

basketball

March 12, 2024

1 Team Project (Title tbd)

1.1 Ideas:

- Does regular season win record impact playoff performance?
 - Is a winning record enough to do well in the postseason?
- Can a player actually “carry” a team?
 -
-

```
[24]: import pandas as pd
import plotly as px
import numpy as np
from sklearn.linear_model import LinearRegression
import matplotlib.pyplot as plt
import plotly.express as pxo
```

```
[25]: per = pd.read_csv("data/per.csv", index_col="Player")
ppg = pd.read_csv("data/ppg.csv")
ppg.drop(columns=["Player-additional", "Rk"], inplace=True)
per = per.rename(columns={"PER": "PER"})
per.drop(columns="Player-additional", inplace=True)
per = per.merge(ppg, left_on="Player", right_on="Player", how="inner")
per = per[per["Tm"] != "TOT"]
per
```

```
[25]:
```

| | Player | Rk | Pos | Age | Tm | PER | 3PAr | PTS |
|-----|-----------------------|-----|-----|-----|-----|------|-------|------|
| 0 | Nikola Jokić | 1 | C | 27 | DEN | 31.5 | 0.146 | 24.5 |
| 1 | Joel Embiid | 2 | C | 28 | PHI | 31.4 | 0.151 | 33.1 |
| 2 | Giannis Antetokounmpo | 3 | PF | 28 | MIL | 29.0 | 0.134 | 31.1 |
| 3 | Luka Dončić | 4 | PG | 23 | DAL | 28.7 | 0.373 | 32.4 |
| 4 | Anthony Davis | 5 | C | 29 | LAL | 27.8 | 0.077 | 25.9 |
| .. | ... | ... | ... | ... | ... | ... | ... | ... |
| 217 | Dillon Brooks | 184 | SF | 27 | MEM | 9.4 | 0.443 | 14.3 |
| 221 | Royce O'Neale | 186 | SF | 29 | BRK | 9.2 | 0.707 | 8.8 |
| 222 | Gabe Vincent | 187 | PG | 26 | MIA | 9.0 | 0.617 | 9.4 |
| 226 | Reggie Bullock | 189 | SF | 31 | DAL | 7.7 | 0.845 | 7.2 |
| 227 | P.J. Tucker | 190 | PF | 37 | PHI | 5.7 | 0.622 | 3.5 |

[171 rows x 8 columns]

```
[26]: three_pt_atmpts = pd.read_csv("data/3pa.csv", index_col="Team")
three_pt_atmpts.rename(index=lambda s: s[:-1] if s.endswith("*") else s,
                        inplace=True)
three_pt_atmpts
```

```
[26]:
```

| | Rk | G | FG | FGA | FG% | 3P | 3PA | 3P% | 2P | \ |
|------------------------|------|----|------|------|-------|------|------|-------|------|---|
| Team | | | | | | | | | | |
| Sacramento Kings | 1.0 | 82 | 43.6 | 88.2 | 0.494 | 13.8 | 37.3 | 0.369 | 29.8 | |
| Golden State Warriors | 2.0 | 82 | 43.1 | 90.2 | 0.479 | 16.6 | 43.2 | 0.385 | 26.5 | |
| Atlanta Hawks | 3.0 | 82 | 44.6 | 92.4 | 0.483 | 10.8 | 30.5 | 0.352 | 33.9 | |
| Boston Celtics | 4.0 | 82 | 42.2 | 88.8 | 0.475 | 16.0 | 42.6 | 0.377 | 26.2 | |
| Oklahoma City Thunder | 5.0 | 82 | 43.1 | 92.6 | 0.465 | 12.1 | 34.1 | 0.356 | 31.0 | |
| Los Angeles Lakers | 6.0 | 82 | 42.9 | 89.0 | 0.482 | 10.8 | 31.2 | 0.346 | 32.1 | |
| Utah Jazz | 7.0 | 82 | 42.5 | 89.8 | 0.473 | 13.3 | 37.8 | 0.353 | 29.2 | |
| Memphis Grizzlies | 8.0 | 82 | 43.7 | 92.1 | 0.475 | 12.0 | 34.2 | 0.351 | 31.7 | |
| Milwaukee Bucks | 9.0 | 82 | 42.7 | 90.4 | 0.473 | 14.8 | 40.3 | 0.368 | 27.9 | |
| Indiana Pacers | 10.0 | 82 | 42.0 | 89.6 | 0.469 | 13.6 | 37.0 | 0.367 | 28.4 | |
| New York Knicks | 11.0 | 82 | 42.0 | 89.4 | 0.470 | 12.6 | 35.7 | 0.354 | 29.4 | |
| Denver Nuggets | 12.0 | 82 | 43.6 | 86.4 | 0.504 | 11.8 | 31.2 | 0.379 | 31.8 | |
| Minnesota Timberwolves | 13.0 | 82 | 42.9 | 87.4 | 0.490 | 12.2 | 33.3 | 0.365 | 30.7 | |
| Philadelphia 76ers | 14.0 | 82 | 40.8 | 83.8 | 0.487 | 12.6 | 32.6 | 0.387 | 28.2 | |
| New Orleans Pelicans | 15.0 | 82 | 42.0 | 87.6 | 0.480 | 11.0 | 30.1 | 0.364 | 31.1 | |
| Dallas Mavericks | 16.0 | 82 | 40.0 | 84.3 | 0.475 | 15.2 | 41.0 | 0.371 | 24.8 | |
| Phoenix Suns | 17.0 | 82 | 42.1 | 90.1 | 0.467 | 12.2 | 32.6 | 0.374 | 29.9 | |
| Los Angeles Clippers | 18.0 | 82 | 41.1 | 86.1 | 0.477 | 12.7 | 33.4 | 0.381 | 28.4 | |
| Portland Trail Blazers | 19.0 | 82 | 40.5 | 85.4 | 0.474 | 12.9 | 35.3 | 0.365 | 27.6 | |
| Brooklyn Nets | 20.0 | 82 | 41.5 | 85.1 | 0.487 | 12.8 | 33.8 | 0.378 | 28.7 | |
| Washington Wizards | 21.0 | 82 | 42.1 | 86.9 | 0.485 | 11.3 | 31.7 | 0.356 | 30.9 | |
| Chicago Bulls | 22.0 | 82 | 42.5 | 86.8 | 0.490 | 10.4 | 28.9 | 0.361 | 32.1 | |
| San Antonio Spurs | 23.0 | 82 | 43.1 | 92.6 | 0.465 | 11.1 | 32.2 | 0.345 | 32.0 | |
| Toronto Raptors | 24.0 | 82 | 41.9 | 91.3 | 0.459 | 10.7 | 32.0 | 0.335 | 31.1 | |
| Cleveland Cavaliers | 25.0 | 82 | 41.6 | 85.2 | 0.488 | 11.6 | 31.6 | 0.367 | 30.0 | |
| Orlando Magic | 26.0 | 82 | 40.5 | 86.3 | 0.470 | 10.8 | 31.1 | 0.346 | 29.8 | |
| Charlotte Hornets | 27.0 | 82 | 41.3 | 90.4 | 0.457 | 10.7 | 32.5 | 0.330 | 30.5 | |
| Houston Rockets | 28.0 | 82 | 40.6 | 88.9 | 0.457 | 10.4 | 31.9 | 0.327 | 30.2 | |
| Detroit Pistons | 29.0 | 82 | 39.6 | 87.1 | 0.454 | 11.4 | 32.4 | 0.351 | 28.2 | |
| Miami Heat | 30.0 | 82 | 39.2 | 85.3 | 0.460 | 12.0 | 34.8 | 0.344 | 27.3 | |
| League Average | NaN | 82 | 42.0 | 88.3 | 0.475 | 12.3 | 34.2 | 0.361 | 29.6 | |

| | 2PA | 2P% | FT | FTA | FT% | AST | TOV | PTS |
|-----------------------|------|-------|------|------|-------|------|------|-------|
| Team | | | | | | | | |
| Sacramento Kings | 50.9 | 0.586 | 19.8 | 25.1 | 0.790 | 27.3 | 13.5 | 120.7 |
| Golden State Warriors | 47.0 | 0.564 | 16.0 | 20.2 | 0.794 | 29.8 | 16.3 | 118.9 |
| Atlanta Hawks | 61.8 | 0.548 | 18.5 | 22.6 | 0.818 | 25.0 | 12.9 | 118.4 |

| | | | | | | | | |
|------------------------|------|-------|------|------|-------|------|------|-------|
| Boston Celtics | 46.2 | 0.567 | 17.5 | 21.6 | 0.812 | 26.7 | 13.4 | 117.9 |
| Oklahoma City Thunder | 58.5 | 0.530 | 19.2 | 23.7 | 0.809 | 24.4 | 13.0 | 117.5 |
| Los Angeles Lakers | 57.8 | 0.555 | 20.6 | 26.6 | 0.775 | 25.3 | 14.1 | 117.2 |
| Utah Jazz | 52.0 | 0.560 | 18.7 | 23.8 | 0.786 | 26.0 | 15.4 | 117.1 |
| Memphis Grizzlies | 57.9 | 0.548 | 17.5 | 23.8 | 0.733 | 26.0 | 13.6 | 116.9 |
| Milwaukee Bucks | 50.1 | 0.557 | 16.6 | 22.4 | 0.743 | 25.8 | 14.6 | 116.9 |
| Indiana Pacers | 52.6 | 0.540 | 18.7 | 23.7 | 0.790 | 27.0 | 14.9 | 116.3 |
| New York Knicks | 53.6 | 0.547 | 19.4 | 25.5 | 0.761 | 22.9 | 13.0 | 116.0 |
| Denver Nuggets | 55.2 | 0.575 | 16.8 | 22.4 | 0.751 | 28.9 | 14.5 | 115.8 |
| Minnesota Timberwolves | 54.1 | 0.568 | 17.9 | 23.7 | 0.755 | 26.2 | 15.3 | 115.8 |
| Philadelphia 76ers | 51.2 | 0.551 | 21.0 | 25.1 | 0.835 | 25.2 | 13.7 | 115.2 |
| New Orleans Pelicans | 57.5 | 0.541 | 19.3 | 24.4 | 0.793 | 25.9 | 14.6 | 114.4 |
| Dallas Mavericks | 43.3 | 0.574 | 19.0 | 25.1 | 0.755 | 22.9 | 12.2 | 114.2 |
| Phoenix Suns | 57.5 | 0.520 | 17.2 | 21.7 | 0.793 | 27.3 | 13.5 | 113.6 |
| Los Angeles Clippers | 52.7 | 0.539 | 18.7 | 23.9 | 0.781 | 23.9 | 14.2 | 113.6 |
| Portland Trail Blazers | 50.1 | 0.550 | 19.6 | 24.6 | 0.796 | 24.2 | 14.5 | 113.4 |
| Brooklyn Nets | 51.3 | 0.559 | 17.7 | 22.1 | 0.800 | 25.5 | 13.7 | 113.4 |
| Washington Wizards | 55.2 | 0.559 | 17.6 | 22.4 | 0.785 | 25.4 | 14.1 | 113.2 |
| Chicago Bulls | 57.9 | 0.555 | 17.6 | 21.8 | 0.809 | 24.5 | 13.4 | 113.1 |
| San Antonio Spurs | 60.4 | 0.529 | 15.8 | 21.2 | 0.743 | 27.2 | 15.3 | 113.0 |
| Toronto Raptors | 59.3 | 0.525 | 18.4 | 23.4 | 0.784 | 23.9 | 11.7 | 112.9 |
| Cleveland Cavaliers | 53.6 | 0.559 | 17.5 | 22.5 | 0.780 | 24.9 | 13.3 | 112.3 |
| Orlando Magic | 55.2 | 0.539 | 19.6 | 25.0 | 0.784 | 23.2 | 15.1 | 111.4 |
| Charlotte Hornets | 57.9 | 0.528 | 17.6 | 23.6 | 0.749 | 25.1 | 14.2 | 111.0 |
| Houston Rockets | 56.9 | 0.530 | 19.1 | 25.3 | 0.754 | 22.4 | 16.2 | 110.7 |
| Detroit Pistons | 54.6 | 0.516 | 19.8 | 25.7 | 0.771 | 23.0 | 15.1 | 110.3 |
| Miami Heat | 50.5 | 0.540 | 19.1 | 23.0 | 0.831 | 23.8 | 13.5 | 109.5 |
| League Average | 54.1 | 0.548 | 18.4 | 23.5 | 0.782 | 25.3 | 14.1 | 114.7 |

```
[27]: abbreviations = pd.read_csv("data/abbreviations.csv")
teams = pd.read_csv("data/teams.csv")

# Rename index and columns to preferred names
abbreviations.rename(columns={"prefix_1": "abbrev"}, inplace=True)

# Regularize abbreviations
abbreviations["abbrev"] = abbreviations["abbrev"].apply(lambda s: s.upper())

# Drop unneeded column
abbreviations.drop(columns="prefix_2", inplace=True)
abbreviations.loc[abbreviations["abbrev"] == "NO", "name"] = "New Orleans_
↳Pelicans"
abbreviations.loc[abbreviations["abbrev"] == "CHA", "name"] = "Charlotte_
↳Hornets"

teams = teams.merge(abbreviations, left_on="Team", right_on="name", how="outer")
teams.drop(columns=["name", "Rk"])
```

```
[27]:
```

| | Team | Overall | abbrev |
|----|------------------------|---------|--------|
| 0 | Atlanta Hawks | 41-41 | ATL |
| 1 | Boston Celtics | 57-25 | BOS |
| 2 | Brooklyn Nets | 45-37 | BKN |
| 3 | Charlotte Hornets | 27-55 | CHA |
| 4 | Chicago Bulls | 40-42 | CHI |
| 5 | Cleveland Cavaliers | 51-31 | CLE |
| 6 | Dallas Mavericks | 38-44 | DAL |
| 7 | Denver Nuggets | 53-29 | DEN |
| 8 | Detroit Pistons | 17-65 | DET |
| 9 | Golden State Warriors | 44-38 | GS |
| 10 | Houston Rockets | 22-60 | HOU |
| 11 | Indiana Pacers | 35-47 | IND |
| 12 | Los Angeles Clippers | 44-38 | LAC |
| 13 | Los Angeles Lakers | 43-39 | LAL |
| 14 | Memphis Grizzlies | 51-31 | MEM |
| 15 | Miami Heat | 44-38 | MIA |
| 16 | Milwaukee Bucks | 58-24 | MIL |
| 17 | Minnesota Timberwolves | 42-40 | MIN |
| 18 | New Orleans Pelicans | 42-40 | NO |
| 19 | New York Knicks | 47-35 | NY |
| 20 | Oklahoma City Thunder | 40-42 | OKC |
| 21 | Orlando Magic | 34-48 | ORL |
| 22 | Philadelphia 76ers | 54-28 | PHI |
| 23 | Phoenix Suns | 45-37 | PHX |
| 24 | Portland Trail Blazers | 33-49 | POR |
| 25 | Sacramento Kings | 48-34 | SAC |
| 26 | San Antonio Spurs | 22-60 | SA |
| 27 | Toronto Raptors | 41-41 | TOR |
| 28 | Utah Jazz | 37-45 | UTA |
| 29 | Washington Wizards | 35-47 | WSH |

1.2 Question 1: How does player efficiency rating (PER) correlate with a team's win-loss record?

```
[28]: model = LinearRegression()
# display(per)

teams[["wins", "losses"]] = teams["Overall"].str.split("-", expand=True).
    .astype(int)
# teams["W/L Ratio"] = round((teams["wins"] / (teams["losses"] +
    teams["wins"])), 2)
# WL = teams["W/L Ratio"]
#display(teams)

top_players = per.groupby("Tm")["PER"].max().reset_index()
top_players.loc[top_players["Tm"] == "CHO", "Tm"] = "CHA"
```

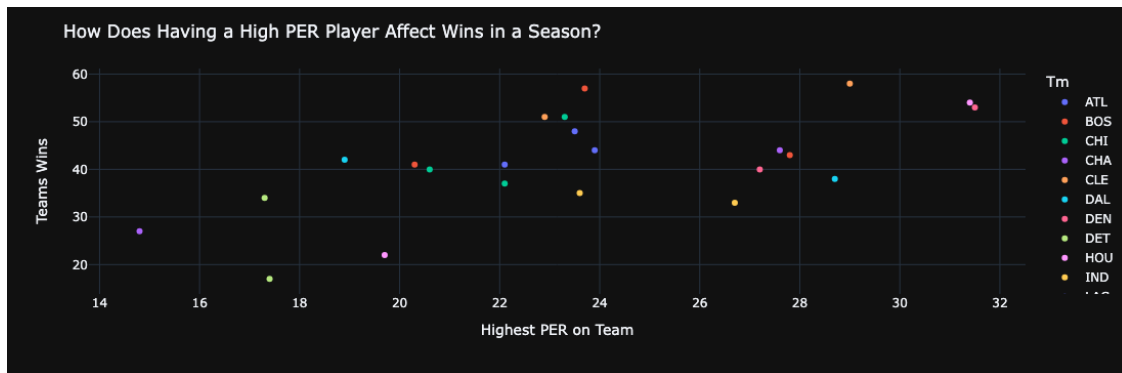
```

top_players = pd.merge(top_players, teams[["abbrev", "wins"]], left_on="Tm",
    ↳right_on="abbrev")
top_players.drop(columns="abbrev", inplace=True)

required_columns = ["Tm", "PER", "wins"]
assert all(column in top_players.columns for column in required_columns),
    ↳"Missing columns"

pxo.scatter(top_players, x = "PER", y = "wins", color="Tm").
    ↳update_layout(xaxis_title = "Highest PER on Team", yaxis_title = "Teams
    ↳Wins", title="How Does Having a High PER Player Affect Wins in a Season?")

```



1.3 Question 2: What impact does age have on PPG?

```

[38]: display(per)
required_columns = ["Player", "Rk", "Pos", "Age", "Tm", "PER", "3PAr", "PTS"]
assert all(column in per.columns for column in required_columns), "Missing
    ↳columns"

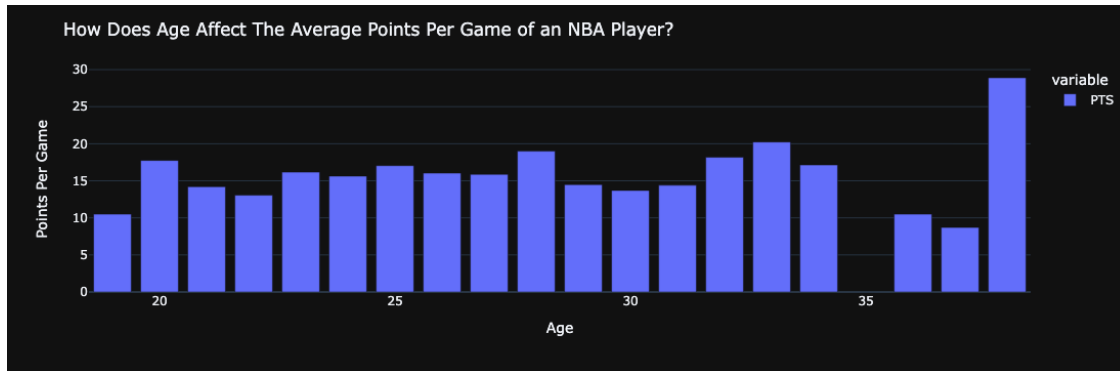
age_avgs = per.groupby("Age")["PTS"].mean()
age_avgs
pxo.bar(age_avgs).update_layout( yaxis_title = "Points Per Game", title="How
    ↳Does Age Affect The Average Points Per Game of an NBA Player?")

```

| | Player | Rk | Pos | Age | Tm | PER | 3PAr | PTS |
|-----|-----------------------|-----|-----|-----|-----|------|-------|------|
| 0 | Nikola Jokić | 1 | C | 27 | DEN | 31.5 | 0.146 | 24.5 |
| 1 | Joel Embiid | 2 | C | 28 | PHI | 31.4 | 0.151 | 33.1 |
| 2 | Giannis Antetokounmpo | 3 | PF | 28 | MIL | 29.0 | 0.134 | 31.1 |
| 3 | Luka Dončić | 4 | PG | 23 | DAL | 28.7 | 0.373 | 32.4 |
| 4 | Anthony Davis | 5 | C | 29 | LAL | 27.8 | 0.077 | 25.9 |
| .. | ... | ... | .. | ... | ... | ... | ... | ... |
| 217 | Dillon Brooks | 184 | SF | 27 | MEM | 9.4 | 0.443 | 14.3 |
| 221 | Royce O'Neale | 186 | SF | 29 | BRK | 9.2 | 0.707 | 8.8 |
| 222 | Gabe Vincent | 187 | PG | 26 | MIA | 9.0 | 0.617 | 9.4 |

| | | | | | | | | |
|-----|----------------|-----|----|----|-----|-----|-------|-----|
| 226 | Reggie Bullock | 189 | SF | 31 | DAL | 7.7 | 0.845 | 7.2 |
| 227 | P.J. Tucker | 190 | PF | 37 | PHI | 5.7 | 0.622 | 3.5 |

[171 rows x 8 columns]



1.3.1 Follow up questions:

- Could this be more indicative of the impact that ppg has on the longevity of a career?

1.4 Question 3: Does the three-point shot frequency correlate with higher team scoring averages and how does this impact Win/Loss record?

```
[41]: ThreePointShotFreq = per["3PAr"]
# display(teams)
teamPPG = three_pt_atmpts["PTS"]
wins = teams["wins"]

team_stats = three_pt_atmpts.merge(teams, left_on="Team", right_on="Team")
team_stats["3PA"] = team_stats["3PA"]

required_columns = ["Team", "Rk_x", "G", "FG", "FGA", "FG%", "3P", "3PA", "3P%", "2P",
                    "2PA", "2P%", "FT", "FTA", "FT%", "AST", "TOV", "PTS", "Rk_y",
                    "Overall", "name", "abbrev", "wins", "losses"]
assert all(column in team_stats.columns for column in required_columns), "Missing columns"

pxo.scatter(team_stats, x = "wins", y = "PTS", size = "3PA", color = "Team")
```

1.4.1 Follow up questions:

- Is this more telling of the state of the league and the emphasis there is on the three-point shot?

[30] :