Introdução

A mobilidade urbana é um dos principais problemas da sociedade atual, especialmente em grandes cidades. Assim sendo, a importância do papel desempenhado pelos aeroportos é maior do que nunca, assegurando a interligação entre cidades, países ou mesmo de continentes, impactando diretamente, tanto a economia, quanto a experiência dos passageiros. Problemas que levem a atrasos de voos, ou até mesmo cancelamentos, tais como atrasos em check-ins, sobrelotação dos aeroportos, ou até mesmo devido a condições climatéricas, afetam não só os passageiros, quanto as companhias, quer de forma monetária, quer a sua reputação.

Compreender os fatores que contribuem para atrasos e cancelamentos de voos, as causas, e como estes influenciam os preços dos bilhetes, pode ajudar na procura de soluções mais eficientes.

**Objetivos do projeto**:

O projeto tem como objetivo analisar as possíveis causas dos atrasos e cancelamentos de voos e as suas possíveis causas com os preços dos mesmos.

* Determinar aeroportos com maior quantidade de atrasos e possíveis causas
* Determinar aeroportos com mais voos cancelados e possíveis causas
* Analisar os preços dos bilhetes nos aeroportos mencionados
* Analisar tempo médio de atraso por aeroporto vs distância média por origem

Metodologia

De forma a conseguir resultados mais consistentes, foram utilizados 2 datasets com dados de voos dos EUA. Um dos datasets contém dados de voos como, trimestre, origem do voo, destino, preço, etc., sendo estes dados do ano 2018. Como segunda fonte de dados, foi utilizado um dataset com detalhes de voos, atrasos e cancelamentos.

Após a importação dos dados, foram selecionadas apenas as features consideradas úteis, em ambos os datasets, eliminando as restantes. No dataset dos atrasos, foi feita a seleção dos voos apenas de 2018, de forma a ir de encontro ao outro dataset. Aplicámos filtros de limpeza em ambos os datasets, eliminando linhas duplicadas ou com valores nulos.