Making The Final Four

Presented by Linh Pham
Last Updated: 03/23/23

TABLE OF CONTENTS

- Objective

- Statistics Comparison

- Model Building: Metrics & Inference

- Questions and Discussion

OBJECTIVE

- **Statistics Comparison**
  - Adjusted Tempo, Offensive and Defensive Ability
  - Two Point and Three Point Shot Percentage
  - Power Rating
  - Adjusted Offensive & Defensive Efficiency
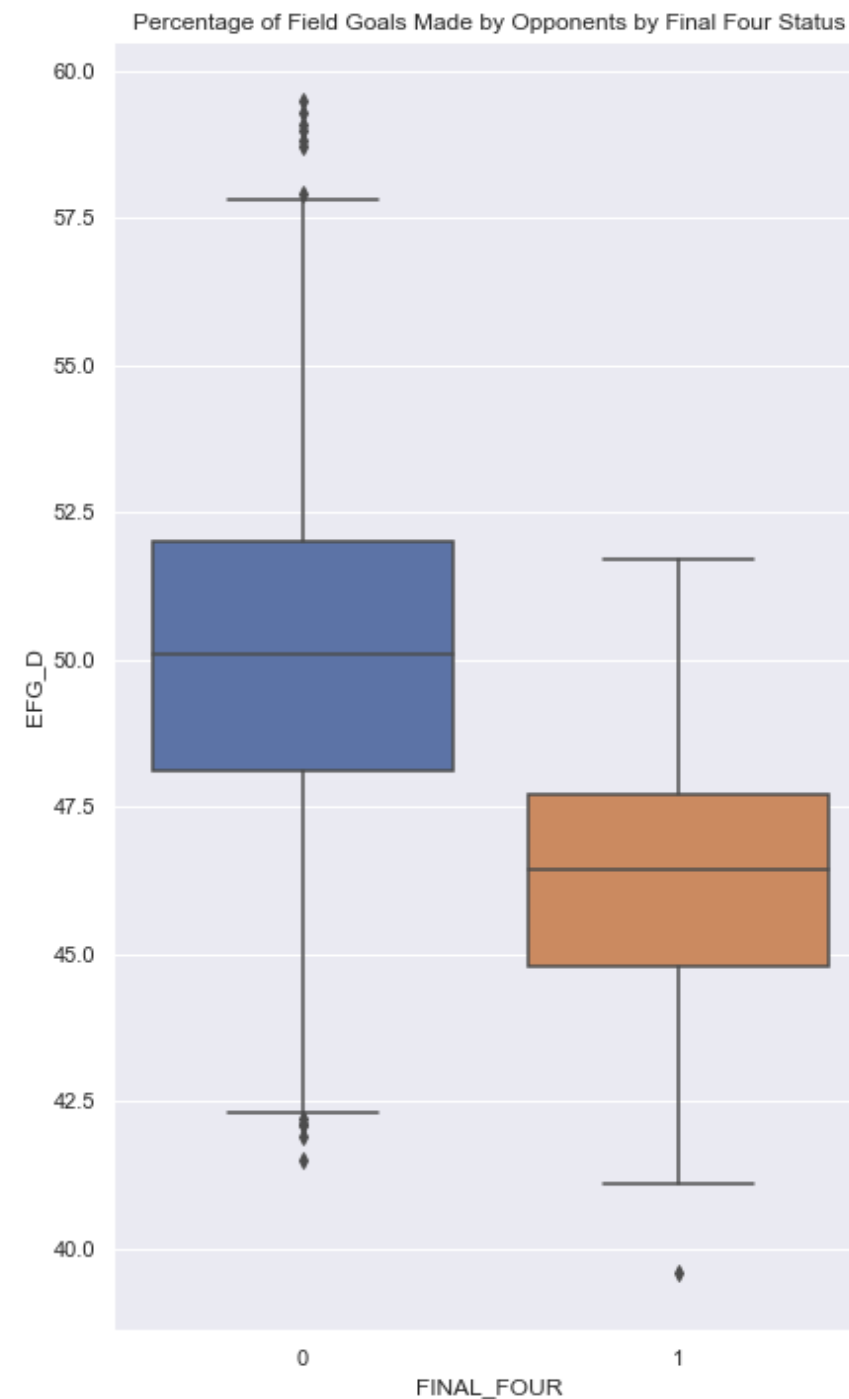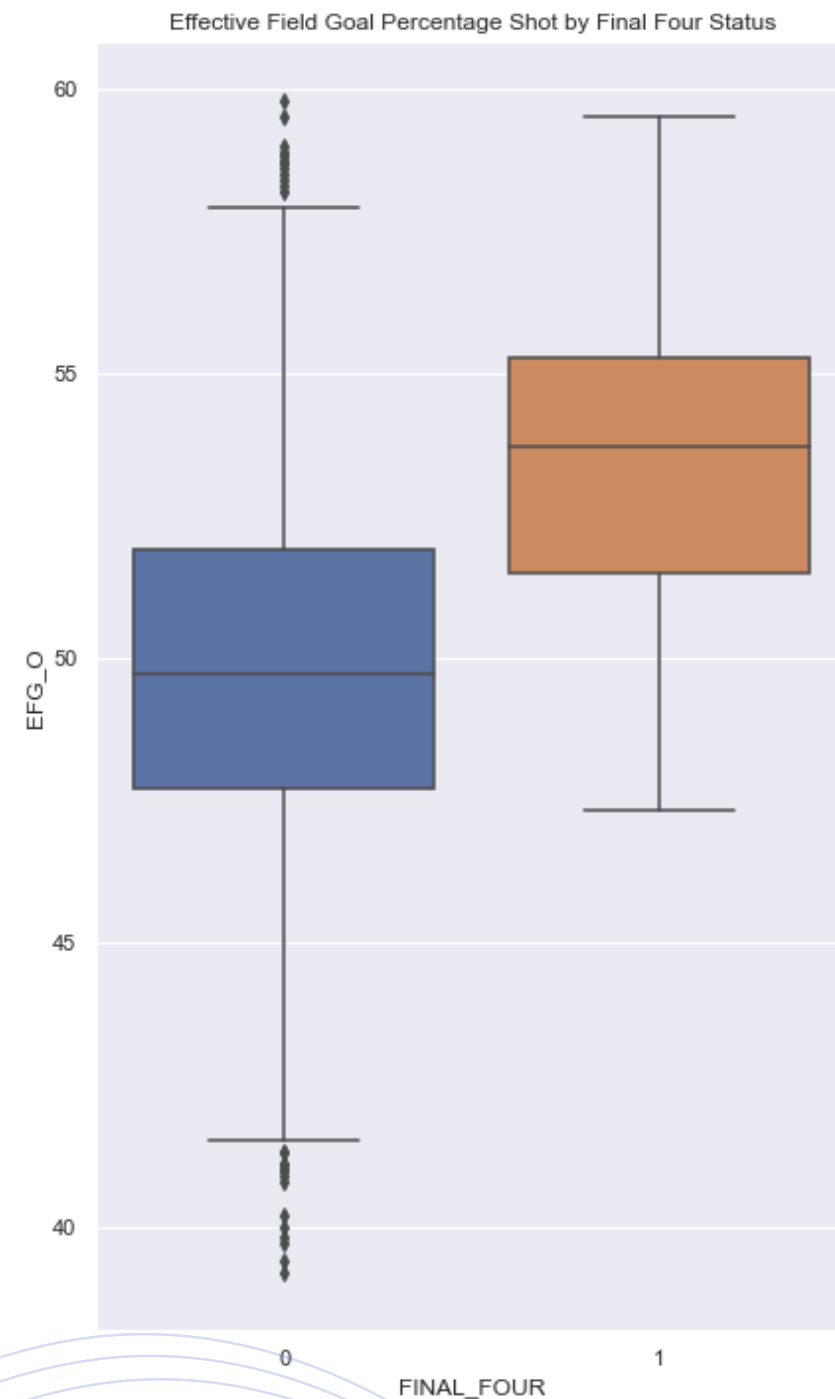
=> **What makes the Final Four Team different?**

- **Machine Learning Model walk-through**
  - Accuracy score
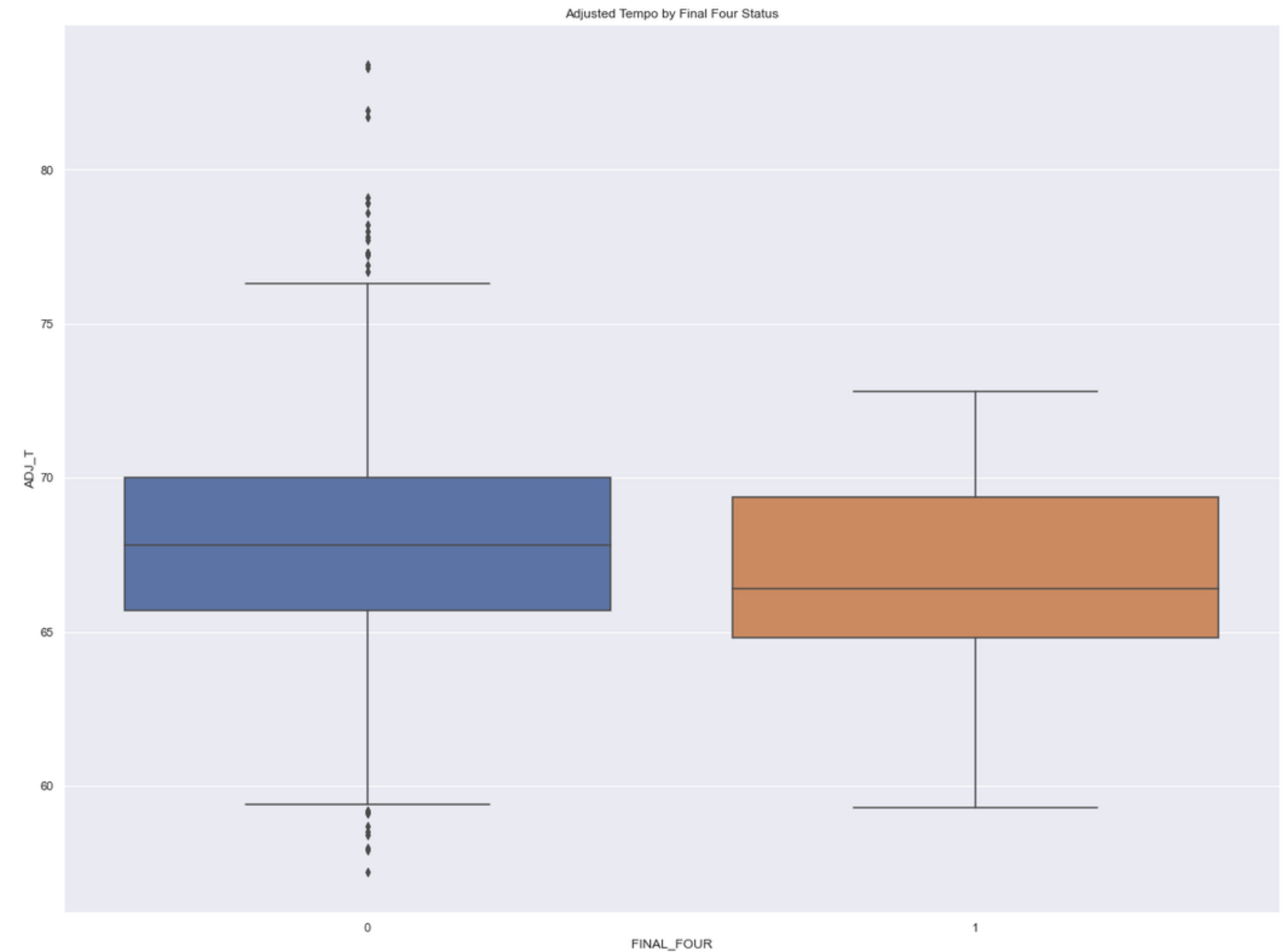 - Important Features

STATISTICS COMPARISON

OFFENSIVE AND DEFENSIVE ABILITY

- Final Four Teams generally have higher effective field goal percentage and a better defensive ability

- The median of effective field goal and defective field goals in other teams are lower

- Final Four Teams have a more consistent range ,while other teams vary more widely
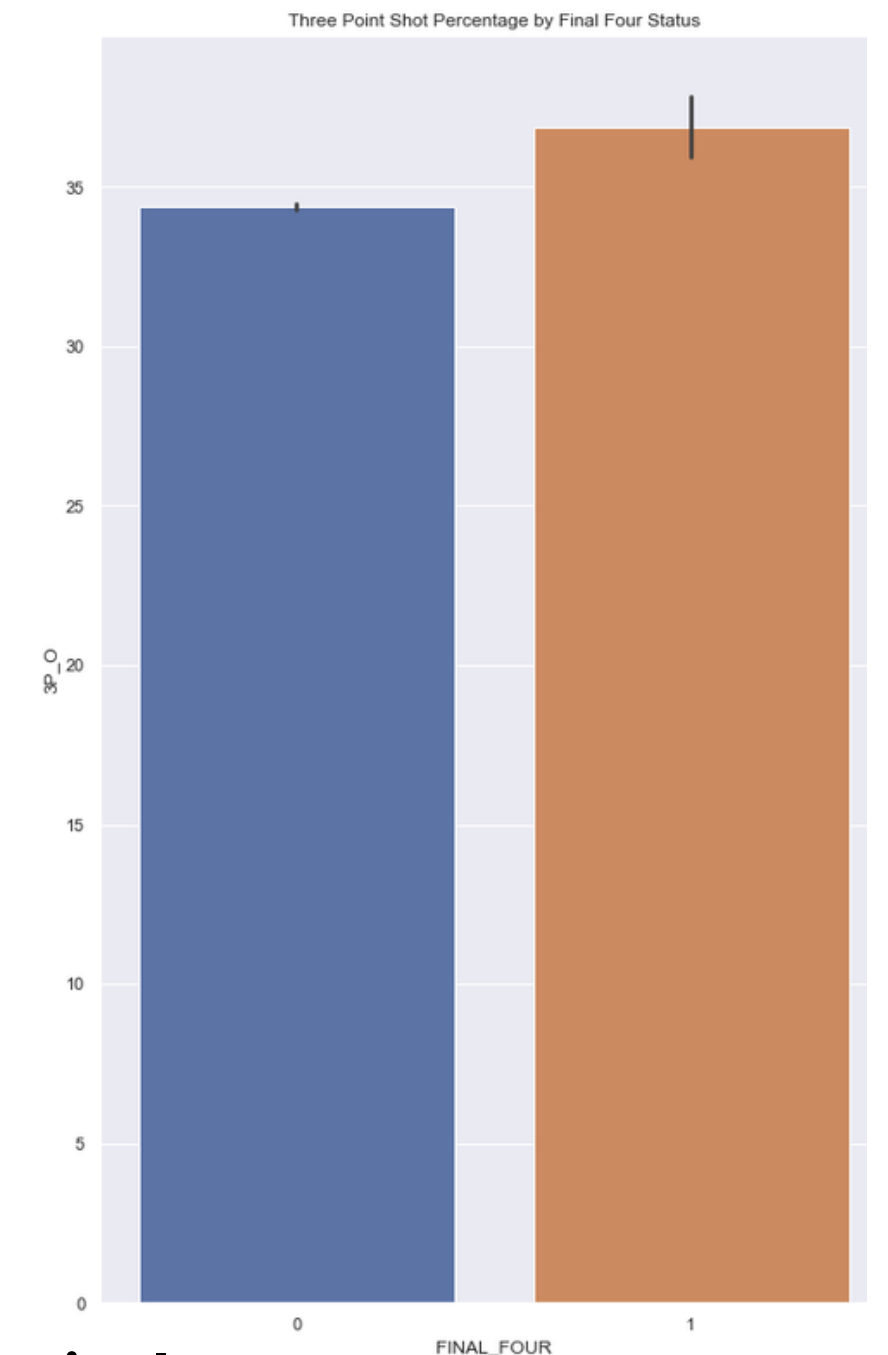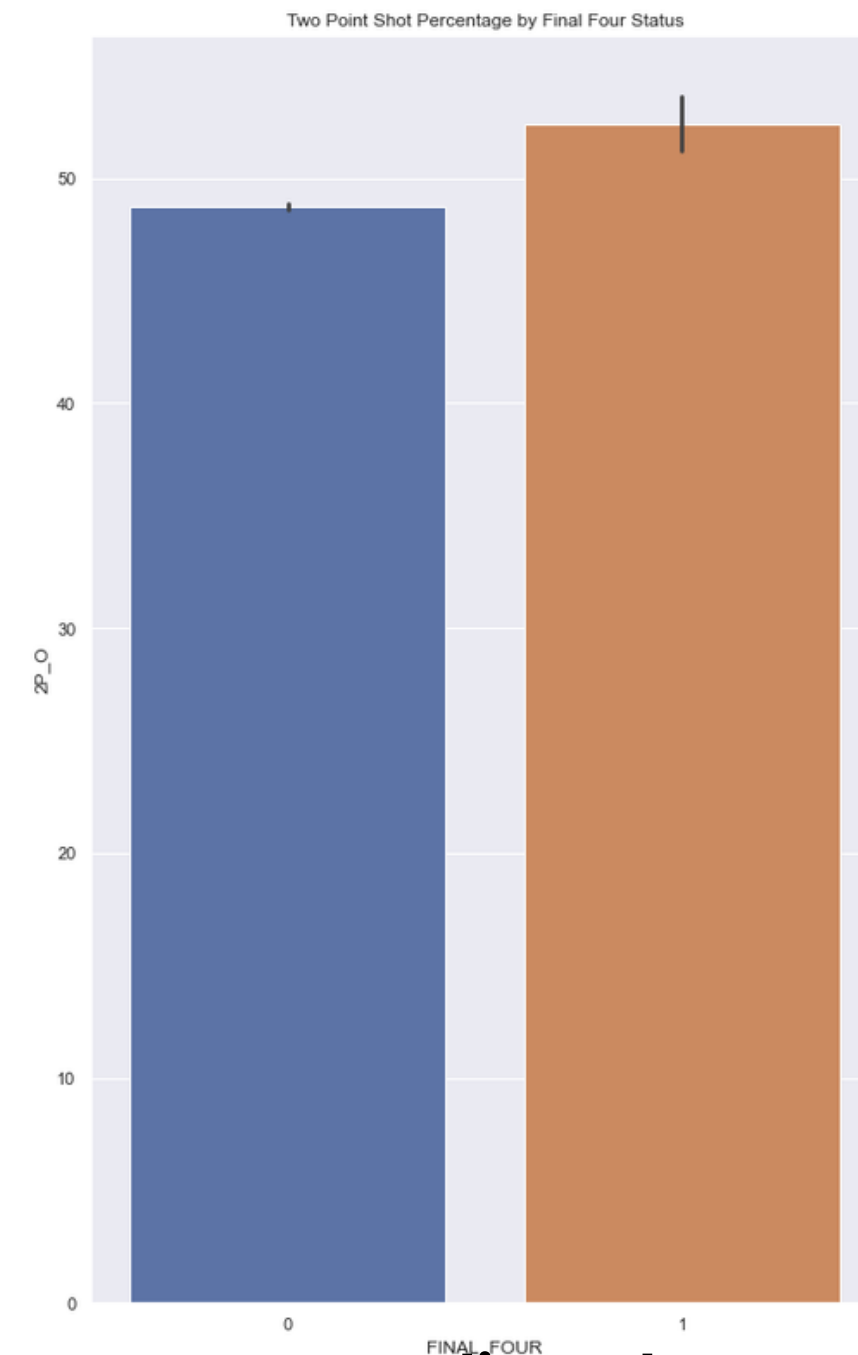
ADJUSTED TEMPO

- Final Four Teams play at a slower tempo than the other teams, normally the tempo ranges from 59 to 73 units

- They also have a more consistent tempo!

Adjusted Tempo by Final Four Status
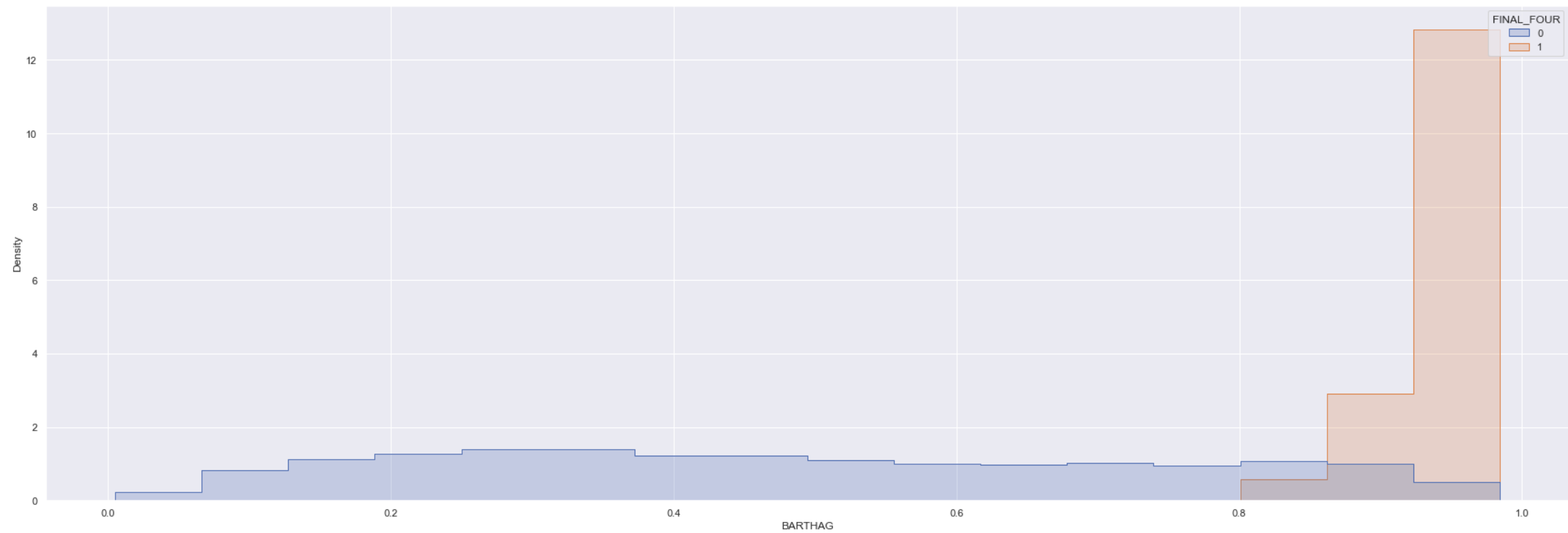
THREE POINT SHOTS & TWO POINT SHOTS

- Final Four Teams perform better in both three point shot percentage and two point shot percentage

- They outplay their opponents by about 7% for both columns

- This suggests that the players for Final Four Teams may be better shooters

Adjusted Tempo by Final Four Status

Two Point Shot Percentage   **Three Point Shot Percentage**

POWER RATING

- The Final Four Team's chances of beating a Division I team is substantially higher than that of other team (0.8 - 1)
- Although the other teams' power rating vary widely, there's higher frequency at the 0.5 mark

MODEL BUILDING

Task: Build a simple model that predicts which teams are most likely to make the Final Four in 2019, using data from 2013-2018

METRICS

- The model yields a 93% accuracy score in comparison to the actual data

- The method I used is a Logistic Regression model, and I implemented SMOTE to handle imbalanced data

- The model succeeds to capture all of the final four teams in the 2019 season!

```
Accuracy of logistic regression classifier on test set: 0.93
Texas Tech        1
Louisville        1
LSU               1
Oregon            1
Michigan          1
Houston           1
Virginia Tech     1
North Carolina    1
Florida St.       1
Kansas St.        1
Iowa St.          1
Wisconsin         1
Syracuse          1
VCU               1
Virginia          1
Florida           1
Oklahoma          1
Kansas            1
Maryland          1
UCF               1
Auburn            1
Michigan St.      1
Gonzaga           1
Kentucky          1
Purdue            1
Duke              1
Tennessee         1
```

THANK YOU
FOR WATCHING