

NFL Statistics - Chicago Bears

Vibhashit Solanki
School of Information Technology
Illinois State University
vsolank@ilstu.edu

October 20, 2014

Abstract

Chicago Bears are a professional *American football* team based in Chicago, Illinois who recorded more regular season and overall victories than any other NFL franchise.

1 Introduction

The **National Football League** or (NFL) is a professional American football league that constitutes one of the four major professional sports leagues in North America. It is composed of 32 teams divided equally between the **National Football Conference** (NFC) and the **American Football Conference** (AFC). The Super Bowl is the annual championship game of the National Football League (NFL), the highest level of professional football in the United States.

Chicago Bears was founded in Decatur, Illinois, in 1919, and moved to Chicago in 1921. Soldier Field, located on Lake Shore Drive in Chicago, is the current home of the Bears. Marc Trestman is currently the head coach of the Chicago Bears. The **Green Bay Packers** are one of the Bears' biggest rivals since their team's inception in 1920 whom they have played 186 times. The Bears currently hold the edge in head-to-head matchups with a record of 93-90-6.

2 OSEM N

OSEMN will involve Obtaining, Scrubbing, Exploring, Modelling and Interpreting data. We will be using *RStudio* as a tool to compare the stats of Chicago Bears over the years and their performance against other teams. We will also be creating a Bar Graph depicting the number of wins per year by the bears and how their points stack up against their opponents by making a Line Graph as well.

2.1 Obtaining Data

We will be using data from **Quandl** which is a search engine for numerical data. For getting Data from *Quandl API*, the User first needs to create an account on **Quandl** and get the authorization code to retrieve data from it. The code for retrieving the data is given below and you will see how the raw data looks like also using head command. However, the head command will only show the first 6 observation.

```
> #Installing quandl packages and using authcode to retrieve data
>
> library(Quandl)
> Bears <- Quandl("PROFB/NFL_CHICAGOBears",authcode="MgtZ1qcfykfk4fixy3Dz")
> head(Bears)
```

	Year	Wins	Losses	Ties	W-L%	Points	Pts0	Pt	Diff	MoV	SoS	SRS	OSRS
1	2014-12-31	0	1	0	0.000	20	23		-3	-3.0	3.0	0.0	0.0
2	2013-12-31	8	8	0	0.500	445	478		-33	-2.1	-2.1	-4.1	3.0
3	2012-12-31	10	6	0	0.625	375	277		98	6.1	0.8	6.9	1.0
4	2011-12-31	8	8	0	0.500	353	341		12	0.8	0.9	1.7	-1.5
5	2010-12-31	11	5	0	0.688	334	286		48	3.0	1.1	4.1	-0.6
6	2009-12-31	7	9	0	0.438	327	375		-48	-3.0	-0.9	-3.9	-1.9

DSRS

1	0.0
2	-7.1
3	6.0
4	3.1
5	4.7
6	-2.0

```
>
>
```

2.2 Scrubbing the Data

We only need the following columns for our analysis: Year, Wins, Points and Opponents Points(Ptso). Scrubbing or Cleaning the data is an important step so that it can be displayed properly in table as well as graph form. The code and steps for scrubbing data is given below:

```
> #manual examination reveals that we don't need columns 3 to 5
> Bears.new <- Bears[,-3:-5]
> #And now again we don't need columns 5 to 11
> Bears.new2 <- Bears.new[,-5:-10]
> #Lets change the name of one of the columnns to a more understandable one
> colnames(Bears.new2)[colnames(Bears.new2)=="Pts0"] <- "OpponentsPoints"
> #View all scrubbed observations
> Bears.new2
```

	Year	Wins	Points	Opponents	Points
1	2014-12-31	0	20		23
2	2013-12-31	8	445		478
3	2012-12-31	10	375		277
4	2011-12-31	8	353		341
5	2010-12-31	11	334		286
6	2009-12-31	7	327		375
7	2008-12-31	9	375		350
8	2007-12-31	7	334		348
9	2006-12-31	13	427		255
10	2005-12-31	11	260		202
11	2004-12-31	5	231		331
12	2003-12-31	7	283		346
13	2002-12-31	4	281		379
14	2001-12-31	13	338		203
15	2000-12-31	5	216		355
16	1999-12-31	6	272		341
17	1998-12-31	4	276		368
18	1997-12-31	4	263		421
19	1996-12-31	7	283		305
20	1995-12-31	9	392		360
21	1994-12-31	9	271		307
22	1993-12-31	7	234		230
23	1992-12-31	5	295		361
24	1991-12-31	11	299		269
25	1990-12-31	11	348		280
26	1989-12-31	6	358		377
27	1988-12-31	12	312		215
28	1987-12-31	11	356		282
29	1986-12-31	14	352		187
30	1985-12-31	15	456		198
31	1984-12-31	10	325		248
32	1983-12-31	8	311		301
33	1982-12-31	3	141		174
34	1981-12-31	6	253		324
35	1980-12-31	7	304		264
36	1979-12-31	10	306		249
37	1978-12-31	7	253		274
38	1977-12-31	9	255		253
39	1976-12-31	7	253		216
40	1975-12-31	4	191		379
41	1974-12-31	4	152		279
42	1973-12-31	3	195		334
43	1972-12-31	4	225		275
44	1971-12-31	6	185		276
45	1970-12-31	6	256		261

```

46 1969-12-31    1    210          339
47 1968-12-31    7    250          333
48 1967-12-31    7    239          218
49 1966-12-31    5    234          272

```

```

> #Make a csv file of the scrubbed data
> write.csv(Bears.new2, "IT 497 Chicago Bears.csv", row.names=FALSE)
>

```

Above code will create a csv file of the data which can be displayed in MS Excel also.

2.3 Tutorial - Analysis of Data

```

> #Summary
> summary(Bears.new2)

```

Year	Wins	Points	OpponentsPoints
Min. :1966-12-31	Min. : 0.000	Min. : 20.0	Min. : 23.0
1st Qu.:1978-12-31	1st Qu.: 5.000	1st Qu.:239.0	1st Qu.:253.0
Median :1990-12-31	Median : 7.000	Median :281.0	Median :282.0
Mean :1990-12-31	Mean : 7.408	Mean :283.8	Mean :292.2
3rd Qu.:2002-12-31	3rd Qu.:10.000	3rd Qu.:334.0	3rd Qu.:346.0
Max. :2014-12-31	Max. :15.000	Max. :456.0	Max. :478.0

A very useful multipurpose function in R is *summary(X)*, where X can be one of any number of objects, including data sets, variables, and linear models, just to name a few. As you can see in above table, a summary has been generated for each column such as Year, Wins, Points and Opponents Points. A summary depicts the min, mean, max,etc.

```

> #Class
> class(Bears.new2)

```

```
[1] "data.frame"
```

The *class function* shows the data type we are using. For example, we are using a data frame here.

```

> #Structure
> str(Bears.new2)

```

```

'data.frame':      49 obs. of  4 variables:
 $ Year           : Date, format: "2014-12-31" "2013-12-31" ...
 $ Wins           : num  0 8 10 8 11 7 9 7 13 11 ...
 $ Points         : num  20 445 375 353 334 327 375 334 427 260 ...
 $ OpponentsPoints: num  23 478 277 341 286 375 350 348 255 202 ...

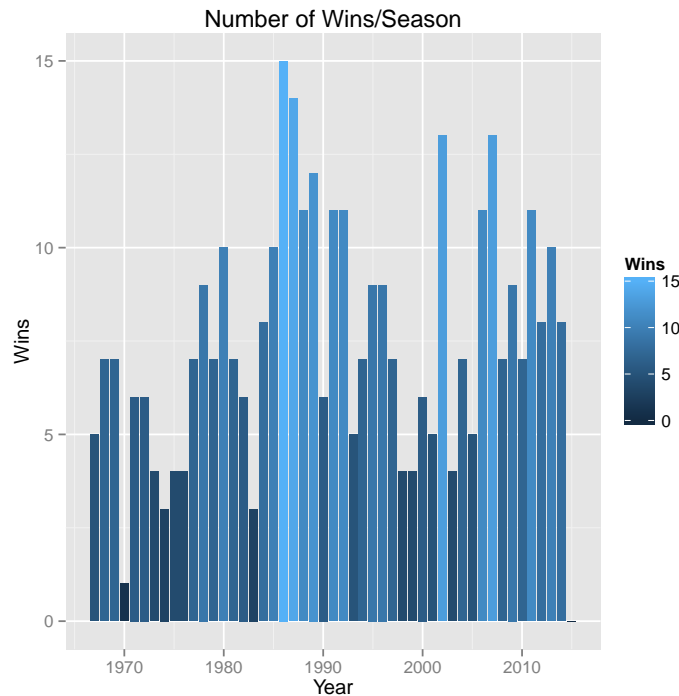
```

The *structure function* shows the the variable names and types. Here we can see the number of observations, variables and the way they are stored. For example, Date is stored in the following format "2013-12-31". Also, we can see other variables such as Wins, Points, Opponents Points and they all are stored in numeric values.

2.4 Plotting Graphs

We will be plotting two graphs or charts. One of which is a Bar Graph and the other one a Line Graph.

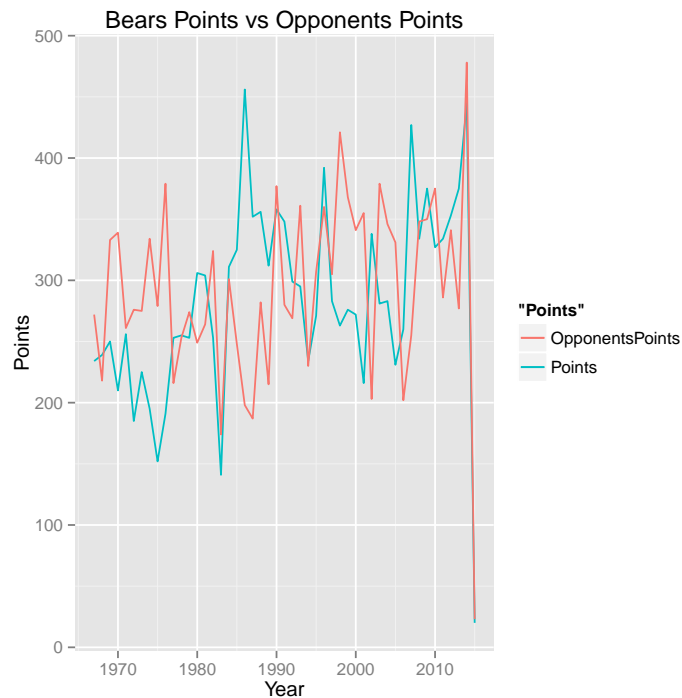
```
> #We will be using ggplot2
> library(ggplot2)
> #To create a bar chart that depicts number of wins per season
> ggplot(data=Bears.new2, aes(x=Year, y=Wins, fill=Wins)) + geom_bar(stat="identity") +
+   ggtitle("Number of Wins/Season")
>
```



The code above the graph will create a graph chart showing *number of wins per season* of Chicago Bears. We will be using Bears.new2 data and keeping Year on X axis and Wins on Y axis in the code. As you may notice, that the highest number of wins were recorded in the 1985 season when they won 15 times and became Super Bowl champions. The season was notable in that the Bears

had only one loss, the "unlucky 13th" game of the season, a Monday night affair in which they were defeated by the Miami Dolphins. Also, the lowest number of wins recorded was in the year 1969 which is also known as their struggle era. It is the worst record in the franchise history and they were only able to defeat one team known as *Pittsburgh Steelers*.

```
> #To create a line chart that shows Points and Opponents Points
> ggplot(data=Bears.new2, aes(x=Year)) + ylab("Points")+
+   geom_line(aes(y = Points, colour = "Points")) +
+   geom_line(aes(y = OpponentsPoints, colour = "OpponentsPoints")) +
+   ggtitle("Bears Points vs Opponents Points")
```



The code above the graph will create a Line chart showing *Chicago Bears's Points and their Opponents's Points*. We will be using Bears.new2 data and keeping Year on X axis and Points on Y axis in the code. As you may notice Chicago Bears have performed quite well against their opponents. If you look at the graph then you will notice that in the year 1982, they scored their lowest of 141 points. However, In the year 1985, they scored their highest of 456 points while allowing only 198 points to be scored against them. Not only did the Bears improve on that record, they completed what is still viewed by many as the best season of any team all-time. The 1985 Chicago Bears are one of the few teams to consistently challenge the undefeated 1972 Miami Dolphins for the unofficial title of the greatest NFL team of all time.