

Análise de dados da Formula1



Introdução:

A Formula1, popularmente abreviada como F1, é a mais prestigiada categoria do automobilismo mundial. Ela combina alta tecnologia, velocidade extrema e competição acirrada, atraindo pilotos e equipes do mais alto calibre!. As corridas são realizadas ao redor do mundo, desde ruas de cidade até pistas especialmente construídas, cada uma apresentando desafios únicos para as equipes. Além da pura competição, a F1 é também um campo de desenvolvimento tecnológico, com inovações frequentes em áreas como aerodinâmica, materiais e eficiência energética. Com uma história rica que remonta aos anos 1950, a Fórmula 1 não é apenas um esporte, mas um espetáculo global que cativa milhões de fãs.

Objetivos:

O objetivo deste estudo é realizar uma Análise Exploratória dos Dados (Exploratory Data Analysis - EDA) do conjunto de dados da Formula1 (F1) adaptado pelo autor e disponível em : <https://data.world/wjsutton/formula-1-race-data/workspace/file?filename=Formula1.sqlite> a fim de levantar dados da modalidade.

Dicionário de dados:

O banco de dados "Formula1.sqlite" contém as seguintes tabelas, que representam diferentes aspectos das corridas de Fórmula 1:

- **circuits:** Informações sobre os circuitos onde as corridas são realizadas.
- **races:** Detalhes sobre cada corrida, incluindo a data e o circuito.

- **driver_standings:** Classificação dos pilotos ao longo das temporadas.
- **drivers:** Informações sobre os pilotos, como nome, nacionalidade e data de nascimento.
- **constructors:** Detalhes sobre as equipes, incluindo nome e nacionalidade.
- **results:** Resultados de cada corrida, incluindo a posição de chegada de cada piloto.
- **constructor_standings:** Classificação das equipes ao longo das temporadas.
- **constructor_results:** Resultados das equipes em cada corrida.
- **laptimes:** Tempos de volta para cada piloto em cada corrida.
- **pitstops:** Detalhes sobre os pitstops realizados durante as corridas.
- **qualifying:** Resultados das sessões de qualificação.
- **seasons:** Informações sobre cada temporada de Fórmula 1.
- **status:** Status dos pilotos em cada corrida (por exemplo, terminou, abandonou).

Plano de ação de análise:

Análise dos Pilotos e Equipes:

Vamos analisar o desempenho dos pilotos e das equipes ao longo do tempo. Isso incluirá olhar para as classificações dos pilotos e das equipes ao longo das temporadas.

Análise de Circuitos:

Investigaremos como o desempenho dos pilotos varia em diferentes circuitos. Isso pode revelar se alguns pilotos ou equipes são mais fortes em certos tipos de circuitos.

Influência das Condições Climáticas:

Se as informações estiverem disponíveis, analisaremos como as condições climáticas afetam os resultados das corridas.

Tabela dos Pilotos (drivers):

Contém informações sobre os pilotos, incluindo nome, nacionalidade e data de nascimento.

Tabela dos Circuitos (circuits):

Fornece detalhes sobre os circuitos, como localização e país.

Tabela dos Resultados (results):

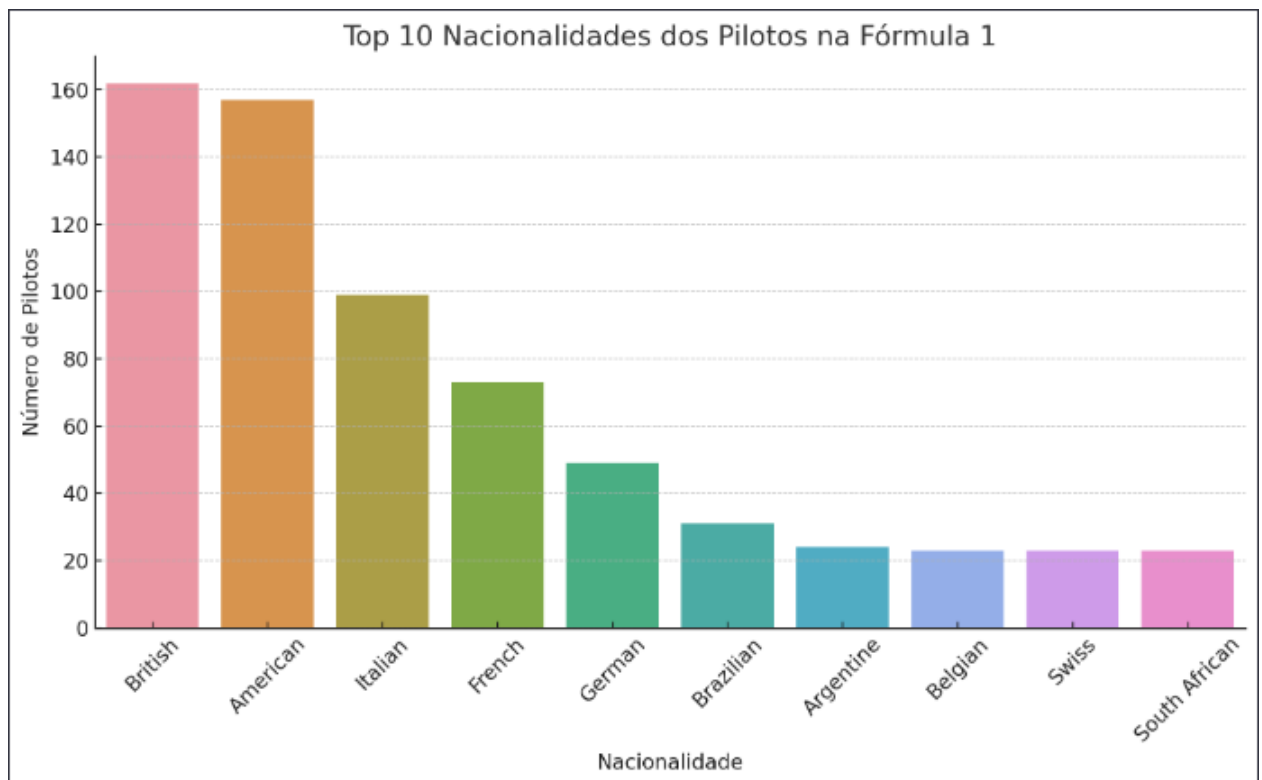
Apresenta os resultados das corridas, incluindo a posição de chegada, pontos conquistados e voltas mais rápidas.

Tabela das Corridas (races):

Detalha cada corrida, incluindo a data e o circuito onde foi realizada.

Com base nesses dados, podemos começar algumas análises. Por exemplo, podemos explorar:

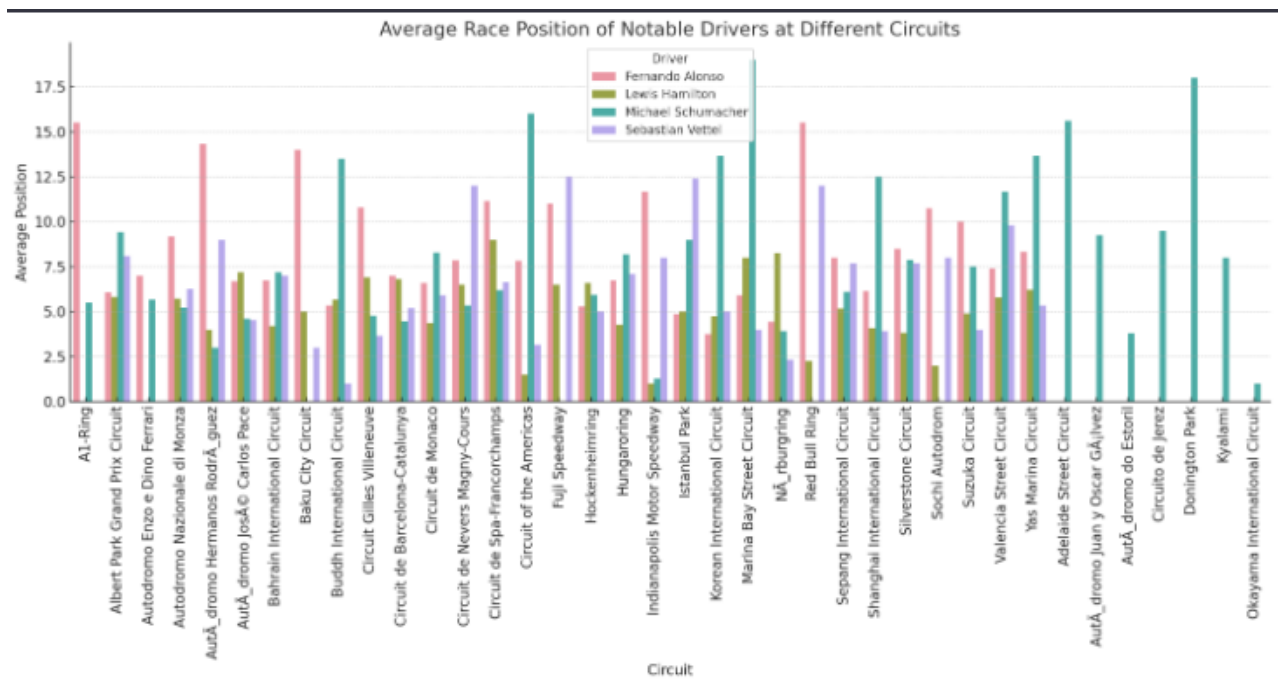
- A distribuição das nacionalidades dos pilotos.
- O desempenho dos pilotos em diferentes circuitos.
- As equipes com mais vitórias ao longo do tempo.



Select utilizado:

```
query_nationalities = """ SELECT nationality, COUNT(*) as num_pilots FROM drivers  
GROUP BY nationality ORDER BY num_pilots DESC; """
```

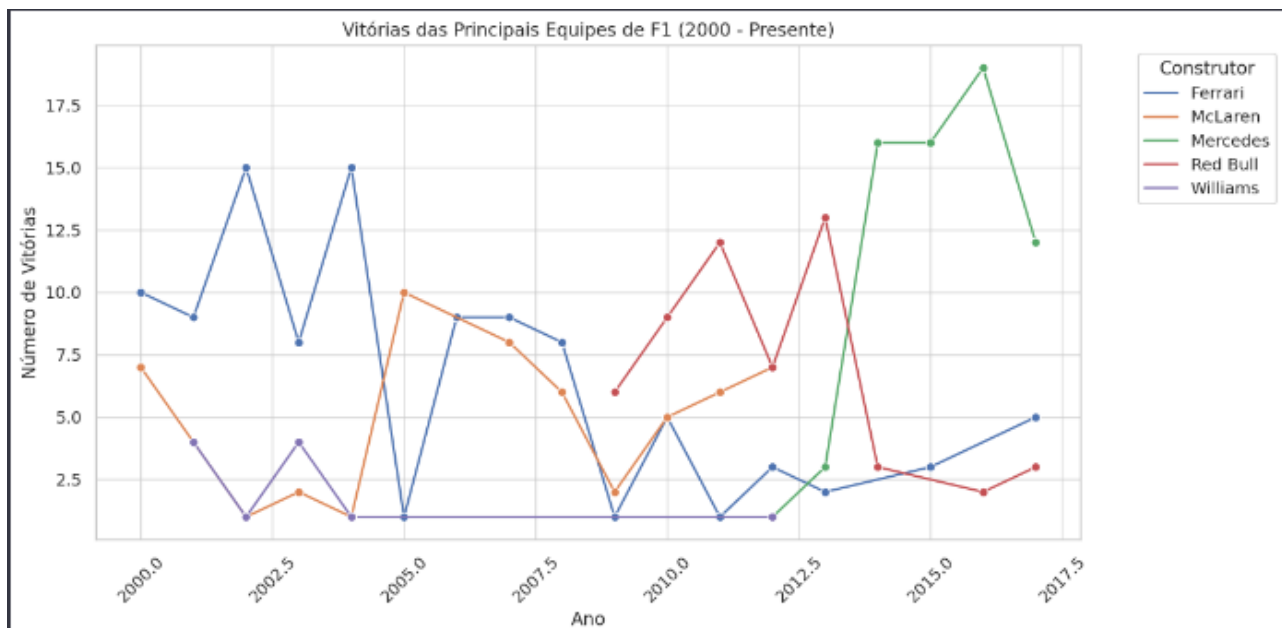
A partir deste dado vemos que os britânicos tomam posse da nacionalidade de pilotos da Fórmula 1, o que será que eles têm no sangue?



Aqui podemos observar 4 pilotos e suas posições nas pistas.

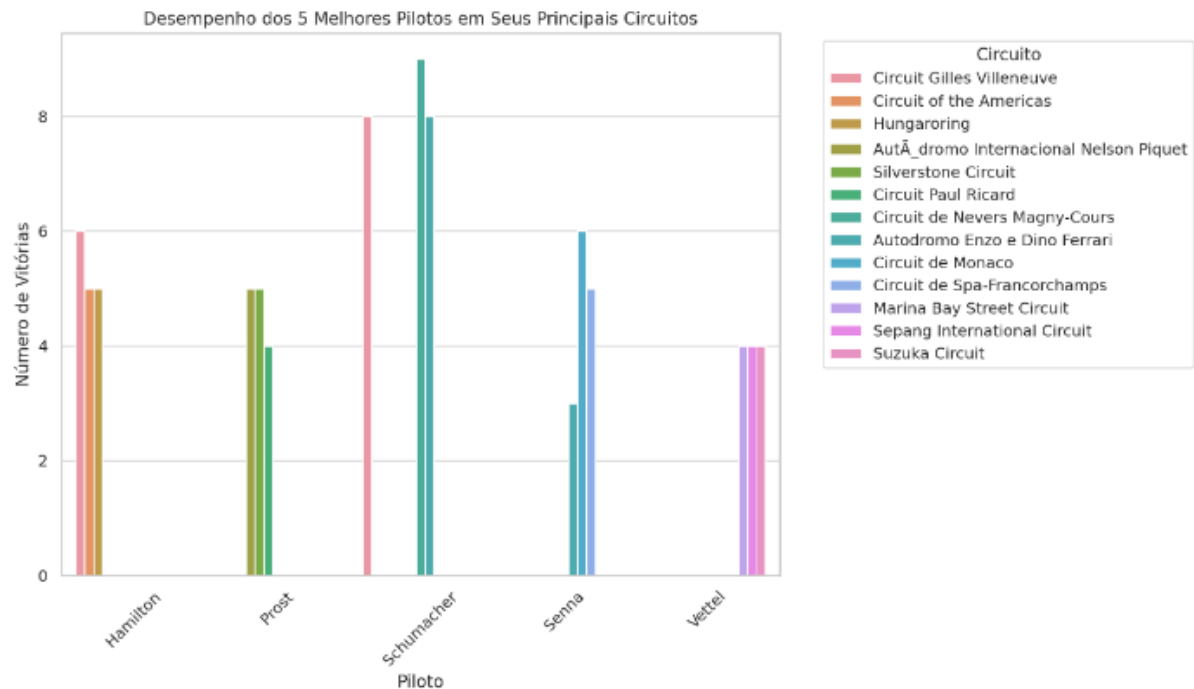
Select utilizado:

```
performance_query = """ SELECT dr.forename || ' ' || dr.surname as Driver, ci.name as
Circuit, r.positionOrder as Position, ra.year as Year FROM drivers dr JOIN results r
ON dr.driverId = r.driverId JOIN races ra ON r.raceId = ra.raceId JOIN circuits ci ON
ra.circuitId = ci.circuitId ORDER BY dr.surname, ra.year; """
```



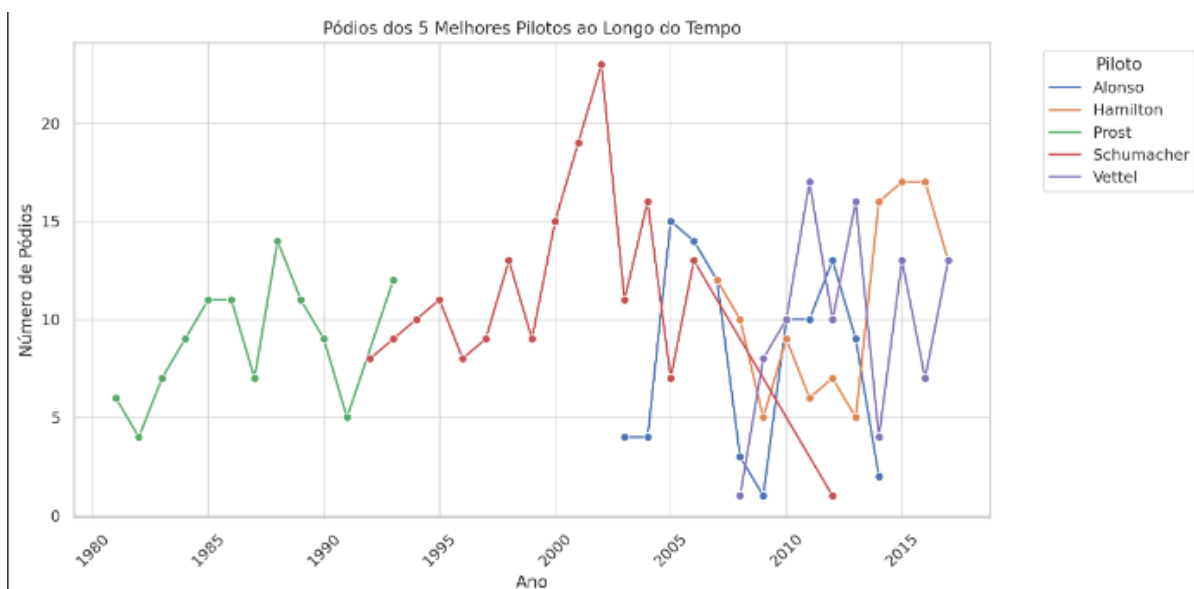
Select utilizado:

```
victories_query = """ SELECT c.name as Constructor, COUNT(r.position) as Victories,
ra.year as Year FROM constructors c JOIN results r ON c.constructorId =
r.constructorId JOIN races ra ON r.raceId = ra.raceId WHERE r.position = 1 GROUP BY
c.name, ra.year ORDER BY c.name, ra.year; """
```



Select utilizado;

```
performance_query = """ SELECT d.surname as Driver, c.name as Circuit, COUNT(r.position) as Wins
FROM drivers d JOIN results r ON d.driverId = r.driverId JOIN races ra ON r.raceId = ra.raceId JOIN
circuits c ON ra.circuitId = c.circuitId WHERE r.position = 1 GROUP BY d.surname, c.name ORDER BY
d.surname, c.name; """
```

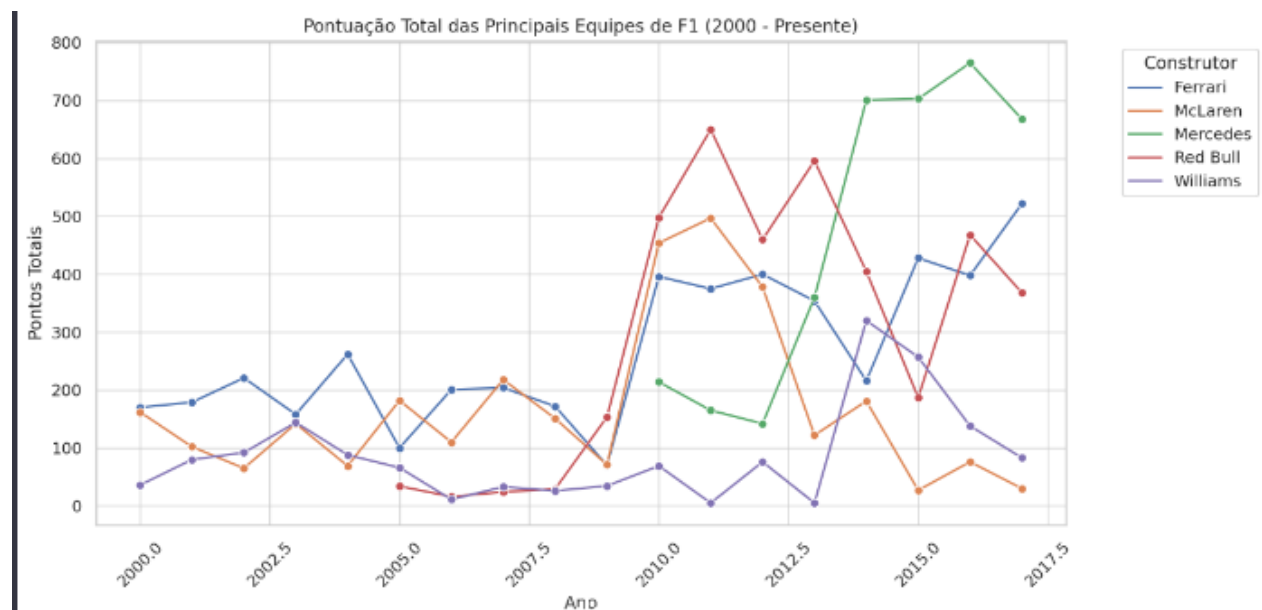


Select utilizado:

```

SELECT d.surname as Driver, ra.year as Year, COUNT(r.position) as Podiums
FROM drivers d
JOIN results r ON d.driverId = r.driverId
JOIN races ra ON r.raceId = ra.raceId
WHERE r.position IN (1, 2, 3)
GROUP BY d.surname, ra.year
ORDER BY d.surname, ra.year;

```



Select utilizado

```

SELECT c.name as Constructor, ra.year as Year, SUM(r.points) as TotalPoints
FROM constructors c
JOIN results r ON c.constructorId = r.constructorId
JOIN races ra ON r.raceId = ra.raceId
GROUP BY c.name, ra.year
ORDER BY c.name, ra.year;

```