



Kevin K. Takano
Universidade Federal do Amazonas



Introdução

Metodologia

Perguntas para análise

Análise exploratória

Conclusão





- Objetivo: Análise exploratória da Base de dados "Ocorrências Aeronáuticas na Aviação Civil Brasileira".



- ▶ Objetivo: Análise exploratória da Base de dados "Ocorrências Aeronáuticas na Aviação Civil Brasileira".
 - ▶ A base contém informações sobre ocorrências de incidentes e acidentes. Envolve:



- ▶ Objetivo: Análise exploratória da Base de dados "Ocorrências Aeronáuticas na Aviação Civil Brasileira".
 - ▶ A base contém informações sobre ocorrências de incidentes e acidentes. Envolve:
 - ▶ Dados de fatores que levaram os acidentes.



- ▶ Objetivo: Análise exploratória da Base de dados "Ocorrências Aeronáuticas na Aviação Civil Brasileira".
 - ▶ A base contém informações sobre ocorrências de incidentes e acidentes. Envolve:
 - ▶ Dados de fatores que levaram os acidentes.
 - ▶ Dados do estado da aeronaves após a ocorrência.



- ▶ Objetivo: Análise exploratória da Base de dados "Ocorrências Aeronáuticas na Aviação Civil Brasileira".
 - ▶ A base contém informações sobre ocorrências de incidentes e acidentes. Envolve:
 - ▶ Dados de fatores que levaram os acidentes.
 - ▶ Dados do estado da aeronaves após a ocorrência.
 - ▶ Características das aeronaves.



- ▶ Objetivo: Análise exploratória da Base de dados "Ocorrências Aeronáuticas na Aviação Civil Brasileira".
 - ▶ A base contém informações sobre ocorrências de incidentes e acidentes. Envolve:
 - ▶ Dados de fatores que levaram os acidentes.
 - ▶ Dados do estado da aeronaves após a ocorrência.
 - ▶ Características das aeronaves.
 - ▶ Localização das ocorrências.



- ▶ Problema: Qual a relação de diferença entre aspectos de causas acidentes e incidentes em aeronaves?
- ▶ Hipótese: É possível encontrar características que diferem causas entre acidentes e incidentes.
- ▶ Hipóteses secundárias:



- ▶ Problema: Qual a relação de diferença entre aspectos de causas acidentes e incidentes em aeronaves?
- ▶ Hipótese: É possível encontrar características que diferem causas entre acidentes e incidentes.
- ▶ Hipóteses secundárias:
 - ▶ Existem vários fatores que causam acidentes e incidentes.



- ▶ Problema: Qual a relação de diferença entre aspectos de causas acidentes e incidentes em aeronaves?
- ▶ Hipótese: É possível encontrar características que diferem causas entre acidentes e incidentes.
- ▶ Hipóteses secundárias:
 - ▶ Existem vários fatores que causam acidentes e incidentes.
 - ▶ Acidentes e incidentes podem ter maiores tendências em certas localizações.



- ▶ Problema: Qual a relação de diferença entre aspectos de causas acidentes e incidentes em aeronaves?
- ▶ Hipótese: É possível encontrar características que diferem causas entre acidentes e incidentes.
- ▶ Hipóteses secundárias:
 - ▶ Existem vários fatores que causam acidentes e incidentes.
 - ▶ Acidentes e incidentes podem ter maiores tendências em certas localizações.
 - ▶ Acidentes e incidentes se diferem em relação a quantidade no tempo.

1. Definição de problema e hipóteses.



1. Definição de problema e hipóteses.
2. Desenvolvimento de perguntas para análise.



1. Definição de problema e hipóteses.
2. Desenvolvimento de perguntas para análise.
3. Tratamento de dados. Remoção de dados irregulares e seleção de colunas de interesse.



1. Definição de problema e hipóteses.
2. Desenvolvimento de perguntas para análise.
3. Tratamento de dados. Remoção de dados irregulares e seleção de colunas de interesse.
4. Desenvolvimento da análise. Uso de gráficos e métodos estatísticos (histogramas, teste de hipótese, etc)



1. Definição de problema e hipóteses.
2. Desenvolvimento de perguntas para análise.
3. Tratamento de dados. Remoção de dados irregulares e seleção de colunas de interesse.
4. Desenvolvimento da análise. Uso de gráficos e métodos estatísticos (histogramas, teste de hipótese, etc)
5. Análise de resultados.



1. Definição de problema e hipóteses.
2. Desenvolvimento de perguntas para análise.
3. Tratamento de dados. Remoção de dados irregulares e seleção de colunas de interesse.
4. Desenvolvimento da análise. Uso de gráficos e métodos estatísticos (histogramas, teste de hipótese, etc)
5. Análise de resultados.
6. Conclusões



Em relação a Aeronave:

- ▶ De onde e quando foram fabricadas as aeronaves acidentadas e no que se difere entre acidentes e incidentes ?
- ▶ Qual tipos de aeronaves acidentadas por tipo de operação e no que se difere entre acidentes e incidentes ?
- ▶ Quais os níveis de danos para aeronaves acidentadas e no que se difere entre acidentes e incidentes ?



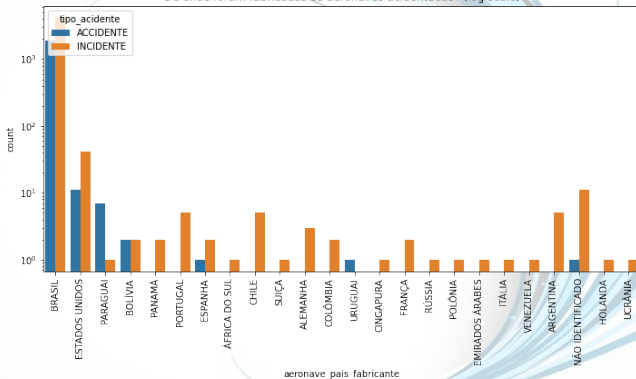
Em relação a fatores de contribuição:

- ▶ Quais fatores e aspectos que levaram as ocorrências? E no que se difere entre acidentes e incidentes ?
- ▶ Que áreas de operação ocorreram as ocorrências? E no que se difere entre acidentes e incidentes ?

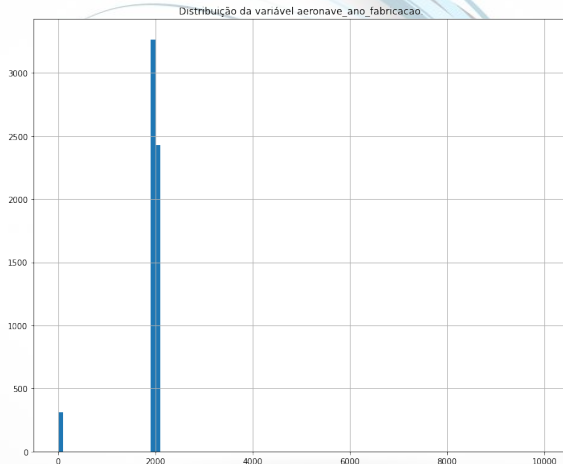
Em relação a ocorrências e tipos:

- ▶ Como ocorreram os acidentes e incidentes em relação ao tempo?
- ▶ Como ocorreram os acidentes e incidentes em relação a localização?
- ▶ Quais principais tipos de ocorrências entre acidentes e incidentes?

De onde foram fabricadas as aeronaves acidentadas ?

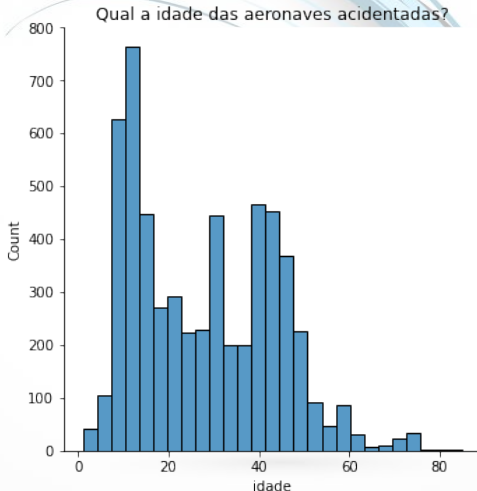


Distribuição da variável aeronave_ano_fabricacao.



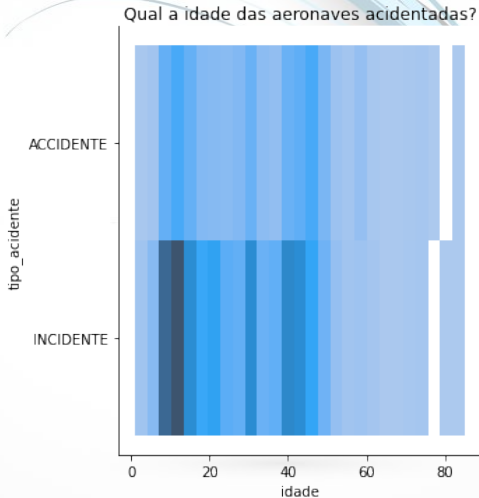
Qual idade das aeronaves acidentadas em anos?

(com remoção de anos irregulares)

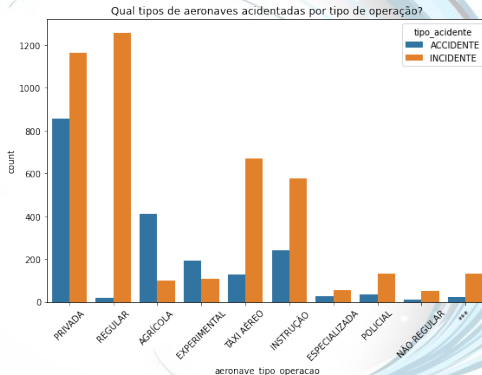


Qual idade das aeronaves acidentadas em anos?

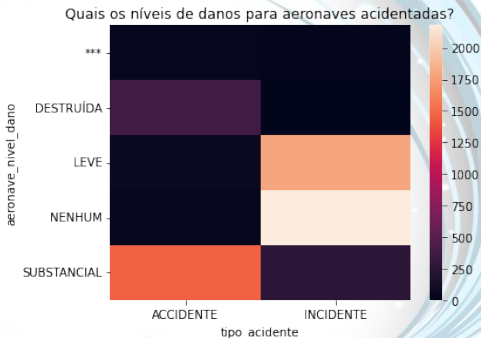
(com remoção de anos irregulares e classes)



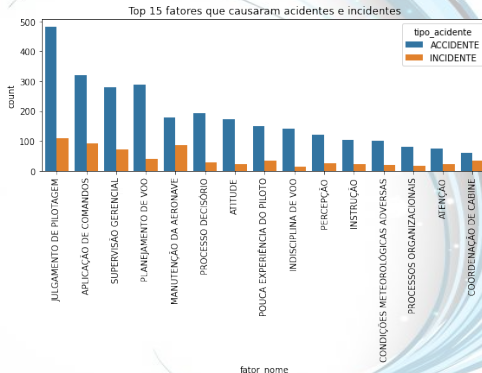
Qual tipos de aeronaves acidentadas por tipo de operação?



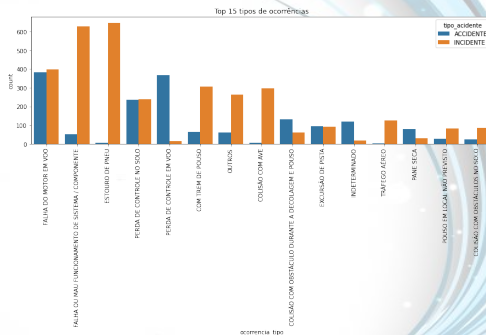
Quais os níveis de danos para aeronaves acidentadas?



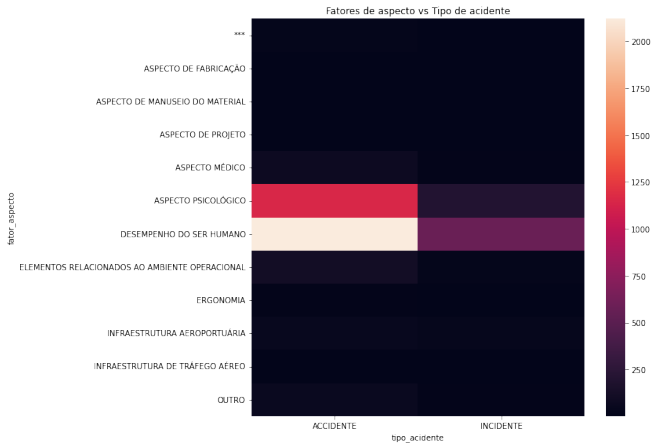
Top 15 fatores que causaram ocorrências.

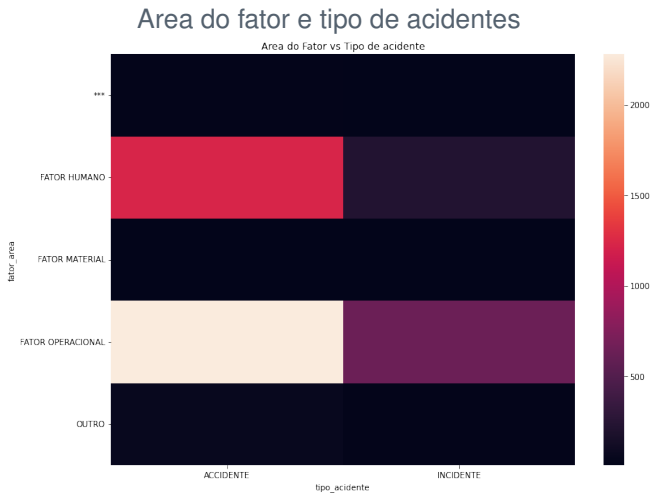


Top 15 tipos de ocorrências

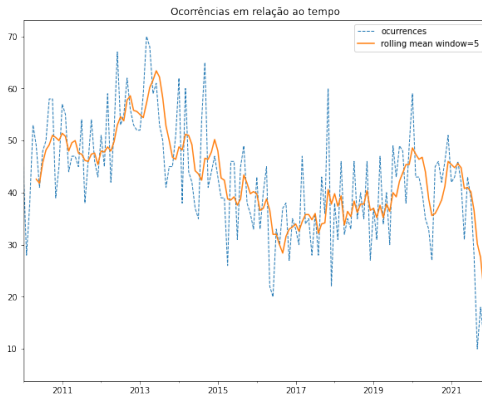


Aspecto de fatores e tipo de acidentes

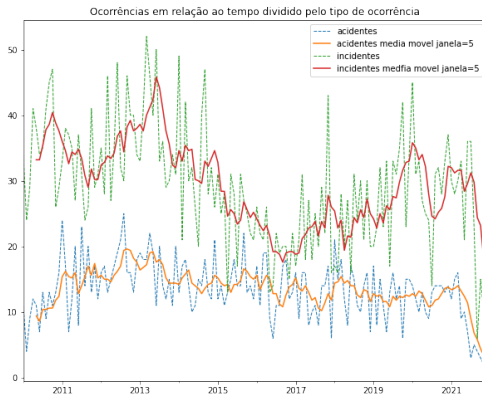




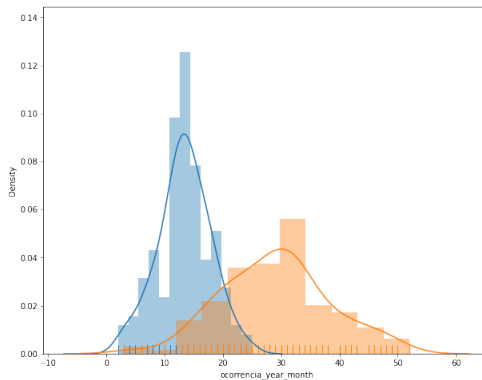
Ocorrências pelo tempo



Ocorrências pelo tempo (com classes)



Distribuição de ocorrências pelo tempo



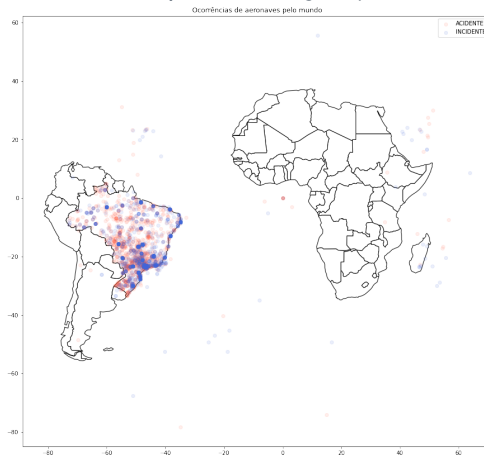
Teste de hipótese `mann_whitney_u_test` (para distribuição não-normal) para verificar se as distribuições são diferentes.

- ▶ Resultado mostra um p-valor < 0.05
- ▶ Logo, podemos dizer que a hipótese nula pode ser rejeitada.

```
: def mann_whitney_u_test(distribution_1, distribution_2):  
    """  
    Perform the Mann-Whitney U Test, comparing two different distributions.  
    Args:  
        distribution_1: List.  
        distribution_2: List.  
    Outputs:  
        u_statistic: Float. U statistic for the test.  
        p_value: Float.  
    """  
    u_statistic, p_value = scipy.stats.mannwhitneyu(distribution_1, distribution_2)  
    return u_statistic, p_value  
  
mann_whitney_u_test(list(dist1), list(dist2))  
  
: (1222.0, 2.352998826097736e-38)
```

- a hipótese nula pode ser rejeitada, isto é, os dois grupos são diferentes.

Distribuição de ocorrências por localização (America latina e África)





- ▶ Através de nossas análises podemos verificar que existem vários fatores que diferenciam acidentes de incidentes. Nossas observações de forma resumida:



- ▶ Através de nossas análises podemos verificar que existem vários fatores que diferenciam acidentes de incidentes. Nossas observações de forma resumida:
 - ▶ No mapa de calor de idade é possível observar que há mais navios com menos de 20 anos em incidentes do que em acidentes.



- ▶ Através de nossas análises podemos verificar que existem vários fatores que diferenciam acidentes de incidentes. Nossas observações de forma resumida:
 - ▶ No mapa de calor de idade é possível observar que há mais naves com menos de 20 anos em incidentes do que em acidentes.
 - ▶ No mapa de calor de danos, mostra-se que há mais danos substanciais em acidentes do que incidentes.

- ▶ Através de nossas análises podemos verificar que existem vários fatores que diferenciam acidentes de incidentes. Nossas observações de forma resumida:
 - ▶ No mapa de calor de idade é possível observar que há mais naves com menos de 20 anos em incidentes do que em acidentes.
 - ▶ No mapa de calor de danos, mostra-se que há mais danos substanciais em acidentes do que incidentes.
 - ▶ No mapa de calor de aspectos de fatores, observa-se que há muito mais aspectos psicológicos e problemas de desempenho de ser humano em acidentes do que em incidentes. Neste gráfico sugere-se que há vários fatores que influenciam as ocorrências.

- ▶ Através de nossas análises podemos verificar que existem vários fatores que diferenciam acidentes de incidentes. Nossas observações de forma resumida:
 - ▶ No mapa de calor de idade é possível observar que há mais naves com menos de 20 anos em incidentes do que em acidentes.
 - ▶ No mapa de calor de danos, mostra-se que há mais danos substanciais em acidentes do que incidentes.
 - ▶ No mapa de calor de aspectos de fatores, observa-se que há muito mais aspectos psicológicos e problemas de desempenho de ser humano em acidentes do que em incidentes. Neste gráfico sugere-se que há vários fatores que influenciam as ocorrências.
 - ▶ No gráficos de ocorrência em relação tempo com classes, é possível observar que as tendências da série temporal das classes diferem. A ocorrência de acidentes tem uma variabilidade muito maior do que incidentes. E além disso, ocorre em frequência muito maior.

- ▶ Através de nossas análises podemos verificar que existem vários fatores que diferenciam acidentes de incidentes. Nossas observações de forma resumida:

- ▶ Através de nossas análises podemos verificar que existem vários fatores que diferenciam acidentes de incidentes. Nossas observações de forma resumida:
 - ▶ O gráfico de distribuição das ocorrências reflete a discrepância entre as amostras. Além disso, essa discrepância é mais reforçada através do teste de hipótese Mann Whitney, em que rejeitamos a hipótese nula com $p\text{-valor} < 0.05$.

- ▶ Através de nossas análises podemos verificar que existem vários fatores que diferenciam acidentes de incidentes. Nossas observações de forma resumida:
 - ▶ O gráfico de distribuição das ocorrências reflete a discrepância entre as amostras. Além disso, essa discrepância é mais reforçada através do teste de hipótese Mann Whitney, em que rejeitamos a hipótese nula com $p\text{-valor} < 0.05$.
 - ▶ No gráfico de localização, podemos ver que há uma densidade muito maior no sudeste de ocorrências do que no resto do país.