

# NBA Team Statistics - Chicago Bulls

Gunjankumar Sharma  
gsharma@ilstu.edu

October 20, 2014

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
1.1	About Basketball . . . . .	2
1.2	About NBA (National Basketball Association) . . . . .	3
1.3	About Chicago Bulls . . . . .	3
1.4	Interesting Facts about Chicago Bulls . . . . .	4
<b>2</b>	<b>Statistics in NBA</b>	<b>4</b>
<b>3</b>	<b>Reproducible Research</b>	<b>4</b>
3.1	Code . . . . .	4
3.1.1	Gathering the Data . . . . .	4
3.1.2	Scrubbing the Data . . . . .	5
3.2	Commands . . . . .	6
3.2.1	Class() . . . . .	6
3.2.2	Str() . . . . .	6
3.2.3	Summary() . . . . .	7
<b>4</b>	<b>Results</b>	<b>7</b>
4.1	Meta Data . . . . .	7
4.2	Data Set . . . . .	8
4.3	Graph . . . . .	10

# 1 Introduction

Have you ever heard about Chicago Bulls? Do you think they are the bulls from Chicago ? This document will certainly shed some light and answer these question. We will go through some data that we gather from the Internet and how we analyse, explore and graphically represent it. We will also discuss about the performance of Chicago bulls. And last but not the least, how you can reproduce this research document using statistical language R.

## 1.1 About Basketball

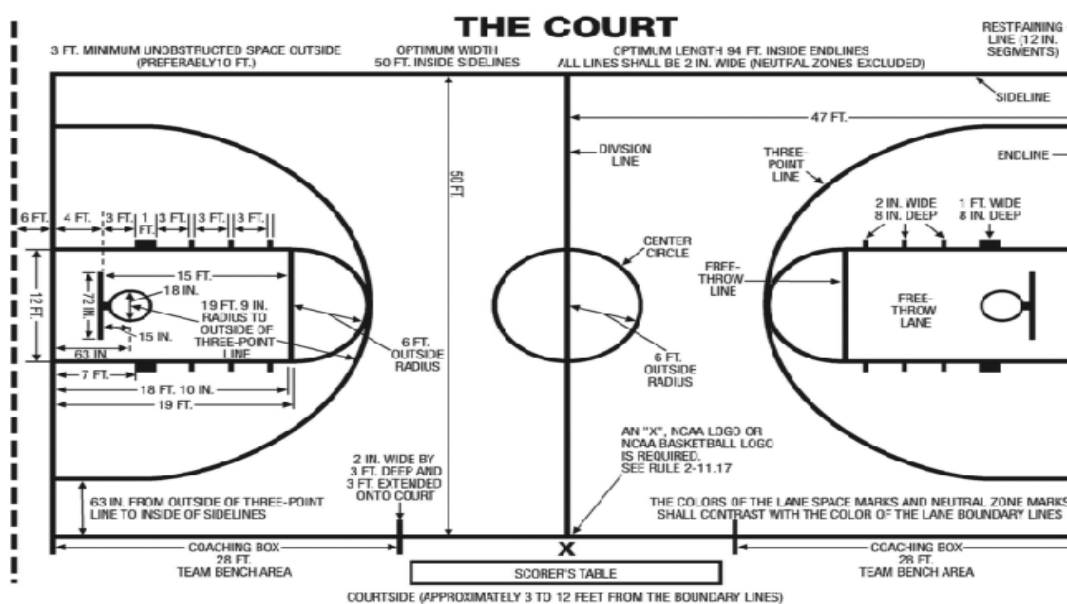


Figure 1: Basketball Court

**Basketball** is a sport played by two teams of five players on a rectangular court. The objective is to shoot a ball through a hoop 18 inches (46 cm) in diameter and 10 feet (3.0 m) high mounted to a backboard at each end. Basketball is one of the world's most popular and widely viewed sports. A team can score a field goal by shooting the ball through the basket during regular play. A field goal scores two points for the shooting team if a player is touching or closer to the basket than the three-point line, and three points (known commonly as a 3 pointer or three) if the player is behind the three-point line. The team with the most points at the end of the game wins, but additional time (overtime) may be issued when the game ends with a draw. The ball can be advanced on the court by bouncing it while walking or running or throwing it to a team mate. It is a violation to move without dribbling the ball, to carry it, or to hold the ball with both hands then resume dribbling.

## 1.2 About NBA (National Basketball Association)



Figure 2: NBA Logo

The **National Basketball Association (NBA)** is the pre-eminent men's professional basketball league in North America, and is widely considered to be the premier men's professional basketball league in the world. It has 30 franchised member clubs (29 in the United States and 1 in Canada), and is an active member of USA Basketball (USAB), which is recognized by FIBA (also known as the International Basketball Federation) as the national governing body for basketball in the United States. The NBA is one of the four major North American professional sports leagues. NBA players are the world's best paid sportsmen, by average annual salary per player.

## 1.3 About Chicago Bulls



Figure 3: Chicago Bulls Logo

The **Chicago Bulls** are an American professional basketball team based in Chicago, Illinois, playing in the Central Division of the Eastern Conference in the National Basketball Association (NBA). The team was founded on January 26, 1966. The Bulls play their home games at the United Center, also known as the "Madhouse on Madison." The Bulls saw their greatest success during the 1990s. They are known for having one of the NBA's greatest dynasties, winning six NBA championships between 1991 and 1998 with two three-peats. All six championship teams were led by Hall of Famers Michael Jordan, Scottie Pippen and coach Phil Jackson. The Bulls are the only NBA franchise to win multiple championships and never lose an NBA Finals in their history.

## 1.4 Interesting Facts about Chicago Bulls

1. They are known for having one of the NBA's greatest dynasties, winning six NBA championships between 1991 and 1998 with two three-peats.
2. The Bulls are the only NBA franchise to win multiple championships and never lose an NBA Finals in their history.
3. The Bulls won an NBA record-72 games during the 1995-96 NBA season and are the only team in NBA history to win 70 games or more in a single season. As of 2013, the Bulls were estimated to be the third most valuable NBA franchise according to Forbes, with an estimated value of \$1 billion, earning an estimated \$52.2 million in operating income in 2013.

## 2 Statistics in NBA

Statistics in basketball are kept to evaluate a player or a team's performance. Some statistics used are

- PSG : Points Scored per Game for a Team
- PAG : Points Scored Against per Game for a Team
- Wins : Number of Wins for a Team
- AST : Number of Assists made by the Team

and so on..

The section ?? talks about it in depth.

## 3 Reproducible Research

### 3.1 Code

#### 3.1.1 Gathering the Data

To install packages use the following command

```
> install.packages(c("Quandl","ggplot2","xtable"))
```

These are the packages being used.

```
> options(width=80)
> library(Quandl)
> library(ggplot2)
> library(xtable)
```

Getting the data from Quandl website using its API To get the authcode you need to signup on the Quandl website which is free. After signing up you can find the Auth token in your Profile > Account Settings.

If we want to run a code chunk once and the output for when you compile the document again, rather than running the code chunk every time, set the option `cache=TRUE`. When the chunk runs for the first time, the output is stored in a subdirectory called `cache` and the same is referred for all the later times, unless the chunk or its options are changed.

```
> bulls = Quandl("BBALL/NBA_CHICAGOBULLS",authcode="Jxdd68xhZxCpy3z-4ri9")
```

If you would like to generate a csv file for the raw data that we just fetched in your local workspace, use the following command :

```
> write.csv(bulls, file = "initialData.csv", row.names = FALSE)
```

Following command shows us the first few rows of the data

```
> head(bulls)
```

	Year	Wins	Losses	W-L%	GB	PS/G	PA/G	SRS
1	2014-12-31	48	34	0.585	8.0	93.7	91.8	1.20
2	2013-12-31	45	37	0.549	4.5	93.2	92.9	-0.01
3	2012-12-31	50	16	0.758	NA	96.3	88.2	7.43
4	2011-12-31	62	20	0.756	NA	98.6	91.3	6.53
5	2010-12-31	41	41	0.500	20.0	97.5	99.1	-1.63
6	2009-12-31	41	41	0.500	25.0	102.2	102.5	-0.16

### 3.1.2 Scrubbing the Data

Manual examination reveals that we want all rows but not columns 3,4 and 5

```
> bulls.1 <- bulls[, -3:-5]
```

We still do not need the SRS column, hence removing it.

```
> bulls.final <- bulls.1 [, -5:-5]
```

Renaming the column names from PS/G , PA/G to PSG and PAG, respectively.

```
> colnames(bulls.final)[3] <- "PSG"
> colnames(bulls.final)[4] <- "PAG"
```

We run the below command to check the first few rows of our final data

```
> head(bulls.final)
```

	Year	Wins	PSG	PAG
1	2014-12-31	48	93.7	91.8
2	2013-12-31	45	93.2	92.9
3	2012-12-31	50	96.3	88.2
4	2011-12-31	62	98.6	91.3
5	2010-12-31	41	97.5	99.1
6	2009-12-31	41	102.2	102.5

Just in case you need this scrubbed data, use the following command to generate a csv file in your local workspace.

```
> write.csv(bulls.final, file = "finalData.csv", row.names = FALSE)
```

We will go through a list of commands to know more about the data that we have.

## 3.2 Commands

### 3.2.1 Class()

In every computer language variables provide a means of accessing the data stored in memory. R does not provide direct access to the computer's memory but rather provides a number of specialized data structures we will refer to as objects. These objects are referred to through symbols or variables. In R, however, the symbols are themselves objects and can be manipulated in the same way as any other object. This is different from many other languages and has wide ranging effects. The following data objects exist in R:

- Vectors
- Lists
- Arrays
- Matrices
- Tables
- Data frames

We will now check the class of our final data object with the help of following command:

```
> class(bulls.final)
```

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

Looking at the above result, we can conclude that the class of "bulls.final" that is an object which we have used to store our data is a Data frame. A data frame is a table, or two-dimensional array-like structure, in which each column contains measurements on one variable, and each row contains one case. As we shall see, a "case" is not necessarily the same as an experimental subject or unit, although they are often the same. Technically, in R a data frame is a list of column vectors, although there is only one reason why you might need to remember such an arcane thing. Unlike an array, the data you store in the columns of a data frame can be of various types. I.e., one column might be a numerical variable, another might be a factor, and a third might be a character variable. All columns have to be the same length (contain the same number of data items)

### 3.2.2 Str()

Compactly display the internal structure of an R object, a diagnostic function and an alternative to summary (and to some extent, dput). Ideally, only one line for each basic structure is displayed. It is especially well suited to compactly display the (abbreviated) contents of (possibly nested) lists. The idea is to give reasonable output for any R object.

```
> str(bulls.final)
```

```
'data.frame':      45 obs. of  4 variables:
 $ Year: Date, format: "2014-12-31" "2013-12-31" ...
 $ Wins: num  48 45 50 62 41 41 33 49 41 47 ...
 $ PSG : num  93.7 93.2 96.3 98.6 97.5 ...
 $ PAG : num  91.8 92.9 88.2 91.3 99.1 ...
```

Looking at the above result, it is quite clear that the first column is "Year" which is in "Date" format and has values like "2014-12-31". The second column is "Wins" which is a "numeric" field and has multiple values like 48,52 and so on. Likewise, we can understand the rest of the columns that are present in the data set.

### 3.2.3 Summary()

Summary is used to get a summary of a model object's contents. The Summary() command is helpful for seeing the basic descriptive statistics for all the variables in a data frame and also the variable types. We use the following command to explore the summary of the data frame.

```
> options(width=80)
> summary(bulls.final)
```

Year	Wins	PSG	PAG
Min. :1970-12-31	Min. :13.00	Min. : 81.9	Min. : 88.2
1st Qu.:1981-12-31	1st Qu.:31.00	1st Qu.: 96.7	1st Qu.: 94.9
Median :1992-12-31	Median :45.00	Median :102.2	Median : 99.1
Mean :1992-12-30	Mean :42.84	Mean :101.2	Mean :100.3
3rd Qu.:2003-12-31	3rd Qu.:51.00	3rd Qu.:106.6	3rd Qu.:105.0
Max. :2014-12-31	Max. :72.00	Max. :114.9	Max. :116.7

Summary also shows a glimpse of statistics, including the mean, standard deviation, range, and percentiles of the values within the columns of a data set.

For example, let's consider the data present in "Wins" column. The minimum value for the data rows in the "Wins" is 13. The calculated Median for all the data rows of Wins is "45.0". The mean for all the data rows of Wins is "42.84"

## 4 Results

### 4.1 Meta Data

**Year** is a date column that can have values from 1st January 1900 to 31st December 2014.

**Wins**, that is a numeric value can have any values between 0 and 82. The maximum number of games that a NBA team can play in a year is 82 and hence, highest possible number of wins can be 82.

**PSG** i.e nothing but points scored by Chicago bulls in an year, that is a numeric field can have values from 0 to 150. It is calculated by dividing the total number of points(team) by number of games(team).

Column Name	Data Type	Minimum Value	Maximum Value
Year	Date	1900-01-01	2014-10-22
Wins	Number	0	82
PSG	Number	0	NA
PAG	Number	1	NA

Table 1: Metadata Information

**PAG** i.e nothing but points scored against Chicago bulls in an year, that is a numeric field can have values from 0 to 150. It is calculated by dividing the total number of points against(team) by number of games(team).

## 4.2 Data Set

Table 2 shows how the data looks like that we have gathered and cleaned. We have collected Data starting from the year 1970 to 2014. Chicago Bulls had the highest number of wins in the year 1972, whereas 1999 being the worst performance year for them. The data also shows decrease in the number of points scored per game after 1997. Moreover, the number of points scored against the Chicago Bulls have reduced gradually.



Year	Wins	PSG	PAG
2014-12-31	48	93.70	91.80
2013-12-31	45	93.20	92.90
2012-12-31	50	96.30	88.20
2011-12-31	62	98.60	91.30
2010-12-31	41	97.50	99.10
2009-12-31	41	102.20	102.50
2008-12-31	33	97.30	100.40
2007-12-31	49	98.80	93.80
2006-12-31	41	97.80	97.20
2005-12-31	47	94.50	93.40
2004-12-31	23	89.70	96.00
2003-12-31	30	95.00	100.10
2002-12-31	21	89.50	98.00
2001-12-31	15	87.60	96.70
2000-12-31	17	84.80	94.20
1999-12-31	13	81.90	91.40
1998-12-31	62	96.70	89.60
1997-12-31	69	103.10	92.30
1996-12-31	72	105.20	92.90
1995-12-31	47	101.50	96.70
1994-12-31	55	98.00	94.90
1993-12-31	57	105.20	98.90
1992-12-31	67	109.90	99.50
1991-12-31	61	110.00	101.00
1990-12-31	55	109.50	106.20
1989-12-31	47	106.40	105.00
1988-12-31	50	105.00	101.60
1987-12-31	40	104.80	103.90
1986-12-31	30	109.30	113.10
1985-12-31	38	108.70	109.60
1984-12-31	27	103.70	108.90
1983-12-31	28	111.00	115.90
1982-12-31	34	106.60	108.60
1981-12-31	45	109.00	107.00
1980-12-31	30	107.50	110.20
1979-12-31	31	104.70	108.70
1978-12-31	40	103.90	104.80
1977-12-31	44	98.90	98.00
1976-12-31	24	95.90	98.80
1975-12-31	47	98.10	95.00
1974-12-31	54	102.00	98.70
1973-12-31	51	104.10	100.60
1972-12-31	57	111.20	102.90
1971-12-31	51	110.60	105.40
1970-12-31	39	114.90	116.70

Table 2: Data set for Chicago Bulls

### 4.3 Graph

Following command is used to plot a line graph based on the data that we have. As the legend clearly states, we have three distinct lines showing Number of wins, Points scored per game by the Bulls team and lastly points scored against the Bulls team per game.

```
> ggplot(data=bulls.final, aes(x=Year)) + ylab("") +  
+   geom_line(aes(y = Wins, colour = "Wins")) +  
+   geom_line(aes(y = PSG, colour = "Points Scored/Game"),size=0.5)+  
+   geom_line(aes(y = PAG, colour = "Points Against/Game"),size=0.5) +  
+   #ggtitle("Performance of Chicago Bulls") +   # Set title  
+   scale_colour_hue(name="Legend ") # Renaming the Title of legend
```

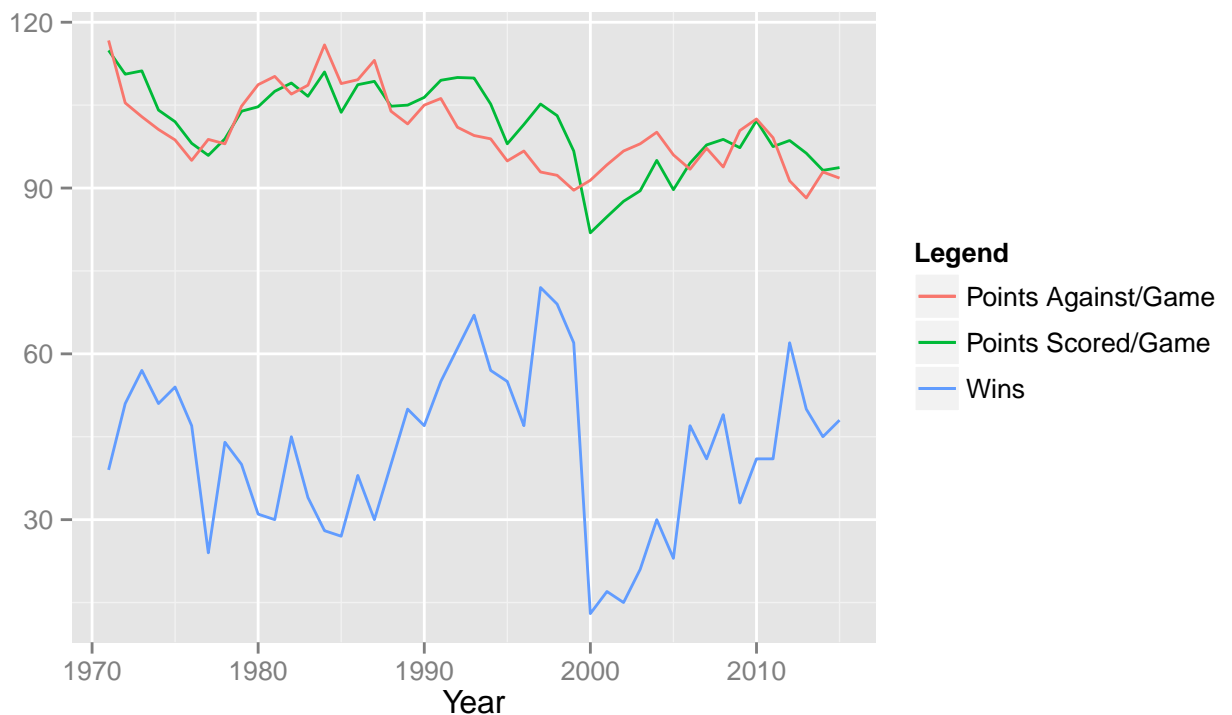


Figure 4: Performance of Chicago Bulls

The graph in Figure 4 tells us about the performance of Chicago Bulls. The Chicago Bulls joined the NBA for the 1966-67 season. The franchise struggled for the better part of a quarter century, occasionally putting excellent teams on the court, such as the tough units of the mid-1970s that featured Bob Love, Norm Van Lier, Jerry Sloan, and Tom Boerwinkle. More often, however, the Bulls worked hard for mediocre results. That all changed in the mid-1980s with the drafting of Michael Jordan, the dominant player of his era and possibly the greatest player of all time.

Jordan won seven straight scoring titles with a combination of breathtaking slam dunks and a bag of thrilling shot-making tricks. He put up some of the biggest numbers in NBA history and wrote some of the most memorable chapters in the annals of the league. The addition of Scottie Pippen, another

Hall of Famer, in 1987 would set the stage for one of sport's great dynasties.

In the early 1990s, the Bulls assembled a strong supporting cast for Jordan and Pippen which won three consecutive NBA titles, becoming only the third franchise in history to string together a trio of crowns. After more than a year of "retirement" to try his hand at professional baseball, Jordan returned to lead the Bulls back to another title in 1996, one more in 1997 and a third in a row in 1998, the Bulls's second three-peat of the decade and their sixth NBA championship trophy.

Ten years later, Chicago beat overwhelming odds to win the NBA Draft Lottery, meaning the Bulls had the top pick in the 2008 NBA Draft. Point guard Derrick Rose was the team's choice, and he has since led the Bulls to the playoffs in each of his professional campaigns, including a trip to the 2011 Eastern Conference Finals.

The list below summarizes the history of Chicago Bulls

1966–1974: Early success

1976–1984: Gilmore and Theus

1984–1998: The Michael Jordan Era

1999–2004: Rebuilding

2004–2007: Resurgence

2007–2008: Missing the playoffs

2008–present: The Derrick Rose era

THE END