Project Report

NFL Salary Cap and Team Success

1. Introduction

To maintain parity in the league, each year the NFL sets a cap on the funds that each team can use to fill out their roster. The reason that they do this is to prevent teams in big markets from acquiring all the best players because they have more money available. If there was no salary cap, teams from New York or Los Angeles would always have the strongest rosters because they have much higher revenues than the other teams. This money is used to pay each player on a team's rosters salary. It is important to note that each player in the NFL is not worth the same amount of money. Teams bid different amounts for each player based on how productive they believe that player will be for them. This means that the best players in the NFL are making significantly more than a replacement level player. Since teams must allocate their funds in a way that allows them to fill out a complete roster, they are unable to acquire as many top valued players as possible as that will leave them without quality depth. Since each player costs money to roster and high-quality players are worth even more, it can be assumed that the teams that spend more money have better rosters than those that do not. Based on this, it can also be assumed that teams that spend more money on their roster will also be more successful that season. I plan on testing if this is the case by finding if there is any correlation between the amount of cap space a team has remaining and their win percentage that season.

2. Data

This project uses two primary sources of data: Spotrac's salary cap data and the NFL's league standings data.

2.1 Salary Cap Data

The first data source that I will be using for this project is Spotrac. Spotrac is a website that tracks all financials regarding NFL teams and their players. They have a ton of data on player salaries as well as each teams' payrolls. For this project, I collected each of the 32 team's name as well as the amount of cap space they had remaining for each season starting from 2011 up until 2023. The only cleaning that I had to do with this data was change the team column from their abbreviation to the team's name and change the formatting of the salary cap column from currency to a number.

2.2 NFL Standings Data

The second data source that I used was NFL.com. This website contains the league standings for each season going all the way back to 1920. On this page of their website, you can find all sorts of data relating to each team's success such as their win/loss record, points scored and against, and their point differential. For this project, I collected each team's name, the number of games that they won, loss, and tied that year, as well as the team's win percentage that year. Since the

Spotrac data only goes back to 2011, I will only be scraping data going back to 2011 from the NFL's website. I also did not have to do much cleaning for this data as the only changes I made were finding the teams that changed their name during this time span and renaming them to their current name.

2.3 Combining Salary Cap and League Standings

Since both data sets contain the same team names and years, I merged both data frames based on both the team and the year. Since both data frames have the same exact number of observations, I did not have to do any further cleaning to the merged file. After merging, the data frame has 416 observations which is the same amount that each of the original two data frames had. A description of each variable is contained in the table below.

Data Dictionary

Column	Type	Source	Description
Team	Text	Both	The name of the NFL team
Year	Numeric	Both	The year of the NFL season
Wins	Numeric	NFL.com	The number of games that team won that year
Losses	Numeric	NFL.com	The number of games that team lost that year
Ties	Numeric	NFL.com	The number of games that team tied that year
PCT	Numeric	NFL.com	The percentage of games that team won that year
Cap Space	Numeric	Spotrac	The amount of money remaing a team has to spend on player contracts

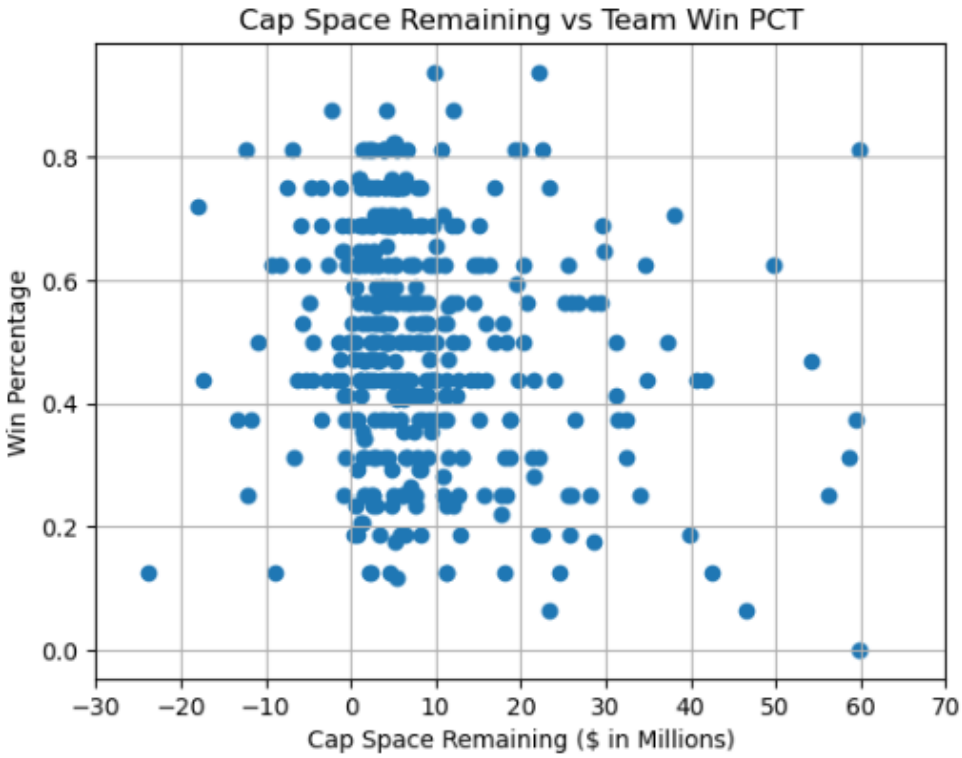
3. Analysis

3.1 Salary Cap Data and Team Win Percentage

I wanted to find out if teams that spend more money on their roster performed better during the regular season. I started by calculating the correlation between Cap Space and PCT to explore the relationship between the two variables. There was a correlation coefficient between the two variables was -.178. This means that the correlation between these two variables is very low. The reason that the correlation is negative is because the variable I used was cap space remaining. This means that the more money a team has available the lower their win percentage will be. I found these results to be a little surprising because it can be assumed that teams that spent more money on players would have a much higher quality roster than those that do not spend as much. I thought that the correlation would be much higher than it ended up being. Since the correlation is so low, teams should not look at these results and think that if they spend all of their cap money they will automatically win a lot of games. I created a scatter plot (Figure 1) to show these results plotting each team's win percentage against the amount of cap space they had remaining. There are a lot of teams that have less than $5 million dollars remaining in cap space that had low win percentages. The plot visualizes how low the correlation is because the plot appears to be completely random. The two plots that are farthest to the right are evidence of this as one of the teams had a 0 win percentage that year and the other had an 80 win percentage. The point farthest to the left is also an indicator of this as the team that spent the most money in this

entire data set had a sub 20 win percentage. One takeaway from this plot is that it does appear that NFL teams do try to utilize as much of their cap space as possible while avoiding going over the limit as most of the points plotted are between $0 and $10 million in remaining cap space.

Figure 1 Scatterplot of the amount of cap space a team had remaining and their win percentage

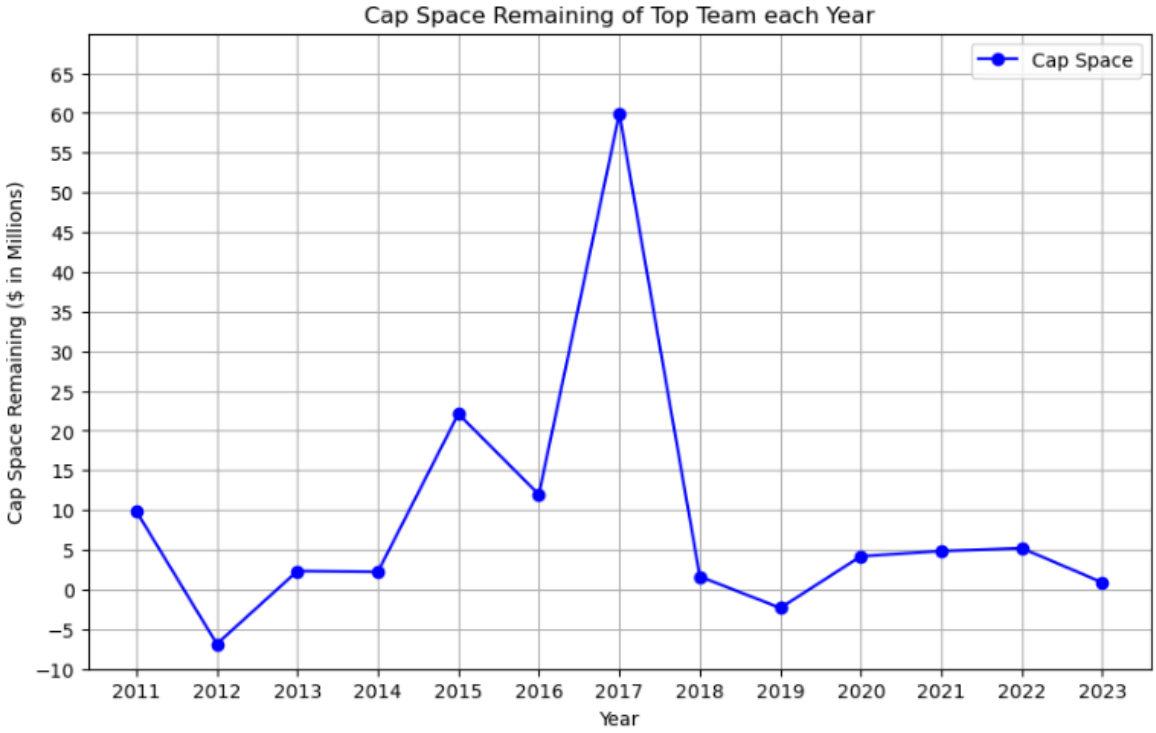


3.2 Cap Space Remaining for the Best Team Each Year

To find out how much cap space the team with the best record had each season I created a data frame that had the team with the highest win percentage for each season as well as how much cap space they had available. This data frame included the year, the best team's name, and how much money they had left over. To find the best team for each season, I wrote a line of code that created groups of teams separated by the year and then found the team in each group with the highest win percentage. These teams were put into a new data frame that I was able to use to plot on a line chart (Figure 2). The reason that I wanted to make this chart was too see if I could find any more context regarding the last part which showed that there was very little correlation between cap space remaining and win percentage. This chart was able to show that there could be more signal between these two variables. Looking at the chart, the majority of these teams had $5 million left over or less. With nine of these thirteen teams, almost 70 percent, having little to no money left over, this variable could still be somewhat predictive despite the low correlation. On the chart, the 2017 season stands out as a huge outlier since its plotted point is much higher than the rest. This point is a massive outlier in the data since at nearly $60 million it is nearly

triple the next highest team in this sample. Teams looking at this data need to understand that this is in fact an outlier and not get hope that they can also have this kind of success just because one team has done it in the past. Based on this chart, teams should be striving to spend their available funds on building a strong roster since the best teams each year tend to do so. By maximizing their roster spending, teams will have a higher chance of becoming the best team in the league than those that do not.

Figure 2 Line chart showing the amount of money the best team each season had available

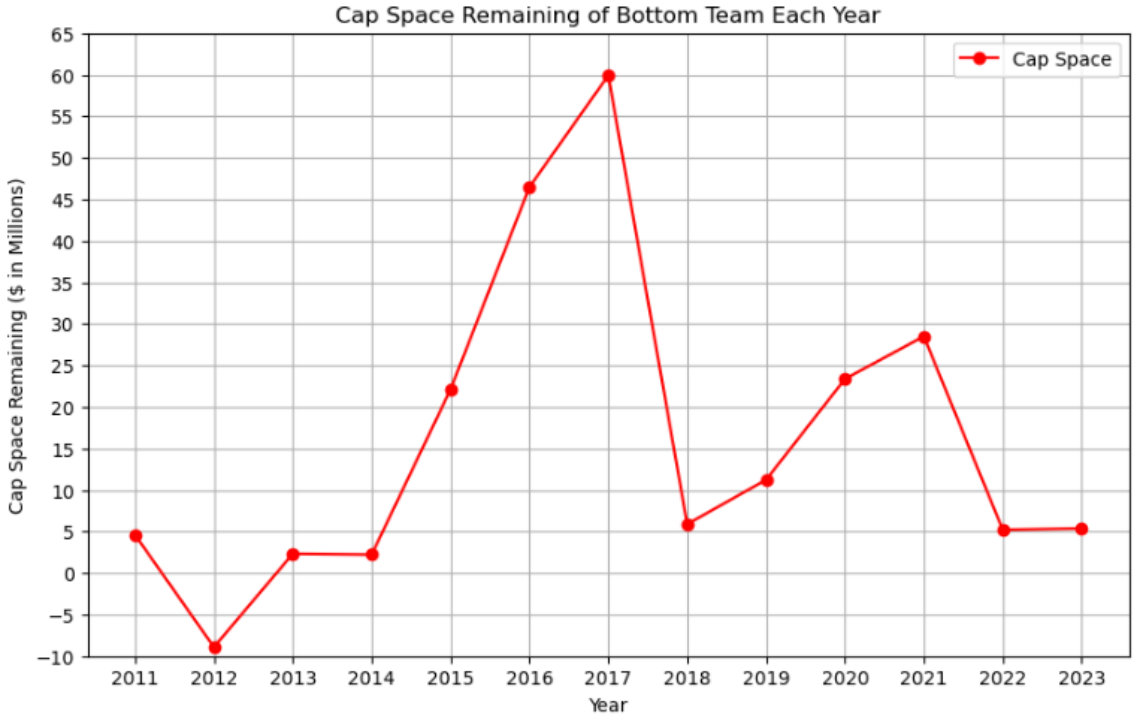


3.3 Cap Space Remaining for the Worst Team Each Year

After the last part showed that there may be more correlation between cap space and win percentage than the coefficient earlier showed, I wanted to investigate further. Since the best team each year tends to have little to no cap space remaining, I wanted to see what the worst team each year looked like on the same chart. If the worst team each season had a lot of money left over, we would be able to assume that cap space remaining can be predictive of the best and worst outcome and that the reason the coefficient is so low is due to the fact that it just is not predictive of the middle outcomes. This process was very similar to 3.2 as the steps to find the data and plot it on a line chart were almost the exact same. First, I created a data frame that found the team with the lowest win percentage in each year as well as how much cap money they had available. This data frame included the year, the worst teams name, and the cap space they had left over. To find the worst team each year, I wrote a line of code that created groups of teams separated by the year. Then, the code looked that the 'PCT' column and found the team with the

lowest value in this column for each NFL season. These teams as well as their win percentages were placed in a new data frame which I used to plot on a line chart. This chart is titled Figure 3 and can be found below. Unfortunately, this chart was not able to prove that the worst team each year had a lot of unutilized money. While the points of this chart are overall higher than those on the last chart, there are still a lot of teams here that spent nearly all of their money. Seven of these thirteen teams had around $5 million or less remaining. This significantly weakens the argument that maximizing your spending will win you more games. However, there are five teams here that had over $20 million remaining compared to just two teams in the last part. While this does not prove that by having spent less money you will win less games, it does show that the weakest teams do in fact tend to spend less than the best teams. This chart also shows that there is a very wide range of cap space for the worst teams, much more than the best teams. The range of cap spaces is about negative $10 million all the way up to $60 million. This further proves that there is little correlation between the two variables. There aren't really any major takeaways from this chart since it does not prove that teams that do not spend much money also lose more games. It also does help the argument that teams that spend more money will win more games as a result since the majority of teams listed here spent most of the money available to them.

Figure 3 Line chart showing the amount of money the worst team each year had available

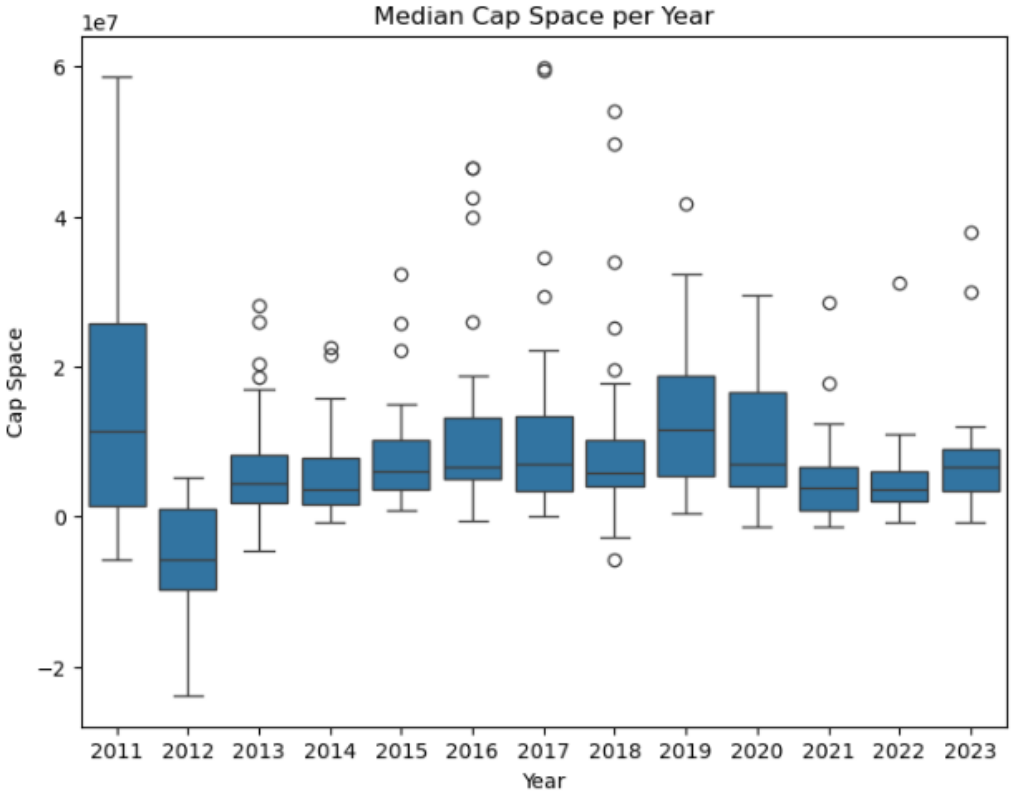


3.4 Median Cap Space per year

The next thing that I wanted to do was see if there were any trends regarding salary cap going on in the NFL. I wanted to see if the amount of salary cap left over has changed over the years. To do this, I decided to find the median amount of money teams had left over for every season in

this data. In order to do this, I had to make a data frame that contained each season as well as each team and their remaining cap space. After I got all of this into a new data frame I was able to find the median for each season. The final step was just to find a way to display the results. I decided to make a box plot to display the medians since it also will provide some additional context. This box plot is named Figure 4 and it can be found below. Looking at this box plot you can see that the median (horizontal line inside the blue box) has been very consistent over this time period. This shows that there are no noticeable trends happening in the NFL regarding how much money teams are spending. Based on this graph it is clear that teams are trying to use up all of the money they have available to field a roster that is as strong as possible. The range of medians is very tight as the lowest median here is 2012 with just under -$1 million and the highest median being 2019 with about $1 million. Outside of 2011, the interquartile ranges are also pretty tight as the blue boxes are not very tall. This tells us that the majority of teams are all in the same range in terms of how much money they have left over. The last thing of note is that there are a lot of outliers mainly on the top part of the graph. There have been quite a few teams that are outside the top whiskers of the bar chart meaning that there are a lot of instances of teams underspending. Comparing this to the bottom of the bar chart and there is only one outlier in this entire data set that went completely overboard with their spending. This shows that teams are much more comfortable underspending rather than overspending which makes sense since the amount a team goes over the salary cap gets an additional tax from the NFL. While I was not able to find any trends relating to the spending of the league overall, this chart does show that teams are afraid to go over the cap which I still find interesting.

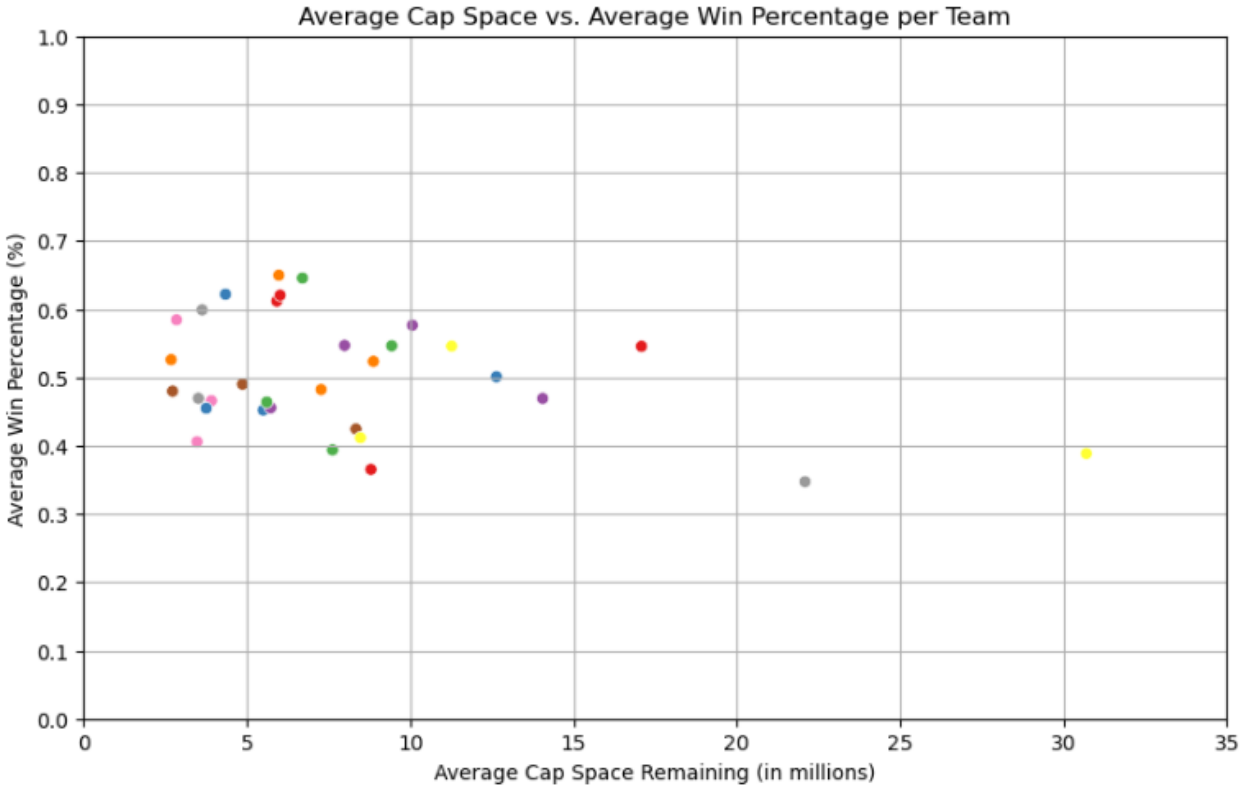
Figure 4 Box plot showing the median amount of money remaining for each year



3.5 Average Cap Space and Average Win Percentage

The last thing I wanted to do with this data was see what the averages looked like for both of these variables. The reason I wanted to do this was to see what teams were the most successful over this time period and also look at what they did with their money. To do this, I had to create another data frame. This data frame contained each teams name and their cap space remaining. I combined all of their seasons data as well as their cap space data under a grouping for each team. By doing this I was able to find the average of both variables for each individual team. Once this was done, I created a scatter plot to visualize the results. This scatter plot is titled Figure 5 and can be found below. Looking at the scatter plot you can see that the majority of the points are on the left side. This did not come as a surprise since the results of the previous parts showed that teams try to utilize all of their cap money. This scatter plot does, however, show that teams that are consistently great also consistently spend their money. Now this does not mean that if you spend most of your money over a period of time you are guaranteed success as there are still quite a few teams that also spent a lot but also have a win percentage that is under 50. What is interesting about this scatter plot is that the teams that have consistently spent much less also won less over this time period. The two teams that fall much further right than the group had some of the lowest win percentages in this sample. The takeaway here is that while spending a lot does not guarantee you success, it does look like it gives you a better shot. Also, if a team consistently does not spend their money, they also will probably not win many games.

Figure 5 Scatter plot of the average win percentage of each team and the average amount of cap space they had remaining over this time period



4. Conclusion

In this project, I analyzed two aspects of NFL teams: their cap space remaining and their win percentage. In summary, from the analysis questions presented in my proposal, I found the following results.

1. *Is there a strong correlation between money spent on player contracts and team success?*

   There is very limited correlation between the amount of money a team spends on player contracts and their teams success. But while the correlation is not very strong, there are some indicators that it is still a good idea for a team to use as much cap space as possible.

2. *Do the teams that spend the most money on their roster usually win the most games?*

   When looking at both individual seasons and averages over this time period (Figures 1 and 5), the teams that have won the most games tend to have also spent a lot of their money. However, there are still more teams that have maxed out their cap space and did not have much success. This means that while successful do usually spend their money, a team using up the majority of their cap space does not guarantee that they will be successful.

3. *Can teams that do not spend a lot of money still be successful?*

   While yes it is possible for teams that do not spend a lot of money to be successful, it definitely decreases their likelihood. Looking back at figure 1 you can see that there are certainly instances of teams with a lot of remaining cap space also winning a lot of games, but there are not nearly as many of these instances as teams that do spend a lot of money. Combine this with the fact that teams that are consistently cheap, consistently have much less success than those that are willing to spend (Figure 5).

4. *Are there any leaguewide trends happening regarding cap space?*

   Looking at Figure 4, I was unable to find any noticeable changes regarding teams and their salary cap over this time period. One thing of note is that teams are consistently trying to max out their cap space as the medians for each year are very low. The last takeaway from Figure 4 is that teams are much more willing to underpay than go over the salary cap. The number of outliers that underspent was significantly higher than the number of outliers that overspent.

This project does have one limitation: sample size. Since Spotrac only started collecting data in 2011, this data only has a sample size of 13 years. While this was perform reliable analysis on, it

does not go back far enough to really showcase how things have changed over the NFL's history. Future work on this project could include finding another website that has contract data like Spotrac's or even going through NFL's league and team's transactions to find the numbers yourself. If a bigger sample was to be found, you would be able to get a much deeper understanding of the minds of NFL front offices and how things have changed over the last century. This also could lead to finding out how salary cap strategy changes have affected the overall success of the NFL.