AVIATION ACCIDENT DATA ANALYSIS

BY: IVET BUYAKI
11TH FEBRUARY 2025

Overview

The project entails use of data cleaning, imputation, analysis, and visualization to generate insights for a business stakeholder.

Outline

- Business Understanding
- Data Understanding
- Data Analysis
- Recommendations
- Next steps
- Overview

Business Understanding

The client aims to expand into new industries, with a focus on purchasing and operating airplanes for commercial and private enterprises. The project entails an analysis of aviation accident data from 1962-2023 with an objective of identifying the aircraft with the lowest risk.

Data Understanding

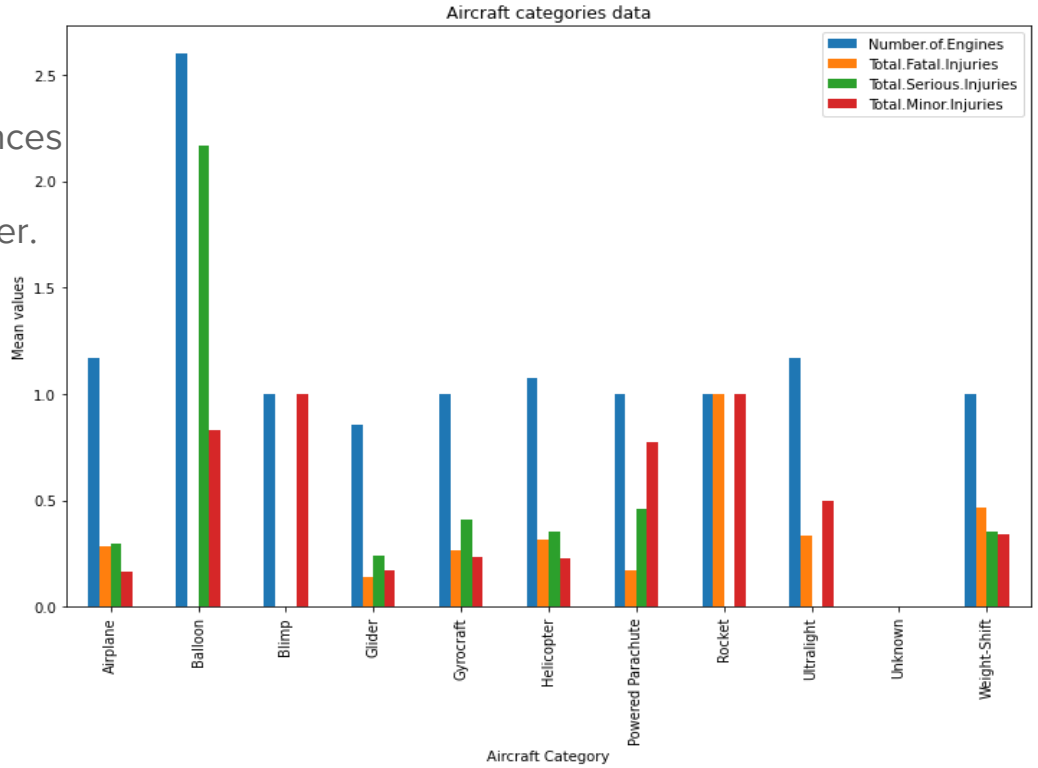
The data set,obtained from Kaggle and issued by the National Transportation Safety Board on accidents in the United States and international waters. It's composed of different types of aircrafts, ranging from airplanes, helicopters, gliders to Gyrocraft. The data set captures the number of fatal, serious and minor injuries that occurred within that period as well as the possible influencing factors and prevailing conditions such as the weather. The raw data had 88,889 rows and 31 columns, on cleaning they reduced to 10011 rows and 25 columns. Some missing columns were dropped such as the schedule as they don't have an impact on the analysis, while other missing values such as the number of injuries were imputed with the mean since the existing values seemed to be clustered around the mean.

Data Analysis

- The aircraft categories data shows

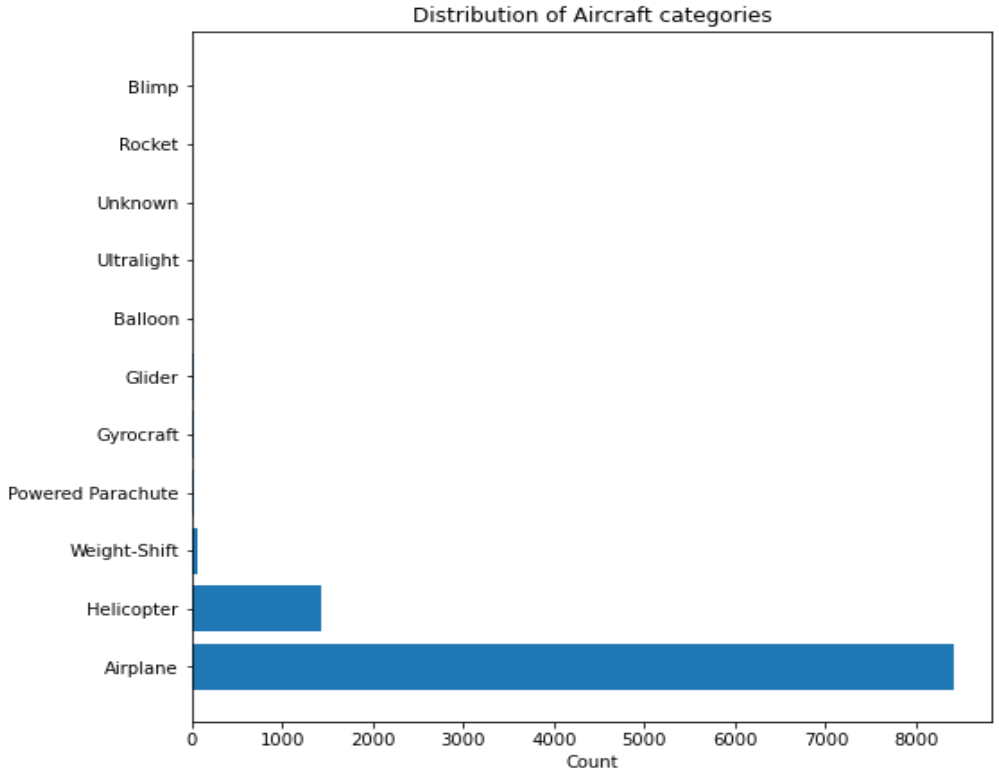
the three lowest average accident occurrences

aircraft are the glider, airplane and helicopter.



Data Analysis

- Airplanes seem to be more common, followed by helicopters then the Weight-shifts.



Insights and Recommendations

INSIGHTS

- The airplane is the most commonly used aircraft from the presented data set, followed by the helicopter.
- The average rate of occurrence/ mean of accidents of the airplane is almost at the same range as the other fewer planes. The three lowest are the glider, airplane and helicopter.
- A lot of accidents occurred when the Visual Meteorological condition was in place.

RECOMMENDATIONS

- The three lowest risk aircraft are the airplane, helicopter and glider. This is from the analysis of the data set provided.
- Further analysis and data be provided on the relation between the Visual Meteorological condition and accidents. Instrument Meteorological conditions seem to be a safer option.
- Further research be done on the relation between the engine type and accident occurrence.

Thank You!

Email: ivet.buyaki@student.moringaschool.com
GitHub: @Ivet566