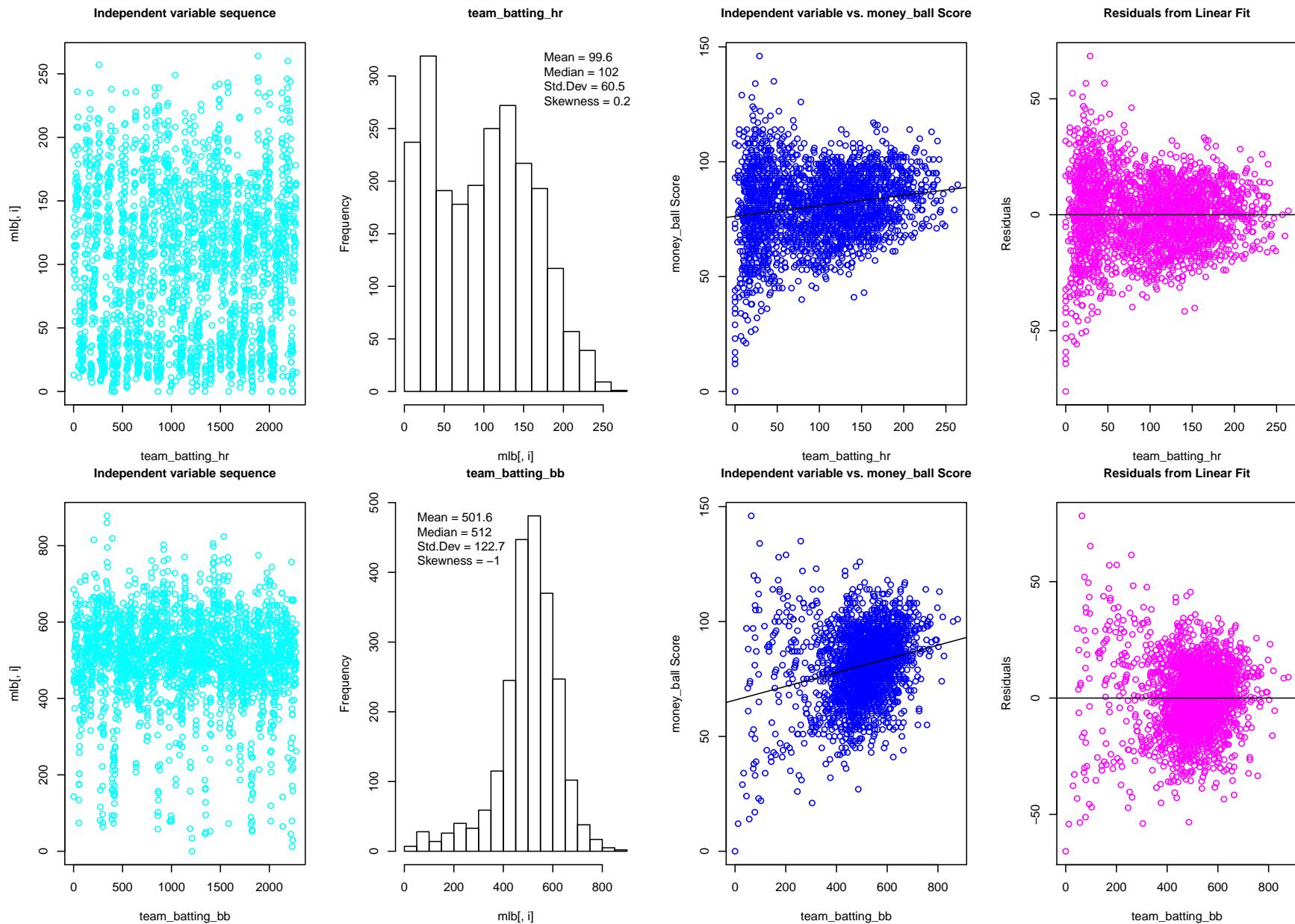
Conclusion

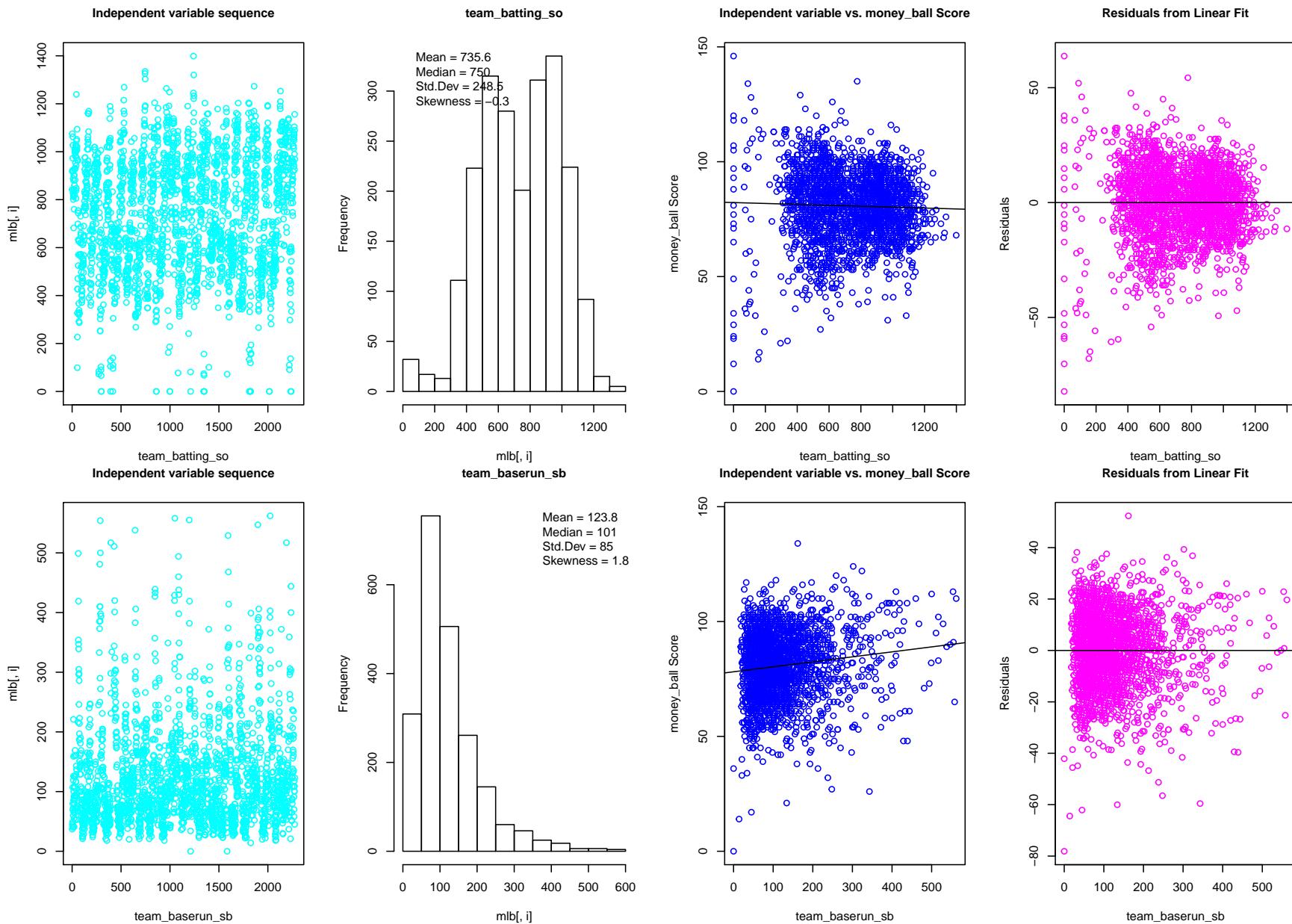
Setup some data directories

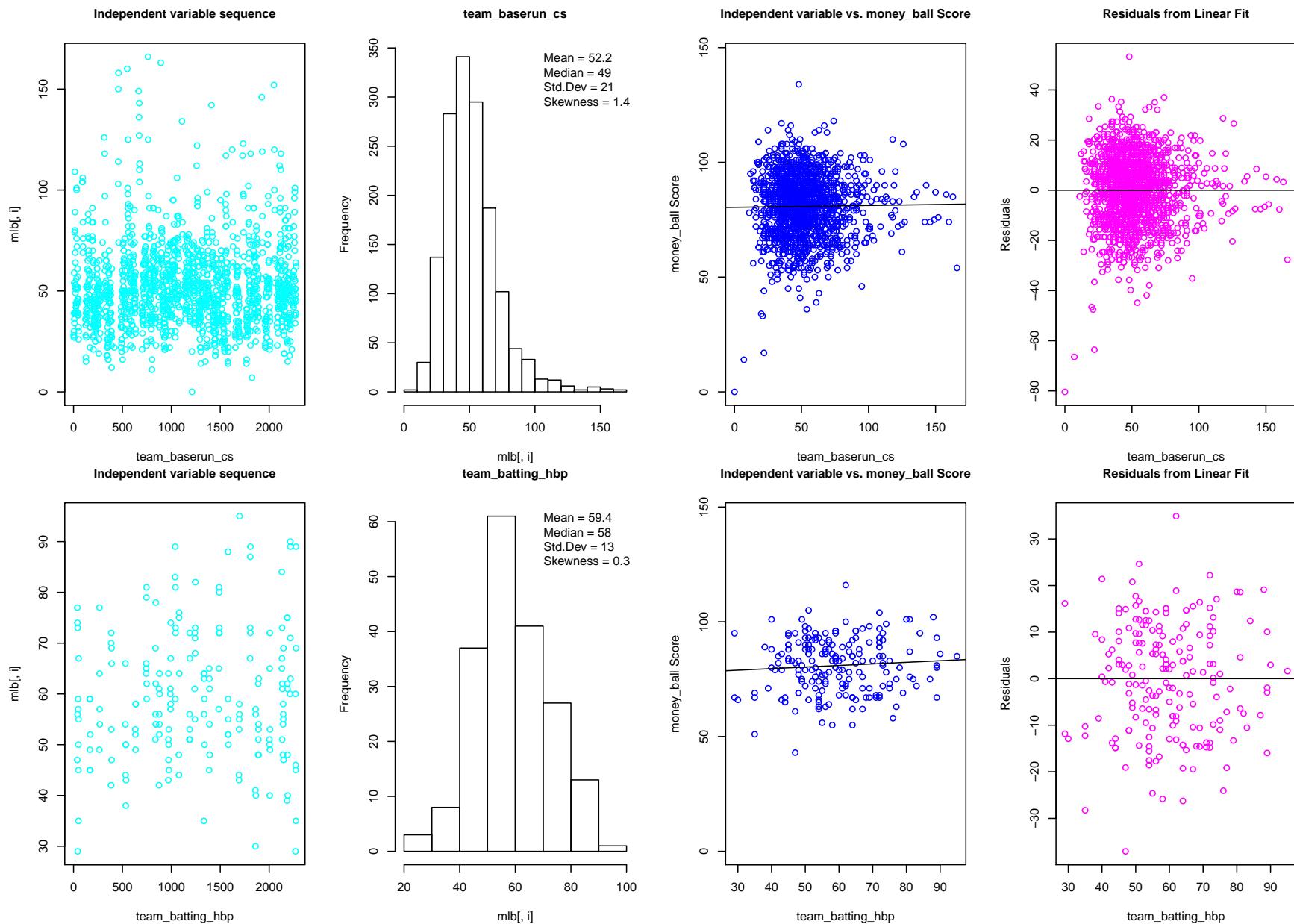
Remove extreme outliers from each column of numerical data

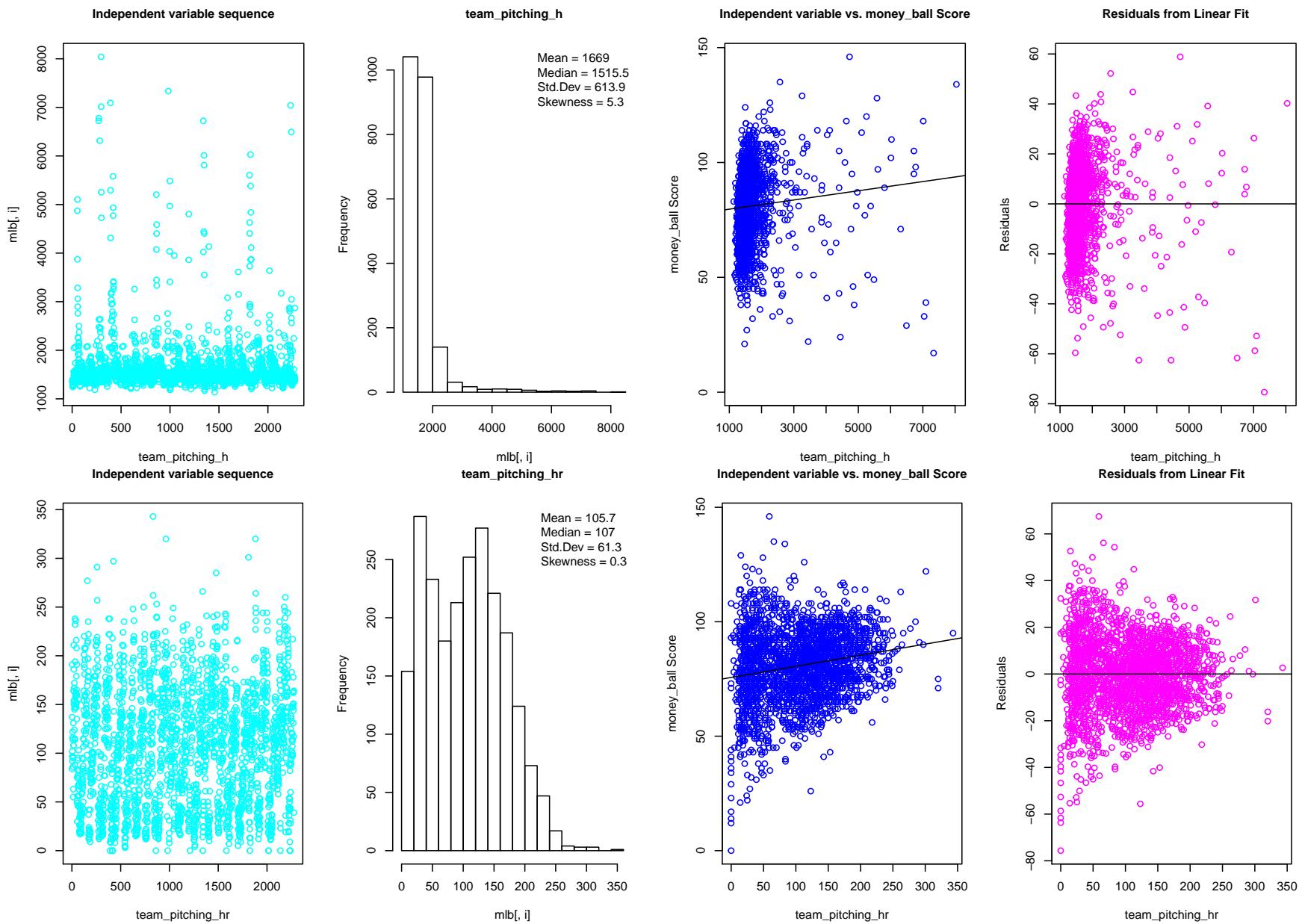


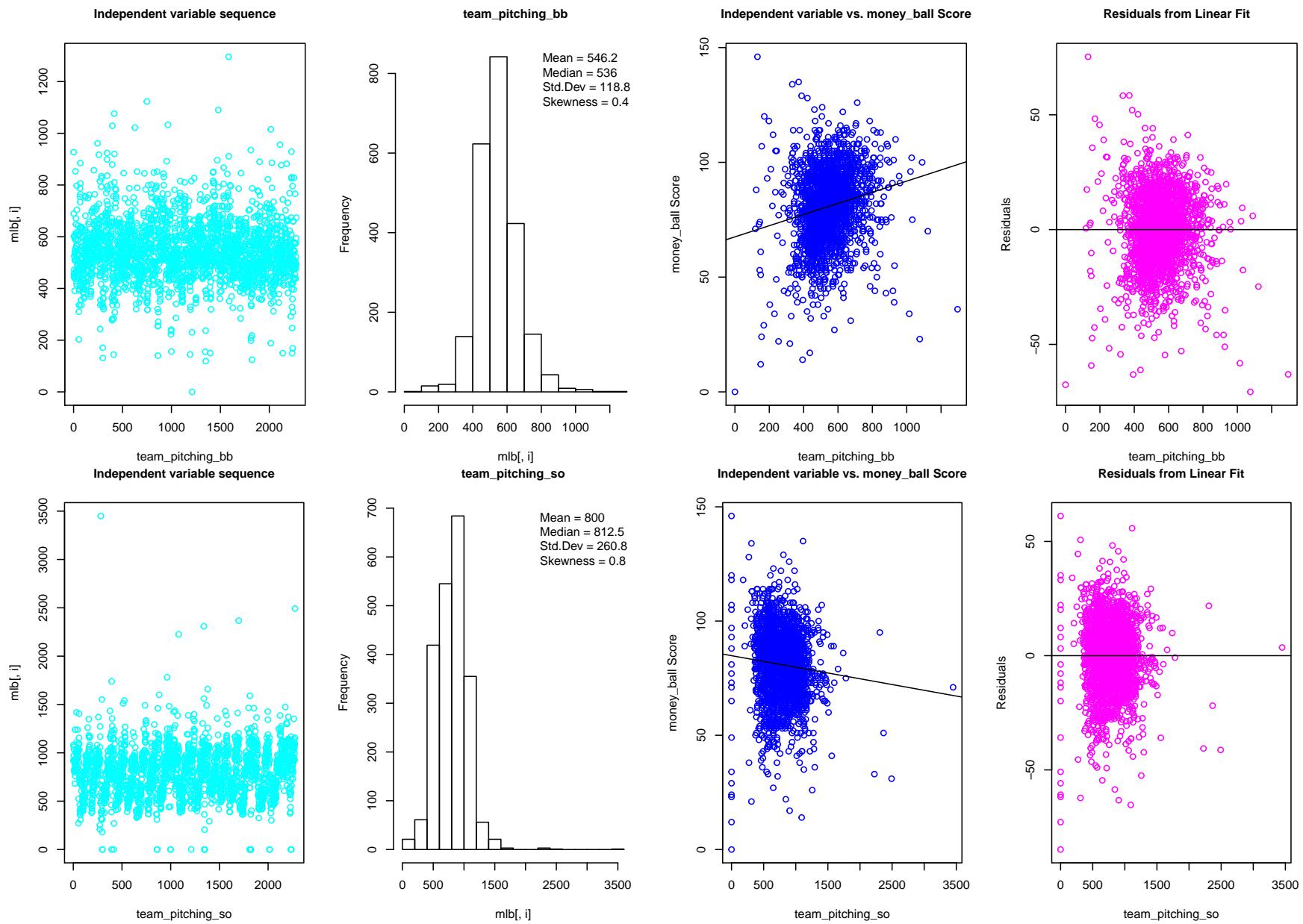


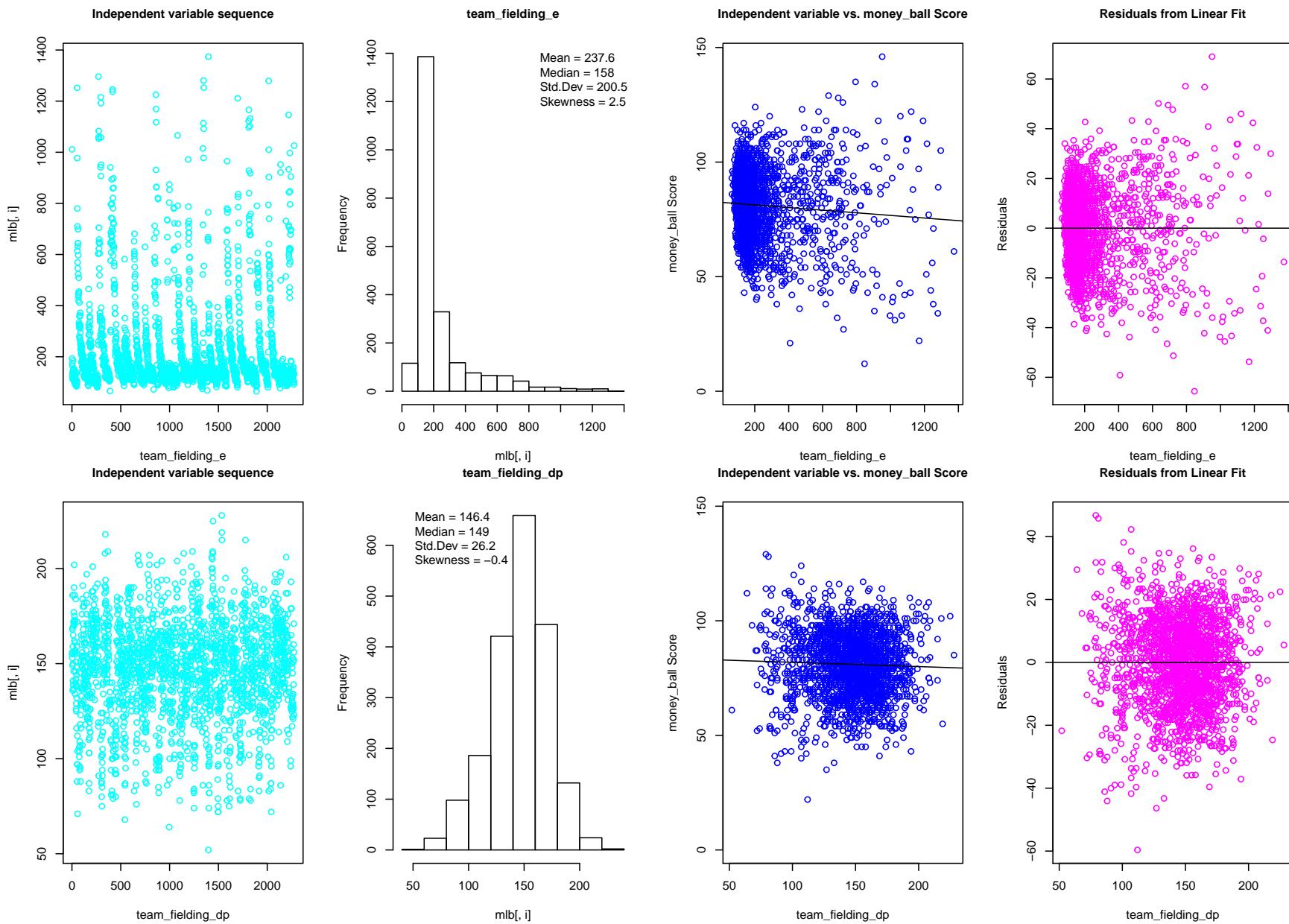


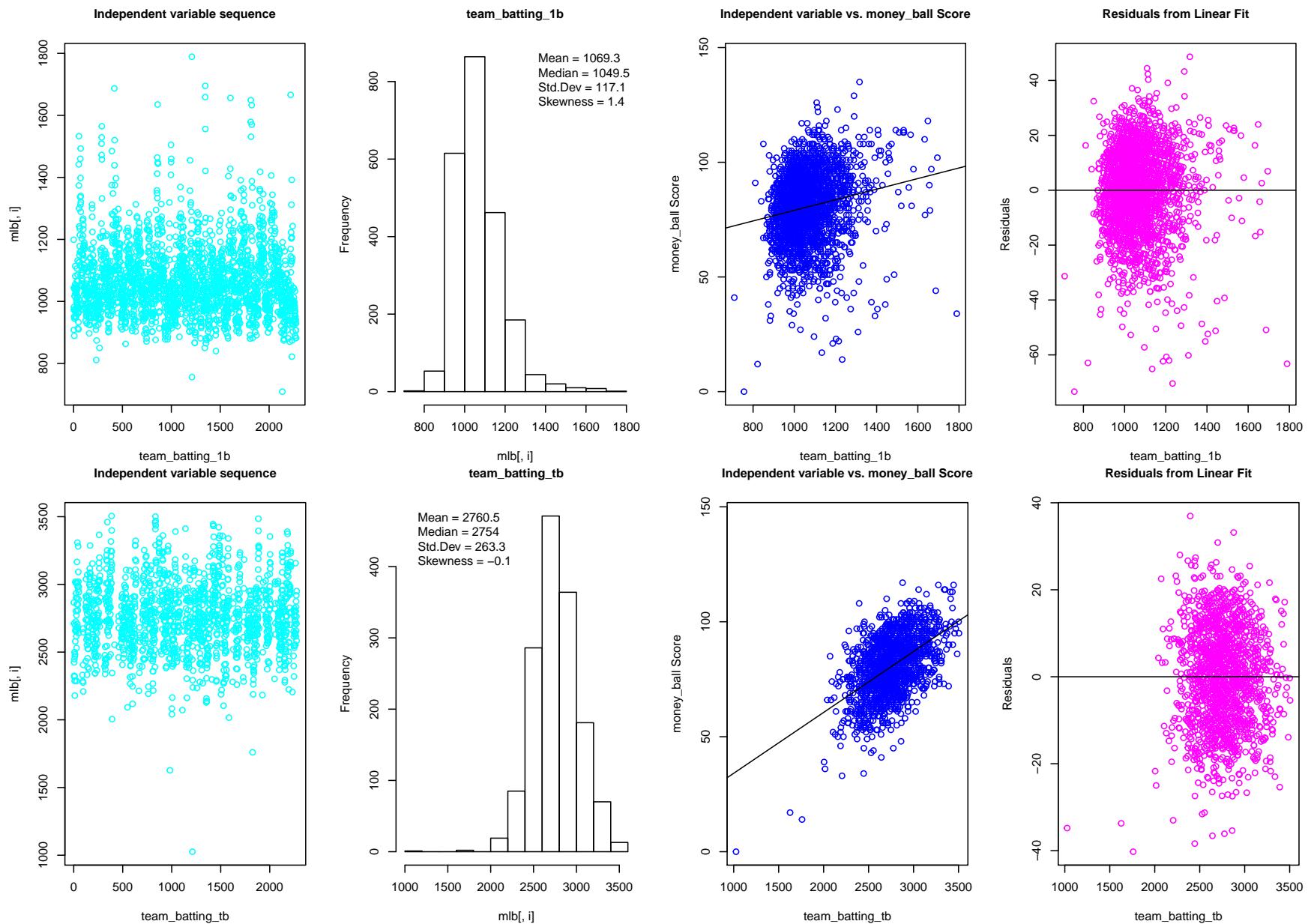






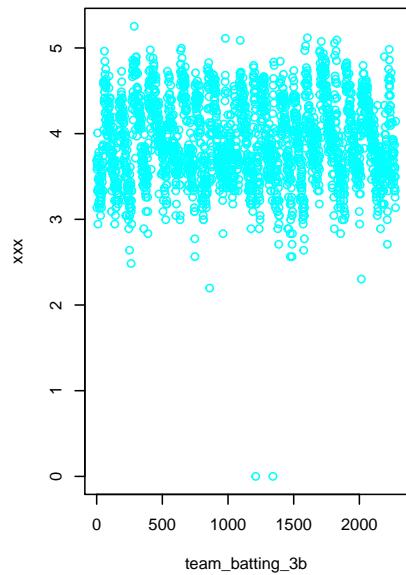




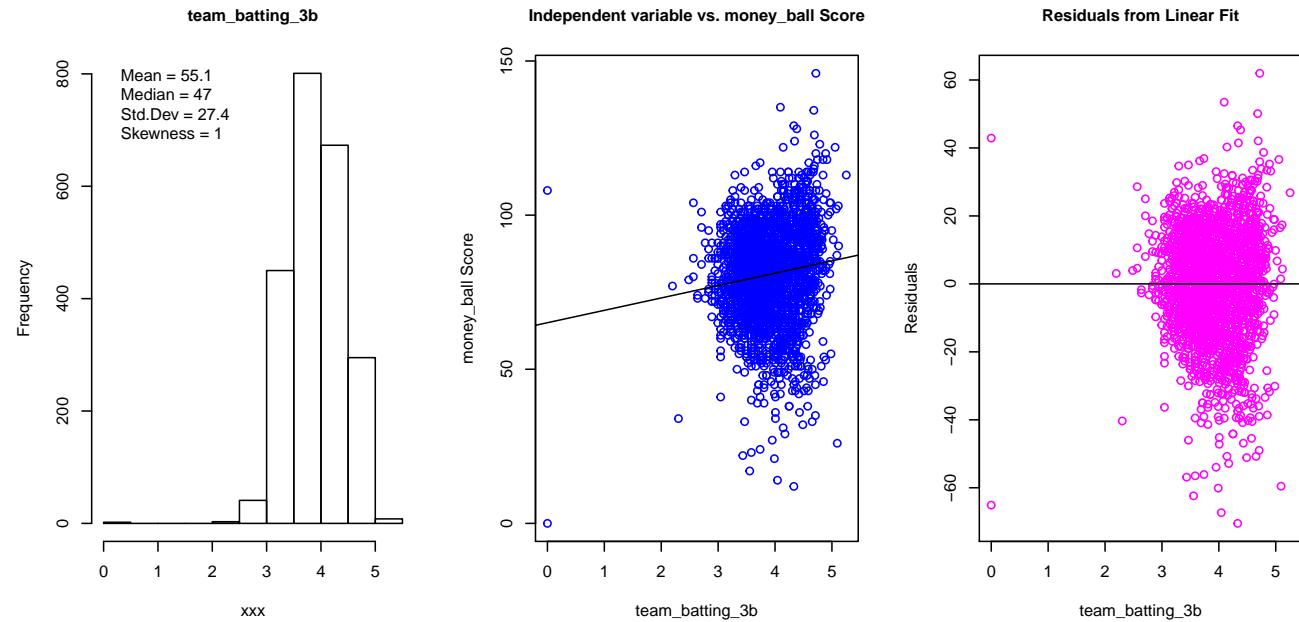


```
[1] 3  
[1] 4  
[1] 5
```

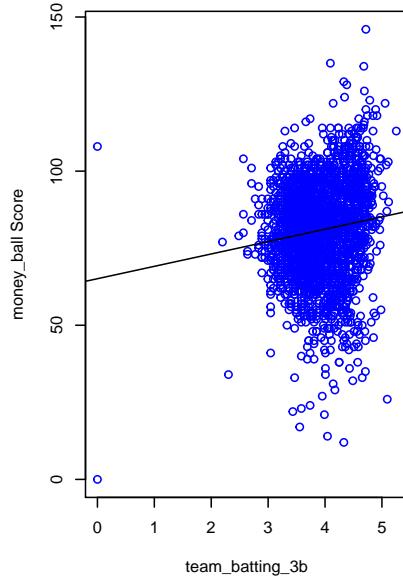
Independent variable sequence



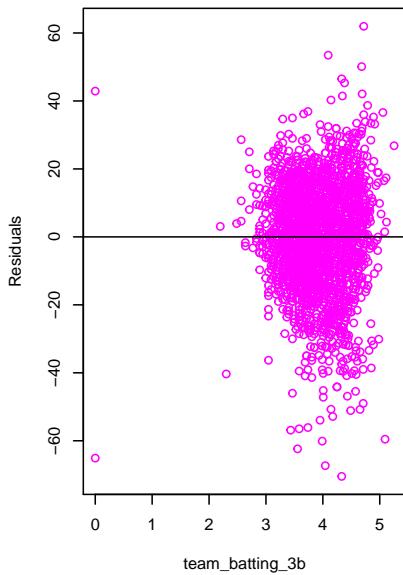
team_batting_3b



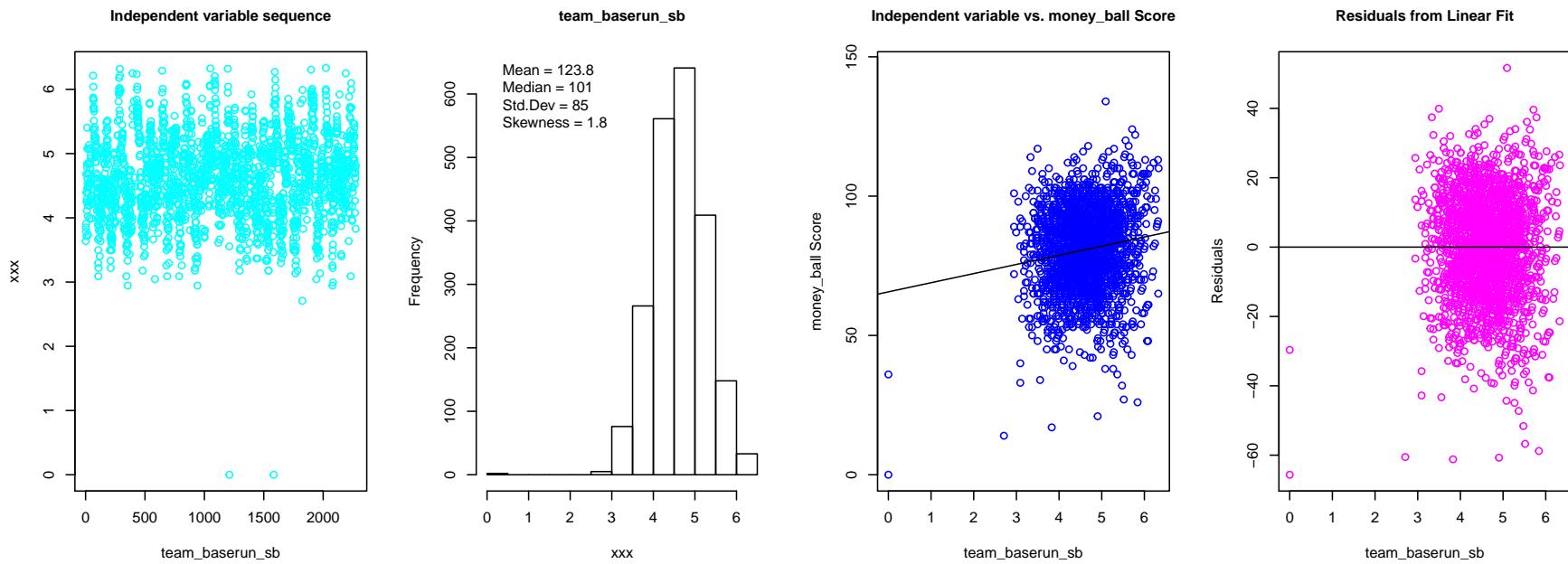
Independent variable vs. money_ball Score



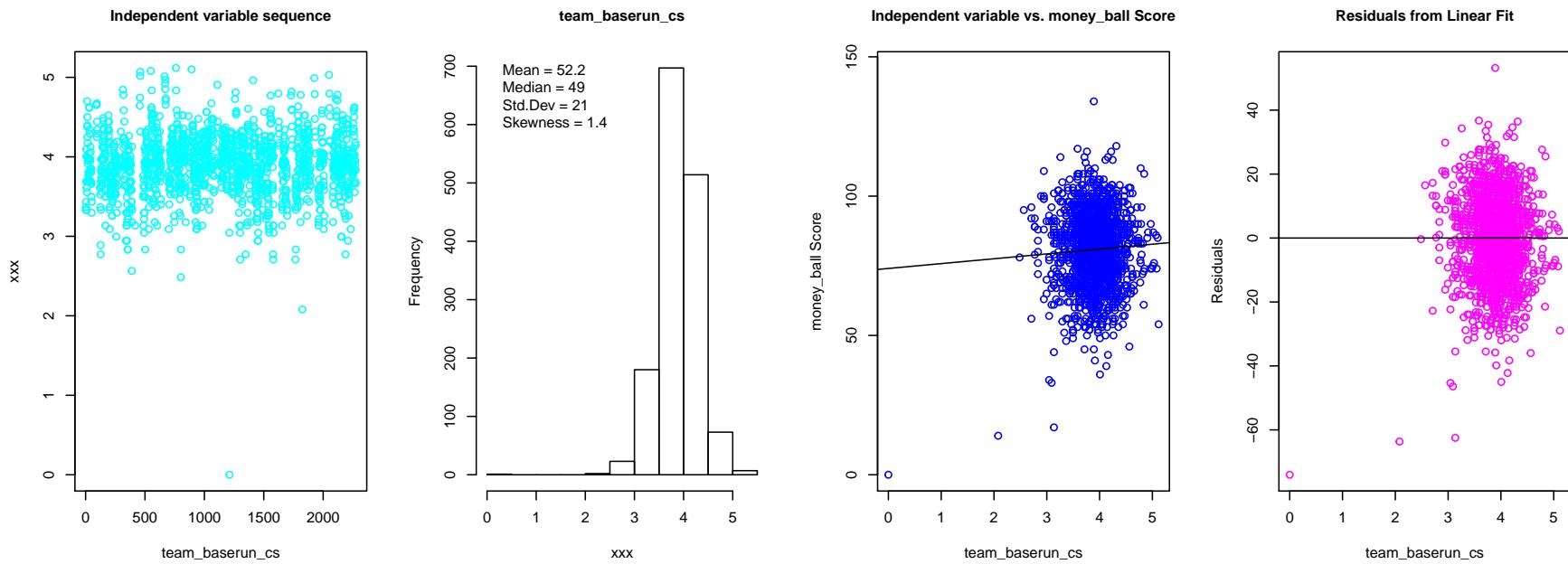
Residuals from Linear Fit



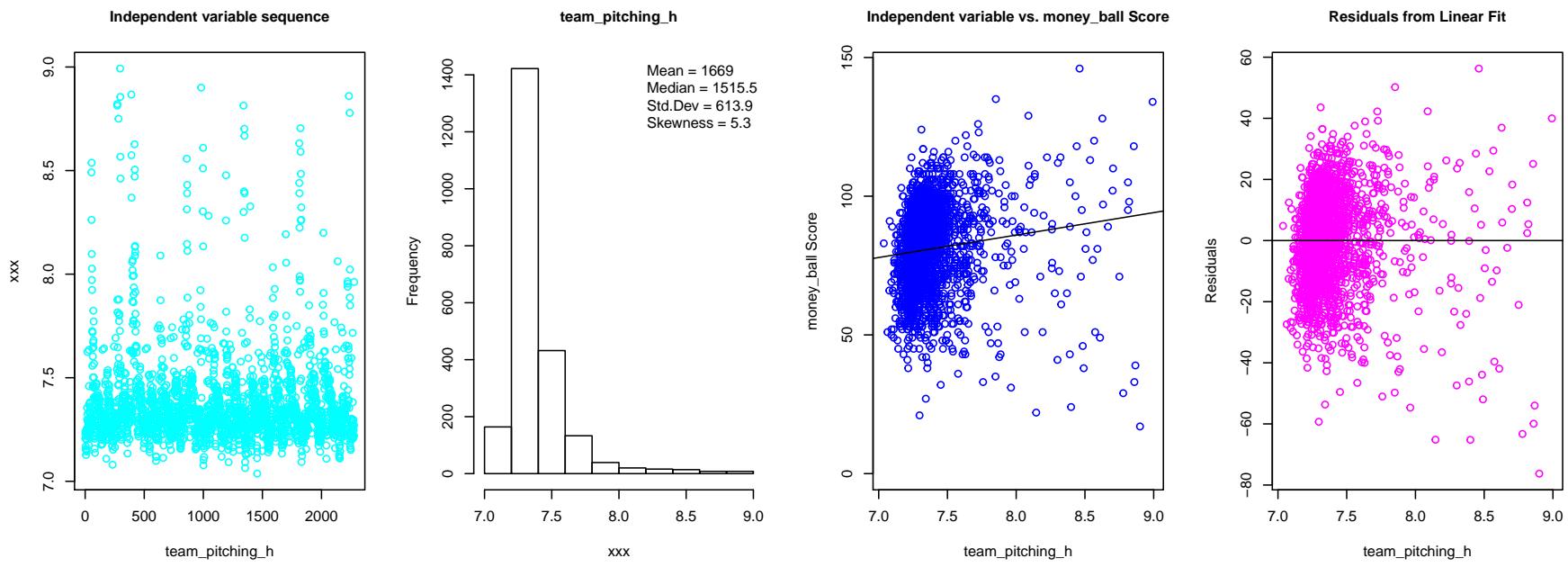
```
[1] 6  
[1] 7  
[1] 8  
[1] 9
```



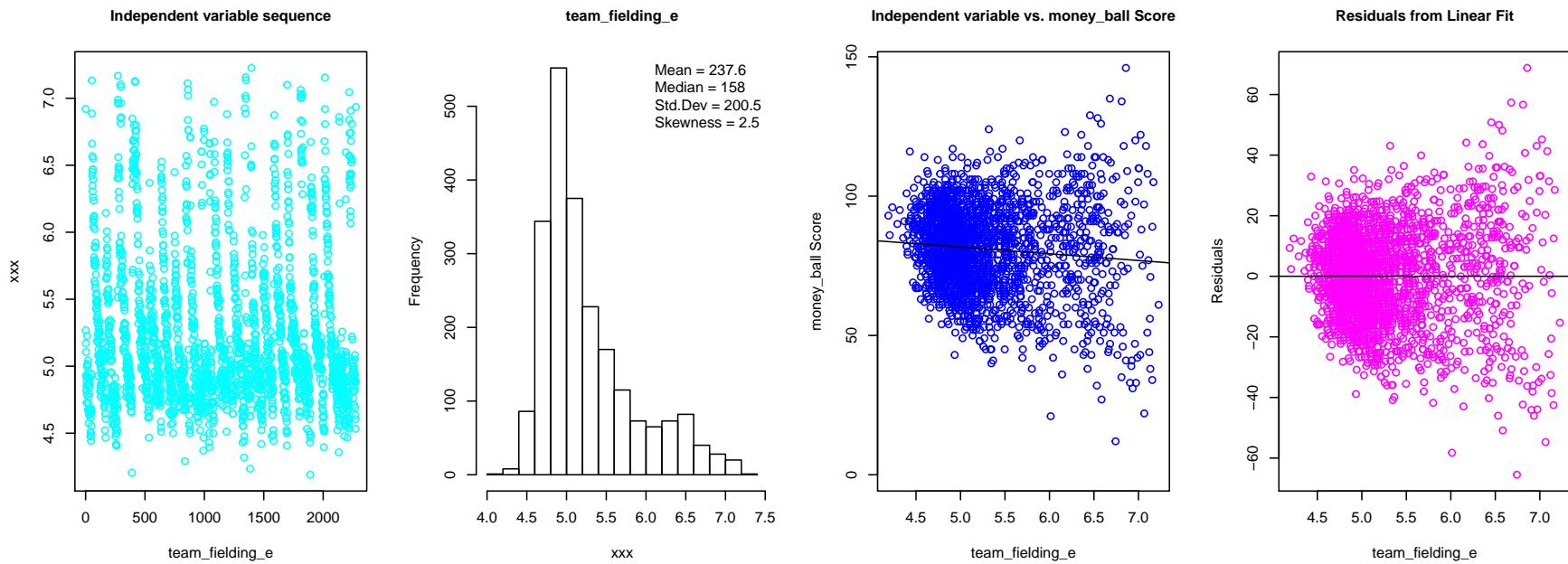
[1] 10



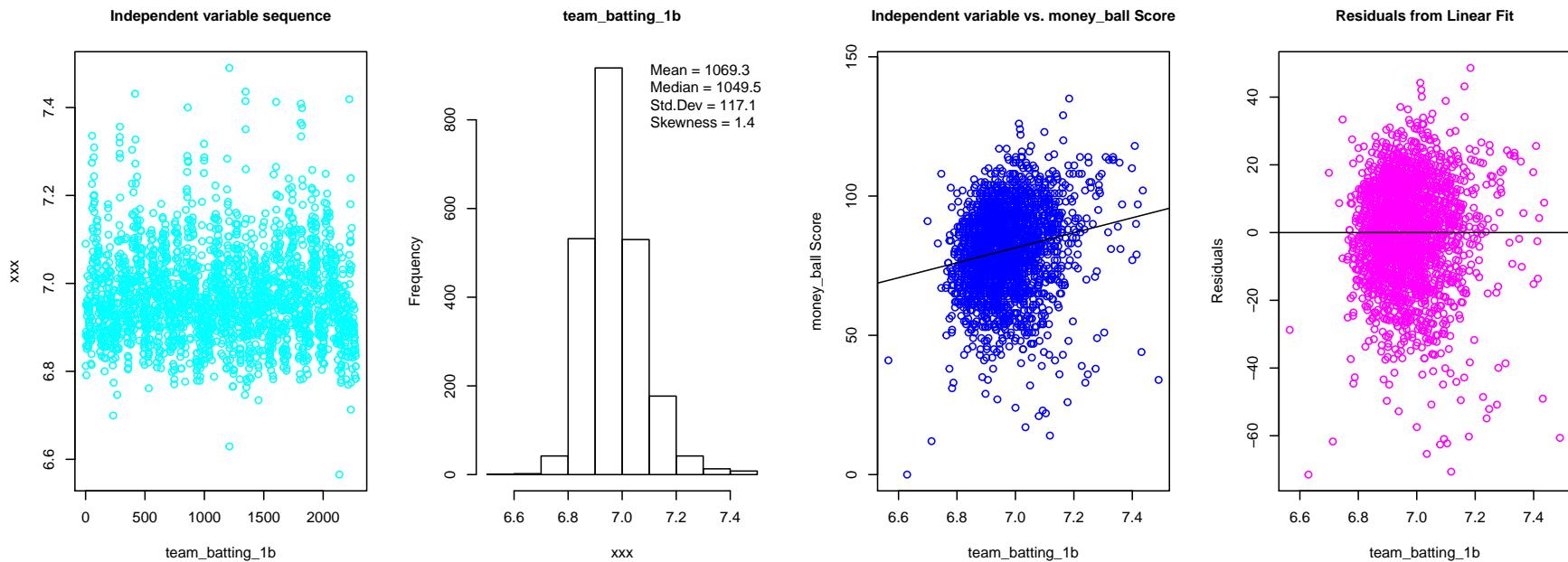
```
[1] 11
[1] 12
```



```
[1] 13
[1] 14
[1] 15
[1] 16
```



```
[1] 17
[1] 18
```



[1] 19

Make first data visualization plot for each column of dataset