The goal of our final project is to utilized player and team information and use Machine learning to predict the 2022 NBA Finals Champions. Currently, the remaining teams in the playoff are Miami Heat, Philadelphia 76ers, Boston Celtics, Milwaukee Bucks in the Eastern Conference. Phoenix Suns, Dallas Mavericks, Memphis Grizzlies and Golden State Warriors in the Western Conference. Our goal is to predict the outcome of each game withing the series. Each series consists of a potential 7-games. With the winning teams meeting up for the Easter and Western conference 7-game series. The winner of the Conference game series will face off for the NBA Conference finals another 7-games series. The winner would be the 2022 NBA Champions. Who will it be?

The following steps will outline the process:

**Identify Relevant Data** – preferably team and player data where we gather the relevant team and player statistics of the remaining team. Data we could use to identify team and player tendencies to predict the final results.

**Clean and Processing the Data** – based on the data we select we will have to clean the data for efficient outcomes. Useable data to analyze tendencies for each matchup series based team and player matchups.

**Feature Engineering** — what additional metrics can we append to our datasets that would help any user or ML model to better understand and predict, respectively, outcomes and trends from the data?

**Data Analysis** — can we determine any collinearity or other relations within the data that may better inform our predictions?

**Predictions** – The results of a 7-game series for the eight remaining teams. Outcome for each game in the series. Then predict both Easter and Western conference finals champion winners. These two teams will then compete for the final series outcome of the NBA Conference final. It will be key to determine which models and features would be most beneficial for us in developing an accurate prediction for each game in the series? Do we focus on team or aggregated individuals’ statistics and matchups?

GitHub Repository. <https://github.com/Save-the-Panda-s/FinalProject.git>

Machine Learning Model we haven’t identified yet.

Database: PgAdmin

Dashboard: NA