



F1 RACING WINS

Nayma Kim, Morgan Prukop,
Leticia Rinaldini

PRESENTATION OUTLINE



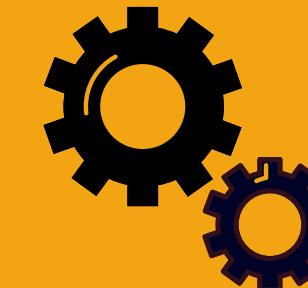
The Data
Slide 1



Project goal
Slide 2



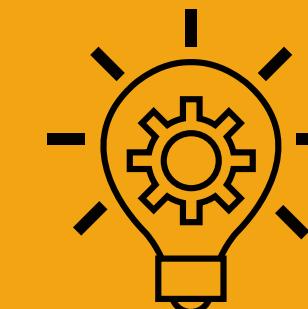
General Trends
Slide 3



Project Tools
Slide 4



Analyzation
Slide 5 - 13



Solutions
Slide 13

THE DATA

- Found on Kaggle
- Originally 13 separate csv files
- Includes information on the constructor, driver, track, date, speed, status, laptimes, and overall results.
- Our data spanned from the 1950 races through 2017
- Featured Engineered columns: age and is_home

[Kaggle link](#)

[F1 Guide](#)





PROJECT GOAL

To understand and analyze our data and models for predicting how a driver will place in a race based on the effects various factors.

GENERAL TRENDS & FINDINGS

- British and Finnish drivers have the best average final positions
- Driver's racing in their home countries have a better chance of winning
- Some constructors do better consistently so driving for them gives you an advantage
- The qualifying rounds can really help or really harm your chances of winning



PROJECT TOOLS

link to our
notebook



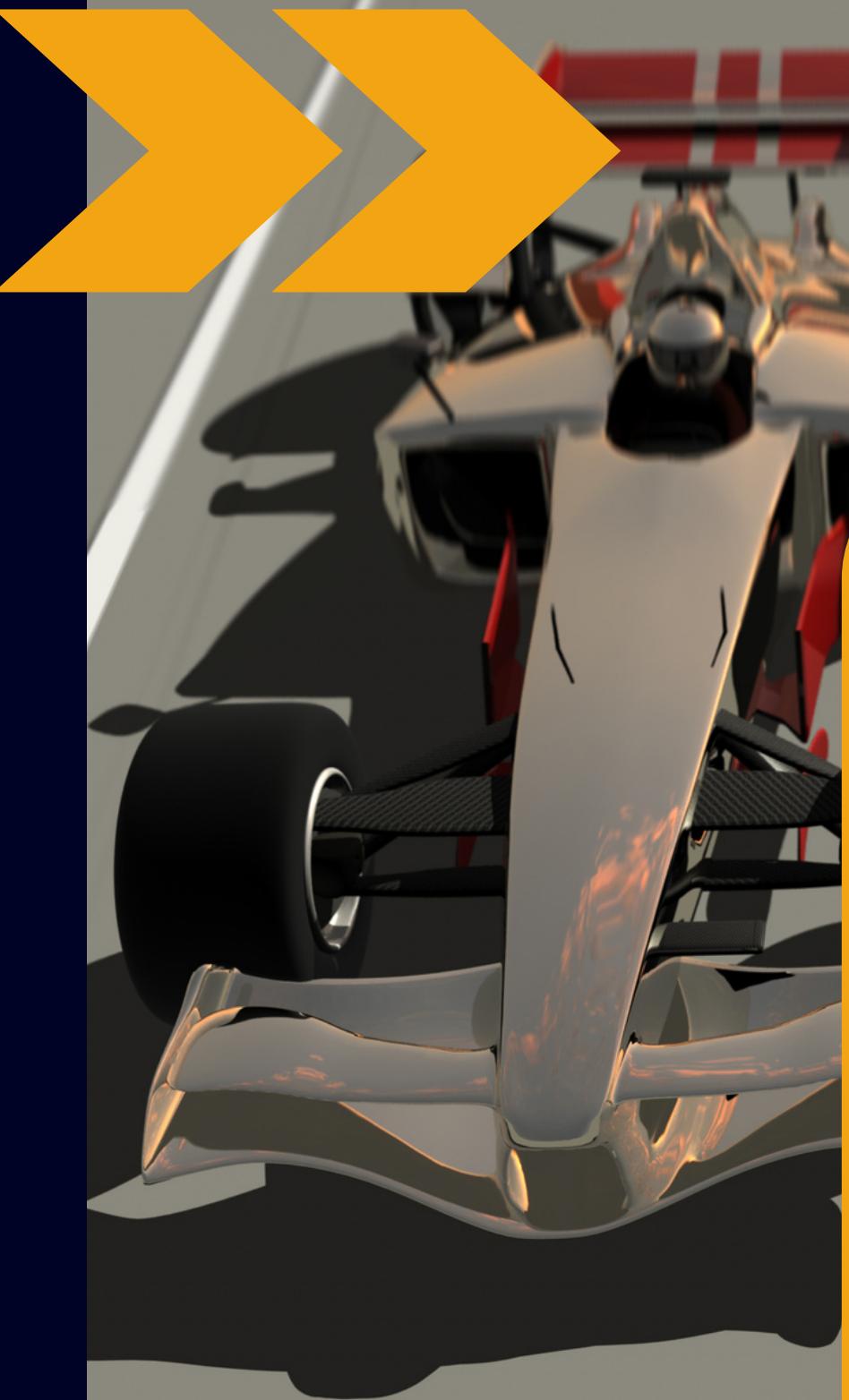
Google Colab Notebook

We used python coding in
google colab notebooks to clean
and analyze the data as well as
create our model.

ANALYSIS

A black and white photograph of a Formula 1 racing scene. On the left, a side view of a race car is shown, featuring prominent 'intel' and 'O2' sponsorship logos. The number '16' is visible on the sidepod. On the right, a driver wearing a 'PETRONAS' helmet and suit is seated in the cockpit of the car, looking forward. The background consists of a dark, textured wall, likely a garage or pit wall.

DEFINITIONS



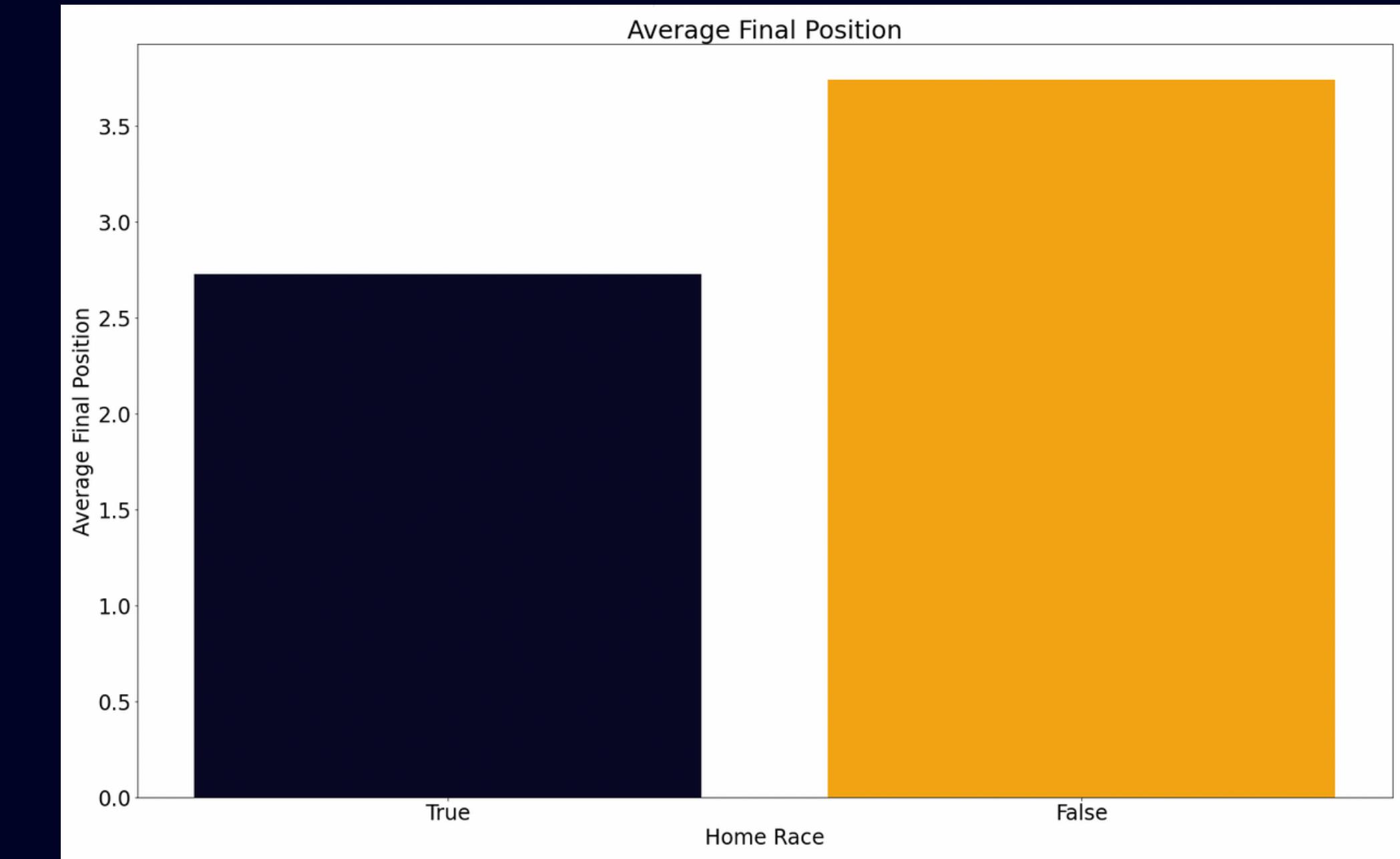
Definitions

- **Constructor:** AKA team. The constructor is the owner of the engine and financially backs the drivers they choose.
- **Circuit:** a long loop of road that must be approved by the FIA. Circuits are usually only a few miles long, but drivers lap around them several times each race.
- **Grid:** the position where the driver will begin their race. Position is determined by qualifying session prior to the race.

How F1 Works

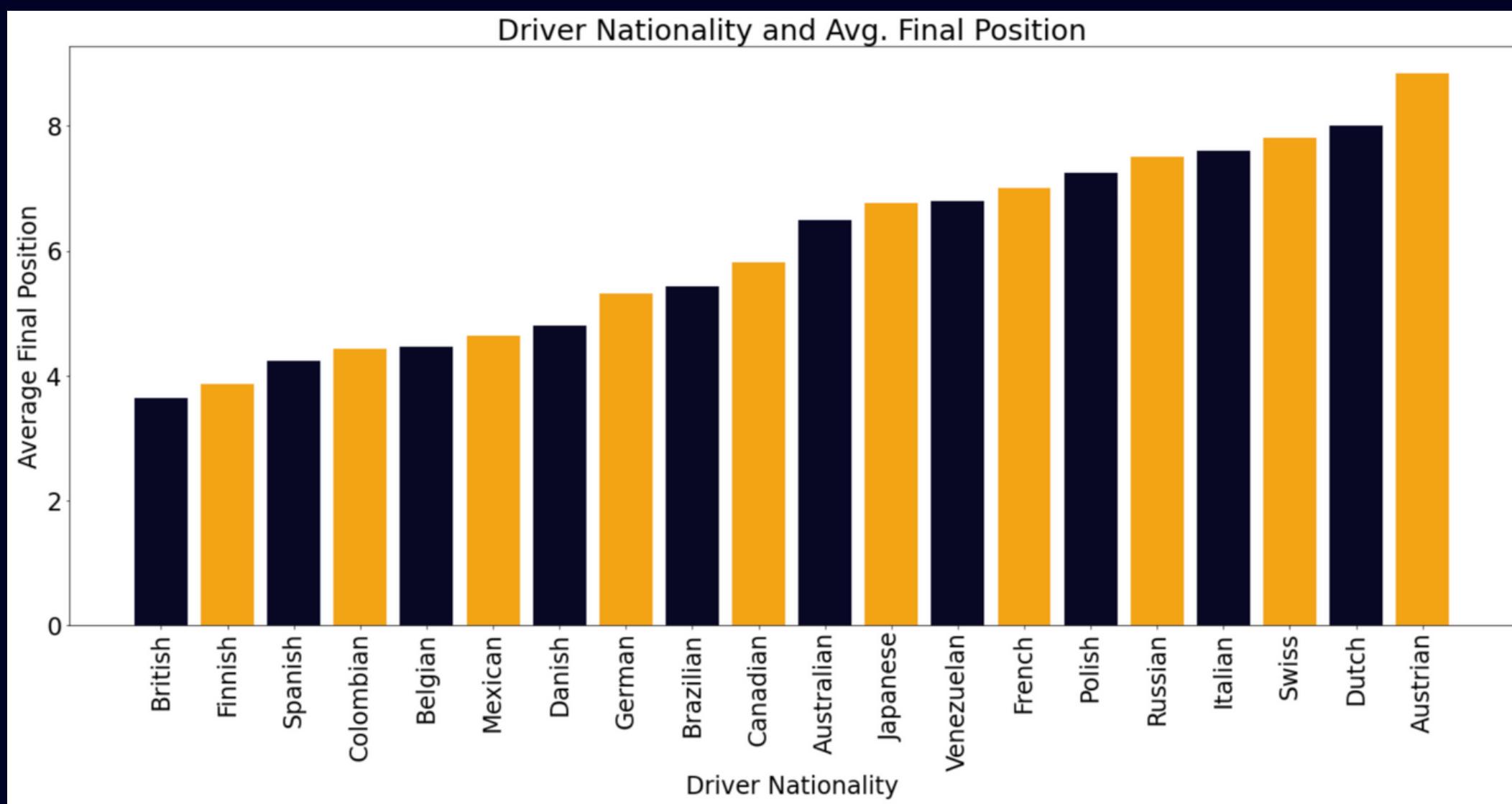
- Formula1 is a series of races held at various global circuits.
- There is A LOT of intricate requirements per vehicle
- Each race totals an average of 305km (190 miles)
- There is a point system, which is the same as position. Only the first 10 to finish are awarded points.
- Between 17–22 races make up the season

HOME RACE & AVERAGE FINAL POSITION



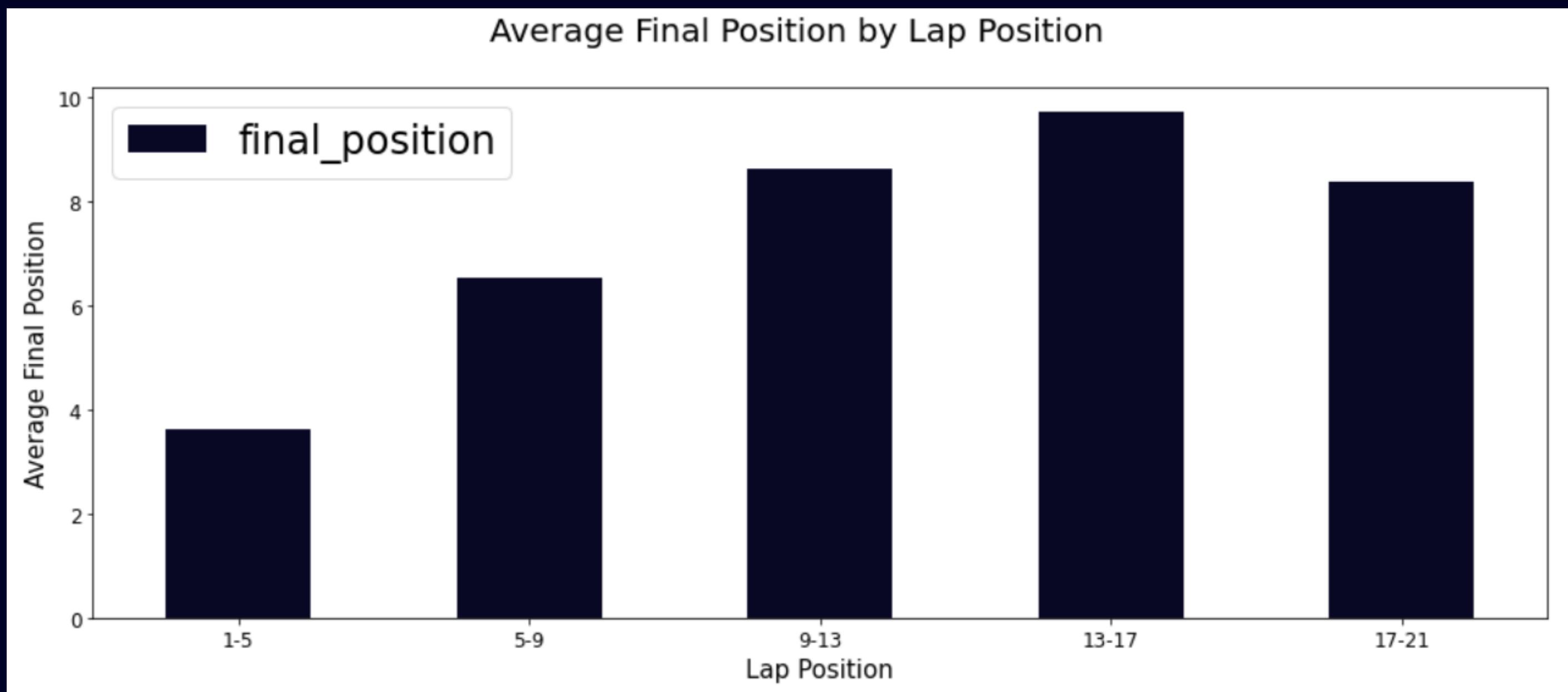
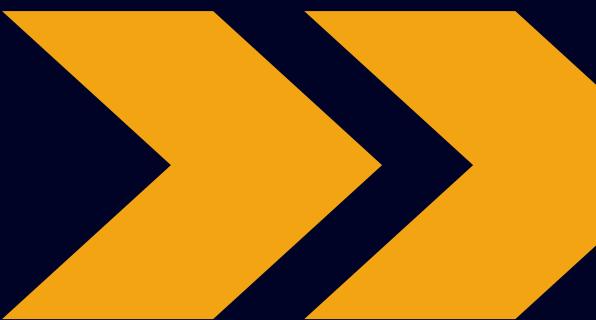
The average final position for those racing in their homeland is better than those who are racing away from home.

DRIVER NATIONALITY & AVERAGE FINAL POSITION



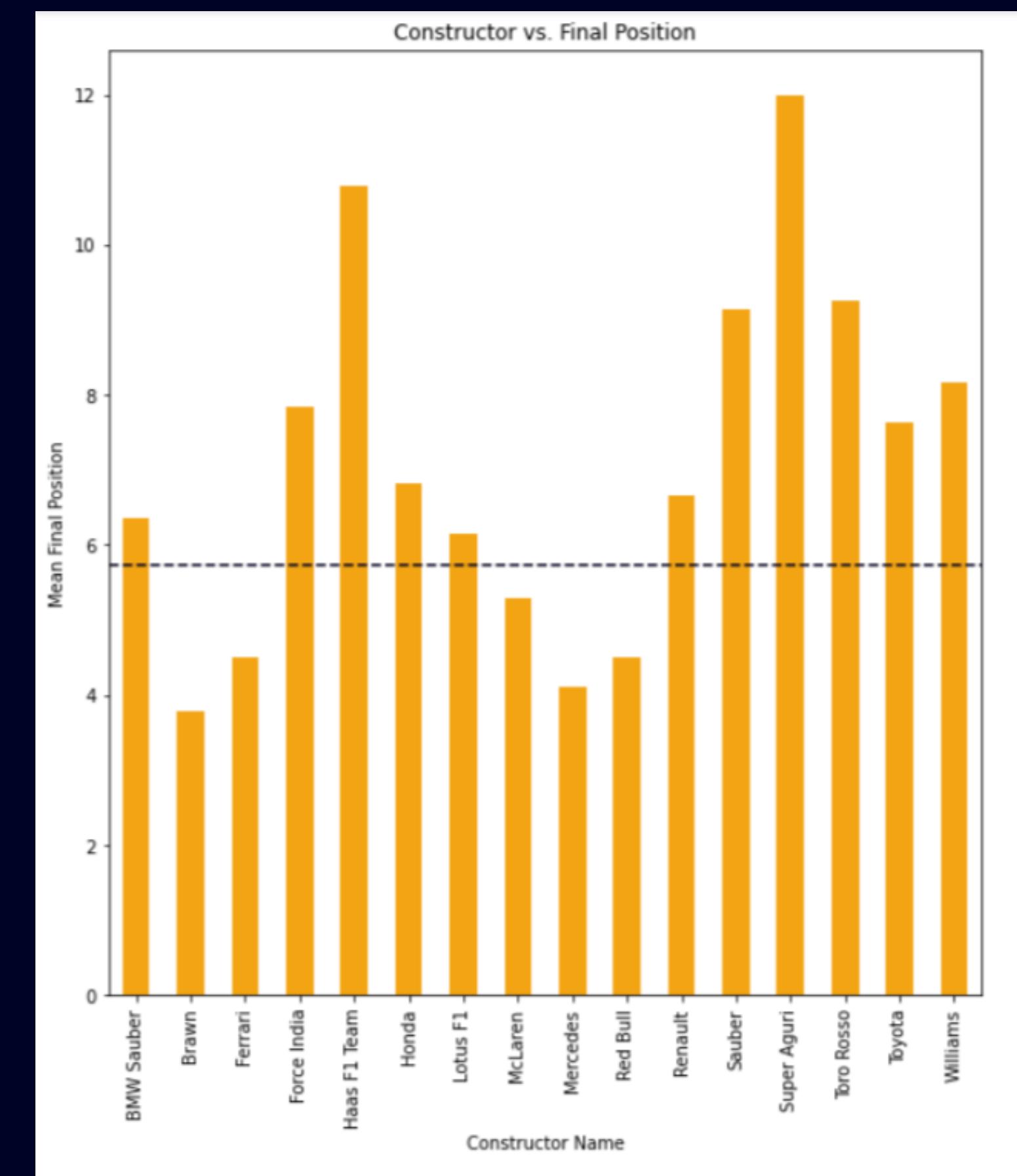
- British drivers have the best average final position
- As of 2019, Britain is home to the second-highest number of F1 sponsors.
 - 20.8% of the 216 companies partnering with the sport were based in Britain

LAP POSITION & AVG. FINAL POSITION

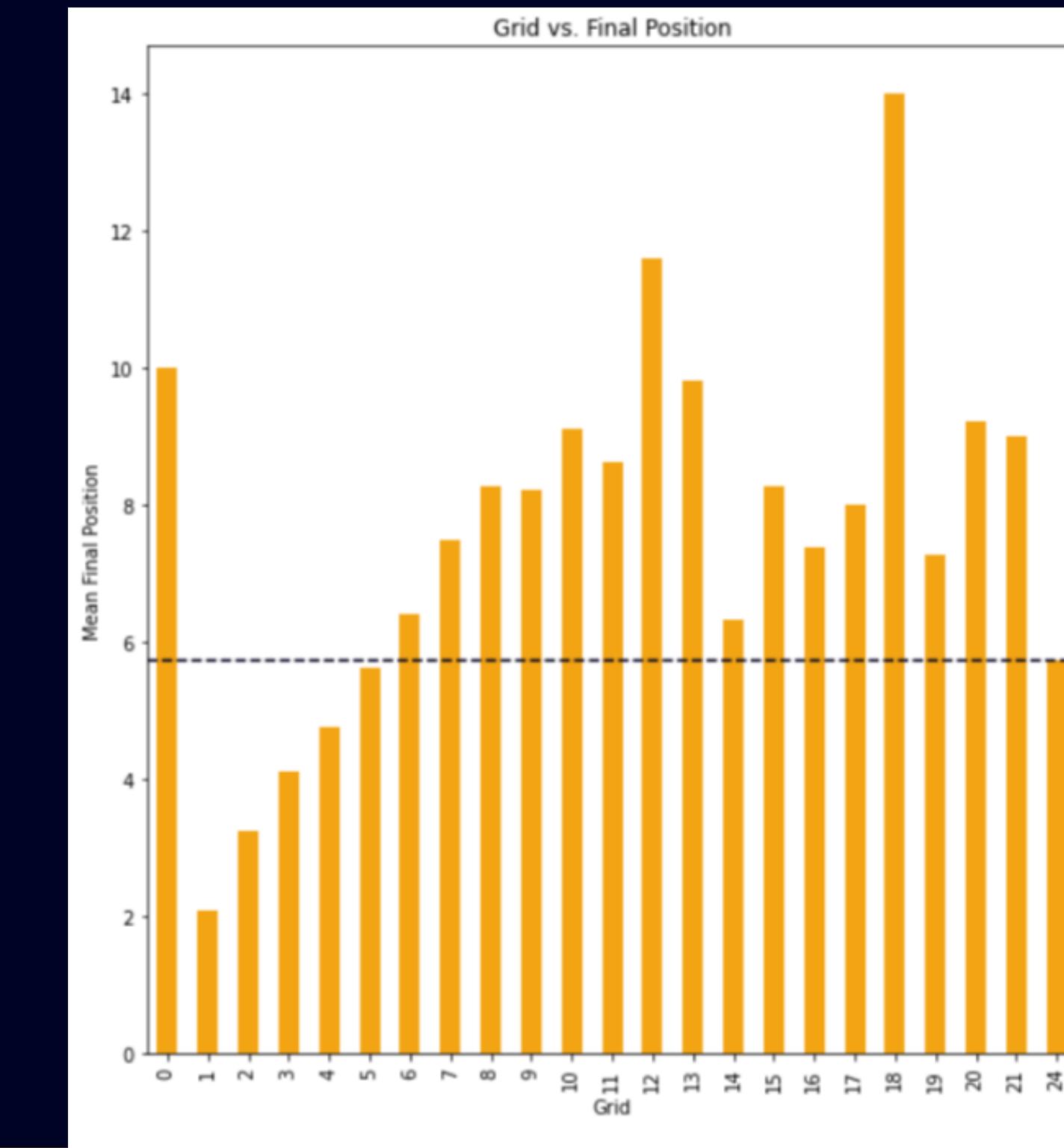


CONSTRUCTOR VS. FINAL POSITION

Certain constructors
win more
consistently than
others.

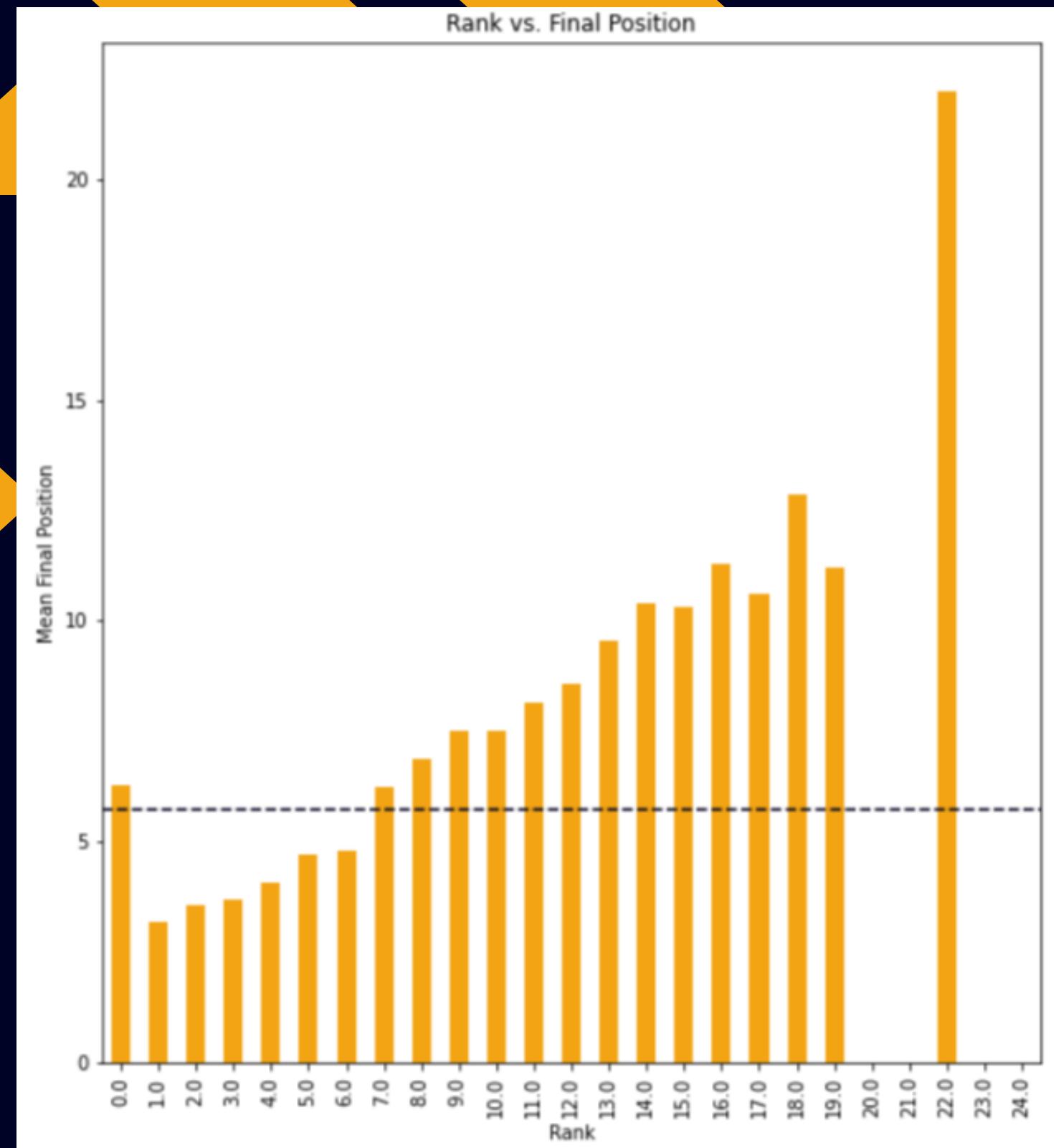


GRID & AVERAGE FINAL POSITION

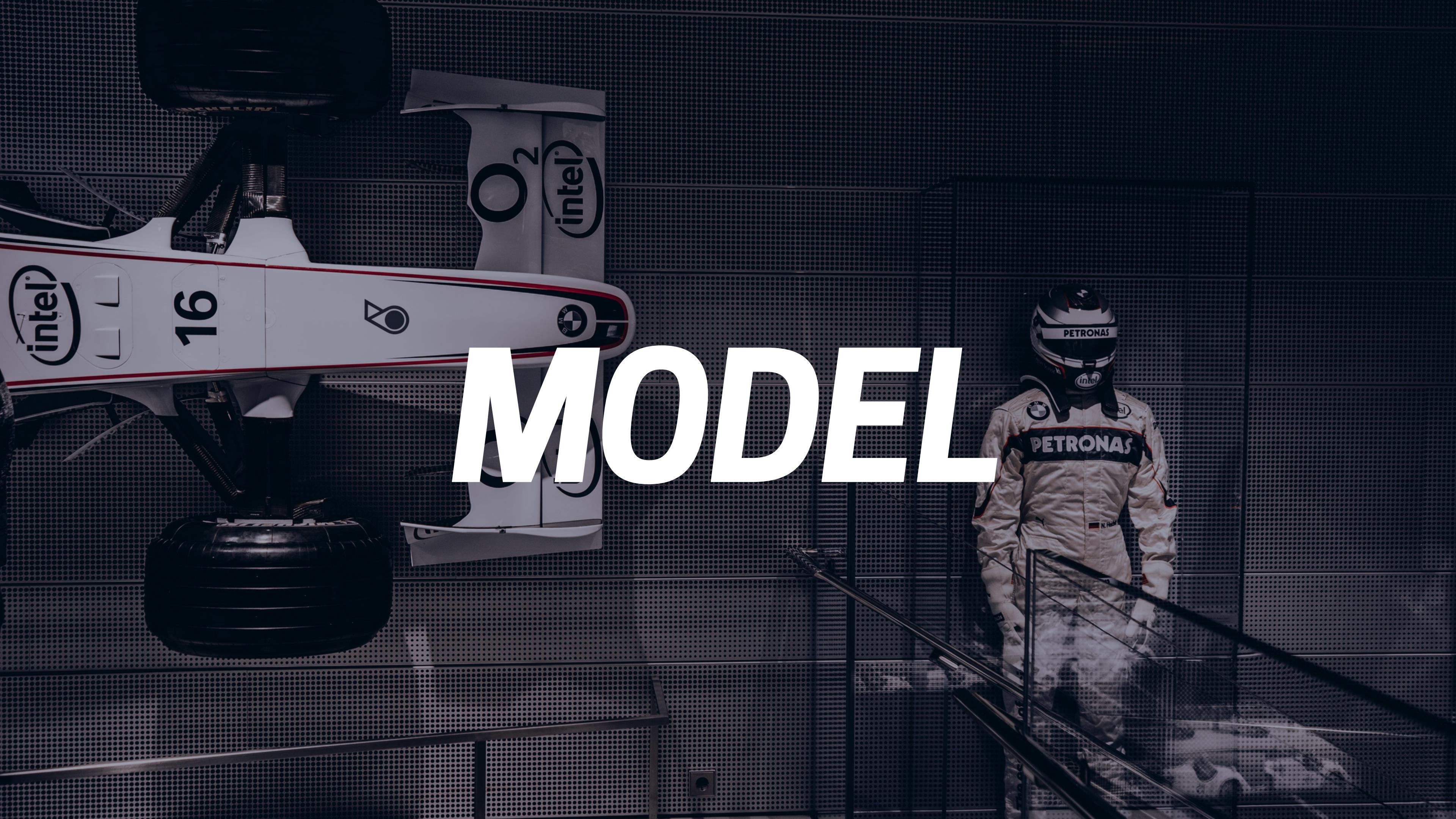


Grid position, determined by the qualifying races, is crucial to the success of a driver. The qualifying rounds aren't considered the most exciting part of the race for nothing!

RANK & AVERAGE FINAL POSITION



- Rank plays a part in success



MODEL



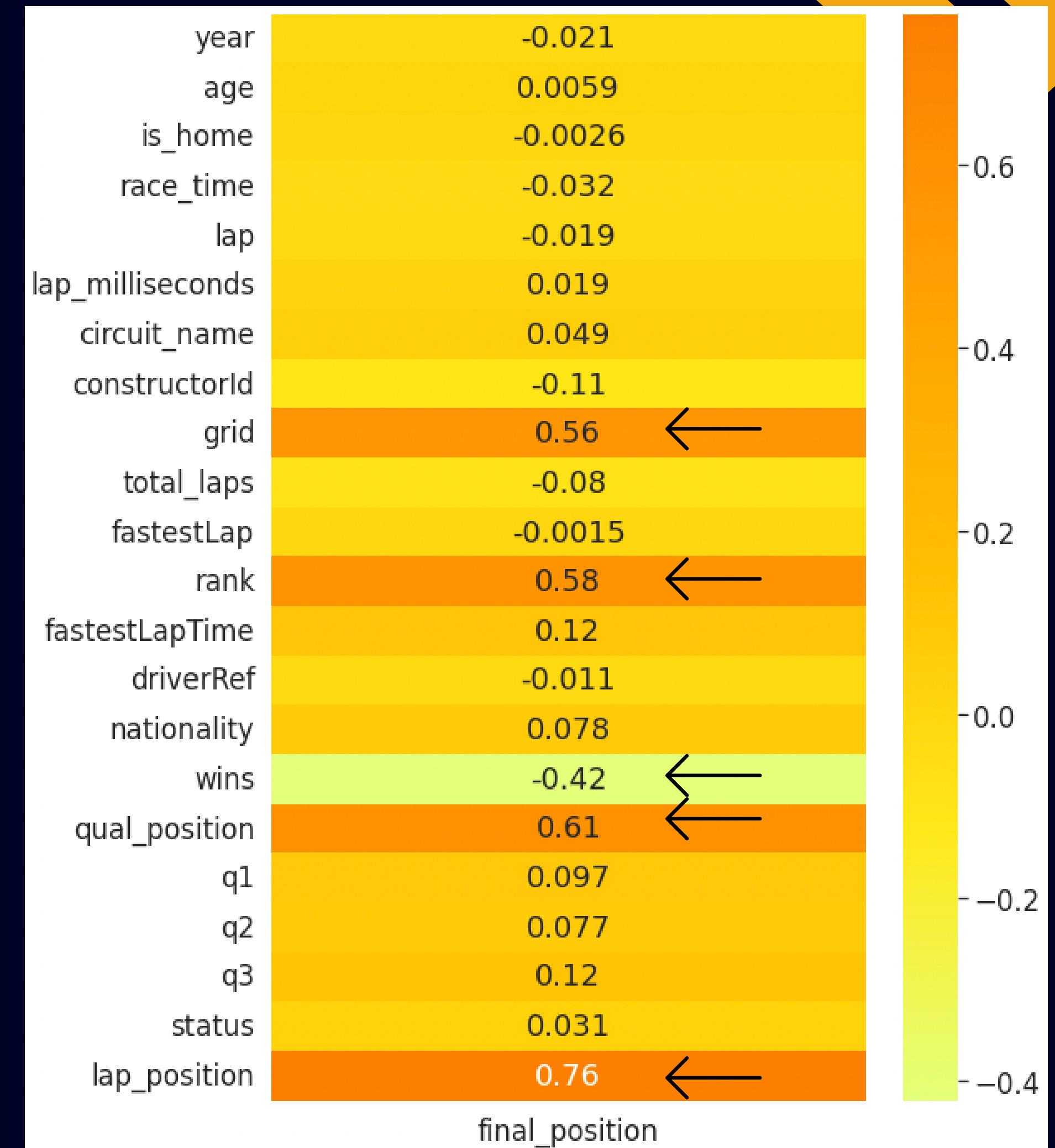
START TALK ON MODEL

We began our model process by finding the columns which best correlated with final position.

We excluded result points as that is the equivalent to the final position.

Columns Used in Model:

year, age, is_home, lap_milliseconds, circuit_name, grid, rank, fastestLapTime, fastestLap, wins, nationality, qual_position, q1, q2, q3, status, lap_position



LINEAR MODEL

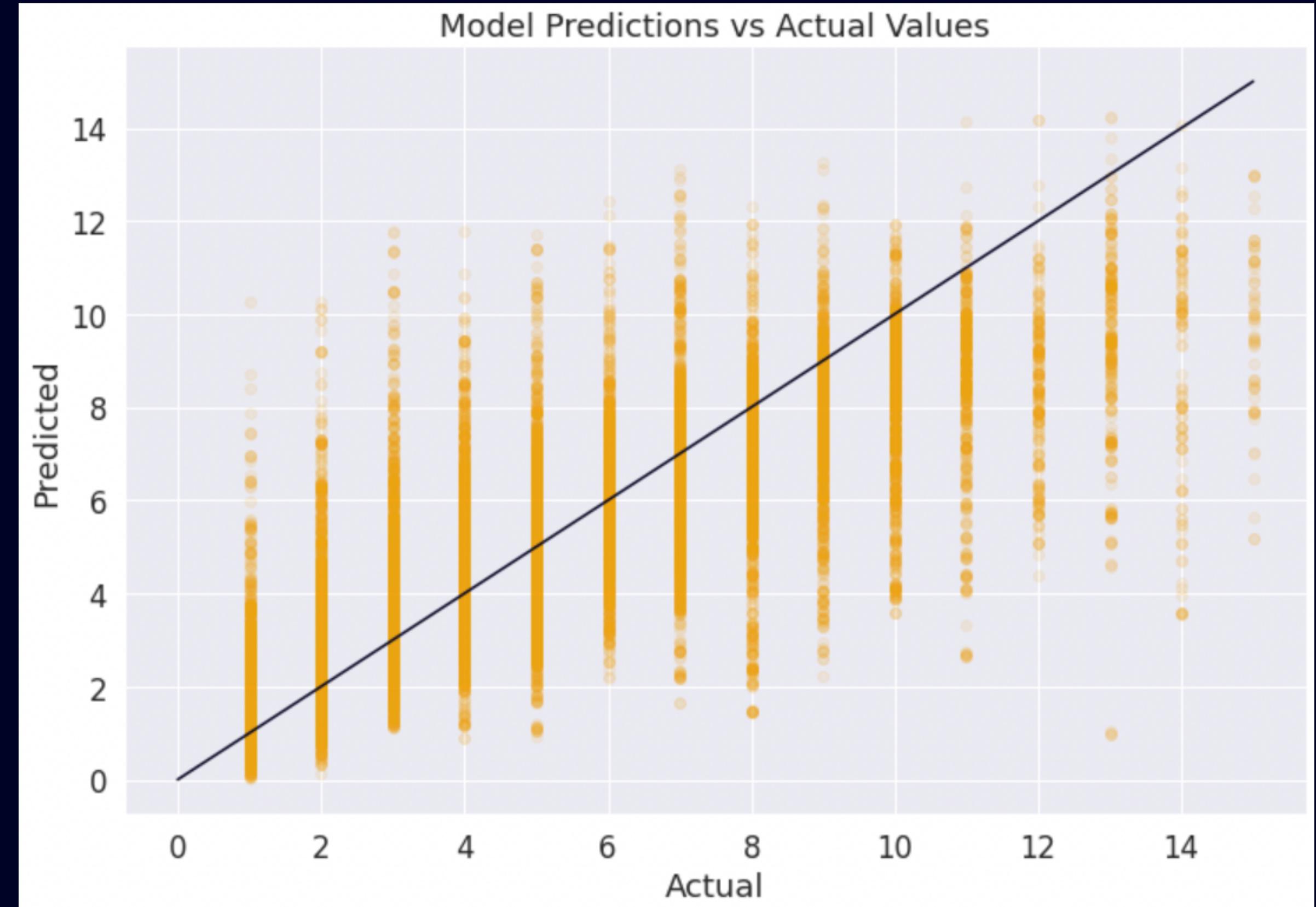
Model 1

This model served as our simple linear regression starting point.

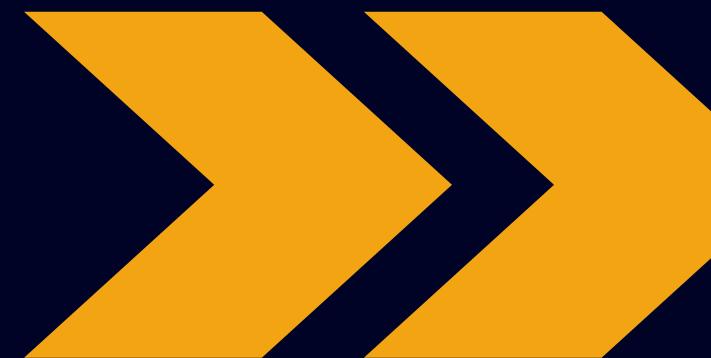
R² = 0.6765

MAE = 1.1581

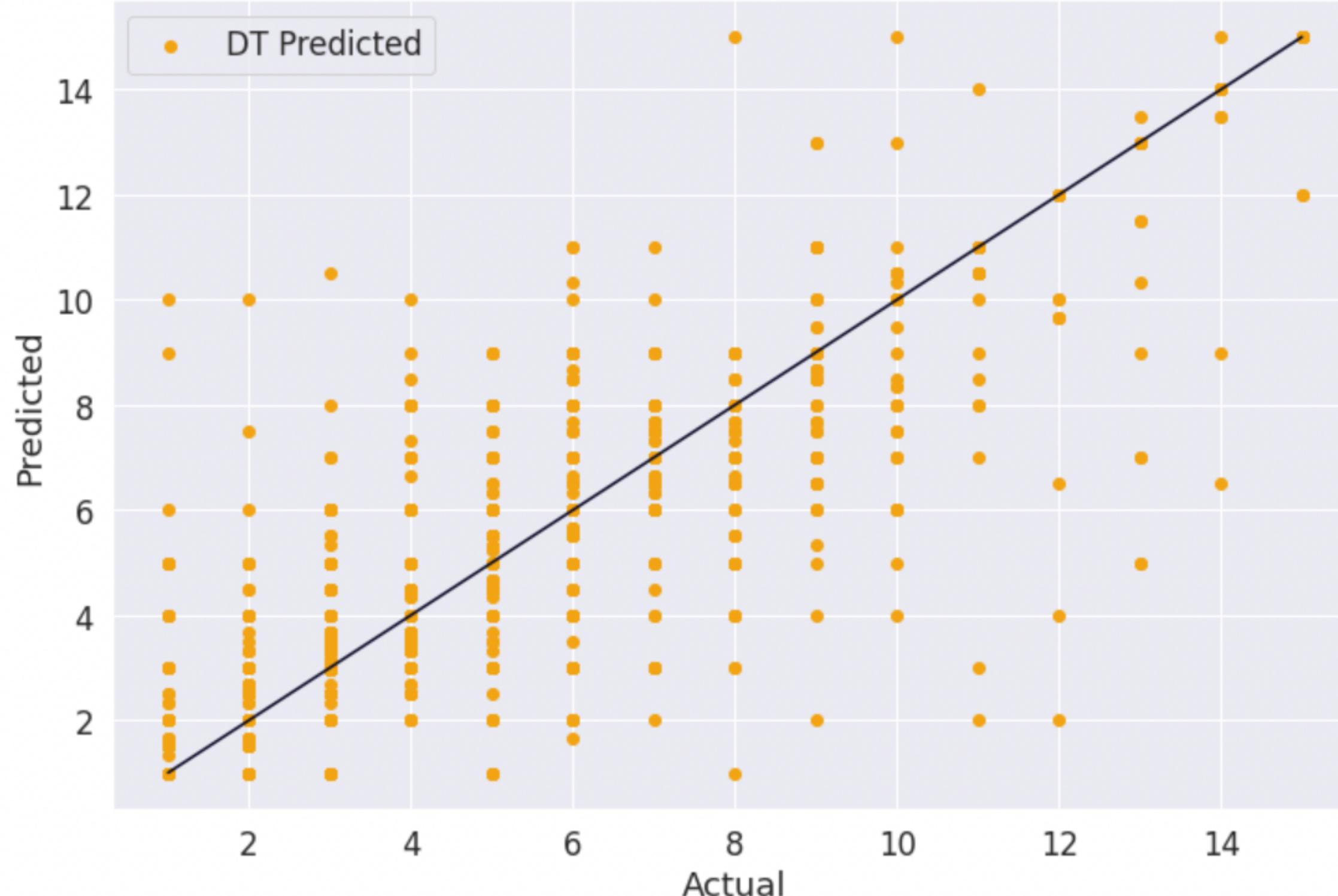
MSE = 2.8839



DECISION TREE MODEL



Decision Tree Model Predictions vs Actual Values



Model 2

We decided to take our findings and create a decision tree model based on our correlations. `max_depth= 30, min_samples_leaf=2, min_samples_split= 4`

$R^2 = 0.9536$

$MSE = 0.2079$

SOLUTIONS TO OUR PROJECT



- Racing in your home country will set you at an advantage.
- Doing well in the qualifying rounds and having a better grid position will help you to be successful
- Some constructors, or teams, are better than others and being a driver for them can also put you at an advantage

THANK YOU

Thank you for listening and we
hope you enjoyed our
presentation and learned
from our work!

VROOM VROOM!

