

Formula 1 - World championship dataset

Source: Dataset from Kaggle: <https://www.kaggle.com/datasets/rohanrao/formula-1-world-championship-1950-2020>

#SUMMARY

1. DDL / Data Definition Language

1.1. Preparation and processing of data / Normalise Tables / Change Datatypes

1.2. Preparation and processing of data / Primary Key, Foreign Key

2. Defining the problem we are trying to solve

3. DML / Data Manipulation Language / Aggregate Functions / Analyzing the data

4. Share the insights from the data (dashboards, graphs)

2. Defining the problem we are trying to solve

General Info about F1

1. Find the country that hosted the F1 races.
2. Find the circuit that hosted the F1 races and which countries join the races.

About the racer

3. Find the avg/max/min time difference between the 1st, 2nd and 3rd runners.
4. Find the constructor that has won in F1 (See if there's any constructor wins the most in F1)
5. Find the driver that has won the races in F1 and the respecting constructor
6. Find the fastest stop in history.

Investigate the relationships

7. Compare the nationality of the champion and the hosting country
8. Find the F1 top speed record, its correlated circuit and its constructor
9. Find the avg/sum/count of the pole position (q1, q2, q3) and the winners
10. Correlation between duration of pit stops and time to win race/reach the end line.
11. Find the avg pit stop duration by each constructor in each year
12. Find the avg pit stop by circuit
13. The curse of number 13*
 - >Does number 13 lost the most races?
 - >Does number 13 have more accidents/disqualify than other racers?