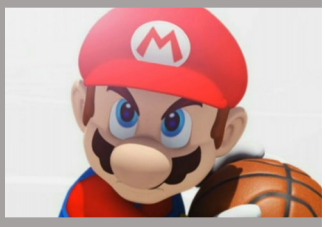# Basketball IV

Produced by Dr. Mario

UNC STOR 538

NBA Salaries

- **Recall Baseball Salary Estimation**
  - Based on Value Over Replacement Player (VORP)
  - Defined Replacement Player as Bottom 20%
  - Assumed Replacement Player Costs $0
  - Team of Replacement Players Cost $0 (44-118 Record)
  - Average Team's Salary was $77M (81-81 Record)
  - $77M Needed for Replacement Team to Get to Average

- **Application to NBA Salaries (Based off 2006-2007)**
  - Average Team Payroll Was Approximately $66M
  - Minimum Player Salary $400,000
  - Define Replacement Player as Bottom 10% in WINVAL
  - WINVAL of Replacement Player = -6
  - Team of Replacement Players Lose to Average Team by 30 Points
  - Average Team Scores 98.7 Points Per Game

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
```
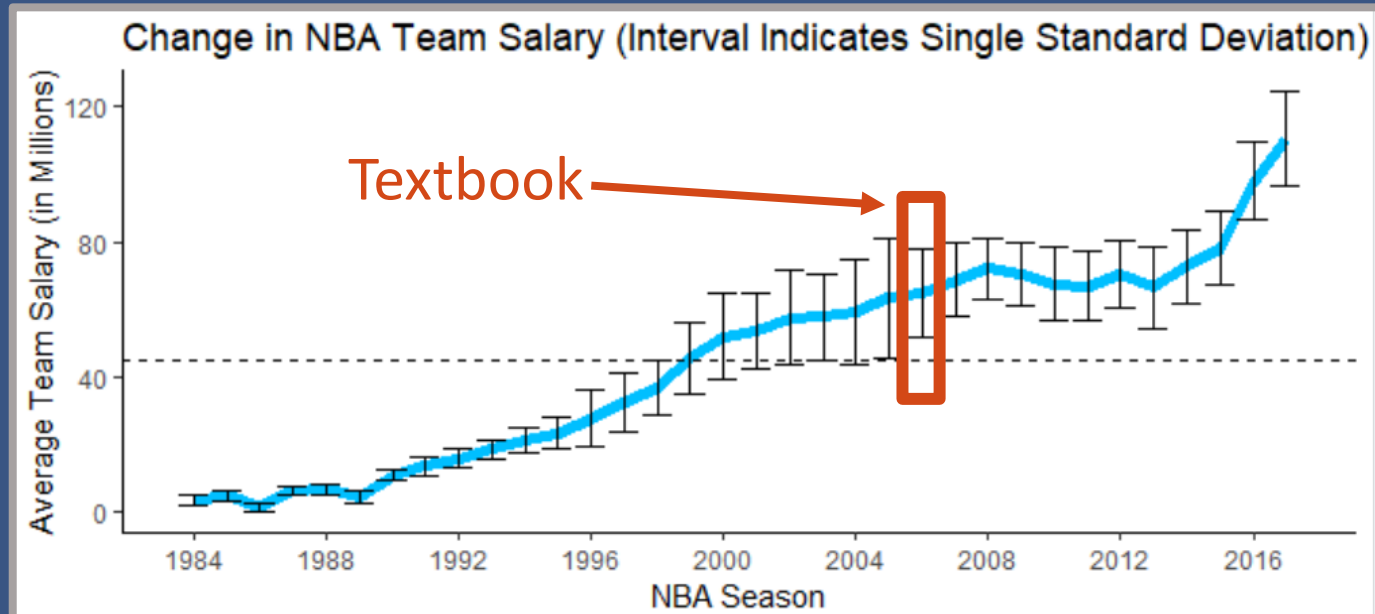
NBA Salaries

- **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**

Change in NBA Team Salary (Interval Indicates Single Standard Deviation)

Textbook

NBA Salaries

- **Pythagorean Theorem For Basketball**
  - **Modeling Win Percentage Using Points**

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

- **From Textbook, α=13.91 Based on Data Before 2007**
- **Question: Has This Result Stayed Consistent?**
- **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 %)**

```r
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**

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

NBA Salaries

- **Application of Pythagorean Theorem to Replacement Players**
  - **Losing 30 Points to Average Team Means Final Score 68.7 to 98.7**
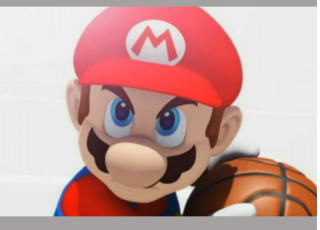 - **Winning Percentage for Team of Replacement Players**

$$Win\ \% \approx \frac{\left(\frac{68.7}{98.7}\right)^{14.4564}}{\left(\frac{68.7}{98.7}\right)^{14.4564} + 1} = 0.005282$$

```
alpha=result$par[1]
R=68.7/98.7
ReplWinPer=(R^(alpha))/(R^(alpha)+1)
```

  - **Expected Number of Wins for Team of Replacement Players**

$$Wins = (Win\ \%) \times (82\ Games) \approx 0.433$$

  - **Team of Replacement Players Would Win 0 Games**

  - **Team of Average Players Would Win 41 Games**

NBA Salaries

- **Cost of Winning (Based on Book)**
  - Cost $2M to Get 0 Wins Over 82 Games
  - $2M is $64M Less Than the Average Salary (2006-2007)
  - Assumption: Costs $64M to Be Average
  - This Implies:

$$Price\ Per\ Win = \frac{\$64M}{41} = \$1,560,976$$

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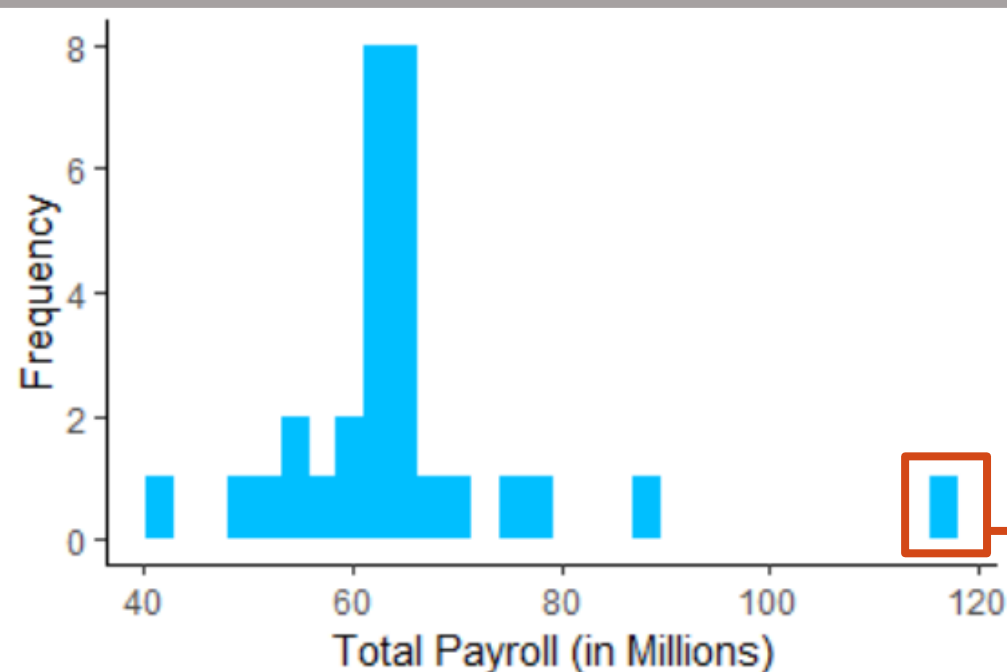 - Recommendation: Regress Salary on Wins and Predict When Wins = 41
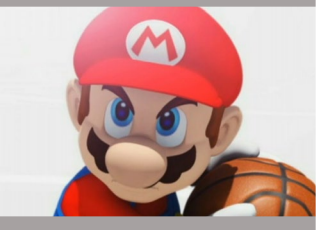
- **Criticism 2**
  - Team Salaries are Highly Skewed and Influenced by Outliers
  - Recommendation: Use Median

NBA Salaries

- ## Visual of Team Payrolls

```
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

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

```
wikipedia="https://en.wikipedia.org/wiki/2006%E2%80%9307_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,".-")
wins3=mutate(wins2,team=str_replace(team,".-",""))
```

```
head(wins3)
                    team wins
      Toronto Raptors   47
     New Jersey Nets    41
  Philadelphia 76ers    35
     New York Knicks    33
      Boston Celtics    24
     Detroit Pistons    53
```

  - ### Merging Datasets

```
salarywins06=inner_join(Salary06,wins3)
salarywins06$wins=as.numeric(as.character(salarywins06$wins))
```

```
head(salarywins06)
A tibble: 6 x 3
team                    total.salary   wins
<chr>                          <dbl>  <dbl>
New York Knicks                 117.     33
Dallas Mavericks               88.4      67
Los Angeles Lakers             77.1      42
Portland Trail Blazers         75.0      32
Philadelphia 76ers             69.1      35
Minnesota Timberwolves         66.8      32
```

NBA Salaries

- **Fix Based on Criticism 1**
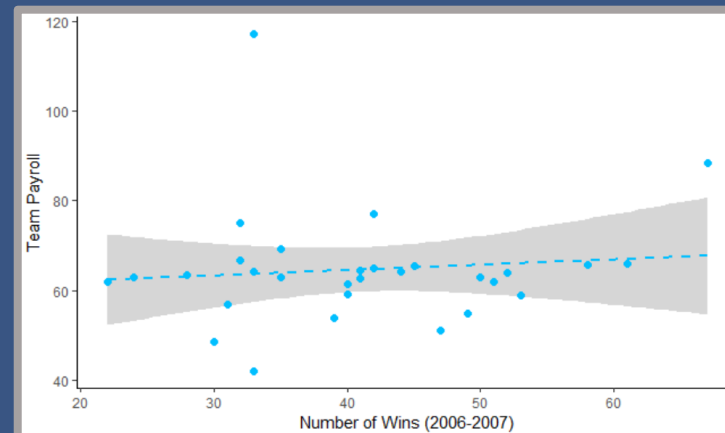 - **Linear Regression Model and Fit**

```
Call:
lm(formula = total.salary ~ wins, data = salarywins06)

Residuals:
    Min      1Q  Median      3Q     Max
-21.690  -4.943  -0.927   0.420  53.597

Coefficients:
            Estimate Std. Error t value  Pr(>|t|)
(Intercept)  59.7207     9.6061   6.217 0.00000102 ***
wins          0.1192     0.2268   0.526     0.603
```
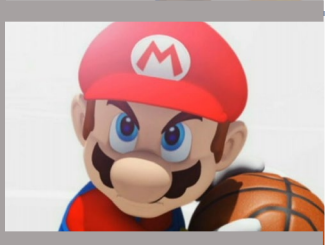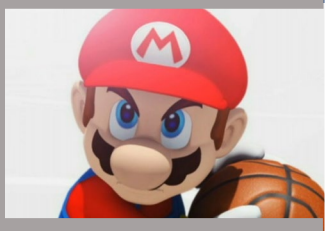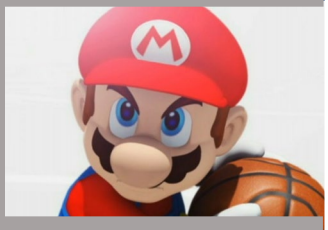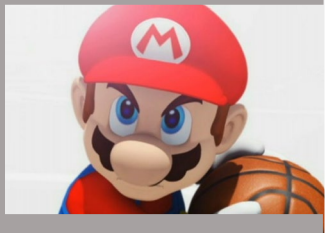

- **Prediction Under Linear Model is $64M for 41 Win Team**
- **Assumption: Costs $62M to Be Average**
- **This Implies:**

$$Price\ Per\ Win = \frac{\$62M}{41} = \$1,512,195$$

NBA Salaries

- **Fix Based on Criticism 2**
  - **Median is Robust Against Outliers (NY Knicks)**
  - **Motivation: We Don't Want One Stupid Team to Get Us to Overpay**
  - **Assumption: Costs $61M to Be Average**

  ```
  median(salarywins06$total.salary)
  1] 63.25874
  ```

  - **This Implies:**

  $$Price\ Per\ Win = \frac{\$61M}{41} = \$1,487,805$$

- **Points Over Replacement Player (PORP)**

$$PORP = (Adjusted +/-) - (-6) = (Adjusted +/-) + 6$$

Individual Player   Replacement Player

NBA Salaries

- **PORP for Average Team**
  - **For Average Player, PORP = 0+6 = +6 Points Per 48 Minutes**
  - **In a Season, There are 3,936 Minutes**
  - **For Team 5 Players, This Equates to 19,680 Minutes**
  - **For Team of 5 Average Players,**

$$PORP \times Minutes = 6 \times 19,680 = 118,080$$

- **Scaled to Number of Wins**

$$\frac{118,080}{41} = 2,880 \ Per \ Win$$

NBA Salaries

- **Determining Fair Salary Based on PORP**
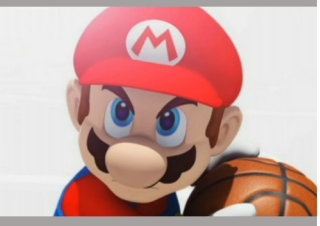 - **Wins Generated by a Player (W)**

$$W = \frac{PORP \times Minutes\ Played}{2,880}$$

  - **Fair Salary (SAL)**

$$SAL = W \times Dollars\ Per\ Win$$

  - **We Found 3 Different Choices for Dollars Per Win**
    - $1,560,976
    - $1,512,195
    - $1,487,805

NBA Salaries

- **Example: King James in 2006-2007**
  - Played 3,190 Minutes and Had an Adjusted +/- of 14.412 Per 48 Minutes
  - Therefore Lebron James Had a PORP of 20.412
  - Wins Generated by Lebron James

$$W = \frac{20.412 \times 3,190}{2,880} = 22.6$$

  - Fair Salary for Lebron James

Method	Wins	Price Per Win	Fair Salary
Book	22.6 Wins	$1,560,976	$35.3M
Regression	22.6 Wins	$1,512,195	$34.2M
Median	22.6 Wins	$1,487,805	$33.6M

Final Inspiration

There is no "I" in team, but there is in win.

- Michael Jordan