Paul Regan
Business Statistics
Midterm Project

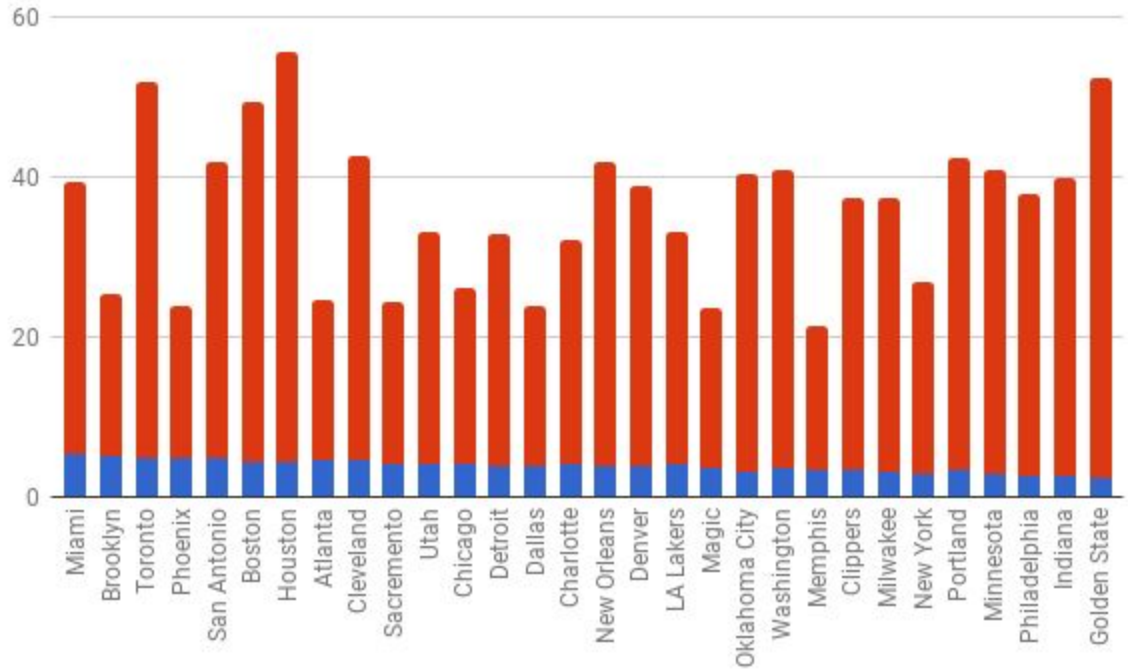
Introduction:

A common criticism of modern NBA is an emphasis on 3 point shooting. When listening to popular talk radio shows or ESPN talk shows many people continuously complain about the aggressive approach to volume 3 point attempts per game. But does shooting and making more three point shots actually correlate with a winning team? Instead of attempting to find a correlation with overall three pointers made by each team I am going to see if there is a correlation with number of three pointers made by each teams bench and number of wins.

1	Miami	65	19.3	38.2	17.2	8.0	2.5	1.8	5.1	7.9	13.7	3.5	13.7-30.4	.449	5.3-14.4
2	Brooklyn	65	21.1	43.7	19.3	8.6	2.7	1.8	6.6	9.6	14.8	4.6	16.0-36.3	.442	5.2-14.3
3	Toronto	64	21.0	41.1	17.7	9.4	4.2	3.3	5.0	10.8	13.3	4.5	15.5-33.5	.463	4.8-14.0
4	Phoenix	66	19.7	37.6	17.4	8.1	2.9	1.8	5.6	9.6	13.3	4.1	13.3-32.3	.412	4.8-13.9
5	San Antonio	64	20.4	41.7	14.9	9.2	3.0	1.4	5.9	8.2	12.1	2.9	15.4-34.6	.445	4.8-13.4
6	Boston	65	19.3	34.2	17.8	8.2	3.1	1.5	5.0	9.0	13.4	4.3	12.1-29.9	.406	4.4-13.2
7	Houston	64	16.7	30.5	13.8	3.3	2.7	1.6	3.9	8.6	10.7	3.1	10.9-24.4	.446	4.4-12.8
8	Atlanta	65	20.8	40.7	17.9	9.3	2.9	1.9	5.2	8.8	13.5	4.4	14.8-33.1	.445	4.5-12.7
9	Cleveland	64	19.5	41.7	15.8	7.2	2.6	1.8	4.4	8.4	12.6	3.2	15.2-32.2	.472	4.6-12.5
10	Sacramento	65	22.4	45.3	18.6	9.5	3.9	1.9	5.7	9.5	14.3	4.3	17.5-38.6	.454	4.6-11.8
11	Utah	65	18.1	35.7	15.1	6.2	3.2	1.6	4.2	7.4	12.4	2.6	12.9-30.3	.426	4.2-11.7
12	Chicago	64	18.6	39.6	18.3	8.6	2.8	1.2	5.0	8.4	14.1	4.2	14.8-33.4	.444	4.0-11.6
13	Detroit	65	17.3	32.8	13.6	7.8	2.5	1.1	4.3	7.3	10.6	3.0	12.1-28.3	.429	4.0-11.4
14	Dallas	65	18.8	36.1	17.1	9.7	3.1	1.9	5.1	9.7	13.2	3.9	13.3-28.7	.463	3.9-11.0
15	Charlotte	65	18.7	37.0	16.0	7.1	2.8	1.5	3.8	8.4	12.3	3.7	13.3-32.4	.412	3.8-10.7
16	New Orleans	64	16.4	26.5	11.1	6.2	1.9	1.0	3.4	7.4	9.2	2.0	10.0-22.2	.448	4.0-10.6
17	Denver	65	16.8	35.6	15.8	7.7	2.6	1.6	5.0	7.0	11.9	3.9	13.5-29.9	.450	3.6-10.4
18	L.A.Lakers	64	18.7	41.8	17.7	8.8	3.0	1.6	6.2	8.7	13.2	4.5	16.1-34.3	.470	3.2-10.3
19	Orlando	65	18.2	33.6	14.8	7.9	2.6	1.9	5.0	8.1	11.7	3.2	12.1-27.6	.440	3.7-10.3
20	Oklahoma City	66	16.2	25.6	11.7	4.7	2.4	1.9	2.9	7.3	8.7	3.0	9.5-22.2	.428	3.4-10.2
21	Washington	65	17.8	34.5	15.1	7.0	2.7	1.5	4.8	10.4	11.2	3.9	12.5-27.6	.453	3.3-9.8
22	Memphis	64	19.2	36.9	15.3	7.2	3.6	2.1	5.6	10.3	11.3	4.1	13.9-30.2	.462	3.2-9.3
23	L.A.Clippers	63	17.8	42.4	14.2	8.2	3.1	1.9	5.1	8.2	10.8	3.3	15.5-32.3	.480	2.9-9.2
24	Milwaukee	65	16.6	25.5	12.1	6.0	2.6	1.7	3.6	8.0	9.4	2.7	9.5-21.9	.432	3.3-9.1
25	New York	65	19.7	37.2	17.2	9.5	3.0	2.1	6.2	9.7	13.0	4.2	14.5-31.3	.464	2.8-8.2
26	Portland	65	17.9	28.3	18.1	5.3	2.5	2.0	4.2	8.6	13.6	4.5	10.7-24.2	.443	2.7-7.9
27	Minnesota	66	13.5	25.8	9.7	6.1	2.4	0.7	2.9	6.0	7.3	2.3	9.7-21.9	.445	2.7-7.6
28	Philadelphia	63	16.6	27.3	14.3	7.5	2.8	1.4	4.4	9.1	10.1	4.2	10.6-23.3	.455	2.4-7.4
29	Indiana	65	16.6	32.0	16.3	7.7	2.2	1.1	4.4	7.5	12.6	3.7	12.7-27.4	.464	2.3-6.6
30	Golden State	64	19.2	33.2	16.9	8.9	3.4	3.5	4.7	9.4	12.5	4.4	13.5-25.9	.521	2.1-6.2

Immediate Observations:

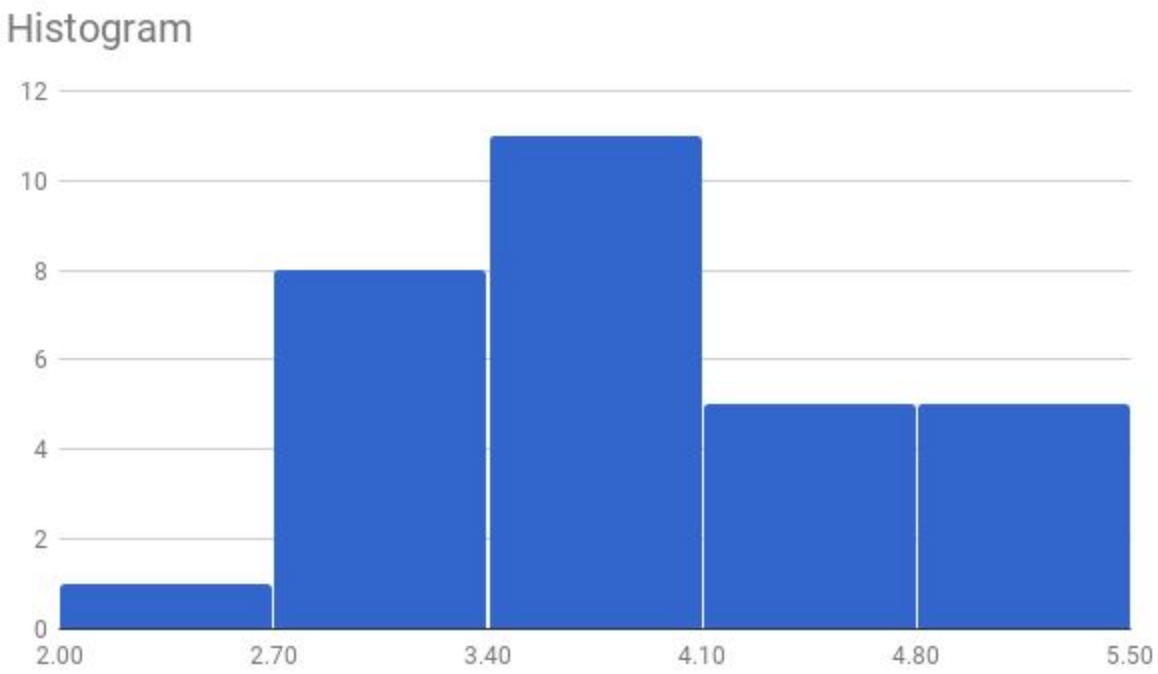
Golden State has dominated the NBA over the past three years with their volume of three point attempts paired with their deadly accuracy. Few would argue that they are not the best team in the NBA and shooting is their calling card. I am interested to see potential patterns that arrive from looking at teams benches instead of their starters.

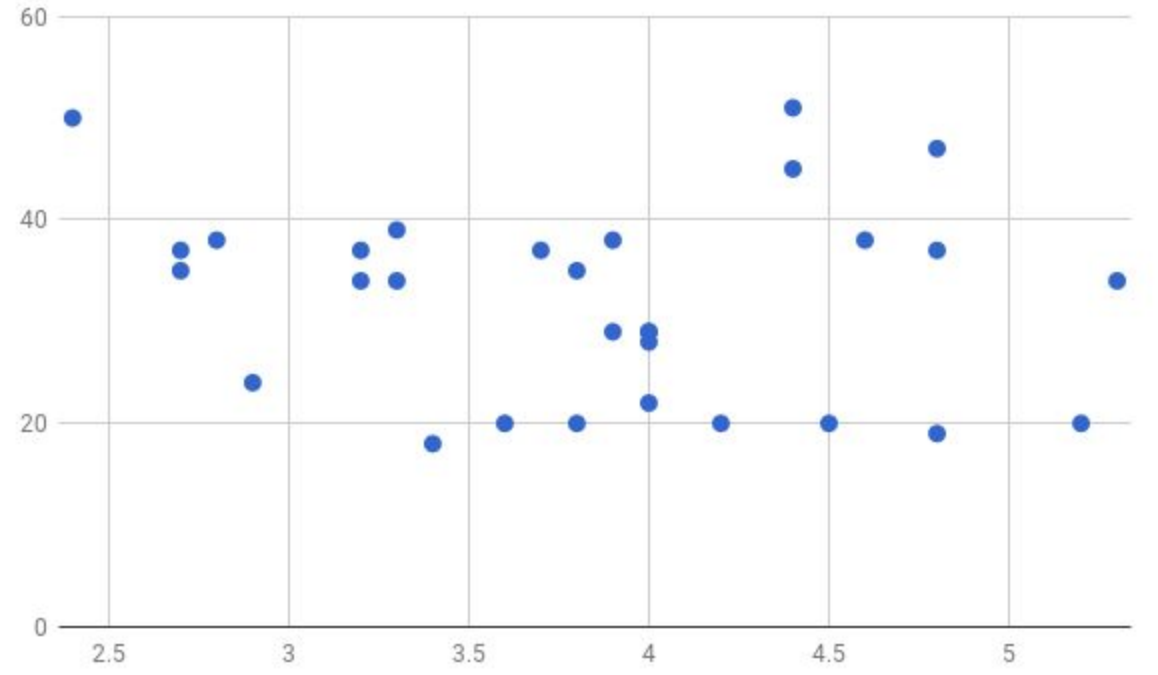


We can see that there does seem to be a cluster of teams with a large number of wins towards the left of the bar graph. As the blue bar, which represents number of 3 pointers made by corresponding team, decreases there does not seem to be a clear pattern indicating that their is any correlation. One interesting note is the worse teams seem to be near the middle of this bar graph. Potentially with a larger sample size this graph may even look bimodal. With this given data it would be hard to say that with confidence but for now we can investigate if the graph is unimodal by examining the teams' bench players that have hit the most threes.

Hypothesis: If my bench is top in the league in three pointers made then my team will win more games.

Test:





These two charts allow a nice visual for the hypothesis potential. The histogram is interesting because it separates the data into sections. These sections are not quartiles because they do

not represent the smallest to the middle number between the smallest and the median. The sections the histogram produce work nicer for my project so I am going to use the ranges they allocate as the ranges I will test and scrutinize.

For this project, especially, it is interesting to observe data within ranges of values because the data does not seem to change much but within ranges or sections of the data there may be valuable data. For example, Only one team hit between 2.0 and 2.7 three point shots. This team is regarded as the best team in the NBA. Intuitively we can predict a potential pattern regarding teams in sections where they hit the most threes and the least threes.

Discussion:

First let's see if being in the section of data representing teams with the most three point shots made on average correlates with more wins.

			Deviation
Miami	**5.3**	**34**	**3^2 = 9**
Brooklyn	**5.2**	**20**	**11^2 = 121**
Toronto	**4.8**	**47**	**16^2 = 256**
Phoenix	**4.8**	**19**	**12^2=144**
San Antonio	**4.8**	**37**	**6^2=36**
Boston	**4.4**	**45**	**14^2 = 196**
Houston	**4.4**	**51**	**20^2 =400**
Atlanta	**4.5**	**20**	**11^2=212**
Cleveland	**4.6**	**38**	**7^2=49**
Sacramento	**4.2**	**20**	**11^2=121**

Average Wins per quartile: (1st : smallest to halfway to median 2nd: halfway to median to median 3rd: median to halfway to largest 4th: halfway from median to largest to largest)

1st: 36.14
2nd: 25.75
3rd: 30.57
4th: 36.37

The first and last quartiles result in more wins than the second and third. These quartiles are worthy of investigation.

Investigation:

This data represents the bars in the Histogram from 4.1-5.5 threes made. In the NBA 42 wins is .500 record and in many years would get a team into the playoffs. We will use 30 wins as a benchmark for a successful season since each team is around 60 games played in the 2017-2018 season. Of the 10 team sample, 6 teams would be considered successful and 4 teams would not.

This sample would imply that if you are in the upper tier of three point shooting you may have a slightly greater chance of being a successful team than not. This sample may be disputed because the sample size is only 10 teams. Although a representative sample, there is a much larger margin of error because the sample size is the minimum.

If we were to assume this sample to be representative of the data there would definitely be something to work with but since teams like Brooklyn, who have minimal amount of wins, can be the number two three point shooting bench there is definitely a chance for type errors and sampling error.

For example,

Number of threes made per this sample:

The mean of this data is 4.7.
Squared Differences: (.6) ^2 =.36
(.5)^2 = .25
(.1)^2= .01
(.1)^2= .01
(.1)^2= .01
(-.3)^2= .09
(-.3)^2= .09
(-.2)^2= .04
(-.1)^2= .01
(-.5)^2= .25
Sum = 1.12/n-1=1.12/9= .12444

95% Confidence Interval: 4.7 ± 0.077
(4.623 to 4.777)

Short Style: 4.7 (95% CI 4.623 to 4.777)

Margin of Error: 0.077
(With more digits: 0.07685)

Sample Size: 10
Sample Mean: 4.7
Sample Standard Deviation: 0.124
Confidence Level: 95%

Number of Wins per this sample:

Sample Standard Deviation = 154.4

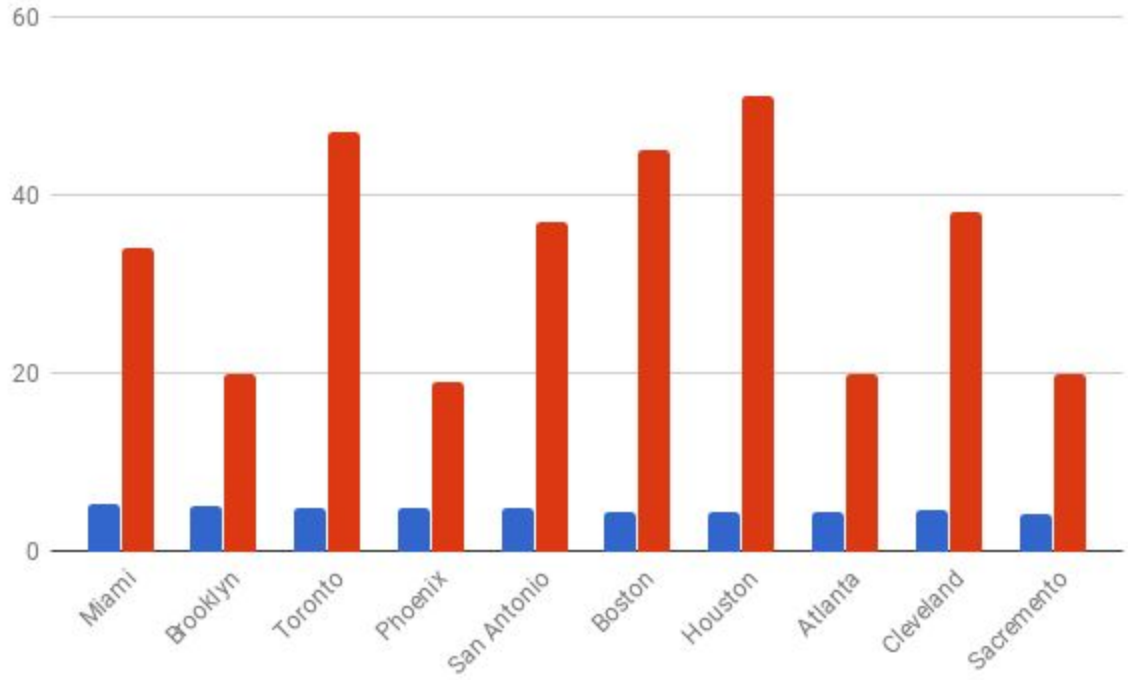
95% Confidence Interval: 31 ± 96
(-65 to 127)

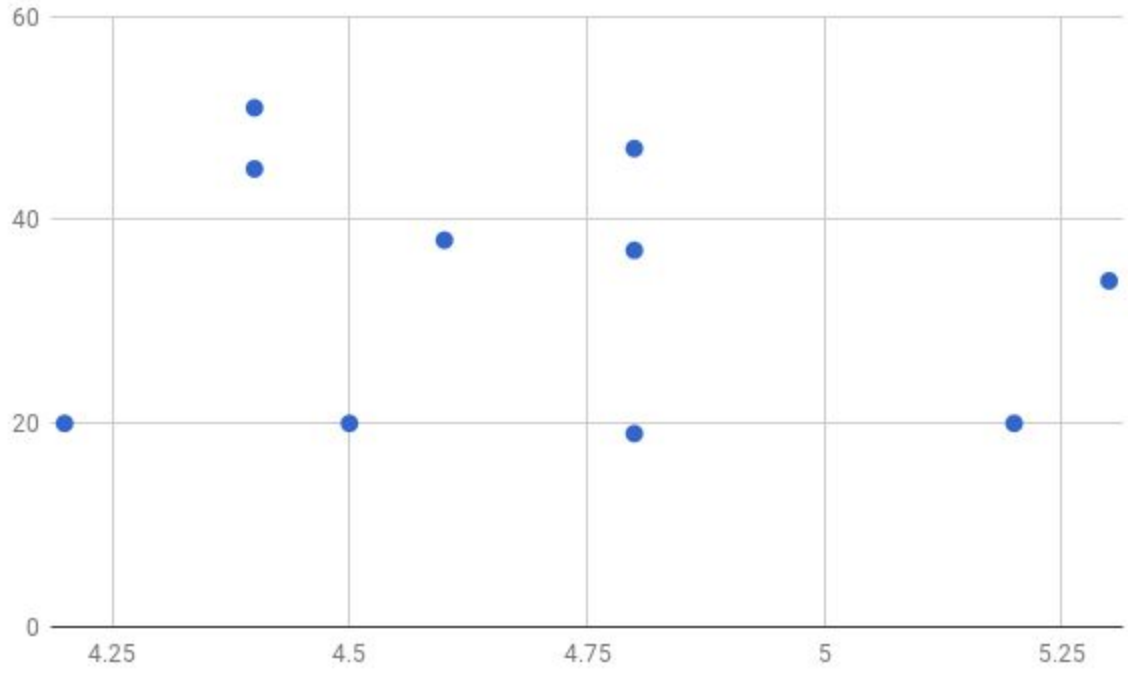
Short Style: 31 (95% CI -65 to 127)

Margin of Error: 96
(With more digits: 95.70)

Sample Size: 10
Sample Mean: 31
Sample Standard Deviation: 154.4
Confidence Level: 95%

This data implies that although the data is clustered well for amount of three pointers per bench team, those teams can range in wins. The z score representing 95 percent confidence significantly enough to help disprove a potential correlation. But we still have some questions considering that the averages for each quartile are greater in the first and last. The investigation does not stop at this perceived road block but some valuable questions are answered.


60
40
20
0
Miami
Brooklyn
Toronto
Phoenix
San Antonio
Boston
Houston
Atlanta
Cleveland
Sacramento


60
40
20
0
4.25
4.5
4.75
5
5.25

Reflection:

These two charts represent the section of data within the sample that we are scrutinizing. One interesting note is that the unsuccessful teams in this sample all have around the exact same record while the successful teams seem to have varying records. The standard deviation and margin of error of the sample regarding good three point shooting benches is very low so the data is properly clustered properly. This allows us to make more accurate predictions regarding the data.

The problem occurs with our second calculation of data. The section's deviation in teams total wins is very large which would indicate no correlation with bench three pointers and total wins.

Some observations that may help keep our hypothesis alive are that the mean amount of wins is noticeably different in the first and fourth quartiles. Also, the unsuccessful teams all have almost the exact same amount of wins. Upon further investigation there maybe some valuable conclusions from this hypothesis but right now I have not developed enough evidence to make any definitive statements regarding the data.

This project did not produce a conclusion but it does raise some important questions regarding this data. Although through statistical analysis strategies we create some conjectures regarding the data we would need a more thorough analysis for anything concrete to develop. I do not see this as a failure, but quite the opposite. I assumed that the correlation between wins and stats is not apparent or every team would know the formula for winning which they do not. Not even close. Basketball games have many variables because the talent level is so high. For one team to find even the slightest advantage through statistical analysis could be the difference between a 5 point loss and a 5 point win.