What Makes a FANTASTIC Grand Prix?

John Omolo

Goals

Clean & Prep the Data

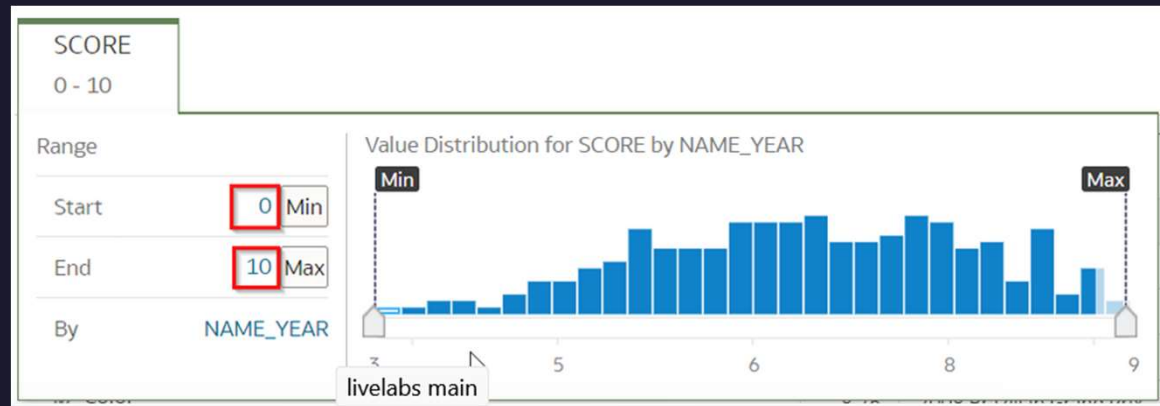Visualize the data

Use ML to make predictions

Introduction

- Our goal is to identify the factors that make a great Formula 1 race. Factors can be the number of overtakes, the weather, the circuit that the race is being held and so on. First of all, we will look for clues in the data, through visualization, to identify candidate variables and then build a model to predict the score fans will give to a race.

What do we have?

- Information on races (1950-2021). For example, the circuit of each race, the lap times, pit stops and the results of each race.

- A fan score per race (2008-2020) (source: racefans.net). These scores were collected from many fans and result in one aggregated average score for each race. We will treat this fan score as the **ground truth** of how "good" a race is. Note that we **don't** have fan scores for 2021.

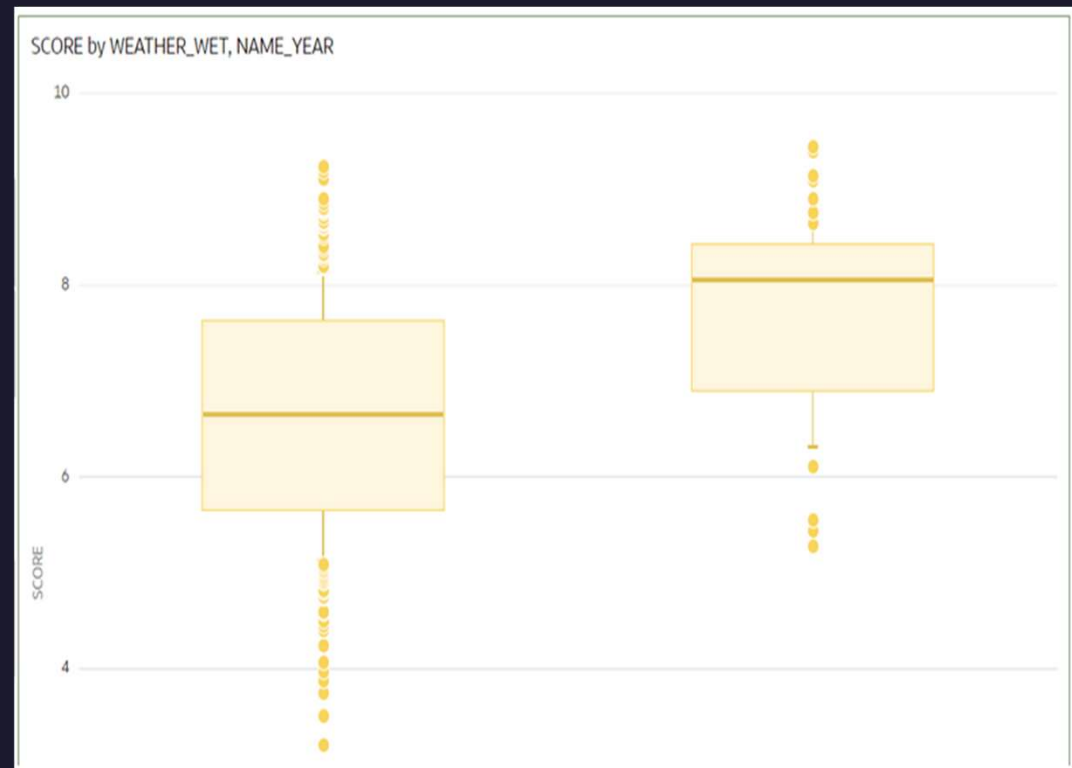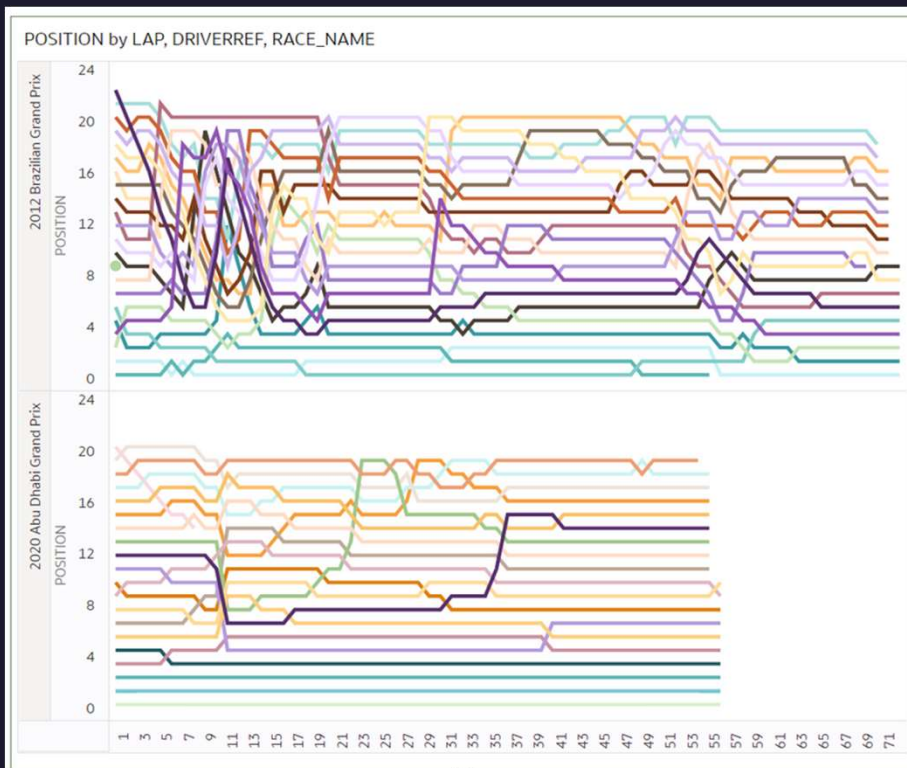# Best & Worst Race Distribution According to fans

DNF's, Overtakes & Weather.

From this graph we can see that the more popular race has more overtakes than the other one. The weather also plays a significant role.
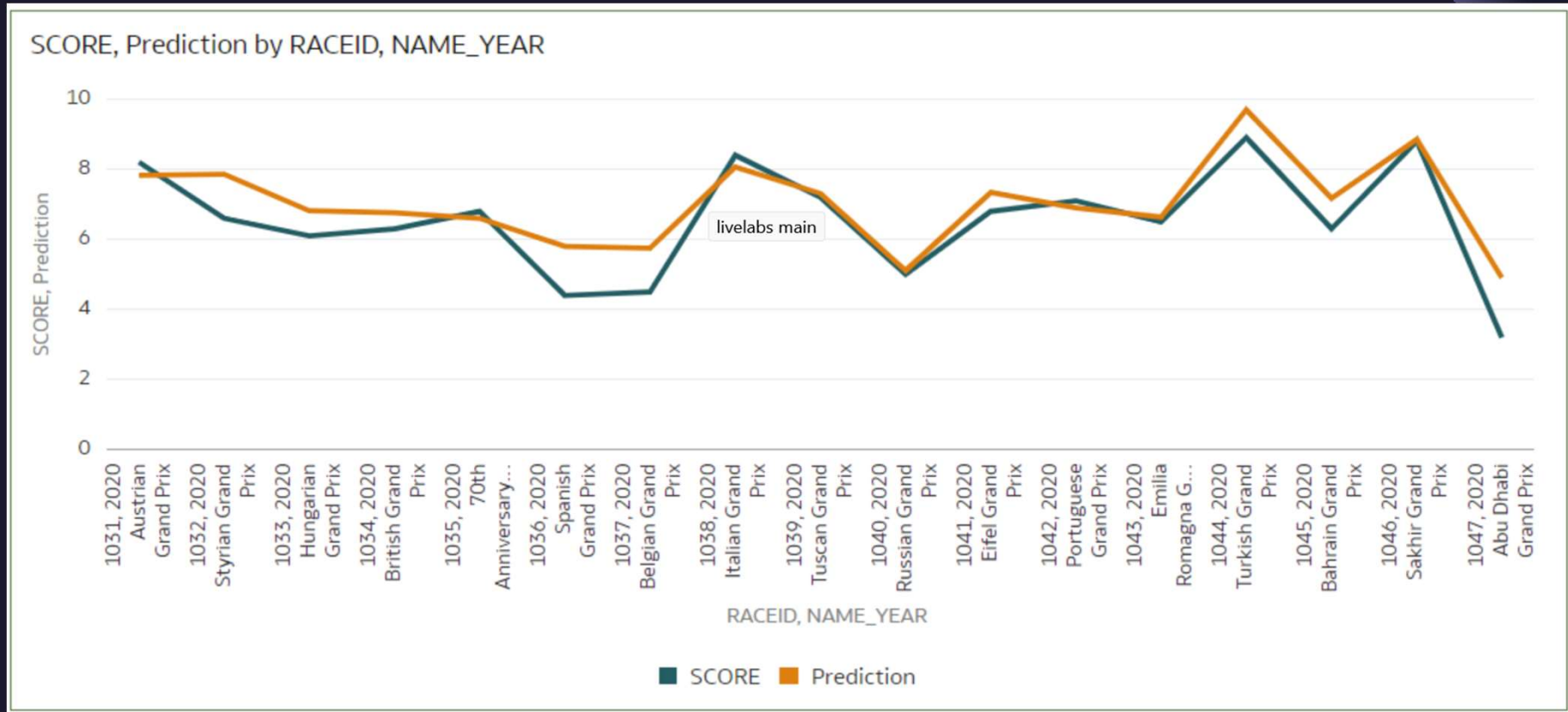
The way to get started is to quit talking and begin doing.

Walt Disney

Using a generalized linear model we can see that our model positively correlates with actual data and may then be used to make predictions.

SCORE, Prediction by RACEID, NAME_YEAR

Conclusions and Next Steps

STARTING POINTS & WAY FORWARD

- Overtakes: We found that take overs are an important factor for the race score. What could we do to improve takeover probability, especially in the top positions?

- Weather: It appears that wet weather adds to the excitement of the races. Maybe we could consider not cancelling a race because of weather conditions?

- Safety car: We should look at how this impacts races.

Thank You

J. Omolo

j.kasimba@gmail.com

www.nyakinda.com