Part 1: Background

Project Roadmap

The goal of our project is to create a simplified team strength model for the modern NBA.

The model will use team statistics (per-game averages) from an NBA season to assign each team a score that will represent the expected number of wins for that season.

In our project, we will do the following:

1. Conduct EDA to find stats that have a moderate to strong correlation with winning.
2. Weight our chosen stats properly to build parameters for the model.
3. Test the model on past seasons and adjust the parameters if necessary.
4. Run our model on the current NBA season to predict what the results will be.

			REGULAR SEASON				PLAYOFFS		
CURRENT RATING	TEAM	CONFERENCE	FULL-STRENGTH RATING	PROJ. RECORD	PROJ. POINT DIFF/G	CHANCE OF MAKING PLAYOFFS	FULL-STRENGTH RATING	CHANCE OF MAKING FINALS	CHANCE OF WINNING FINALS
1691	Celtics 17-4	East	1717	61-21	+7.3	>99%	1749	44%	30%
1582	Grizzlies 12-8	West	1666	53-29	+4.1	98%	1695	28%	13%
1620	Bucks 14-5	East	1658	53-29	+4.1	97%	1688	19%	11%
1573	Suns 14-6	West	1652	52-30	+4.5	97%	1678	22%	10%
1578	76ers 12-9	East	1639	50-32	+3.6	92%	1682	14%	8%
1583	Nuggets 13-7	West	1617	52-30	+3.3	96%	1654	16%	6%
1584	Warriors 11-10	West	1616	47-35	+2.3	87%	1654	13%	5%
1499	Heat 10-11	East	1612	46-36	+1.9	79%	1643	7%	3%
1588	Raptors 11-9	East	1616	48-34	+2.8	86%	1624	6%	3%
1560	Pelicans 12-8	West	1589	48-34	+3.0	86%	1606	7%	2%
1581	Mavericks 9-10	West	1586	45-37	+2.1	76%	1604	6%	2%
1534	Cavaliers 13-8	East	1581	48-34	+3.0	83%	1599	4%	1%
1432	Clippers 12-9	West	1568	42-40	-0.1	64%	1596	4%	1%
1556	Nets 11-11	East	1566	43-39	+1.0	62%	1598	3%	1%
1548	Hawks 11-10	East	1567	44-38	+0.9	67%	1580	2%	0.8%
1511	Timberwolves 10-11	West	1532	40-42	-0.6	45%	1539	1%	0.3%

FiveThirtyEight's Team Rating Model

Data Collection

- We chose to analyse statistics that are freely available on basketball-reference.com.
- We downloaded the relevant statistics into Excel, combined and cleaned the data there, and then imported each file into Python for further analysis.

```
In [4]: df_2021.head()
```

Out[4]:

	TEAM	PLAYOFFS	AGE	W	L	FG	FGA	FG RATE	3P	3PA	...	DEF EFG	DEF TOV	DEF DRB	DEF FT/FGA	DIST	LAYUP FREQ	CORNER FREQ	CORNER RATE	AST/TOV	Corner 3-pt Per Game
0	Atlanta Hawks	Made Playoffs	26.1	43	39	41.5	88.3	0.470	12.9	34.4	...	0.543	11.5	76.9	0.177	14.5	0.293	0.221	0.426	2.067227	3.238622
1	Boston Celtics	Made Final Four	26.1	51	31	40.7	87.4	0.466	13.2	37.1	...	0.502	12.5	77.3	0.183	14.8	0.334	0.231	0.389	1.823529	3.333769
2	Brooklyn Nets	Made Playoffs	29.1	44	38	42.0	88.4	0.475	11.5	31.7	...	0.521	11.7	75.1	0.201	13.9	0.304	0.235	0.393	1.794326	2.927653
3	Charlotte Hornets	Missed Playoffs	25.5	43	39	42.8	91.4	0.468	13.9	38.2	...	0.544	13.1	74.8	0.187	13.8	0.297	0.273	0.418	2.112782	4.359155
4	Chicago Bulls	Made Playoffs	26.3	46	36	41.7	86.9	0.480	10.6	28.8	...	0.541	11.9	78.3	0.199	13.9	0.364	0.287	0.372	1.867187	3.074803

Part 2:
Exploratory Data Analysis

EDA Process

Using the dataset for the 2021-22 NBA season (the most recent completed one), we took the following steps for each statistic:

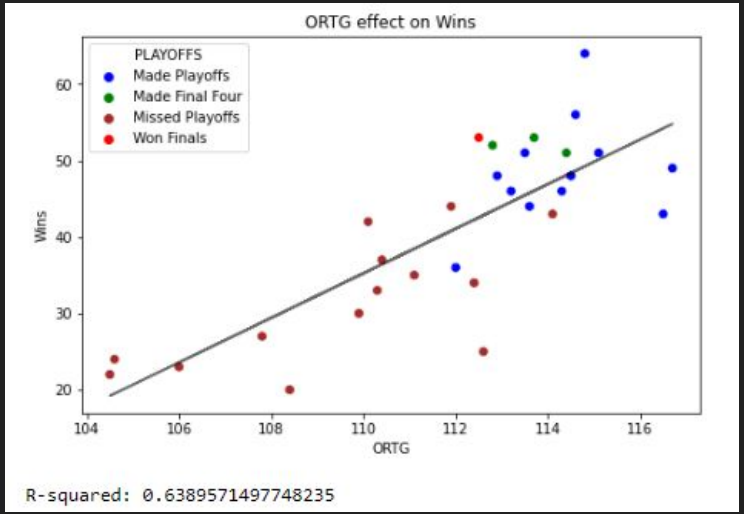
1. Created a scatterplot with the statistic on the x-axis and the team's wins on the y-axis
2. Fit a linear regression model to the data, and calculated the r^2 value.

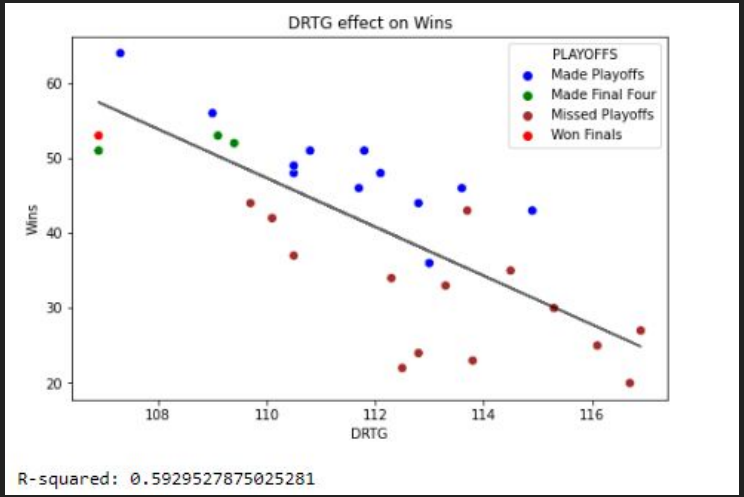
After observing the results, we then created some new statistics (that combined existing ones) to try to find a stronger correlation, and repeated the above steps.

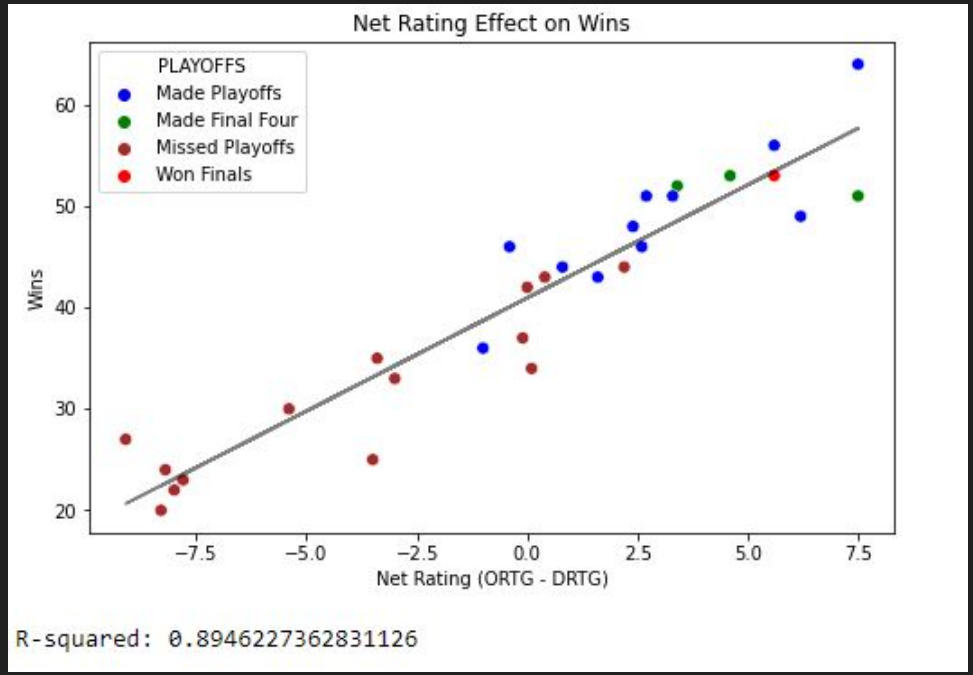
To speed up the process, we wrote a python script that did all of this given only the statistic and year as an input.

In [7]:
```
# Function to make scatterplot and generate regression line and r^2 value given dataframe and statistic column

def test_statistic(df, stat, figsize = (8, 5), y = "W", playoffs = "PLAYOFFS", s = 50,
                   title = 'Default', xlabel = 'Default', ylabel = 'Default'):

    if title == 'Default':
        title = f'{stat} effect on Wins'
    if xlabel == 'Default':
        xlabel = stat
    if ylabel == 'Default':
        ylabel = 'Wins'

    plt.figure(figsize = figsize)
    scatter = sns.scatterplot(x = stat, y = y, data = df, hue = playoffs, palette = colors, s = s)
    scatter.set(title = title, xlabel = xlabel, ylabel = ylabel)

    x = df[stat].values.reshape(-1, 1)
    y = df[y].values
    model = LinearRegression().fit(x, y)
    plt.plot(x, model.predict(x), color = 'black', alpha = 0.6)
    r_2 = r2_score(y, model.predict(x))

    r2_dict[stat] = r_2

    plt.show()

    print(f'R-squared: {r_2}')
```

Example: ORTG, DRTG, NRTG







Net Rating (NRTG) = ORTG - DRTG

	TEAM	PLAYOFFS	ORTG
0	Utah Jazz	Made Playoffs	116.7
1	Atlanta Hawks	Made Playoffs	116.5
2	Milwaukee Bucks	Made Playoffs	115.1
3	Phoenix Suns	Made Playoffs	114.8
4	Memphis Grizzlies	Made Playoffs	114.6
5	Denver Nuggets	Made Playoffs	114.5
6	Boston Celtics	Made Final Four	114.4
7	Minnesota Timberwolves	Made Playoffs	114.3
8	Charlotte Hornets	Missed Playoffs	114.1
9	Miami Heat	Made Final Four	113.7
10	Brooklyn Nets	Made Playoffs	113.6
11	Philadelphia 76ers	Made Playoffs	113.5
12	Chicago Bulls	Made Playoffs	113.2
13	Toronto Raptors	Made Playoffs	112.9
14	Dallas Mavericks	Made Final Four	112.8
15	Indiana Pacers	Missed Playoffs	112.6
16	Golden State Warriors	Won Finals	112.5
17	San Antonio Spurs	Missed Playoffs	112.4
18	New Orleans Pelicans	Made Playoffs	112.0
19	Cleveland Cavaliers	Missed Playoffs	111.9
20	Washington Wizards	Missed Playoffs	111.1
21	New York Knicks	Missed Playoffs	110.4
22	Los Angeles Lakers	Missed Playoffs	110.3
23	Los Angeles Clippers	Missed Playoffs	110.1
24	Sacramento Kings	Missed Playoffs	109.9
25	Houston Rockets	Missed Playoffs	108.4
26	Portland Trail Blazers	Missed Playoffs	107.8
27	Detroit Pistons	Missed Playoffs	106.0
28	Oklahoma City Thunder	Missed Playoffs	104.6
29	Orlando Magic	Missed Playoffs	104.5

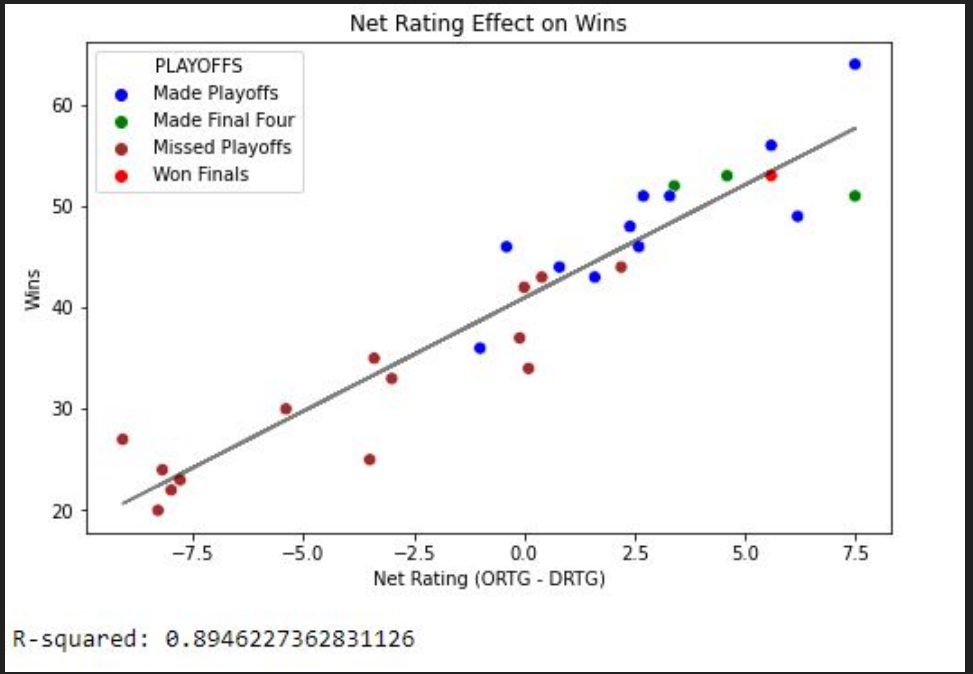
	TEAM	PLAYOFFS	DRTG
0	Boston Celtics	Made Final Four	106.9
1	Golden State Warriors	Won Finals	106.9
2	Phoenix Suns	Made Playoffs	107.3
3	Memphis Grizzlies	Made Playoffs	109.0
4	Miami Heat	Made Final Four	109.1
5	Dallas Mavericks	Made Final Four	109.4
6	Cleveland Cavaliers	Missed Playoffs	109.7
7	Los Angeles Clippers	Missed Playoffs	110.1
8	Toronto Raptors	Made Playoffs	110.5
9	Utah Jazz	Made Playoffs	110.5
10	New York Knicks	Missed Playoffs	110.5
11	Philadelphia 76ers	Made Playoffs	110.8
12	Minnesota Timberwolves	Made Playoffs	111.7
13	Milwaukee Bucks	Made Playoffs	111.8
14	Denver Nuggets	Made Playoffs	112.1
15	San Antonio Spurs	Missed Playoffs	112.3
16	Orlando Magic	Missed Playoffs	112.5
17	Brooklyn Nets	Made Playoffs	112.8
18	Oklahoma City Thunder	Missed Playoffs	112.8
19	New Orleans Pelicans	Made Playoffs	113.0
20	Los Angeles Lakers	Missed Playoffs	113.3
21	Chicago Bulls	Made Playoffs	113.6
22	Charlotte Hornets	Missed Playoffs	113.7
23	Detroit Pistons	Missed Playoffs	113.8
24	Washington Wizards	Missed Playoffs	114.5
25	Atlanta Hawks	Made Playoffs	114.9
26	Sacramento Kings	Missed Playoffs	115.3
27	Indiana Pacers	Missed Playoffs	116.1
28	Houston Rockets	Missed Playoffs	116.7
29	Portland Trail Blazers	Missed Playoffs	116.9

	TEAM	PLAYOFFS	Net Rating
0	Phoenix Suns	Made Playoffs	7.5
1	Boston Celtics	Made Final Four	7.5
2	Utah Jazz	Made Playoffs	6.2
3	Memphis Grizzlies	Made Playoffs	5.6
4	Golden State Warriors	Won Finals	5.6
5	Miami Heat	Made Final Four	4.6
6	Dallas Mavericks	Made Final Four	3.4
7	Milwaukee Bucks	Made Playoffs	3.3
8	Philadelphia 76ers	Made Playoffs	2.7
9	Minnesota Timberwolves	Made Playoffs	2.6
10	Toronto Raptors	Made Playoffs	2.4
11	Denver Nuggets	Made Playoffs	2.4
12	Cleveland Cavaliers	Missed Playoffs	2.2
13	Atlanta Hawks	Made Playoffs	1.6
14	Brooklyn Nets	Made Playoffs	0.8
15	Charlotte Hornets	Missed Playoffs	0.4
16	San Antonio Spurs	Missed Playoffs	0.1
17	Los Angeles Clippers	Missed Playoffs	0.0
18	New York Knicks	Missed Playoffs	-0.1
19	Chicago Bulls	Made Playoffs	-0.4
20	New Orleans Pelicans	Made Playoffs	-1.0
21	Los Angeles Lakers	Missed Playoffs	-3.0
22	Washington Wizards	Missed Playoffs	-3.4
23	Indiana Pacers	Missed Playoffs	-3.5
24	Sacramento Kings	Missed Playoffs	-5.4
25	Detroit Pistons	Missed Playoffs	-7.8
26	Orlando Magic	Missed Playoffs	-8.0
27	Oklahoma City Thunder	Missed Playoffs	-8.2
28	Houston Rockets	Missed Playoffs	-8.3
29	Portland Trail Blazers	Missed Playoffs	-9.1

Part 3:
Building the Model

Stat 1: Net Rating

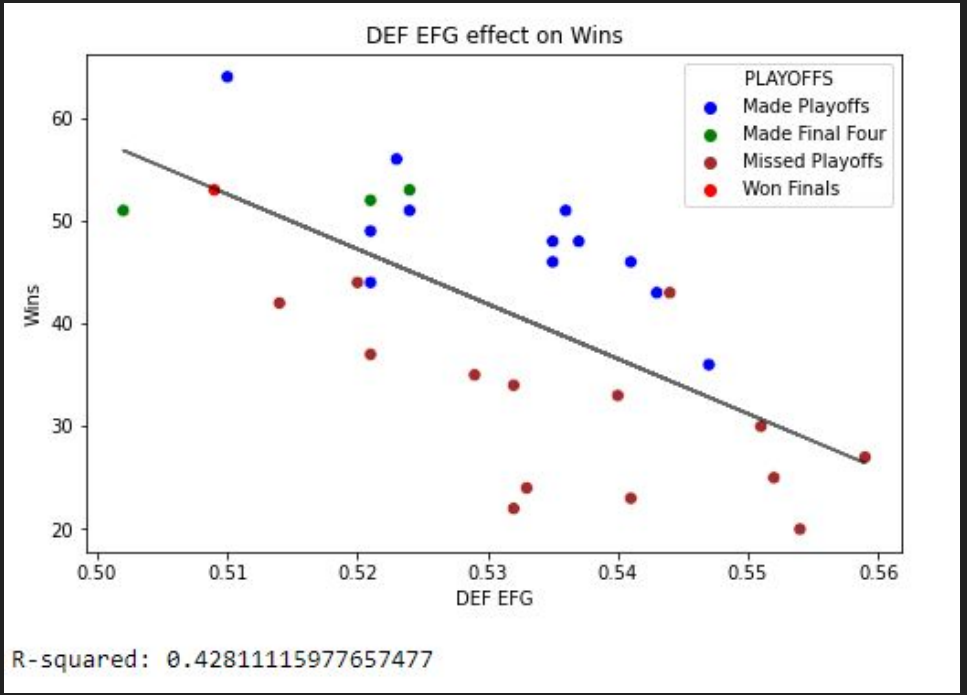
Model Weight: 70%



Stat 2: Defensive EFG%

Model Weight: 13%

Defensive EFG% (effective field goal percentage) is a weighted average of the opposing team's shooting percentage, based on the difficulty of the shot taken.

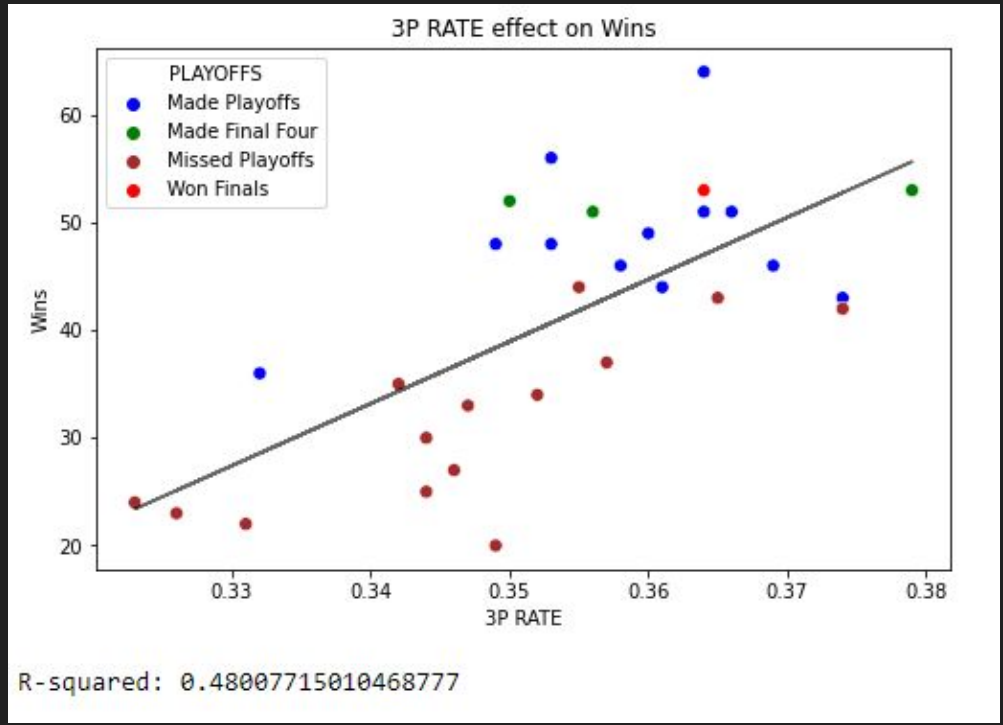


Stat 3: 3-point Percentage

Model Weight: 11%

3-point percentage is simply the percentage of 3-point shots that a team makes.

While a great 3-point shooter can exceed 40%, it is rare for an entire team to shoot that well.

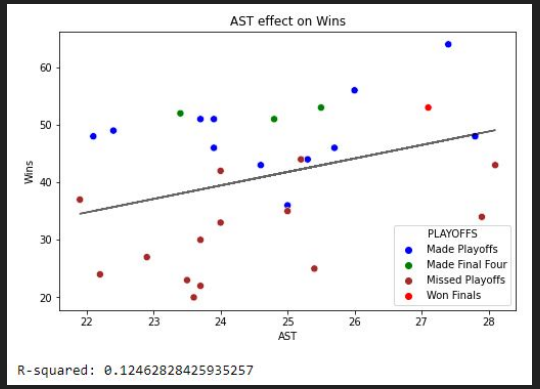


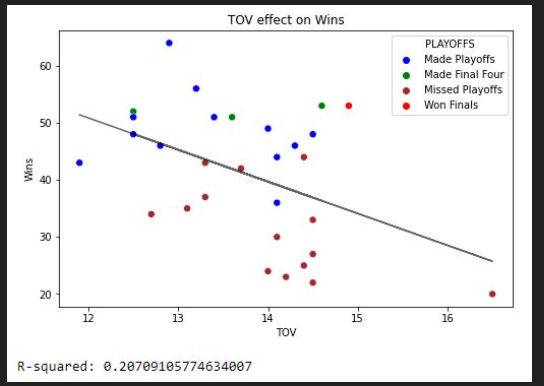
Stat 4: AST/TOV Ratio

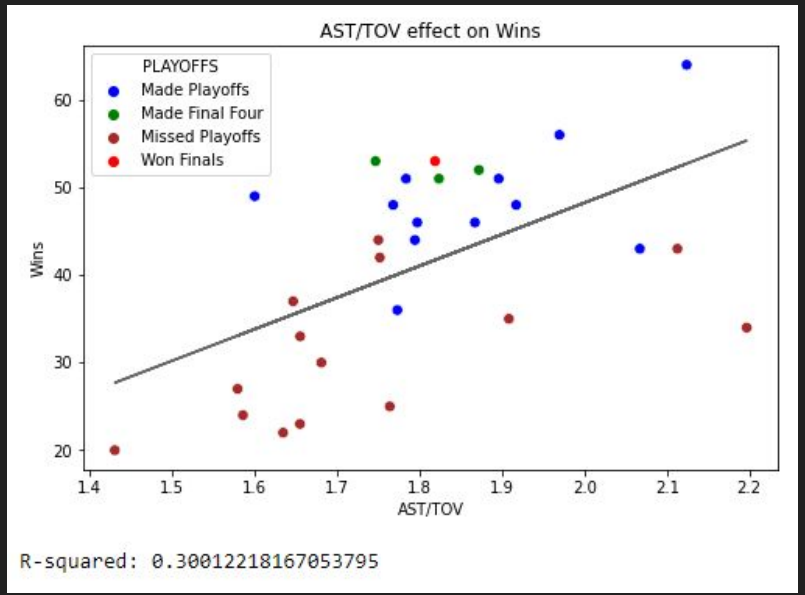
Model Weight: 4%

The AST/TOV ratio is the number of assists divided by the number of turnovers - effectively, it refers to how good a team is at passing.

This statistic removes magnitude from the equation, and looks solely at efficiency (how often a pass will lead to a mistake)





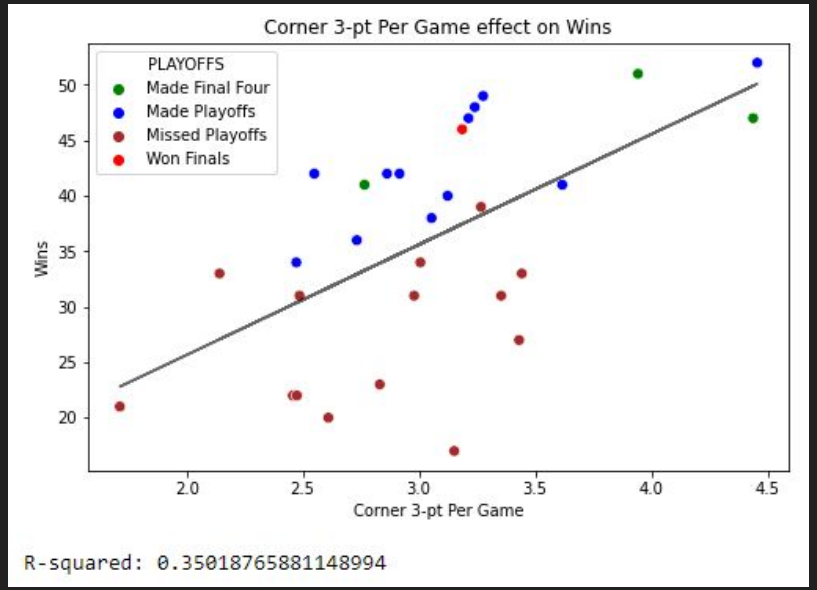


Stat 5: Corner 3-pt Shots Per Game

Model Weight: 2%

This statistic refers to the average number of corner 3-pointers a team makes per game.

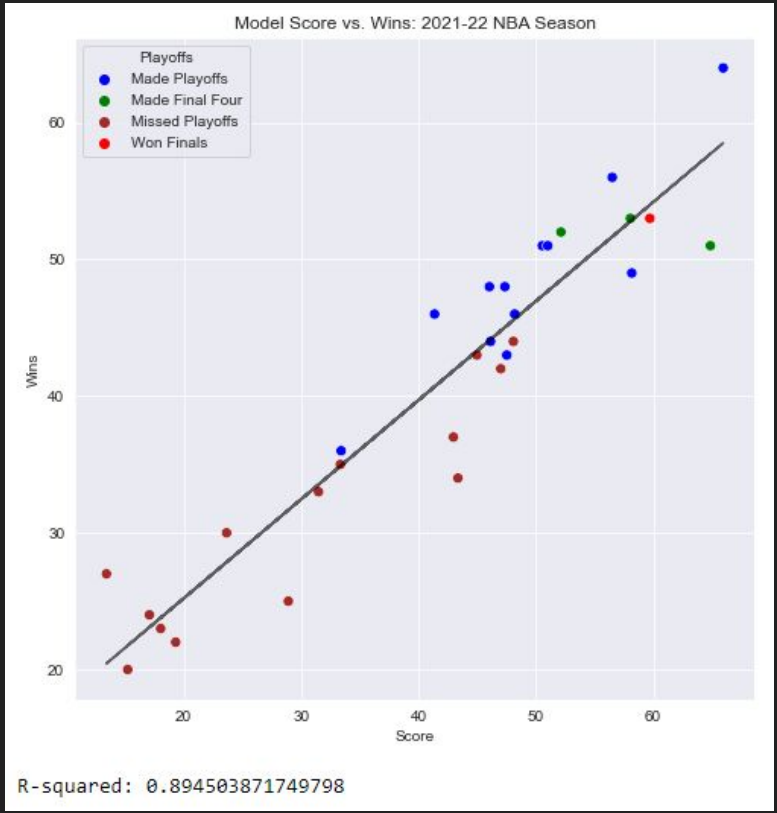
This shot is slightly shorter than a regular 3-pointer, and has been shown mathematically to be the best shot on the floor besides a dunk or layup.



Calculating the Score

Score = (# of games in season) * (((NRTG + 12.5) / 25) * 0.70) +
(((DEF EFG - 0.5) / 0.06) * 0.13) +
(((3P RATE - 0.31) / 0.07) * 0.11) +
(((AST/TOV RATE - 1.5) / 0.6) * 0.04) +
(((CORNER 3PT/G - 2.5) / 1.5) * 0.02)

Results: 2021-22 Season



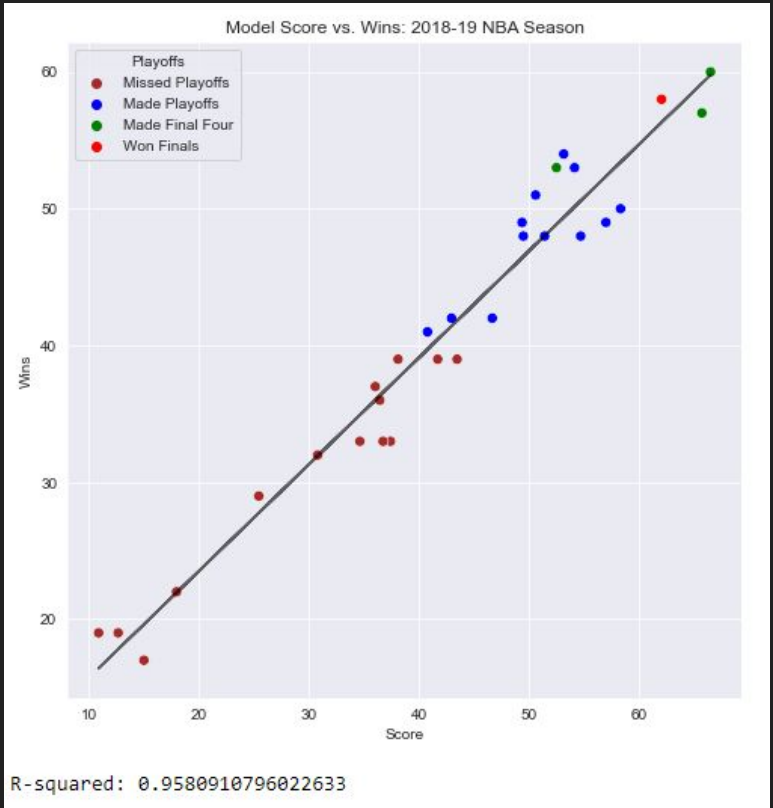
Team	Wins	Rank	Score	Modeled Rank	Difference
Phoenix Suns	64	1	65.9608	1	1.9608
Memphis Grizzlies	56	2	56.5062	6	0.5062
Golden State Warriors	53	3	59.7124	3	6.7124
Miami Heat	53	3	58.0478	5	5.0478
Dallas Mavericks	52	5	52.1356	7	0.1356
Boston Celtics	51	6	64.8702	2	13.8702
Philadelphia 76ers	51	6	51.0122	8	0.0122
Milwaukee Bucks	51	6	50.5448	9	-0.4552
Utah Jazz	49	9	58.1708	4	9.1708
Toronto Raptors	48	10	46.0348	16	-1.9652
Denver Nuggets	48	10	47.3550	13	-0.6450
Chicago Bulls	46	12	41.3608	20	-4.6392
Minnesota Timberwolves	46	12	48.1914	10	2.1914
Brooklyn Nets	44	14	46.1332	15	2.1332
Cleveland Cavaliers	44	14	48.0766	11	4.0766
Atlanta Hawks	43	16	47.5190	12	4.5190
Charlotte Hornets	43	16	44.9688	17	1.9688
Los Angeles Clippers	42	18	47.0024	14	5.0024
New York Knicks	37	19	42.9680	19	5.9680
New Orleans Pelicans	36	20	33.3986	21	-2.6014
Washington Wizards	35	21	33.3248	22	-1.6752
San Antonio Spurs	34	22	43.3534	18	9.3534
Los Angeles Lakers	33	23	31.4716	23	-1.5284
Sacramento Kings	30	24	23.6324	25	-6.3676
Portland Trail Blazers	27	25	13.3906	30	-13.6094
Indiana Pacers	25	26	28.8968	24	3.8968
Oklahoma City Thunder	24	27	17.0560	28	-6.9440
Detroit Pistons	23	28	18.0154	27	-4.9846
Orlando Magic	22	29	19.2946	26	-2.7054
Houston Rockets	20	30	15.2028	29	-4.7972

Notes

Our model overrated all four teams that made it to the conference finals, especially the Celtics (+13.9) and Warriors (+6.7).

Part 4:
Testing the Model

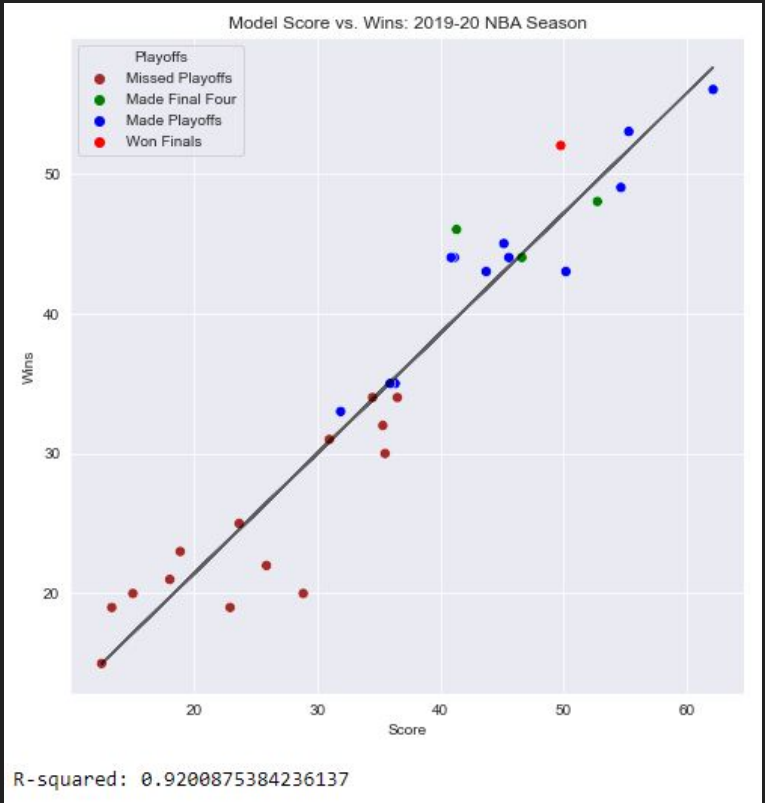
Results: 2018-19 Season



	Team	Wins	Rank	Score	Modeled Rank	Difference
16	Milwaukee Bucks*	60	1	66.5184	1	6.5184
27	Toronto Raptors*	58	2	62.0658	3	4.0658
9	Golden State Warriors*	57	3	65.7312	2	8.7312
7	Denver Nuggets*	54	4	53.1688	8	-0.8312
24	Portland Trail Blazers*	53	5	52.5046	9	-0.4954
10	Houston Rockets*	53	5	54.1528	7	1.1528
22	Philadelphia 76ers*	51	7	50.6104	11	-0.3896
28	Utah Jazz*	50	8	58.3348	4	8.3348
1	Boston Celtics*	49	9	57.0064	5	8.0064
20	Oklahoma City Thunder*	49	9	49.3804	13	0.3804
12	Los Angeles Clippers*	48	11	49.4952	12	1.4952
11	Indiana Pacers*	48	11	54.7104	6	6.7104
26	San Antonio Spurs*	48	11	51.4304	10	3.4304
2	Brooklyn Nets*	42	14	42.9598	16	0.9598
21	Orlando Magic*	42	14	46.6662	14	4.6662
8	Detroit Pistons*	41	16	40.7786	18	-0.2214
15	Miami Heat	39	17	43.4682	15	4.4682
3	Charlotte Hornets	39	17	38.0972	19	-0.9028
25	Sacramento Kings	39	17	41.7052	17	2.7052
13	Los Angeles Lakers	37	20	36.0308	23	-0.9692
17	Minnesota Timberwolves	36	21	36.4408	22	0.4408
6	Dallas Mavericks	33	22	37.4084	20	4.4084
18	New Orleans Pelicans	33	22	36.7360	21	3.7360
14	Memphis Grizzlies	33	22	34.6286	24	1.6286
29	Washington Wizards	32	25	30.7910	25	-1.2090
0	Atlanta Hawks	29	26	25.4364	26	-3.5636
4	Chicago Bulls	22	27	17.9498	27	-4.0502
5	Cleveland Cavaliers	19	28	10.8650	30	-8.1350
23	Phoenix Suns	19	28	12.6444	29	-6.3556
19	New York Knicks	17	30	14.9896	28	-2.0104

Notes

Results: 2019-20 Season

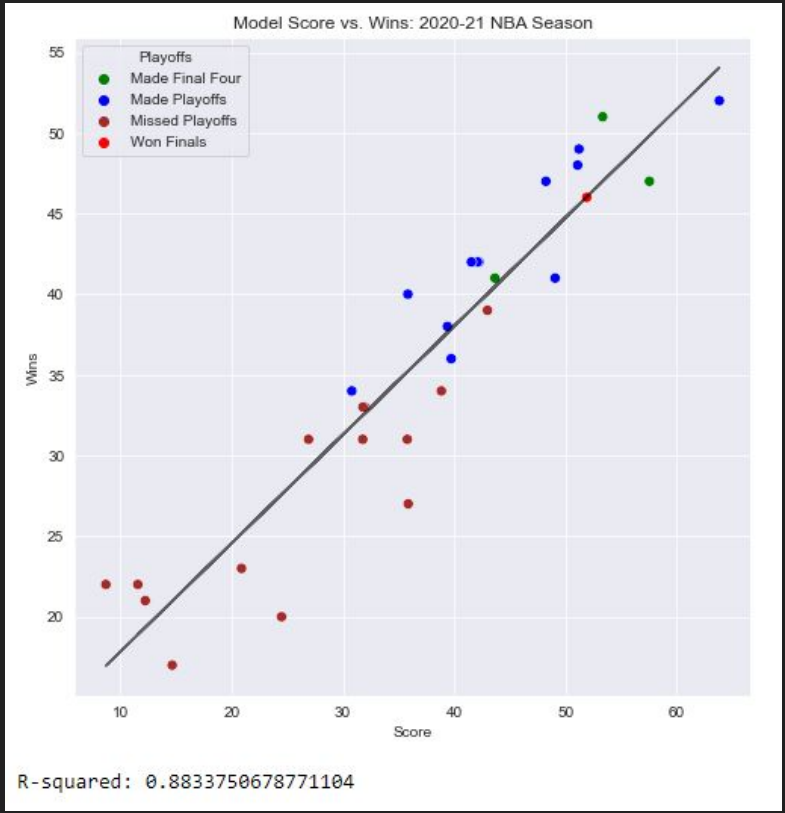


	Team	Wins	Rank	Score	Modeled Rank	Difference
16	Milwaukee Bucks*	56	1	62.1216	1	6.1216
27	Toronto Raptors*	53	2	55.2816	2	2.2816
13	Los Angeles Lakers*	52	3	49.7592	6	-2.2408
12	Los Angeles Clippers*	49	4	54.6408	3	5.6408
1	Boston Celtics*	48	5	52.7328	4	4.7328
7	Denver Nuggets*	46	6	41.3064	11	-4.6936
11	Indiana Pacers*	45	7	45.1440	9	0.1440
10	Houston Rockets*	44	8	41.1264	12	-2.8736
15	Miami Heat*	44	8	46.6056	7	2.6056
28	Utah Jazz*	44	8	45.5472	8	1.5472
20	Oklahoma City Thunder*	44	8	40.8672	13	-3.1328
6	Dallas Mavericks*	43	12	50.1768	5	7.1768
22	Philadelphia 76ers*	43	12	43.7112	10	0.7112
24	Portland Trail Blazers*	35	14	35.9136	16	0.9136
2	Brooklyn Nets*	35	14	36.3240	15	1.3240
23	Phoenix Suns	34	16	36.5040	14	2.5040
14	Memphis Grizzlies	34	16	34.4880	19	0.4880
21	Orlando Magic*	33	18	31.9032	20	-1.0968
26	San Antonio Spurs	32	19	35.3232	18	3.3232
25	Sacramento Kings	31	20	30.9816	21	-0.0184
18	New Orleans Pelicans	30	21	35.5032	17	5.5032
29	Washington Wizards	25	22	23.6736	24	-1.3264
3	Charlotte Hornets	23	23	18.8856	26	-4.1144
4	Chicago Bulls	22	24	25.8840	23	3.8840
19	New York Knicks	21	25	18.0360	27	-2.9640
8	Detroit Pistons	20	26	28.8648	22	8.8648
0	Atlanta Hawks	20	26	15.0408	28	-4.9592
17	Minnesota Timberwolves	19	28	22.9320	25	3.9320
5	Cleveland Cavaliers	19	28	13.3272	29	-5.6728
9	Golden State Warriors	15	30	12.4992	30	-2.5008

Notes

Lakers rank (6) was the lowest of any champion. Our model really didn't like their lack of 3-point shooting and average defense, but both of those turned around drastically in the playoffs that year.

Results: 2020-21 Season



	Team	Wins	Rank	Score	Modeled Rank	Difference
28	Utah Jazz*	52	1	63.8424	1	11.8424
23	Phoenix Suns*	51	2	53.3520	3	2.3520
22	Philadelphia 76ers*	49	3	51.2208	5	2.2208
2	Brooklyn Nets*	48	4	51.0984	6	3.0984
7	Denver Nuggets*	47	5	48.2472	8	1.2472
12	Los Angeles Clippers*	47	5	57.5568	2	10.5568
16	Milwaukee Bucks*	46	7	51.9192	4	5.9192
24	Portland Trail Blazers*	42	8	41.5440	13	-0.4560
6	Dallas Mavericks*	42	8	42.2280	11	0.2280
13	Los Angeles Lakers*	42	8	42.0912	12	0.0912
0	Atlanta Hawks*	41	11	43.6608	9	2.6608
19	New York Knicks*	41	11	49.0752	7	8.0752
15	Miami Heat*	40	13	35.8344	18	-4.1656
9	Golden State Warriors	39	14	42.9912	10	3.9912
14	Memphis Grizzlies*	38	15	39.3840	15	1.3840
1	Boston Celtics*	36	16	39.7296	14	3.7296
11	Indiana Pacers	34	17	38.8440	16	4.8440
29	Washington Wizards*	34	17	30.7872	23	-3.2128
3	Charlotte Hornets	33	19	31.9896	20	-1.0104
26	San Antonio Spurs	33	19	31.7880	21	-1.2120
18	New Orleans Pelicans	31	21	31.7736	22	0.7736
4	Chicago Bulls	31	21	35.7768	19	4.7768
25	Sacramento Kings	31	21	26.8992	24	-4.1008
27	Toronto Raptors	27	24	35.8632	17	8.8632
17	Minnesota Timberwolves	23	25	20.8728	26	-2.1272
5	Cleveland Cavaliers	22	26	11.5488	29	-10.4512
20	Oklahoma City Thunder	22	26	8.6760	30	-13.3240
21	Orlando Magic	21	28	12.2184	28	-8.7816
8	Detroit Pistons	20	29	24.4728	25	4.4728
10	Houston Rockets	17	30	14.6376	27	-2.3624

Notes

Part 5:
Predicting the Future

Through the first ~20 games of the 2022-23 NBA season, our model predicts the following results:

East:

TEAM	REC	PCT
Boston Celtics	66-16	0.805
Cleveland Cavaliers	61-21	0.744
Milwaukee Bucks	54-28	0.659
Philadelphia 76ers	54-28	0.659
Brooklyn Nets	48-34	0.585
Indiana Pacers	44-38	0.537
Chicago Bulls	39-43	0.476
Atlanta Hawks	38-44	0.463
Toronto Raptors	37-45	0.451
Washington Wizards	37-45	0.451
Miami Heat	35-47	0.427
New York Knicks	31-51	0.378
Charlotte Hornets	26-56	0.317
Orlando Magic	21-61	0.256
Detroit Pistons	17-65	0.207

West:

TEAM	REC	PCT
Phoenix Suns	64-18	0.780
New Orleans Pelicans	56-26	0.683
Utah Jazz	46-36	0.561
Golden State Warriors	45-37	0.549
Memphis Grizzlies	44-38	0.537
Denver Nuggets	44-38	0.537
Dallas Mavericks	43-39	0.524
Sacramento Kings	42-40	0.512
Los Angeles Clippers	42-40	0.512
Portland Trail Blazers	39-43	0.476
Minnesota Timberwolves	36-46	0.439
Oklahoma City Thunder	35-47	0.427
Los Angeles Lakers	31-51	0.378
Houston Rockets	21-61	0.256
San Antonio Spurs	8-64	0.098

Champion:

Runner Up:

#1 Pick:

Part 6:
Comments and Improvements

Methods for Improvement

- Use of more detailed and complex stats, beyond what is freely available online
  - Teams have access to newly developed aggregate statistics and real-time player data (location, speed, etc.) that can be incorporated into similarly structured models.

- Include individual stats (player-based model) or consider matchups between teams (ELO model)
  - Most detailed models used by the teams are based on one of these.