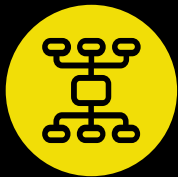


# NY FLIGHTS

**Estudo de atraso de voos**  
**Análise estatística com Python**  
**[1h50]**

# Dinâmica New York Flights

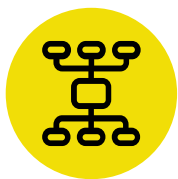


## CONTEXTUALIZAÇÃO

Vocês atuam na NY Flights, uma companhia aérea de nível global, e foram questionados pela gestão sobre o desempenho das linhas aéreas.

A empresa está extremamente preocupada com atraso em seus voos, alguns dados estatísticos sobre o tema foram solicitados para o time de ciência de dados da companhia (vocês)

O diretor do time solicitou algumas análises para a reunião, pois planos de ações serão definidos a partir destes dados da análise.



# INSTRUÇÕES

Vocês farão as análises em 3 etapas e em cada uma delas estarão respondendo as informações solicitadas pelos gestores da área

## ETAPA 1

### **VOOS QUE ATRASAM MAIS DE 2 HORAS**

Qual comportamento  
médio e a variância desses  
dados?

## ETAPA 2

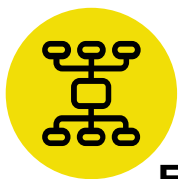
### **DESTINOS COM MAIOR NÚMERO DE ATRASOS**

- Quais são os 5 piores?  
- Considerar 5 minutos um  
atraso aceitável

## ETAPA 3

### **VELOCIDADE MÉDIA DOS VOOS ATRASADOS**

Qual a velocidade desses 5  
destinos do item anterior



# Dicas

## ETAPA 1

Qual o comportamento médio e a variância dos dados dos voos que atrasam mais de duas horas?

- a. Filtrar a tabela com os registrados que atrasam mais de duas horas.
- b. Filtrar a coluna 'dep\_delay' e aplicar as funções describe() e .T
  - i. Com o describe() algumas medidas de tendência central, mas não temos a variância.
- c. Adicionar uma coluna para a amplitude: `df['range'] = df['max'] - df['min']`
- d. Adicionar uma coluna para variância: `dfT['var'] = pow(dfT_1['std'], 2)`



# Dicas

## ETAPA 2

Quais destinos geraram voos com o maior número de atrasos na partida?

(5 piores | Considerar 5 minutos um atraso aceitável)

- a. Filtrar a tabela com os registrados que atrasam mais de 5 minutos.
- b. Filtrar a coluna de destino, fazer uma contagem dos valores e apresentar os 5 últimos
- c. Filtrar a coluna de destino, fazer uma contagem dos valores e apresentar os 5 primeiros



# Dicas

## ETAPA 3

Quais são as velocidades médias dos voos nos 5 destinos que mais ocorrem atrasos.

- a. Filtrar a tabela com os registrados que mais atrasam
- b. Calcular a velocidade média
  - i. Criar uma coluna com os valores dessas velocidades médias usando as colunas de ['distance'] e ['air\_time']
  - ii. Agrupar a tabela com os piores destinos

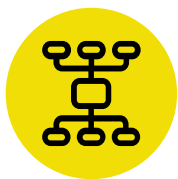
# ENTREGA



Uma pessoa representante do grupo deverá:

1. Fazer o download do arquivo no **formato .ipynb**
2. Submeter no link informado no Slack.





# Timelog

