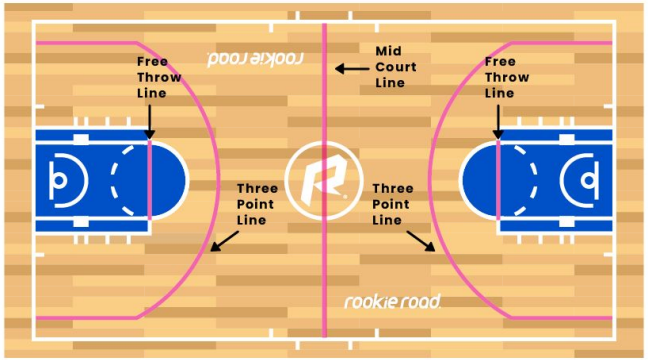
The Weight of the 3-Pointer in the Modern NBA

Jackson Conte
Andrew Nguyen
Anish Rajeshkumar
Baraa Zekeria

COGS 108 Fall 2022 Group 018

Overview



Basketball Scoring

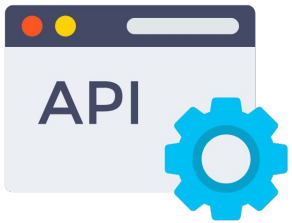
Question

Is there a relationship between how much an NBA team uses 3-point field goals (compared to 2-point field goals) and how well they do in the playoffs in the past 20 years?

Hypothesis

- We predicted that there would be a relationship
- The 3-point strategy became famous when it was used by notable playoff teams (e.g. the Warriors dynasty) — and we think they used it for a reason
- The playstyle is still relatively new in the NBA, so teams may not have fully figured out how to counter it

Data



pandera

```
schema = pa.DataFrameSchema(
    {
        "TEAM_ID": Column(int, Check.isin(team_ids), nullable = False),
        "TEAM_CITY": Column(str),
        "TEAM_NAME": Column(str, Check.isin(team_nickname)),
        "YEAR": Column(int, Check.greater_than(1999), Check.less_than(2022)),
        "GP": Column(int, Check.greater_than(0), Check.less_than(83)),
        "WINS": Column(int, Check.greater_than(-1), Check.less_than(83)),
        "LOSSES": Column(int, Check.greater_than(-1), Check.less_than(83)),
        "WIN_PCT": Column(float, Check.greater_than(0), Check.less_than(1.0)),
        "CONF_RANK": Column(int),
        "DIV_RANK": Column(int),
        "PO_WINS": Column(int, Check.greater_than(-1), Check.less_than(29)),
        "PO_LOSSES": Column(int, Check.greater_than(-1), Check.less_than(29)),
        "NBA_FINALS_APPEARANCE": Column(str),
        "FGM": Column(int, Check.greater_than(-1)),
        "FGA": Column(int, Check.greater_than(-1)),
        "FG_PCT": Column(float, Check.greater_than(0), Check.less_than(1.0), lambda x: x == (x["FGM"] / x["FGA"])),
        "FG3M": Column(int, Check.greater_than(-1)),
        "FG3A": Column(int, Check.greater_than(-1)),
        "FG3_PCT": Column(float, Check.greater_than(0), Check.less_than(1.0), lambda x: x == (x["FG3M"] / x["FG3A"])
        "FTM": Column(int, Check.greater_than(-1)),
        "FTA": Column(int, Check.greater_than(-1)),
        "FT_PCT": Column(float, Check.greater_than(0), Check.less_than(1.0), lambda x: x == (x["FTM"] / x["FTA"])),
        "PTS": Column(int),
        "PTS_RANK": Column(int, Check.greater_than(0), Check.less_than(31)),
    }
)
```

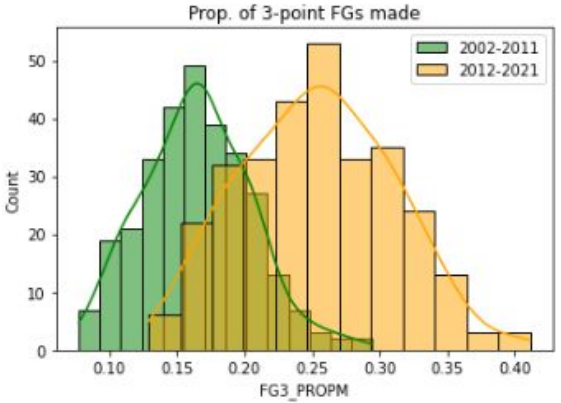
```
df = pd.read_csv("data/nba_stats_cleaned.csv")
df.head()
```

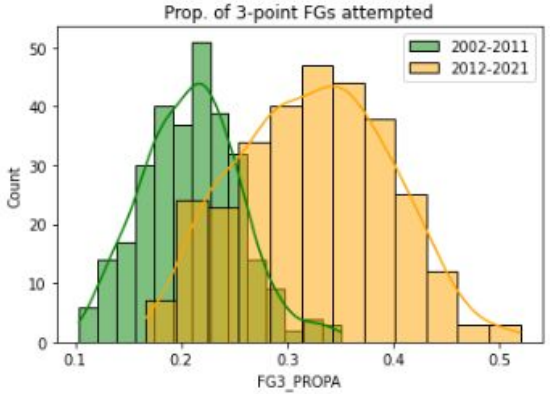
	TEAM_ID	TEAM_CITY	TEAM_NAME	YEAR	GP	WINS	LOSSES	WIN_PCT	CONF_RANK	DIV_RANK	...	FGA	FG_PCT	FG3M	FG3A	FG3_PCT	FTM
0	1610612737	Atlanta	Hawks	2000	82	25	57	0.305	13	7	...	6668	0.431	333	933	0.357	1374
1	1610612737	Atlanta	Hawks	2001	82	33	49	0.402	12	6	...	6610	0.439	423	1194	0.354	1486
2	1610612737	Atlanta	Hawks	2002	82	35	47	0.427	11	5	...	6434	0.444	402	1141	0.352	1594
3	1610612737	Atlanta	Hawks	2003	82	28	54	0.341	12	7	...	6529	0.433	419	1249	0.335	1534
4	1610612737	Atlanta	Hawks	2004	82	13	69	0.159	15	5	...	6672	0.441	304	973	0.312	1417

Exploratory Data Analysis

Proportion of 3 Points attempted/made is higher in the last 10 years compared to the 10 years before that

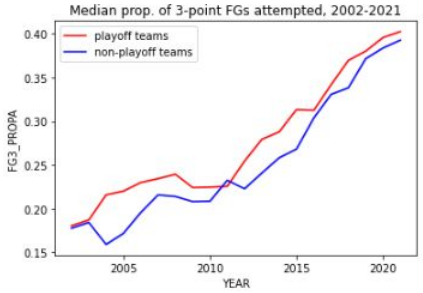
Distribution is roughly bimodal so we will consider each decade separately

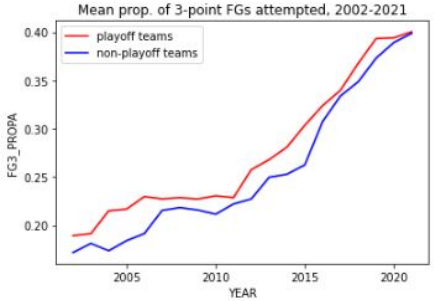


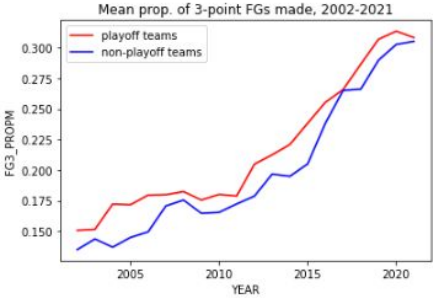


Exploratory Data Analysis

Mean/Median Proportion of 3 Pointers attempted/made has risen for both playoff and non-playoff teams, with playoff teams having a higher proportion on average

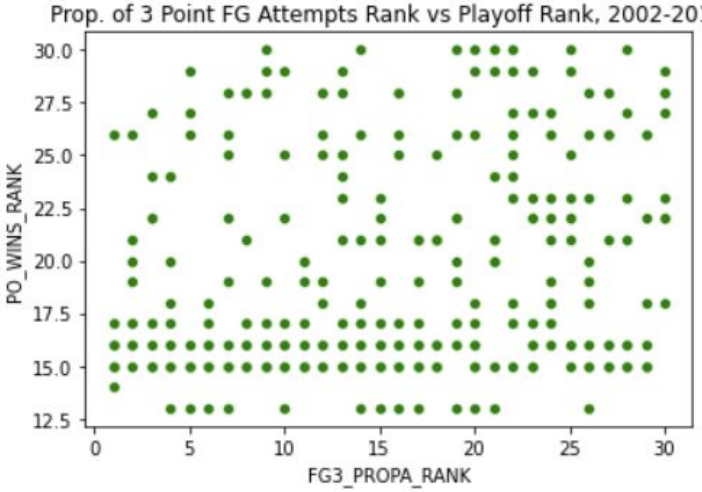


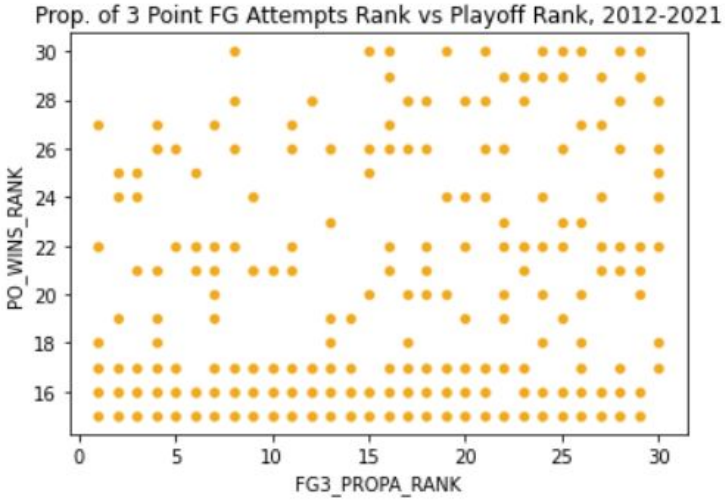


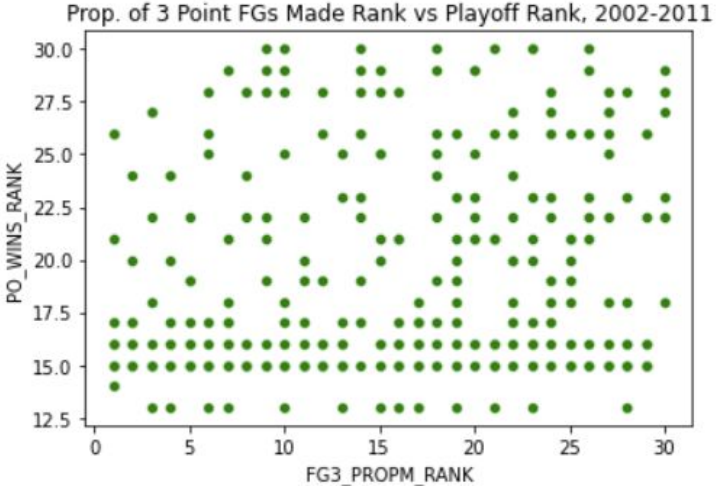


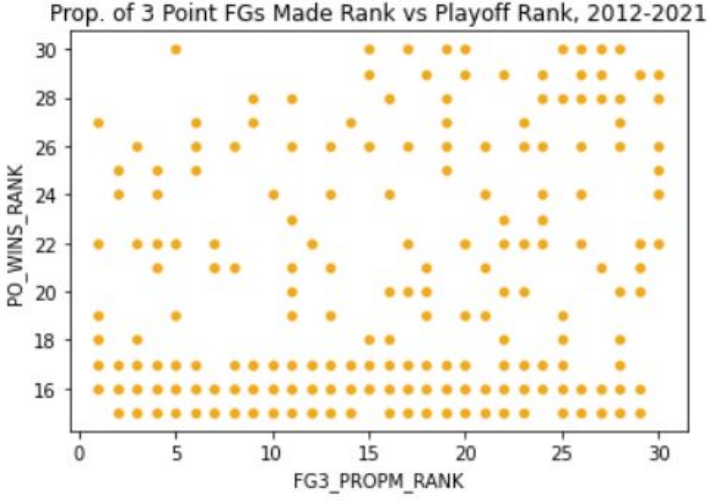
Exploratory Data Analysis

- Quantifying Playoff Success
  - # of playoff wins is NOT normal since playoffs are bracket based and many teams have 0 wins
  - Rank teams based on number of playoff wins (ties ok)
  - Compare with ranking of 3 point prop.


Prop. of 3 Point FG Attempts Rank vs Playoff Rank, 2002-2011
PO_WINS_RANK
FG3_PROPA_RANK


Prop. of 3 Point FG Attempts Rank vs Playoff Rank, 2012-2021
PO_WINS_RANK
FG3_PROPA_RANK


Prop. of 3 Point FGs Made Rank vs Playoff Rank, 2002-2011
PO_WINS_RANK
FG3_PROPM_RANK


Prop. of 3 Point FGs Made Rank vs Playoff Rank, 2012-2021
PO_WINS_RANK
FG3_PROPM_RANK

Data Analysis and Results (2002-2011)

Prop. of 3-point field goals attempted ~ Playoff win rank (2002-2011)

```
from scipy.stats import kendalltau
tau, p_val = kendalltau(df_02_11["FG3_PROPA_RANK"], df_02_11["PO_WINS_RANK"])
print("tau = " + str(tau))
print("p = " + str(p_val))
```

```
tau = 0.15280982479311364
p = 0.00022113078826839074
```

Prop. of 3-point field goals made ~ Playoff win rank (2002-2011)

```
tau, p_val = kendalltau(df_02_11["FG3_PROPM_RANK"], df_02_11["WINS_RANK"])
print("tau = " + str(tau))
print("p = " + str(p_val))
```

```
tau = 0.1859912438598741
p = 3.367832085856858e-06
```

Data Analysis and Results (2012-2021)

Prop. of 3-point field goals attempted ~ Playoff win rank (2012-2021)

```
tau_recent, p_recent = kendalltau(df_12_21["FG3_PROPA_RANK"], df_12_21["PO_WINS_RANK"])
print("tau = " + str(tau_recent))
print("p = " + str(p_recent))
```

```
tau = 0.1574109436143241
p = 0.00014251775870618988
```

Prop. of 3-point field goals made ~ Playoff win rank (2012-2021)

```
tau_recent, p_recent = kendalltau(df_12_21["FG3_PROPM_RANK"], df_12_21["PO_WINS_RANK"])
print("tau = " + str(tau_recent))
print("p = " + str(p_recent))
```

```
tau = 0.16910226178476012
p = 4.3835926709159635e-05
```

Conclusion/Discussion

- There is a *slight* positive correlation between the proportion of the 3 point field goals (made / attempted) and playoff success based on the Kendall-Tau Test
  - p-value was below 0.05 and our tau value (correlation coefficient) was ~0.15
- This does NOT imply that the strategy is the reason for this success
  - That would be causation

Future Work

- 3-Point Net Gain (3NG)
  - + (make) OR - (miss)
- Can be applied to any NBA team/player
- Possible data science topics with this metric
  - Career Averages/Totals
  - Top Teams in NBA History/Certain Decade
  - Best Seasons All Time

Appendix

- NBA Statistics
- NBA API
- Pandera
- 3-Point Net Gain
  - Points Per Possession (PPP)
- GitHub Repository

THANK YOU!