

# AVIATION RISK ASSESSMENT: DATA-DRIVEN RECOMMENDATIONS FOR AIRCRAFT PURCHASES

# PROJECT

*Presented by: Njiru Logan Kimathi*



# ROCINANTE GLOBAL LTD

## *ABOUT THE COMPANY*

Rocinante Global is an innovative aviation enterprise focused on providing reliable and safe air transportation solutions to both commercial and private sectors.

With a commitment to operational excellence, cutting-edge technology, and sustainable practices, Rocinante Global aims to redefine the future of air travel by ensuring a seamless and efficient experience for customers across the globe.





# OVERVIEW

## Business Problem

Rocinante Global is diversifying 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.

## Project Goal

Determining which aircraft are the lowest risk for the company to start this new business endeavor. Translate the findings into actionable insights that the head of the new aviation division can use to help decide which aircraft to purchase.

## Approach

This project is an in-depth Aviation Accident Risk Assessment aimed at evaluating historical aviation accident data to inform Rocinante Global's strategic decision-making .



# DATA

## The Dataset

The NTSB aviation accident database contains information from 1962 to 2022 about civil aviation accidents and selected incidents within the United States, its territories and possessions, and in international waters.

## Features

The dataset contains over 80,000 records of individual aviation accidents or incidents and 33 different variables relating to each accident

# EDA STEPS

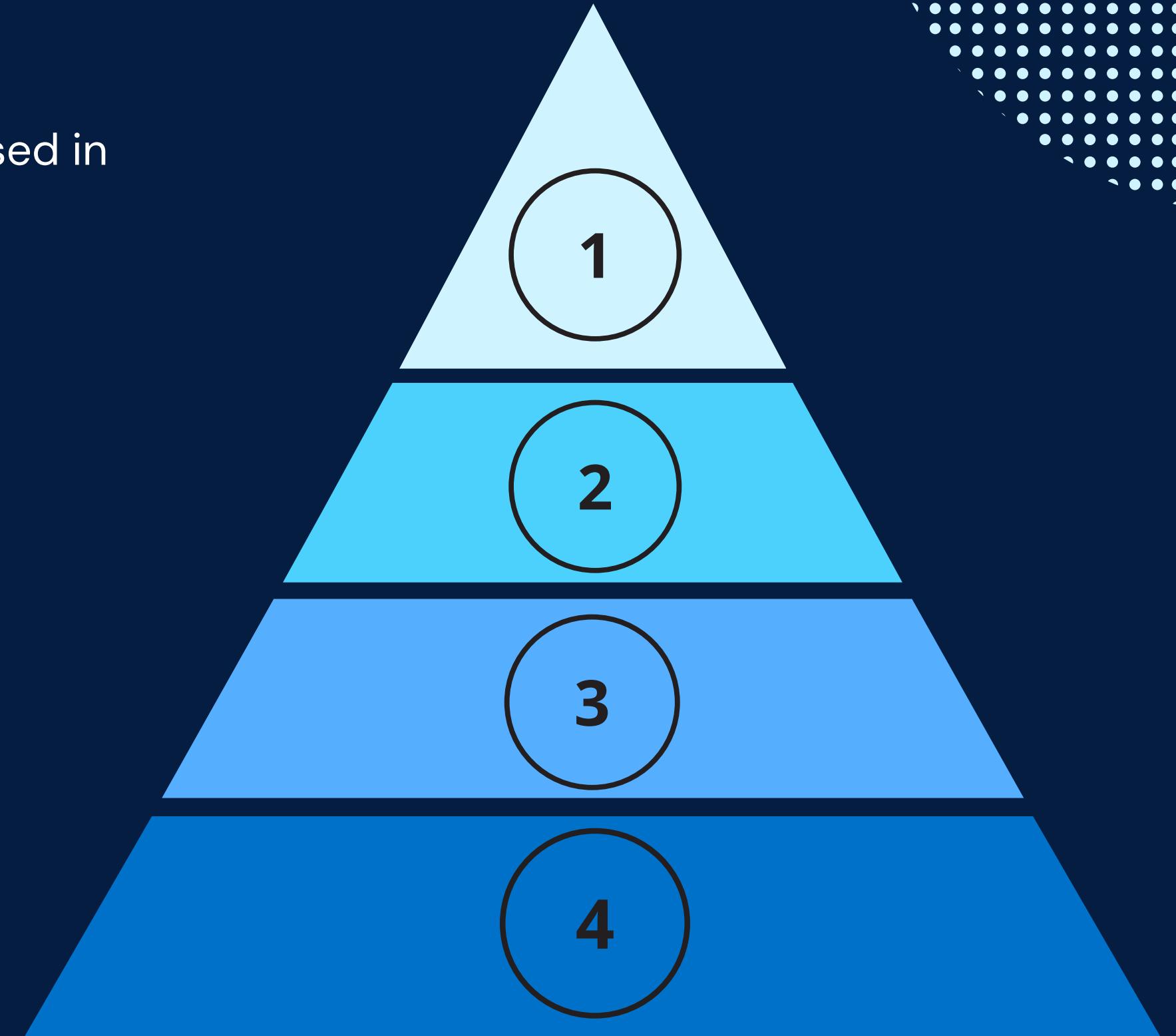
The following are the general steps that I used in exploratory data analysis of the dataset.

**01** DATA CLEANING

**02** DATA MODELING AND VISUALIZING

**03** RESULTS AND OBSERVATIONS FROM THE VISUALIZATIONS

**04** CONCLUSIONS AND RECOMMENDATIONS



# DATA CLEANING

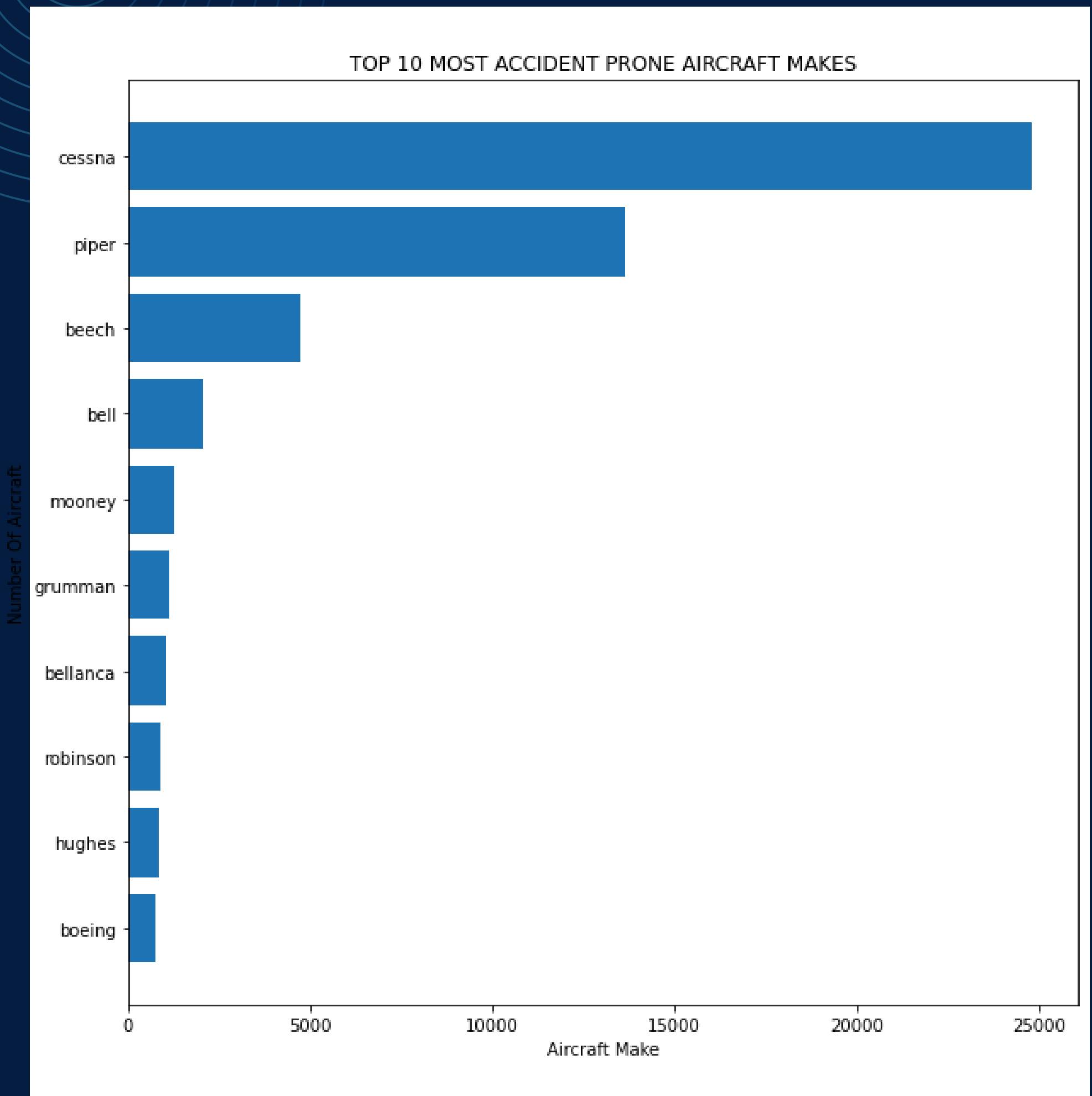
- I inspected the dataset to understand its features and properties.
- I identified missing values and duplicates and began the process of handling them.
  - Variables that had largely missing values (over 70% NaNs) and were also not relevant to the study in context were dropped.
  - For those variables that I deemed to be relevant to the study, I imputed the missing values with appropriate values.
- The result was a dataset that has less than 7% missing values in only a few variables.
- I then dropped all the remaining rows with missing values
- The end result was a cleaned database with no missing values, over 75,000 records and 24 relevant variables.

# Data Analysis

The top 10 most accident prone aircraft makers were:

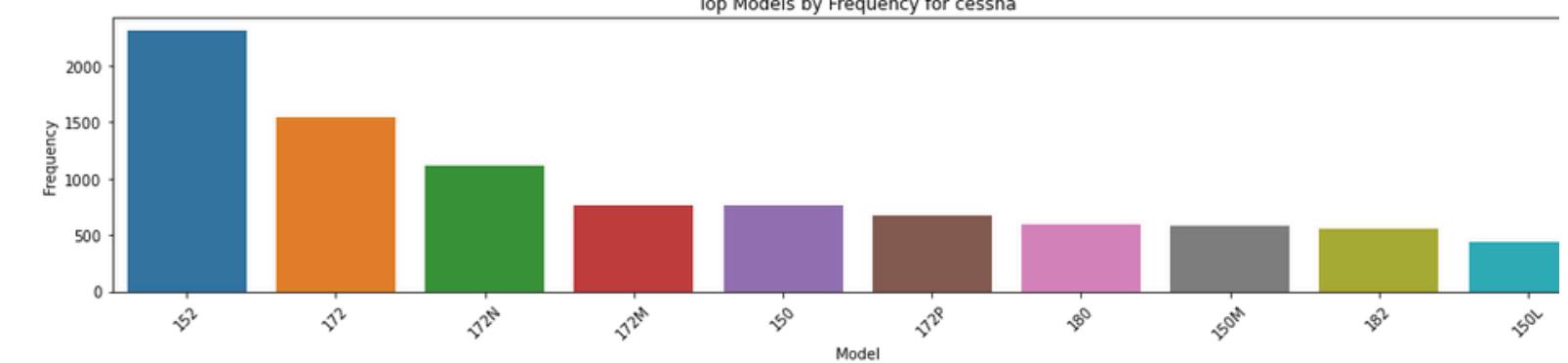
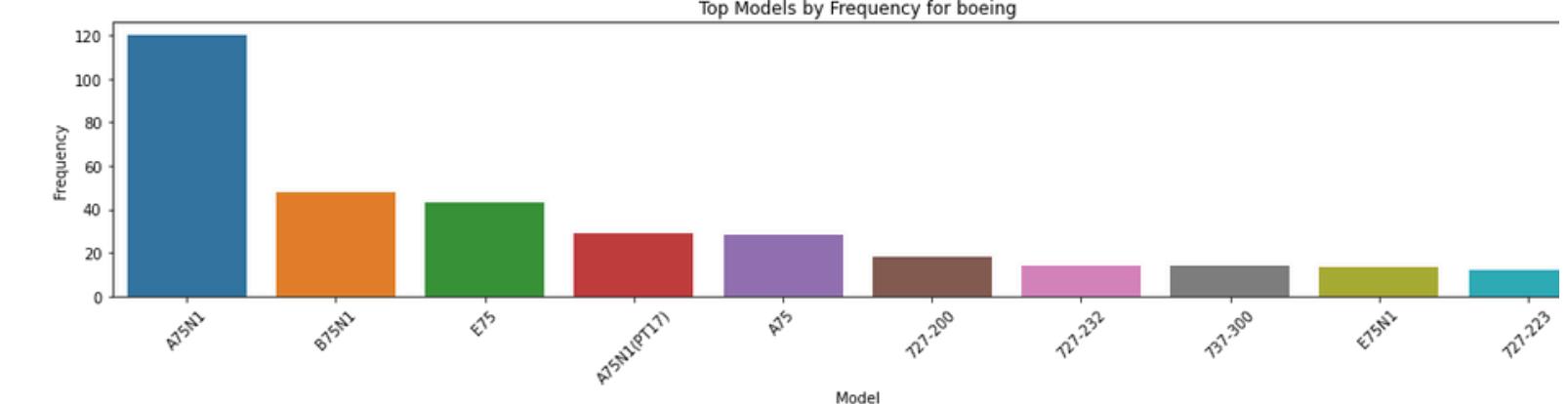
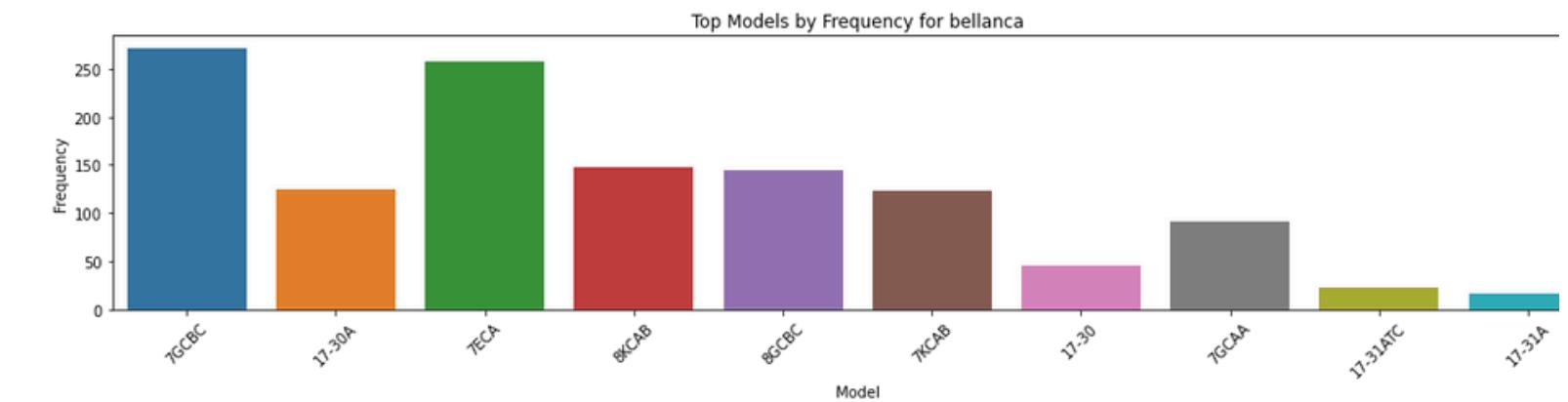
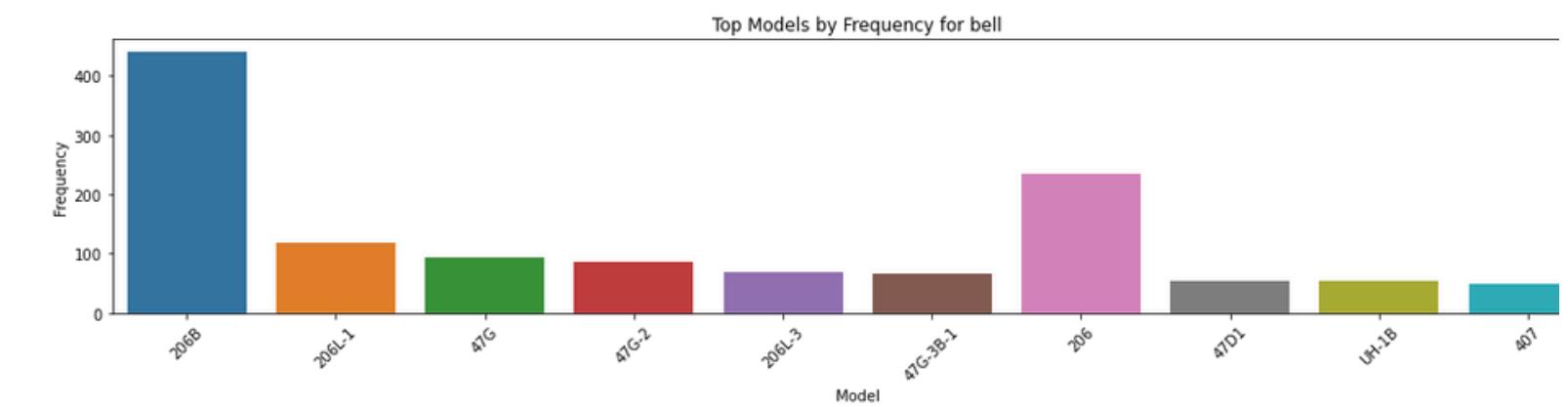
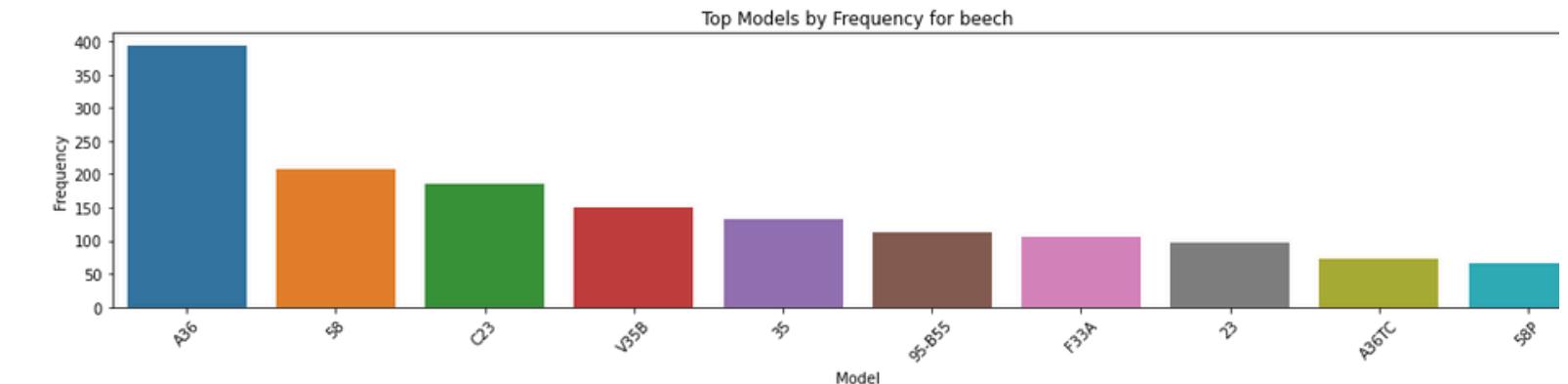
1. Cessna
2. Piper
3. Beech
4. Bell
5. Mooney
6. Grumman
7. Bellanca
8. Robinson
9. Hughes
10. Boeing

The higher numbers of Cessna and Piper aircraft might suggest that they are more popular in use around the world.



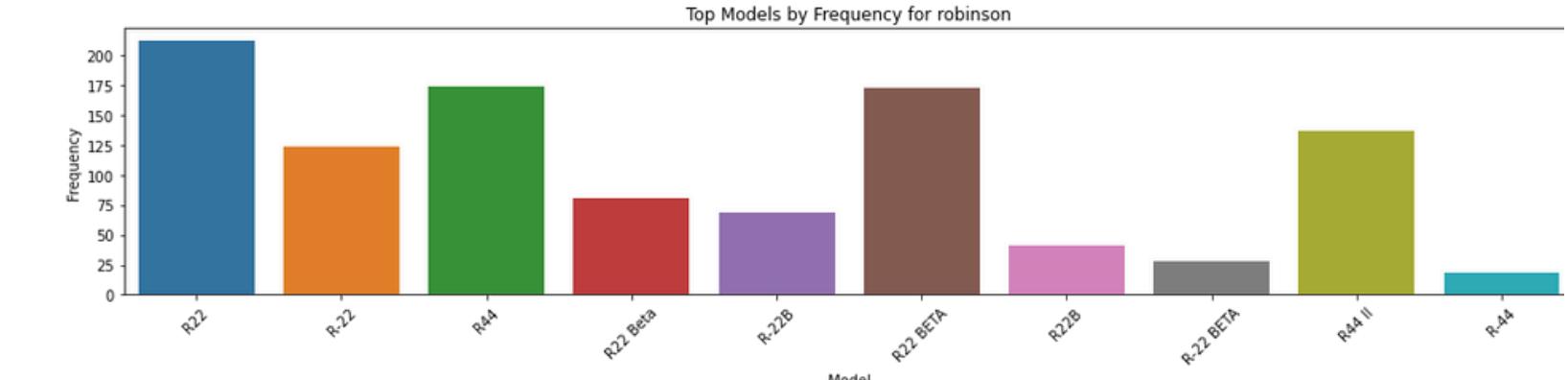
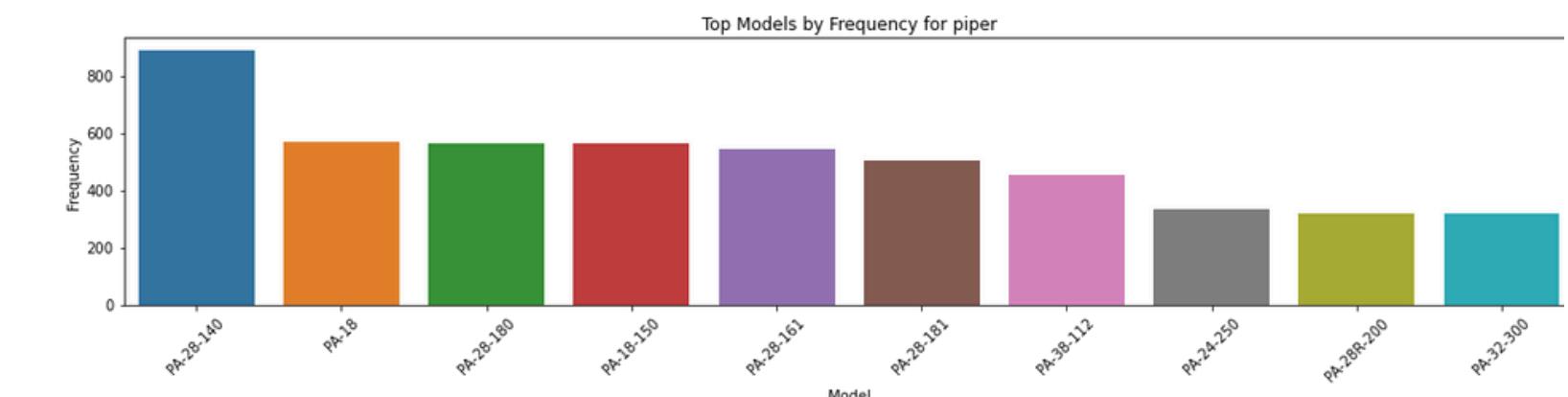
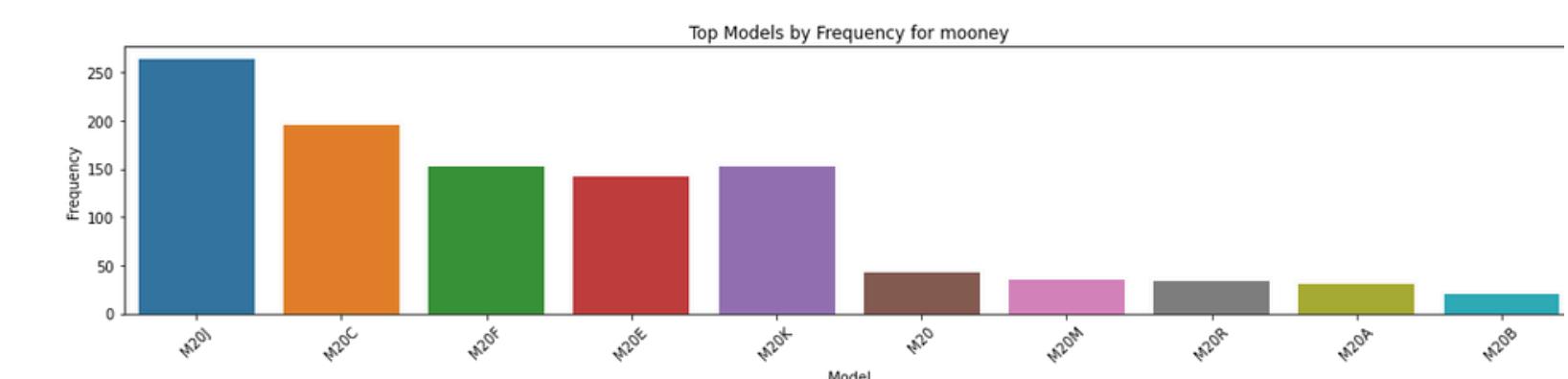
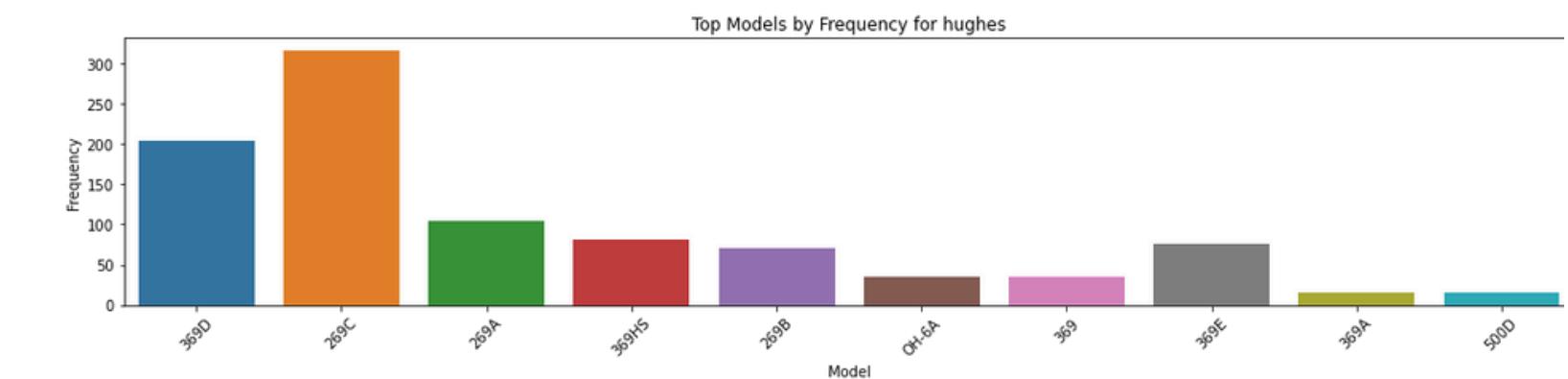
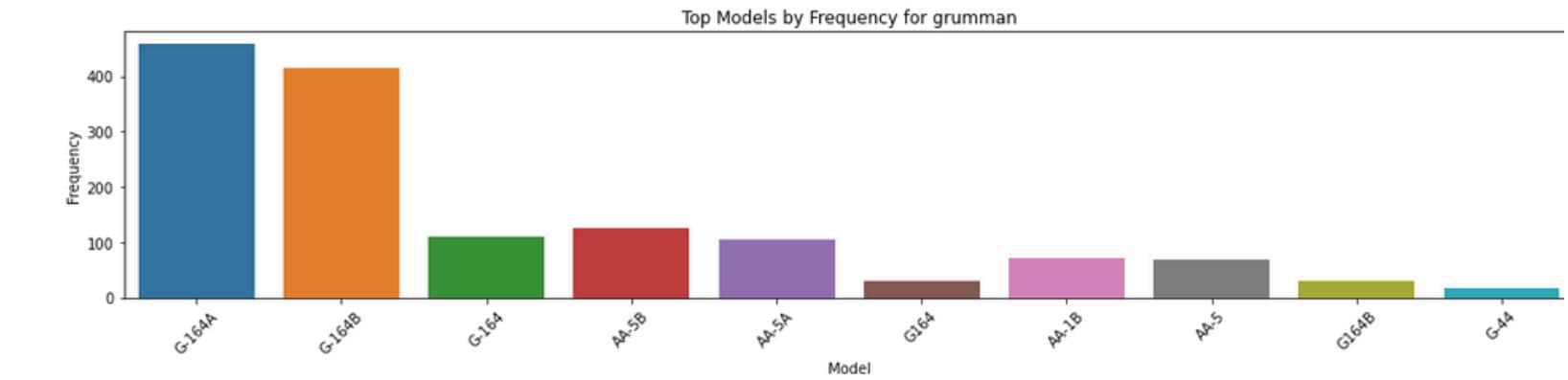
# Data Analysis

- I then identified the top models for each of the top 10 make that were involved in accidents
- This was to help narrow down the potentially higher risk models for each make.
- Some models have relatively low accident rates like Bellanca 17-31A with less than 50 instances.
- Other models like the Cessna 152 had very high accident rates with over 2000 instances.
- This could also be an indication that those models with extremely high accident rates are more popular.



# Data Analysis

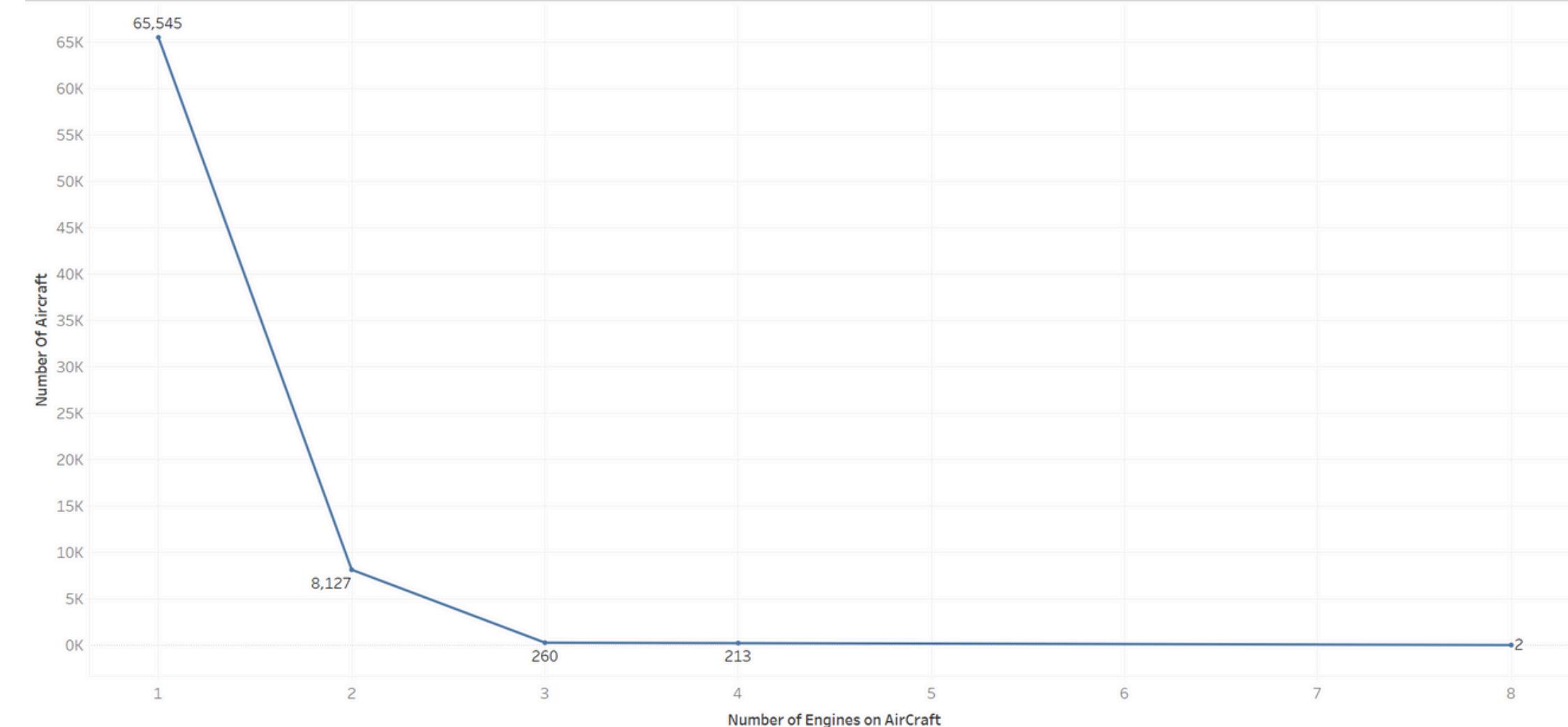
- Same trend can be seen for the other makes and their models.



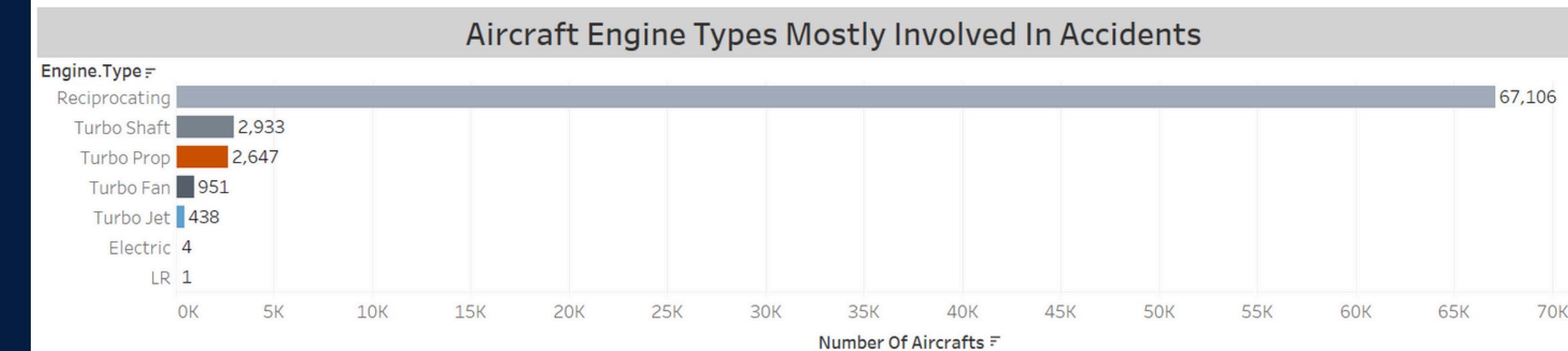
# Data Analysis

- Closer look at some of the features of the aircrafts involved in accidents showed that single engine aircraft had the most number of accidents.
- The Reciprocating engine type also had very high accident rates.
- Some engine types like Electric had very low number of instances but this can be attributed to it being an emerging technology in the aviation sector.

No. of Engines On Aircrafts Involved In Accidents

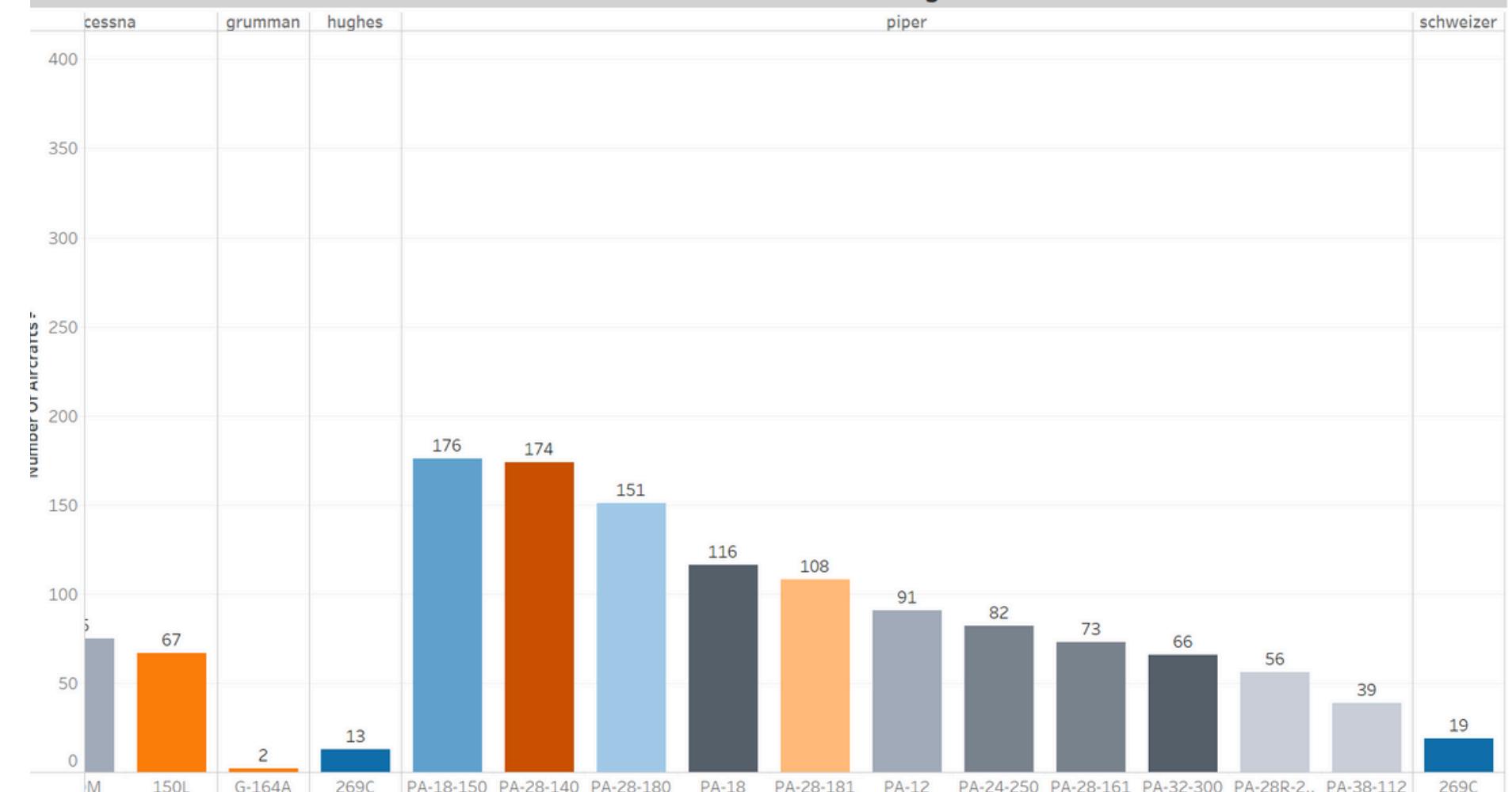
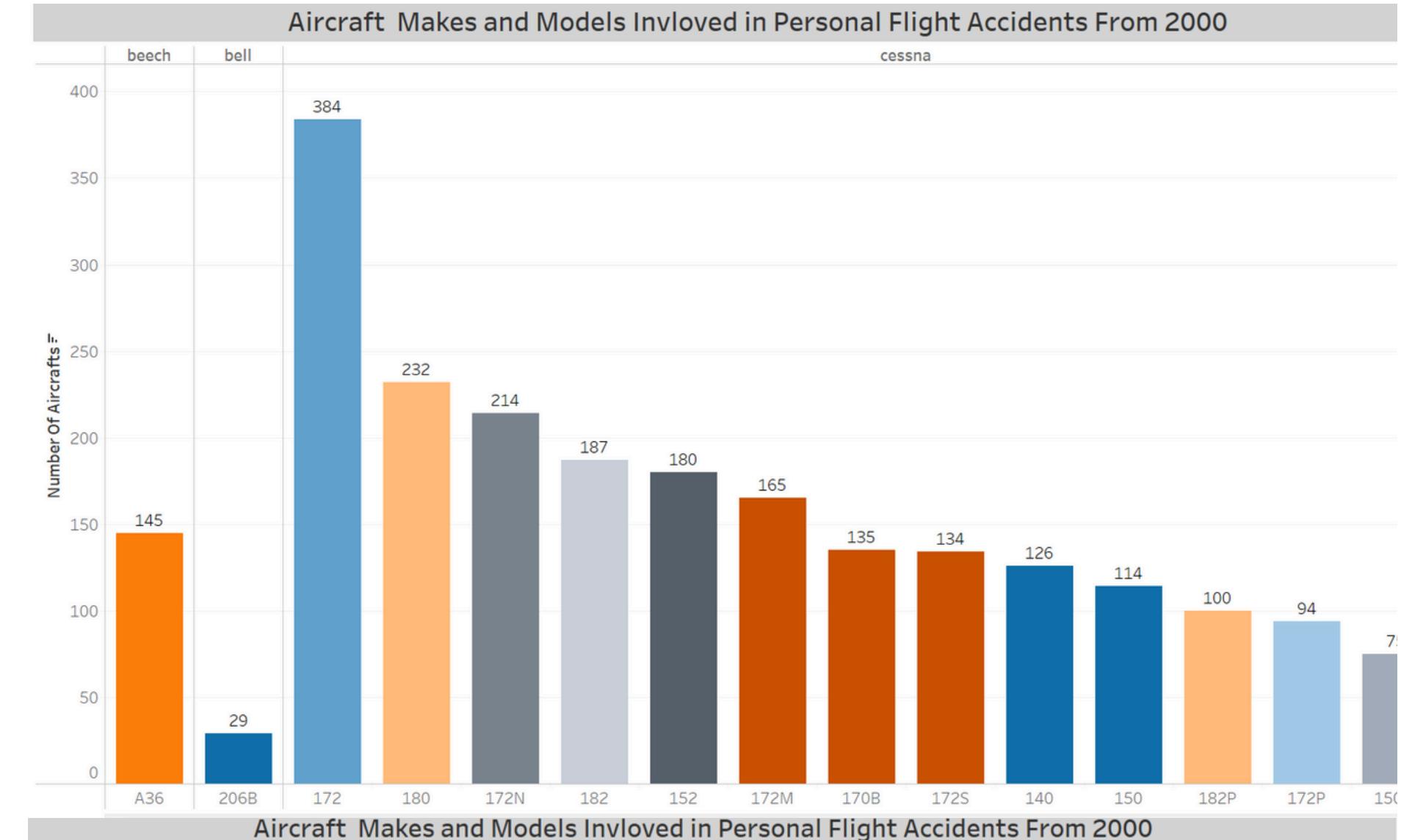


Aircraft Engine Types Mostly Involved In Accidents



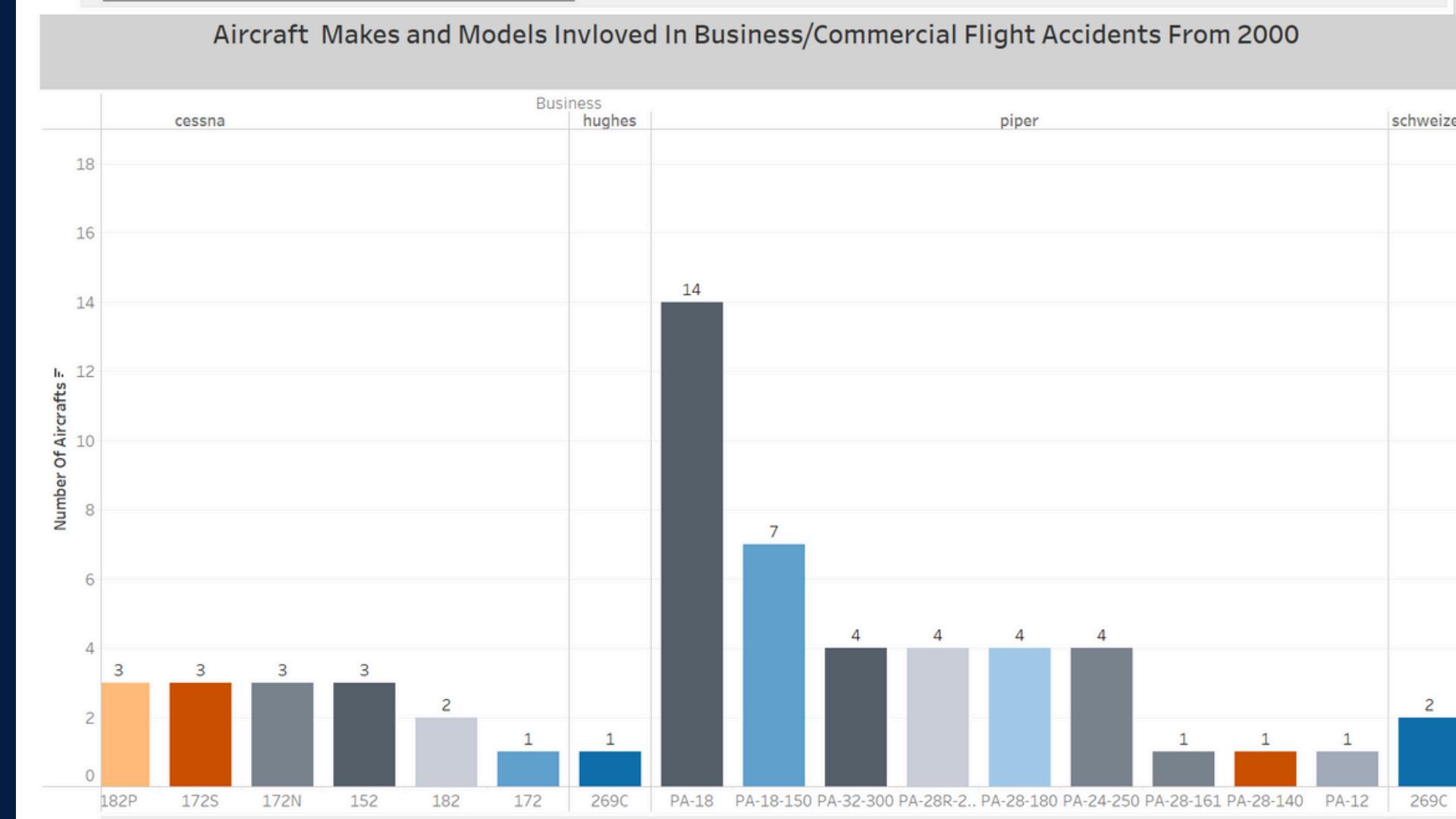
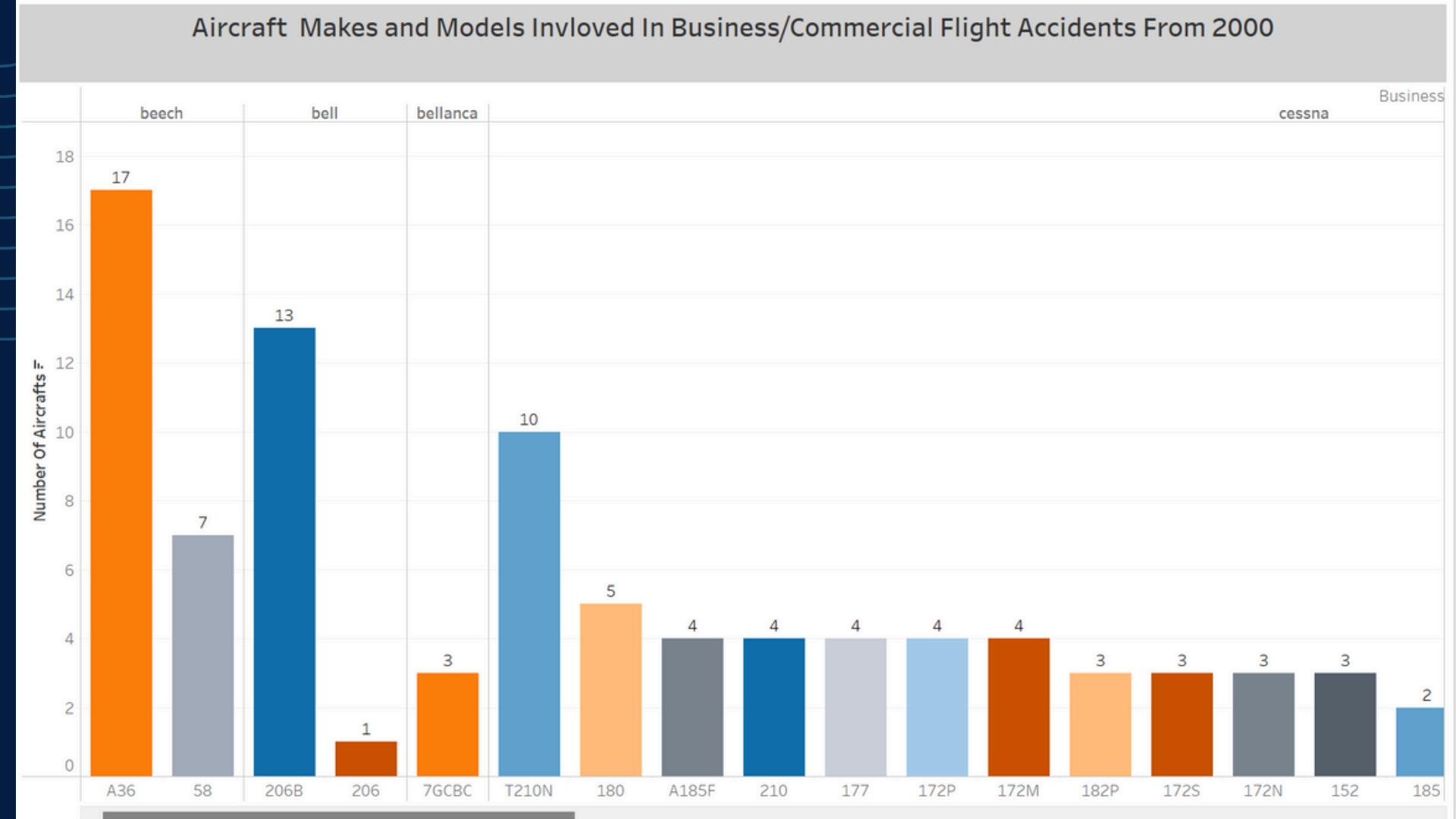
# Data Analysis

- Since Personal flights and Business/commercial flights are the main focus points for Rocinante Global, it is important to know the different models used for these purposes in the data and how they fair.
- I filtered the data from the year 2000 onwards in order to filter out old models that might not still be in production.
- For private or personal flights there were some specific models that were more common for private flights in the dataset.
- Cessna models 172, 180 and 172N were the most common aircraft for that make with the highest accident rates.
- Piper models PA-18-150, PA-28-140, PA-28-180 were most common for that make.



# Data Analysis

- There were fewer accidents involving business flights than private flights over the years.
- The most common models for business flights included Beech A36, Bell 206B helicopter and Piper PA-18.
- Cessna models T21ON, 340A were also the most common for the make.



# OBSERVATIONS

From these visualizations and others not included in this presentation, I was able to make some key observations.

## 1. Cessna and Piper aircraft makes are the most popular.

Their disproportionately large numbers in the dataset might be attributed to the popularity in use of their various models. This might be influenced by other external factors like price of the aircrafts, cost of servicing and maintenance, fuel economy e.t.c.

However, there are particular models of these makes that are high risk due to their appearance in various risk assessment categories.



# OBSERVATIONS

2. Single engine aircraft particularly with the Reciprocating Engine type are also very high risk.
3. The trend indicated that the more engines an aircraft has, the less probability it has of having an accident.
4. Amateur built aircraft sustained relatively more severe damage levels to non-amateur build aircraft, considering the ratio of number of records for both types of builds in the data.
5. There has been a slight uptrend in accidents involving personal and business flights from the year 2021.
6. Generally, the most common damage level was Substantial.
7. The Top 10 makes had over 40,000 aircrafts involved in accidents combined, while the bottom 20 had just 20 aircraft involved.



# RECOMMENDATIONS

---

Here is the summary of those models I assess to be overall very high risk :

**Beech:** ['A36', '58', 'C23', 'V35B', '35']

**Bell:** ['206B', '206L-1', '47G', '47G-2', '206L-3']

**Bellanca:** ['7GCBC', '17-30A', '7ECA', '8KCAB', '8GCBC']

**Boeing:** ['A75N1', 'B75N1', 'E75', 'A75', 'A75N1']

**Cessna:** ['152', '172', '172N', '172M', '150']

**Grumman:** ['G-164A', 'G-164B', 'G-164', 'AA-5B', 'AA-5A']

**Hughes:** ['369D', '269C', '269A', '369HS', '269B']

**Mooney:** ['M20J', 'M20C', 'M20F', 'M20E', 'M20K']

**Piper:** ['PA-28-140', 'PA-18', 'PA-28-180', 'PA-18-150', 'PA-28-161']

**Robinson:** ['R22', 'R-22', 'R44', 'R22 Beta', 'R-22B']

*Looking at those models popularly used for Private flights, and eliminating those that also appear in the high risk categories, I derived a list of low risk models popular for private use by market demand. These models have low accident rates and can be deemed safer to operate:*

**Cessna:** [182P, 172P]

**Piper:** [PA-28R-200, PA-32-300, PA-38-112 ]

**Bell:** [206B]

**Grumman:** [G-164A]

**Hughes** [269C]

*Using the same criteria as with private purpose planes, I filtered out the planes also present in the high risk lists to achieve the most common low risk aircrafts for commercial use:*

**Cessna:** [T206H, 525A, 421C, 208B]

**Piper:** [PA-30, PA-46-31OP ]

**De Havilland:** [DHC-2]

**Hughes:** [369-D]

**Cirrus:** [SR22]

# IMPROVEMENTS



There could be better selections from the dataset to address the business problem, but they would need more supporting external data.

The suggested models are NOT the only models or makes suitable, they are just the most commonly used, lowest risk aircrafts preferred for those purposes. The data suggests that they are more popular hence they may be more attractive to customers and clients.

To improve this project in the future, a machine learning algorithm can be trained to better analyse and give recommendations from the dataset.



# Thank's For Watching

Connect with us.



0741858977



Logankim62@gmail.com

