

Ciência de dados e Big Data

Análise e Limpeza dos Dados

- Leitura da base de dados (CSV/XLSX).
- Tratamento de valores faltantes.
- Conversão de colunas de datas.
- Criação de colunas derivadas.

Modelagem

- Técnica selecionada: Regressão Linear para prever o preço com base na distância.
- Features utilizadas: distância, tipo de veículo, dia da semana.
- Variável target: preço da corrida (R\$).

Foi construído um dashboard com os seguintes gráficos:

- Dispersão: Preço vs Distância
- Média de Preço por Tipo de Veículo
- Distribuição de Corridas por Veículo (pizza)
- Evolução do preço ao longo do tempo
- Frequência de faixas de preço
- Boxplot de preços por tipo de veículo
- Evolução acumulada do preço

Filtros disponíveis:

- Tipo de veículo
- Dia da semana

Filtrar por tipo de veículo:

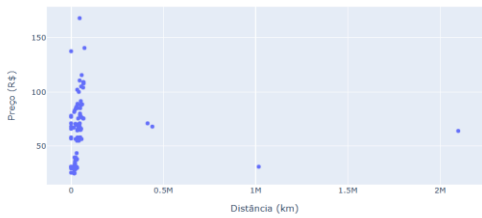
uberX

Filtrar por dia da semana:

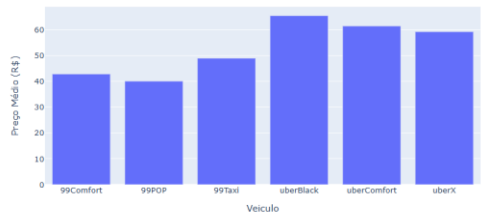
Terça-feira

Dashboard de Corridas

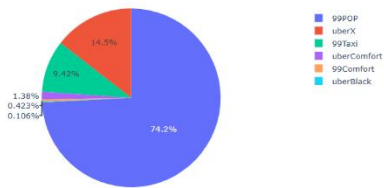
Distância vs Preço (uberX - Terça-feira)



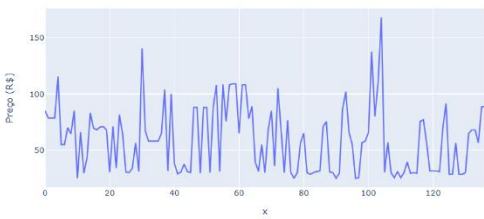
Média de Preço por Veículo (Terça-feira)



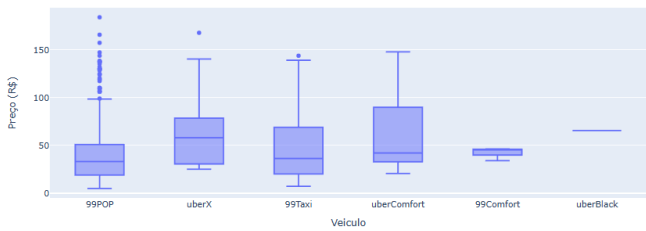
Distribuição das Corridas (Terça-feira)



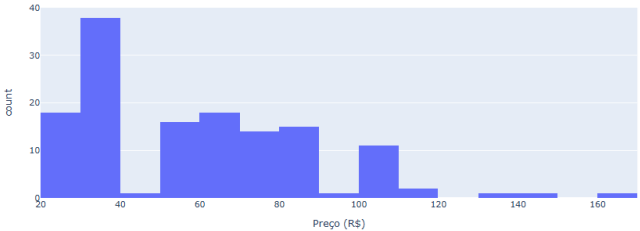
Preço ao longo do tempo (uberX - Terça-feira)



Distribuição de Preços por Veículo (Terça-feira)



Frequência de Corridas por Faixa de Preço (uberX - Terça-feira)



Evolução Acumulada de Preço (uberX - Terça-feira)

