Aviation Accident Data Analysis

INTRODUCTION

This project aims to analyze aviation accident data to enhance safety in aircraft purchasing decisions. Utilizing the **Aviation Accident Database & Synopses** up to 2023, we investigate factors contributing to aircraft accidents, focusing on key questions such as which aircraft models have the highest accident rates, the reliability of different engine types, and how injury severity varies by aircraft model.

Through data cleaning and analysis, we employ **Tableau** to create an interactive dashboard that visualizes trends and patterns in the data, providing insights into accident causes and risk factors. The outcomes of this analysis will support stakeholders in making informed decisions about acquiring aircraft with lower associated risks, ultimately contributing to improved aviation safety.

Project Overview

Objective: To analyze aviation accident data to determine the lowest-risk aircraft for purchase.

Scope: Evaluation of accident frequency, severity, and factors contributing to accidents.

Business Understanding

- **Purpose:**
  1. To enhance safety and minimize risks associated with aircraft operations.
  2. To inform decision-making regarding aircraft purchases.
- **Key Questions:**
  1. Which aircraft models cause a lot of accidents?
  2. What is the most reliable engine type in terms of accident rates?
  3. How does injury severity vary by aircraft model?
  4. Which aircraft models have the least incidence of accidents based on their make?

Data Description

The dataset includes aviation accidents and incidents up to 2023, with the following key columns: Event.Id, Accident.Number, Event.Date, Location, Country, Injury.Severity, Aircraft.Damage, Aircraft.Category, Make, Model, Engine.Type, Purpose.of.Flight, Weather.Condition, Broad.Phase.of.Flight, Total.Fatal.Injuries, Total.Serious.Injuries, Total.Minor.Injuries, Total.Uninjured. Data Cleaning Process The raw data required extensive preprocessing, including: Duplicate Removal: Removed duplicate entries for accidents. Missing Data Handling: Used mean, median, and mode imputation to fill missing values. Unnecessary Columns Dropped: Columns not relevant to the analysis were removed to streamline the dataset. The final cleaned dataset allowed for focused analysis on aircraft models, locations, injury severity, and other critical factors.

Data Understanding

Dataset: Aviation Accident Database & Synopses (up to 2023)

Key Variables:

- Event.Id, Investigation.Type, Location, Injury Severity, Aircraft Model, Engine Type, etc.

Data Cleaning Steps:

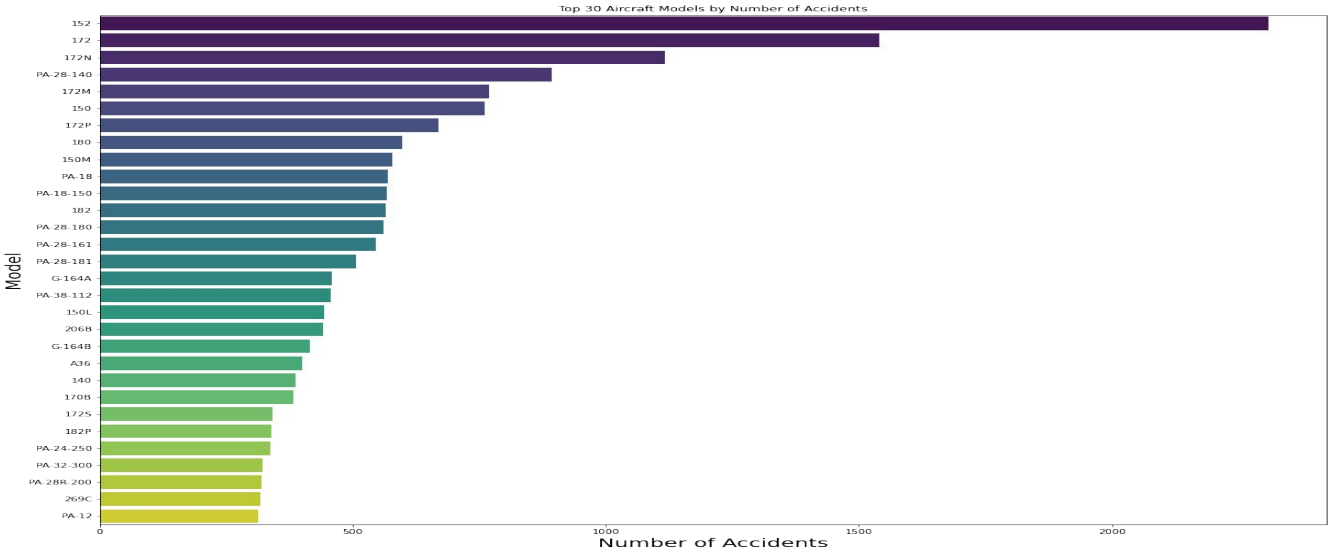
- Removed duplicates.
- Handled missing values using mean, median, and mode.
- Dropped unnecessary columns.

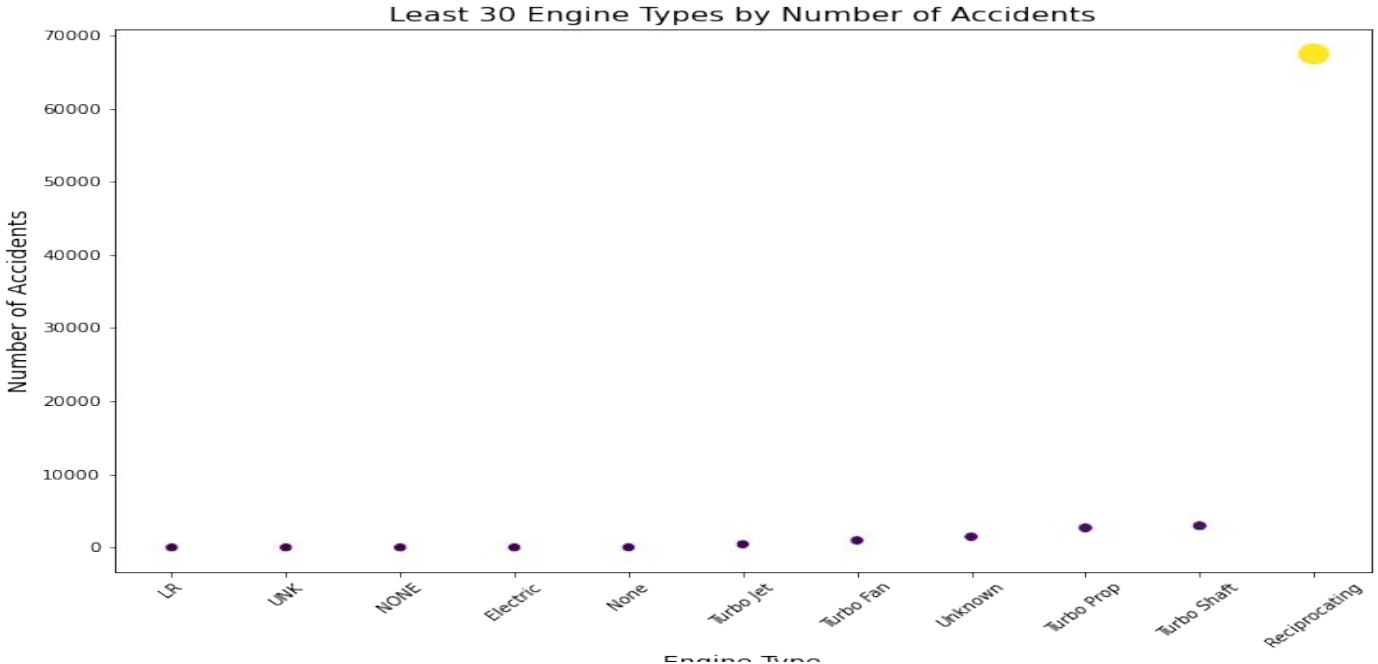
Data Analysis

Analysis Techniques:

- **Tableau Visualizations:**
  - Box Plots and Bar Charts to analyze accident rates by aircraft model and injury severity.
  - Filters applied to focus on significant data points.
- **Key Findings:**
  - Trends in accident frequency by aircraft model.
  - Correlation between engine types and accident rates.

1. **Which aircraft model causes a lot of accidents?**




Least 30 Engine Types by Number of Accidents
Number of Accidents
70000
60000
50000
40000
30000
20000
10000
0
LR
UNK
NONE
Electric
None
Turbo Jet
Turbo Fan
Unknown
Turbo Prop
Turbo Shaft
Reciprocating

RECCOMENDATIONS

Purchase Considerations:

- Favor aircraft models with lower accident rates.
- Consider engine types with higher reliability.

Safety Enhancements:

- Implement training and maintenance programs based on accident data insights.

Next Steps

- **Future Research:**
  - Expand analysis to include factors like pilot experience and weather conditions.
- **Implementation:**
  - Collaborate with safety teams to integrate findings into purchasing decisions.

Thank you

- **Questions?**
- **Contact Information:**
  - [Samuel Yashua]
  - [Your LinkedIn Profile]
  - [samuel.yashua@student.moringaschool.com
  -]

FINISHED