

Visualização de Dados



Integrantes:

Luís Fernando Bastos Rego - 470043

Samuel Vieira de Paula Farias - 498844

SUMÁRIO

3 INTRODUÇÃO

4 DESCRIÇÃO DOS DADOS

5 TECNOLOGIAS USADAS

6 DIVISÃO DO TRABALHO

7 DIFICULDADES

8 LIÇÕES APRENDIDAS

9 CONCLUSÃO

INTRODUÇÃO

A Formula 1 é uma competição de corrida realizada pela Federação Internacionais de Automóveis (FIA), que ocorre desde 1950.

É um evento que possui muita história por todos os anos que ele existiu e que todos conhecem.

Nesse trabalho abordamos os dados estatísticos da Formula 1 para mostrar um panorama geral sobre eles.



DESCRIÇÃO DOS DADOS

O dataset foi composto por Rohan Rao, utilizando dados disponibilizados pela Ergast Developer API, e está disponível para no site Kaggle.

São no total 14 arquivos do formato .csv, com mais de 1000 corridas catalogadas, e quase 1000 pilotos, com seus devidos circuitos e equipes. Os dados de cada arquivo estão conectados por meio de chaves estrangeiras, que ligam as tabelas por id.

Neles é possível encontrar diversas informações sobre circuitos, corridas, equipes e pilotos.

TECNOLOGIAS UTILIZADAS

- Python (Pandas)
- Javascript (Vega-Lite)
- Observable
- Google Colab
- Github



DIVISÃO DO TRABALHO

Os membros se dividiram da seguinte maneira:

Luis: Planejar as visualizações;

Samuel: Realizar pre-processamento ;

Ambos fizeram a execução das visualizações e a criação da webpage.

DIFICULDADES

- Manipular as tabelas entre si, pois cada junção de tabela gerava arquivos pesados, próximos do limite do site Observable.
- Produzir a interatividade, por conta de comportamentos inesperados do Vega-lite que não soubemos tratar.
- Construção dos mapas, que precisou de um pré-processamento e coordenação das tabelas e com o arquivo topojson utilizado.

LIÇÕES APRENDIDAS

- Aprendemos a filtrar e tratar os dados que precisam de pré-processamento por meio do Pandas.
- Aprendemos a analisar melhor os dados e aumentamos nossas criatividades em construir sobre eles visualizações mais interessantes e informativas.

SCREENSHOTS

Localização dos Circuitos e Seu Número de Corridas

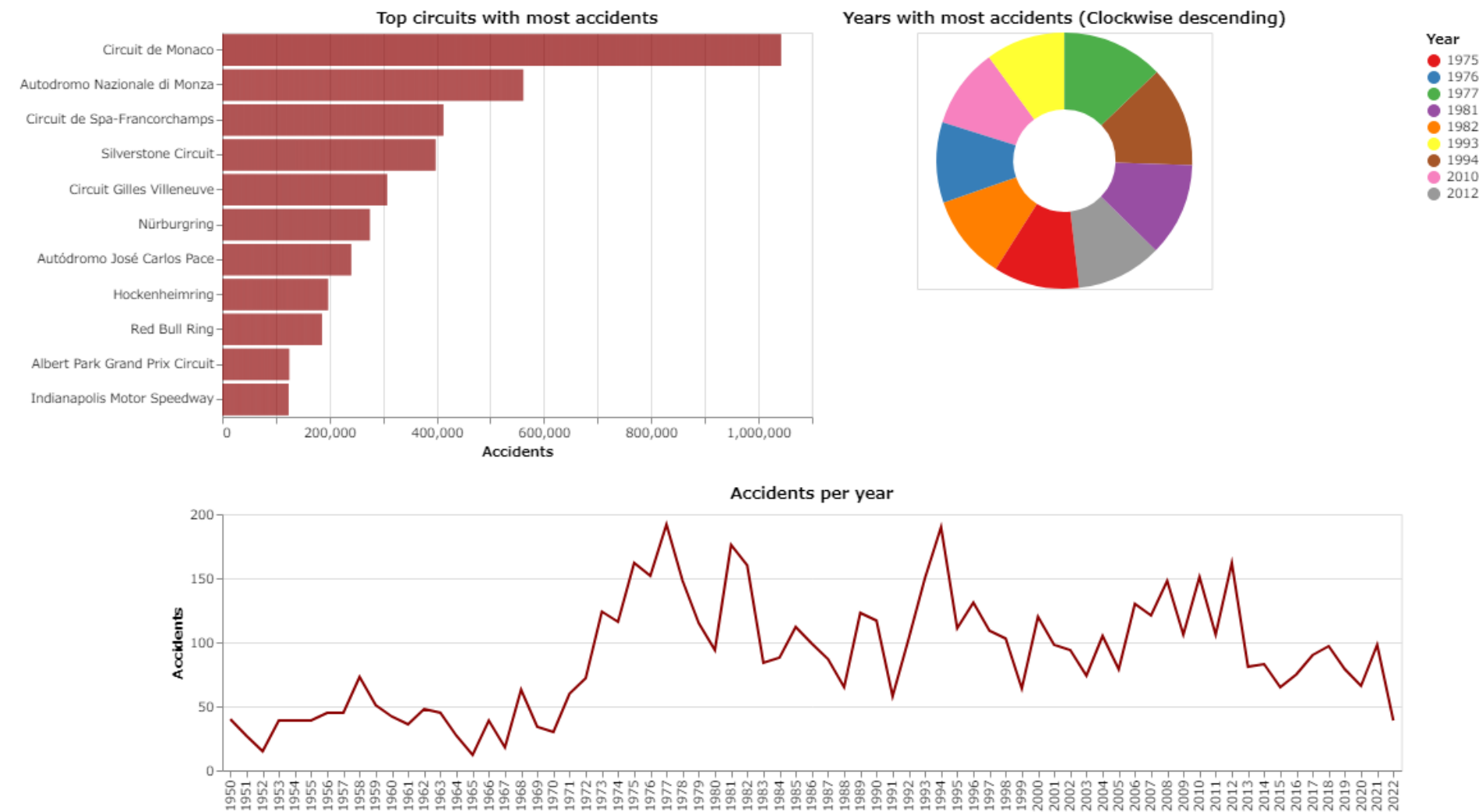
Aqui mostramos os circuitos por localização geográfica, mostrando também a imagem de como é seu traçado, como também o número de corridas do circuito.



SCREENSHOTS

Circuitos com mais acidentes:

Agora, abordamos algumas visualizações a mais sobre os circuitos e seus acidentes. Para calcular esse valor, utilizamos as tags 3 e 4 que indicam acidentes e colisões, dentro da tabela *status*.



SCREENSHOTS

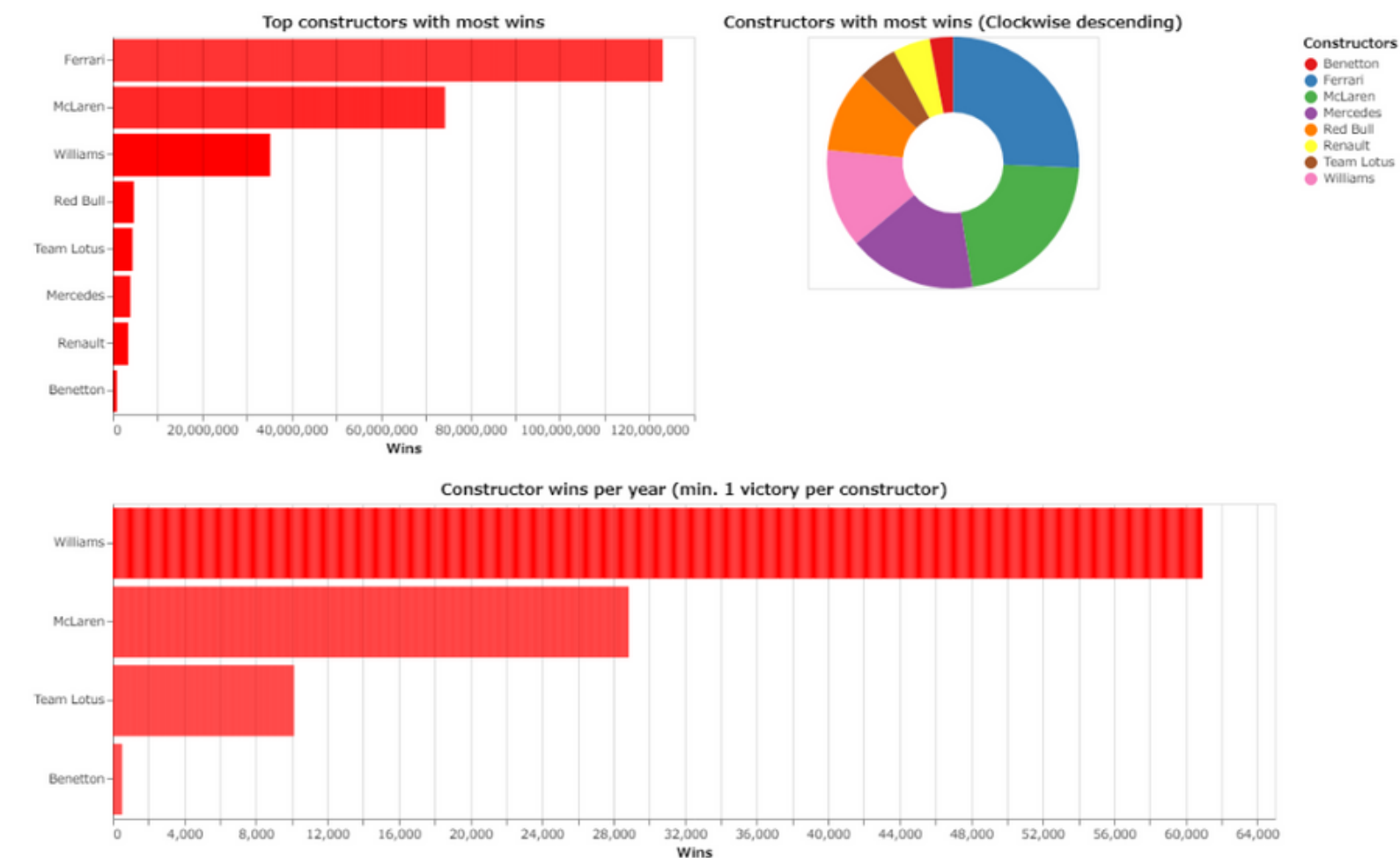
Equipes com mais Vitórias

Agora, mostramos as equipes com mais vitórias em geral, com também uma visão parte-todo, além de uma visualização interativa com as equipes mais vencedoras de cada ano. É possível c

Ano:



Entre 1950 e 2022



SCREENSHOTS

Pilotos com mais vitórias

Similar a tabela anterior, visualizamos os pilotos com suas vitórias (encontradas em driver_standings), com vitórias em geral e com uma visualização interativa que mostra o piloto com mais vitórias a cada ano.

Ano:



Entre 1950 e 2022

