NFL Cap Space Analysis

Michael Kozloff

Data Set

```
Team <- c("Bills","Dolphins","Patriots","Jets","Ravens","Bengals","Browns",
          "Steelers","Texans","Colts","Jaguars","Titans","Broncos","Chiefs",
          "Raiders","Chargers","Cowboys","Eagles","Giants","Redskins","Bears",
          "Lions","Vikings","Packers","Falcons","Panthers","Saints","Bucs","Cardinals",
          "49ers","Seahawks","Rams")
Division <- c(rep("AFC East",4),rep("AFC North",4),rep("AFC South",4),
              rep("AFC West",4),rep("NFC East",4),rep("NFC North",4),
              rep("NFC South",4),rep("NFC West",4))
Wins_2018 <- c(6,7,11,4,10,6,7.5,9.5,11,10,5,9,6,12,4,12,
               10,9,5,7,12,6,8.5,6.5,7,7,13,5,3,4,10,13)
Carry_Over <- c(8.7,5.9,3.1,12.2,4.5,7.4,56.5,18.2,18.8,49.1,11.6,
                25.6,8,.72,1.9,1,11.7,6.2,5.9,5,3.6,6.4,4.1,7.8,1.1,
                2.5,1.8,4.5,6.4,35,2.5,.47)
Cap_2018 <- c(7.1,  7.1,  6.1, 18.2,9.9, 8.2,
              53.7, 19.8, 15.0, 49.4,  6.0, 24.9,  6.4,
              5.3,  7.7,  2.3,  8.5,  7.7,  6.9, 5.0,  5.1,  6.8,
              1.1,  5.0,  4.2, -1.0,  2.2,  2.5,  6.0, 34.9,  3.7,  3.4)
d <- data.frame(Team,Division,Wins_2018,Cap_2018,Carry_Over)
head(d)
```

```
##       Team  Division Wins_2018 Cap_2018 Carry_Over
## 1    Bills  AFC East         6      7.1        8.7
## 2 Dolphins  AFC East         7      7.1        5.9
## 3 Patriots  AFC East        11      6.1        3.1
## 4     Jets  AFC East         4     18.2       12.2
## 5   Ravens AFC North        10      9.9        4.5
## 6  Bengals AFC North         6      8.2        7.4
```

```
tail(d)
```

```
##         Team  Division Wins_2018 Cap_2018 Carry_Over
## 27    Saints NFC South        13      2.2       1.80
## 28      Bucs NFC South         5      2.5       4.50
## 29 Cardinals  NFC West         3      6.0       6.40
## 30     49ers  NFC West         4     34.9      35.00
## 31  Seahawks  NFC West        10      3.7       2.50
## 32      Rams  NFC West        13      3.4       0.47
```

DATA EXPLANATION

The data selected contains information on all 32 NFL teams (column=Team) partnered with their division (column=Division), wins from this past season (column=Wins_2018), salary cap at the end of the season (column=Cap_2018), and the carry over cap space that each team will have at the start of 2019 (column=carry.over). Every category is observational.

Cap space is a complicated metric. The NFL cap space is a measure of the salary cap, which is a rule that limits the amount of money teams can put towards their players' salary. If the owners spend past the salary cap, there are fines associated. Each team has the same amount of cap space to work with, but every year this cap space gets bigger.

The carry over cap space is dependent on our cap space from 2018 and other adjustments. This money is the amount of left over cap space (our Cap_2018 column) as well as some adjustments that are complicated and more financially based. This all seems complicated, but going forward we will use these assumptions and facts to guide statistical analysis and exploration.

DATA CHARACTERISTICS

We are analyzing this data as population data: analyzing all 32 teams from the NFL.

Understanding how the team, division, and win data are gathered is simple and doesn't need any explanation. I went out and found the data myself: it all can be found on the NFL site or from a quick google search. However, the win data has a catch; some teams have a decimal amount for wins. These teams had a tie at some point during the season.

Both the carry over cap space and 2018 cap space were pulled in a similar manner. There are websites that track the cap space and carry over cap space– I pulled my data from there.

ISSUES

A big issue to keep in mind is that cap space changes after free agency starts. Free agency is when owners are able to sign contracts and release contracts, which changes the total

available cap space. Free agency in the NFL happens on March 13, 2019. This data likely will not be viable or helpful for very long.

Assumptions and things to keep in mind:

1.Carry over cap space and cap space after the 2018 season are dependent, and therefore positively correlated.

2.The column "Wins_2018" is a sum of the amount of wins and ties of a given team.

3.Ties take on a value of 0.5, while wins take on a value of 1.

4. The columns "Cap_2018" and "Carry_Over" are in millions, scaled down to a more succinct number and rounded to one decimal point
5. The "Division" column helps make the data set easier to read
6. Each team plays 16 games in the regular season, and for this data, we will not be analyzing playoff wins.

GRAPHICAL SUMMARY #1: Confirming Assumption 1

```
library(tidyverse)
```

```
## Warning: package 'tidyverse' was built under R version 3.5.2
```

```
## -- Attaching packages ---------------------------------------------------------
```

```
## v ggplot2 3.1.0     v purrr   0.3.0
## v tibble  1.4.2     v dplyr   0.7.8
## v tidyr   0.8.2     v stringr 1.3.1
## v readr   1.3.1     v forcats 0.3.0
```

```
## Warning: package 'ggplot2' was built under R version 3.5.2
```

```
## Warning: package 'tidyr' was built under R version 3.5.2
```

```
## Warning: package 'readr' was built under R version 3.5.2
```

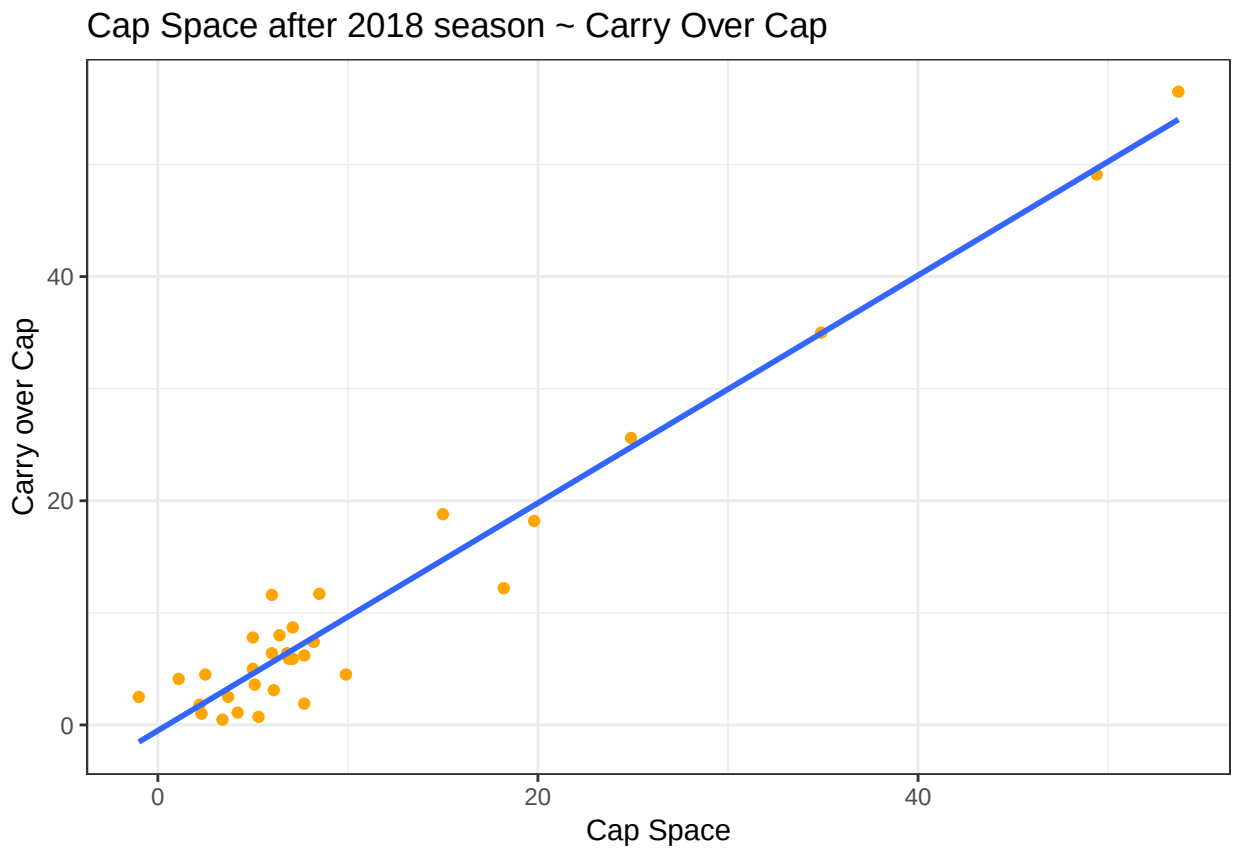
```
## Warning: package 'purrr' was built under R version 3.5.2
```

```
## Warning: package 'dplyr' was built under R version 3.5.2

## Warning: package 'forcats' was built under R version 3.5.2

## -- Conflicts ---------------------------------------------------------------- tidyv
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()

base <- ggplot(d, aes(x=Cap_2018, y=Carry_Over))
base+geom_point(color="orange")+geom_smooth(method=lm,se=F)+
  labs(title="Cap Space after 2018 season ~ Carry Over Cap",
       x="Cap Space",y="Carry over Cap")+theme_bw()
```

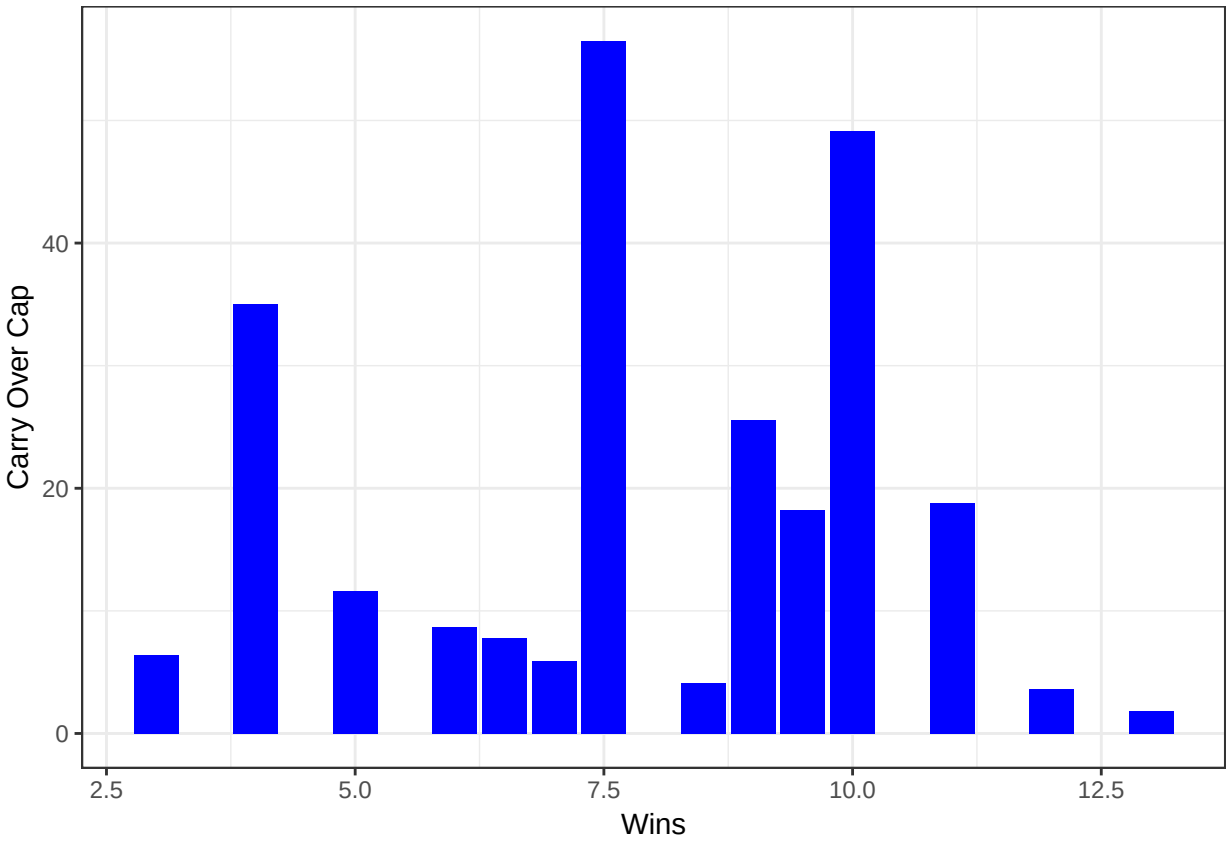


The regression line displays a strong, positive correlation between carry over between cap space and carry over cap. We expect this by the nature of what the carry over cap represents. Due to the baseline data of both variables being similar, we can be fairly confident that our omitted variable bias is limited in its scope. Homoskedasticity seems to hold at first glance as well.

This graph is important because we want to use the Carry Over Cap instead of the Salary Cap when we do more exploratory analysis, as it more recent and doesn't have negative values.

GRAPHICAL SUMMARY #2 Are Wins and Cap Space related?

```
bar <- ggplot(d, aes(x=Wins_2018, y=Carry_Over))
bar+ geom_bar(stat="identity",position="dodge",fill="blue")+
  labs("Carry Over Cap by Wins",x="Wins",y="Carry Over Cap")+theme_bw()
```



We expect teams with lots of carry over cap space to have a lower win total. This makes sense economically– a firm not utitlizing all of their resources operates suboptimally. The opposite should be true as well: teams who are good (have more wins) should have used all of their resources in an attempt to sign the best teams possible. Owners make mistakes though; sometimes players get lucrative contracts and get injured or don't contribute the way they do. But, is it that black and white?

From this graph, that might not be the case. It doesn't really seem like there is an apparent

trend between wins and carry over cap; however, this opens up the possibility to confirm this suspicion through statistical inference.

NUMERICAL SUMMARY #1 Mean & Carry Over Cap Space by Division

```
group_by(d, Division)%>%summarize(AverageWins=mean(Wins_2018),
                                  AverageCarryOverCap=mean(Carry_Over))
```

```
## # A tibble: 8 x 3
##   Division  AverageWins AverageCarryOverCap
##   <fct>           <dbl>               <dbl>
## 1 AFC East         7                   7.48
## 2 AFC North        8.25               21.6
## 3 AFC South        8.75               26.3
## 4 AFC West         8.5                 2.90
## 5 NFC East         7.75                7.2
## 6 NFC North        8.25                5.48
## 7 NFC South        8                   2.48
## 8 NFC West         7.5                11.1
```

This numerical summary gives us average wins and average carry over cap for each division. From first glance, it doesn't look like anything reasonable can be deduced. The average wins are all pretty close, and the Carry Over Cap is all over the place, with maximum and minimum values of both average carry over cap and average wins coming from AFC divisions. If we broadened the average to conference instead of division, maybe we could have some more definitive results.

NUMERICAL SUMMARY #2 Mean & Carry Over Cap Space by Conference

```
Conference <- c(rep("AFC",16),rep("NFC",16))
dconf <- data.frame(Team,Conference,Wins_2018,Cap_2018,Carry_Over)

group_by(dconf, Conference)%>%summarize(AverageWins=mean(Wins_2018),
                                        AverageCarryOverCap=mean(Carry_Over))
```

```
## # A tibble: 2 x 3
##   Conference AverageWins AverageCarryOverCap
##   <fct>            <dbl>               <dbl>
## 1 AFC               8.12               14.6
## 2 NFC               7.88                6.56
```

These results seem to be more definitive, but we should be wary about drawing any statistically significant conclusions without going into more inferential depth. AFC conference teams had higher average wins than the NFC teams