mwakad /
Data-Driven-Decision-Support-for-Aircraft-Procurement

Code Issues Pull requests Actions Projects Wiki Security Insights S

The repository is for a project that aims to support a company's data-driven decision to procure relatively low-risk airplane models to operate the fleet in the commercial and private enterprise sectors. It analyzes a dataset compiled by the National Transportation Safety board

0 stars **0** forks **1** watching Branches Activity
Tags

Public repository

1 Branch **0** Tags Go to file t Go to file Add file Code

mwakad Update README.md 41f110d · now

Data	update	8 hours ago
Images	update	8 hours ago
.gitignore	update .gitignore	yesterday
Index.ipynb	updated	5 hours ago
Notebook.pdf	updated	5 hours ago
Presentation.pdf	update	1 hour ago
README.md	Update README.md	now

README

Data-Driven-Decision-Support-for-Aircraft-Procurement

Project Summary

The commercial aviation sector is a highly sensitive industry. Although air travel is the safest means of transport, passenger survival rate is the lowest for aircraft accidents. This fact amplifies the correlation between a company's competitiveness and its aircraft accidents/ incidents record. Prevention is better than cure, and the foundational strategy in guaranteeing the longevity of the company's aviation division is to procure safe aircraft. This project proposes three data-supported recommendations on the safest aircraft models the company should procure and operate after entering the commercial aviation industry.

Data Understanding

This project aims to support the company's data-driven decisions to procure relatively low-risk airplane models to operate the fleet in the commercial and private enterprise sectors. It analyzes a dataset compiled by the National Transportation Safety board and retrieved from https://doi.org/10.34740/KAGGLE/DSV/4875576. Recommendations on the safest, practical, and logistically viable aircraft (model and make) for various aviation services the company can venture into to diversify its portfolio are backed up by descriptive statistics and justified by interactive visualizations created using Tableau Public.

Features: Prior to data cleaning, the original dataset comprises 31 columns and 88889 rows. The columns are meticulously organized to capture variables of interest in aircraft accidents and incidents. Extracting, analyzing, and visualizing relevant data from the dataset is vital to shedding insight into the safest, low-risk airplane models for a new entrant to the commercial aviation industry.

Target: For this project, the target variables categorize into Dimensions and Measures. The Dimension variables include: Event.Date, Investigation.Type, Aircraft.Damage, Location, Make, Model, Engine.Type, Weather.Condition, and Purpose.of.flight. The Measure variables include: Number.of.Engines, Total.Fatal.Injuries, Total.Serious.Injuries, Total.Minor.Injuries, and Total.Uninjured. The other columns are deemed inapplicable for this project and dropped prior to data cleaning processes.

Problem Statement

Operating airplanes for commercial and private enterprises is a potentially profitable portfolio diversification strategy for the company. However, venturing into such a highly sensitive sector necessitates data-driven decisions, strategic implementation, and formative performance evaluation. Thankfully, digital maturity in leveraging novel data analytics is a key driver for aircraft safety. The insights yielded by this project have a significant impact in supporting the company's data-driven decisions in procuring a safe aircraft fleet to operate in a highly sensitive sector.

Business Objectives

1. **Goal:** To make data-driven decisions on the purchase of airplanes with high safety ratings to operate commercial flights and offer aviation services to private enterprises.
2. **Specific Objectives:**

- To identify the least safe aircraft the company should refrain from purchasing.
- To identify the relatively safer single-engine, and multi-engine aircraft the company can consider purchasing.
- To recommend the safest and most operationally feasible aircraft models the company should procure to operate in specific applications (`Purpose.of.flight`).

Requirements to Meet Objectives

1. **Load the Data**

- Load a .csv dataset, create the working `df` and inspect imported data.

2. **Perform Extract Transform Load (ETL) processes**

- Apply lambda functions to perform data transformations (drop entries with unknown values, eliminate duplicates, convert data types, and data aggregation).

3. **Imputing missing values**

- Perform descriptive statistics on (float and int) variables to determine the appropriate measure of central tendency to impute for missing values.

4. **Analyzing the cleaned dataset**

- Applying the **.groupby()** method to create subgroups of a DataFrame suitable for analyzing and visualizing the relationship between specific aircraft makes and models and their potential correlation to fatality or survival rates in the event of accidents/ incidents.

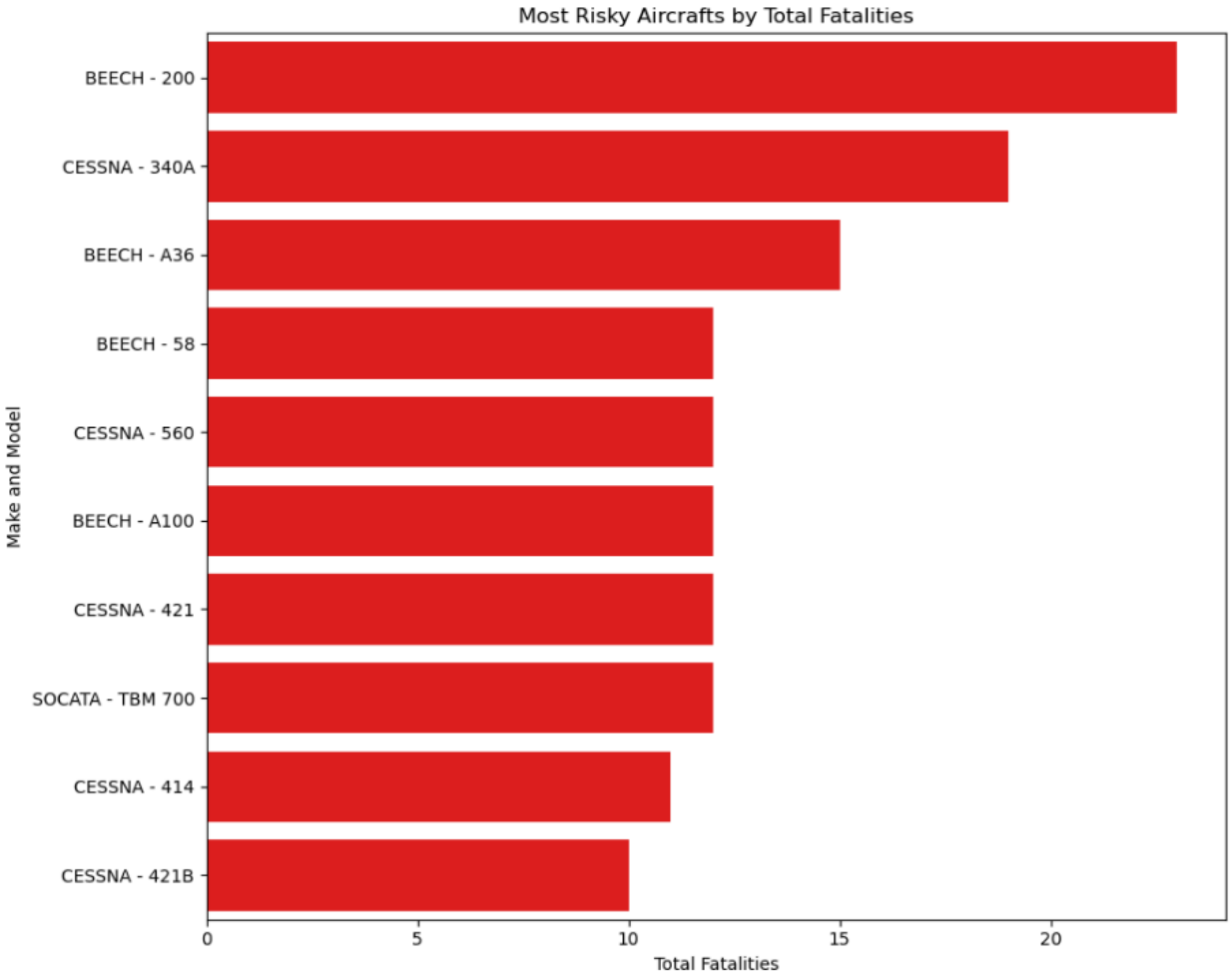
5. **Deducing recommendations based on data analysis findings**

- The grand aim for the project is met by recommending the safest aircraft the company should procure based on respective applicability, and operational feasibility for specific aviation services the business will be offering upon entry to the industry.

Objective 1

Least Safe Aircraft Models

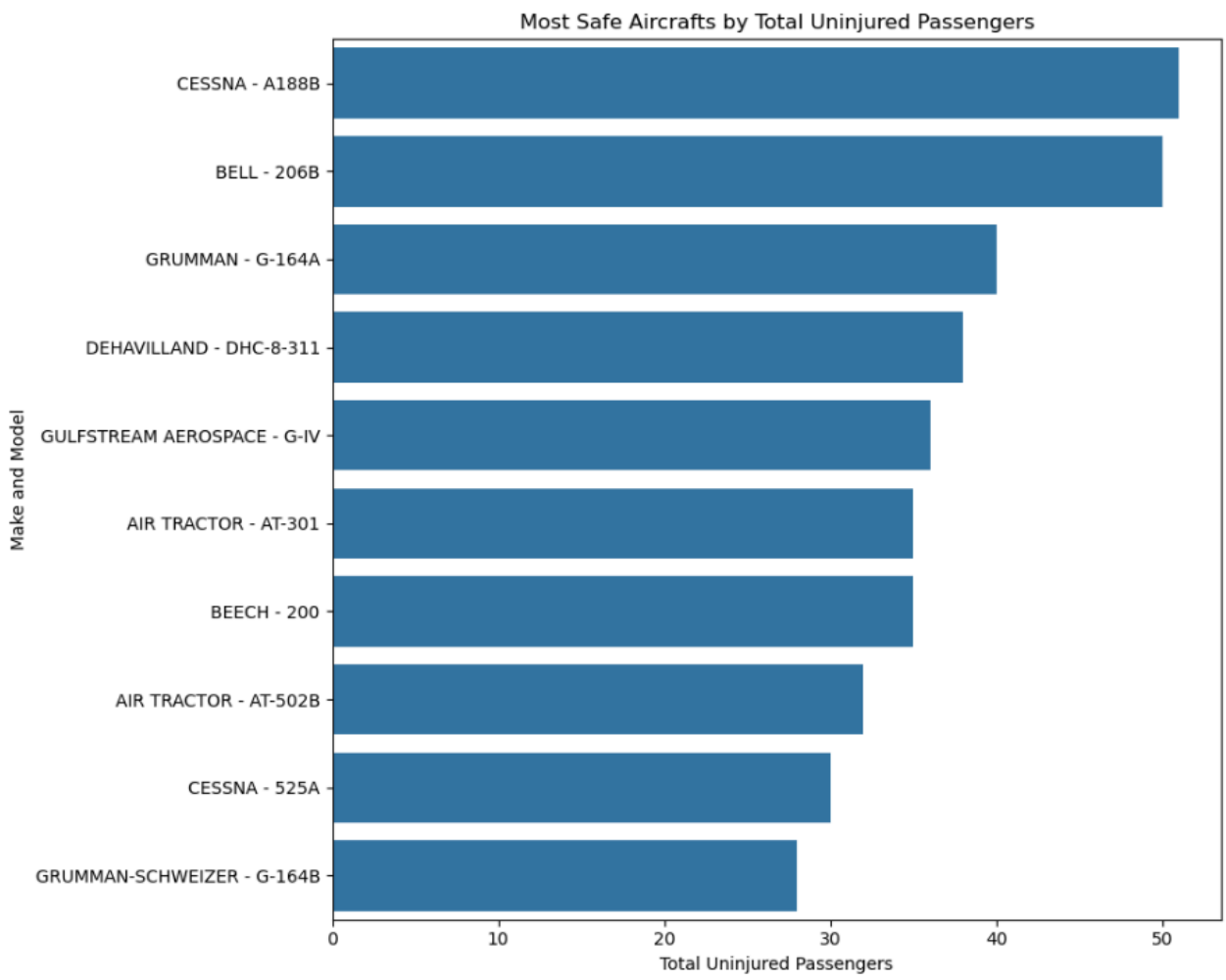
- The `.groupby()` method is used to group the cleaned DataFrame (`df_clean`) by `Make` (the manufacturer) and `Model` (a specific aircraft model).
- Each unique combination of an aircraft's Make and Model is treated as a single group to avoid ambiguity since aircraft models made by different manufacturers can share the exact model name or a similar str-num identifier.
- Grouping an aircraft's `Make` and `Model` optimizes the readability of plotted visualizations.
- The aggregate sum of fatalities `Total.Fatal.Injuries` for each `Make` and `Model` are sorted in descending order (ascending = False) and filtered (head = 10) to select the top-10 least safe aircraft in terms of respectively attributed summation of fatalities.



Objective 2

Most Safe Aircraft Models

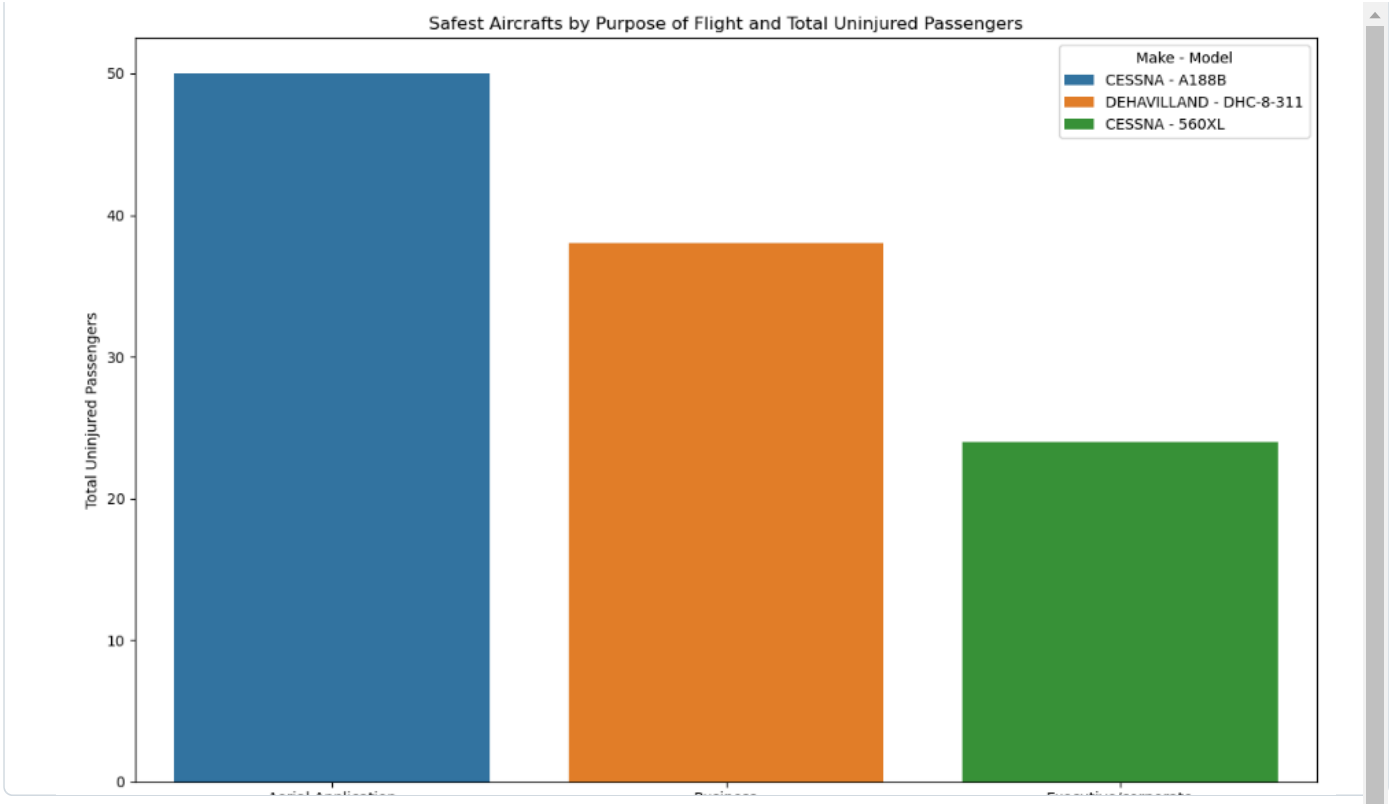
- The aggregate sum of survivors `Total.Uninjured` for each `Make` and `Model` is sorted in descending order (ascending = False) and filtered (head = 10) to select the top-10 safest aircraft in terms of cumulative survivors from aircraft accidents and incidents.



Objective 3

Recommended Aircraft Model to Purchase

- The company's portfolio diversification strategy involves purchasing and operating an airplane fleet for commercial and private enterprises.
- Per the company's target market, its new aviation division will majorly offer services spanning Executive/corporate flights, Business flights, and Aerial applications.
- Owing to unique operational complexities and logistical considerations, the safest aircraft models are not universally applicable.
- Thus, the dataset is filtered for the three relevant `Purpose.of.flight` aviation services the company will typically offer its clients.
- The `.groupby()` method is used to group each of the filtered `Purpose.of.flight` aviation service, `Make`, and `Model` to determine the safest aircraft model for each application.
- The `make` and `Model` with the highest aggregate sum of survivors `Total.Uninjured` for each of the selected `Purpose.of.flight` is selected for visualization.



Releases

No releases published
Create a new release

Packages

No packages published
Publish your first package

Languages

- **Jupyter Notebook** 100.0%