

# Teste Técnico

# Cientista de Dados



Klemerson da Silva Gomes

# Conteúdo

1. Desafio
2. Processo Utilizado
3. Hipóteses
4. Análise Exploratória
5. Conclusão e Insights

# Desafio

- O teste técnico para a vaga de Cientista de Dados da A3Data consiste na exploração da base dados "Ocorrências Aeronáuticas da Aviação Civil Brasileira" dos dados abertos do governo (<https://dados.gov.br/dataset/ocorrencias-aeronauticas-da-aviacao-civil-brasileira>).
- O objetivo deste teste é conseguir explorar e demonstrar habilidades como Cientista de Dados. Tendo como premissa ser criativo com as análises.

# Processo Utilizado

- Diante do desafio proposto foi decidido fazer uma Análise Exploratória dos Dados, com o objetivo de levantar insights através dos dados fornecidos;
- Para a obtenção dos resultados foram levantadas nove hipóteses;
- As ferramentas escolhidas para o desenvolvimento da Análise Exploratória foram:
  - Linguagem de programação Python utilizando a distribuição Anaconda;
  - Ambiente de desenvolvimento Jupyter Notebook.
- As bases utilizadas foram:
  - ocorrencia.csv
  - ocorrencia\_tipo.csv
  - aeronave.csv

# Hipóteses

H1: Os Incidentes representam mais de 50% das ocorrências e os Acidentes representam menos de 40% das Ocorrências.

H2: Em todas as classificações de ocorrências os relatórios levam em média até 365 dias para serem publicados a partir da data da ocorrência.

H3: A ocorrência de Acidentes cresce ao longo dos anos.

H4: A maior quantidade de acidentes acontece durante o período da noite.

H5: Os cinco estados que mais tiveram ocorrências registradas também são os cinco estados que mais registraram acidentes.

H6: O tipo de aeronave “Avião” é responsável por mais de 50% das ocorrências.

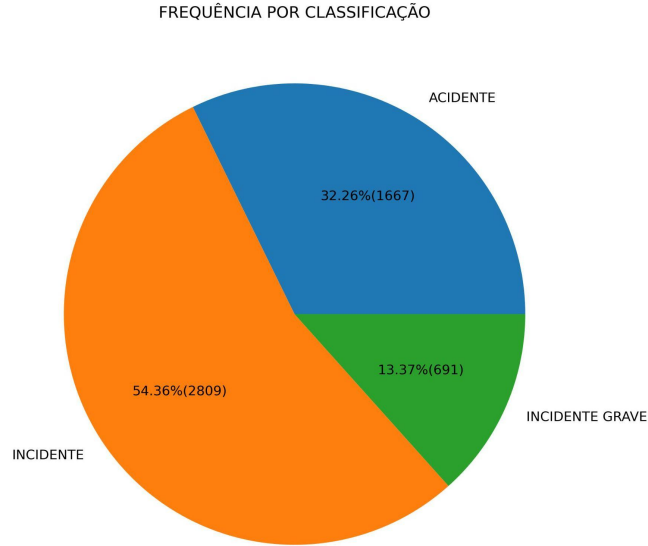
H7: Nos registros de acidentes a maior quantidade de nível de dano registrado está na categoria Substancial e Destruída.

H8: O motor de pistão é o que apresenta maior ocorrência de falha em voo.

H9: As aeronaves que apresentaram a maior quantidade de ocorrência por falha de motor em voo, são as aeronaves mais antigas.

# **ANÁLISE EXPLORATÓRIA**

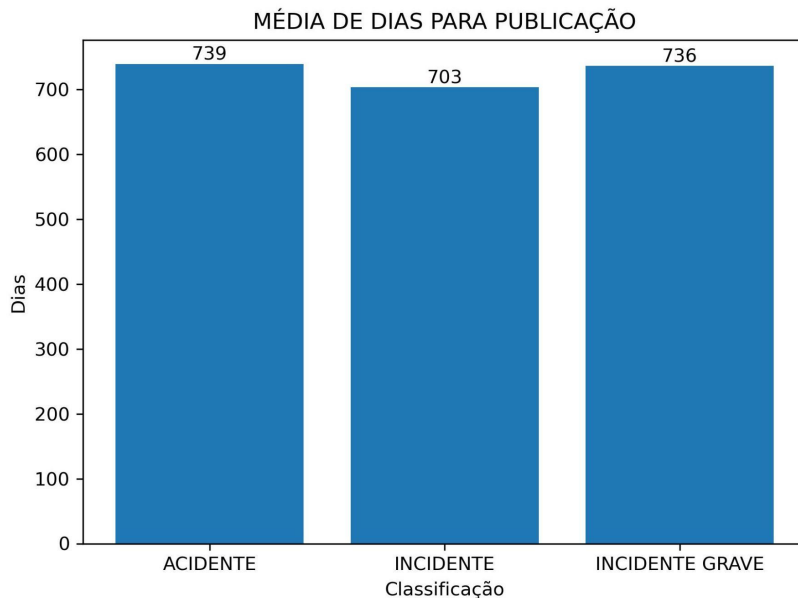
# **H1: Os Incidentes representam mais de 50% das ocorrências e os Acidentes representam menos de 40% das Ocorrências.**



- Diante da análise realizada nota-se que os incidentes representam **54.36%** das ocorrências totais e os acidentes representam **32.26%** , validando a hipótese levantada.

## H2: Em todas as classificações de ocorrências os relatórios levam em média até 365 dias para serem publicados a partir da data da ocorrência.

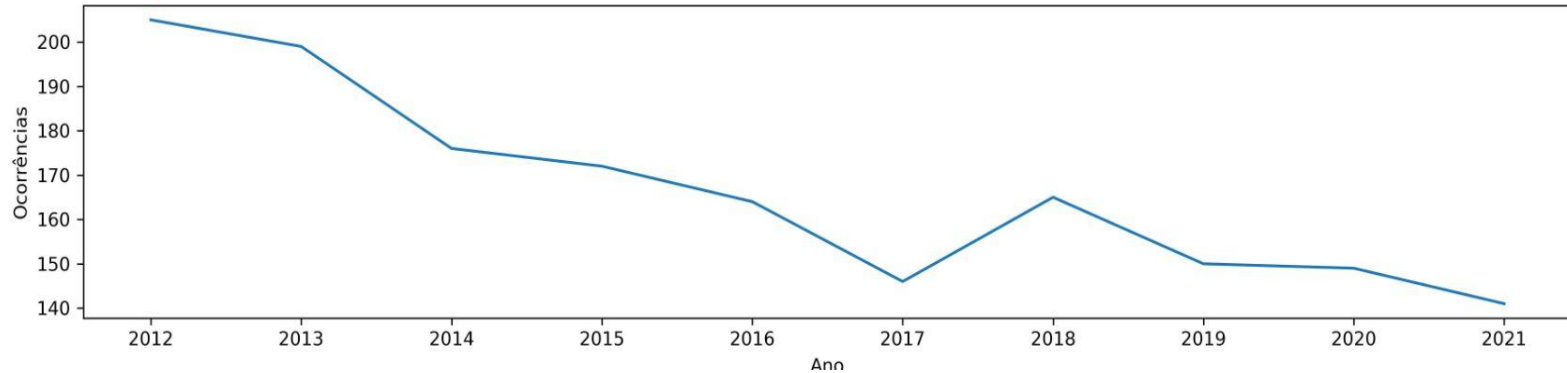
	ocorrencia_classificacao	diferenca_dias
0	ACIDENTE	739
1	INCIDENTE	703
2	INCIDENTE GRAVE	736



- Diante da análise realizada observa-se que os relatórios são publicados depois de 700 dias em média em todas as classificações, o que caracteriza quase o dobro da hipótese levantada e diante disso, a hipótese é invalidada.

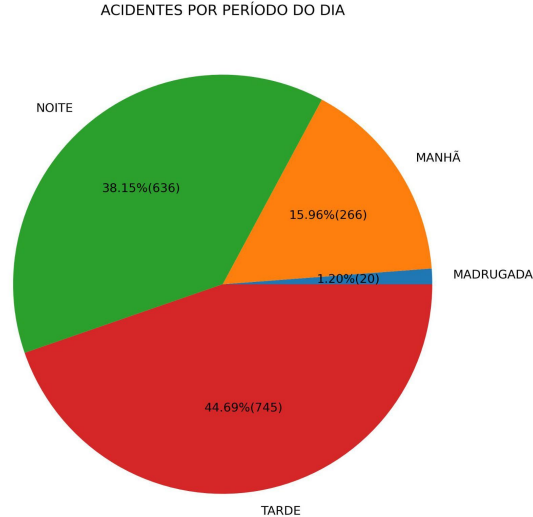


### H3: A ocorrência de Acidentes cresce ao longo dos anos.



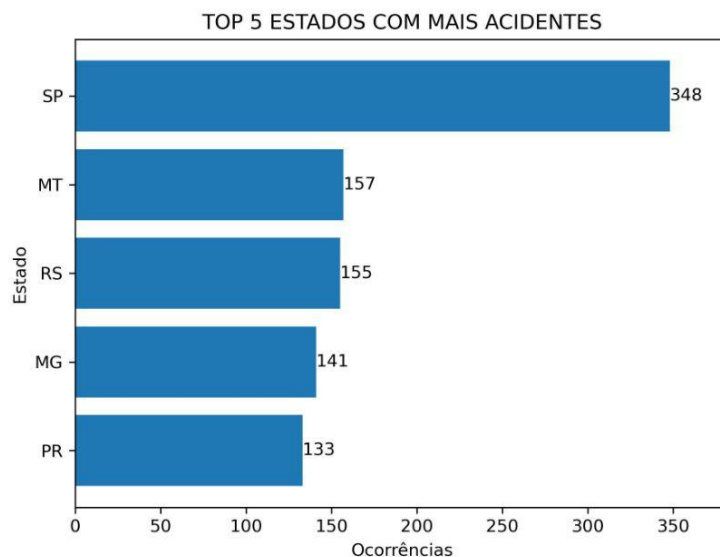
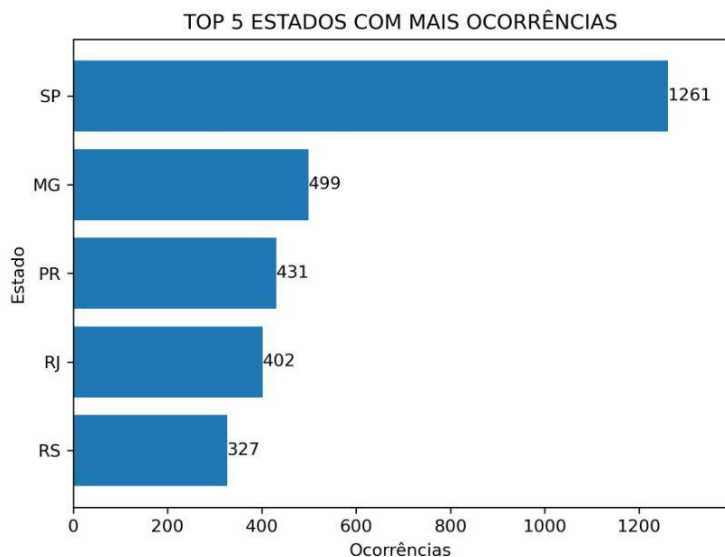
- Diante da análise realizada foi possível observar que as ocorrências seguiram uma tendência de queda ao longo dos anos, com exceção do ano de 2018. Diante desse cenário a hipótese levantada é invalidada.

## H4: A maior quantidade de acidentes acontece durante o período da noite.



- Diante da análise realizada nota-se que o período da noite representa uma alta porcentagem dos acidentes (38.15%), porém o período com maior quantidade de acidentes é o período da tarde, representando 44.69% de todos os acidentes. Diante desse cenário a hipótese levantada é invalidada.

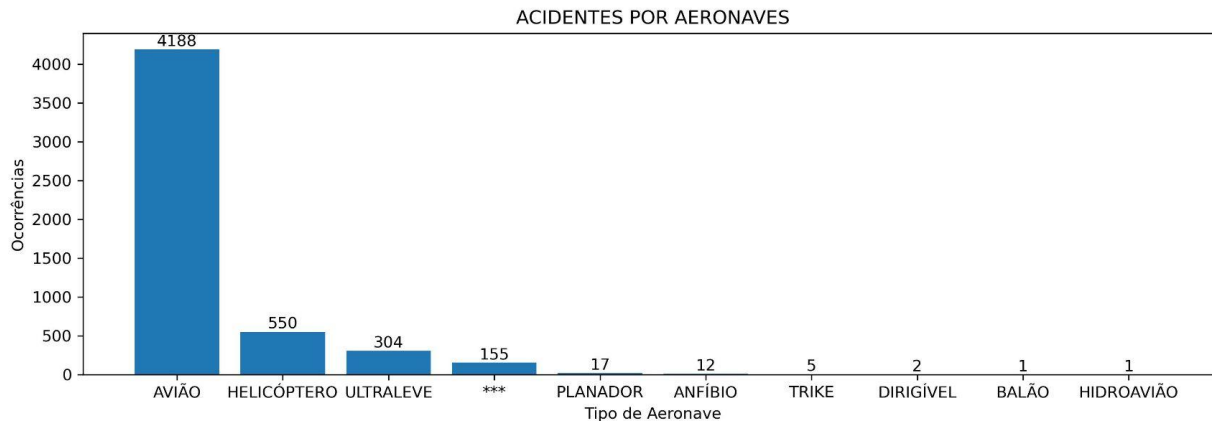
## H5: Os cinco estados que mais tiveram ocorrências registradas também são os cinco estados que mais registraram acidentes.



- Diante da análise realizada observa-se que quatro dos cinco estados com maior número de ocorrências também estão entre os cinco com maior número de acidentes, porém o estado do Rio de Janeiro não aparece entre os cinco com mais acidentes, invalidando assim a hipótese levantada.

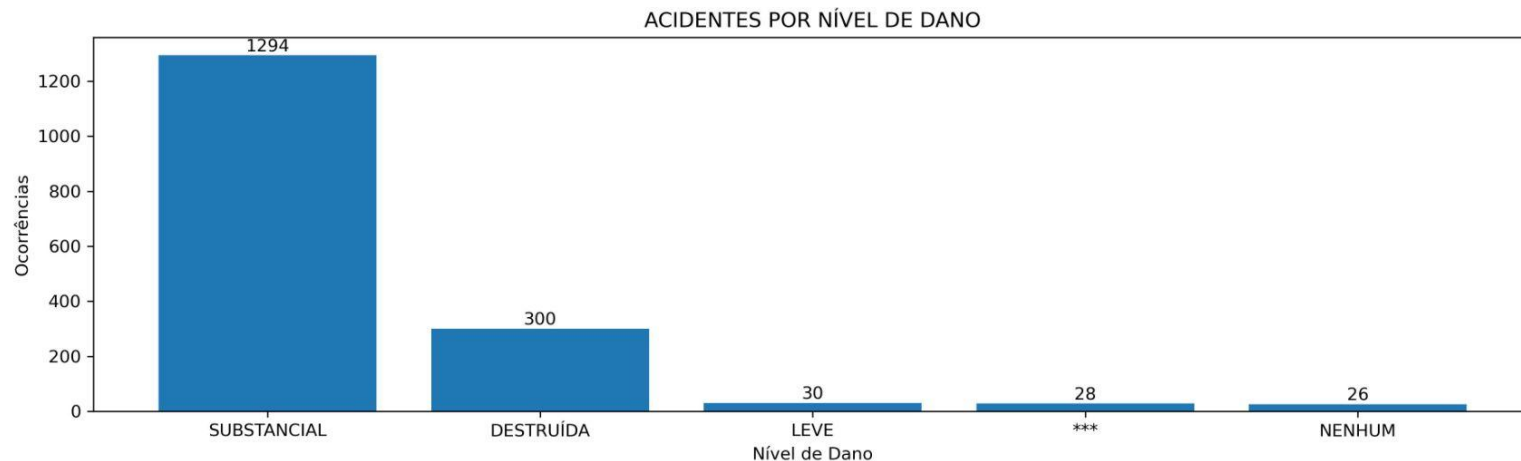
## H6: O tipo de aeronave “Avião” é responsável por mais de 50% das ocorrências.

	tipo	contagem	Porcentagem
0	AVIÃO	4188	80.00
1	HELICÓPTERO	550	10.51
2	ULTRALEVE	304	5.81
3	***	155	2.96
4	PLANADOR	17	0.32
5	ANFÍBIO	12	0.23
6	TRIKE	5	0.10
7	DIRIGÍVEL	2	0.04
8	BALÃO	1	0.02
9	HIDROAVIÃO	1	0.02



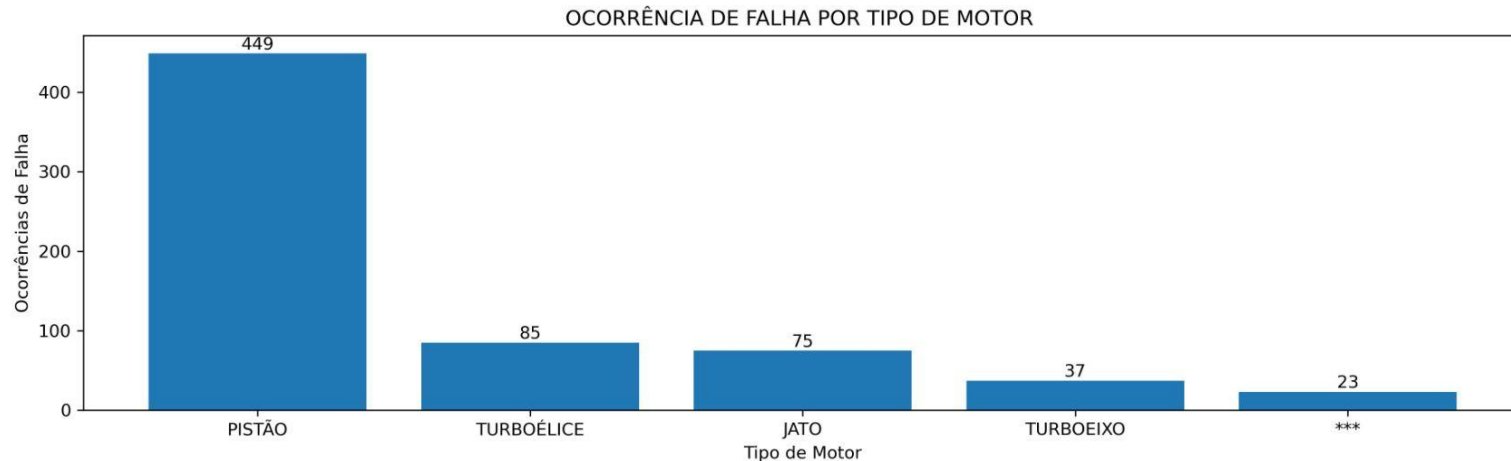
- Diante da análise realizada foi possível observar que a aeronave do tipo Avião está presente em 80% das ocorrências. Diante desse resultado a hipótese levantada é validada.

## **H7: Nos registros de acidentes a maior quantidade de nível de dano registrado está na categoria Substancial e Destruída.**



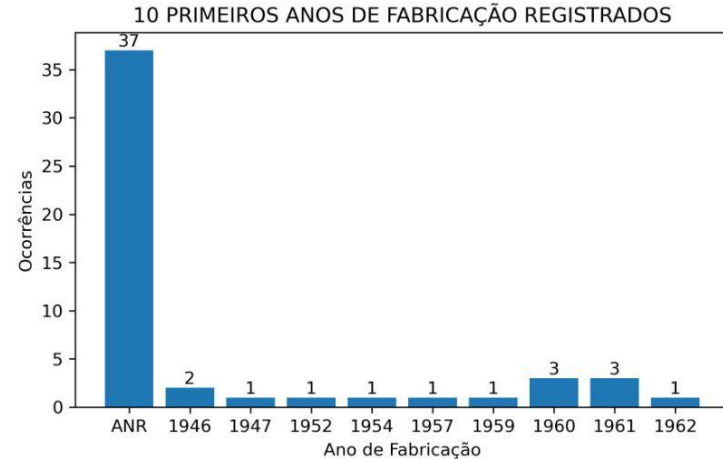
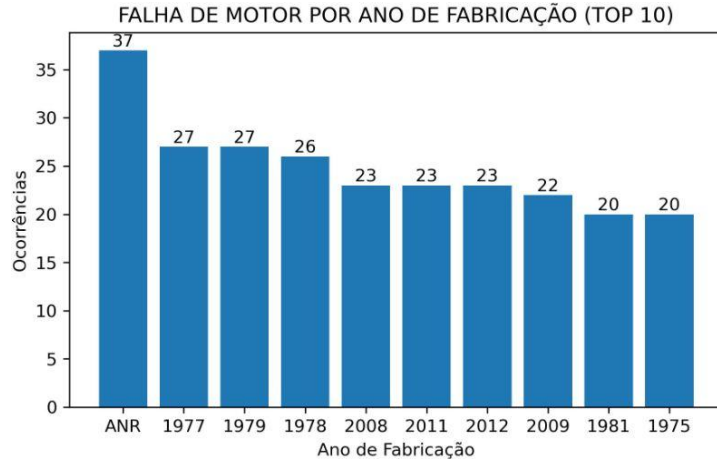
- Diante da análise realizada notou-se o comportamento esperado dos dados, onde as categorias de dano substancial e destruída apresentaram maior quantidade de ocorrências de acidentes. Diante do resultado obtido a hipótese levantada foi validada.

## H8: O motor de pistão é o que apresenta maior ocorrência de falha em voo.



- Diante da análise realizada foi possível notar que o comportamento dos dados foi de acordo com o esperado, onde o motor do tipo pistão esteve presente na maior quantidade de ocorrências de falha de motor em voo. Diante desse resultado a hipótese levantada foi validada.

## H9: As aeronaves que apresentaram a maior quantidade de ocorrência por falha de motor em voo, são as aeronaves mais antigas.



- Diante da análise realizada, notou-se que as aeronaves mais antigas não são as mais presentes em ocorrências que tiveram a falha do motor. Com esse resultado a hipótese levantada foi invalidada.

# Conclusão e Insights

- Os resultados obtidos apresentam a possibilidade de diversas tomadas de decisões, algumas únicas, em cima de cada hipótese, e algumas utilizando informações em conjunto, como na hipótese do tipo de motor e fabricação das aeronaves.
- Os resultados obtidos também oferecem um insight quanto à riqueza de informações dos dados fornecidos, possibilitando o desenvolvimento de mais análise e em alguns casos a possibilidade da aplicação de modelos de classificação e modelos de séries temporais.