

# 2024 NBA Finals Prediction Final Project

---

A Data Science Project By  
Nate, Johan, Bin, Jelver



# Table Of Contents

---

1. Project Overview
2. Data Exploration
3. Data Visualization
4. Machine Learning
5. Correlation Matrix
6. Results / Predictions
7. Conclusions

# Project Overview

---

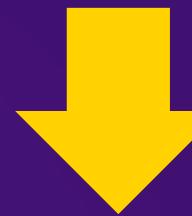
For our final project, we were asked to design a system to predict this year's 2024 NBA Championship winner. The NBA playoffs is the post-season tournament held at the end of the regular NBA season. This tournament takes the top 8 teams from each conference, 16 teams total into a best-of-7-game elimination tournament. Many fans make predictions. This is our attempt utilizing machine learning and the skills we've learned through this year of data science to predict a winning team.

```
In [1]: import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
from PIL import Image, ImageOps
from PIL import ImageOps
```



# Data Exploration

```
: df= pd.read_csv("NBA Champion LR Data.csv")
print(list(df.columns))
```



```
In [4]: three_point_attempts_per_season = df.groupby('Season')['3PA'].sum()

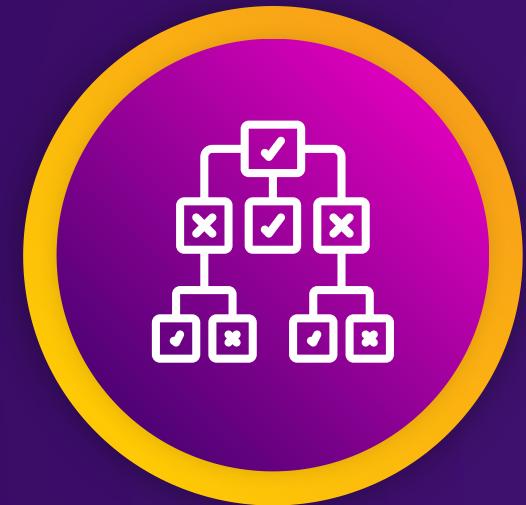
plt.figure(figsize=(20, 6))

three_point_attempts_per_season.plot(kind='bar', xlabel='Season', ylabel='Total 3-Point Attempts', title='Total 3-Po
```

To access the data and utilize it, we had to import different libraries to use the tools within those libraries to create visualizations that make the data much more digestible. Libraries such as Pandas make it much easier to code the visualization as it has a list of available methods. With Pandas and many other libraries, we could visualize most of the data we had gathered from the CSV file to show how the model would predict some outcomes.

# Criteria For Machine Learning Prediction

---



## Strong Historical Playoff Experience

Consistency in overall efficiency with post-game play. This factors in number of times a team has gone to the playoffs and how they performed (conference semi-finalists, western conference finalists).



## Good Regular Season Performance

During regular season we also check whether or not teams are just overall good and or dominant in general.



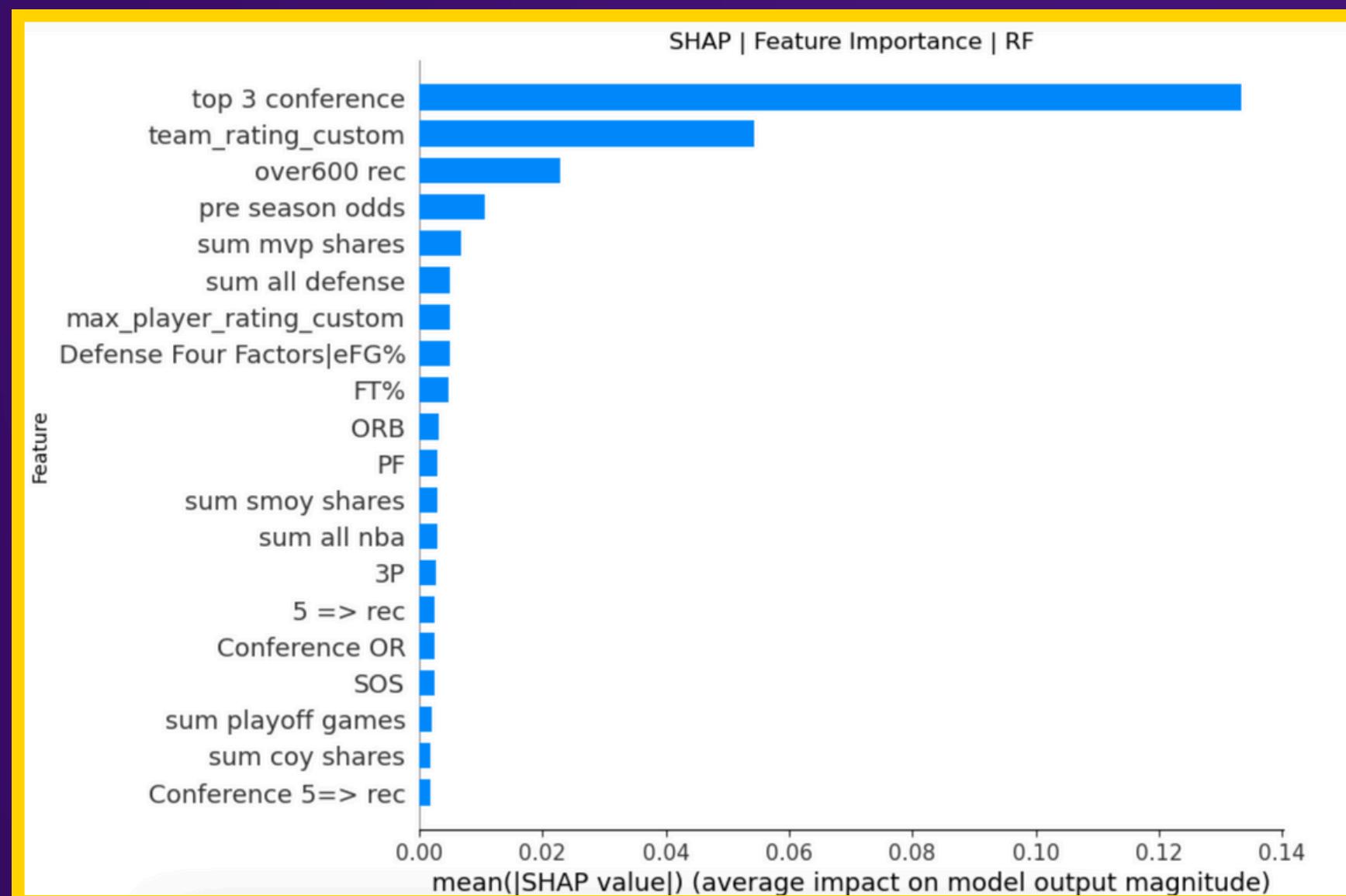
## Dominance In Team Efficiency

Another key factor in determining whether or not a team will succeed is their team efficiency. Some teams have players with better chemistry than others and that can determine the outcome of some games.

# Machine Learning

## Model

### Feature Importance



The model determined that the top 3 teams in the conference are the most crucial, as they are more likely to advance further in the playoffs.

### Feature Importance:

top 3 conference	0.706709
team_rating_custom	0.063694
pre season odds	0.043323
over600 rec	0.019165
sum dploy shares	0.018356
ORB	0.012328
max_player_rating_custom	0.012319
sum mvp shares	0.009400
FG%	0.007029
5 => rec	0.006747
sum smoy shares	0.006622
sum franchise L1Y cs	0.006225
Conference OR	0.006132

# Machine Learning Model (Accuracy)

## Error Metrics

### Error Metrics:

PT MAE: 0.1416206011187136

PT MSE: 0.04976642330582578

PT RMSE: 0.22308389297711698

01

### MAE (Mean Absolute Error)

The Mean Absolute Error measures the model's performance, indicating the average prediction error at 14.16%.

02

### MSE (Mean Squared Error)

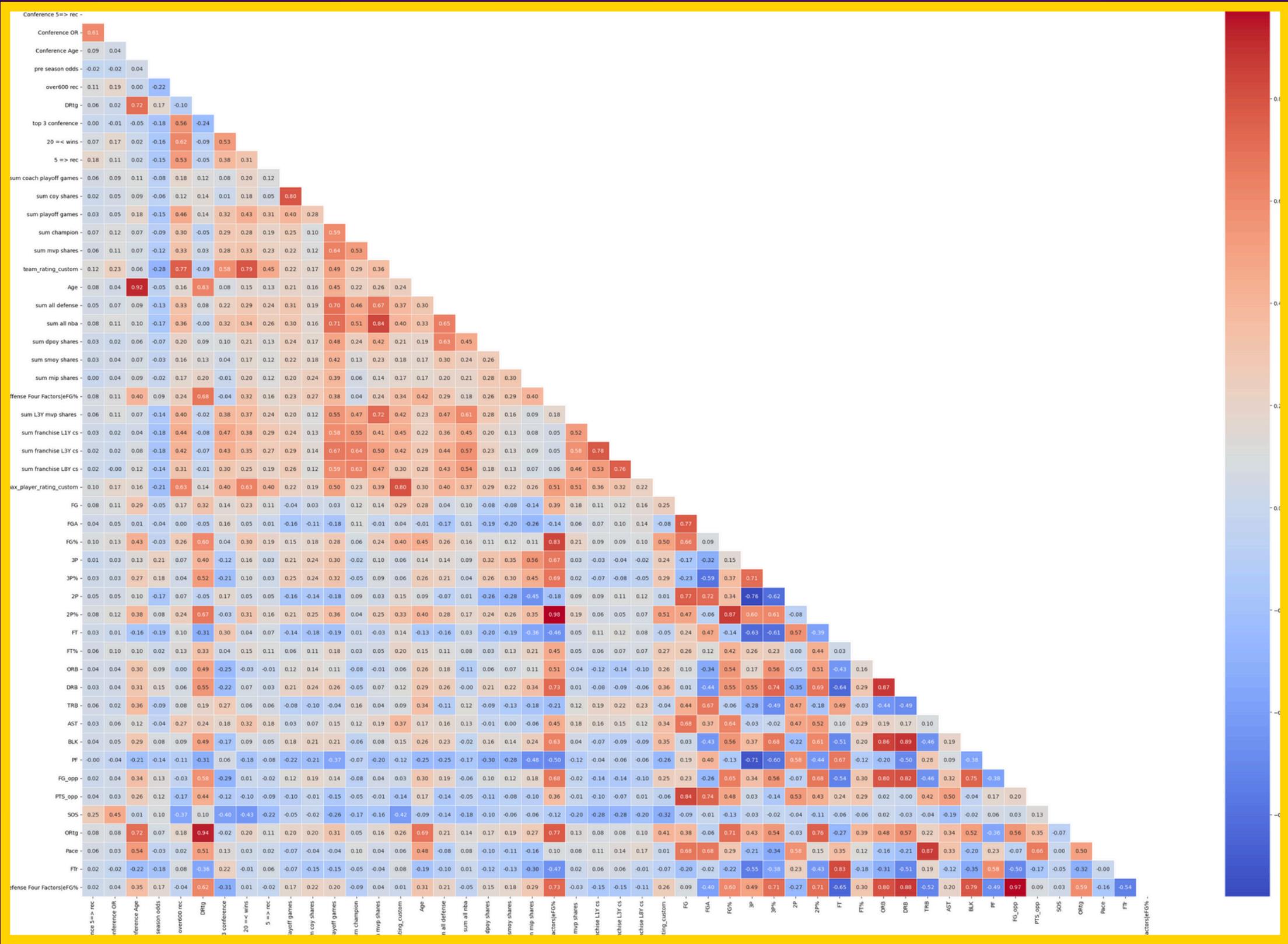
The Mean Squared Error value indicates fewer large errors. 4.98% shows that most errors are relatively small.

03

### RMSE (Root Mean Squared Error)

The Root Mean Squared Error indicates the typical magnitude of errors. 22.31% indicates relatively small errors.

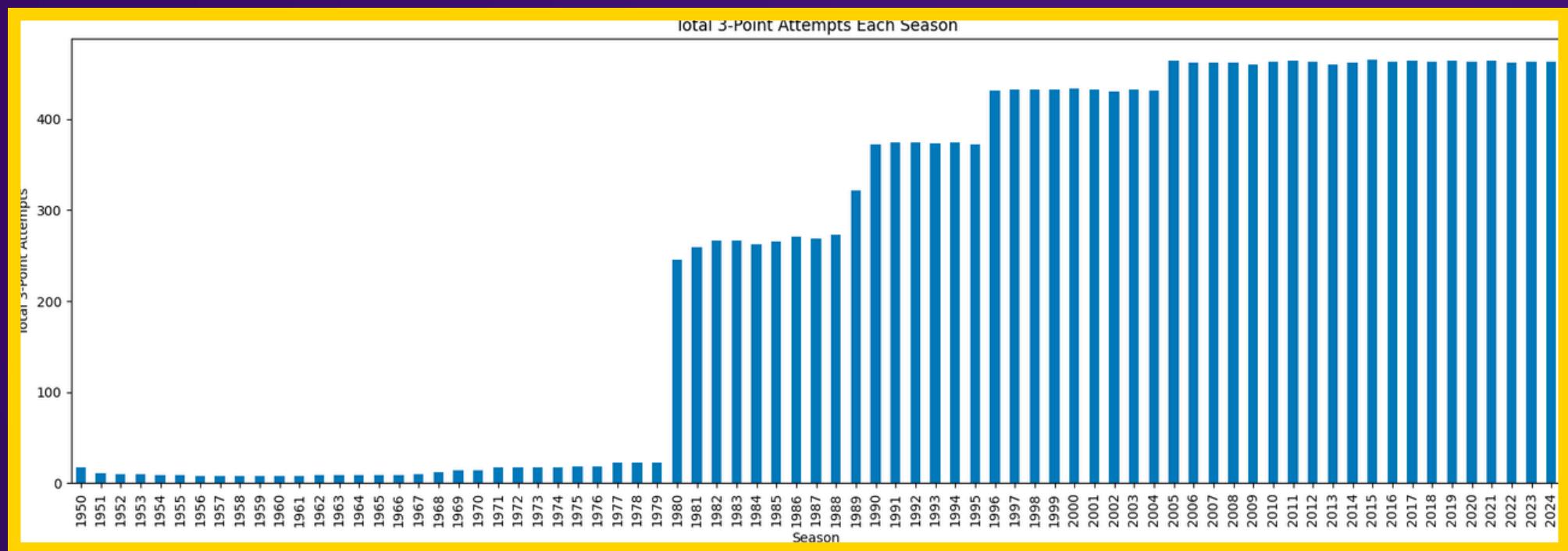
# Correlation Matrix



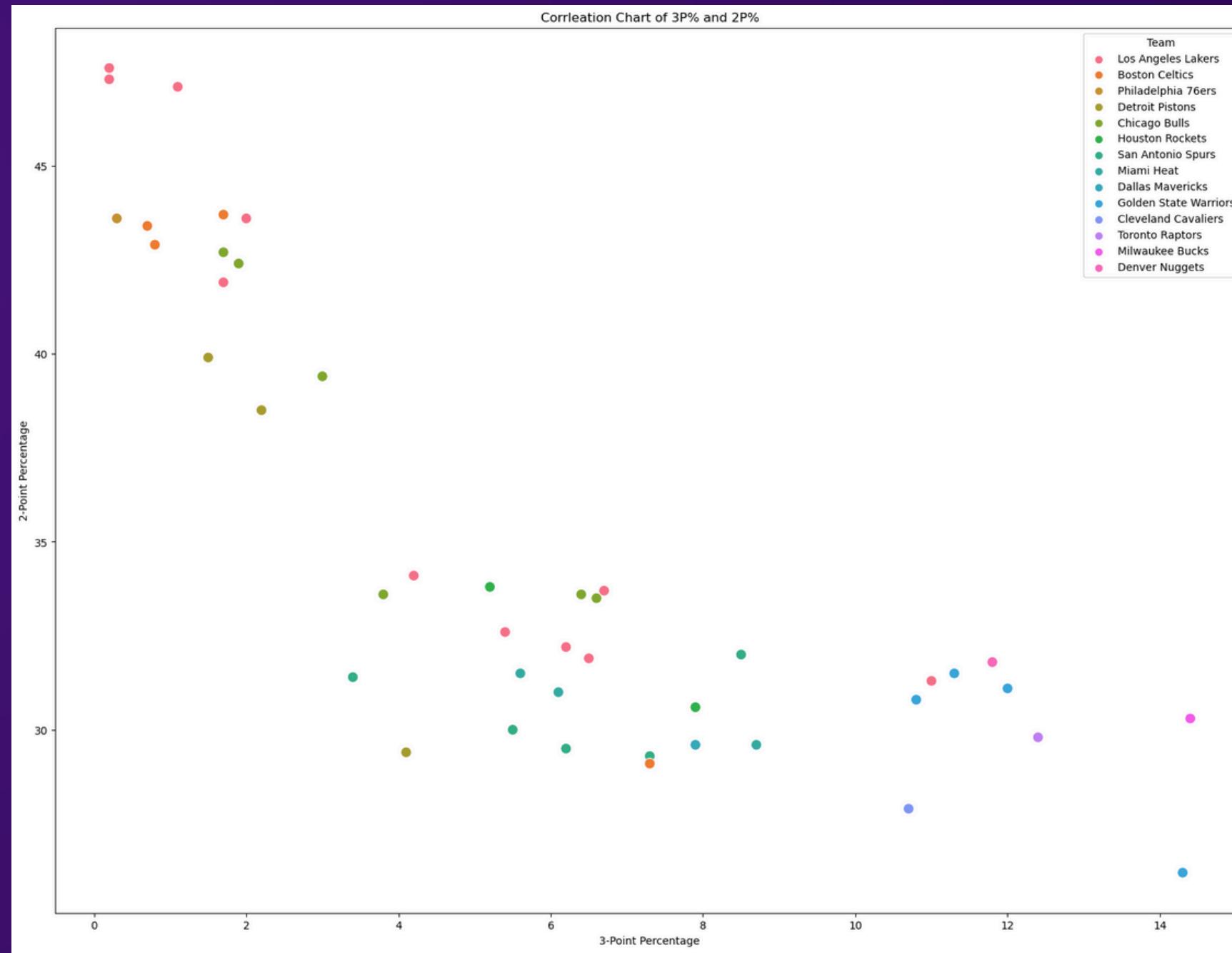
Positive Correlation	Negative Correlation
Defense Four Factors & 2P%	3P & PF
FGOpp & Defense Four Factors	3P & 2P
BLK & DRB	FT & Defense Four Factors
DRTG & ORTG	Offense Four Factors & PF

# Data Visualization

## 3-Point Attempts Each Season by Year

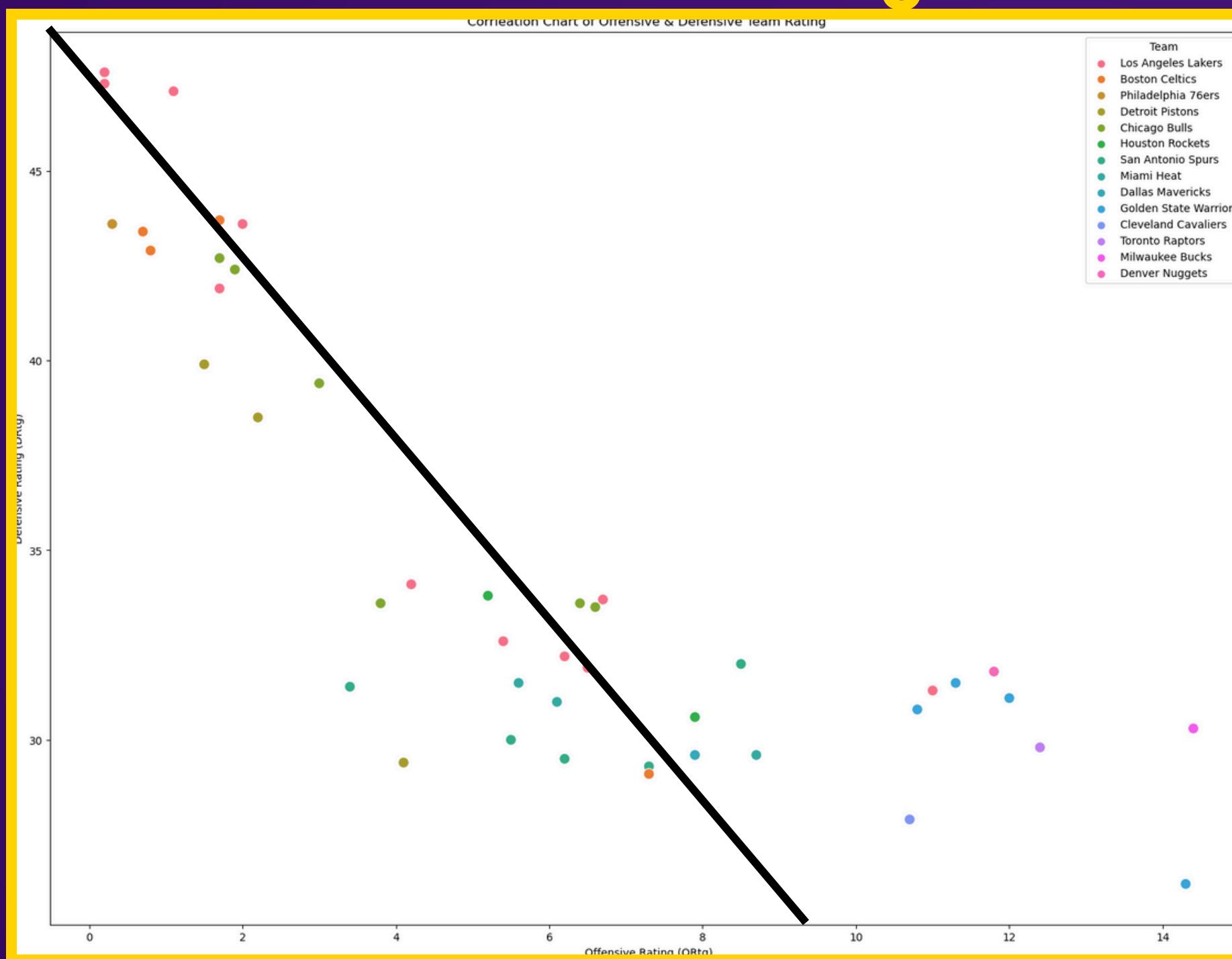


## Correlation Chart of 3P% and 2P%

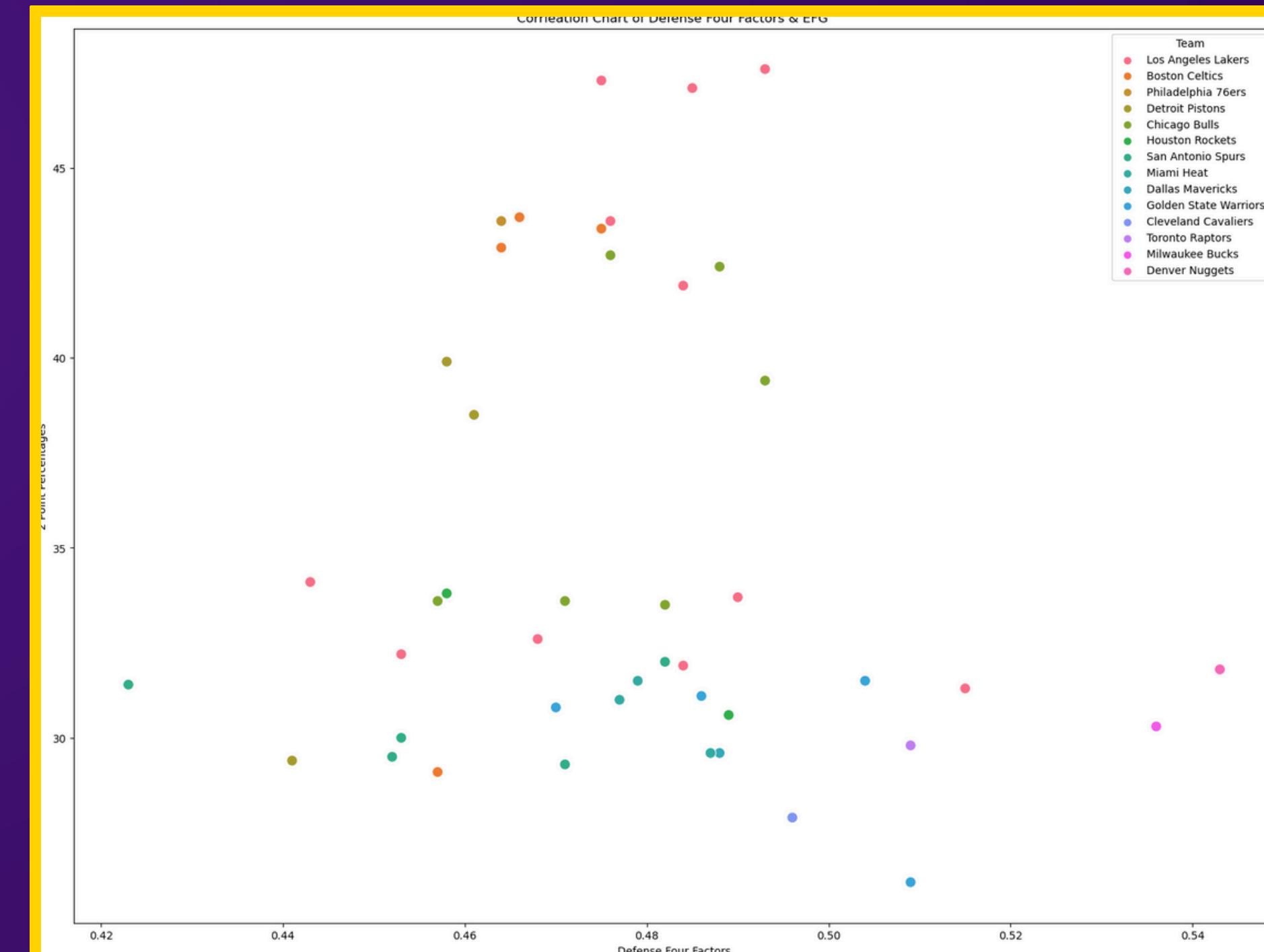


# Correlation Matrix

## Negative Correlation Chart of Offensive and Defense Team Ratings

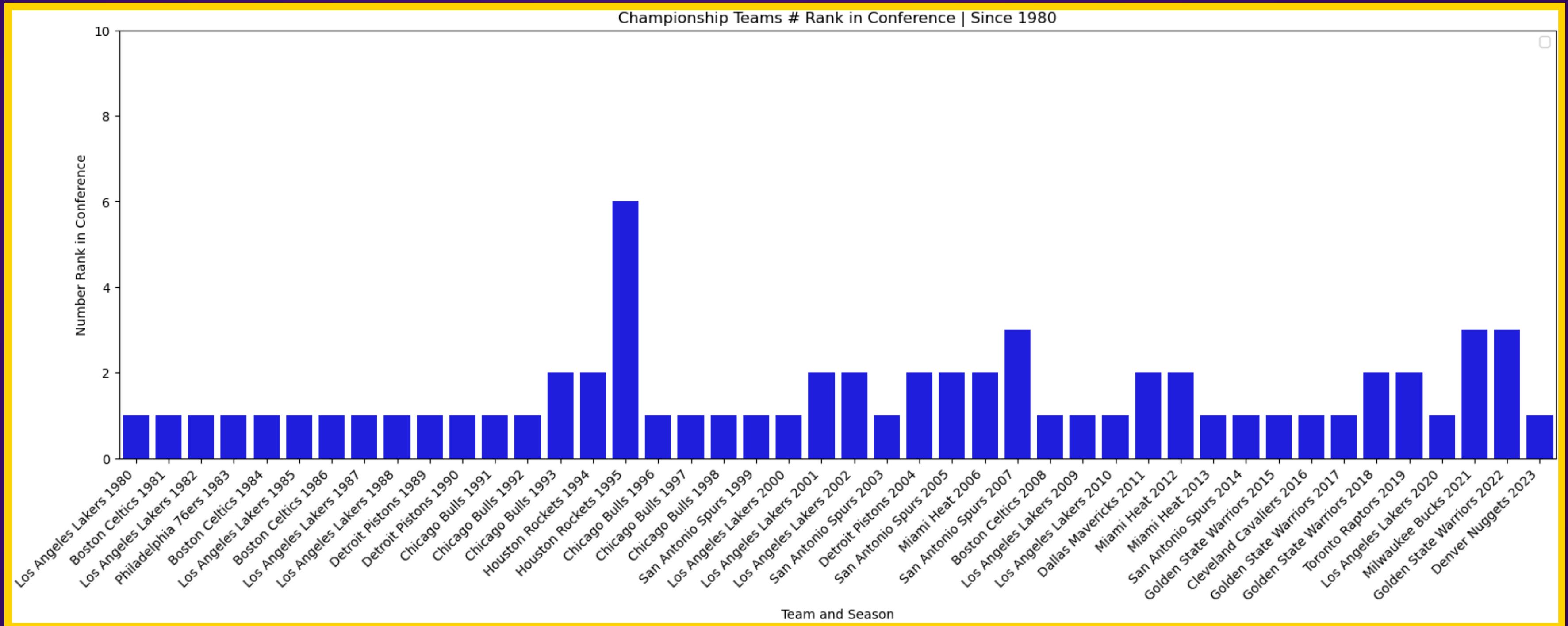


## \*Positive\* Correlation Chart of Defense Four Factors and 2 Pointers



# Good Regular Season Performance

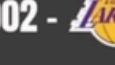
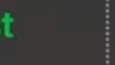
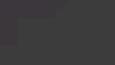
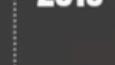
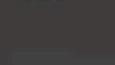
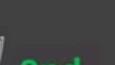
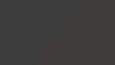
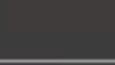
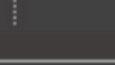
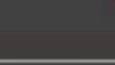
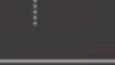
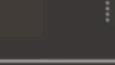
\*Top 3 Conference\*



Credit: [JA Swoosh](#) | [If JxmyHighRoller Doesn't Make This Video, Then I Will...](#)

# Good Regular Season Performance

## TOP 3 CONFERENCE

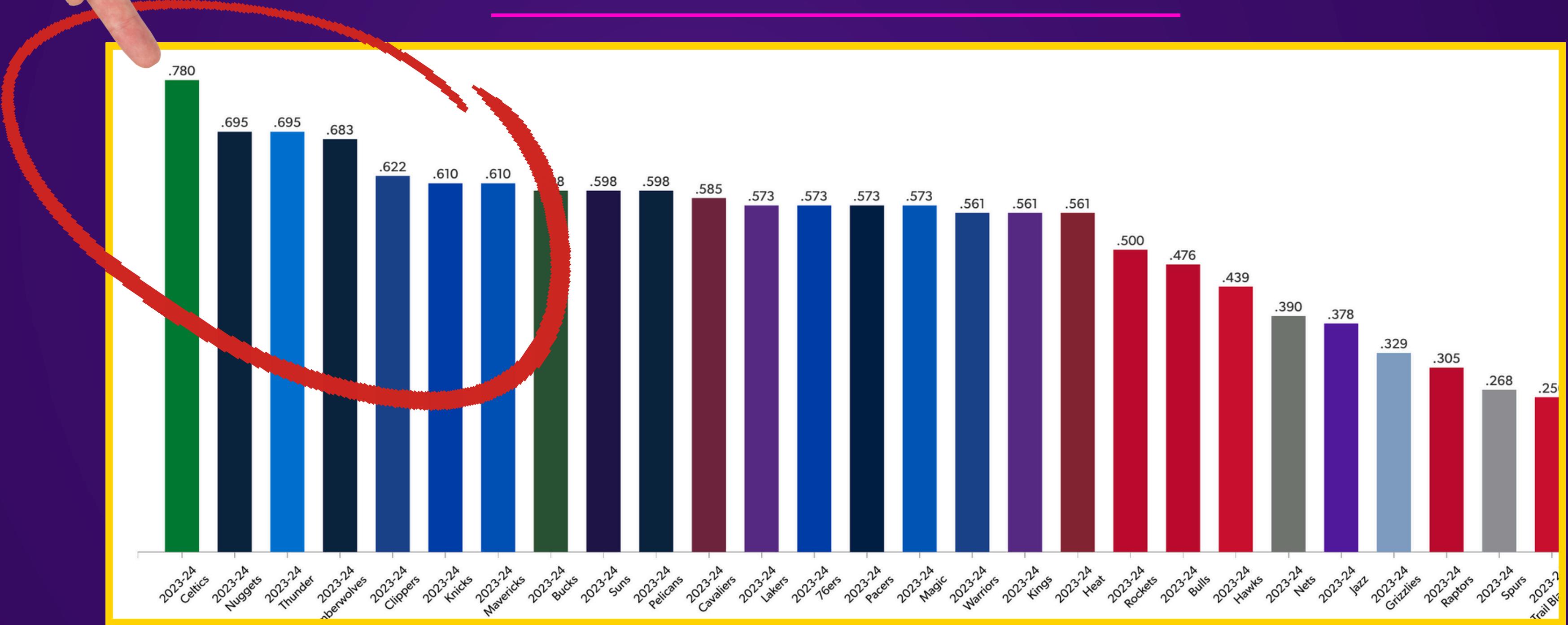
2023 -  1st	2021 -  3rd	2015 -  1st	2009 -  1st	2003 -  1st	1997 -  1st	1991 -  1st	1985 -  1st
2022 -  3rd	2020 -  1st	2014 -  1st	2008 -  1st	2002 -  2nd	1996 -  1st	1990 -  1st	1984 -  1st
2019 -  2nd	2013 -  1st	2007 -  3rd	2006 -  2nd	2001 -  2nd	1995 -  6th	1989 -  1st	1983 -  1st
2018 -  2nd	2012 -  2nd	2005 -  2nd	2000 -  1st	1994 -  2nd	1988 -  1st	1982 -  1st	
2017 -  1st	2011 -  2nd	1999 -  1st	1993 -  2nd	1987 -  1st	1981 -  1st		
2016 -  1st	2010 -  1st	2004 -  2nd	1998 -  1st	1992 -  1st	1986 -  1st	1980 -  1st	



JA

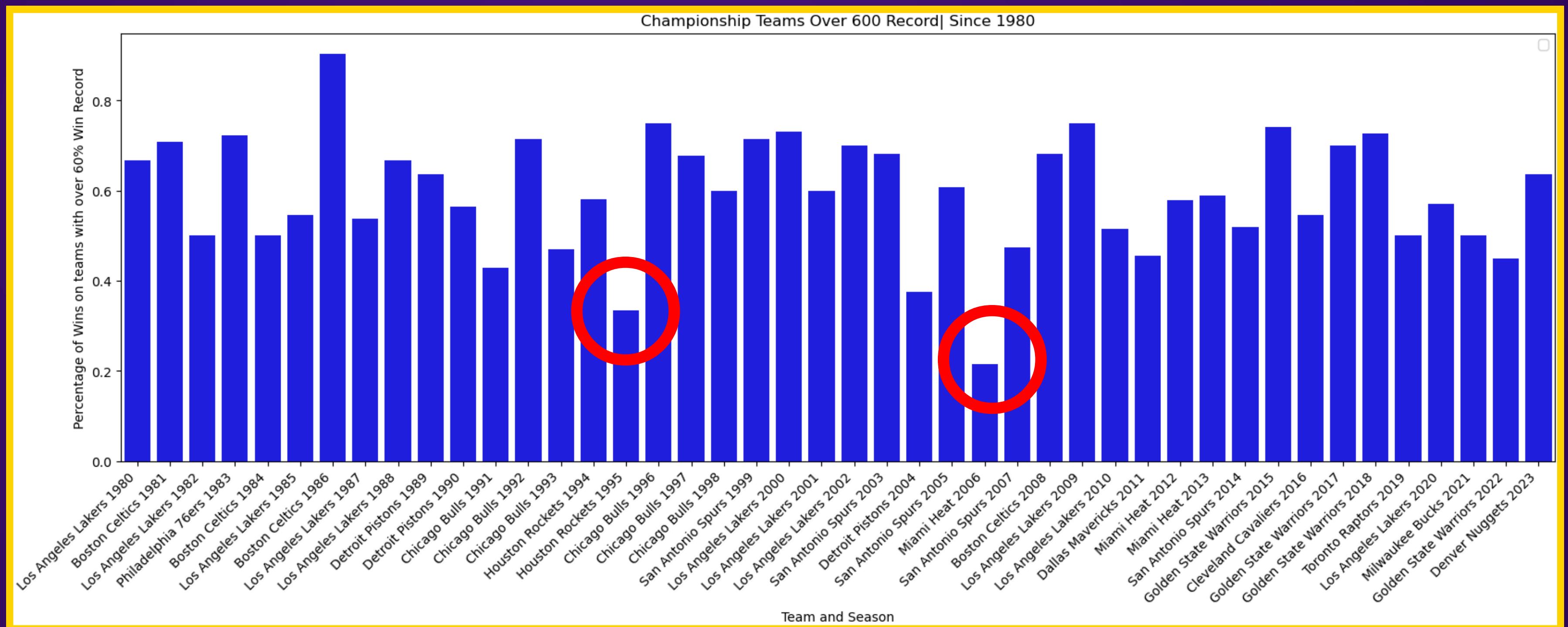
# Good Regular Season Performance

\*Over 600 Record\*

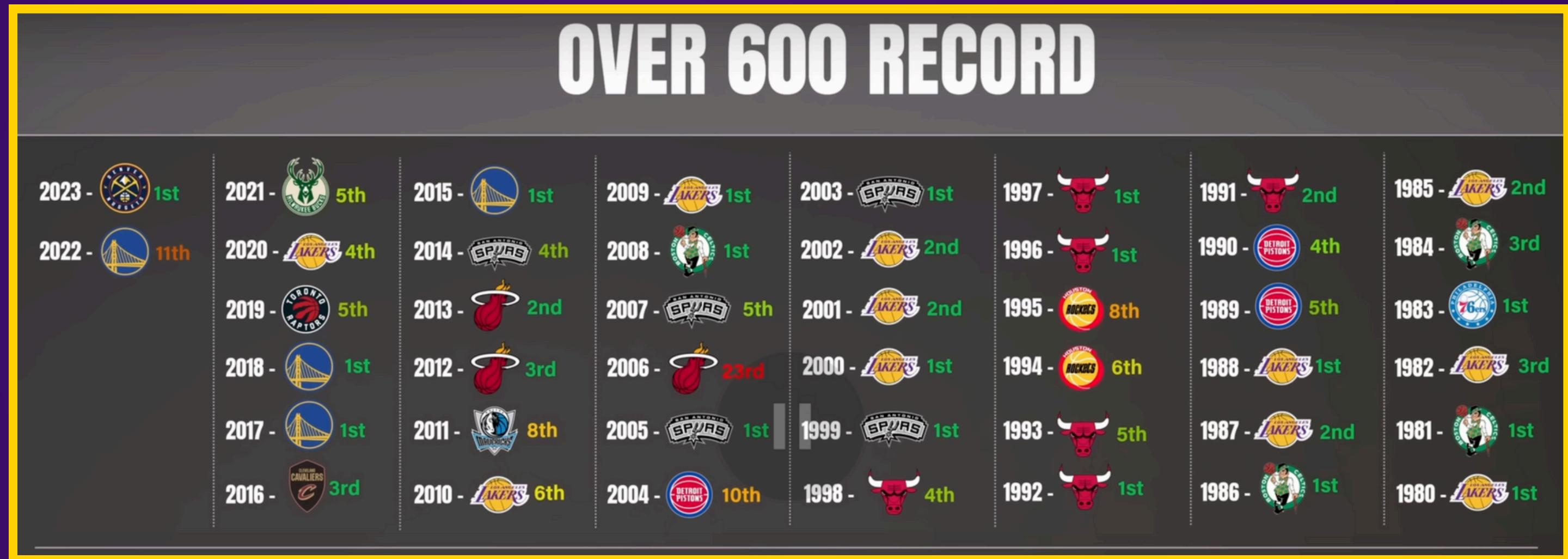


# Good Regular Season Performance

\*Over 600 Record\*

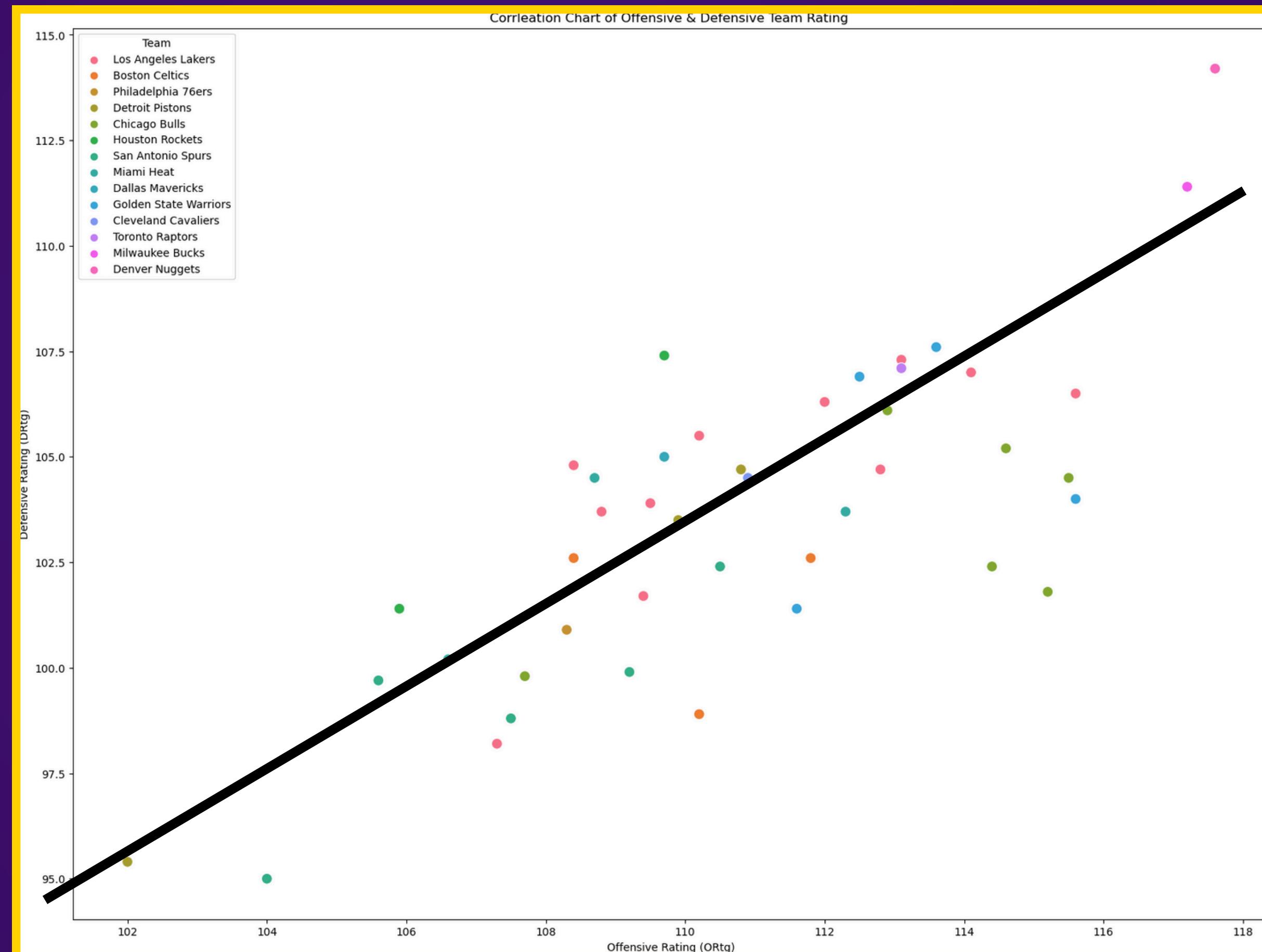


# Good Regular Season Performance



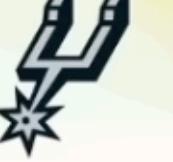
- **Nearly 50% of All NBA Champions Ranked #1 in Over 600 Record**
- **Roughly 84% of All NBA Champions Top 5 in Over 600 Record**

# Dominant in Team Efficiency



# Dominant in Team Efficiency

## Offensive + Defensive RTG

2023		Off - 5th Def - 15th	2017		Off - 1st Def - 2nd	2011		Off - 8th Def - 8th	2005		Off - 8th Def - 1st
2022		Off - 17th Def - 1st	2016		Off - 3rd Def - 10th	2010		Off - 11th Def - 4th	2004		Off - 18th Def - 2nd
2021		Off - 6th Def - 10th	2015		Off - 2nd Def - 1st	2009		Off - 3rd Def - 6th	2003		Off - 7th Def - 3rd
2020		Off - 11th Def - 3rd	2014		Off - 7th Def - 3rd	2008		Off - 10th Def - 1st	2002		Off - 2nd Def - 7th
2019		Off - 5th Def - 5th	2013		Off - 2nd Def - 9th	2007		Off - 5th Def - 2nd	2001		Off - 2nd Def - 21st
2018		Off - 3rd Def - 11th	2012		Off - 8th Def - 4th	2006		Off - 7th Def - 9th	2000		Off - 5th Def - 1st

# Measuring Team & Playoff Experience

The table below shows the data from the spreadsheet:

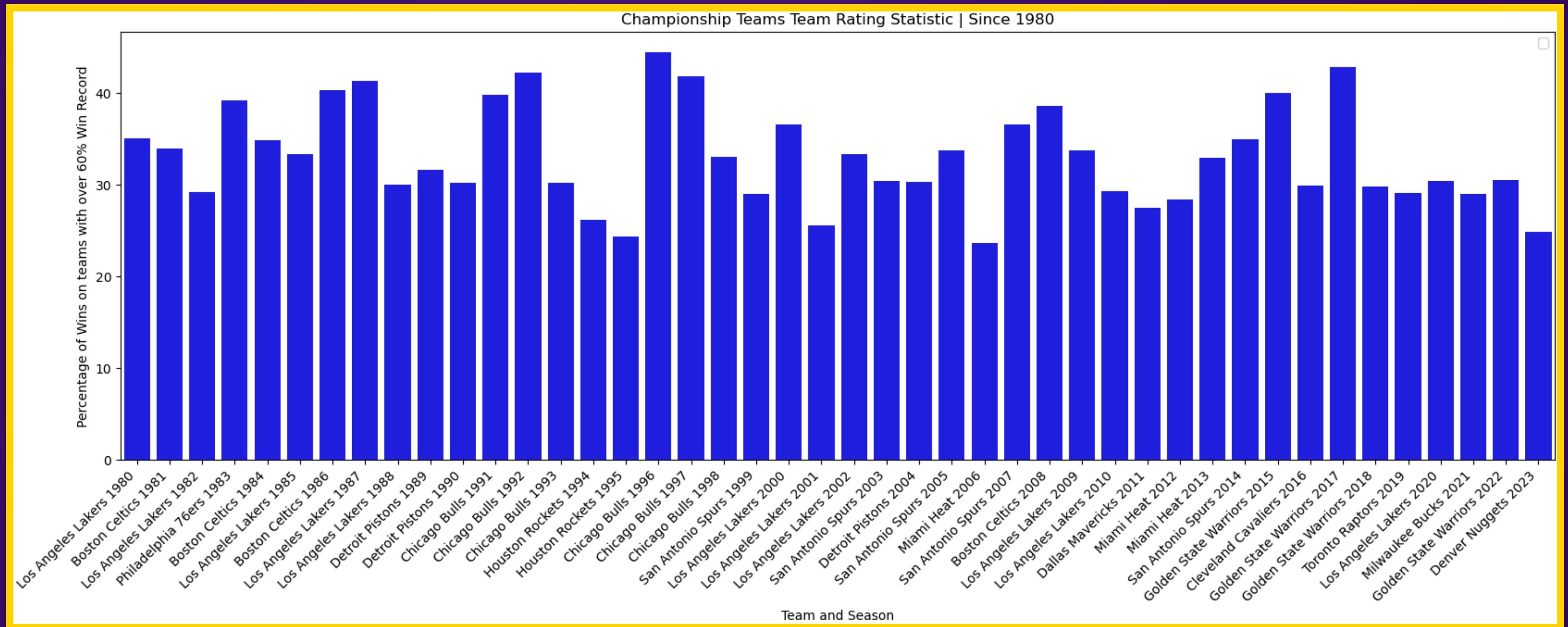
Season	Player	player_rating_custom	PTS	VORP	SRS	sum playoff games	sum mvp shares	sum dpoy shares	sum L5Y mvp shares	League Rank
2024	Kevin Durant	66.60055	27.6	3.8	2.2	166	3.21	0.014	0.026	12
2024	Devin Booker	59.26008	27.1	2.5	2.2	43	0.216	0	0.216	19
2024	Grayson Allen	34.587	13.7	1.8	2.2	24	0	0	0	73
2024	Bradley Beal	33.435	18.3	0.3	2.2	45	0	0	0	78
2024	Jusuf Nurkić	27.6975	11.4	1.2	2.2	16	0	0.002	0	111
2024	Eric Gordon	24.03	11.9	0.2	2.2	60	0	0	0	141
2024	Royce O'Neale	21.372	7.8	0.9	2.2	44	0	0	0	164
2024	Bol Bol	15.466	5.5	0.4	2.2	7	0	0	0	229
2024	Josh Okogie	14.2555	4.9	0.4	2.2	11	0	0	0	242
2024	Drew Eubanks	13.35	5.3	0.2	2.2	0	0	0	0	249
2024	David Roddy	10.528	7	-0.9	2.2	6	0	0	0	289
2024	Nassir Little	9.4265	3.6	-0.1	2.2	3	0	0	0	312
2024	Saben Lee	9.2	3.2	0	2.2	0	0	0	0	318
2024	Udoka Azubuike	8.5255	2.5	0	2.2	1	0	0	0	328
2024	Théo Maledon	7.15	3.5	-0.5	2.2	0	0	0	0	352
2024	Isaiah Thomas	4.4129	0	0	2.2	25	0.08	0	0	390

Player Rating

Weighting Breakdown:

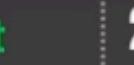
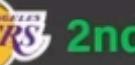
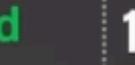
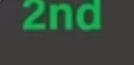
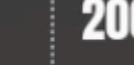
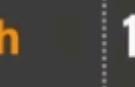
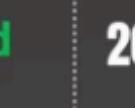
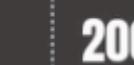
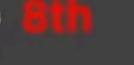
=SUM( (x) 1.5 + (x) 5 + (x) 2 + (x) .0005 + (x) .005 + (x) 1 + (x) 1 )

# Measuring Team & Playoff Experience



# Measuring Team & Playoff Experience

## TEAM RATING CUSTOM

2023 -  8th	2021 -  4th	2015 -  1st	2009 -  2nd	2003 -  3rd	1997 -  1st	1991 -  2nd	1985 -  3rd
2022 -  1st	2020 -  3rd	2014 -  1st	2008 -  1st	2002 -  2nd	1996 -  1st	1990 -  4th	1984 -  1st
2019 -  3rd	2013 -  2nd	2007 -  1st	2001 -  7th	1995 -  8th	1989 -  5th	1983 -  1st	
2018 -  3rd	2012 -  4th	2006 -  6th	2000 -  1st	1994 -  6th	1988 -  3rd	1982 -  4th	
2017 -  1st	2011 -  8th	2005 -  2nd	1999 -  1st	1993 -  5th	1987 -  1st	1981 -  3rd	
2016 -  4th	2010 -  4th	2004 -  3rd	1998 -  2nd	1992 -  1st	1986 -  1st	1980 -  2nd	

- Nearly every single NBA Champion Team has ranked in the Top 6 for TRC

5  
6

# Machine Learning Model 2023 Results

= ~83%

## NBA Finals 2018

Top 7 Predicted Teams for Season 2018:

	Team	predicted	champion share	champion share
182	Golden State Warriors	0.796031	1.0000	
180	Houston Rockets	0.642423	0.6875	
181	Toronto Raptors	0.491615	0.2500	
184	Philadelphia 76ers	0.475192	0.3125	
186	Portland Trail Blazers	0.448566	0.0000	
185	Cleveland Cavaliers	0.415445	0.7500	
183	Boston Celtics	0.369593	0.6875	

Top 7 Predicted Teams for Season 2019:

	Team	predicted	champion share	champion share
154	Houston Rockets	0.638474	0.3750	
151	Toronto Raptors	0.611118	1.0000	
152	Golden State Warriors	0.598830	0.8750	
150	Milwaukee Bucks	0.582524	0.6250	
155	Portland Trail Blazers	0.481418	0.5000	
153	Denver Nuggets	0.442969	0.4375	
156	Philadelphia 76ers	0.370685	0.4375	

## NBA Finals 2020

Top 7 Predicted Teams for Season 2020:

	Team	predicted	champion share	champion share
121	Toronto Raptors	0.641049	0.4375	
122	Los Angeles Lakers	0.632914	1.0000	
120	Milwaukee Bucks	0.587848	0.3125	
123	Los Angeles Clippers	0.554929	0.4375	
124	Boston Celtics	0.461929	0.6250	
125	Denver Nuggets	0.450912	0.5625	
130	Miami Heat	0.420076	0.8750	

## NBA Finals 2021

Top 7 Predicted Teams for Season 2021:

	Team	predicted	champion share	champion share
93	Brooklyn Nets	0.671379	0.4375	
91	Phoenix Suns	0.576985	0.8750	
90	Utah Jazz	0.523043	0.3750	
96	Milwaukee Bucks	0.520809	1.0000	
92	Philadelphia 76ers	0.487691	0.4375	
94	Los Angeles Clippers	0.477317	0.6250	
95	Denver Nuggets	0.467131	0.2500	

## NBA Finals 2022

Top 7 Predicted Teams for Season 2022:

	Team	predicted	champion share	champion share
62	Golden State Warriors	0.636340	1.0000	
66	Milwaukee Bucks	0.623722	0.4375	
60	Phoenix Suns	0.530458	0.4375	
67	Philadelphia 76ers	0.512417	0.3750	
65	Boston Celtics	0.477965	0.8750	
63	Miami Heat	0.458039	0.6875	
61	Memphis Grizzlies	0.407992	0.3750	

## NBA Finals 2023

Top 7 Predicted Teams for Season 2023:

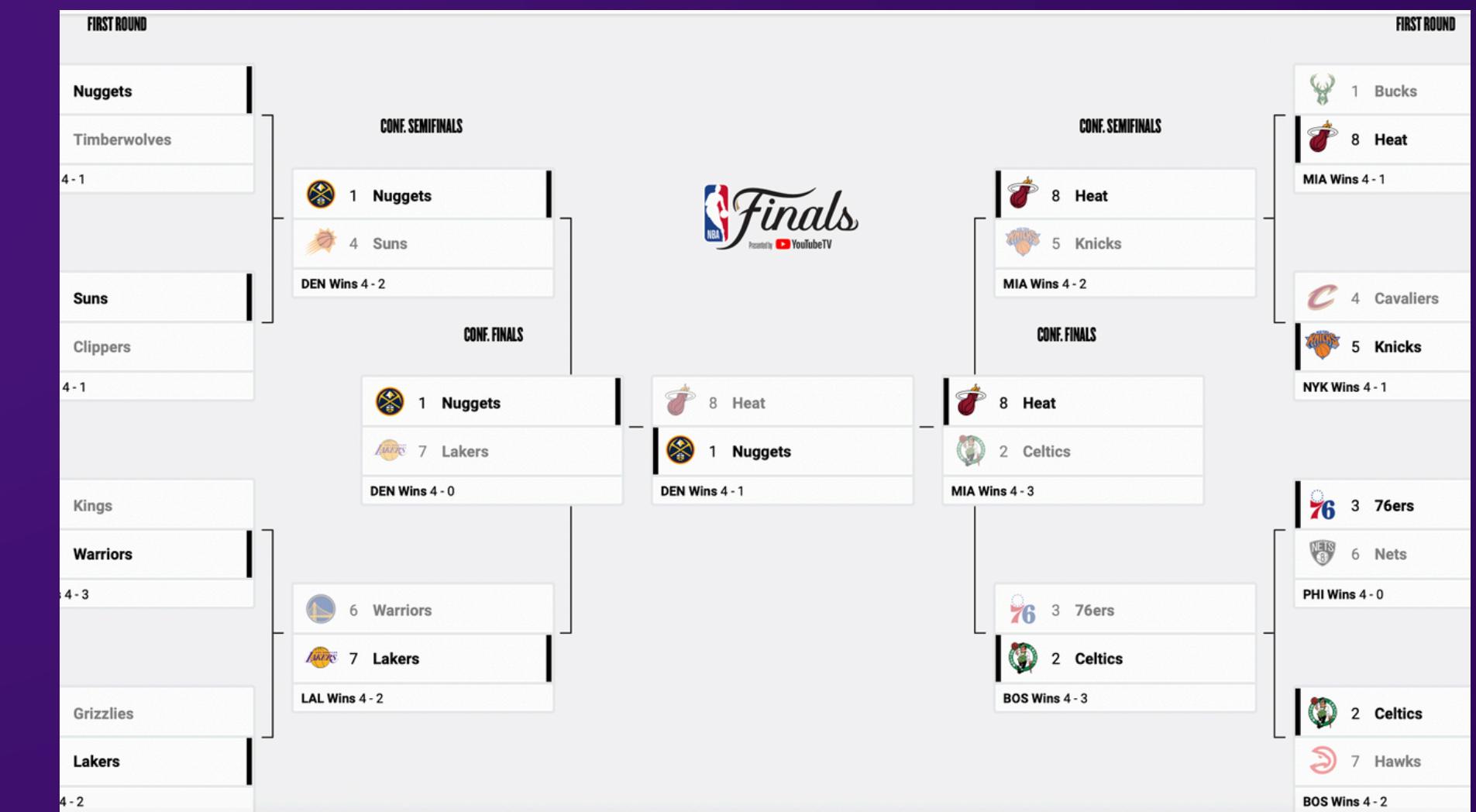
	Team	predicted	champion share	champion share
33	Denver Nuggets	0.648830	1.0000	
31	Boston Celtics	0.593126	0.6875	
30	Milwaukee Bucks	0.557548	0.0625	
32	Philadelphia 76ers	0.508318	0.4375	
34	Memphis Grizzlies	0.439924	0.1250	
36	Sacramento Kings	0.368554	0.1875	
37	New York Knicks	0.198096	0.3750	

Championship Share Score = Calculated Predicted Odds of Winning Finals

# 2023 NBA Finals Brackets based on Regression Model Prediction

VS

# \*REAL\* 2023 NBA Finals Brackets



# Machine Learning Model

## What Did The Model Predict

Full Results (Playoff Teams Only):

Season	Team	champion share	predicted champion share
1	Boston Celtics	0.0	0.659909
3	Milwaukee Bucks	0.0	0.631875
13	Denver Nuggets	0.0	0.611160
2	Oklahoma City Thunder	0.0	0.534656
17	Minnesota Timberwolves	0.0	0.488652
18	New York Knicks	0.0	0.357763
0	Indiana Pacers	0.0	0.257880
5	Los Angeles Lakers	0.0	0.250211
25	Miami Heat	0.0	0.232331
14	Philadelphia 76ers	0.0	0.208788
11	Los Angeles Clippers	0.0	0.192316
12	New Orleans Pelicans	0.0	0.179672
7	Golden State Warriors	0.0	0.159819
8	Sacramento Kings	0.0	0.147016
23	Orlando Magic	0.0	0.124653
19	Cleveland Cavaliers	0.0	0.115590

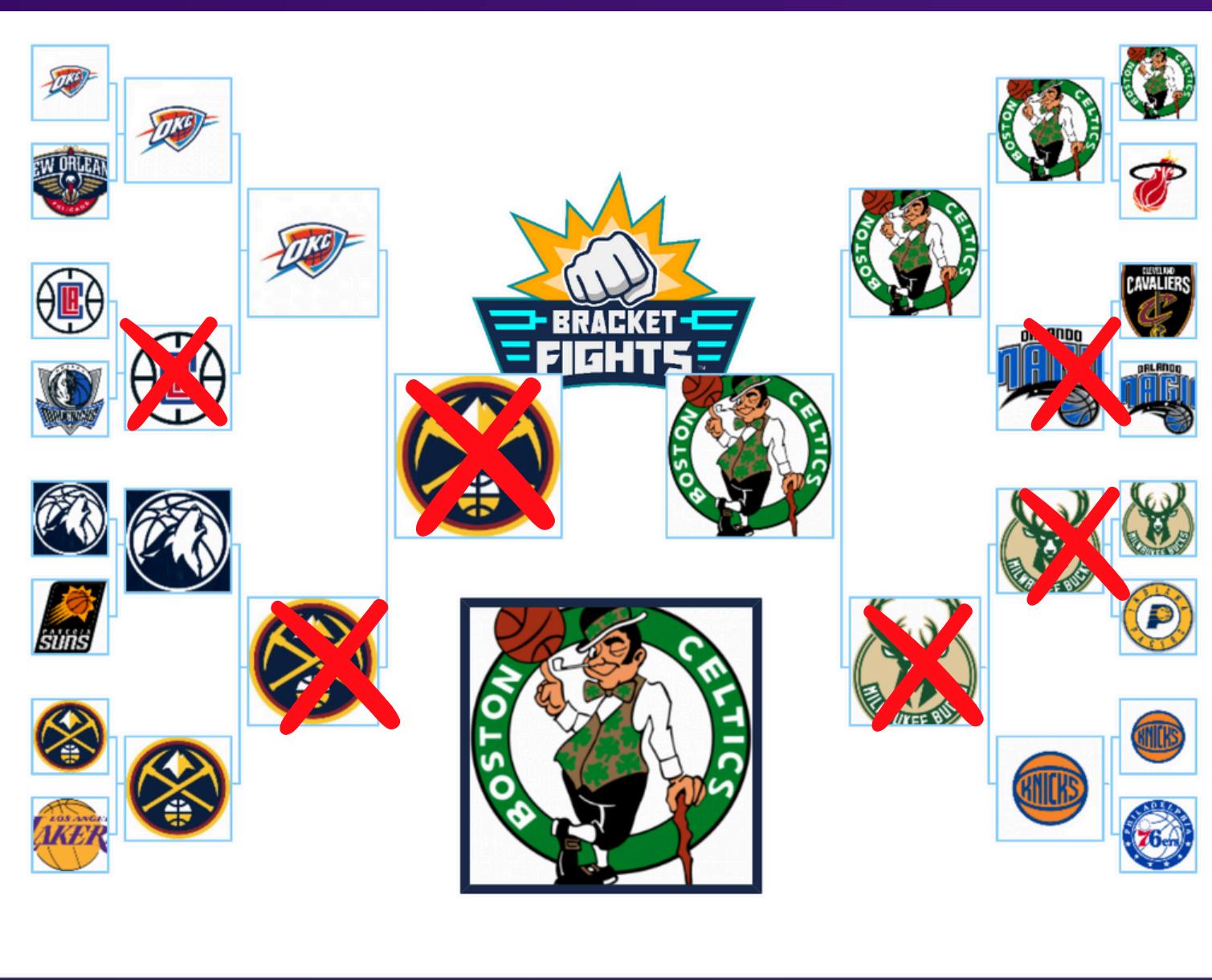
## Predictions

The model predicted that the Top 5 teams will be the Boston Celtics, Milwaukee Bucks, Denver Nuggets, Oklahoma City Thunder and the Minnesota Timberwolves

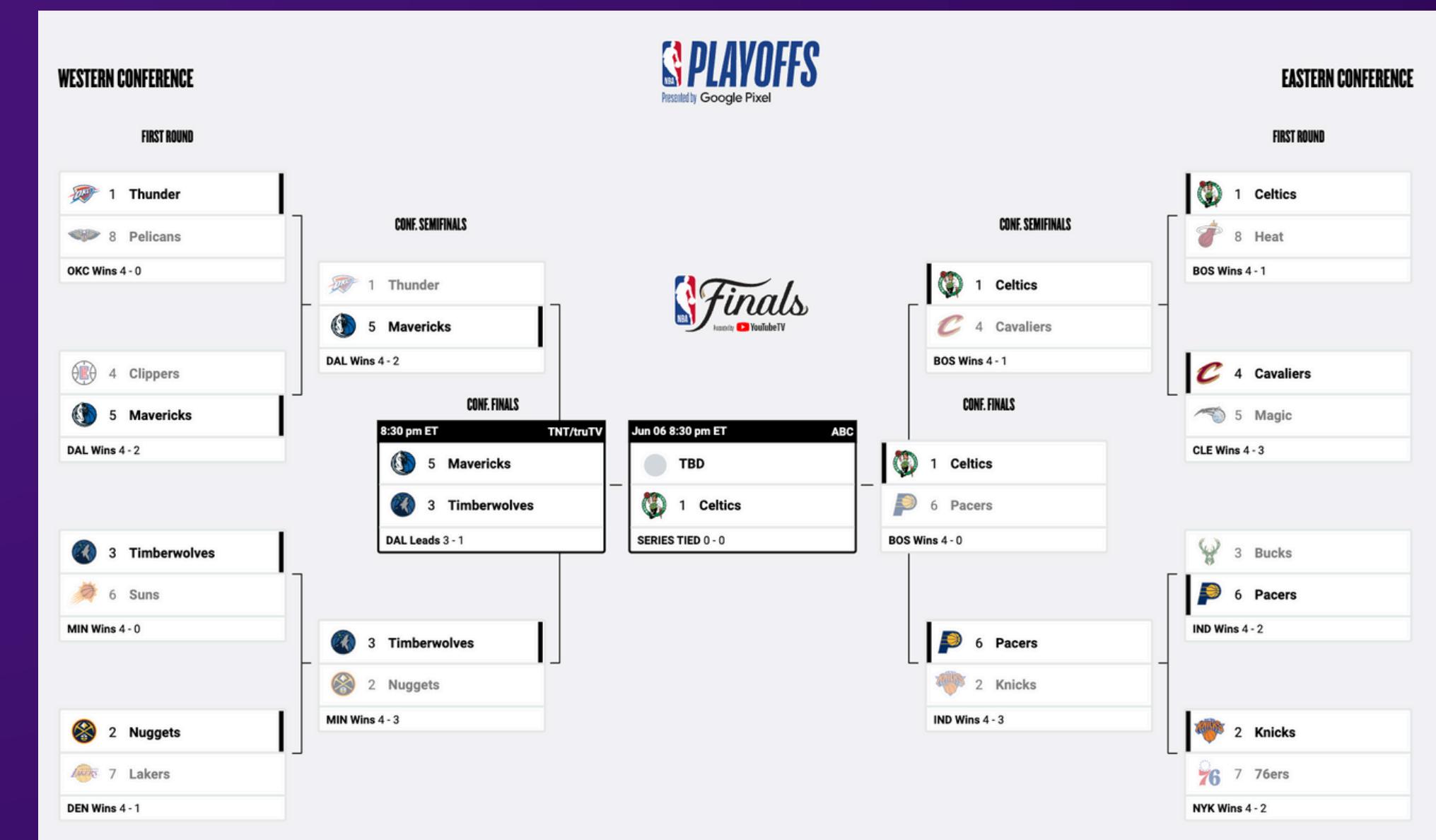
## What Does This Mean

- The model predicts that the Boston Celtics will win the NBA Finals.
- It was also predicted that the Milwaukee Bucks would finish high in the standings, but they are now out of the playoffs, so nothing is certain.

# 2024 NBA Finals Brackets based on Regression Model Prediction



# 2024 NBA Finals VS <sup>\*Current\*</sup> Brackets (As of May 30th 2024)



# Reflections on NBA Predictions

---

1

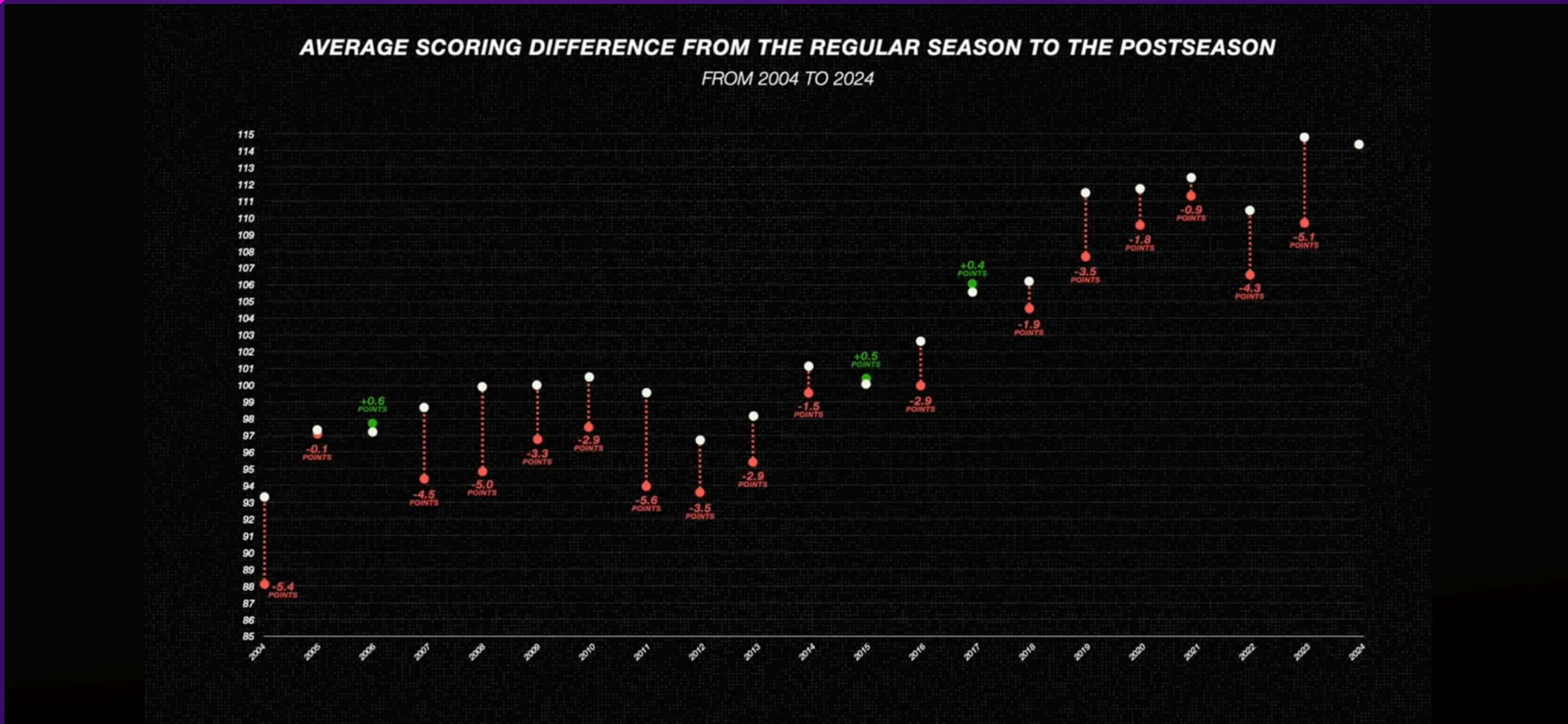
Player Injuries are **UNPREDICTABLE**



# Reflections on NBA Predictions

2

## Season games are **NOTHING** like Playoff Games

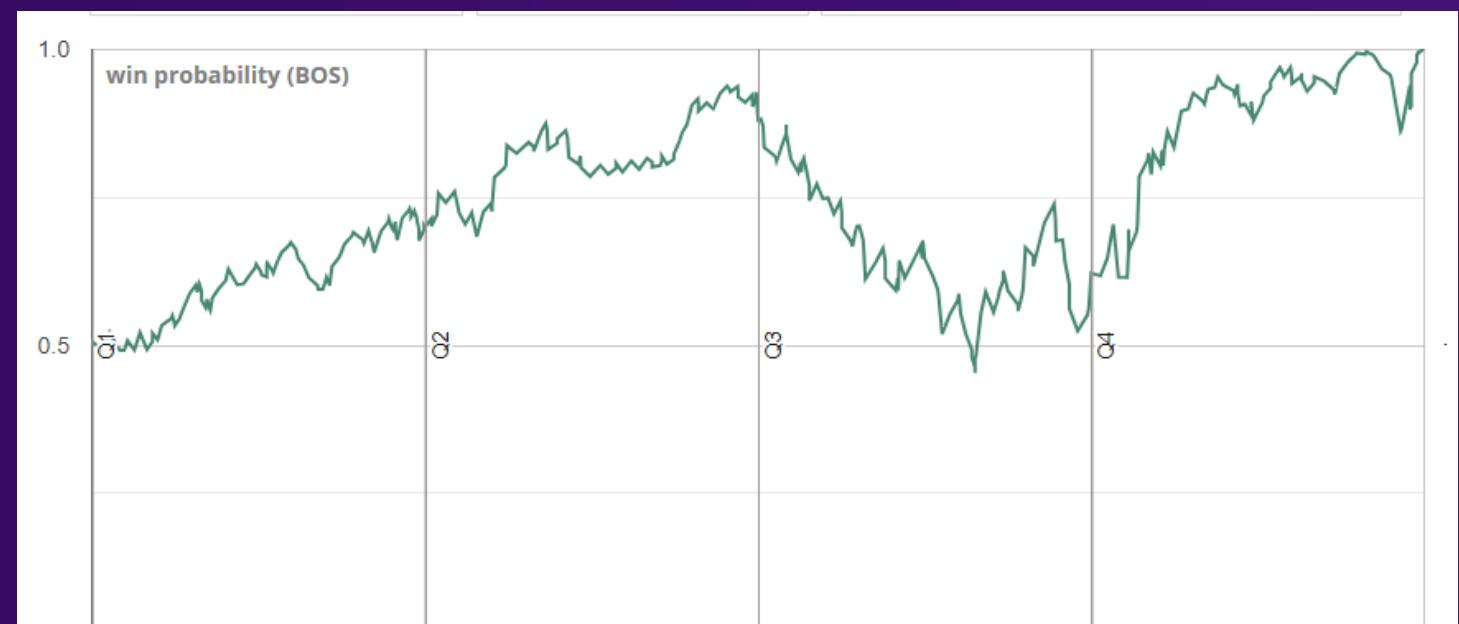


# Reflections on NBA Predictions

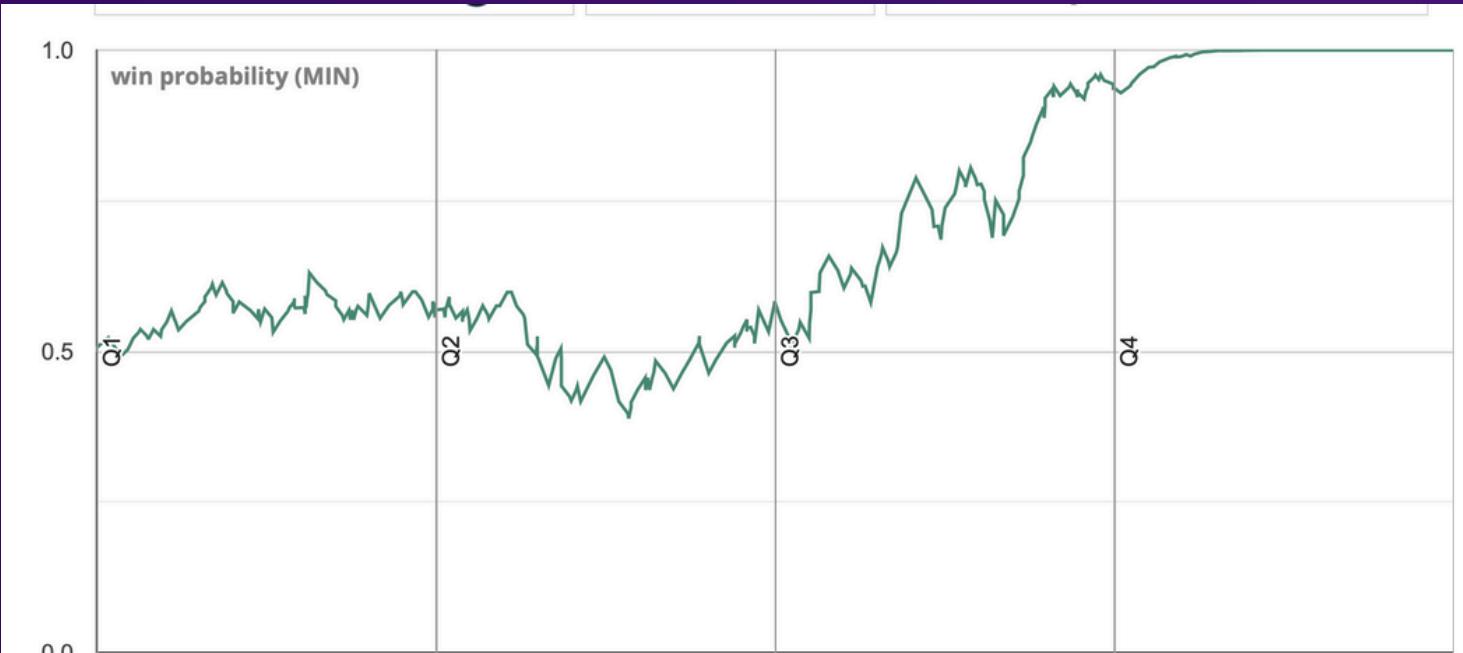
2

Season games are **NOTHING** like Playoff Games

Celtics vs. Pacers Last **SEASON** Game



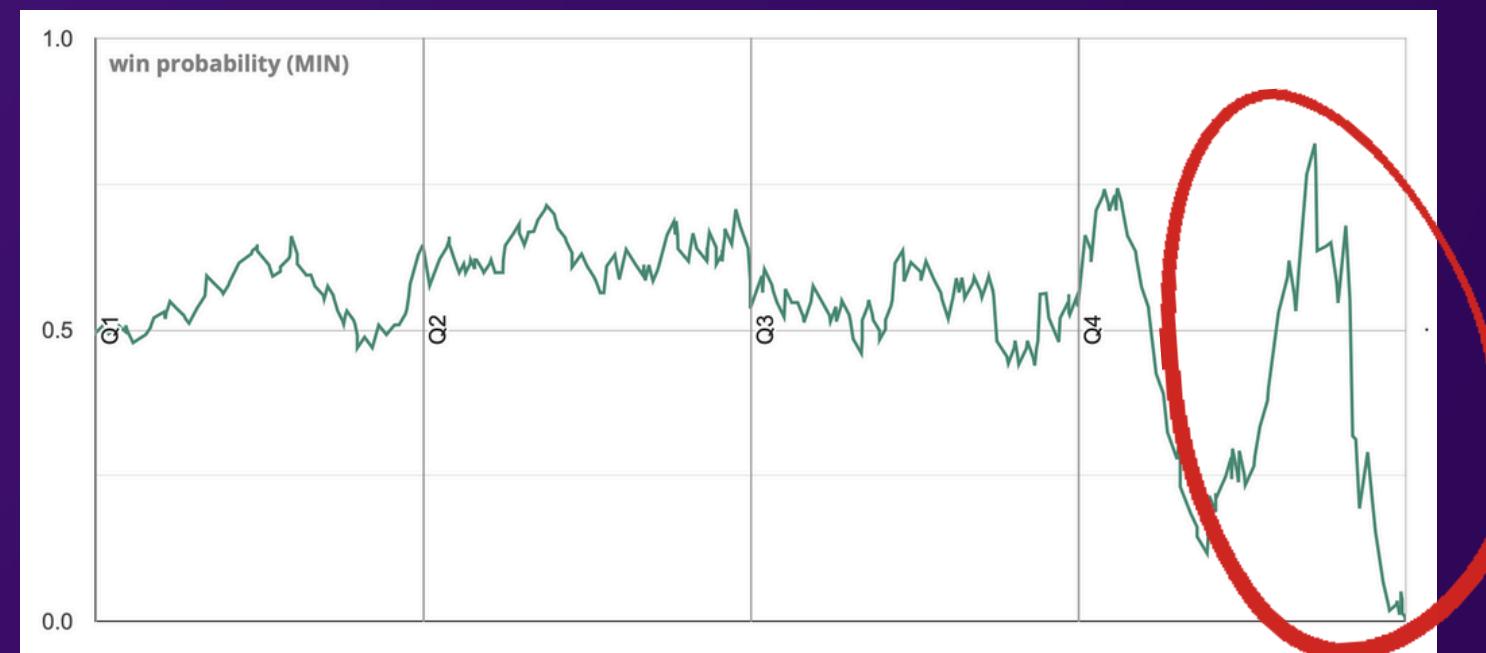
Timberwolves vs. Mavericks Last **SEASON** Game



Celtics vs. Pacers Last **FINALS** Game



Timberwolves vs. Mavericks Last **FINALS** Game



# Reflections on NBA Predictions

3

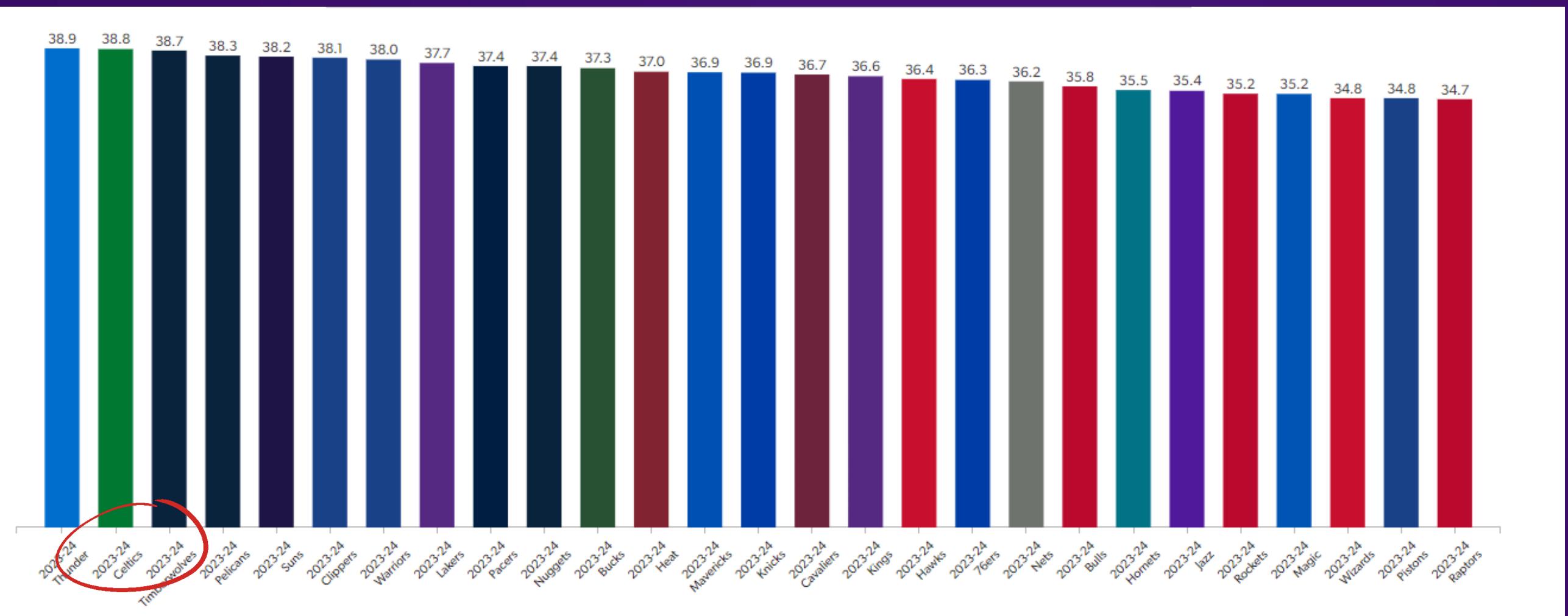
The Era of 3-Pointers has **REVOLUTIONIZED** the trajectory of NBA Games



# Reflections on NBA Predictions

3

The Era of 3-Pointers has **REVOLUTIONIZED** the trajectory of NBA Games



TEAM	3PA	3PM
Boston Celtics	42.5	16.5
Golden State Warriors	38.9	14.8
Dallas Mavericks	39.5	14.6
Sacramento Kings	39.3	14.4
Milwaukee Bucks	38.1	14.2
Atlanta Hawks	37.7	13.7
Cleveland Cavaliers	36.8	13.5
Oklahoma City Thunder	34.2	13.3

# Reflections on NBA Predictions

Where can we improve?

---

01

## Measuring the value of Role Players

Weighing in games where All Star caliber players are out. These are the opportunities where the great offensive player, the strong defensive player, the viable rebounder, or the quick and efficient passer can show off their impact on the game.

02

## Factoring In/Quantifying Coaching Experience

Last NBA Finals (2022–2023) the Miami Heat took an incredibly unforeseen lead to the Championship Finals credited most to the coaching of Erik Spoelstra. He offered immense support for the team through his past coaching experience during playoffs.

03

## Predicting Conference Rounds First

An algorithm that can predict the outcome of matchups before the finals game is likely going to show the better results as it breaks down the more minute pieces of the overall equation. It's best to start at the first round and then semi-finals conf., then finals. conf. That way the algorithm takes into account previous matchups within the season to make its predictions.

# Thank You!

**This Concludes Our Presentation**

---

Does anyone have  
questions?

---

**Special thank you to:**

**Data Science  
Instructor**



**Andre Sealy**

**Program  
Manager**



**Abrar Hasnat**

