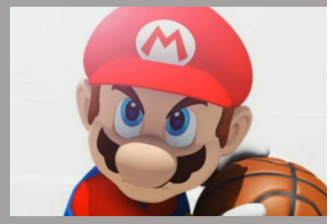
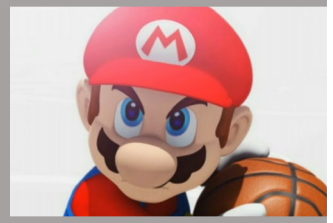
Basketball IV

Produced by Dr. Mario

UNC STOR 538

NBA Salaries

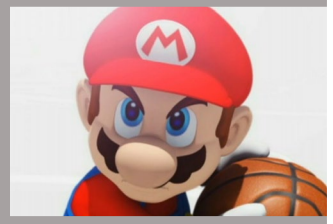


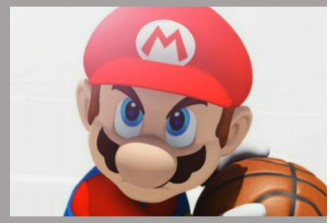


- Recall Baseball Salary Estimation
  - Based on WAR in Baseball
  - Assumed Replacement Player Costs $500,000
  - Team of Replacement Players Cost $12.5M (48-114 Record)
  - Average Team's Salary was $114M (81-81 Record)
  - $101.5M Needed for Replacement Team to Get to Average
- Real Plus-Minus (RPM)
  - Designed by Jermias Engelmann and Steve Ilardi
  - Utilized Modified Ridge Regression to Shrink Coefficients Toward the Box Plus-Minus of the Player
  - Leaders in 2018-2019

RANK	NAME	TEAM	GP	MPG	ORPM	DRPM	RPM
1	Paul George, SG	OKC	77	36.9	4.55	3.08	7.63
2	James Harden, SG	HOU	78	36.8	7.4	0.02	7.42
3	Stephen Curry, PG	GS	69	33.8	5.99	0.85	6.84
4	Giannis Antetokounmpo, PF	MIL	72	32.8	3.16	3.53	6.69

NBA Salaries





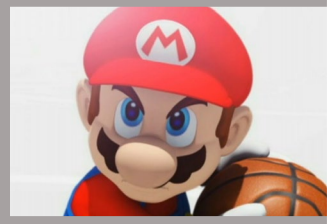
- **Interpretation of RPM**
  - Numbers are Per 100 Possessions
  - Giannis RPM = 6.69
  - If Giannis Replaced an Average Player, then his Team Improves by 6.69 Points Over the Opponent Per 100 Possessions
  - RPM of an Average Player = 0
  - RPM of a Replacement Player = -3.1 (Equivalent to 10 Percentile)
- **Team of Replacement Players**
  - Deficit Versus an Average Team

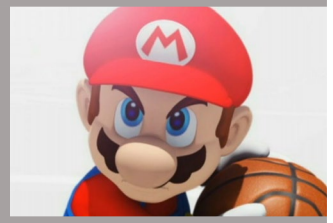
$$5(-3.1) = -15.5 \text{ Points Per 100 Possessions}$$

  - Average Pace in 2017-2018 = 96 Possessions Per Game
  - Conversion of Deficit Per 100 Possessions to Per Game

$$\left(-\frac{15.5}{100}\right) * 96 = -14.88 \text{ Points Per Game}$$

NBA Salaries





- **Replacement Team Versus Average Team**
  - Average Team Scored 105.6 Points Per Game
  - Expected Final Score: 90.72 to 105.6 (Difference of 14.88)
  - Scoring Ratio

$$\frac{90.72}{105.6} = 0.86$$

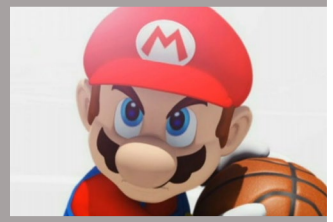
  - Basketball Pythagorean Theorem From Chapter 1 ($\alpha = 14$)

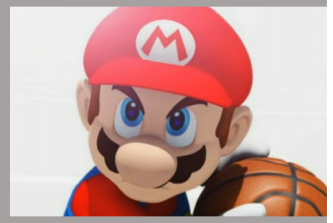
$$\frac{0.86^{14}}{0.86^{14} + 1} = 10.7\%$$

  - Conclusion: Expect Replacement Team to Win 10.7% of Games

$$Final\ Record = 8.7\ Wins\ and\ 73.3\ Losses$$

NBA Salaries



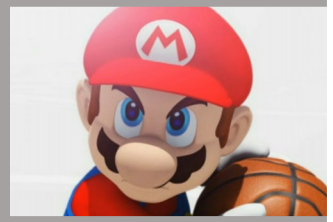


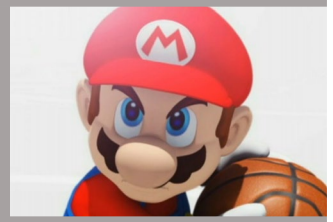
- Application to NBA Salaries (Based off 2017-2018)
  - Average Team Payroll Was Approximately $93M
  - Minimum Player Salary Between $500K and $1.5M
  - Assume Average Minimum= $1M
  - Payroll of Replacement Team = $12M
  - Costs $93M - $12M = $81M to Go From Replacement to Average
  - This is Equivalent to Go From 9 Wins to 41 Wins
  - Equivalent:

$$32\ Wins = \$81M$$

  - For Simplicity/Laziness, $32\ Wins = \$80M$
  - Each Win Above Replacement is Worth $2.5 Million

NBA Salaries





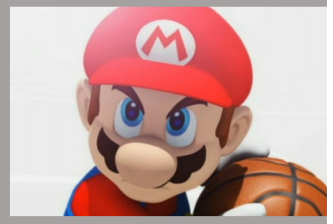
- Calculation of Fair Salary
  - Suppose Player Generated 20 Wins in 2016-2017

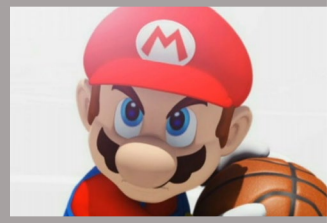
$$Fair\ Salary = 20 * 2.5 = \$50M$$

  - Descriptive Versus Predictive Metrics
  - Book Calculates Fair Salary of a Player in a 2018 Based off Generated Wins From Previous Season
  - What is the Problem Here?

	Player Wins This Year	Player Wins Next Year
Team Salaries This Year	Fair Pay in Previous Year	Not Helpful?
Team Salaries Next Year	Determining Fair Salary	Fair Pay in Next Year

NBA Salaries





- NBA Salary Information Across the Years
  - Data from Basketball-Reference.com
  - Data Preview

```
head(Salary)
A tibble: 6 x 7
league player_id  salary season  season_end season_start team
<chr>  <chr>       <int> <chr>        <int>        <int> <chr>
NBA    abdelal01  395000 1990-91       1991         1990 Portland Trail Bl~
NBA    abdelal01  494000 1991-92       1992         1991 Portland Trail Bl~
NBA    abdelal01  500000 1992-93       1993         1992 Boston Celtics
NBA    abdelal01  805000 1993-94       1994         1993 Boston Celtics
NBA    abdelal01  650000 1994-95       1995         1994 Sacramento Kings
NBA    abdulka01 1530000 1984-85       1985         1984 Los Angeles Lakers
```

  - Salary Summarized by Season

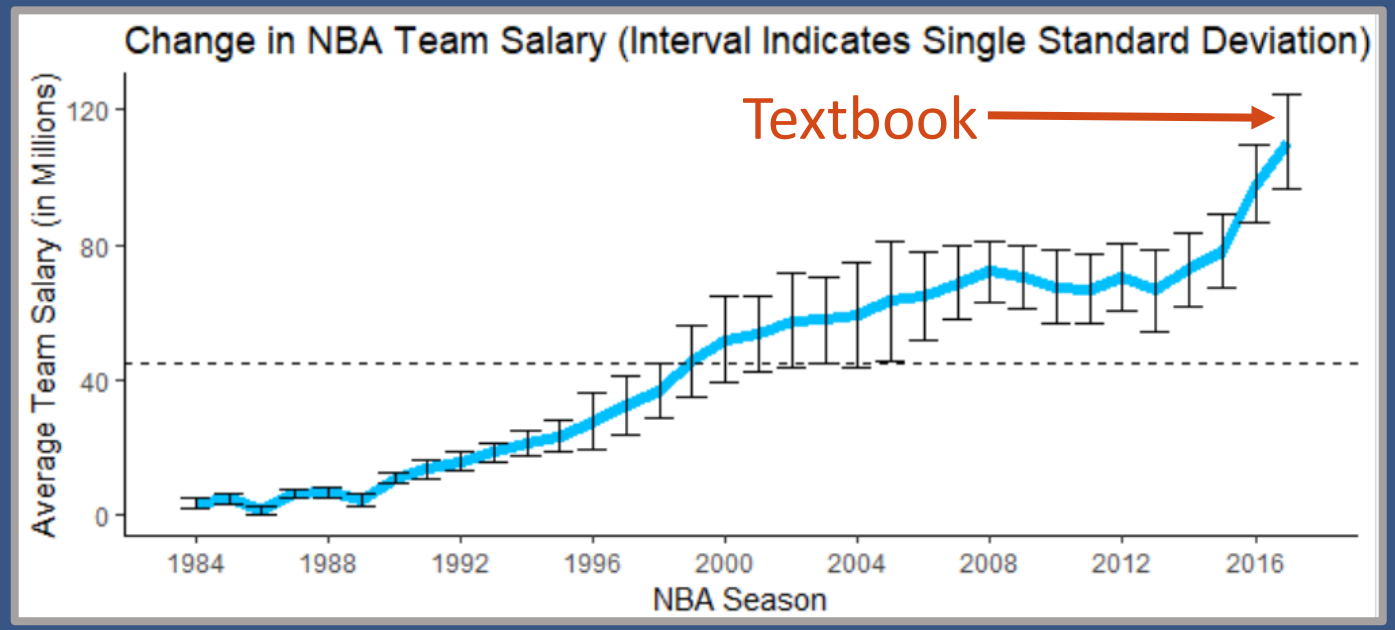
```
head(Salary.Data)
A tibble: 6 x 4
season_start     n mean.salary sd.salary
       <int> <int>       <dbl>     <dbl>
        1984    24        3.49      1.62
        1985    23        4.76      1.41
        1986    16        1.36      1.26
        1987    23        6.05      1.26
        1988    25        6.78      1.45
        1989    24        4.46      2.13
```

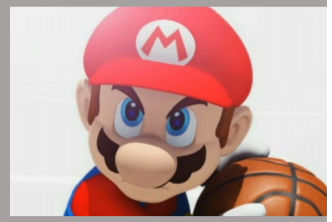
NBA Salaries

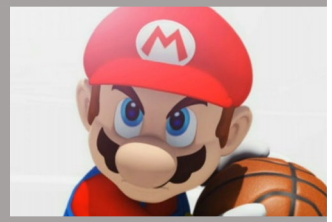
- NBA Salary Information Across the Years
  - Code for Summary Table

```
Salary.Data=Salary %>%
  group_by(team,season_start) %>%
  summarize(total.salary=sum(salary)/1000000) %>%
  ungroup() %>%
  group_by(season_start) %>%
  summarize(n=n(),mean.salary=mean(total.salary),
            sd.salary=sd(total.salary))
```

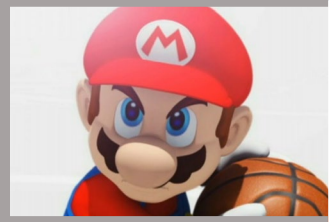
  - Figure Showing Change

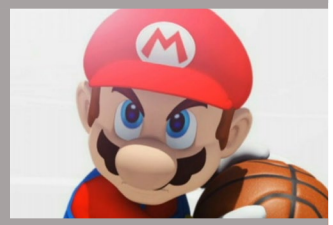






NBA Salaries





- Pythagorean Theorem For Basketball
  - Modeling Win Percentage Using Points

$$Win\ \% \approx \frac{\left(\frac{Points\ Scored}{Points\ Allowed}\right)^{\alpha}}{\left(\frac{Points\ Scored}{Points\ Allowed}\right)^{\alpha} + 1}$$

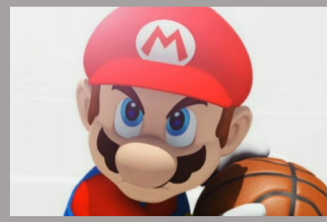
  - From Textbook, α=14 Based on Data
  - Question: Can We Confirm This?
  - Data from 2014 to 2018 Found on Kaggle

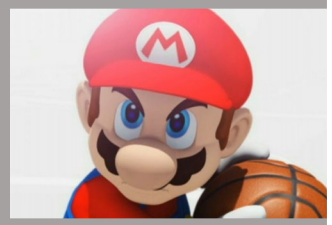
```
head(Games[,2:9])
```

A tibble: 6 x 8

Team	Game	Date	Home	Opponent	WINorLOSS	TeamPoints	OpponentPoints
<chr>	<int>	<date>	<chr>	<chr>	<chr>	<int>	<int>
ATL	1	2014-10-29	Away	TOR	L	102	109
ATL	2	2014-11-01	Home	IND	W	102	92
ATL	3	2014-11-05	Away	SAS	L	92	94
ATL	4	2014-11-07	Away	CHO	L	119	122
ATL	5	2014-11-08	Home	NYK	W	103	96
ATL	6	2014-11-10	Away	NYK	W	91	85

NBA Salaries



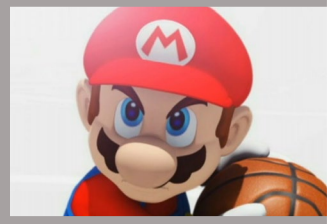


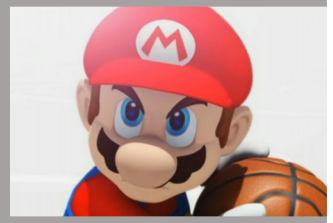
- Pythagorean Theorem For Basketball
  - Modifying Data for Estimating α

```
Games2 = Games %>%
            mutate(Season=rep(c(2014,2015,2016,2017),each=82*30)) %>%
            group_by(Team,Season) %>%
            summarize(Win.Per=mean(WINorLOSS=="W"),
                      Scored=mean(TeamPoints),
                      Allowed=mean(OpponentPoints))
```

```
head(Games2)
A tibble: 6 x 5
Groups:   Team [2]
Team  Season Win.Per Scored Allowed
<chr>  <dbl>   <dbl>  <dbl>   <dbl>
ATL     2014   0.732   103.    97.1
ATL     2015   0.585   103.    99.2
ATL     2016   0.524   103.   104.
ATL     2017   0.293   103.   109.
BOS     2014   0.488   101.   101.
BOS     2015   0.585   106.   103.
```

NBA Salaries





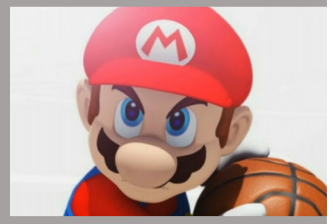
- Pythagorean Theorem For Basketball
  - Minimize Sum of Squares (Predicted Win % Versus Actual Win %)

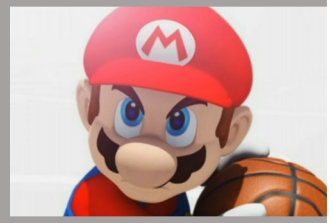
```
pythag.func=function(data,par){
  R=data$Scored/data$Allowed
  y=data$Win.Per
  resid=y-(R^(par[1]))/(R^(par[1])+1)
  return(sum(resid^2))
}

result=optim(par=c(13),fn=pythag.func,data=Games2,method="BFGS")
```

  - Based on Recent Data, Best α is 14.4564

```
print(result$par[1])
] 14.4564
```

NBA Salaries





- **Cost of Winning (Based on Book)**
  - Cost $12M to Get 9 Wins Over 82 Games
  - $12M is $81M Less Than the Average Salary (2017-2018)
  - Assumption: Costs $81M to Be Average
  - This Implies:

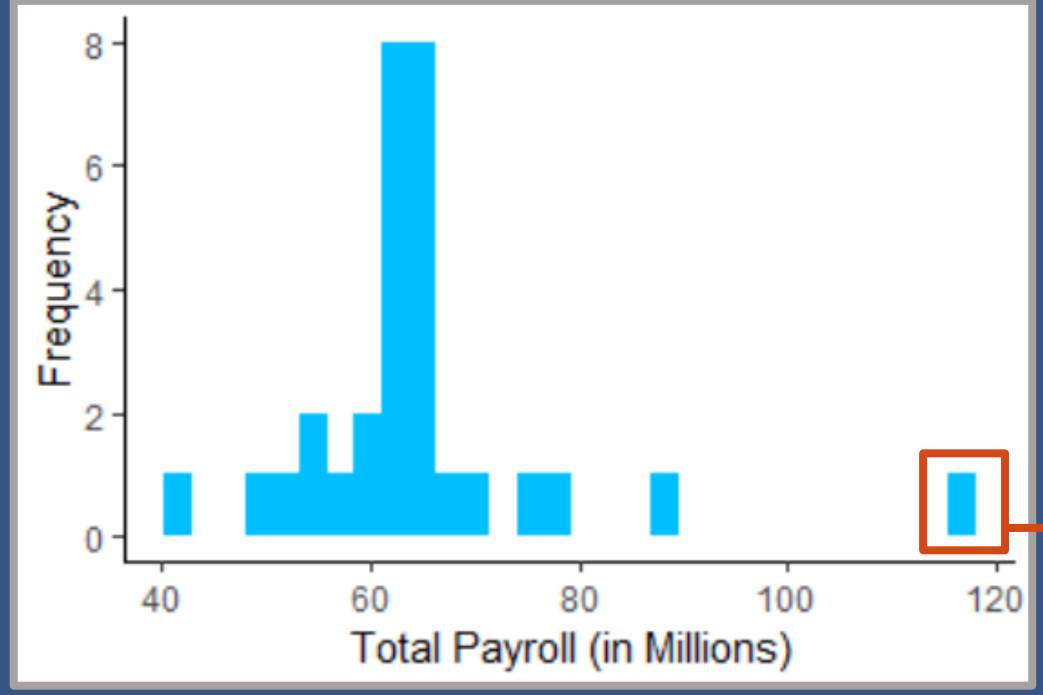
$$Price\ Per\ Win = \frac{\$81M}{41 - 9} = \$2.5M$$

- **Criticism 1**
  - Team Salaries are Highly Skewed and Influenced by Outliers
  - Recommendation: Use Median
- **Criticism 2**
  - Average Salary may not be the Salary of an Average Team
  - Recommendation: Regress Salary on Wins and Predict When Wins = 41

NBA Salaries

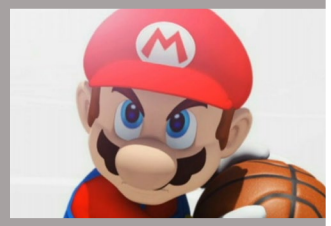
- Observe Interesting Data From 2006

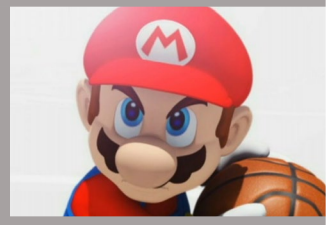
```
Salary06 = Salary %>%
  filter(season_start==2006) %>%
  group_by(team) %>%
  summarize(total.salary=sum(salary)/1000000) %>%
  arrange(desc(total.salary))
```




```
head(Salary06)
A tibble: 6 x 2
team                        total.salary
<chr>                              <dbl>
New York Knicks                    117.
Dallas Mavericks                    88.4
Los Angeles Lakers                  77.1
Portland Trail Blazers              75.0
Philadelphia 76ers                  69.1
Minnesota Timberwolves              66.8
```

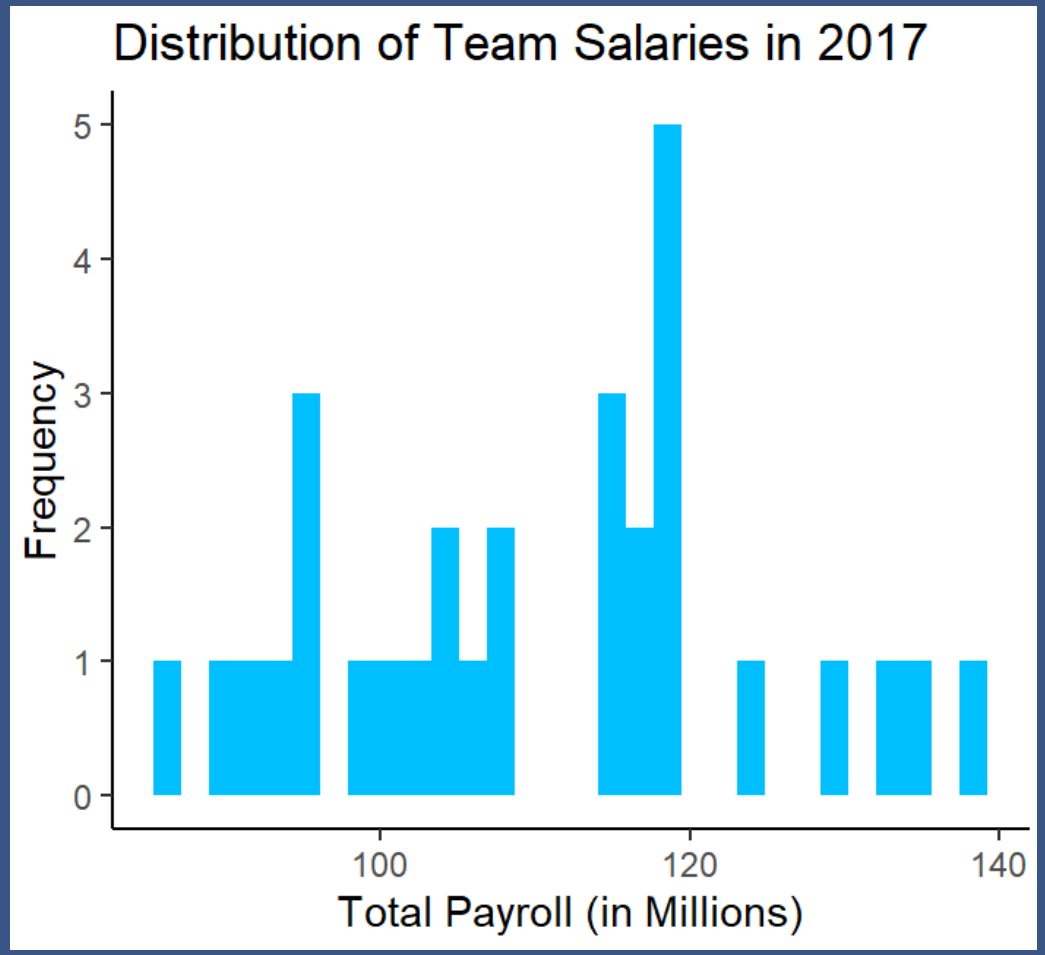
Team: New York Knicks
Payroll: $117M
Record: 33-49
Conclusion: Idiots

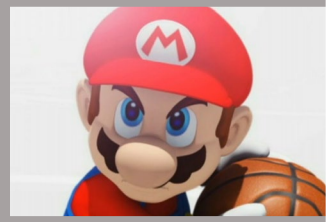


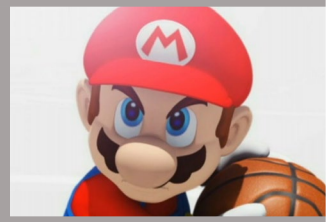


NBA Salaries

- **Fix Based on Criticism 1**







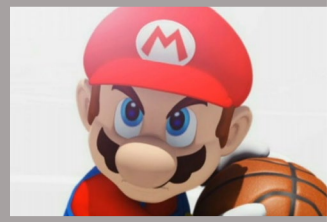
2017-2018 Season

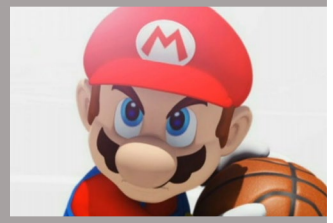
Average Salary: $110M

Fix

Median Salary: $111M

NBA Salaries





- Getting Wins and Losses Into Data
  - Scraping Team Records From Wikipedia

```
wikipedia="https://en.wikipedia.org/wiki/2017%E2%80%9318_NBA_season"
wins = wikipedia %>%
        read_html() %>%
        html_table(fill=T)

wins2=as.data.frame(rbind(as.matrix(wins[[4]]),as.matrix(wins[[5]]),
                    as.matrix(wins[[6]]),as.matrix(wins[[7]]),
                    as.matrix(wins[[8]]),as.matrix(wins[[9]])))[,1:2]
names(wins2)=c("team","wins")
str_detect(wins2$team,"..\\p{Pd}.")
wins3=mutate(wins2,team=str_replace(team,"..\\p{Pd}.",""))
```

```
head(wins3)
                    team wins
    Toronto Raptors    59
     Boston Celtics    55
 Philadelphia 76ers    52
    New York Knicks    29
      Brooklyn Nets    28
Cleveland Cavaliers    50
```

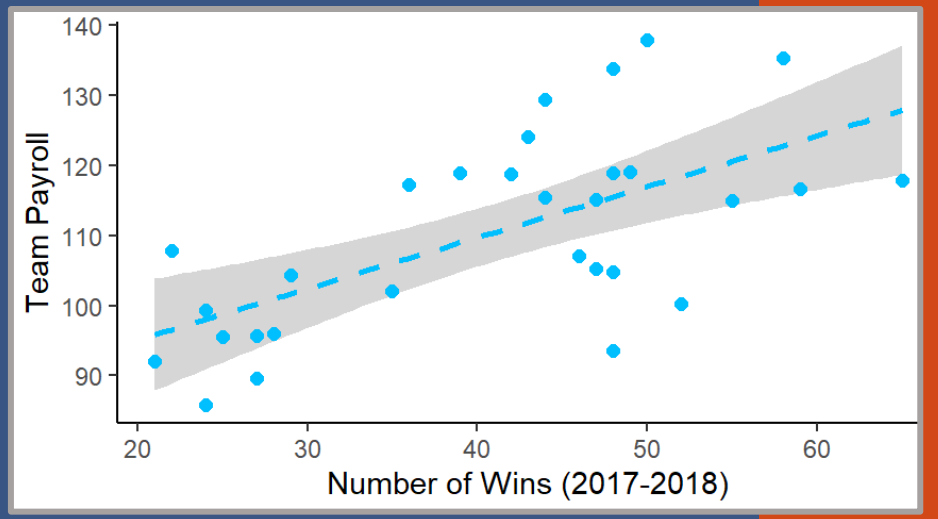
  - Merging Datasets

```
salarywins17=inner_join(Salary17,wins3)
)ining, by = "team"
salarywins17$wins=as.numeric(as.character(salarywins17$wins))
head(salarywins17)
A tibble: 6 × 3
team                    total.salary  wins
<chr>                          <dbl> <dbl>
Cleveland Cavaliers             138.    50
Golden State Warriors           135.    58
Oklahoma City Thunder           134.    48
Miami Heat                      129.    44
Washington Wizards              124.    43
Portland Trail Blazers          119.    49
```

NBA Salaries

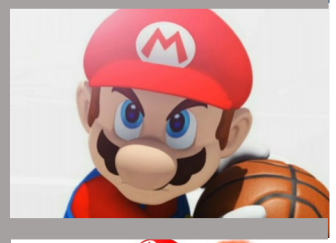
- Fix Based on Criticism 2
  - Linear Regression Model and Fit

```
lm(formula = total.salary ~ wins, data = salarywins17)

Residuals:
    Min      1Q  Median      3Q     Max
-21.930  -7.004  -1.433   9.382  20.911

Coefficients:
            Estimate Std. Error t value            Pr(>|t|)
(Intercept)  80.5493     7.1444  11.274 0.00000000000639 ***
wins          0.7281     0.1672   4.354            0.000161 ***
```




```
predict(linearwts,newdata=data.frame(wins=41),interval="confidence")
      fit      lwr      upr
110.4022 106.2861 114.5183

predict(linearwts,newdata=data.frame(wins=41),interval="prediction")
      fit      lwr      upr
110.4022 87.48456 133.3198
```

- Prediction for 41 Wins is Almost Identical to Actual Average Salary
- What is the Value of Knowing the Lower and Upper Limits?




Final Inspiration

There is no “I” in team,
but there is in win.

- Michael Jordan