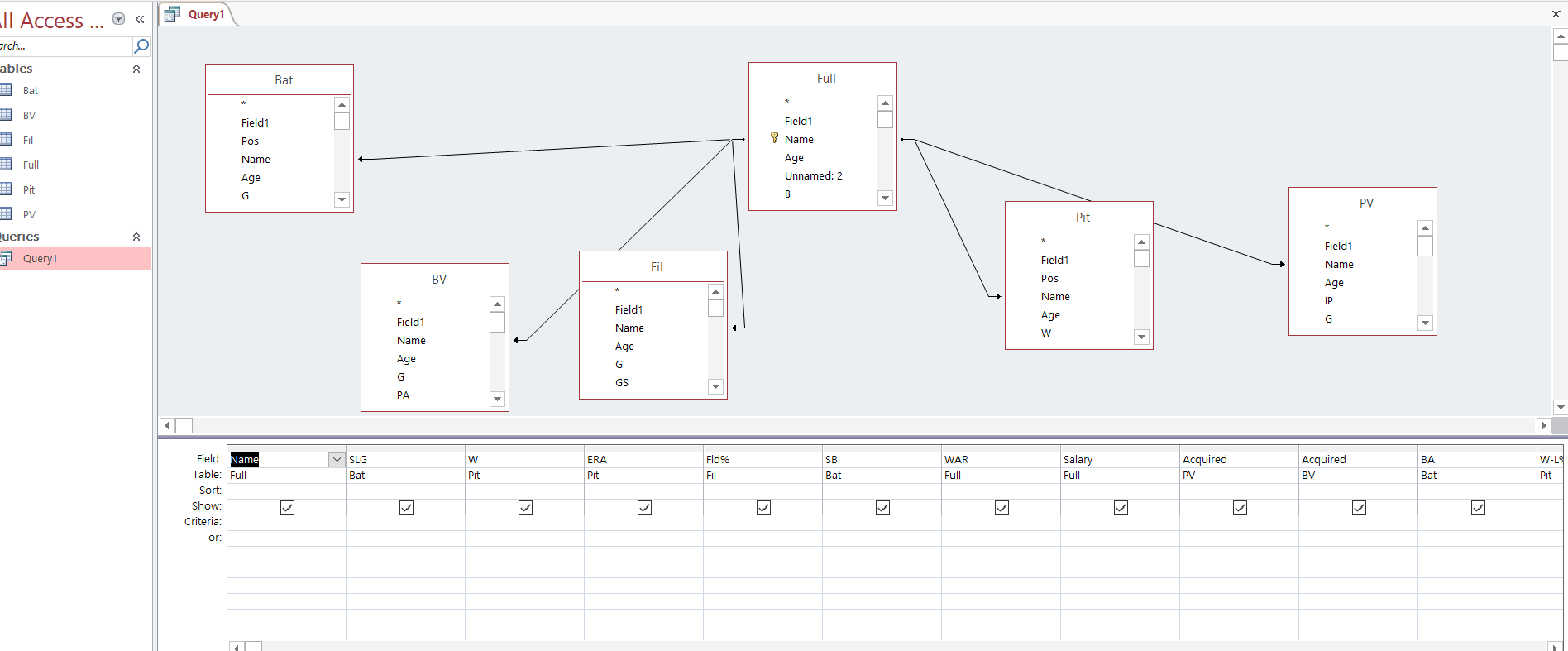
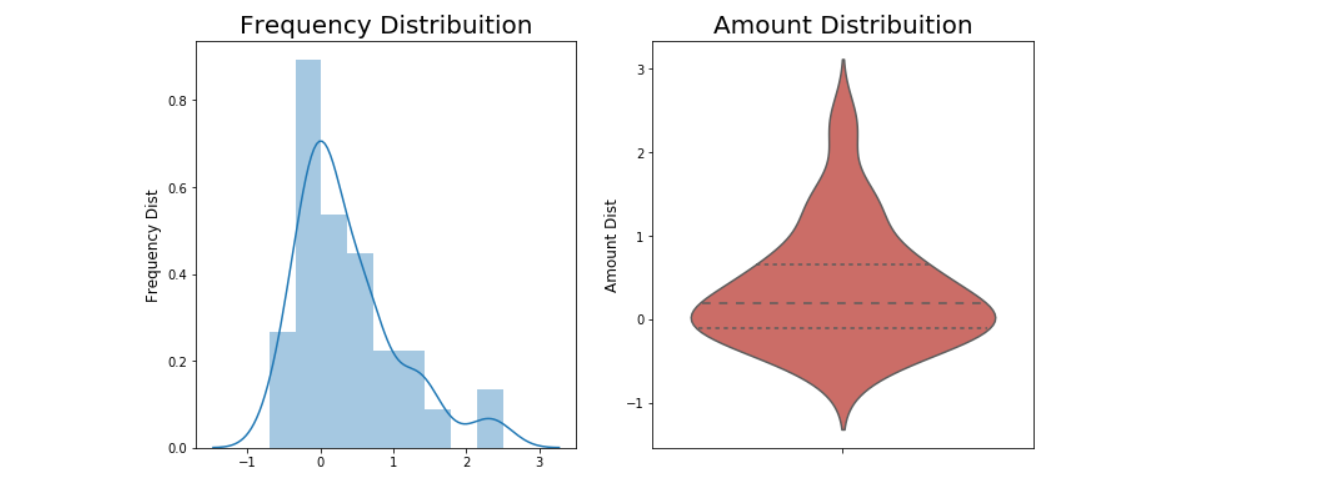
Create a new dataset by joining the table using an Access.

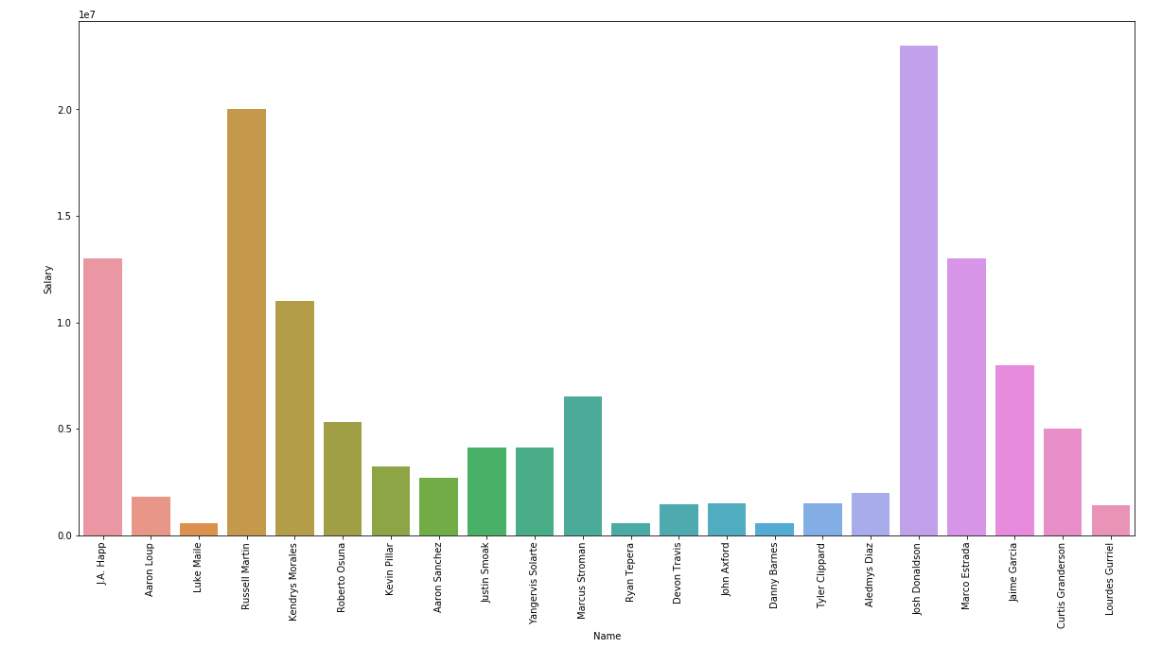


First of all, let's analyze the players' records.

WAR

Wins Above Replacement (WAR) is an attempt by the sabermetric baseball community to summarize a player’s total contributions to their team in one statistic. You should always use more than one metric at a time when evaluating players, but WAR is all-inclusive and provides a useful reference point for comparing players. Let's take a look at the distribution of data. You can see that there is a gathering around zero. 0-2 means reserve class level. We can feel right away that results were not good this season.



Let's look at the Salary graph.

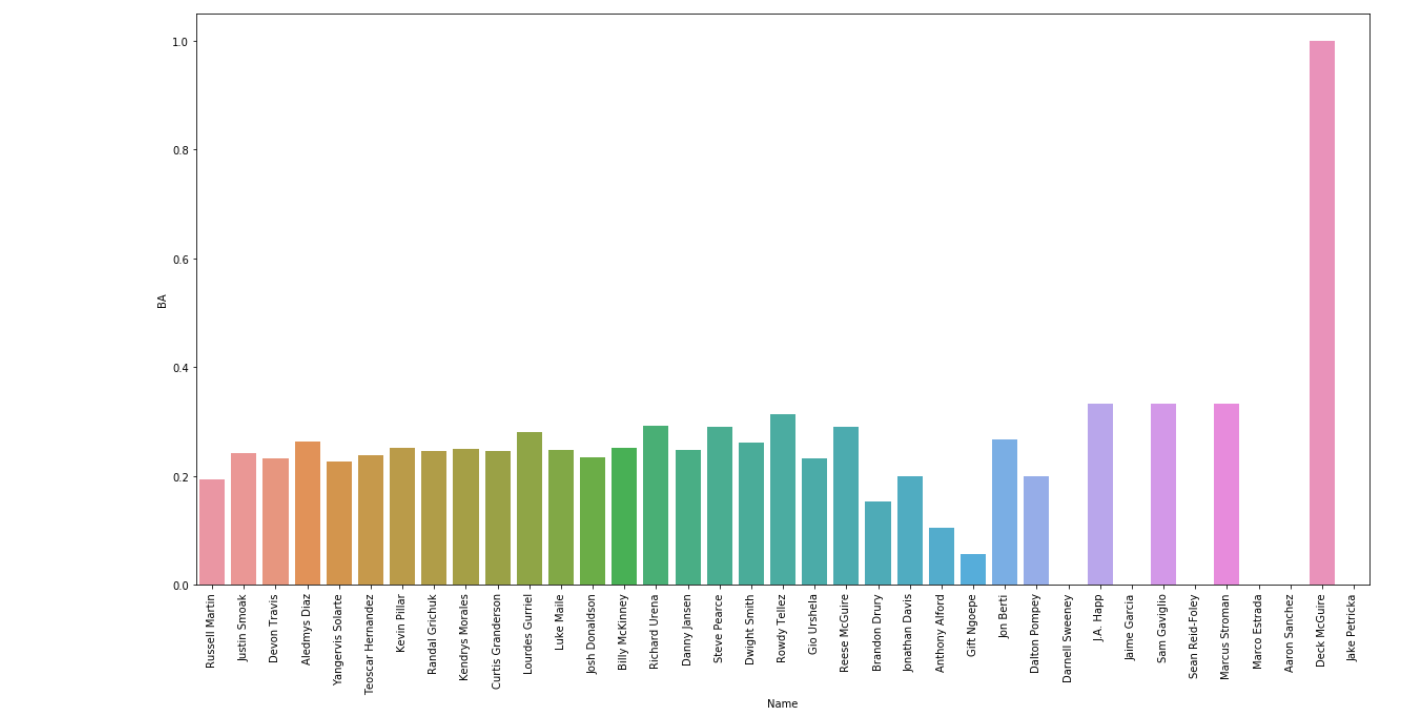
A high salary does not guarantee a good grade. This is the problem.



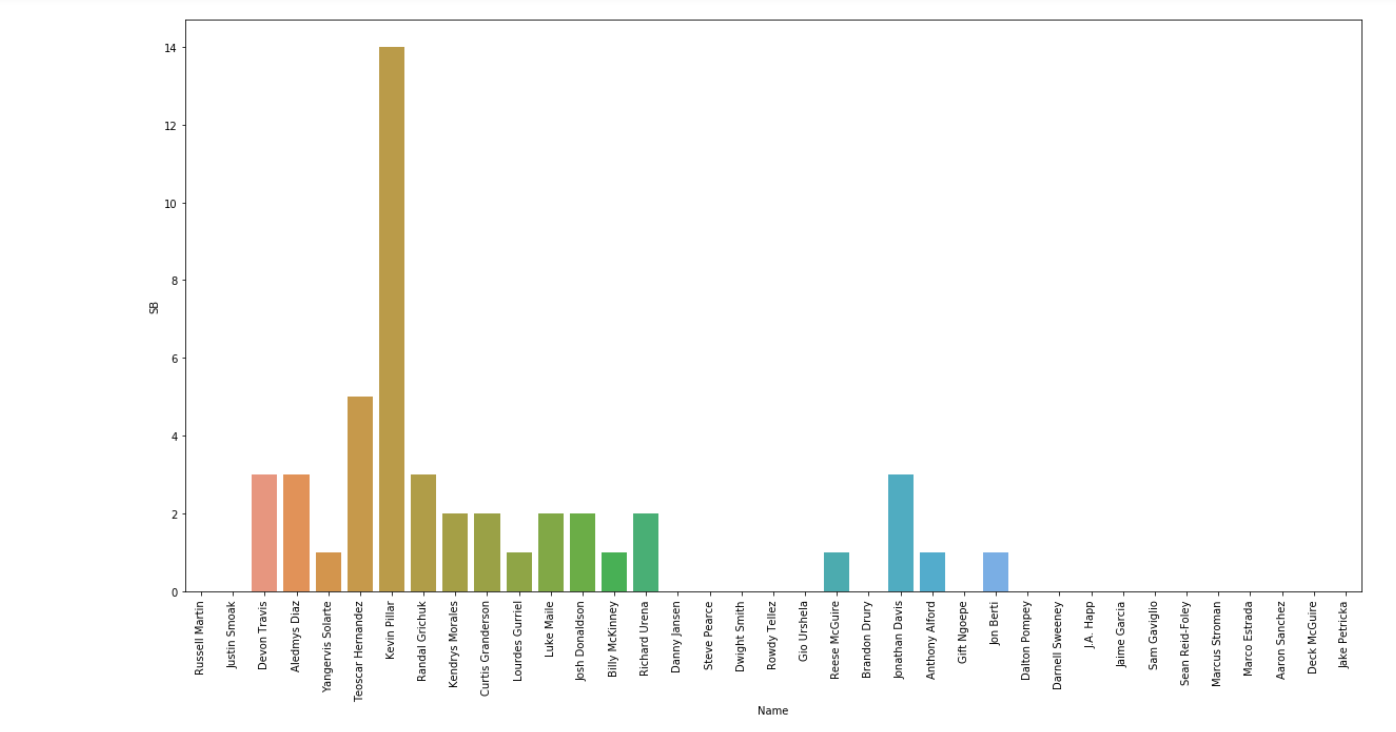
Let's analyze the batter first. This is batting average. Look at the Deck McGuire. A great batter came out.

But real is … 

He is only 1 Plate Appeances.

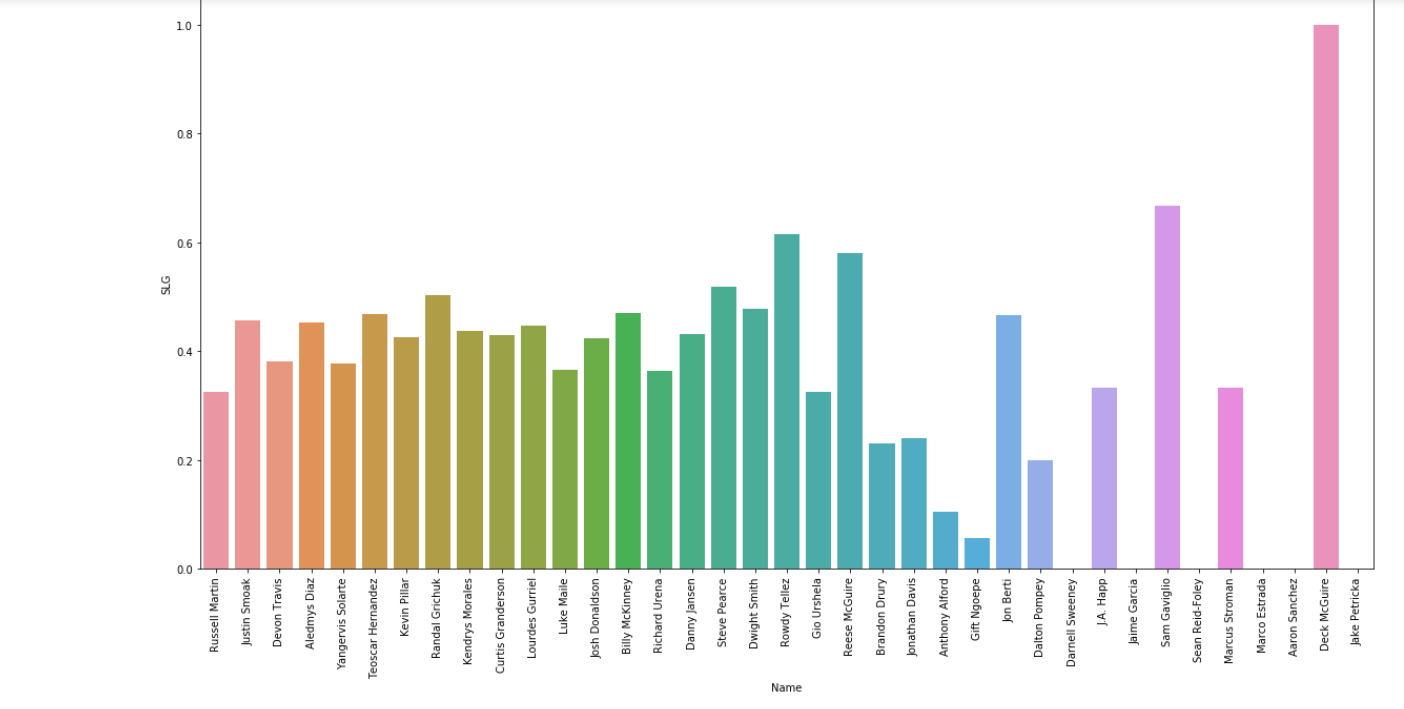


This is Stolen bases. Focus on Kevin Pillar. Why he is No7 hitter?

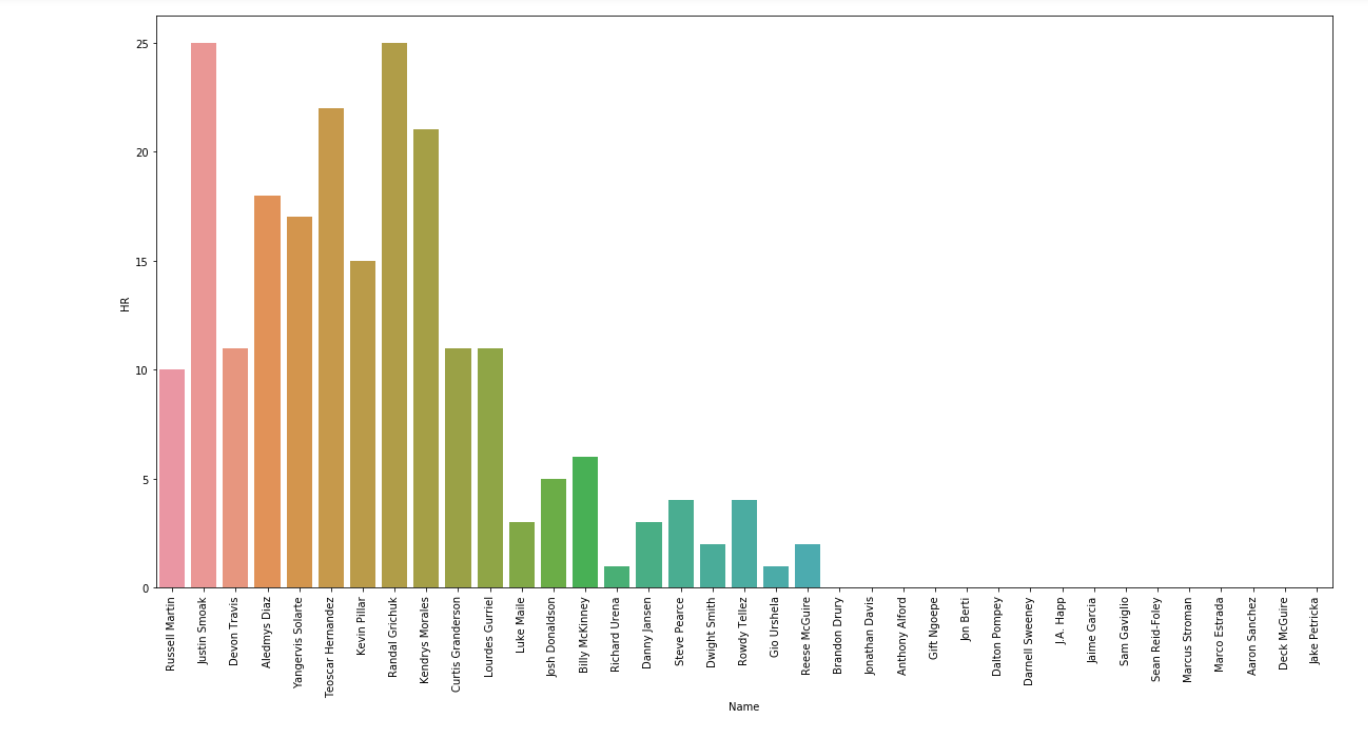


Let's see who's the slugger.

Let's not get confused with Deck McGuire again.

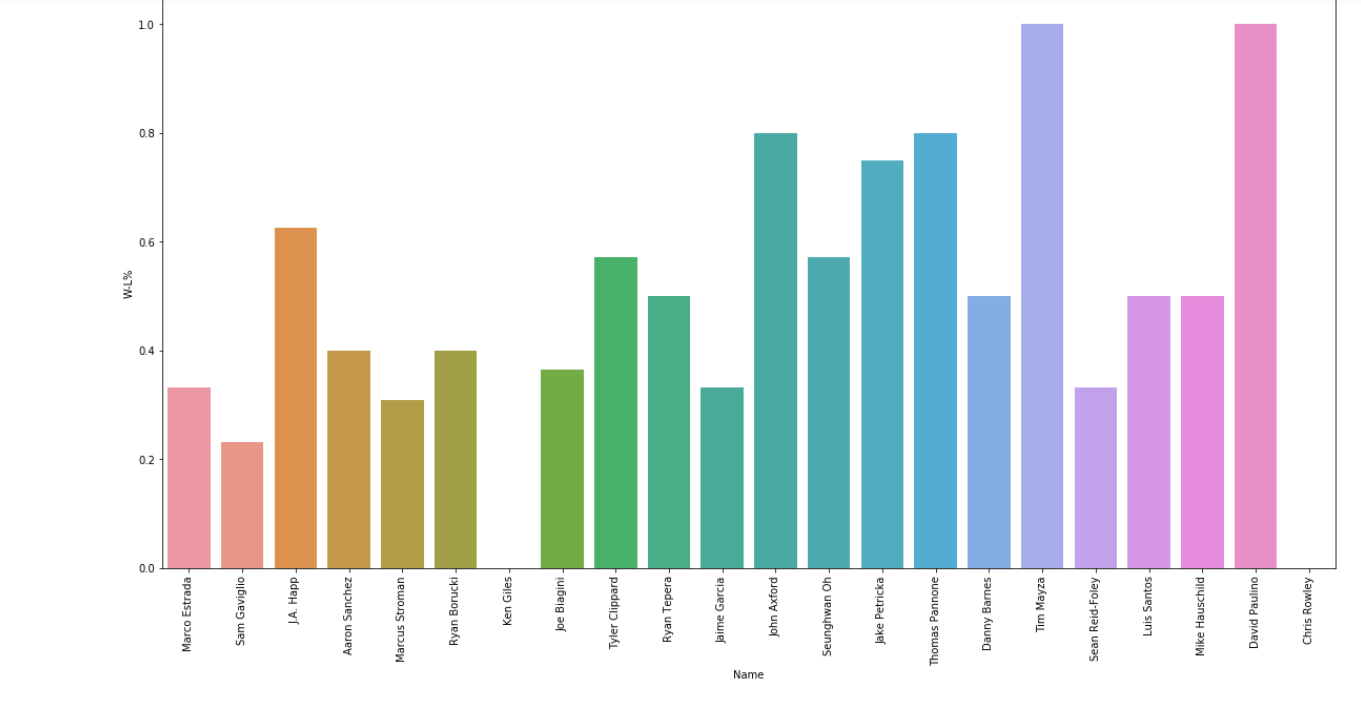


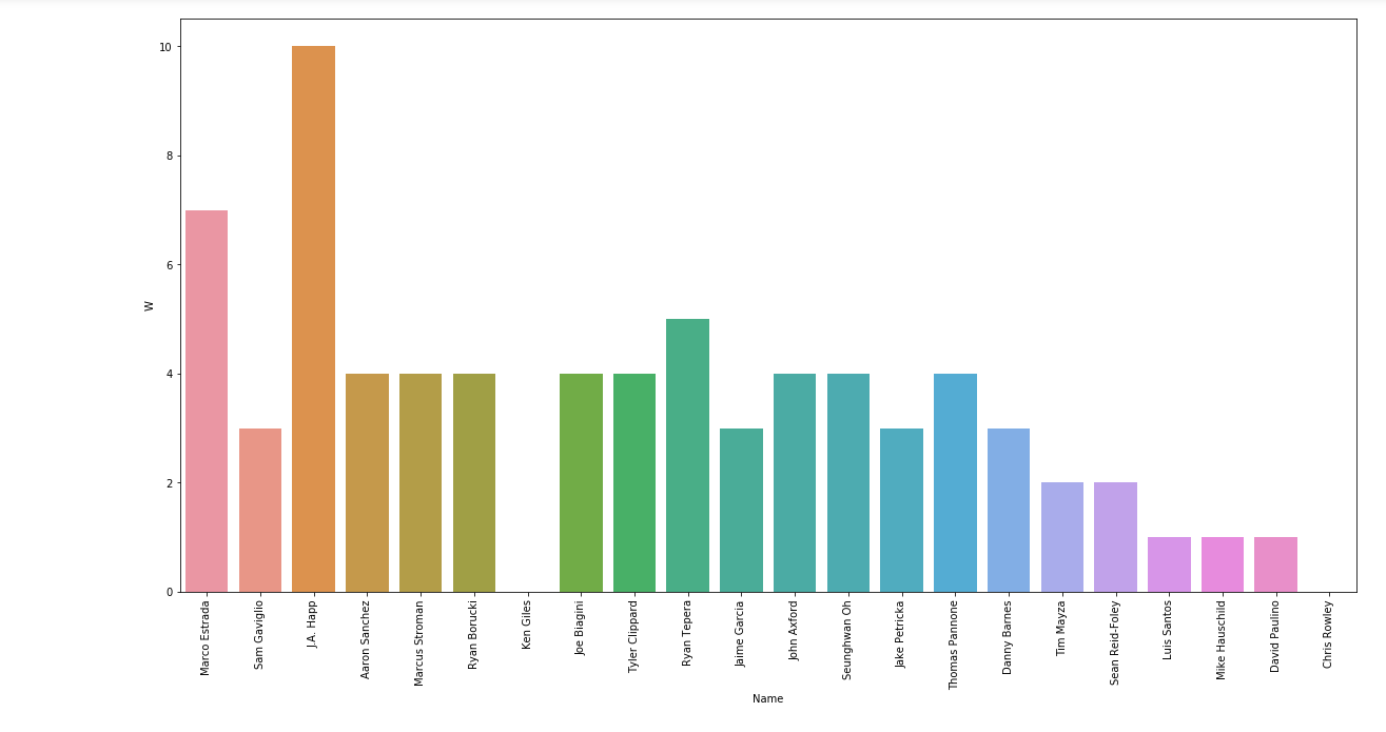
This is Number of HR. Justin Smoak, Randal Grichuk is share the lead. They are No2, 8 hitter..



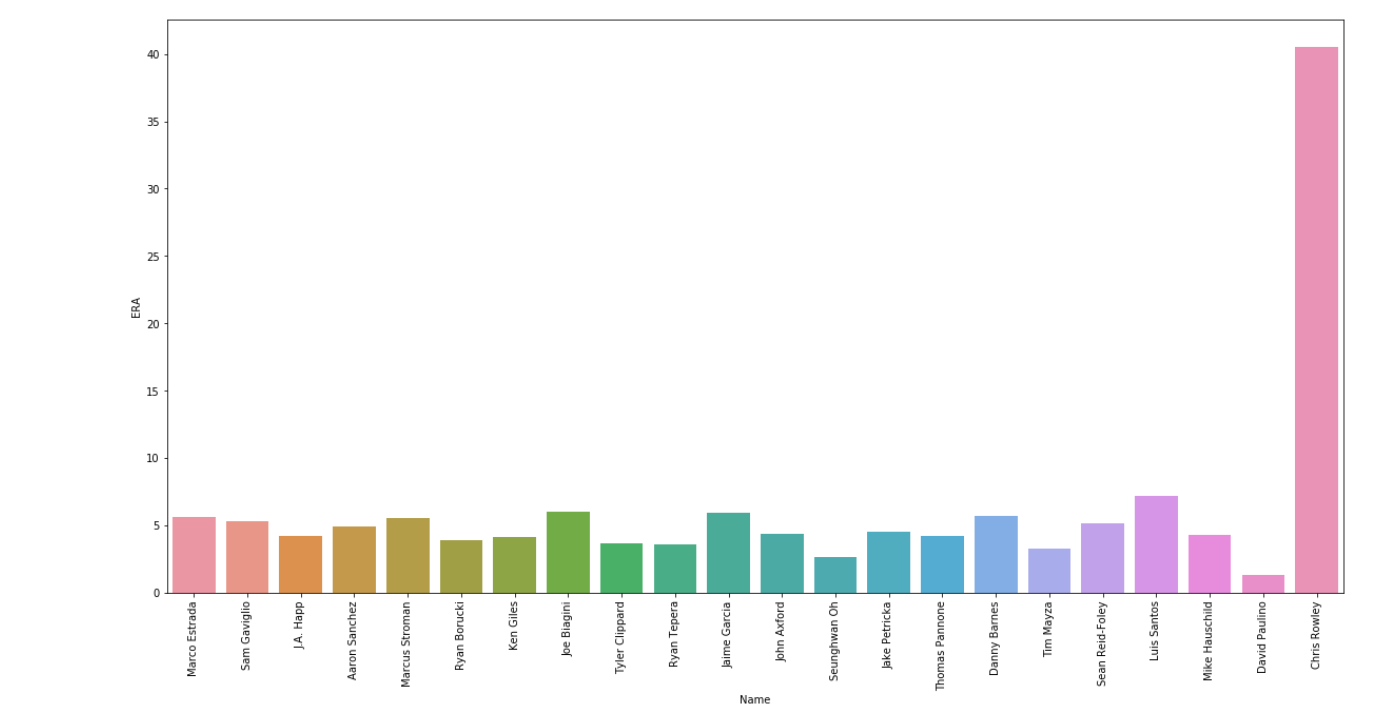
Now time to see pitch. Let’s see Win Rate(%) first. Tim Mayza, David Paulino has 100% win rate.

We have to check how many inning pitched. Tim Mayza (37 IP), David Paulino(7 IP).

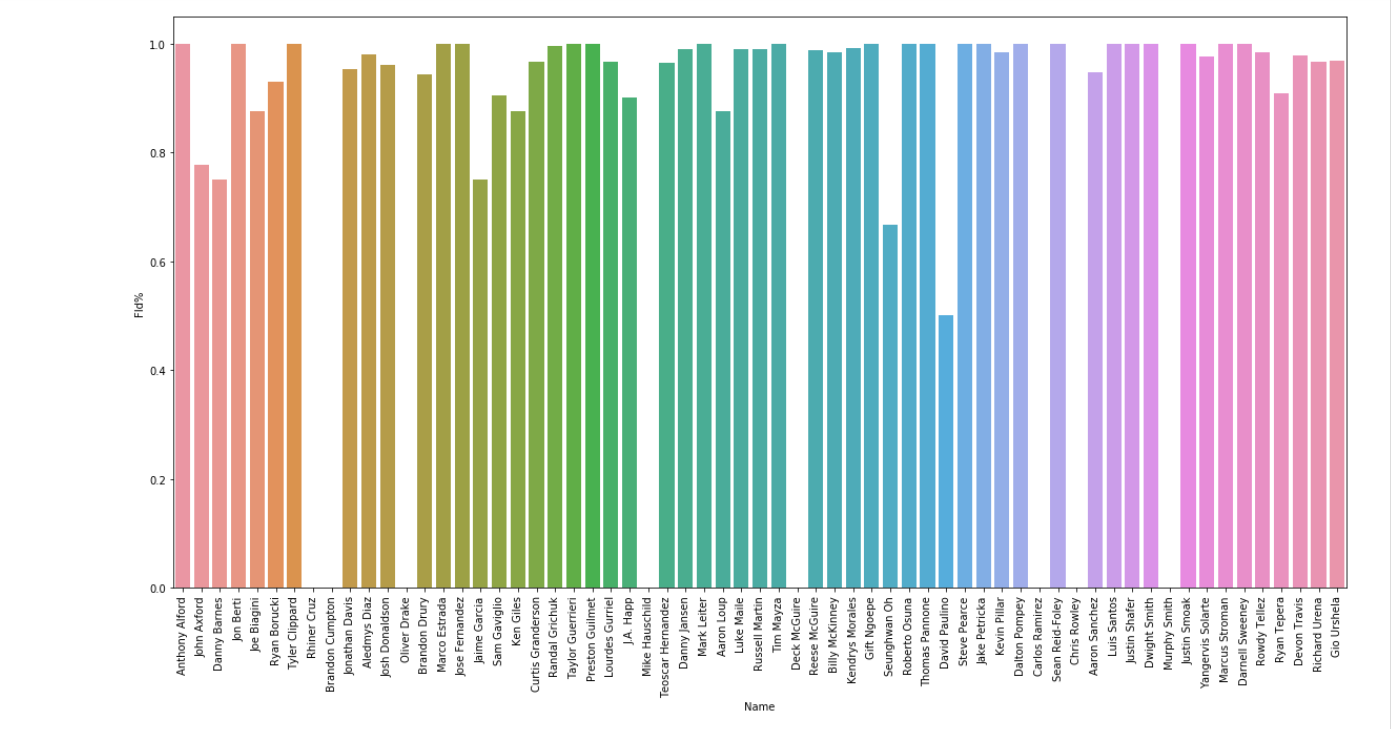


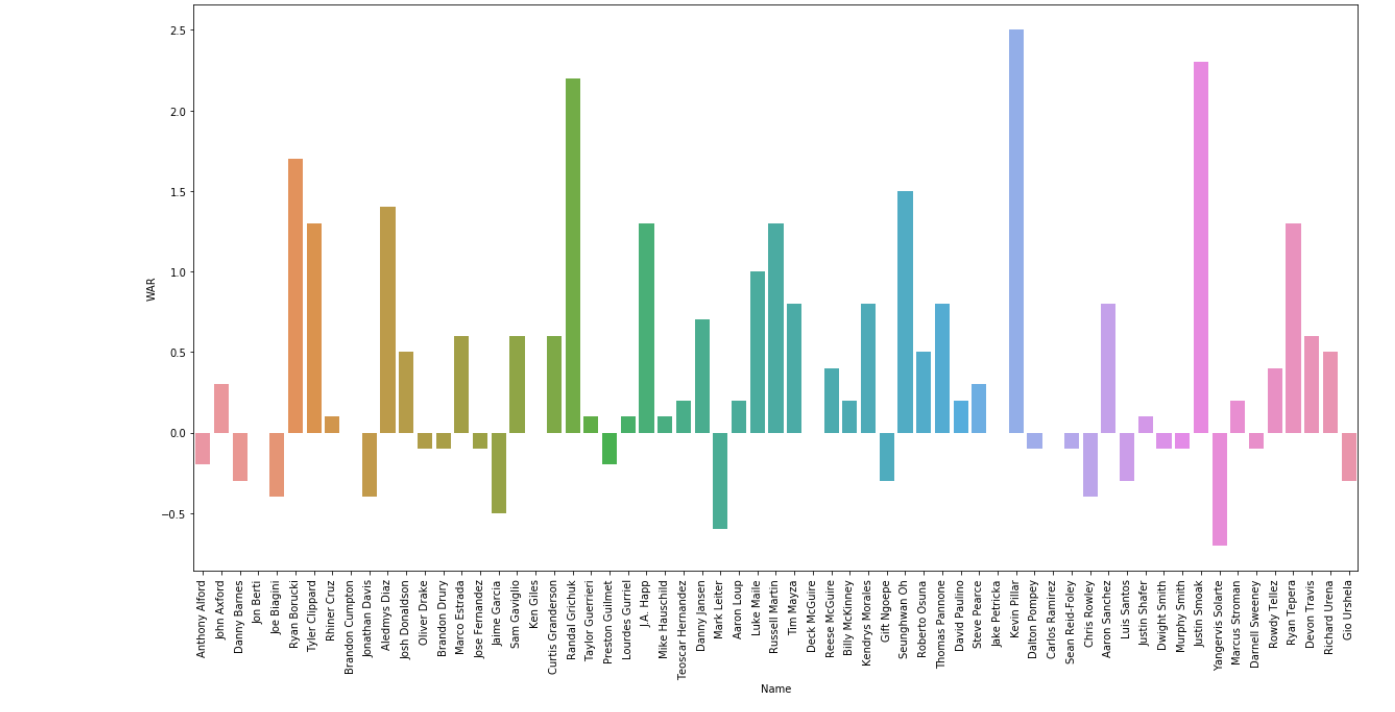
Look at the number of win. J.A. Happy led the team with 10 wins. 

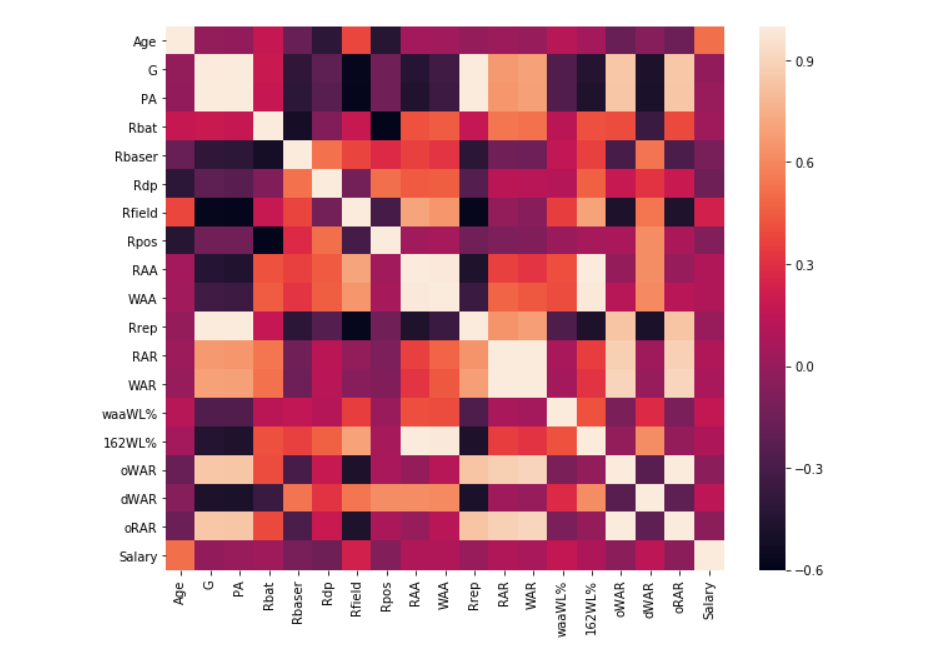
Let’s see EAR. earned run average (ERA) is the mean of earned runs given up by a pitcher per nine innings pitched. One of the best important factor of pitch. The ERA is mostly in the 5-point range and there is also 40 point. Most of the time, they lost five points per game. This seems to be the biggest problem of poor team performance.

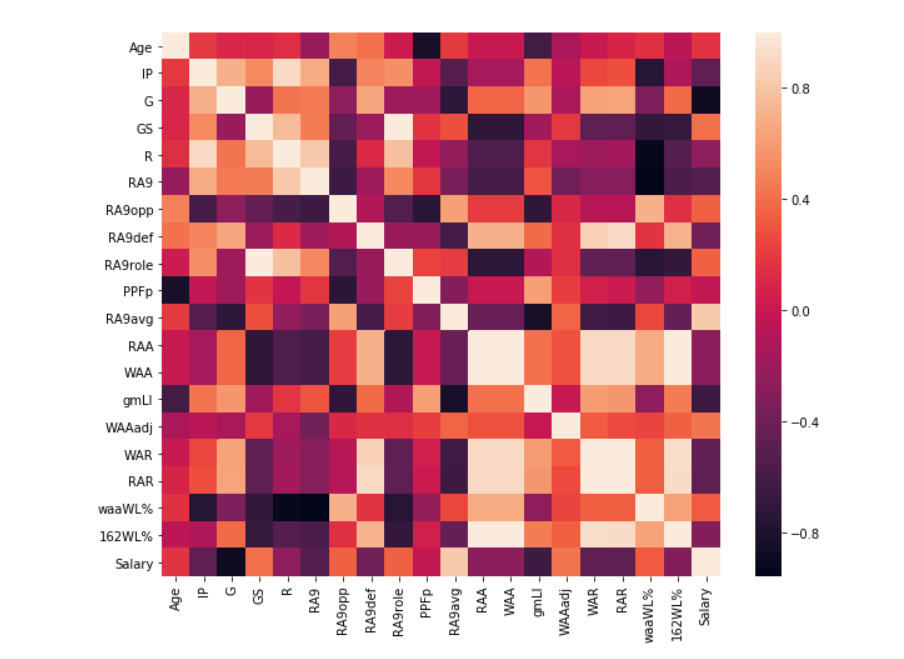


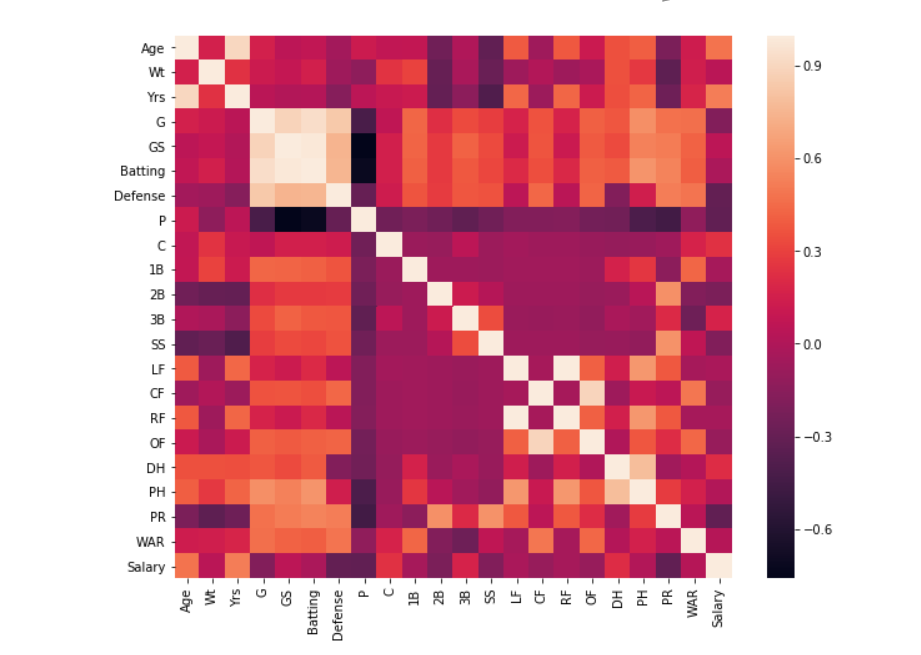
Let's look at the Fielding record. FLD% is fielding percentage, also known as fielding average, is a measure that reflects the percentage of times a defensive player properly handles a batted or thrown ball. It is calculated by the sum of putouts and assists, divided by the number of total chances (putouts + assists + errors). There are several frequent mistakes, but most of them are not.



We were a simple look at the WAR distribution above. There are no players over 2.5. Even if there are no MVP levels of 8 or higher, there are few 2+ athletes considered to be basic starting members.

Let's look at the relationship between batting records and salary. Surprisingly, no record this season has anything to do with salary except that an older player has a high salary relatively. This means that all the high-salarying players were sluggish.

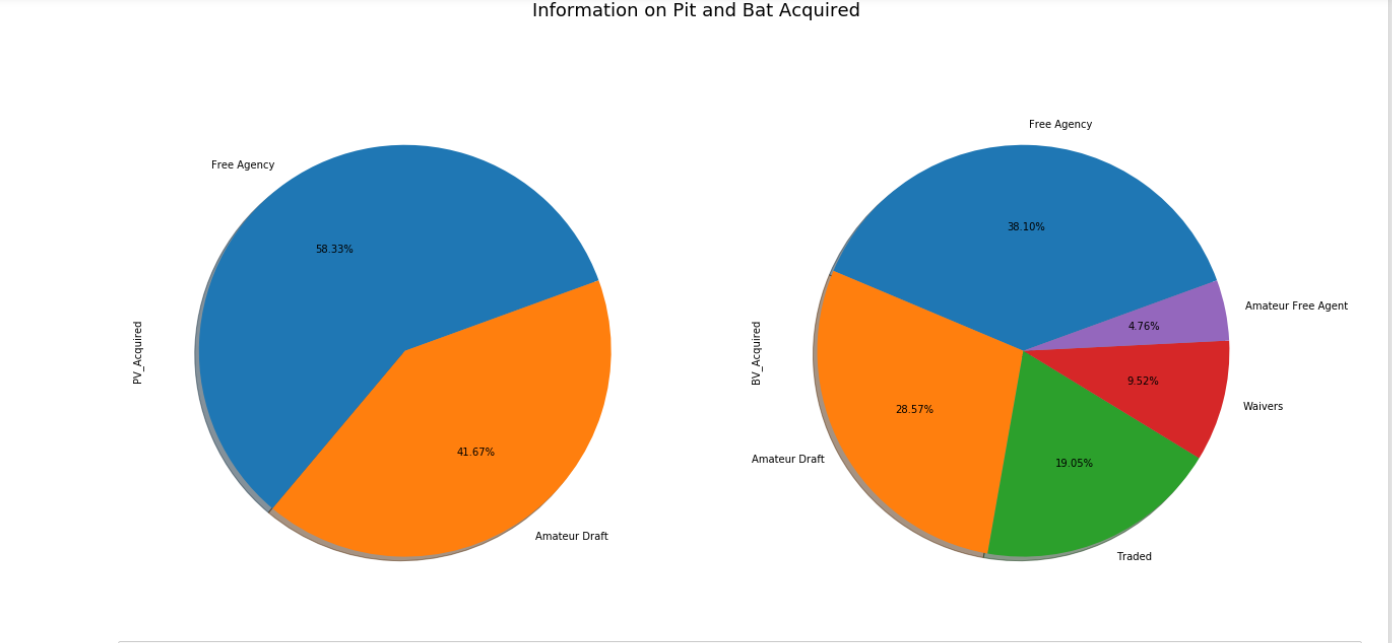
This time, let's look at the relationship between the pitcher's record and his salary. The pitcher who played less in the game has a low salary. RA9avg has a strong correlation. The gmLI has a strong inverse relationship.

Let's look at the relationship between position and salary. Pitchers and pinch runners seem to have relatively small annual salaries. However, the salary data should not be generalized by this because there are many players who have not been provided with the salary data, and they lack samples.

Just as you can see that the number of games played and the number of game started has a strong correlation, you can also see what similar records are.

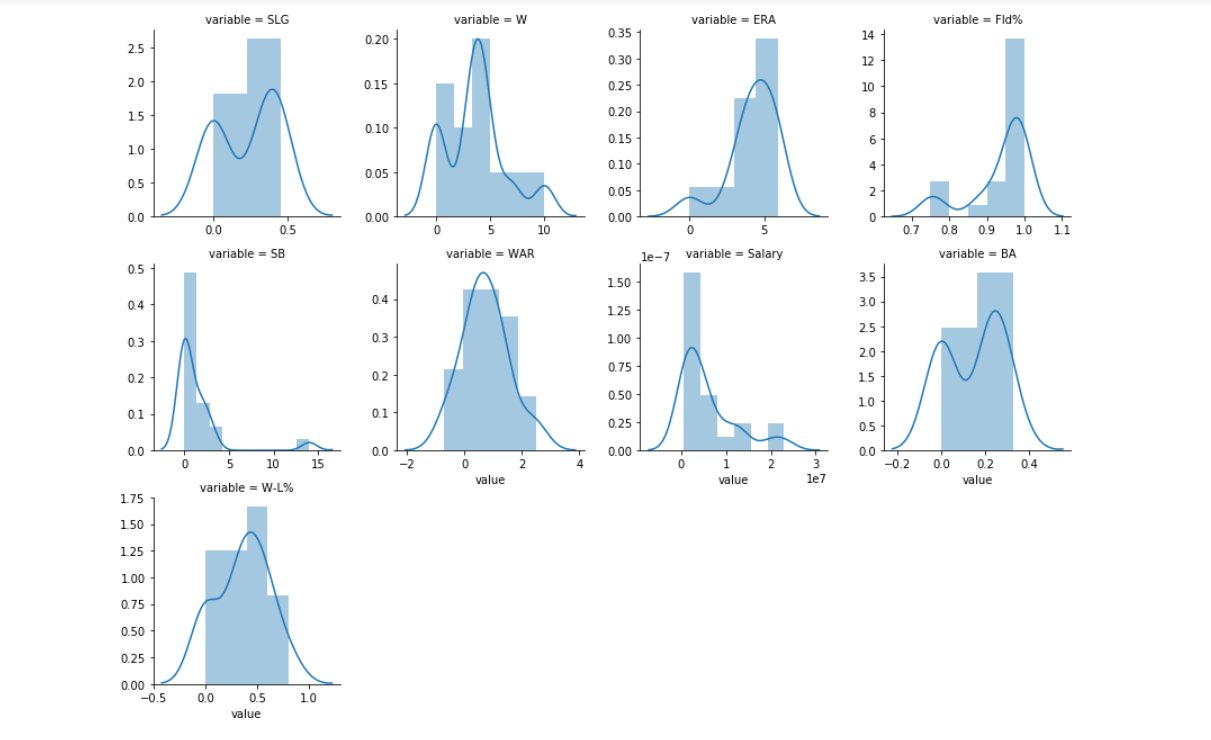
Let’s see how player was acquired. This is Information on Pit and Bat Acquired.

The pitcher was obtained using an amateur draft and a free agency, and the batter used various methods.



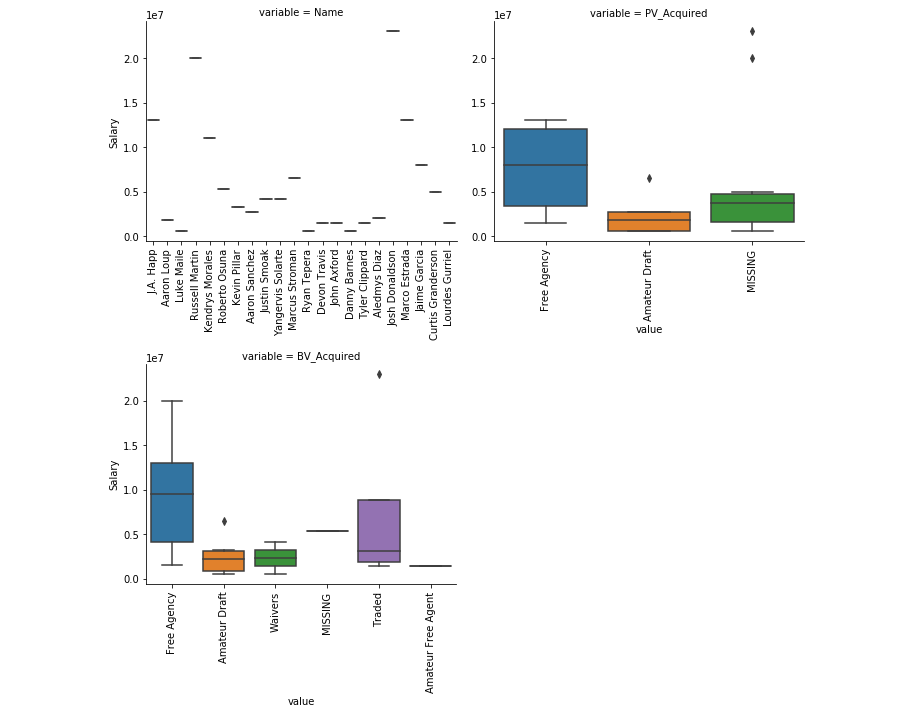
Let's look at the distribution of probability density of the main variables.

The biggest problem is the five-point team's ERA and a 2.5 percent low batting average.



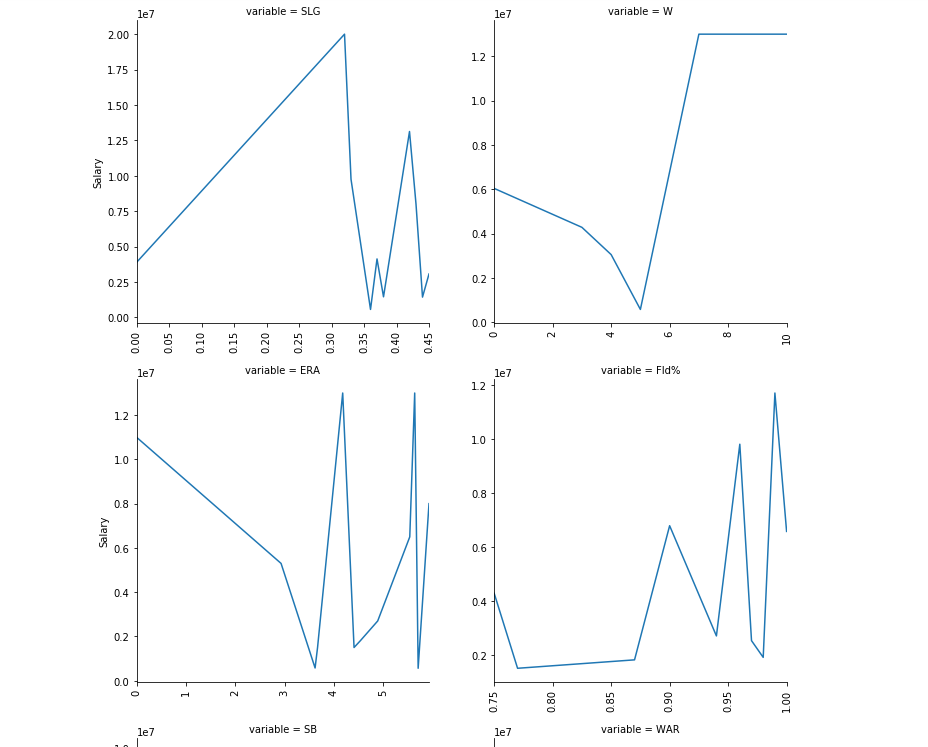
Let's analyze the difference in salary according to how was players acquired.

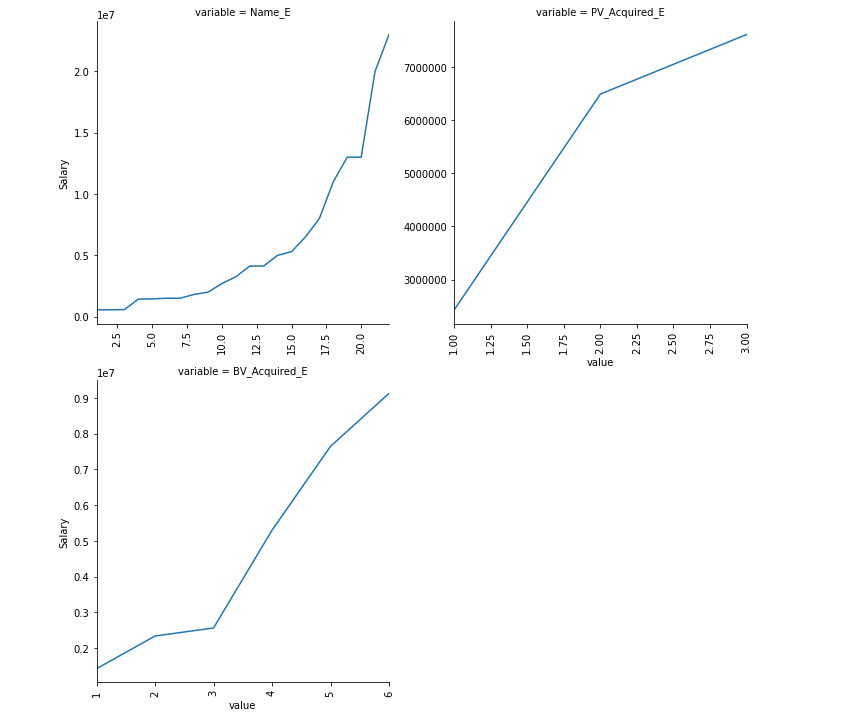
The players who brought him to the Free Agency have a high salary, You can see that players brought in amateur drafts have a low salary.

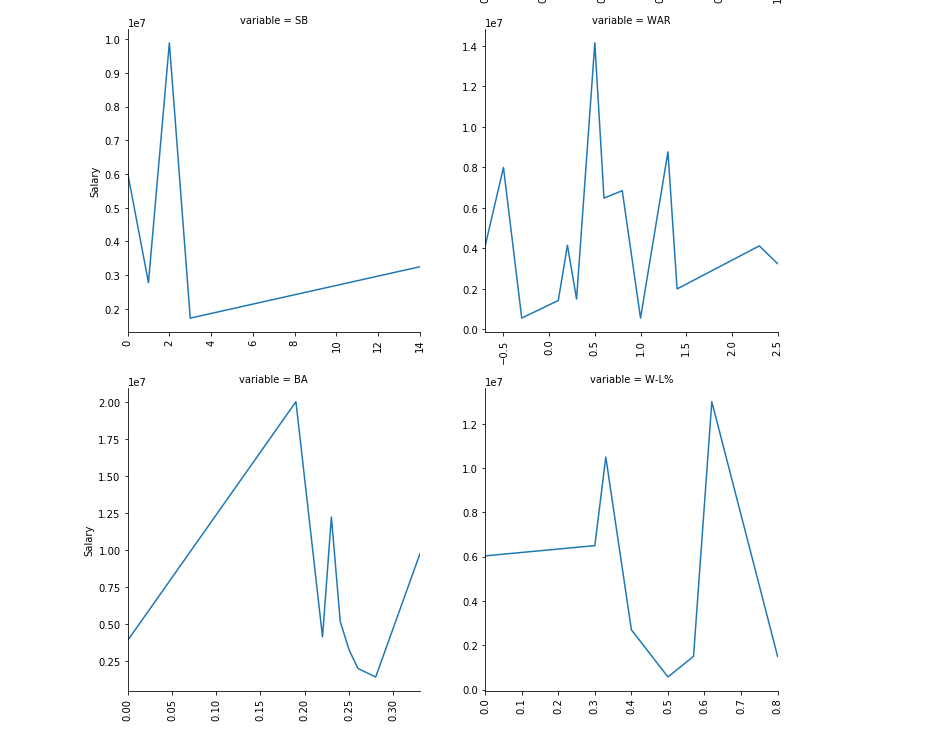


Let's analyze the relationship between the important variables and Salary. The unique values of the categorical variables were encoded according to the order of the average values excluding the anomalies of the dependent variables.

It must be close to a positive linearity except for the ERA. But in many cases, we can see that the high-salaried player has been sluggish.

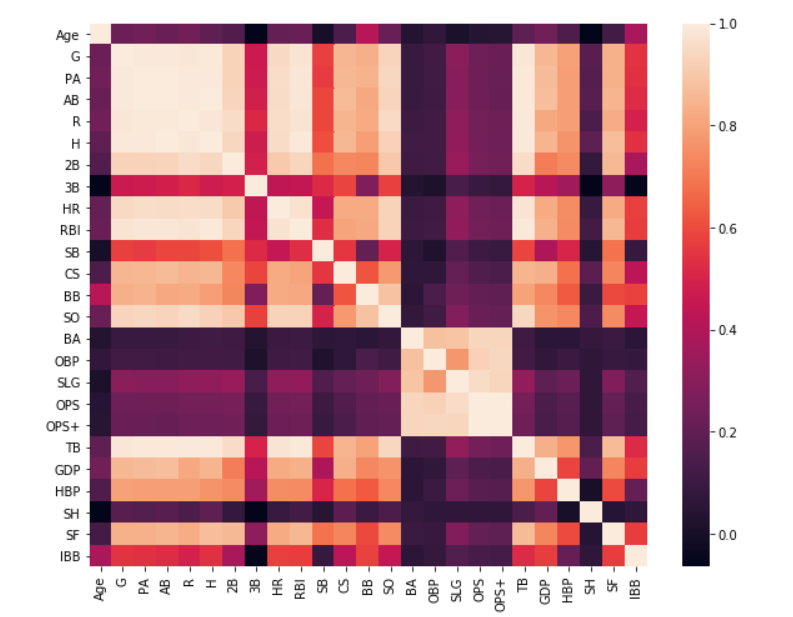




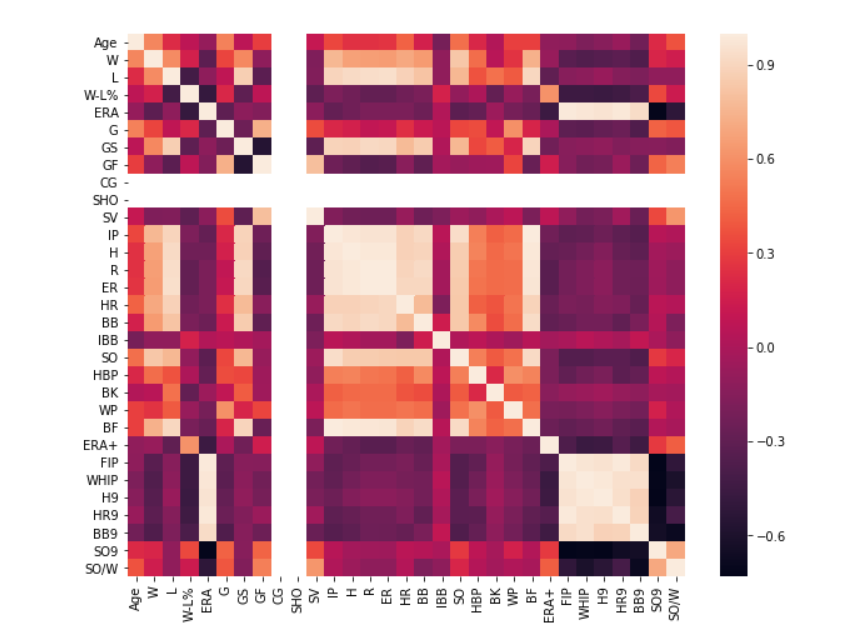


You can see the correlation between variables such as batting, pitching, position, and fielding. It is also possible to ascertain what similar records are.

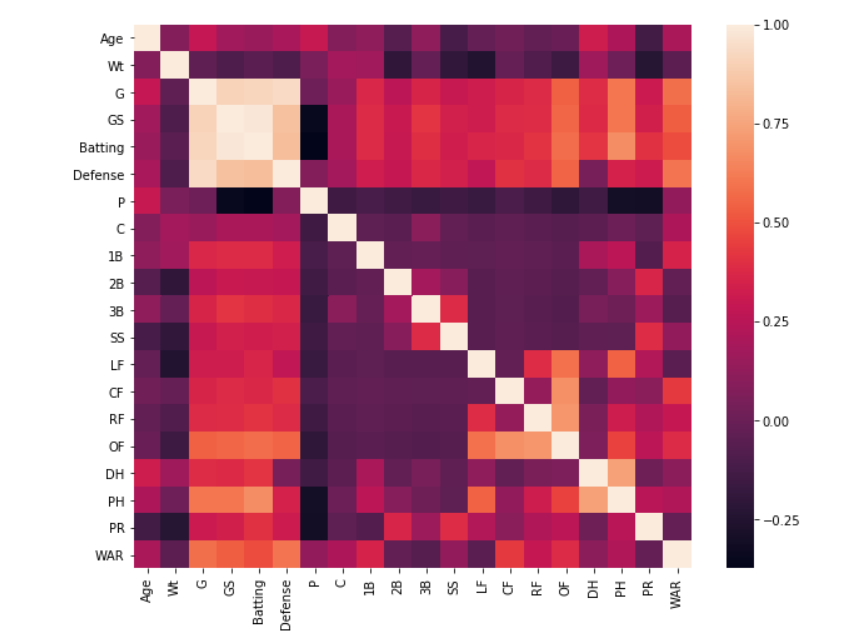
Batting



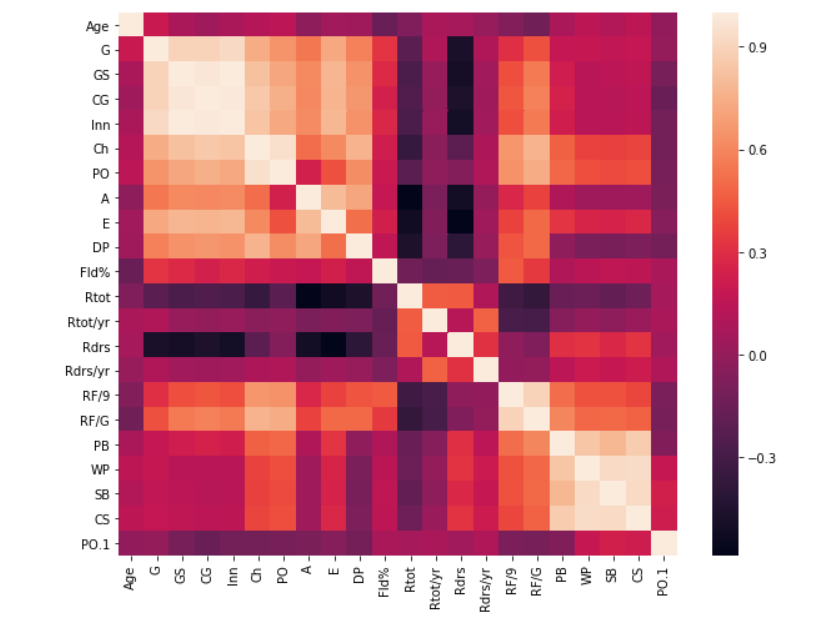
Pitching



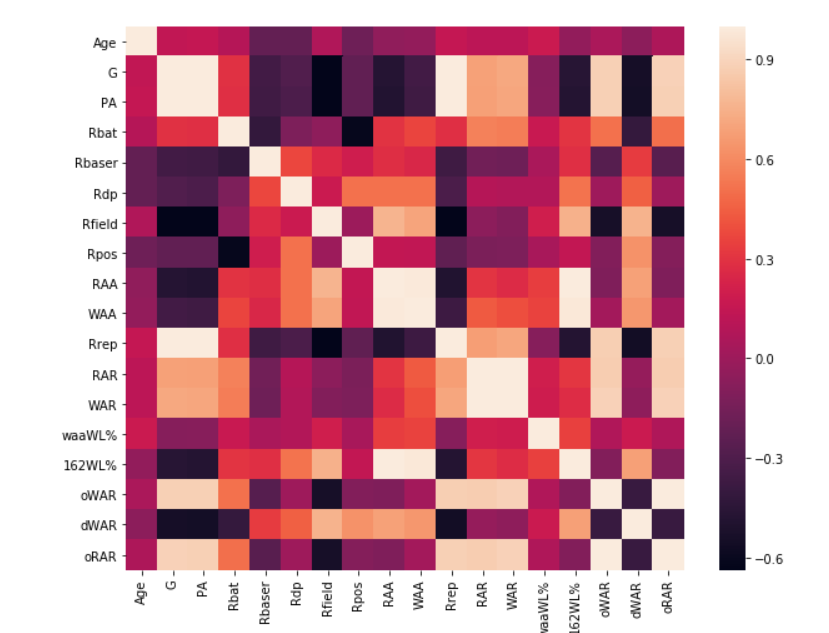
Position



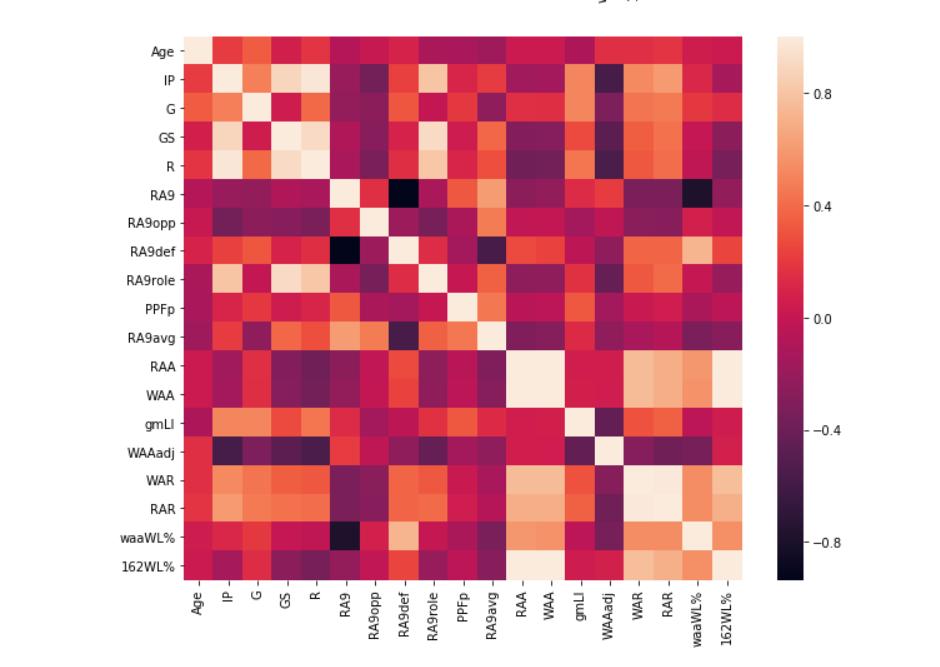
Filelding



Batter

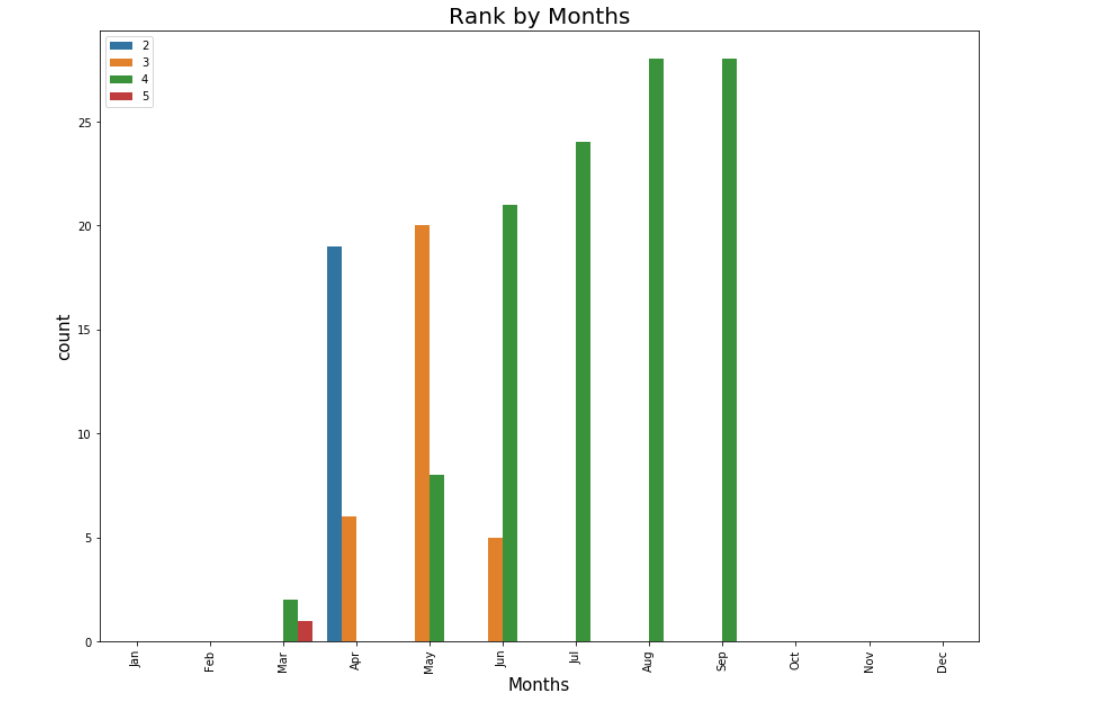


Pitcher

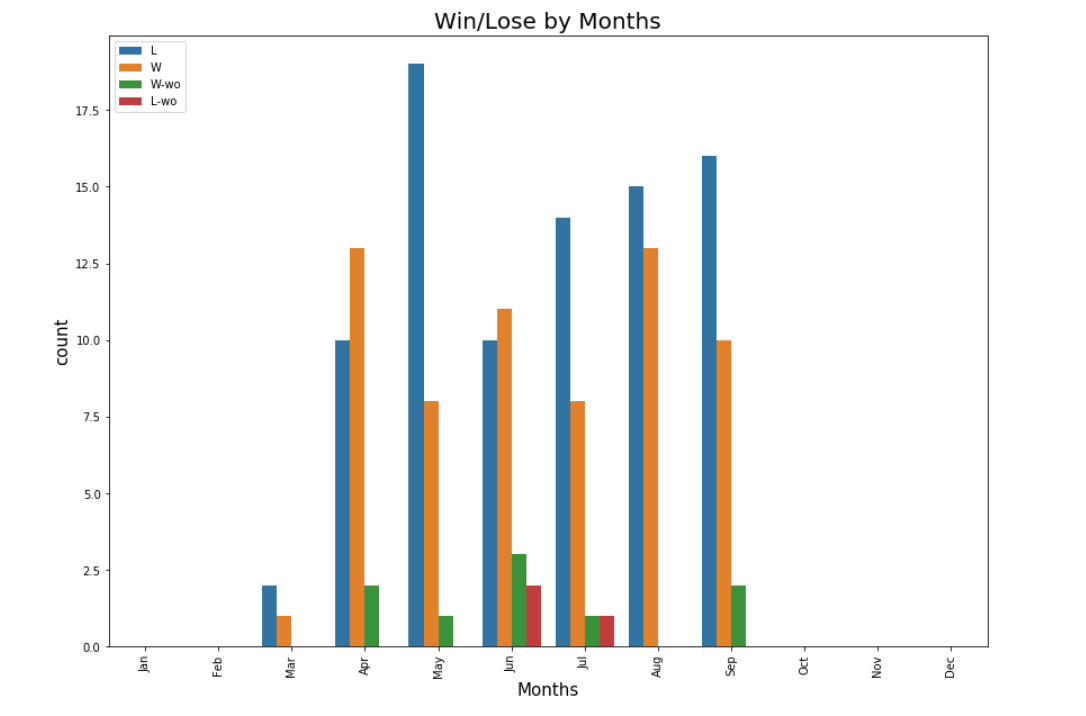


Now let's look at the team records and analyze them.

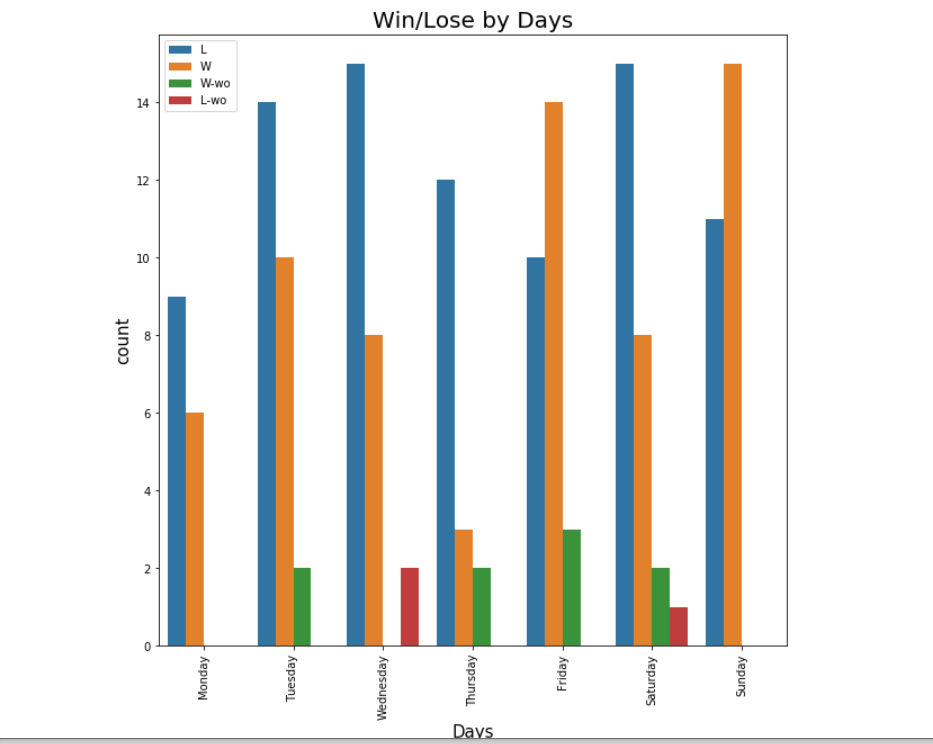
It is the change of the rank according to the months. You can see results deteriorate with time.



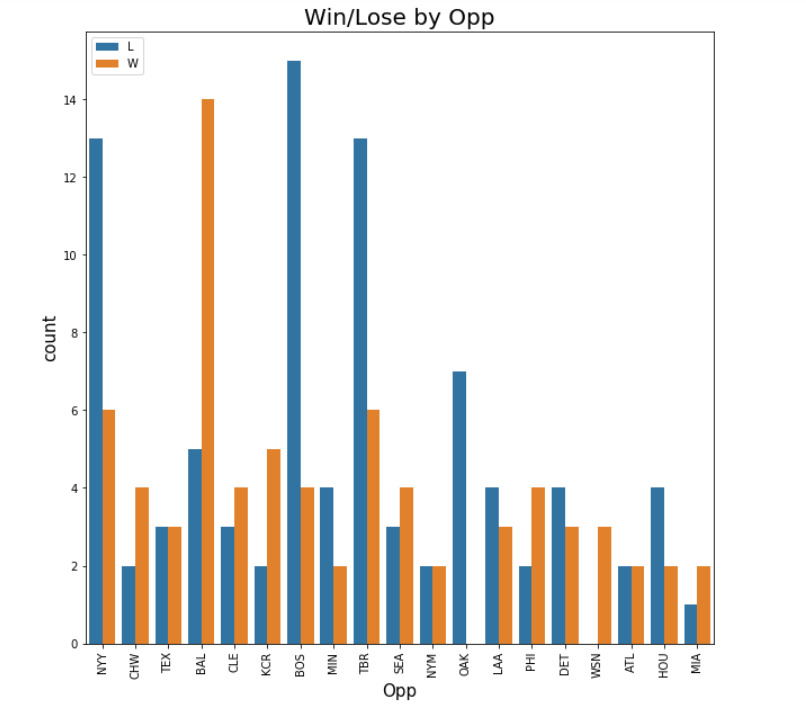
It is a record of victory and defeat according to the months. Wo is walkout win or loss.



It is a record of victory and defeat according to the day of the week. There is no particular difference.

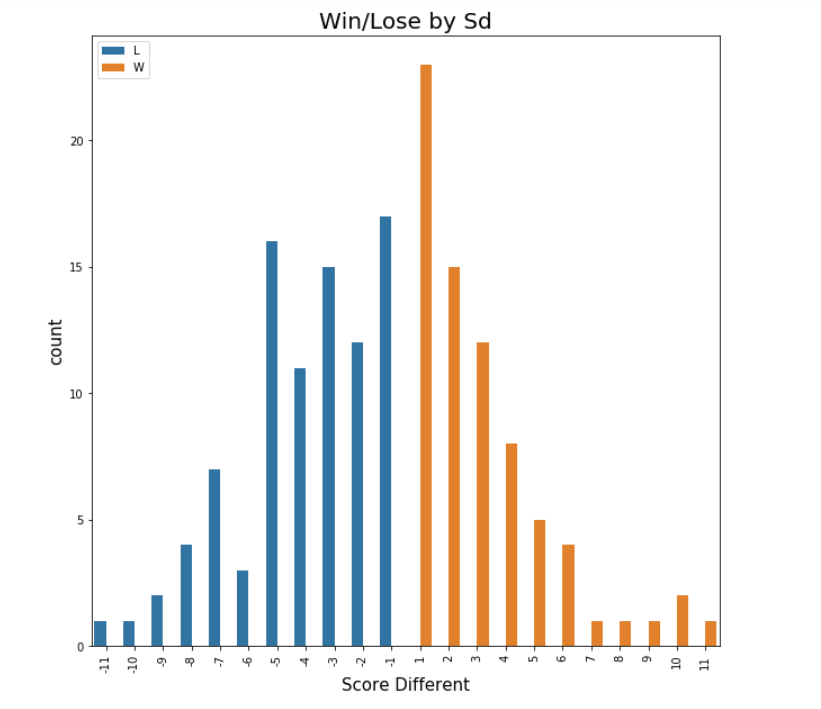


It is a record of victory and defeat by the other team. They lost a lot to NYY, BOS and TBR. They’ve never won an OAK.



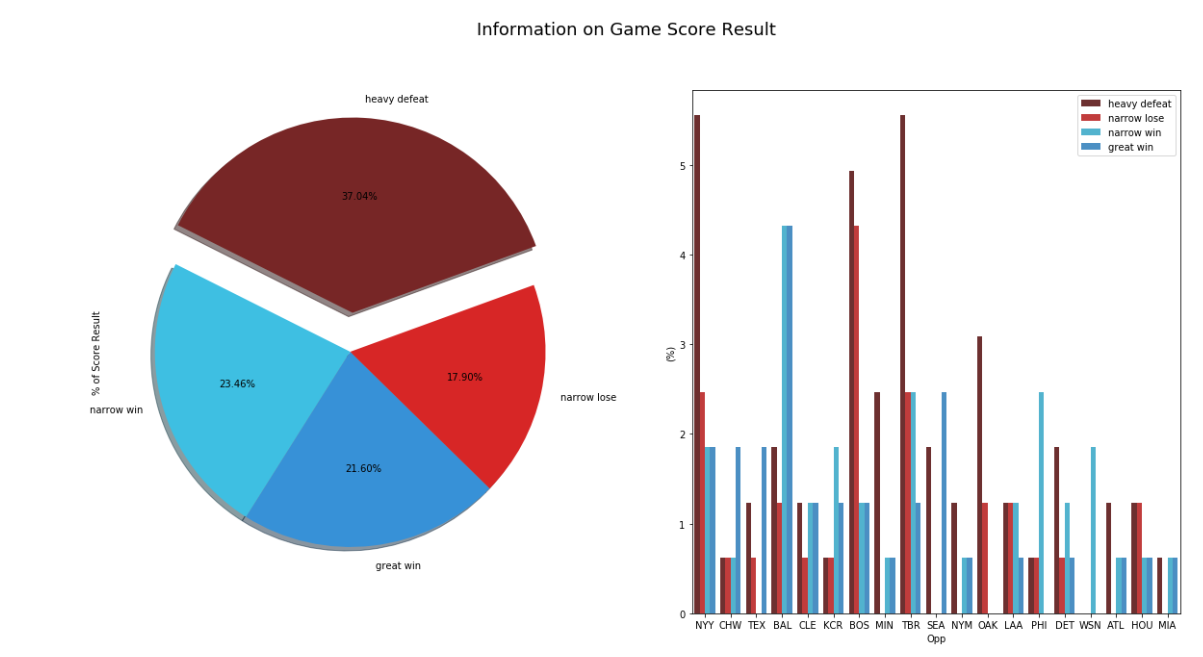
I made a new variable and analyzed the difference between the gains and losses.

When you win, you win by only one or two points and when you lose, you can see that you lose big.

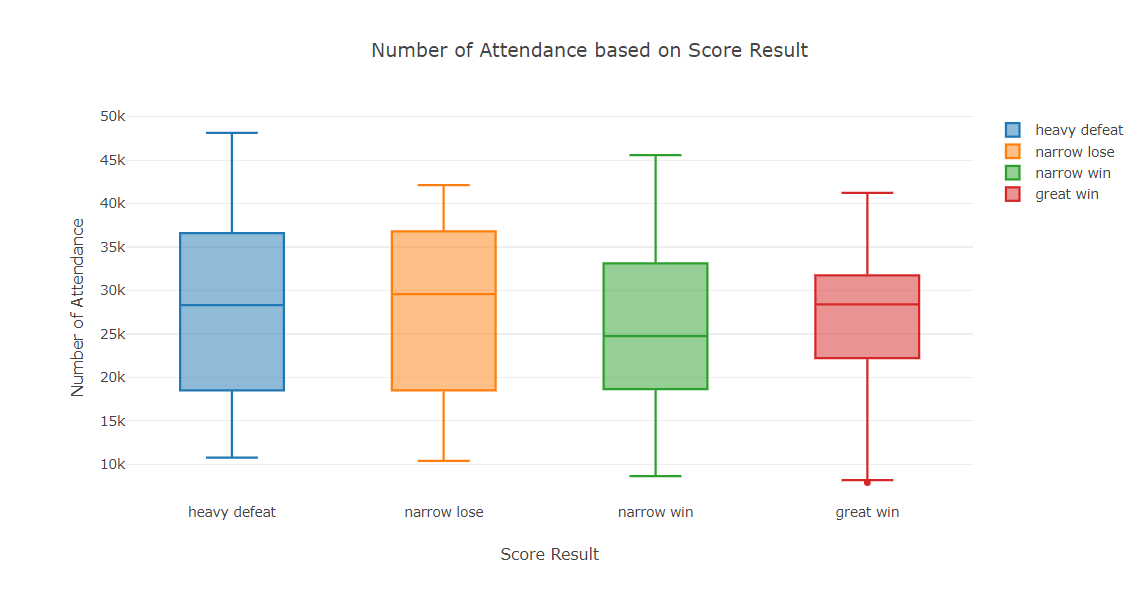


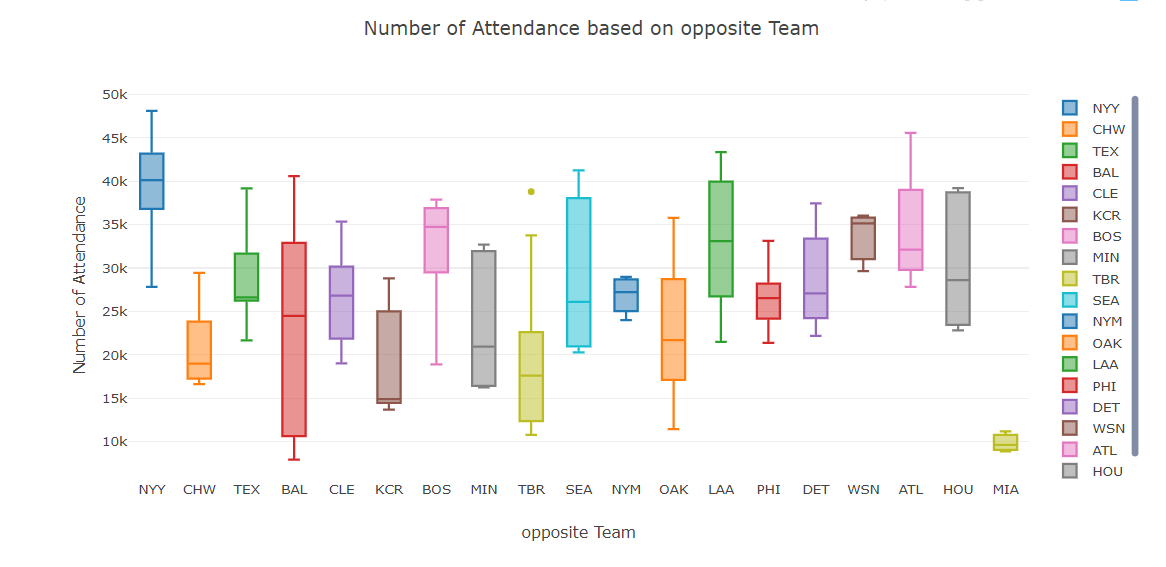
We compared the cases we won by a narrow margin or lost by a narrow margin, won or lose by a large margin. You can see that a big loss is the most common. It is a hope that 63% of the victory can be won if 1 or 2 can be won by a gradual.

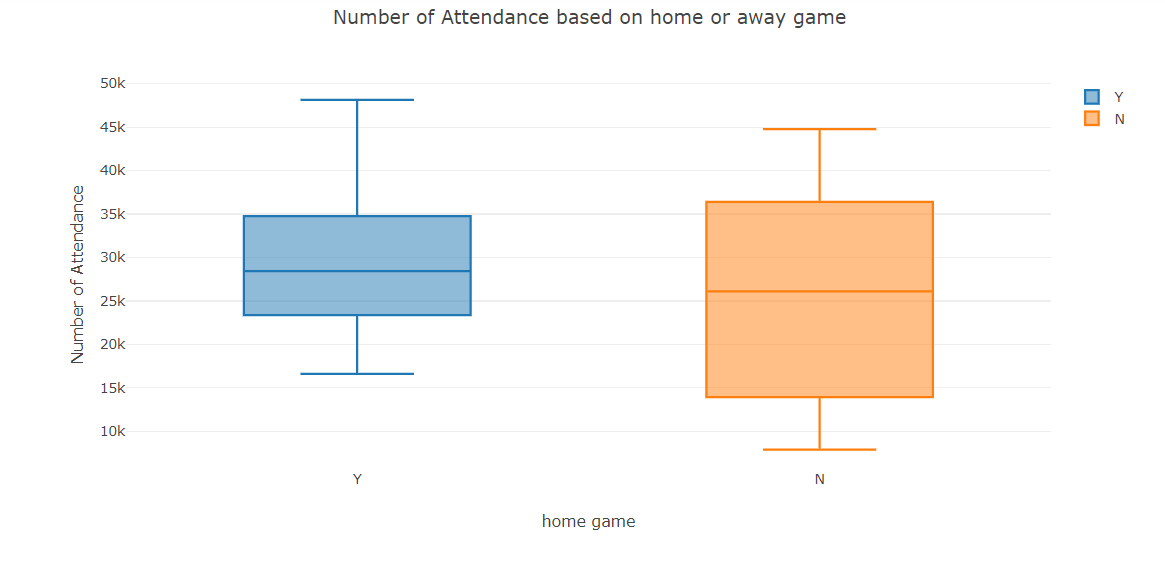
The right graph is a classification according to the opponent team. You can see that NYY, BOS, TBR, and OAK are the biggest bullies.



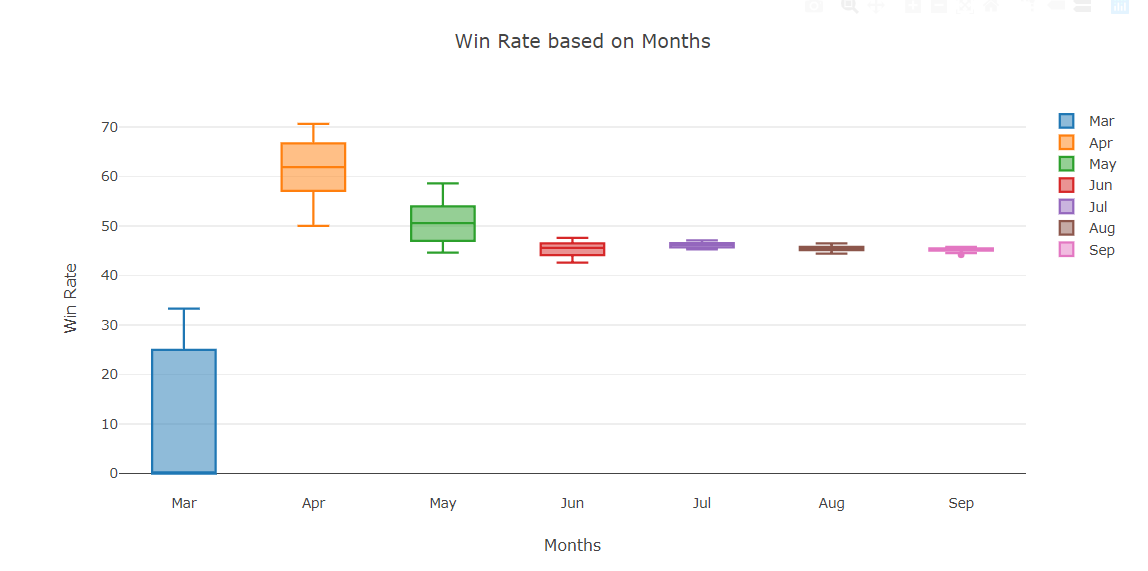
We looked at the number of Attendance according to the various circumstances.

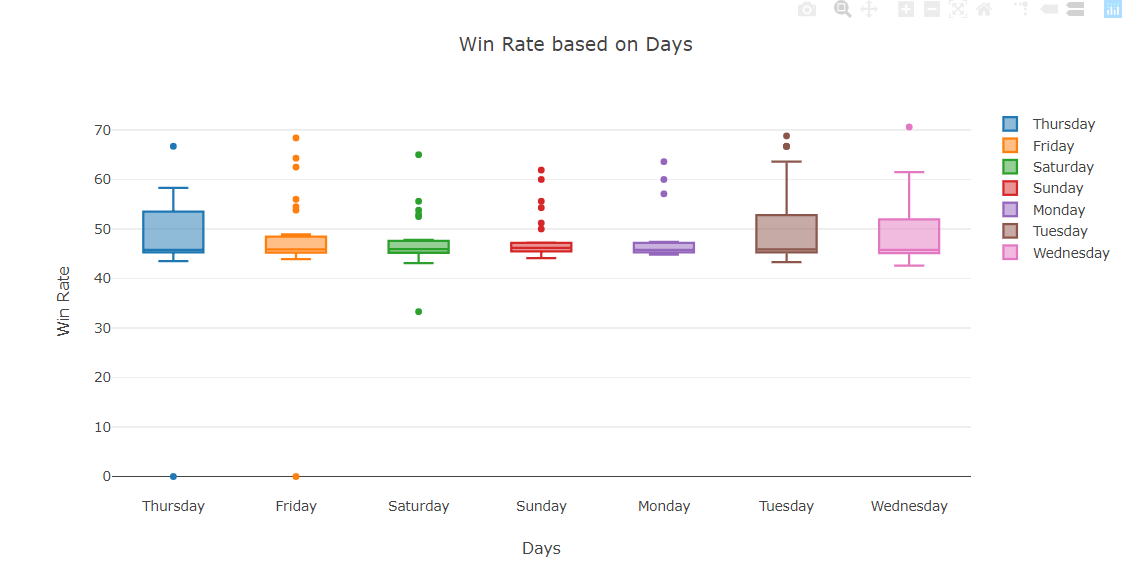




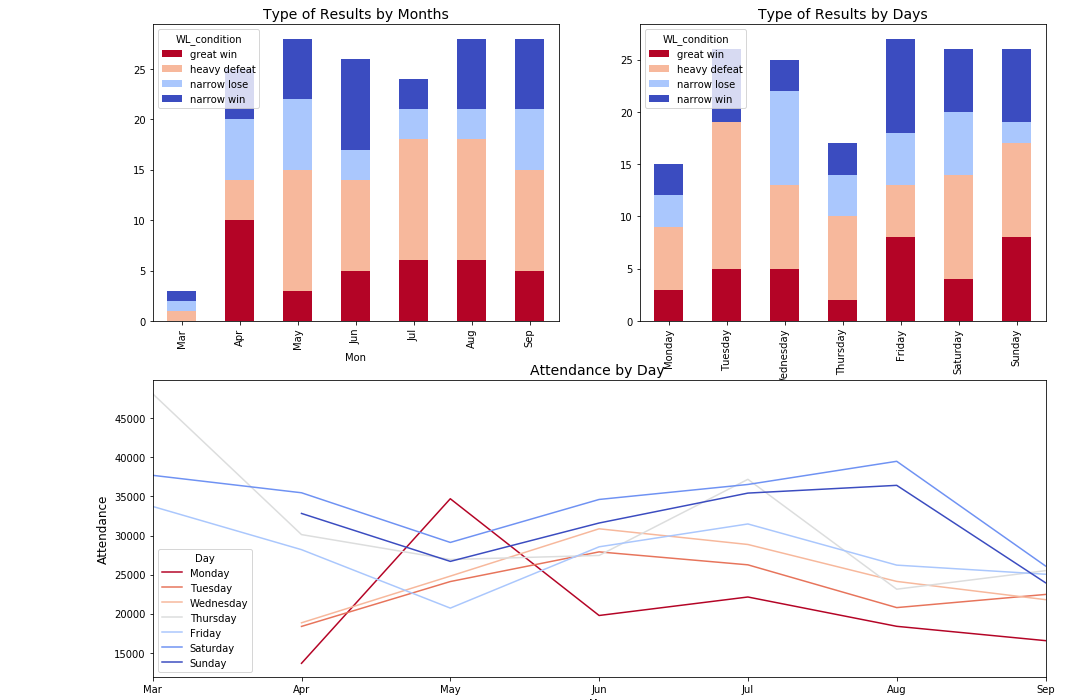


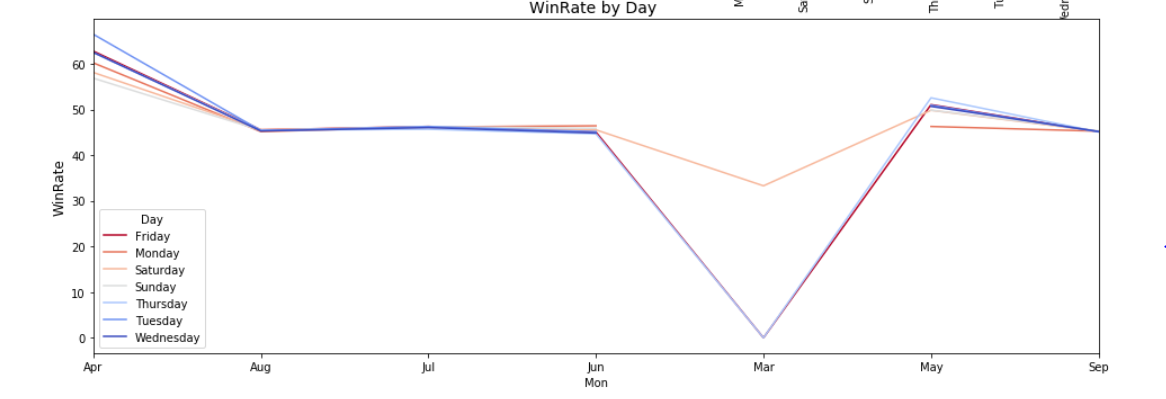
Let's look at the winning rate by month and day.



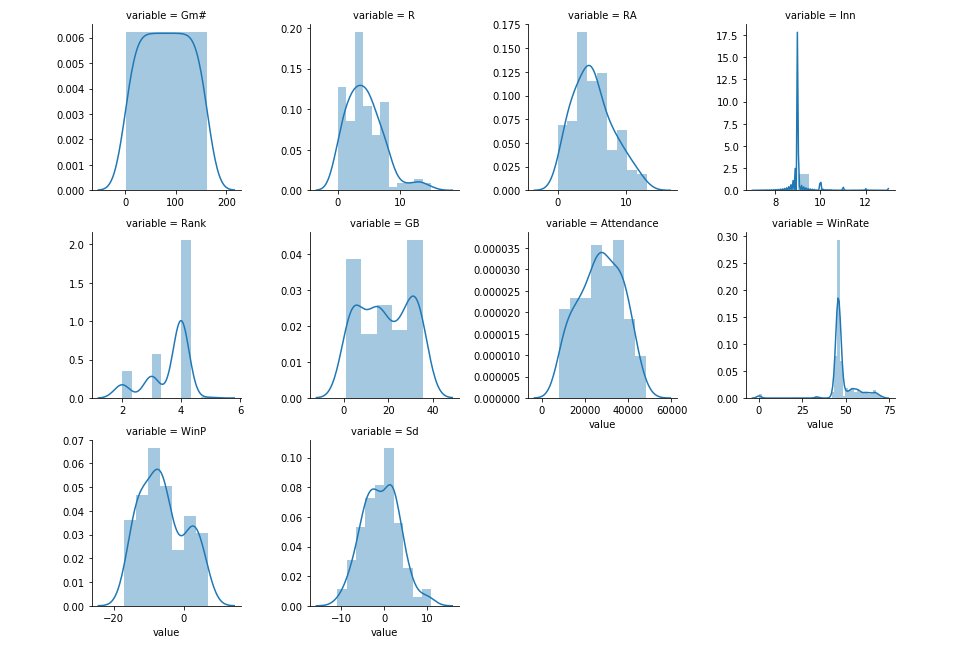


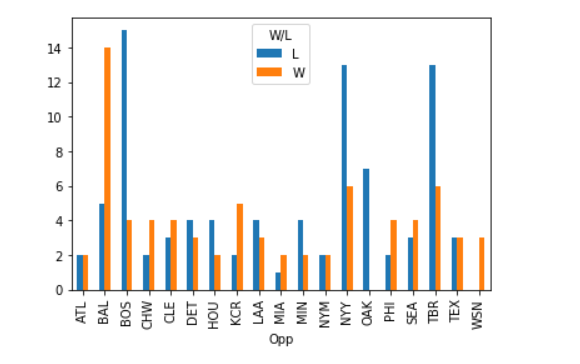
Let's look at the results of the game and the number of attendance according to the month and the day of the week. Lets see win rate by day.





Let’s see quantitative data about team. There are more lost than scores. Even if we win, it is disturbing that we have narrow wins. The collapse of the pitcher's team made the team nervous when looking at the player's data. Attacks are also sluggish, so you can see a lot of defeats.



View wins and losses by opponent team

Outcome for a home game.