Impact of the Bubble on NBA Player Performance

By: Aryaman Sinha
Daryl Foo
Jason Lin
Jonny Tran
Nathan Brodie

Image:
https://media.istockphoto.com/vectors/basketball-orange-grunge-background-vector-id945874236?k=6&m=945874236&s=612x612&w=0&h=8j

Background

- 2019-2020 NBA Season was postponed due to COVID-19.
- NBA came up with the idea of creating a bubble to resume play.
- Players and essential staff members from the majority of teams attended.
- Everyone stayed in Walt Disney World Resort in Orlando, Florida.
- There were no fans in attendance for the vast majority of games.
- The 3-point shot important over the past few season

https://images.app.goo.gl/kmjx2kY455LJmwXK8

References:
https://www.sportingnews.com/us/nba/news/nba-bubble-rules-teams-schedule-orlando/zhap66a9hcwq1khmcex3ggabo
https://shottracker.com/articles/the-3-point-revolution

Research Question

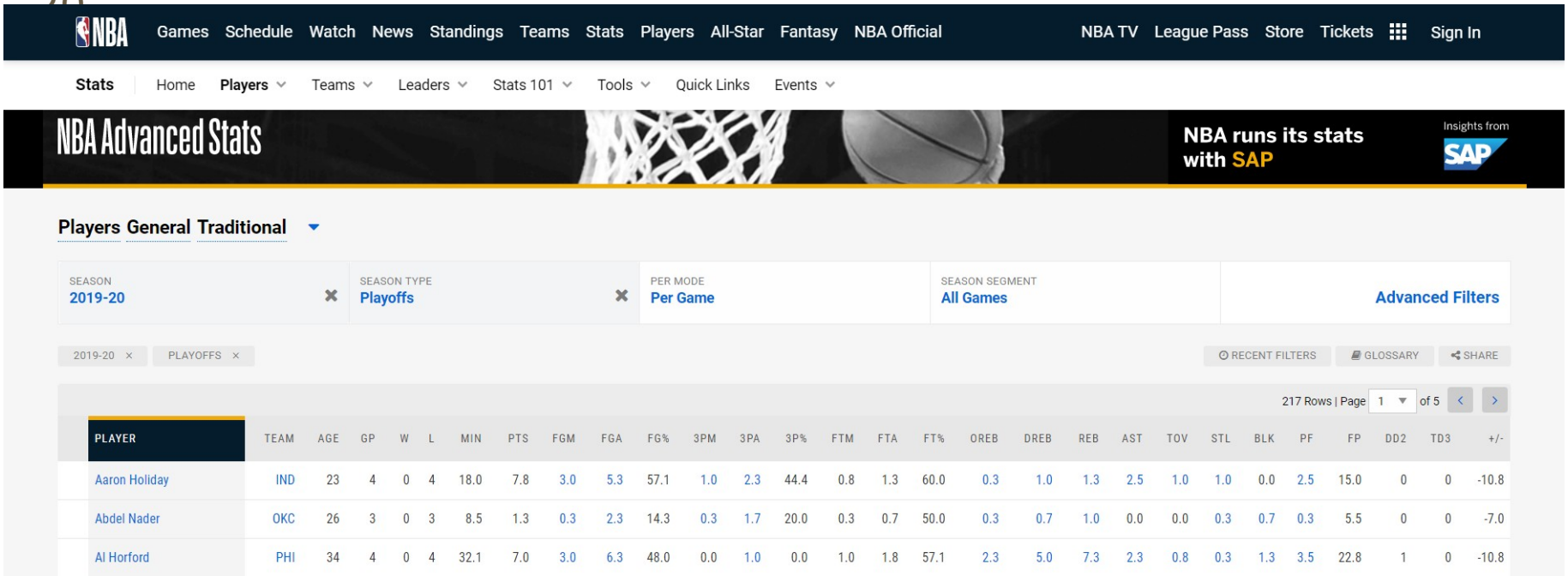
How did the 2019-20 NBA Bubble affect player shooting performance in the the 2019-20 playoffs compared to that of the previous 5 years?

Hypothesis

We believe the 2019-20 NBA Bubble negatively impacted player shooting performance in the 2019-20 playoffs compared to that of the previous 5 years due to a lack of support from fans.

Dataset(s)

All of our datasets were obtained from the NBA official website. The data was on NBA player data in the NBA playoffs from 2014-15 to 2019-20



PLAYER	TEAM	AGE	GP	W	L	MIN	PTS	FGM	FGA	FG%	3PM	3PA	3P%	FTM	FTA	FT%	OREB	DREB	REB	AST	TOV	STL	BLK	PF	FP	DD2	TD3	+/-
Aaron Holiday	IND	23	4	0	4	18.0	7.8	3.0	5.3	57.1	1.0	2.3	44.4	0.8	1.3	60.0	0.3	1.0	1.3	2.5	1.0	1.0	0.0	2.5	15.0	0	0	-10.8
Abdel Nader	OKC	26	3	0	3	8.5	1.3	0.3	2.3	14.3	0.3	1.7	20.0	0.3	0.7	50.0	0.3	0.7	1.0	0.0	0.0	0.3	0.7	0.3	5.5	0	0	-7.0
Al Horford	PHI	34	4	0	4	32.1	7.0	3.0	6.3	48.0	0.0	1.0	0.0	1.0	1.8	57.1	2.3	5.0	7.3	2.3	0.8	0.3	1.3	3.5	22.8	1	0	-10.8

https://www.nba.com/stats/players/traditional/?sort=PLAYER_NAME&dir=-1&Season=2019-20&SeasonType=Playoffs

Data Setup

We then converted the data from the website into .csv format and read them in as pandas data frames (about 200 rows each).

In [1]:
```
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
```

In [2]:
```
#Pre-Covid Data
players1415 = pd.read_csv("DataSetup/nba_1415.csv")
players1516 = pd.read_csv("DataSetup/nba_1516.csv")
players1617 = pd.read_csv("DataSetup/nba_1617.csv")
players1718 = pd.read_csv("DataSetup/nba_1718.csv")
players1819 = pd.read_csv("DataSetup/nba_1819.csv")

#Bubble data
players1920 = pd.read_csv("DataSetup/nba_1920.csv")
```

In [3]:
```
players1415.head()
```

Out[3]:

	PLAYER	TEAM	AGE	GP	W	L	MIN	PTS	FGM	FGA	...	REB	AST	TOV	STL	BLK	PF	FP	DD2	TD3	+/-
0	Aaron Brooks	CHI	30	12	6	6	11.0	4.5	1.8	5.1	...	1.5	0.9	0.8	0.3	0.1	1.3	7.9	0	0	-3.2
1	Al Horford	ATL	29	16	8	8	32.6	14.4	6.7	13.2	...	8.6	3.7	1.6	0.8	1.4	2.1	35.5	7	0	2.5
2	Al-Farouq Aminu	DAL	24	5	1	4	30.0	11.2	3.4	6.2	...	7.2	1.2	1.2	2.0	1.6	3.0	31.2	1	0	3.6
3	Alan Anderson	BKN	32	6	2	4	23.6	11.0	4.2	6.8	...	3.5	1.2	1.3	0.7	0.2	1.5	18.1	0	0	2.3
4	Alexis Ajinca	NOP	27	3	0	3	3.5	2.7	1.3	1.3	...	0.3	0.3	0.0	0.3	0.0	0.7	4.6	0	0	-2.7

5 rows × 29 columns

In [4]:
```
players1415.shape
```

Out[4]:
```
(208, 29)
```

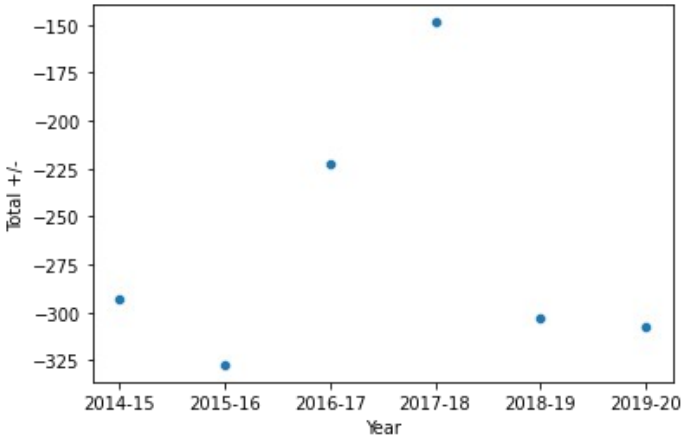
Glossary
GP: Games Played
FG: Field goals
3P: 3-point field goals
FT: Free throw
+/-: Statistic representing player impact

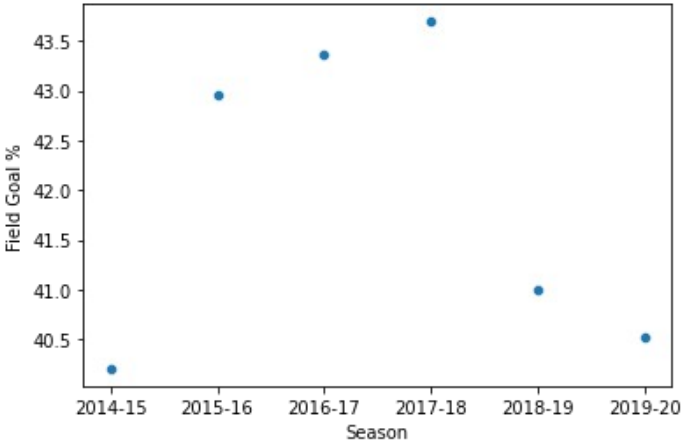
Data Cleaning

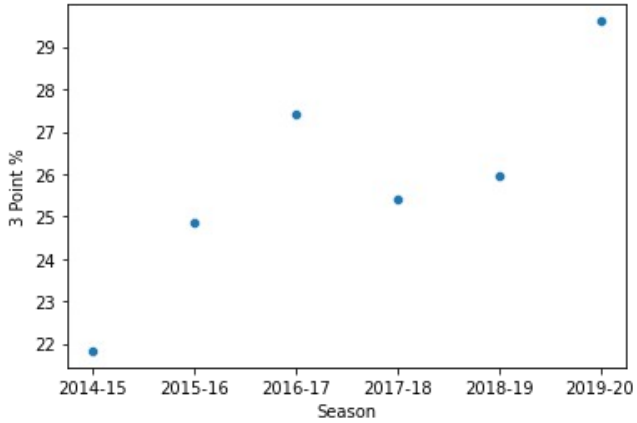
No data cleaning was necessary as there were no missing values in our data frame, and the columns were named well enough for our interpretation.

```
In [9]: players1415.isnull().values.any()
Out[9]: False

In [10]: players1516.isnull().values.any()
Out[10]: False

In [11]: players1617.isnull().values.any()
Out[11]: False

In [12]: players1718.isnull().values.any()
Out[12]: False

In [13]: players1819.isnull().values.any()
Out[13]: False

In [14]: players1920.isnull().values.any()
Out[14]: False
```

Exploratory Data Analysis

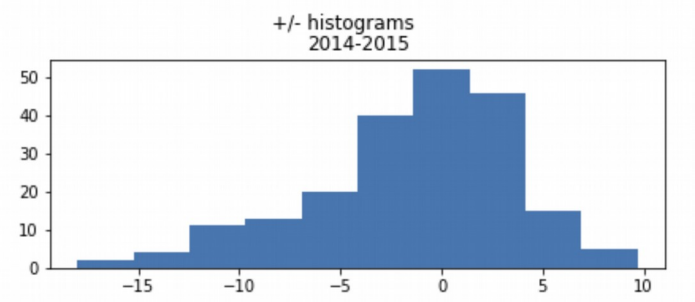


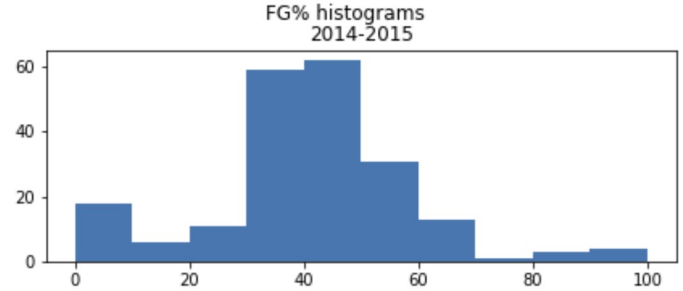




+/-, Field Goal % and 3 Point %

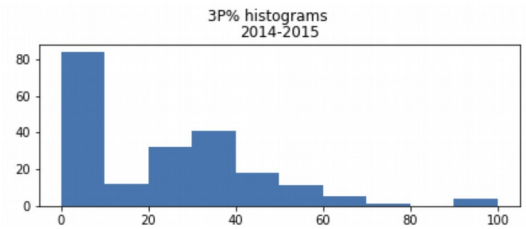
No Changes Needed

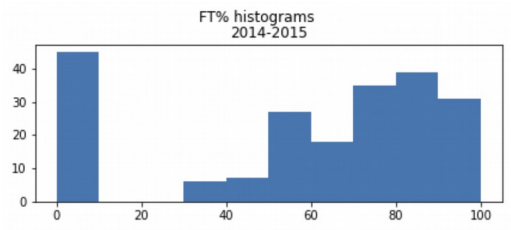




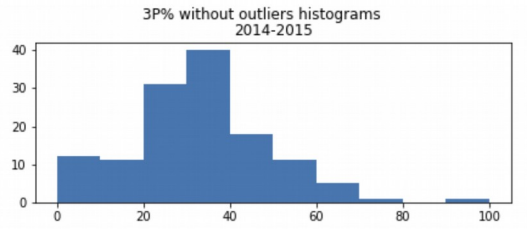
Removed Outliers

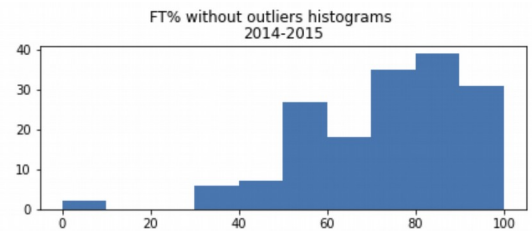
Before



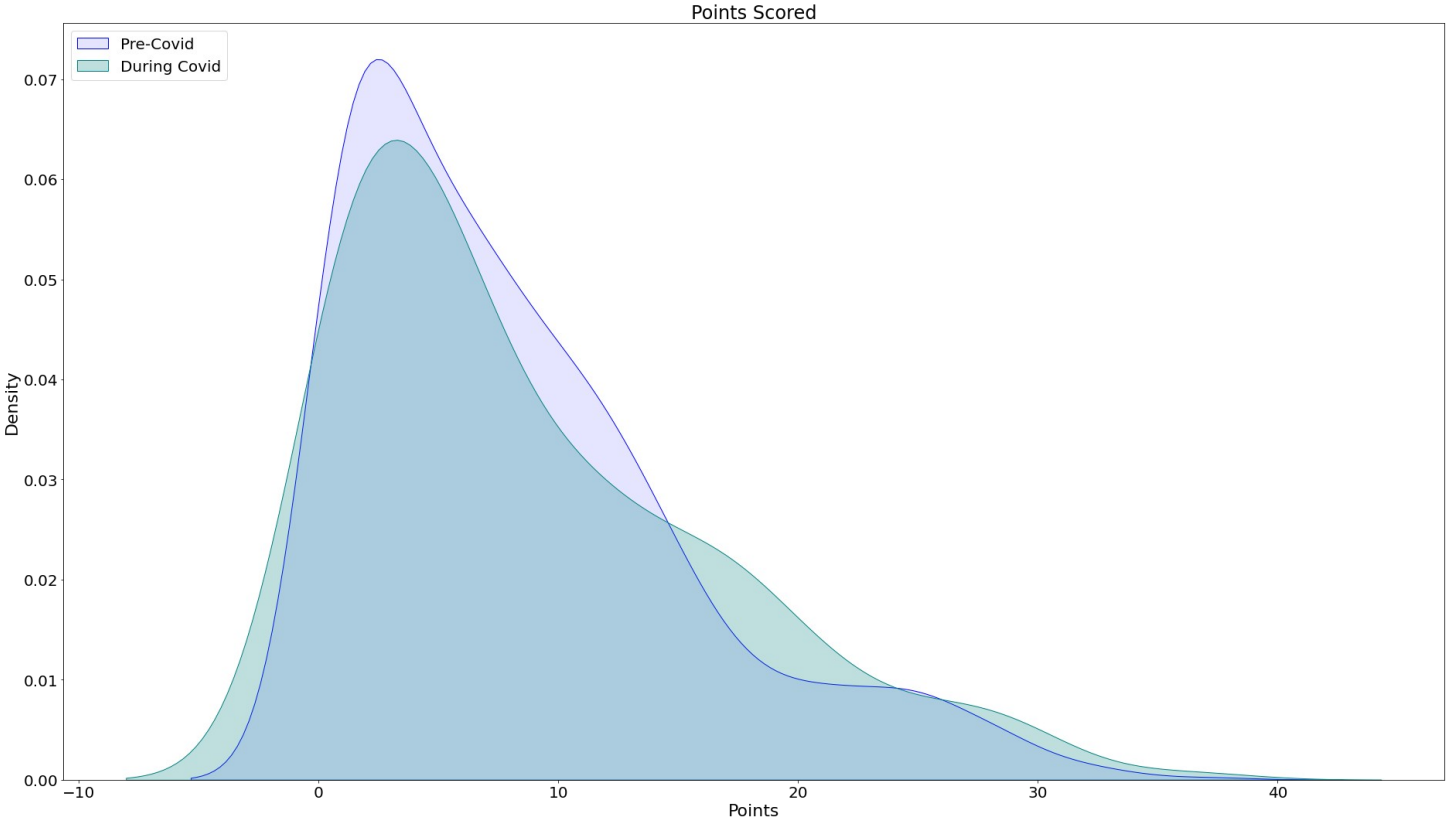


After





EDA continued



Density curves of Points scored

Data Analysis & Results

```
pre_covid_FG = pre_covid['FG%']
during_covid_FG = during_covid['FG%']
stats.ttest_ind(pre_covid_FG, during_covid_FG, equal_var=False)
```

```
Ttest_indResult(statistic=1.3054581869463904, pvalue=0.19271503323294054)
```

```
pre_covid_3P = pd.concat([players1415_3P['3P%'], players1516_3P['3P%'], players1617_3P['3P%'],
                          players1718_3P['3P%'], players1819_3P['3P%']], axis=0)
during_covid_3P = players1920_3P['3P%']
stats.ttest_ind(pre_covid_3P, during_covid_3P, equal_var=False)
```

```
Ttest_indResult(statistic=-1.5869530457326695, pvalue=0.11391326190398245)
```

```
pre_covid_FT = pd.concat([players1415_FT['FT%'], players1516_FT['FT%'], players1617_FT['FT%'],
                          players1718_FT['FT%'], players1819_FT['FT%']], axis=0)
during_covid_FT = players1920_FT['FT%']
stats.mannwhitneyu(pre_covid_FT, during_covid_FT, alternative='two-sided')
```

```
MannwhitneyuResult(statistic=72693.5, pvalue=0.4544912914531958)
```

All 3 tests fail to reject the null hypothesis

Ethics and Privacy

All data used is freely available to the public and our analysis did not involve the use of any personal information.

https://media.istockphoto.com/vectors/stampround2red-vector-id1156274375?k=6&m=1156274375&s=612x612&w=0&h=vk1cHlvR4vcKBOKK9_EZJVOvUssCqFhIyi2LUACsmm8=

https://www.jing.fm/clipimg/detail/22-227066_mens-basketball-clipart-basketball-clip-art.png

Conclusion

After conducting analyses on the statistics like the total 3 Point , Field Goal and Free Throw percentages, we found that there were some deviations pre-covid and during covid.

However, none of these were significant enough based on the hypothesis tests we conducted.

It is reasonable to say that there was no marked difference in player performance and that any deviations could be attributed to random chance and not necessarily due to the bubble.