Aviation Data cleaning and Visualization

—

By Mwangi John Eric Chege

Evaluating the Aviation data to help our business stakeholders to make the best decision in investing to an aircraft

Introduction

Project Overview

The project focuses on utilizing data cleaning, imputation, analysis, and visualization methods to extract valuable insights for a business stakeholder. The aim is to ensure data quality and present findings that facilitate effective decision-making.

Business Problem

The company aims to diversify by entering the aviation industry through the acquisition and operation of aircraft. However, they face a critical challenge due to their lack of expertise in aircraft-related risks. Your role is to assess and identify the low-risk aircraft options and provide actionable recommendations to guide the head of the aviation division in making informed purchase decisions

Business Understanding

Our business expanding to something greater, we are looking forward to guide our stakeholders to buying a new aircraft one that has the lowest amount of risk and causes the least amount of accidents

- Expansion into New Industries: The company is diversifying its portfolio by entering the aviation sector.

- Lack of Knowledge: There is a significant gap in understanding the risks associated with aircraft operations.

- Risk Assessment: The organization needs to ascertain which aircraft carry the lowest risk.

- Actionable Insights Required: Findings must be translated into practical recommendations for decision-makers.

Business Solution

Being able to identify the risks and problems at hand, being able to offer various data visualizations to give a clear indication of what we can deal with

- With the proper data visualization tools we are able to portray the possible aircrafts that give the lowest possible risk

- Being able to assess the data with the tools that we have we can clear off and narrow down to the specific aircrafts that benefits us accordingly

Tools for Solution

With the proper tools such as matplotlib and tableau we are going to be able to work with the data that we have collected, to help give solutions to our business problems

- In Matplotlib we are going to use both scatterplot and and bar graphs to enhance our narrowing down of what is expected. With this tool we are going to get the proper visualization and data cleaning and filtering processes easily managed and done

- With tableau it is going to essentially help us create good data visualizations for our bosses to asses and also convince them on what aircraft is to be invested in next

Hypothesis

Data and relevant variables

Hypothesis support

With the use of matplotlib...

We are going to narrow down our search from large pool of data that has been presented to us and also give a visualization of total number of injuries against the number of engines to be able to know how many number of engines that we need for our aircraft.

Furthermore, we narrowed down our search with a bar graph plotting our aircraft model and the number engines, this way we will identify which specific model we need

With Tableau...

We are going to get better visualizations of how the aircrafts that we have selected after using matplotlib to clean our data and this way we are going to have a clearer view of what is taking place and which aircrafts and their models perform the best and come with the least amount of risks, making them the best to invest in and purchase

Hypothesis support

With the use of matplotlib...

With matplotlib, it shows us the relationship of engines and total fatal injuries caused and this way we identified that the more the engines the safer it is and less prone to risks

Furthermore, with matplotlib and data filtering and cleaning tools we are able to identify that most of the models with more engines offer the least amount of riks
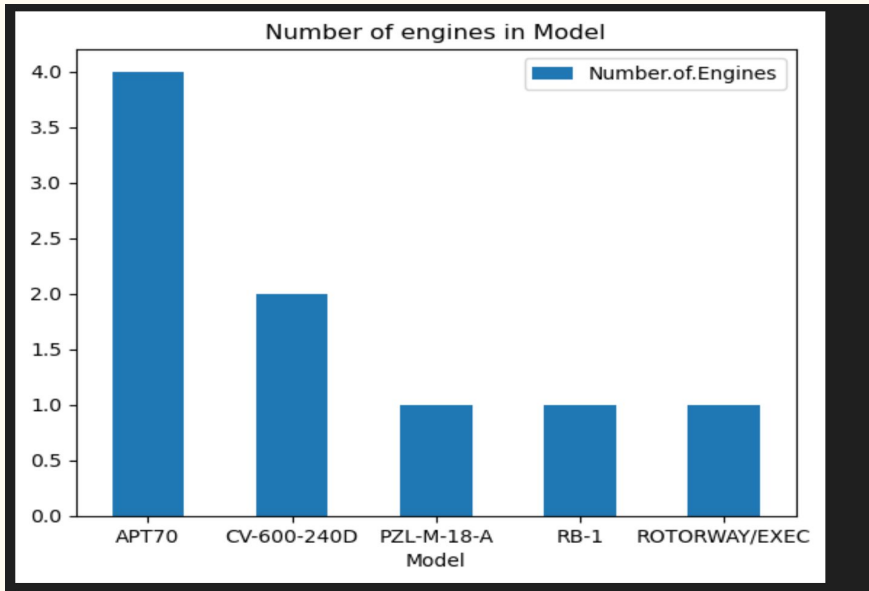
With Tableau...

We are going to get better visualizations of how the aircrafts perform from different models and with different number of engines. This way we are going to have a clearer view of what is taking place and which aircrafts and their models perform the best and come with the least amount of risks, making them the best to invest in and purchase
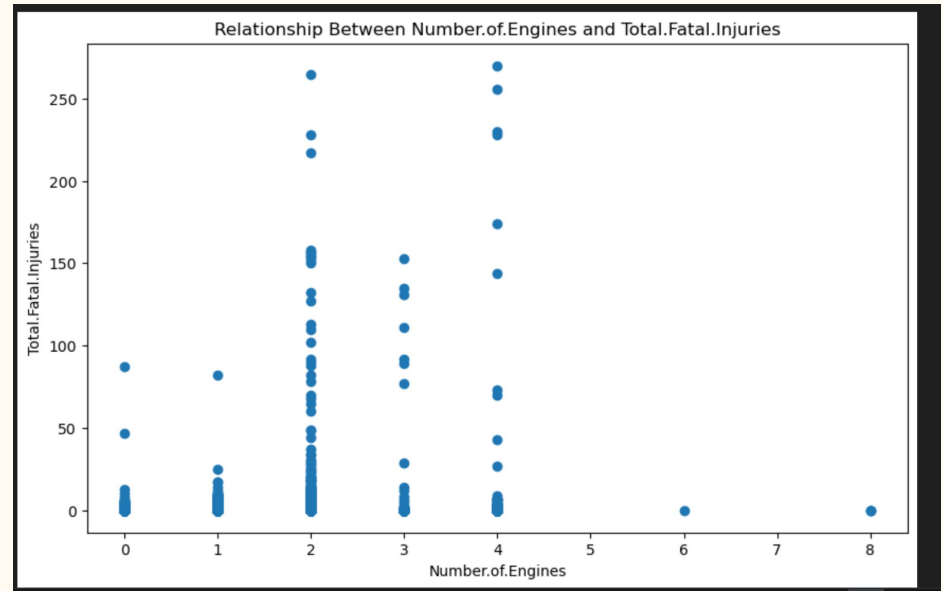
Hypothesis support
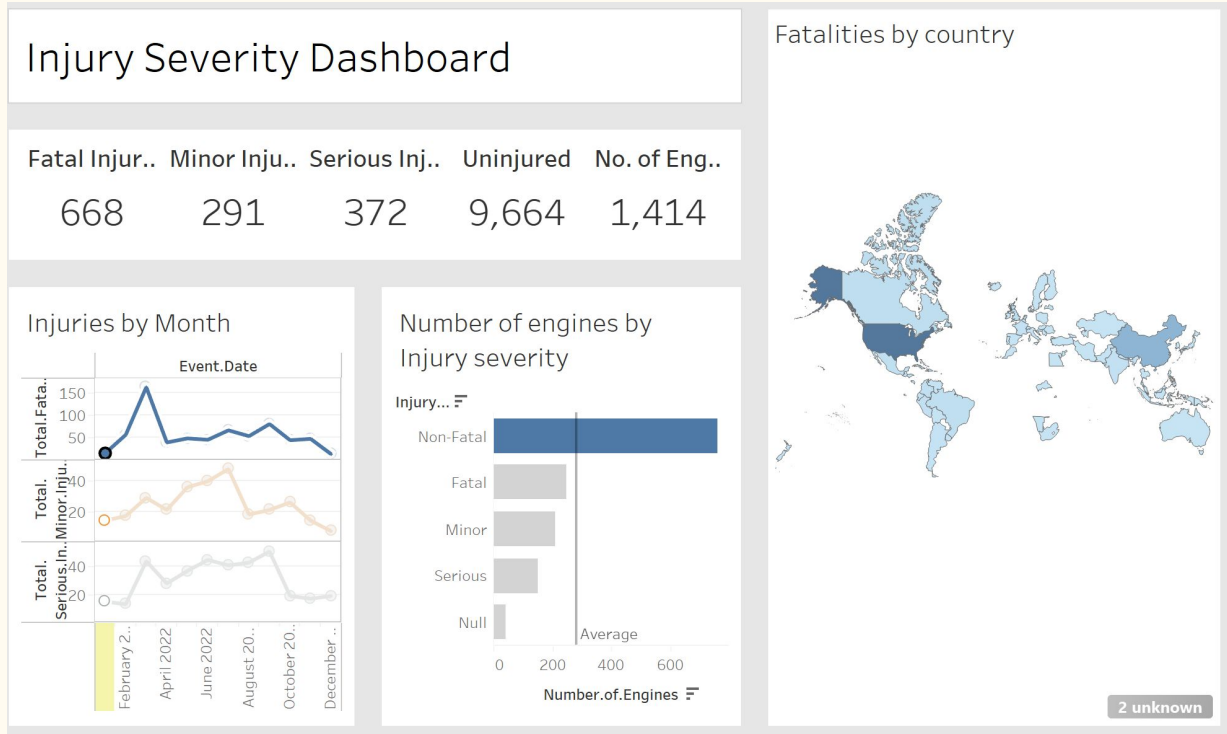
With the use of matplotlib...

Bar graph visualization...   **Scatter plot visualization...**

Hypothesis support

With the use of tableau...

Conclusion

With the solutions we have at hand it is going to:

1. Help us narrow down to the specific model and number of engines we require for our aircraft
2. We are going to establish a make and model that bears very low risks because of the findings that we have it is going to give us an upper hand in deciding what specific aircraft in terms of model and number of engines to purchase

The solutions offered are beneficial in terms using a minimalistic way to identify what is required hence reducing the amount of time required to make a decision. In addition to that with the proper visualizations we are able to get options in terms of choosing the aircraft we require.

Conclusion Cont'

What mostly gives confidence in this solution that it gives us options from the aircrafts that are not prone to risk. This way we have a variety to choose from, but specifically from which models and the number of engines that we need. This offers the best solutions for both the company and the stakeholders to choose from.

THANK YOU