

Case técnico - Data Person - Cumbuca

Lucas Ketzer

Mai/2022

Processando e explorando dados em um banco relacional - nycflights13

em media, o atraso medio diario tende a ser menor em setembro, outubro e novembro (outono)

tende a ser maior em junho, julho (inicio do verao) e dezembro (onde o atraso medio diario tende a ter mais valores em uma faixa maior, possivelmente por causa das festas)

de janeiro a maio tende a se distribuir de maneira similar

(boxplot e grafico em barras)

desvio padrao p/ atraso medio tende a ser particularmente alto nos meses com menor atraso, tende a ser menor nos meses com maior atraso e relativamente constante nos outros

o desvio padrao eh maior q o atraso medio pra todos os dias do ano

entre os 10% maiores atrasos, 10% sao de mais de 217 minutos, e entre o 1% dos maiores atrasos, apenas 10% destes sao de mais de 471 minutos

segundas e quintas tendem a ter o % de dias com mais atrasos extremos

a terceira semana de junho tende a ter um percentual alto de atrasos extremos, assim como a primeira semana de julho

entre 25% dos atrasos que ocorrem em um dia, apenas 10% nao sao negativos (ou seja, 90% de um quarto dos tempos de voos registrados chegam adiantados)

entre um meio dos tempos de voos registrados, aproximadamente 30% chegam adiantados

essa distribuicao explica a alta variabilidade nos meses ao longo do ano, com os meses com maior atraso possivelmente registrando mais casos extremos do que o normal (o mesmo vale para meses com menor atraso)

algumas limitacoes: a variabilidade dos dados (apesar da amostra ser relativamente uniforme para os meses), nao diferenciar alguns fatores entre os atrasos que poderiam influenciar os meses (alguns poderiam ter mais viagens internacionais) e fatores climaticos

REFERENCIAL ATRASOS: 41% REFERENCIAL ADIANTAMENTOS: 58%

menos atrasos Alaska Airlines Inc. (26%) American Airlines Inc. (33%) Delta Airlines Inc. (34%) Endeavor Air Inc. (39%) Hawaiian Airlines Inc. (25%) SkyWest Airlines Inc. (27%) US Airways Inc. (33%) Virgin America (36%)

mais atrasos AirTrain Airways Corporation (58%) Envoy Air (43%) ExpressJet Airlines Inc. (48%) Frontier Airlines Inc. (60%) JetBlue Airways (44%) Mesa Airlines Inc. (47%) Southwest Airlines Co. (48%) United Air Lines Inc. (42%)

criando grupos para as companhias aereas:

grupos 1 (Frontier Airlines Inc. e AirTran Airways Corporation) concentra os maiores atrasos

2 (Southwest Airlines Co., ExpressJet Airlines Inc., Mesa Airlines Inc.) e 3 (JetBlue Airways, Envoy Air, United Air Lines Inc., Endeavor Air Inc.) concentram um percentual similar de atrasos, tendendo a ser mais altos

4 (American Airlines Inc., US Airways Inc., Delta Air Lines Inc, Delta Air Lines Inc., Virgin America) concentra um nível um pouco abaixo de atrasos

5 (SkyWest Airlines Inc., Alaska Airlines Inc.) e 6 (Hawaiian Airlines Inc.) concentram um nível consideravelmente baixo de atrasos

ao calcular a estatística IV, existe evidência fraca de que há discriminação entre o valor dos grupos