Phase 1 Project

BY ANDREW NYAKIBA

Business Overview

Our business has recently moved into the direction of the aviation industry for both commercial and private enterprises in order to expand its portfolio.

However, the aviation industry is well known aviation carries unique catastrophic risks—not only in terms of asset loss e.g. the airplanes, but also legal liability, brand reputation, and regulatory scrutiny.

To mitigate these risks, the company has taken a "Safety-First" kind of strategy by analyzing 60 years of NTSB accident data(1962-2023) with the aim to identify the aircraft makes and models which have a proven record of reliability and safety.

This approach that the company's first purchase will be a high performance and risk resilience aircraft.

Business Understanding

The business aims for;

- Low risk – Thus after analyzing the data the business can deduce the common causes for accidents for the specific models.

- Financial Security – Aircrafts that with a high "repairability" shall be prioritized.

- Good Fleet selection – Aircrafts models with "low risk" shall be used for both private and commercial enterprises

Data Understanding

- The source of the data is the National Transportation Safety Board (NTSB) which covers aviation accidents from 1962 to 2023.

- The dataset before filtering had 31 columns and 90348 rows.

- During the analysis certain rows were focused on such as "make" and "model".

- Most numerical columns were correctly formatted as float type of data ,however, other columns required conversion from object type data to float e.g. "number.of.engines"

- It was noted that the data had high number of missing values and inconsistent text. These problems were dealt with during the analysis.

Analysis and Results

- The data was first filtered to remove the data that are older than 30 years in order to deal with more modern aircrafts.

- Aircrafts with an amateur build were removed next in order to deal with those that have been professionally manufactured.

- The analysis was made to solve three problems;

1. Manufacturer Survivability

The data was grouped by "make" to find the safest companies by calculating the fatality rate.

It was then discovered that companies like Cessna have high survival rate due to their widespread use despite having many incidents.

2. Mechanical Redundancy

 Here the impact of number of engines on survivability was analyzed.

 The survivability between one and two engine airplanes were compared. There was an increase in the survivability when using a twin-engine airplane.

3. Asset Selection

 By combining the first two findings, the top 10 specific airplane models could be found.

The filtered twin-engine models were sorted by the highest survival rate with a minimum accident count to ensure statistical reliability.

• The statistical methods used to perform the data analysis were;

I. Aggregation- which turned individual accidents into manufacture-level statistics using

.groupby() and .agg() .

II. Feature Adding- creating a new column(Total.Injuries) by summing injured and

uninjured counts.

III. Data Standardization- which ensured that data with mixed cases were treated

equally using .str.upper()

Recommendations

Make	Model	Historical Fatal Injuries	Historical Injuries	Accident_Count	Survival_Rate
CESSNA	402C	0	76	15	100
Boeing	737	0	3051	33	100
CESSNA	560XL	7	82	19	91.46341463
BEECH	76	3	35	16	91.42857143
CESSNA	550	6	57	16	89.47368421
BEECH	200	8	72	17	88.88888889
CESSNA	414	4	36	15	88.88888889
Piper	PA-44-180	2	15	15	86.66666667
CESSNA	421B	11	59	18	81.3559322

1) Based on survival rates and manufacturer consistency, we recommend prioritizing the acquisition of the **Cessna 402C** or the **Boeing 737**
2) Prioritize the acquisition of **multi-engine aircraft** for the business fleet. This mitigates the risk of fatal outcomes during mechanical failures, providing a "safety buffer" essential for corporate operations.
3) Limit the procurement search to manufacturers with **proven market experience** . High-volume manufacturers have more robust maintenance networks, better parts availability, and more predictable safety profiles than newer brands.

Thank You for your attention

ANY QUESTIONS?