NBA MODELING

JT Flume

Raymond Li

Anant Gupta

Emily McCullough

CONTENTS OF THE REPORT

Problem Statement

Exploratory Data Analysis

Models

Insights

PROBLEM STATEMENT

DO NBA TEAMS HAVE A HOME COURT ADVANTAGE?

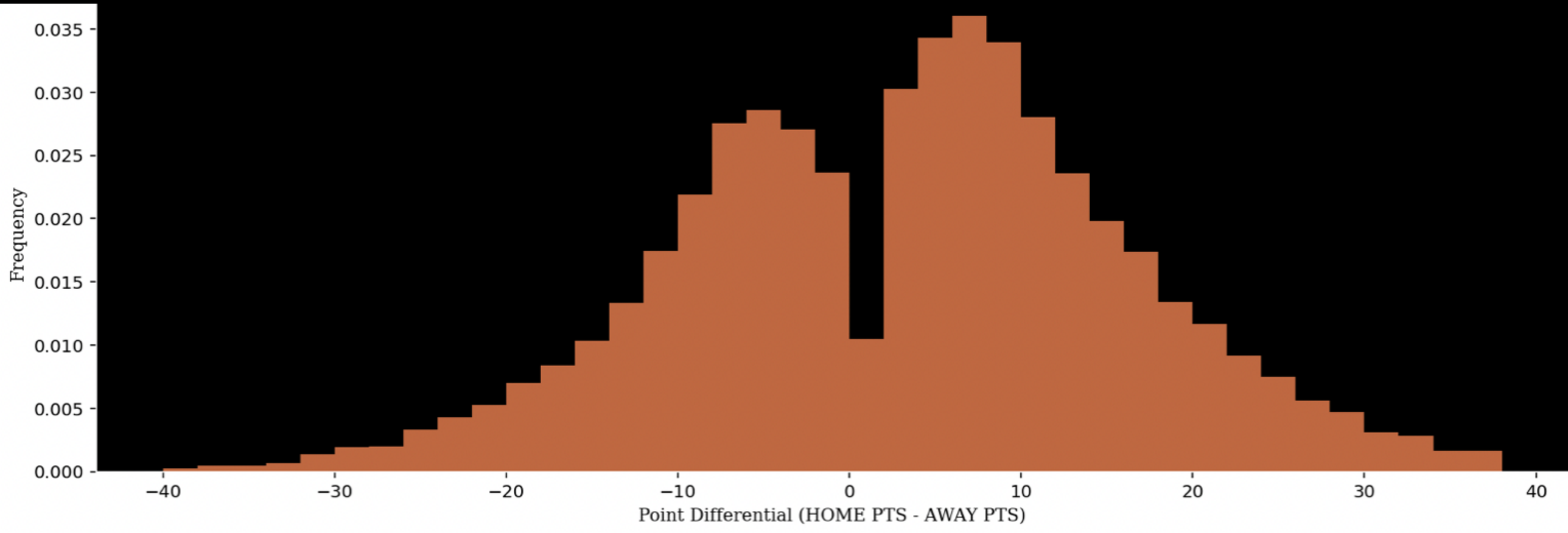
THE DATASET

KEY PERFORMANCE METRICS FROM EVERY NBA GAME SINCE 2004 TO 2020

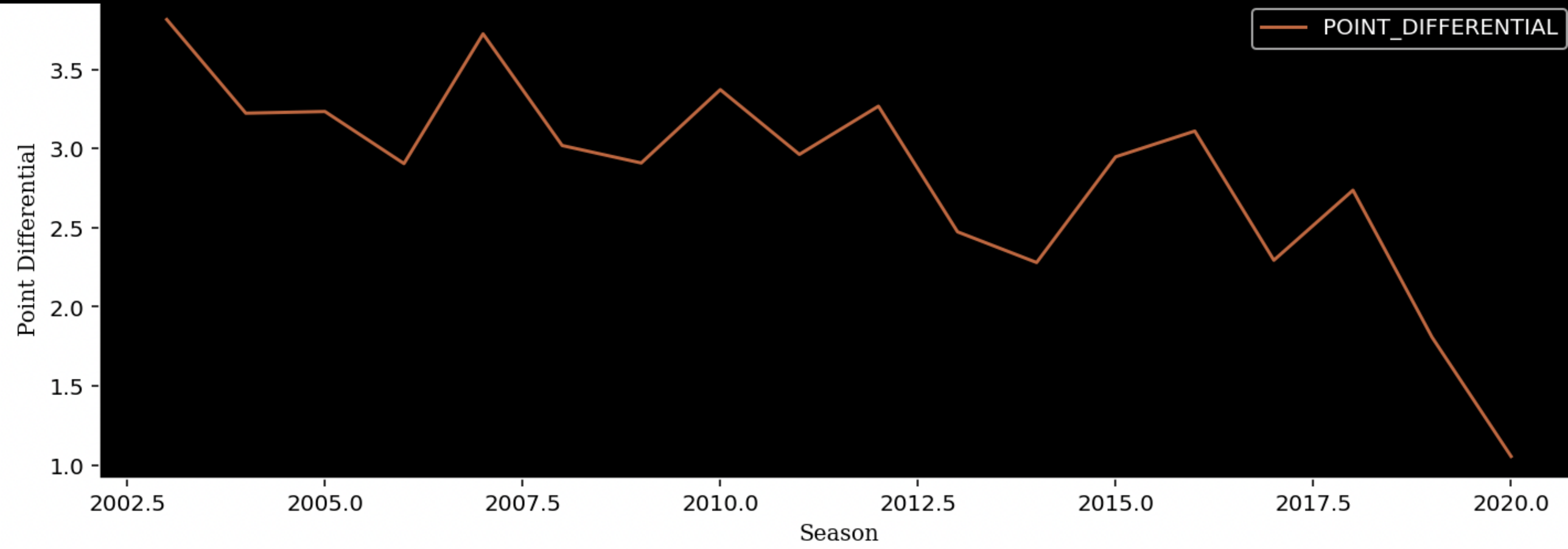
INFORMATION ON EVERY TEAM IN THE NBA

INDIVIDUAL PLAYER PERFORMANCE FOR EVERY NBA GAME SINCE 2004

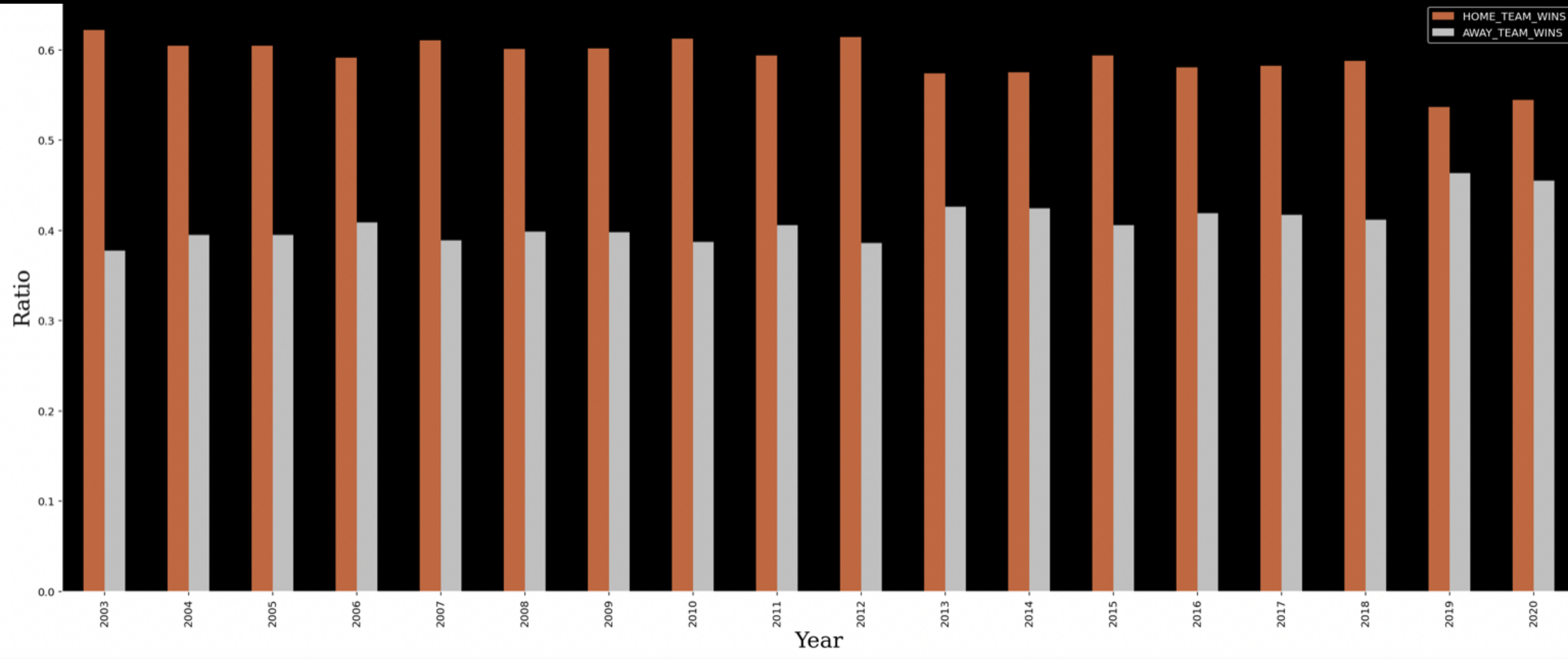
Point Differential Frequency Distribution



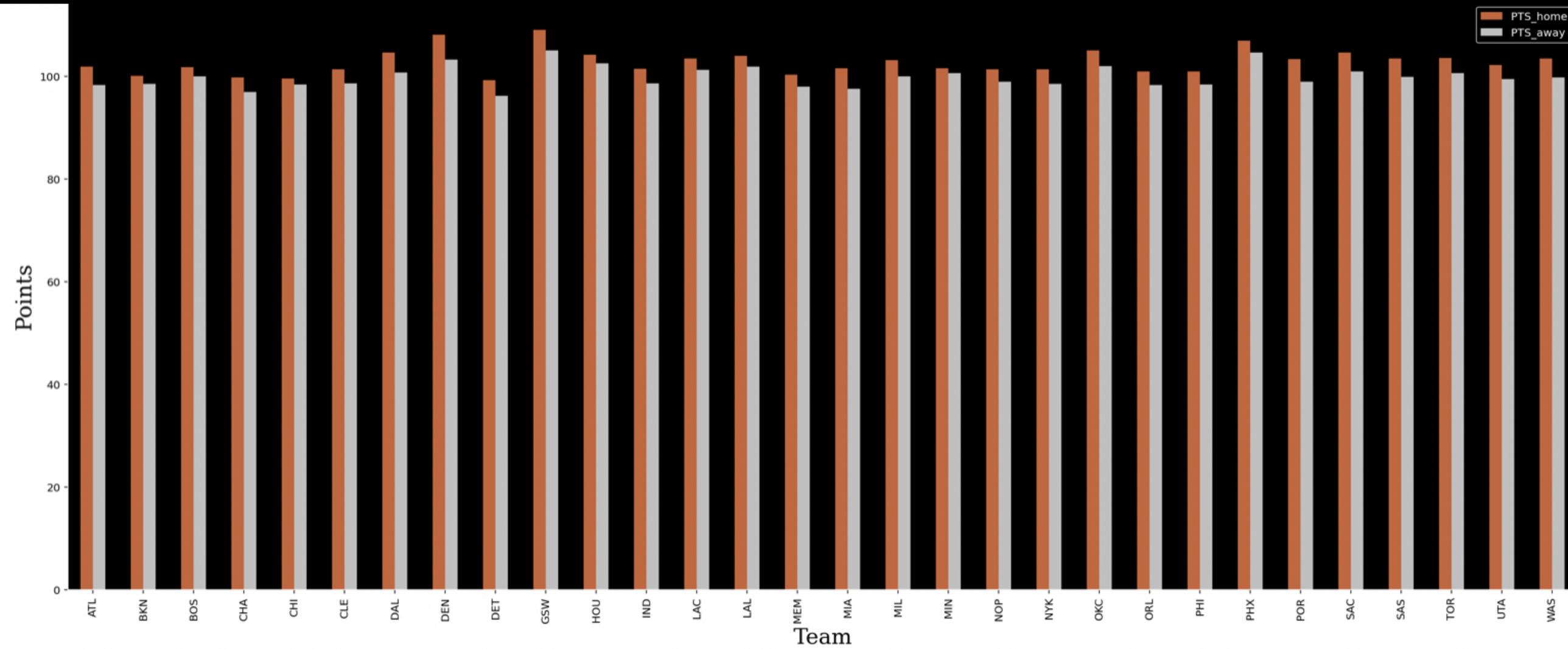
HOME TEAM POINT DIFFERENTIAL OVER SEASONS



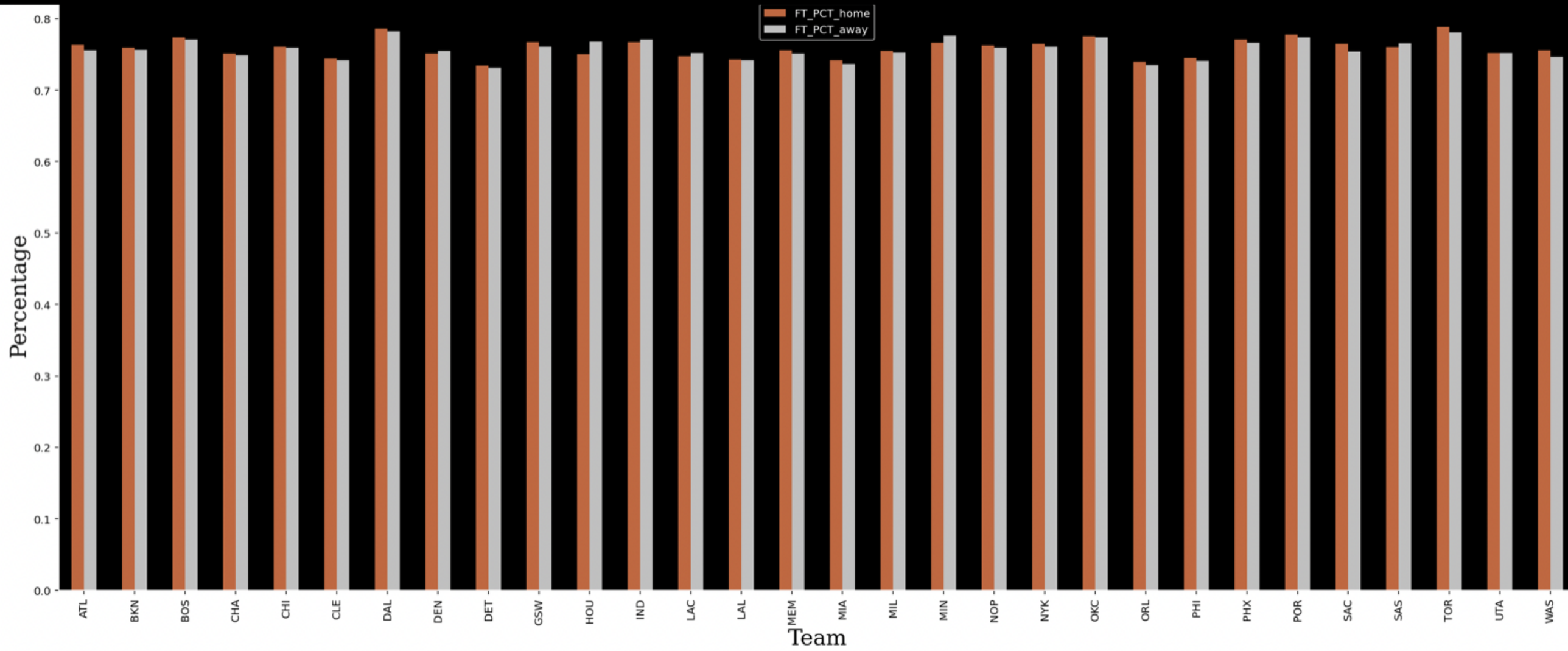
WIN RATIO - HOME VS AWAY



POINTS SCORED HOME VS AWAY

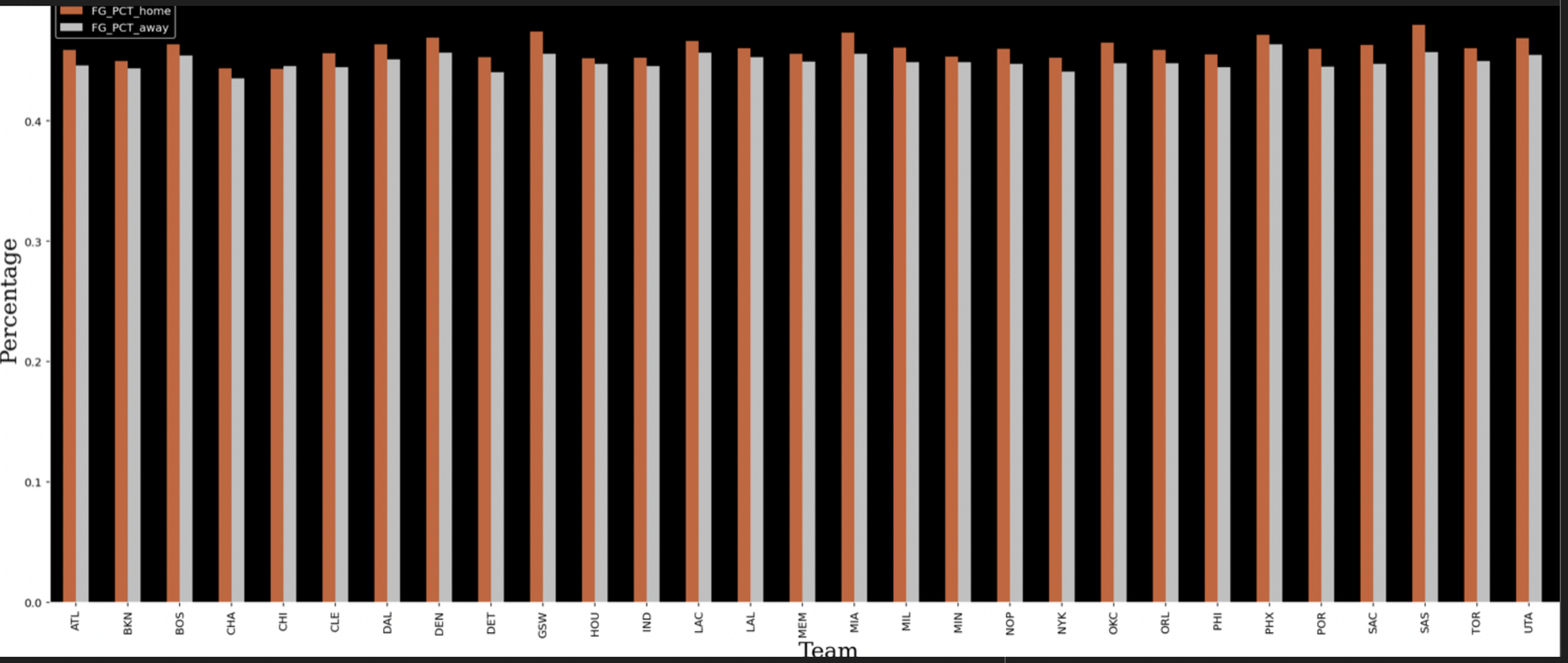


FREE THROW PERCENTAGE - HOME VS AWAY



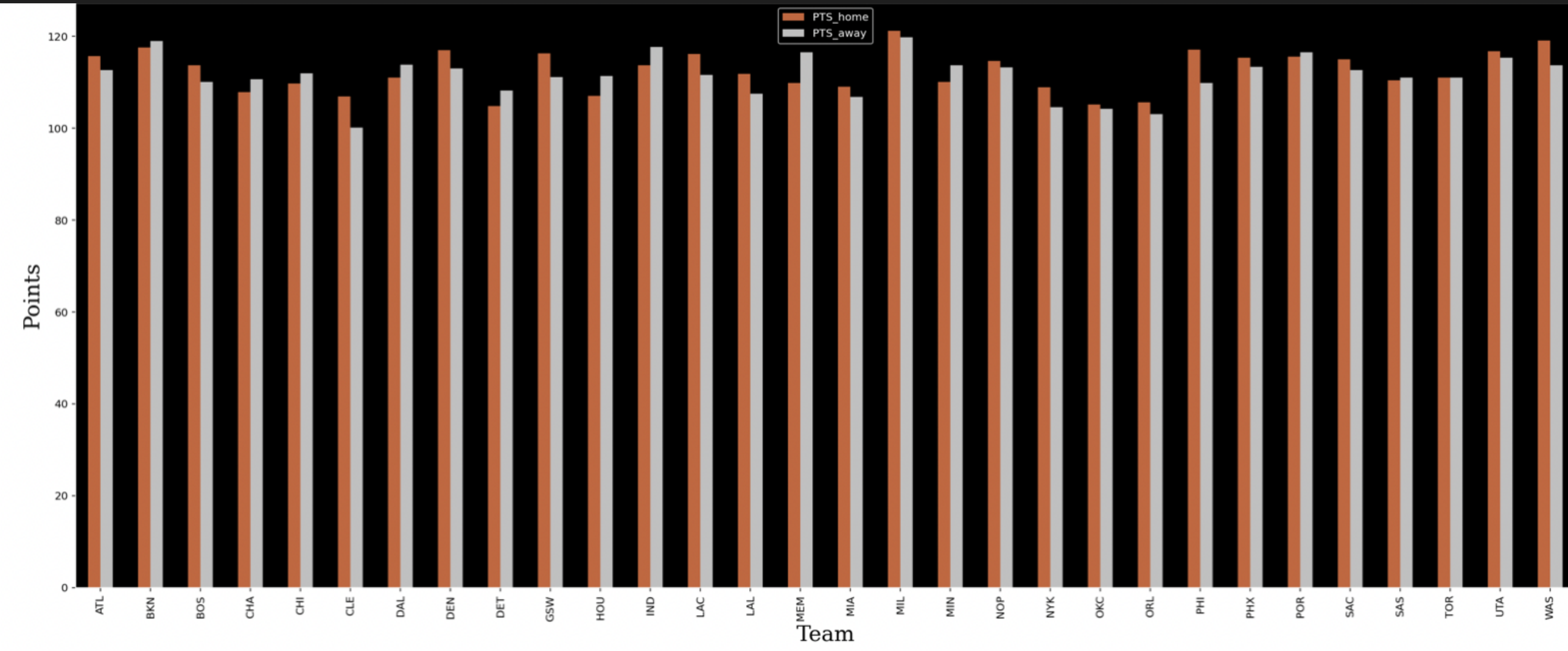
THIS TREND PERSISTS FOR EACH PERFORMANCE METRIC

THE HOME TEAM TENDS TO SCORE MORE POINTS, HAVE A HIGHER FIELD GOAL PERCENTAGE, ETC



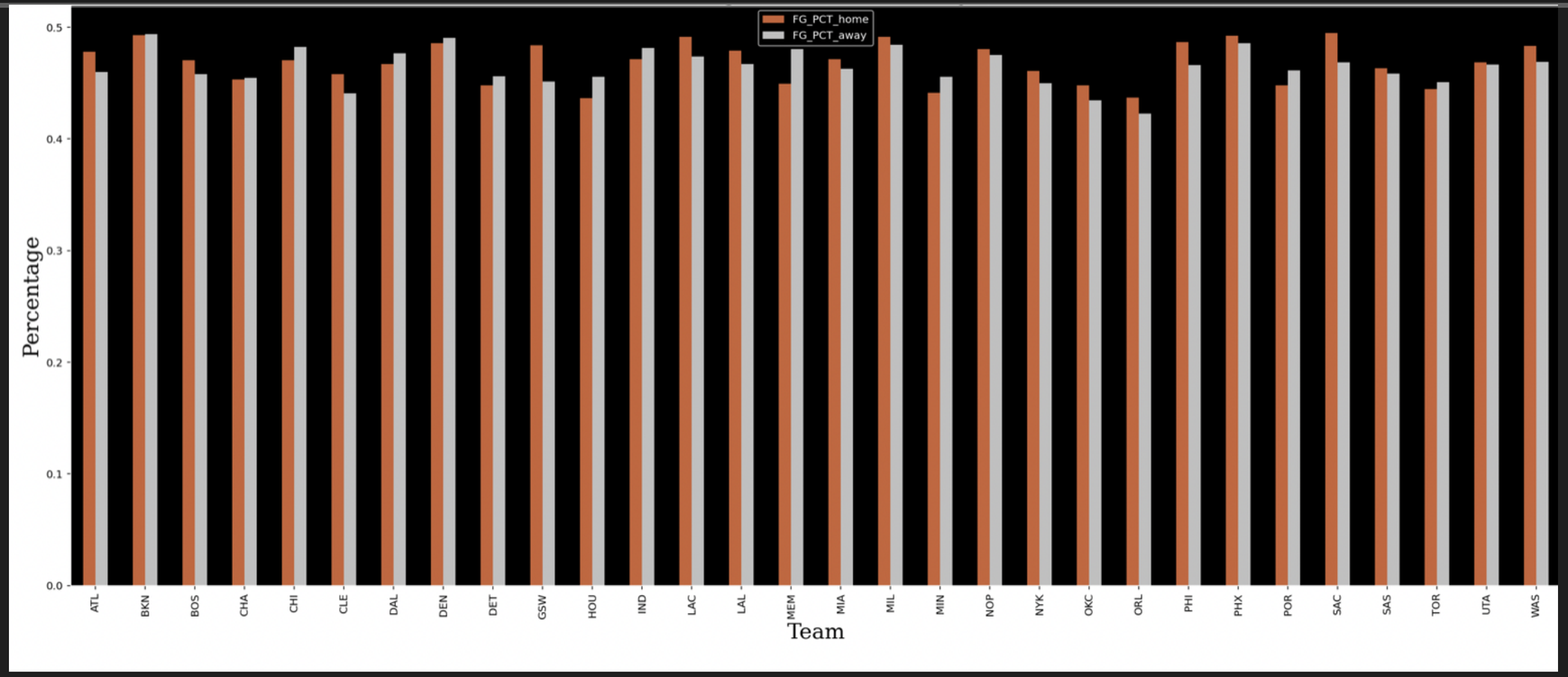
- POINTS
- FREE THROW PERCENTAGE
- ASSISTS
- REBOUNDS
- FIELD GOAL 3 PERCENTAGE

POINTS - HOME VS AWAY (2020)



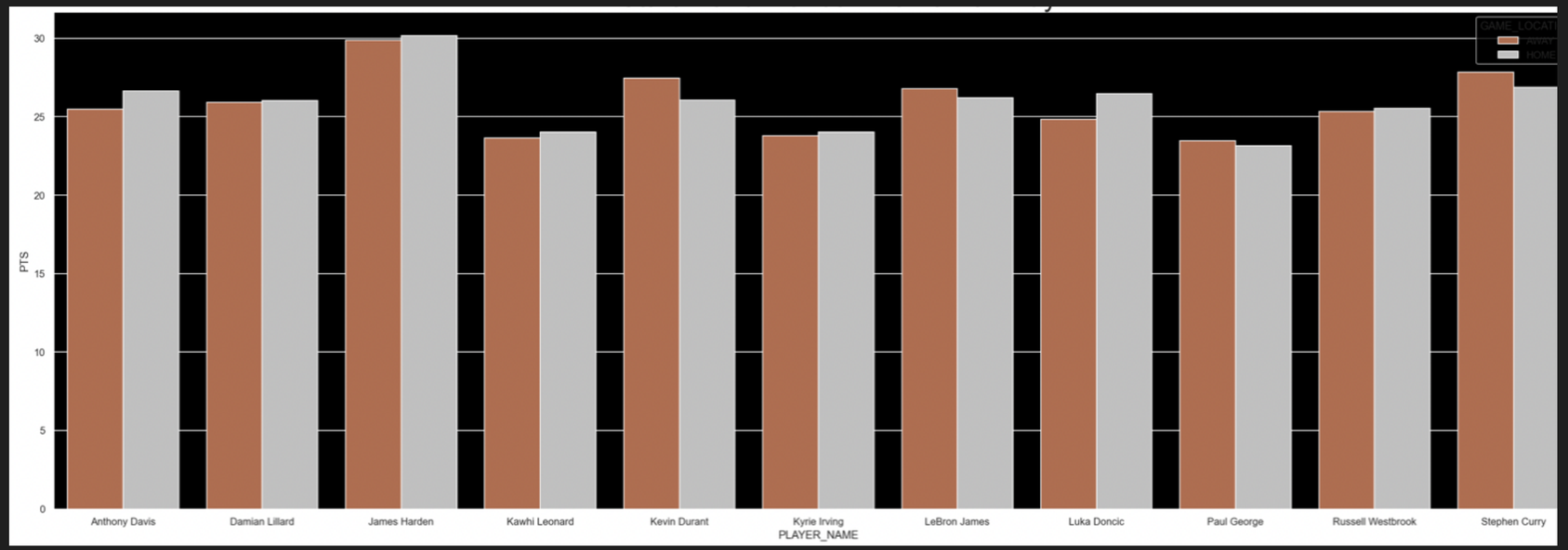
NO HOME COURT ADVANTAGE TREND IN THE 2020 SEASON

The same teams that perform better at home over all seasons, did not have a performance difference between home and away games

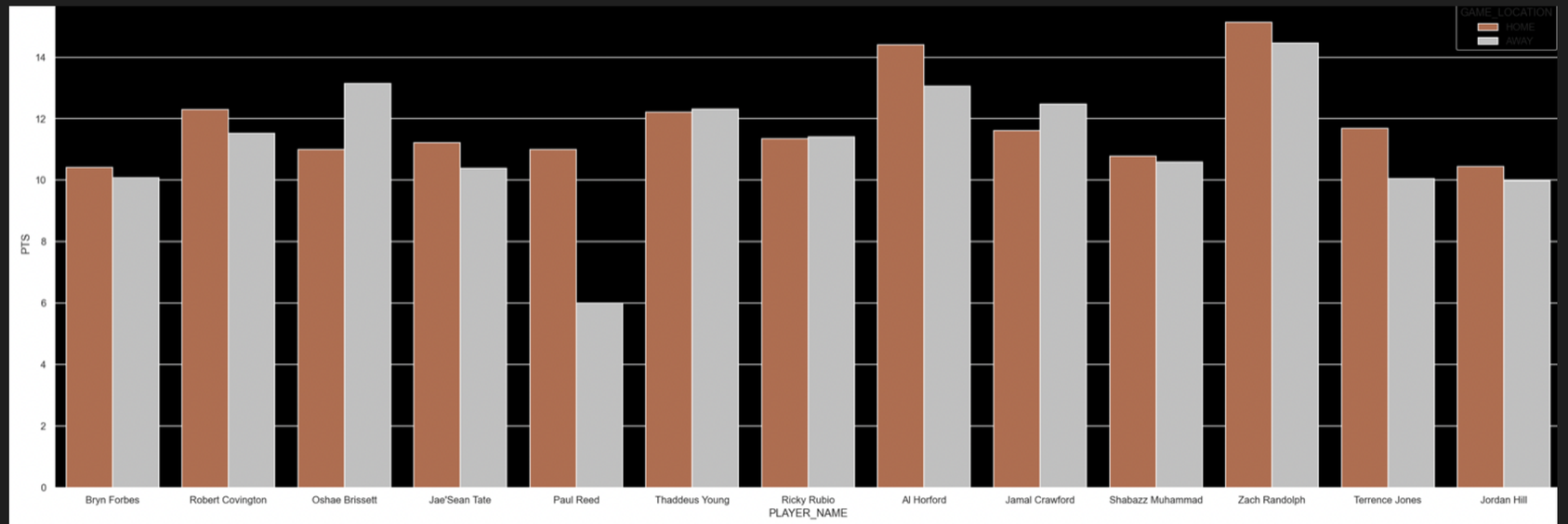


In 2020, when the COVID-19 pandemic interrupted the season, crowds were not allowed in the stadiums for games. The decline of the home court advantage trend for this season suggests that the crowd size or crowd noise level affects the team performance.

ALL STARS



ROLE PLAYERS



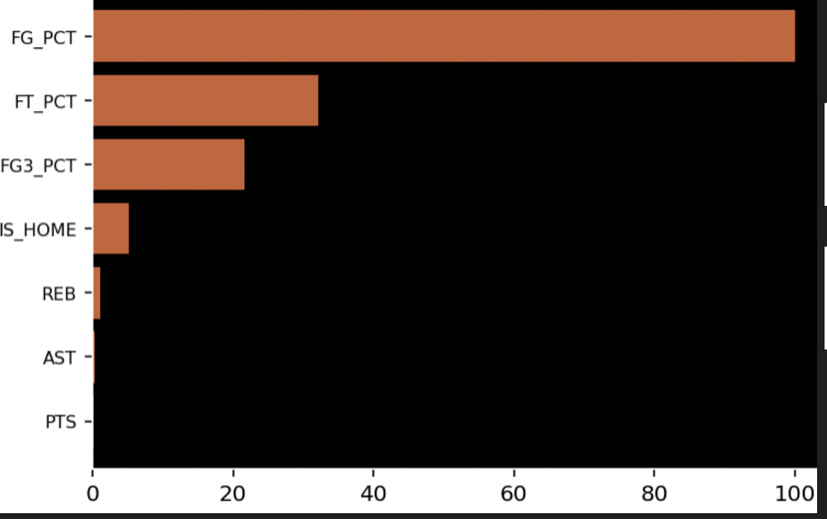
MODELS

THE PREDICTORS OF EACH MODEL INCLUDED THE POINTS, FIELD GOAL PERCENTAGE, FREE THROW PERCENTAGE, FIELD GOAL 3 PERCENTAGE, ASSISTS, AND REBOUNDS OF A TEAM, AS WELL AS A DUMMY VARIABLE TO INDICATE WHETHER THE TEAM WAS PLAYING ON THEIR HOME COURT.

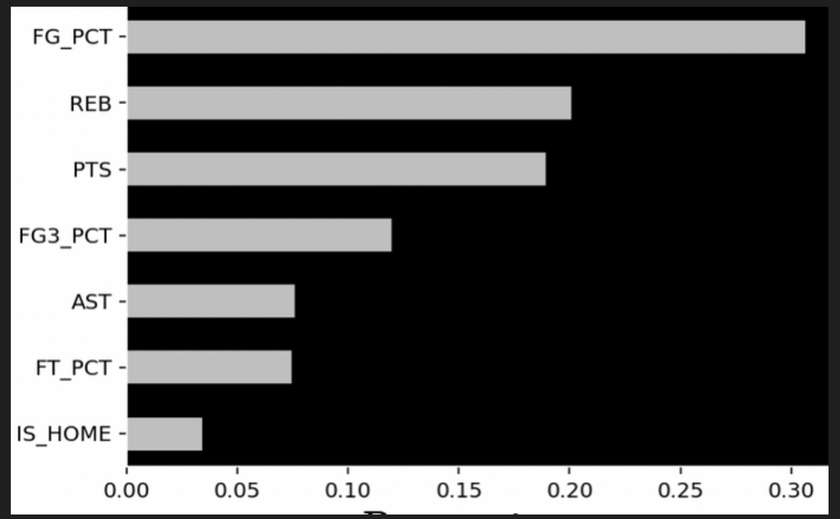
THE RESPONSE VARIABLE IS WHETHER THE TEAM WON THE GAME.
.

	LOGISTIC REGRESSION	RANDOM FOREST	BOOSTING
TRAINING ACCURACY	76.38%	82.87%	78.56%
TESTING ACCURACY	76.30%	75.91%	76.34%
MOST IMPORTANT FEATURE	FIELD GOAL PERCENTAGE	FIELD GOAL PERCENTAGE	FIELD GOAL PERCENTAGE

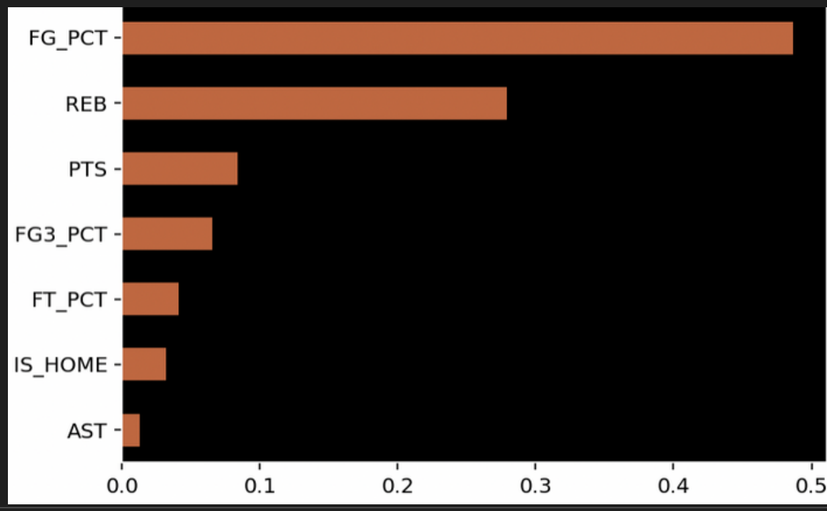
LOGISTIC REGRESSION



RANDOM FOREST



GRADIENT BOOSTING



INSIGHTS

FIELD GOAL PERCENTAGE

All three models rated the location flag low in feature importance. However, field goal percentage was the most important feature, indicating a team's field goal percentage is a good predictor of winning a game.

ALL STAR PLAYERS

All star players tend to perform about the same in both home and away games. Adding a count for all star players on a team does not increase prediction accuracy. This isuggests that if home court advantage exists, it affects those players with less skill and experience.

FURTHER ANALYSIS

Due to our exploratory data analysis, it is clear that teams perform better in home games. The lack of correlation to the location flag indicates that further analysis is needed with attributes such as crowd size and crowd noise level

THANK YOU!

QUESTIONS?