Aviation Risk Analysis

Project

Summary

- This project gets the aircrafts with the highest risks and the lowest risk aircrafts.

Business Problem

1. **Objective:**

Identify patterns in aviation accidents to determine the aircraft makes and models with the highest and lowest risk levels, helping identify which aircraft to Purchase.

2. **Key Questions:**

- Which aircraft models have the highest and lowest total injuries, fatalities, and damage levels?

- How can data visualization highlight the most and least risky aircrafts to purchase?

- Stakeholders: The Company

Data

1. Source

National Transportation Safety Board that includes aviation accident data from 1962 to 2023 about civil aviation accidents and selected incidents in the United States and international waters.

2. Features

- 88889 rows and 31 columns
- Make and Model of the aircraft
- Total.Fatal.Injuries – Number of fatal injuries per event
- Total.Serious.Injuries - Number of serious injuries per event
- Total.Minor.Injuries – Number of minor injuries per event
- Total.Uninjured – Number of uninjured per event

Data Preparation and Cleaning

- Dropped 18 columns
- Converted Event_Date to datetime format for time-based analysis
- Dropped Duplicates
- Standardized column names
- Dropped columns with more than 50% missing values
- Input missing categorical columns with unknown and numeric columns with median

Feature Engineering

- Created a column of Fatalities using from Injury Severity columns had categories with value counts indicating a fatal incident in parentheses.
- Created a column of total injuries, adding Total Fatal Injuries, Total Serious Injuries and Total Minor injuries
- Created a column of Damage Level from Aircraft Damage, mapping 'destroyed' to be 3, 'substantial' to be 2, 'minor' to be 1 and 'unknown' to be 0.
- Created a column Aircraft by combining make and model to get full name of Aircraft
- Created a column of Total Metric to get the riskiest aircraft by adding Total Injuries, Total fatalities and Damage Level

Methods

Exploratory Data Analysis (EDA)

- Grouped data by Aircraft to compute aggregated statistics
- Visualized distribution and trends using bar charts

Sorting and Filtering:

- Identified top 10 aircraft models with the highest and lowest aggregated risk metrics.
- Extracted subsets for aircraft with zero fatalities, injuries, and damages.

Visualization

- Used Python (Matplotlib, Seaborn) for preliminary visualizations.
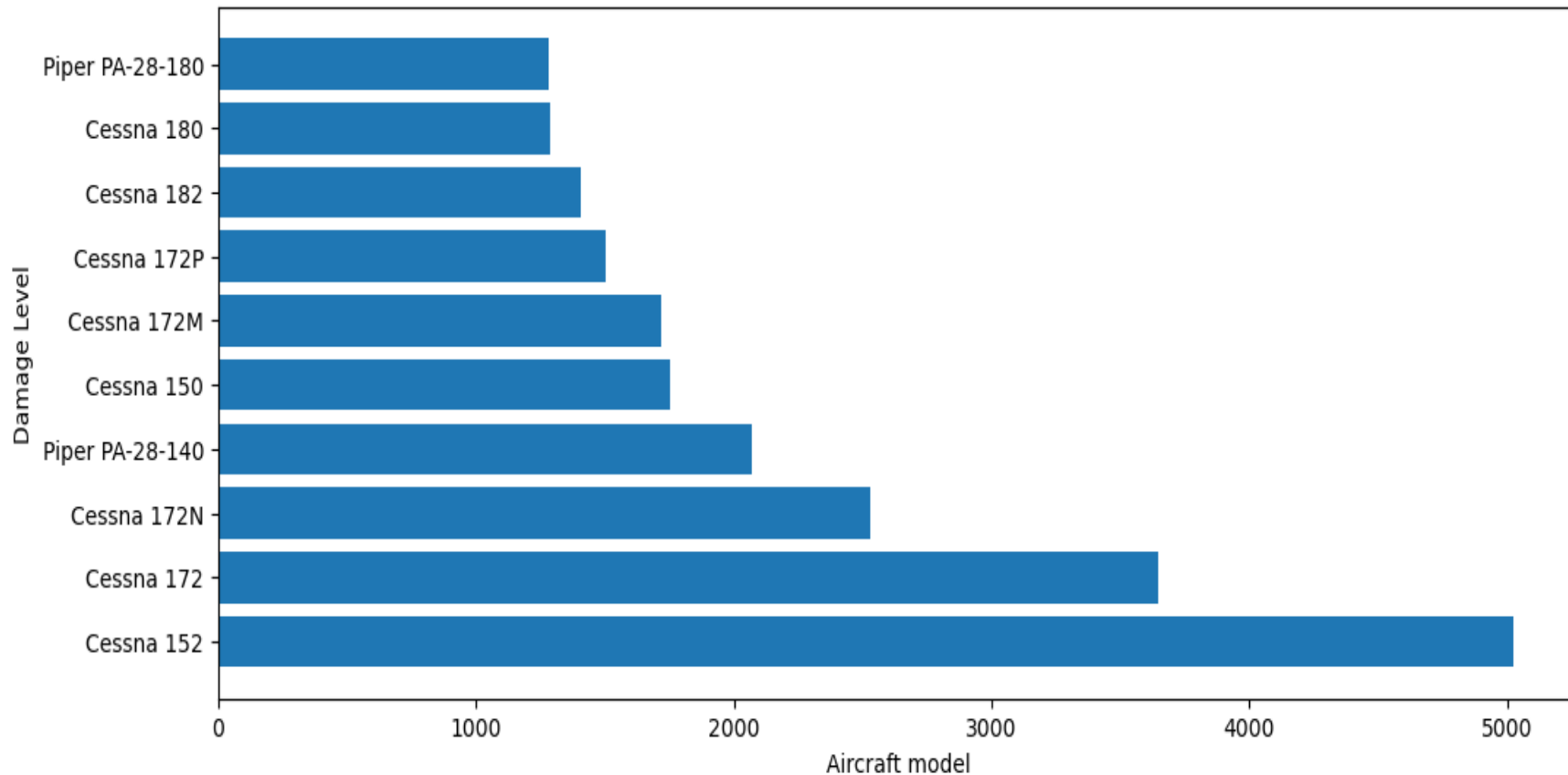- Prepared a Tableau dashboard for interactive data exploration and presentation.

Top 10 Aircraft with Highest Total Injuries
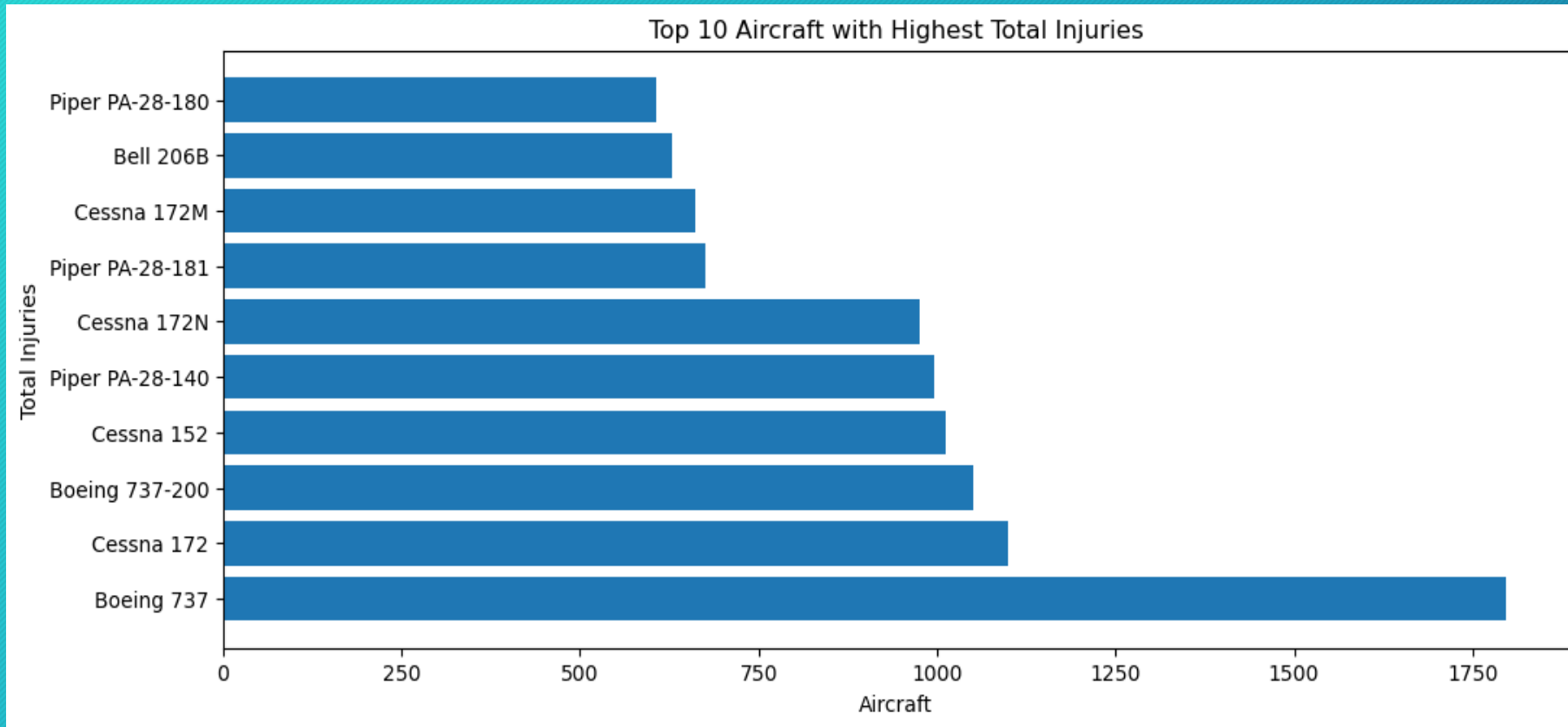
The Boeing 737 has the highest total number of injuries per event
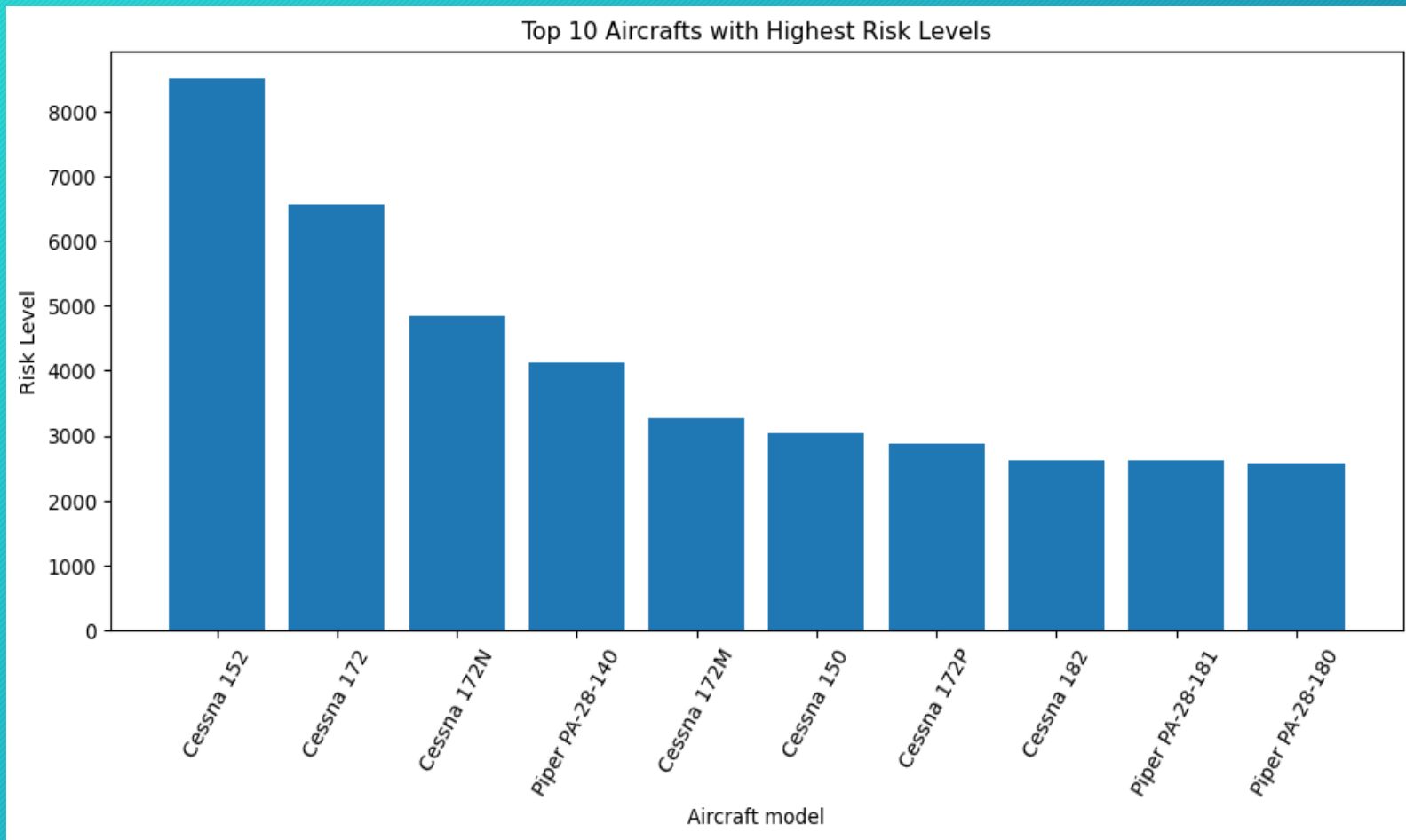
Top 10 Aircrafts with Highest Fatalities
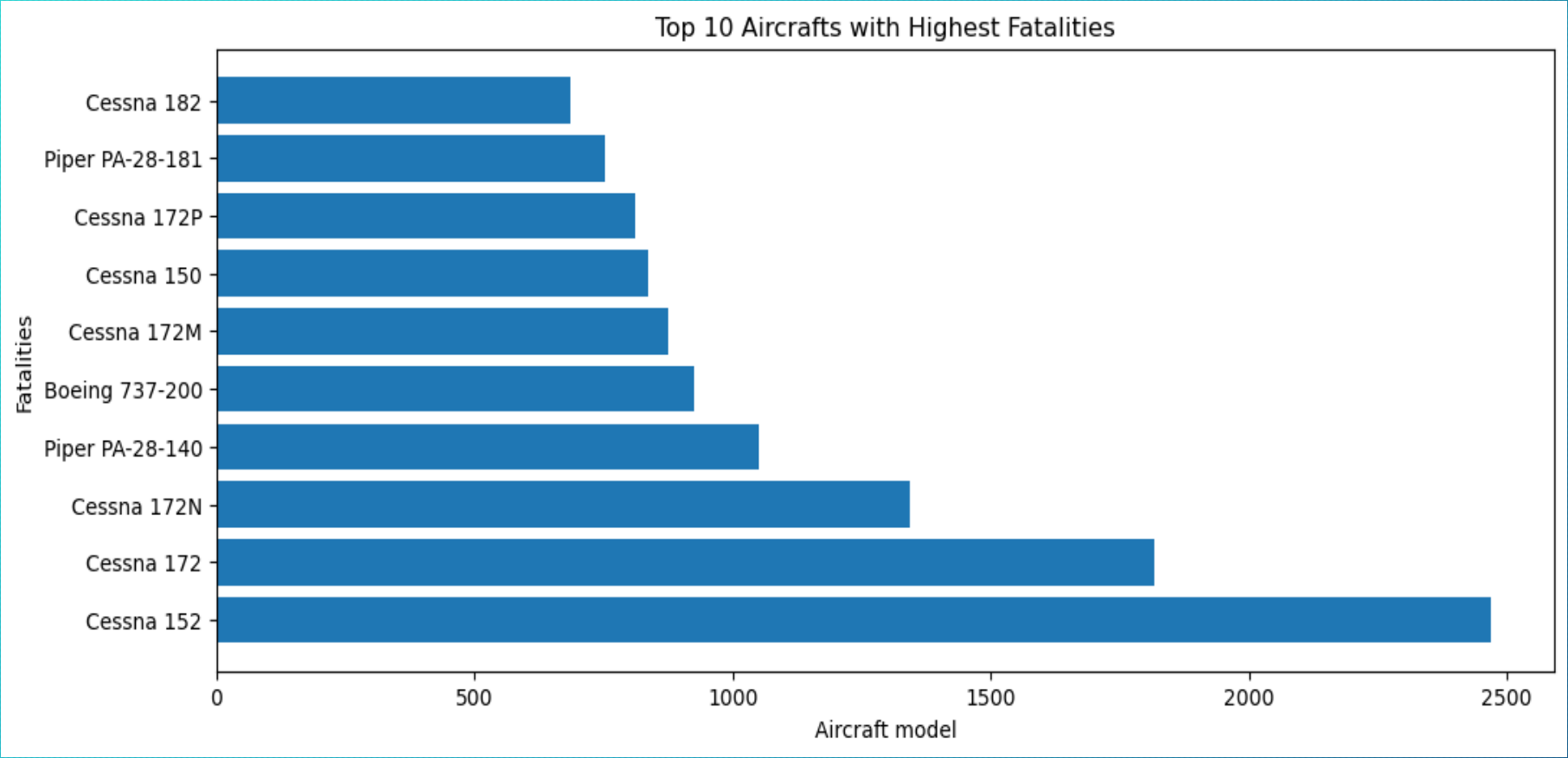
The Cessna 152 has the highest number of Fatalities

Top 10 Aircrafts with Highest Damage Levels

Cessna 152 again has the highest levels of Damege

Top 10 Aircrafts with Highest Risk Levels

Buy adding total injuries, Fatalities and damage level for each aircraft

Results

Top 10 aircrafts we should avoid
- Cessna 152
- Cessna 172
- Cessna 172N
- Piper PA-28-140
- Cessna 172M
- Cessna 150
- Cessna 172P
- Cessna 182
- Piper PA-28-181
- Piper PA-28-180

Top 10 Safest aircrafts to purchase

- Cessna 180
- Gulfstream G150
- Bombardier BD 700 1A10
- Hawker Siddeley 800XP
- Syracuse V F/Syracuse C D KITFOX
- Boeing 777 - 200
- Found Acft Canada Inc FBA-2C1
- Uvify IFO
- American Legend Aircraft Compa AL18
- Boeing 737-4B6

These aircrafts had zero total fatalities, injuries and damage level

Conclusion

- Aircraft models like Cessna 152 are associated with higher risks, warranting loses.

- Choose any of the top 1

- Models with zero fatalities or low metrics could represent best practices or superior designs.

- Limitations include potential underreporting or data inconsistencies; further validation with additional datasets is recommended.