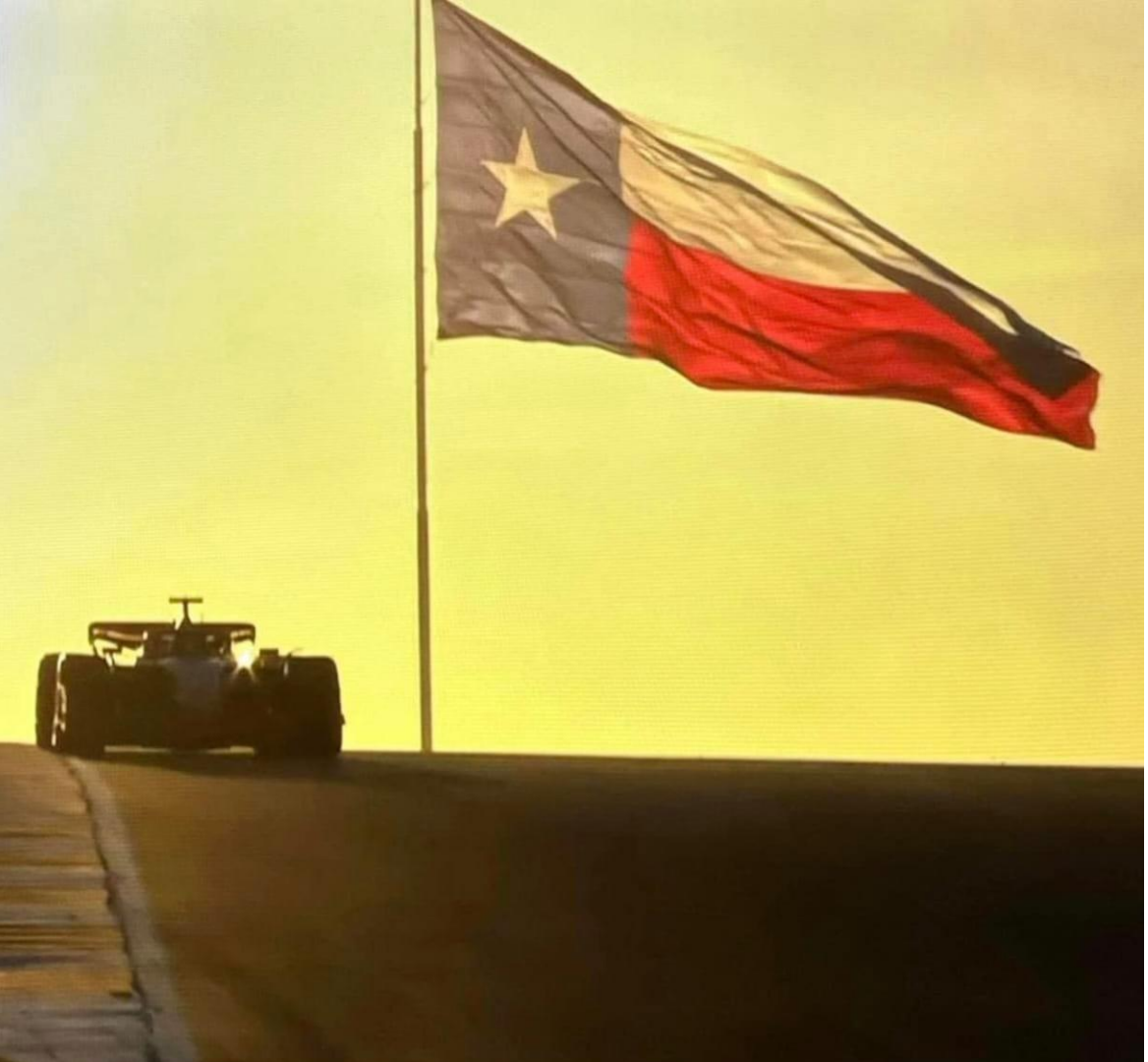


Formula 1™

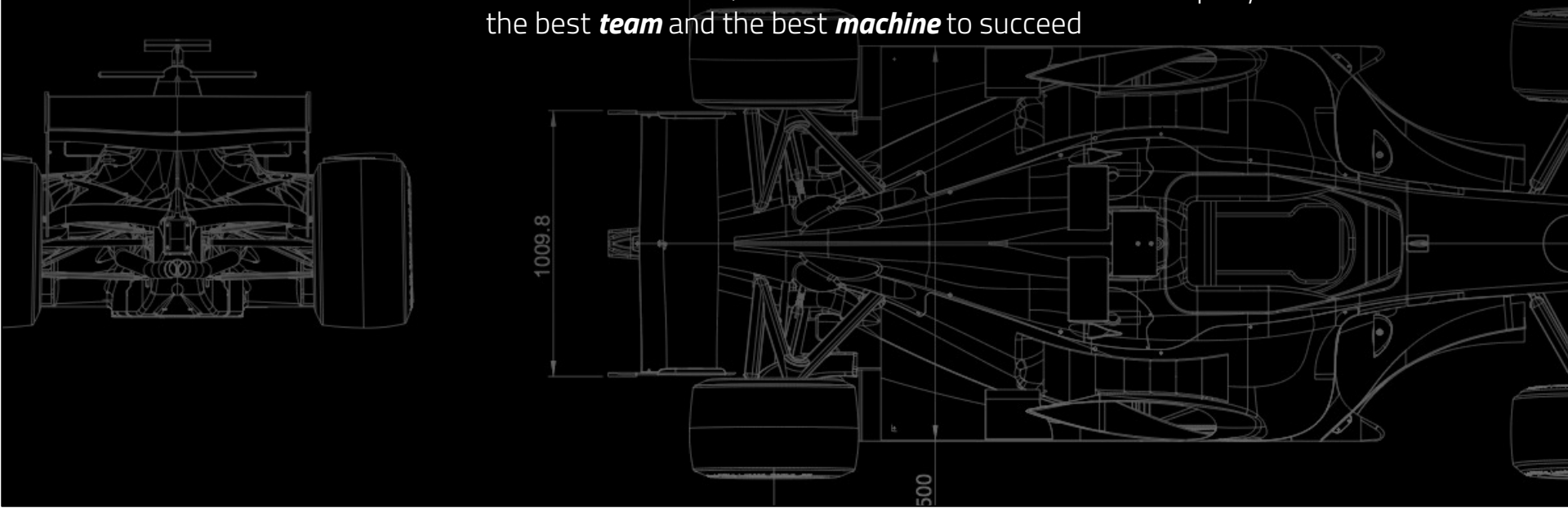
Knowledge Management system and Trend analysis

GROUP 9 - TEAM

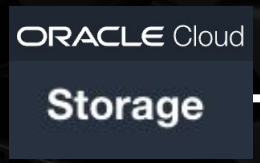
Aniket Patil	AAP3788
Apurva Audi	AA85254
Chyavan M Chandrashekar	CM65624
Kavya Angara	KA32577
Kyle Tobia	KJT887
Soumya Agarwal	SA55638

GOAL

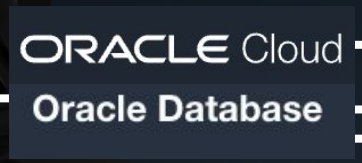
When Audi enters Formula-1 in 2026, we want to make sure that the company has the best ***team*** and the best ***machine*** to succeed



SETUP



Data Lake
(Bucket)



Database
(Workload type: Data Warehouse)

Application-1

Driver Analysis

Application-2

Race Engineering

Application-3

Machine & Race Strategy

DATA LAKE

Created a bucket in Oracle Cloud - Storage

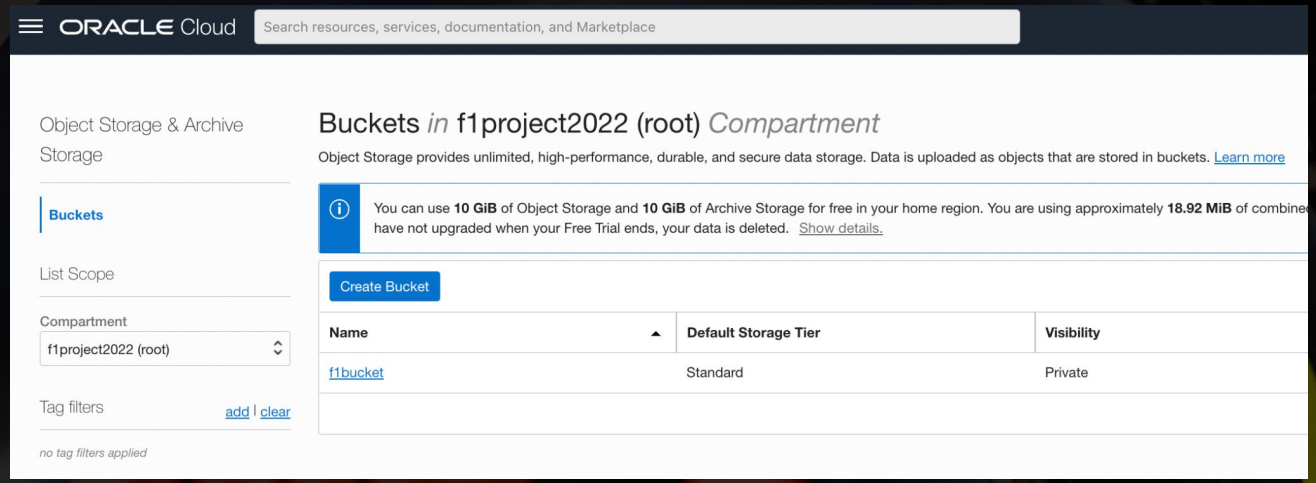


Fig 1. Bucket in Oracle Storage

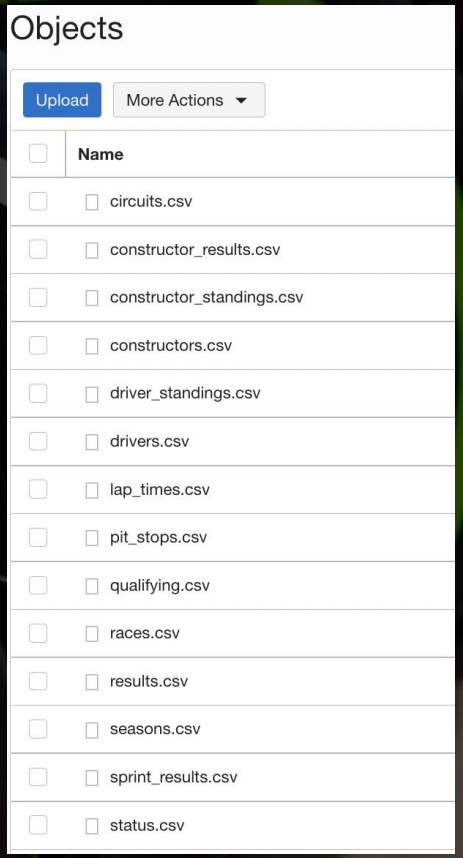


Fig 2. Files in Bucket

DATA WAREHOUSE & DATABASE

- Created autonomous database in Oracle Cloud - Database
- Data Warehouse is the selected Workload type
- ‘f1database’ is what we are using as our SSOT

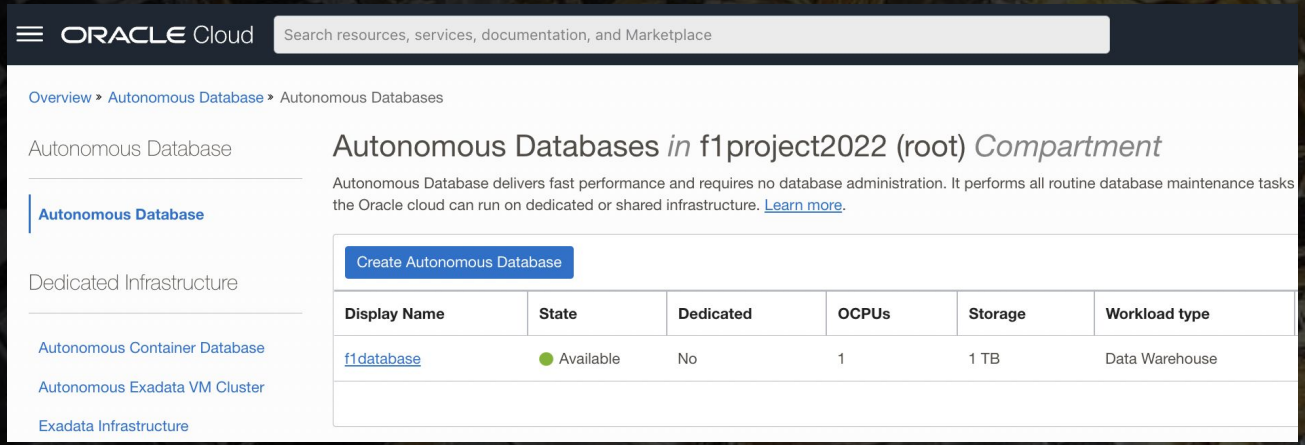


Fig 1. Autonomous Databases created in Oracle - Database

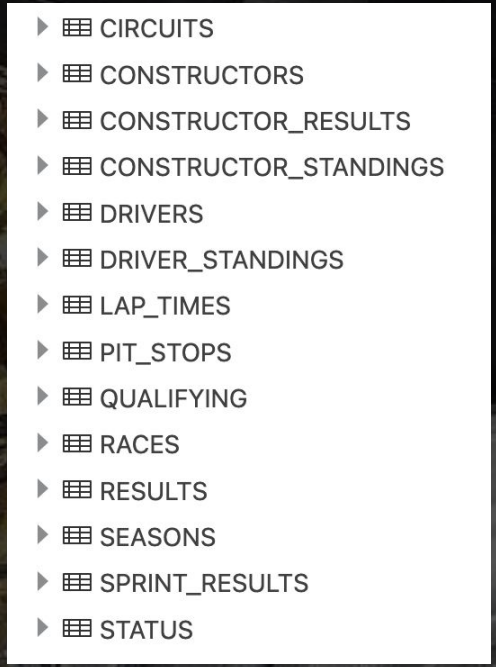
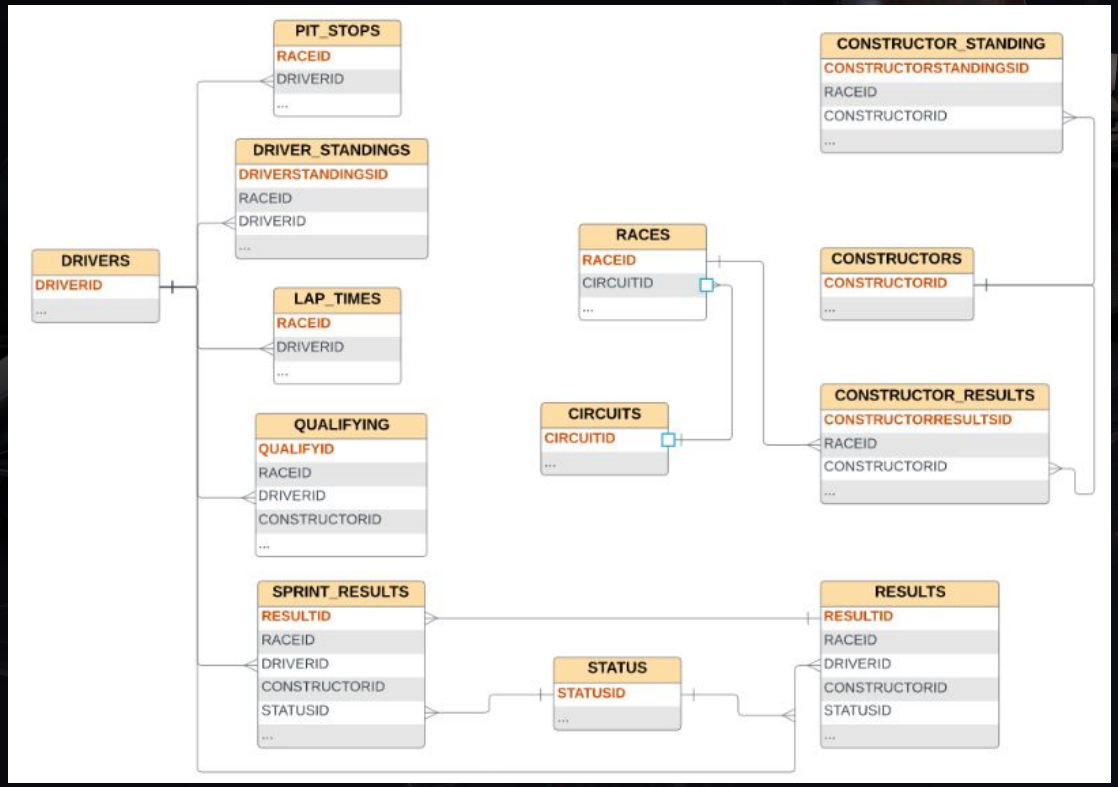


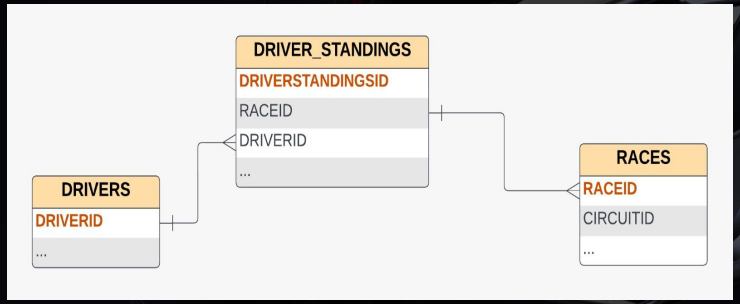
Fig 2. Files in database

DATA SCHEMA

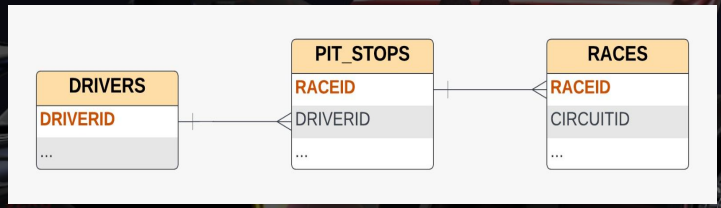

PIT_STOPS
RACEID
DRIVERID
...
CONSTRUCTOR_STANDING
CONSTRUCTORSTANDINGSID
RACEID
CONSTRUCTORID
...
DRIVER_STANDINGS
DRIVERSTANDINGSID
RACEID
DRIVERID
...
RACES
RACEID
CIRCUITID
...
DRIVERS
DRIVERID
...
CONSTRUCTORS
CONSTRUCTORID
...
LAP_TIMES
RACEID
DRIVERID
...
CONSTRUCTOR_RESULTS
CONSTRUCTORRESULTSID
RACEID
CONSTRUCTORID
...
CIRCUITS
CIRCUITID
...
QUALIFYING
QUALIFYID
RACEID
DRIVERID
CONSTRUCTORID
...
SPRINT_RESULTS
RESULTID
RACEID
DRIVERID
CONSTRUCTORID
STATUSID
...
RESULTS
RESULTID
RACEID
DRIVERID
CONSTRUCTORID
STATUSID
...
STATUS
STATUSID
...

DATA MODELS

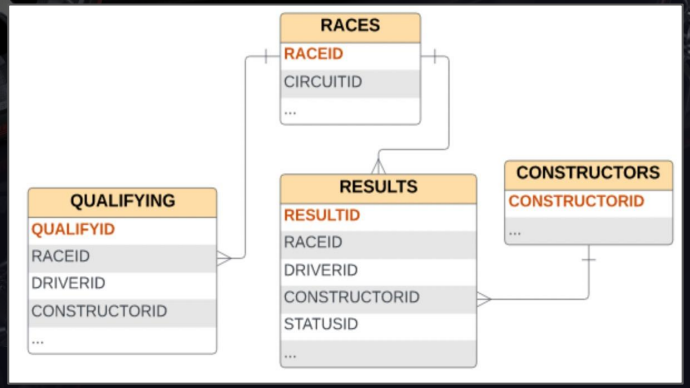
Application 1 : Driver Perspective



Application 2 : Race Engineering Perspective

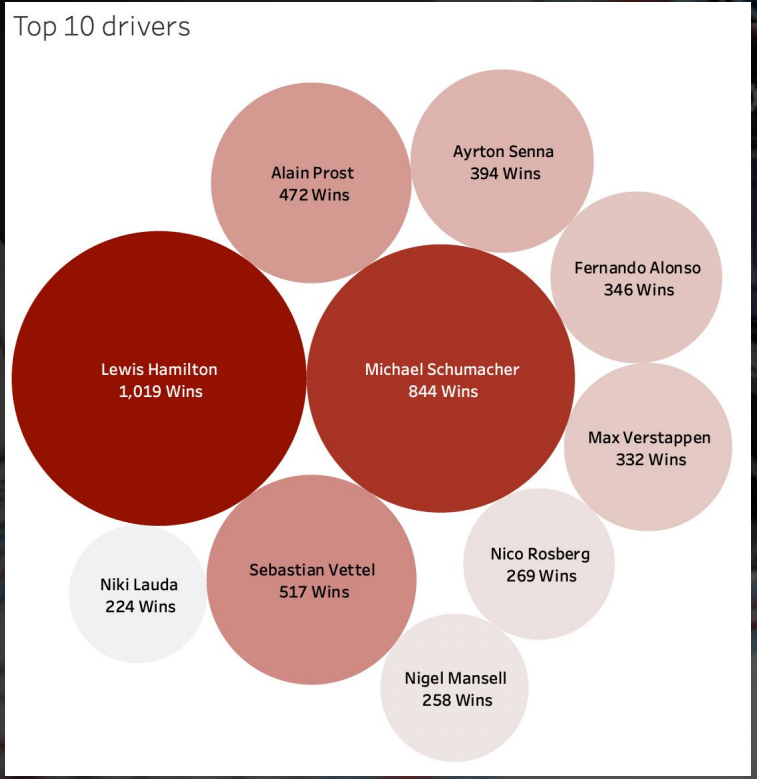


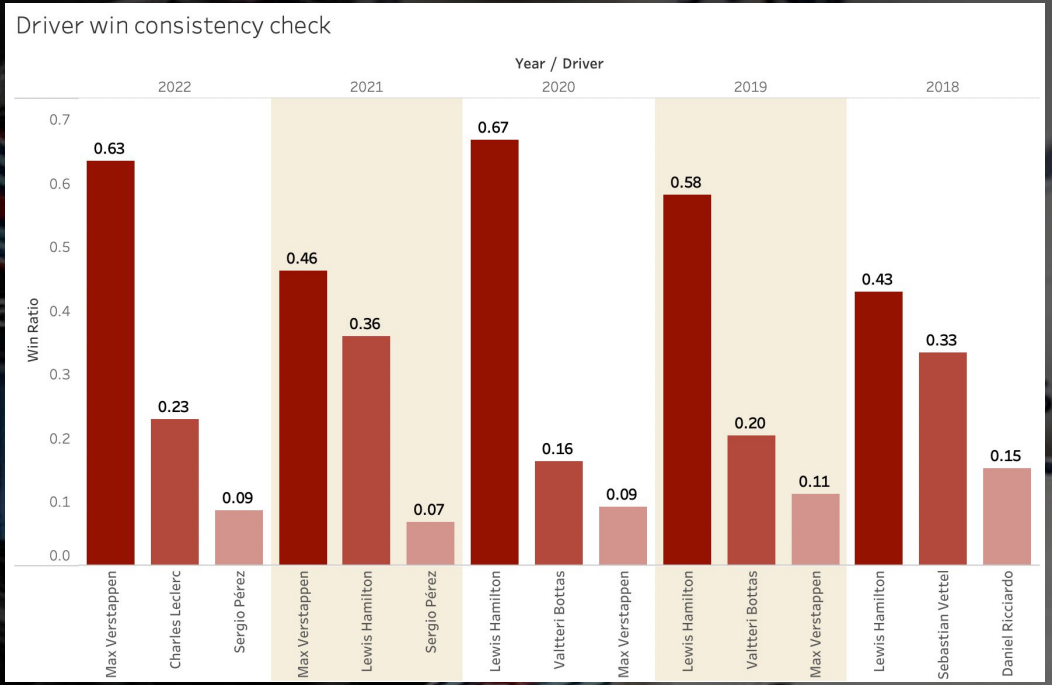
Application 3 : Race Strategy Perspective



DRIVER PERSPECTIVE

Who should we sign?

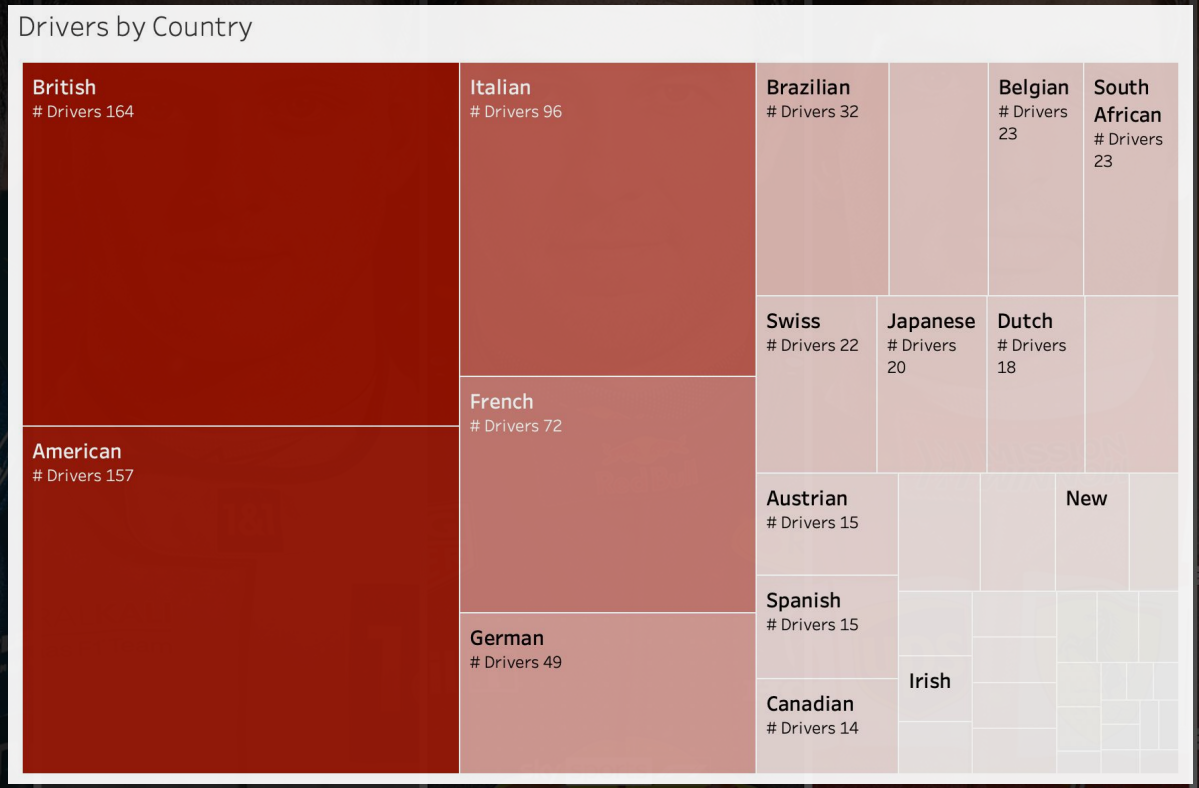




- The drivers with the most wins are Lewis Hamilton, Michael Schumacher, and Sebastian Vettel, respectively
- Lewis Hamilton dominates F1 from 2018-2020, but from 2021 onward, it seems that the torch has been passed to Max Verstappen, on a win ratio basis

DRIVER PERSPECTIVE

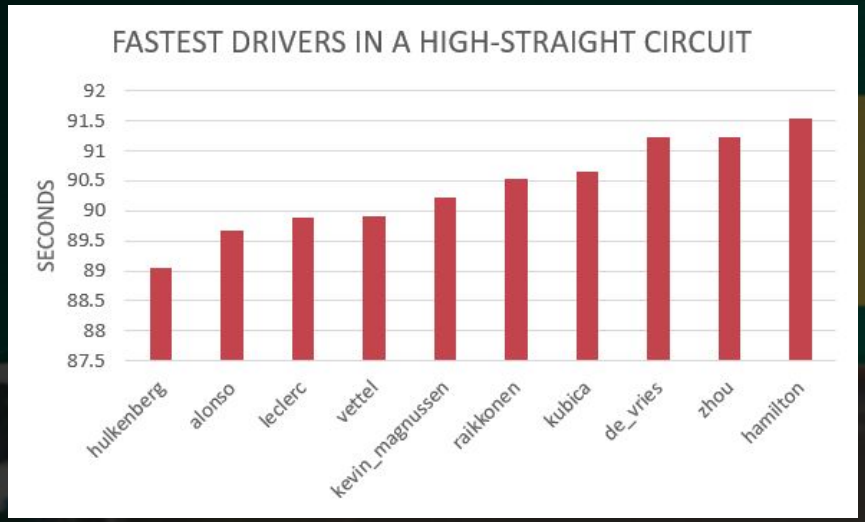
Who should we sign?



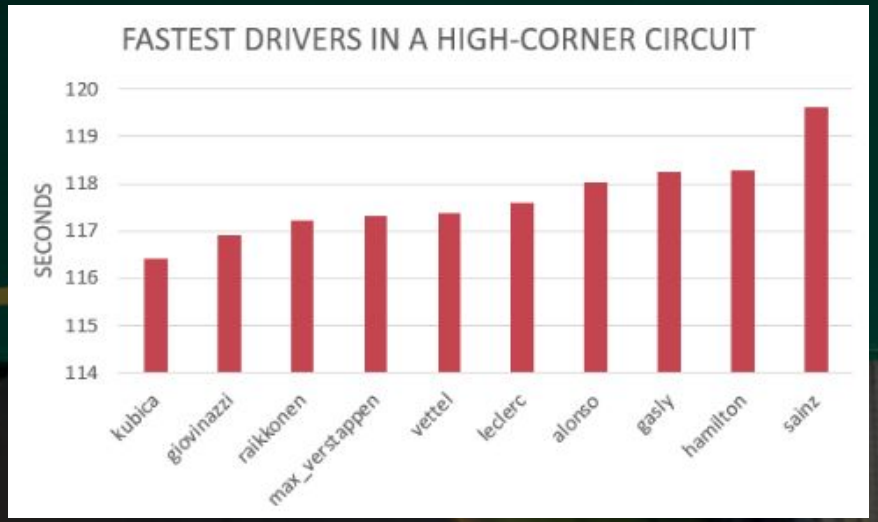
DRIVER PERSPECTIVE

Who should we sign?

Baku - Azerbaijan



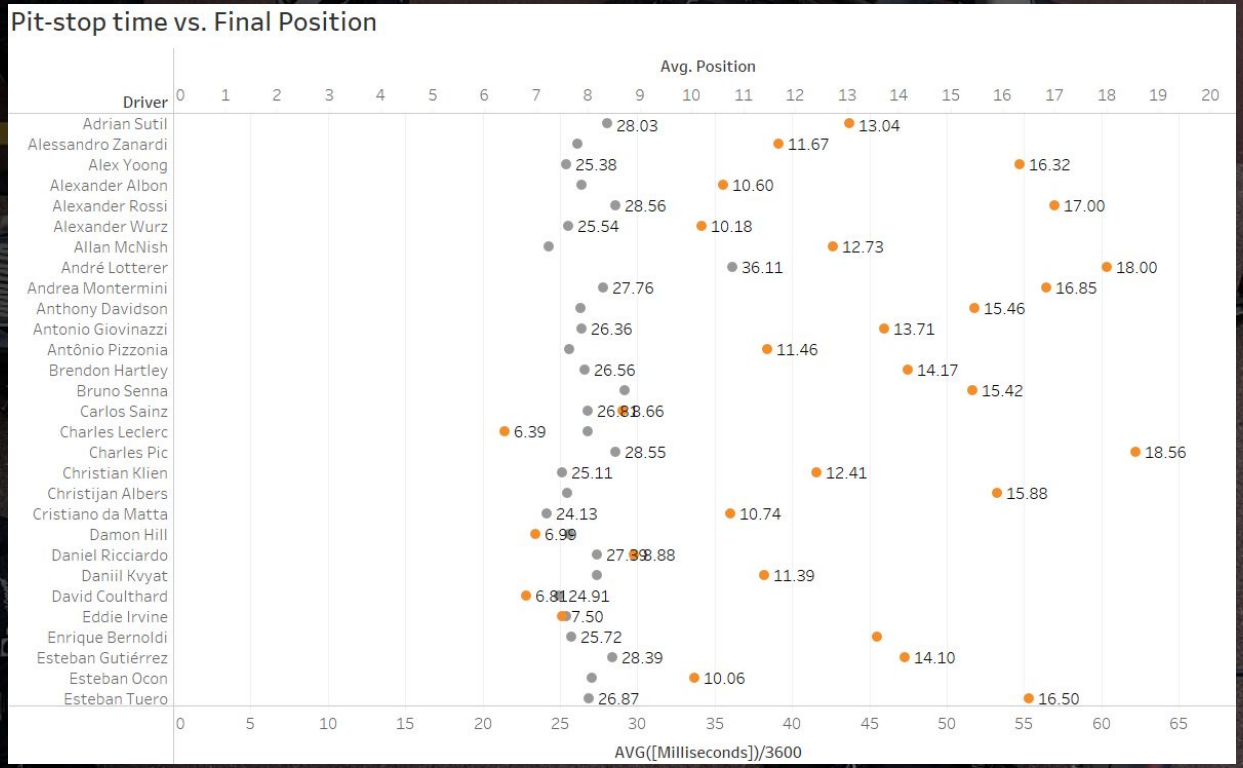
Monza - Italy



RACE ENGINEERING PERSPECTIVE

Should we focus on pit-stop?

- Plot shows if faster pit-stops translate to better final result
- Only a few drivers are able to achieve better result with a faster pit-stop
- Only focus efforts on pit-stops if we buy a driver who can perform well on fast pit-stops

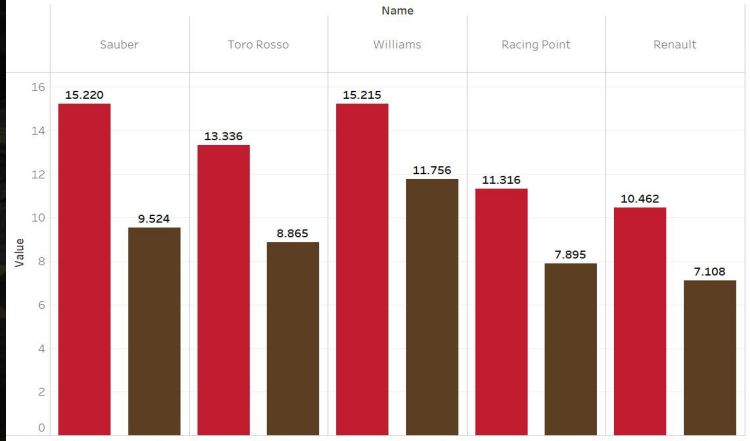


RACE STRATEGY PERSPECTIVE

Determining success in Qualifying round based on the actual race?

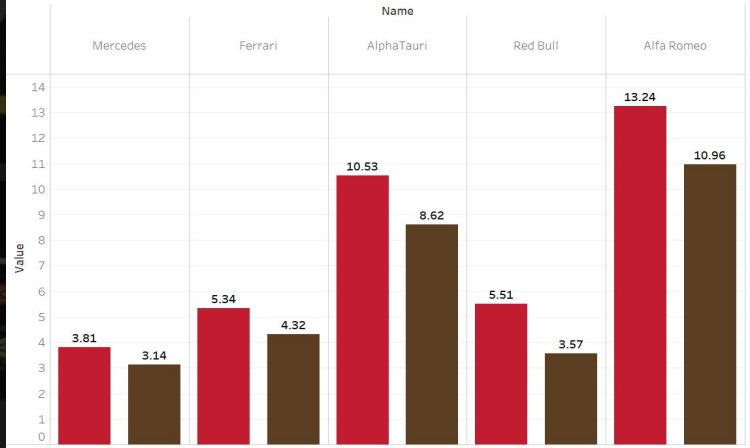
- In last 5 years, the average qualifying position called grid and average position in final race is used to compute the relative position
- Sauber won the final race irrespective of the loss in the qualifying round
- Mercedes and Ferrari lost the final race despite win in the qualifying round

Top 5 Constructors - Performance in Final Round relative to performance in Qualifying Round



Avg. Grid and Avg. Position for each Name. Color shows details about Avg. Grid and Avg. Position. The marks are labeled by Avg. Grid and Avg. Position. The view is filtered on Name, which keeps Racing Point, Renault, Sauber, Toro Rosso and Williams.

Bottom 5 Constructors - Performance in Final Round relative to performance in Qualifying Round



Avg. Grid and Avg. Position for each Name. Color shows details about Avg. Grid and Avg. Position. The marks are labeled by Avg. Grid and Avg. Position. The view is filtered on Name, which keeps Alfa Romeo, AlphaTauri, Ferrari, Mercedes and Red Bull.

CHALLENGES

- Our data was huge, so we were not able to use the ORACLE SQL Server provided in the class
- We were unable to use Google Cloud platform since it does not support referential integrity and is only used for LDM (Logical Data Modeling).
- Since we are using Oracle Autonomous Database, we were not able to communicate across databases. Hence we had to create our aggregated and raw tables in the same databases.
- We had to perform a lot of data preprocessing to extract meaningful insights out of it
    - When a driver doesn't finish a race, the position is noted as null
    - Date format was not supportive

LEARNING

CONCLUSION

Driver Perspective

- Using the last five years on an equal basis, we would buy Lewis Hamilton as our driver
- However, if we view the most recent year's performance as an indicator of future performance, we would instead choose to buy Max Verstappen
- If we were to do this analysis again, we would try to see how consistent a change is in the leaderboard year after year

Race Engineering Perspective: Pit-stop Timing v. Final result

- The relationship between pit-stop time and average position cannot easily be detected
  - In certain scenarios, a small decrease in pit-stop time may result in a change in results
  - However, there are cases in which someone has a very quick pit-stop, but they remain in last place, so that very fast pit-stop has little-to-no impact on the driver's final position
- If we were to do this analysis in the future, it may be best to find patterns between pit-stop time and final position of only the drivers in the top-10

Race Strategy Perspective: Importance of Qualifying Round

- We found that among the top teams, qualifying times are an important indicator of their final race positions
- However, for the midfield, qualifying is not a strong indicator of final race position


THANK YOU