PHASE ONE PROJECT

Business Understanding

With respect to diversification, the organization has entered the aviation industry seeking to purchase and manage aircraft to provide commercial and private air transport. Growth in the aviation industry has largely attracted investors but at the same time bringing in various operational and financial limitations terms of reliability and safety of aircraft. It is imperative to state that the company is concerned with aircraft investment in low-risk aircraft models to be held as their assets for starters to limit liabilities, minimize operational costs as well as providing safety of operation.

Task at hand is to elicit and examine the critical risk determinants of various models of aircraft which involves looking at the past records of accidents, the chances of the aircraft experiencing mechanical failures, the extent of it's maintenance and other such information which would help the organization's new venture into aviation.

These findings will be. Used to assist in making purchase decisions regarding the types of aircrafts in consideration of the safety operations, cost of acquisition and future returns as the company ventures into the new business.

Problem Statement

The main goal of the company at hand is diving into the aviation industry to diversify their portfolio.
With this project I will help the compnay to consider and achieve a number of things including;

- Identifyingaeroplanes with the least historical rating of risks
- Understand trens, patterns and causes of aircrafts based accidents
- Provide a comprehensive purchase and operation procedures datasheet of aircrafts

Metric of Success

Some of the considerations I'm putting in place to ensure the project is a success include

Ensuring accurate risk assessment. Accurate rate of the predictive models of the analytical tools and also targeting a high accuracy rate.

By plotting graphs with actual data this will give a clear picture of what to consider then investing in an aero plane and it will help in clearly indicating the do's and don'ts while engaging in the aircraft industry.

Data Understanding

The dataset contains 88,889 entries with 31 columns, which include information about aviation incidents. Here's a summary of the key columns and their contents:

- Event.Id: A unique identifier for each aviation event.
- Investigation.Type: Indicates if the event was an accident or an incident.
- Accident.Number: A number associated with each accident or incident.
- Event.Date: The date when the event occurred.
- Location: The geographic location where the event took place.
- Country: The country where the event happened.
- Latitude/Longitude: Geographical coordinates for some events (many entries are missing).
- Airport.Code/Name: Information about the airport, if the event occurred near one.
- Injury.Severity: Severity of the injuries (if any) resulting from the event.
- Aircraft.damage: Indicates the extent of damage to the aircraft.
- Aircraft.Category: Category/type of the aircraft.
- Registration.Number: Aircraft registration number.
- Make/Model: Manufacturer and model of the aircraft involved.
- Amateur.Built: Whether the aircraft was built by an amateur.
- Number.of.Engines: Number of engines on the aircraft.
- Engine.Type: Type of engine used.
- Purpose.of.flight: Purpose of the flight (personal, commercial, etc.).
- Total.Fatal.Injuries/Serious Injuries/Minor Injuries/Uninjured: Injury counts from the event.
- Weather.Condition: Weather conditions at the time of the event.
- Broad.phase.of.flight: Phase of flight during the event (e.g., takeoff, landing, cruise).
- Report.Status: Status of the investigation report.
- Publication.Date: Date when the report was published.
- Head (First 5 Rows): The dataset starts with entries from aviation accidents ranging from 1948 to 1979, showing fields like location, country, injury data, weather conditions, and flight phase.
- Shape: The dataset contains 88,889 rows and 32 columns.
- Tail (Last 5 Rows): The most recent records include accidents from 2022, with similar fields showing details about the event, injuries, and other contributing factors.

Data Preparation

Cleaning the data
Understand the type of dataset you are working with.

- converting any data that needs conversion for easy access and readability
- detect outliers and work on them accordingly
- Handle missing data (imputation or removal).
- Dropping missing data
- Clean and standardize data
- Derive new features
- Detect and handle outliers.
- standardize the data if necessary
- Remove duplicates.
- Select relevant features and drop unnecessary ones.

Plotting graphs for visualization

For better graphical illustration different types of graphs need to be done

This will also aid in making a sound decision dictated by visible data metrics.

This will also give insights to easily grasp and understand visual patterns.

Engaging visuals will capture the attention of the companies decision makers