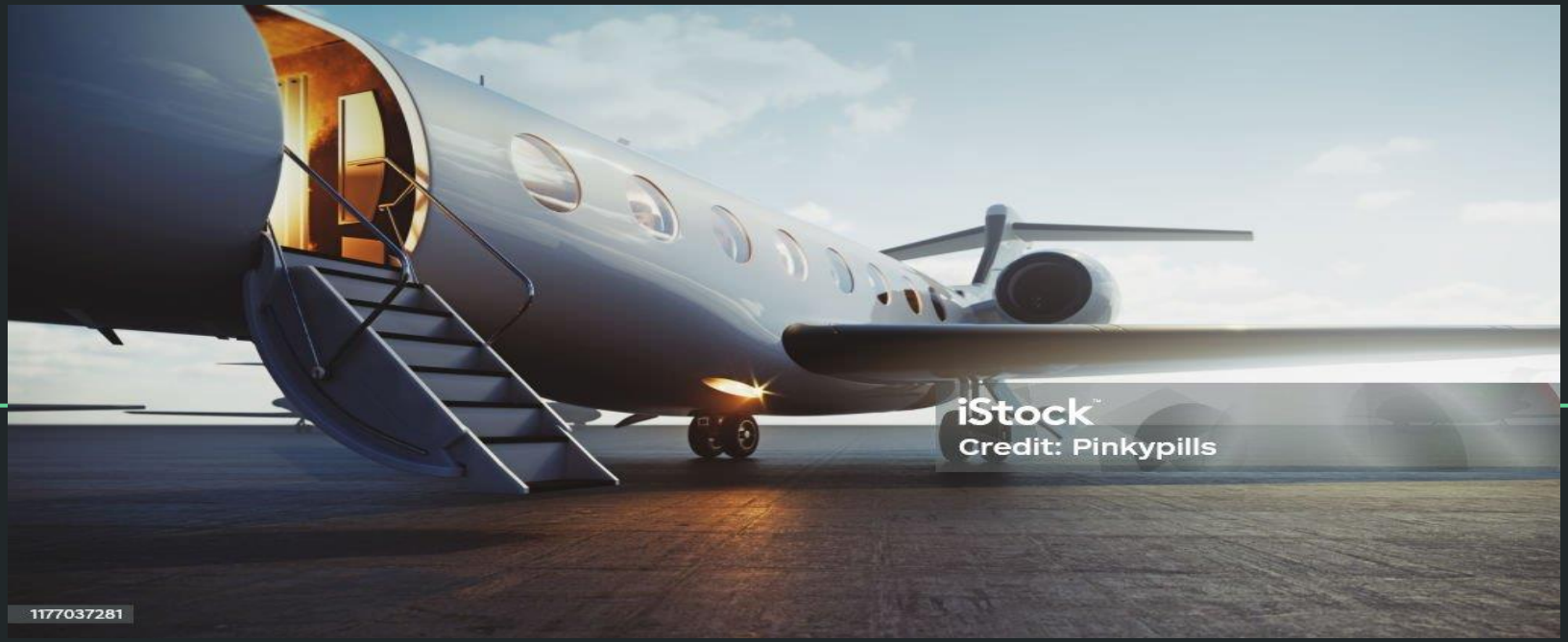


Data Analysis on the Aviation Business for both Private and Commercial Enterprise

Summary

Safety Scores across different Aircraft Make_Model were analyzed to evaluate the potential risks for both commercial and private Enterprise.

From the analysis I was able to assess the trends in the Safety Scores over the Event Year to determine the improvement as well as the declines with the Aircraft Safety

There was no clear correlation between the Number of Engines and the Safety Scores which further implies that increasing the Number of Engines does not primarily seem to lead to consistents increase in the Safety Scores

Outline

- Business Problem
- Data
- Methods
- Results
- Conclusions

Business Problem

The company is expanding in to new industries to diversify its portfolio. Specifically, they are interested in purchasing and operating airplanes for commercial and private enterprises, but do not know anything about the potential risks of aircraft. You are charged with determining which aircraft are the lowest risk for the company to start this new business endeavor. You must then translate your findings into actionable insights that the head of the new aviation division can use to help decide which aircraft to purchase

Data

From the two data sets, df (USState_codes.csv) and df2(AviationData.csv)

Have both categorical and numercial data.

The two data were pulled from Kaggle Pulled from Kaggle dataset, here is the link:
https://www.kaggle.com/datasets/khsamaha/aviation-accident-database-synopses

The various datatypes include: object, float64, int64, int32.

Methods

Describe the methods you used here. Can include data preparation, analysis, and/or modeling

Importing necessary Python Libraries

Loading the two dataset. USState_codes.csv, 'AviationData.csv'

Cleaning the dataset by ensuring the missing values are sorted.

Grouping data by 'Make' and 'Model' then summing up with the injury related columns.

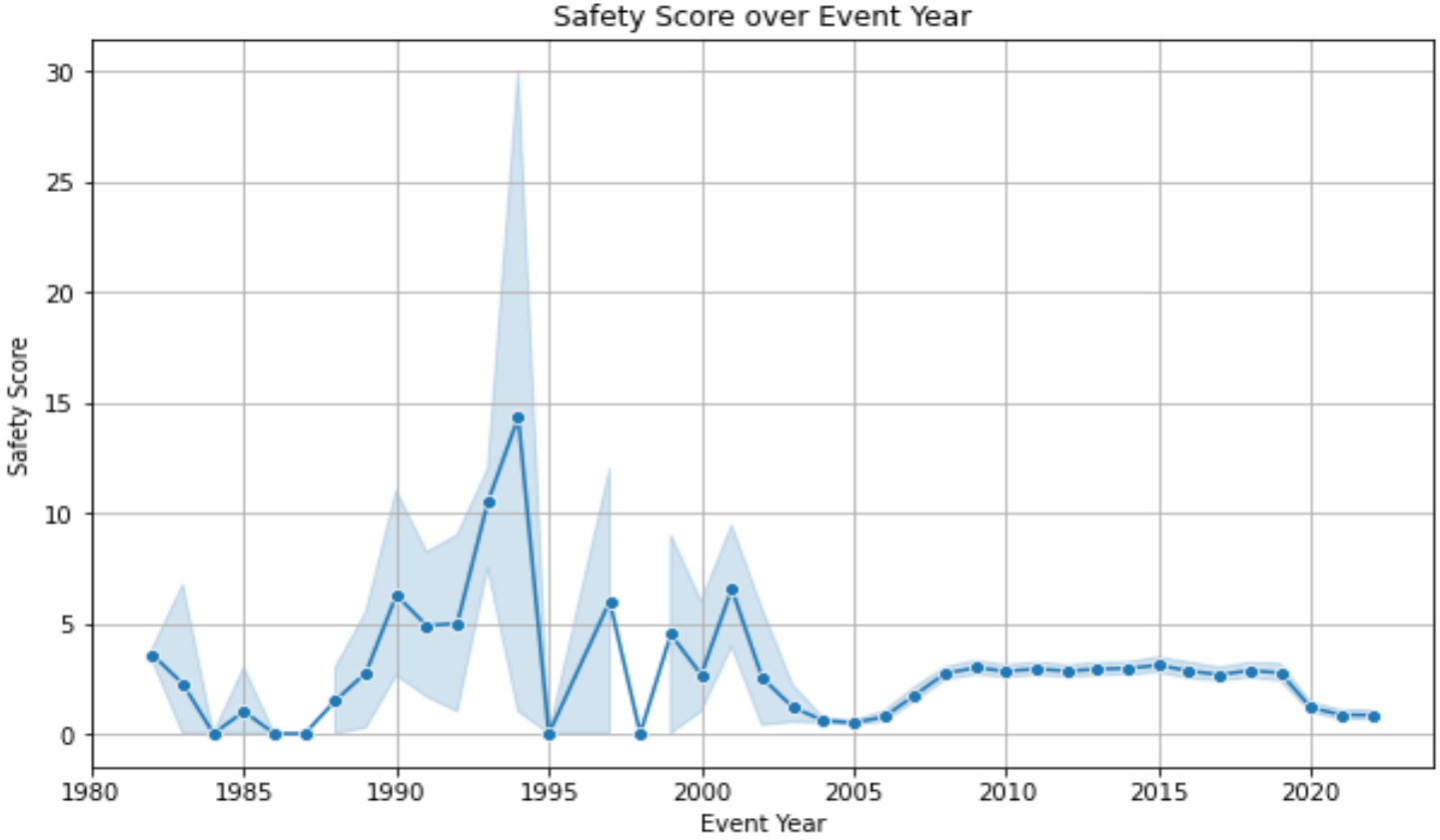
Filter the dataset to include the rows in the Purpose of f

Perform Univariate analysis for the Distribution of the Number of Engines

Creating Visualization using Matplotlib and Seaborn

Use .merge() method to join the two dataframe within their index.

Results


Safety Score over Event Year
Safety Score
30
25
20
15
10
5
0
1980
1985
1990
1995
2000
2005
2010
2015
2020
Event Year

Conclusions

Business Recommendation:

1. For the analysis do not rely solely on the Number of Engines when assessing the Aircraft safety.

It would be necessary to utilize multiple factors in the analysis inorder to make appropriate data driven decision about the safety improvement on the new Airplane business venture.

Project Limitation

From the analysis we should be able to investigate how the number of engines correlates with other performance factors like speed, range, and luggage capacity. This could give a more comprehensive insights into which aircraft make/models provide the best overall performance for either Private and Commercial Enterprise.

Future Improvement Ideas

By considering the Beech Aircraft Make having the highest Safety Score it could be prioritized for the Private Enterprise where the safety and comfort are the key metrics from my analysis.

Thank You!

Email: mika.benson@student.moringaschool.com
GitHub: @mikabenson
LinkedIn: linkedin.com/in/MikaBenson/