



MORINGA
SCHOOL

AIRCRAFT PROJECT

DATA SCIENCE : DATA
CLEANING AND
VISUALIZATION

Presented by: Immaculate Kithei Kyalo



OVERVIEW

This research project aims to provide insightful data-driven approaches to support the air transport industry in expanding to new industries and diversifying its portfolio. The project will help the company in determining the most efficient aircraft for either private or commercial enterprises. This project will use data cleaning techniques, imputation, analysis, and visualization to generate insights for a business stakeholder. I analyzed the Aviation Dataset from the National Transportation Safety Board, which includes aviation accident data from 1962 to 2023 about civil aviation accidents and selected incidents in the United States and international waters. The analysis suggests that in the Aviation industry, there has been a decreasing trend in the number of accidents as more aircraft are designed. This indicates that the Aviation industry is an effective focus and it is worth investing as accidents are expected to reduce in the future thus the sustainability of the industry.



BUSINESS PROBLEM

The business's main issues include

1. Uncertainty on the future of the Aviation industry and whether it is a good venture

2. Identify the Low-risk Aircraft for business to purchase

3. Understanding the risk factors to consider in the aviation industry

The following questions will be used for the analysis

1. Is it safe to invest in Aviation Industry?

2. What sector should the company invest in between commercial and personalized flights?

3. What make should the company invest in compared to the number of accidents?

a. Which type of aircraft Should the company invest in(amateur built)?

4. Which is the best weather condition for airlines to pass through?

5. ♦Which is the best engine-type to consider with reduced fatalities and accidents?



DATA

The research project used data from the National Transportation Safety Board that includes aviation accident data from 1962 to 2023 about civil aviation accidents and selected incidents in the United States and international waters. This dataset contains summaries of aviation accidents and incidents spanning multiple decades, providing detailed information on various aspects of civil aviation accidents.

01

Dataset: AviationData.csv Description: This is the main dataset that contains detailed records of aviation accidents and incidents. Each row represents a unique aviation event with attributes such as the aircraft type, accident date, location, severity, weather conditions, and details of injuries or fatalities.

Relationship to data analysis: This file provides the data necessary for analyzing various risk factors such as aircraft type, flight conditions, and injury outcomes.

02

Dataset 2: AviationAccidentNarratives.csv Description: This file contains textual summaries or narratives of aviation accidents and incidents. These narratives are written descriptions of the events, typically including information about what led to the accident, the conditions at the time, and any relevant findings from investigations.



DATA

Description of relevant columns

1. Event Id: A unique identifier for each aviation accident or incident
2. Event Date: The date on which the accident or incident occurred.
3. Location: The city or general area where the accident or incident took place.
4. Country: The country where the accident or incident occurred.
5. Injury Severity: Describes the severity of injuries sustained in the event.
6. Aircraft Damage: Indicates the extent of the damage sustained by the aircraft.
7. Make: The manufacturer or company that produced the aircraft.
8. Model: The specific model of the aircraft involved in the event.
9. Amateur Built: Indicates whether the aircraft was built by an amateur or a certified manufacturer.
10. Number of Engines: The number of engines on the aircraft.
11. Engine Type: The type of engine used by the aircraft.
12. Purpose of Flight: Describes the intended mission of the flight.
13. Total Fatal Injuries: The total number of fatal injuries resulting from the accident or incident.
14. Total Serious Injuries: The total number of serious but non-fatal injuries resulting from the accident.
15. Total Minor Injuries: The total number of minor injuries resulting from the event.
16. Total Uninjured: The total number of people on board the aircraft who were uninjured.

17. Weather Condition: The weather conditions at the time of the accident or incident.

DATA

A photograph of a man with dark hair and glasses, wearing a dark long-sleeved shirt and jeans, sitting in a white chair and looking at a tablet. He is positioned on the left side of the slide, partially obscured by a large blue triangle.

The following data questions will be used for the analysis

1. Is it safe to invest in Aviation Industry?
2. What sector should the company invest in between commercial and personalized flights?
3. What make should the company invest in compared to the number of accidents?
4. Which type of aircraft ahould the company invest in(amateur built)?
5. Which is the best weather condition for airlines to pass through?
6. Which is the best engine-type to consider with reduced fatalities and accidents?

Data Preparation

Missing Data Handling: Imputation for categorical and numerical fields, with a focus on critical fields like Injury Severity and Aircraft Damage.

Transformation: Date parsing, encoding of categorical data, creation of new composite risk features.

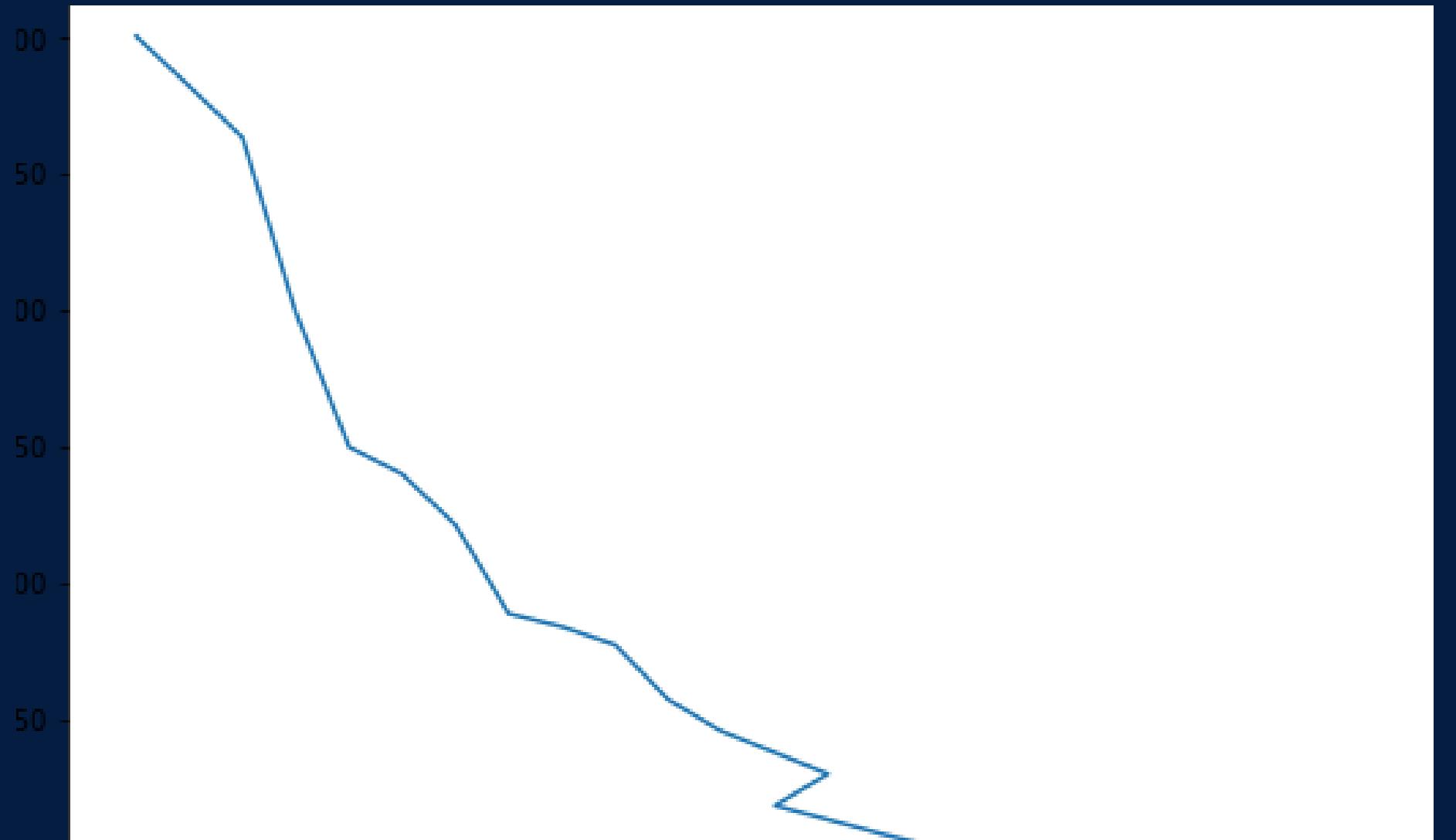
Data Validation: Consistency checks and removal of erroneous data.

Aggregation: Summarizing data by aircraft model, time, and location.

Visualization Preparation: Aggregating data and transforming features for effective exploratory data analysis and visualization.

DATA ANALYSIS: RESULTS

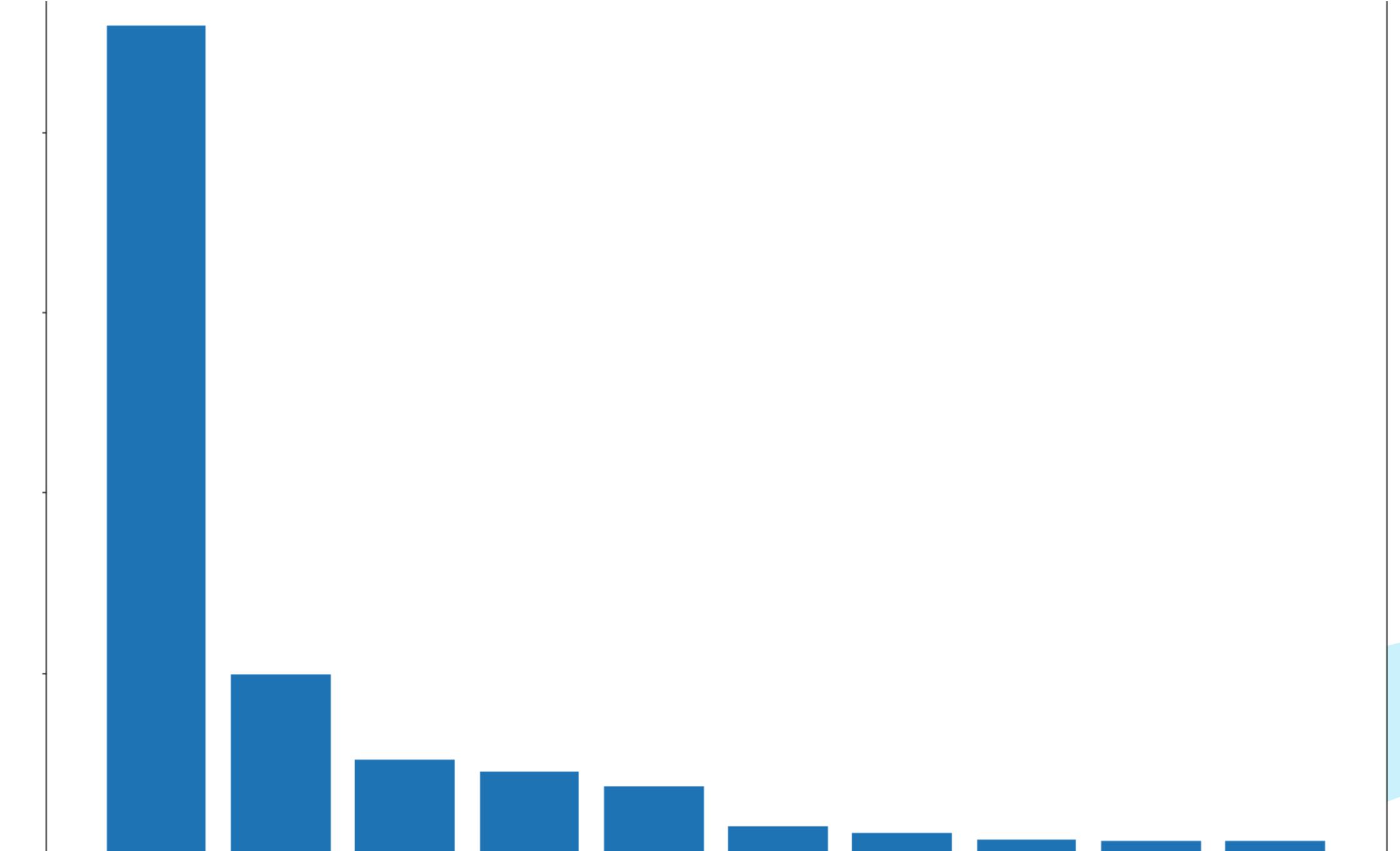
Question 1: Is it safe to invest in Aviation Industry?



Recommendation: In the Aviation industry, there has been a decreasing trend in the number of accidents as more aircrafts are designed. This indicates that the Aviation industry is an effective focus and it is worth investing as accidents are expected to reduce in the future thus the sustainability of the industry



What sector should the company invest in between commercial and personalized flights?

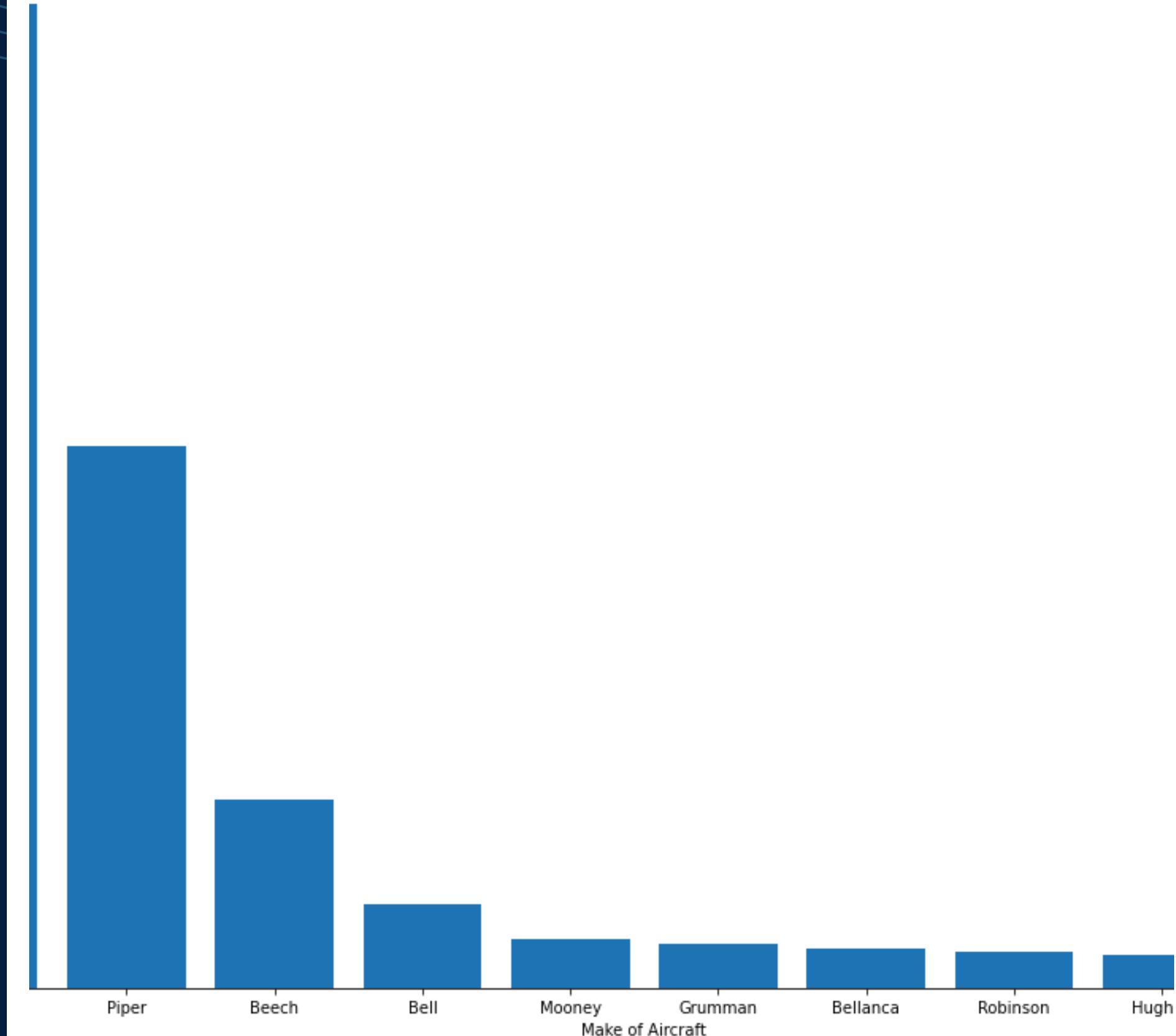


The aircraft should invest in public Aircraft because it has recorded the least number of accidents and incidents thus has a low accident risk compared to the other purposes.

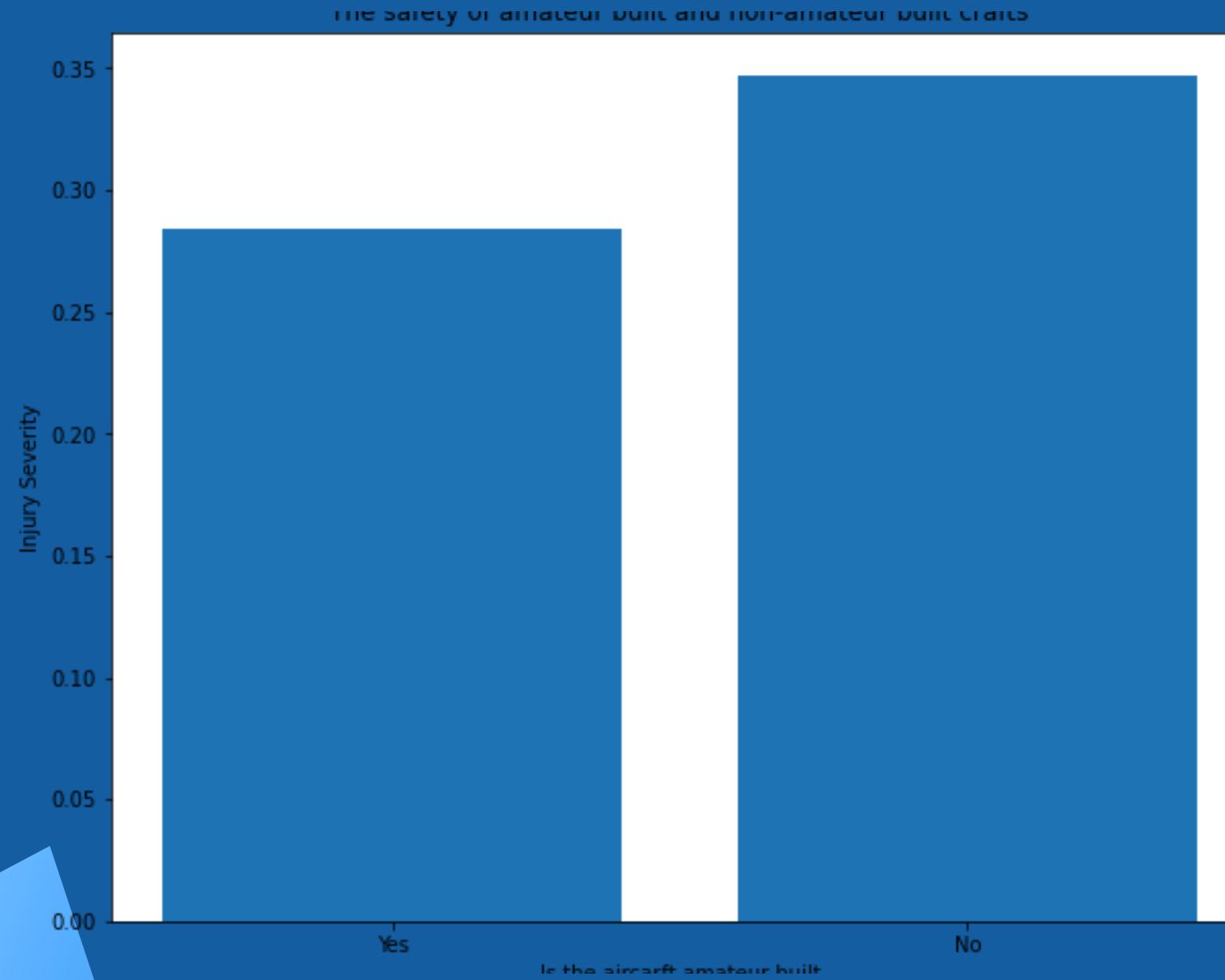
What make should the company invest in compared to the number of accidents?

While the visualization offers the leading companies in the number of accidents, the information might be limited to time constraints. The above visualization limits the company from investing in the above makes since they have a trend of causing accidents. However, the recommendation is not reliable with the current industry. The company should invest in the Boeing Make as they convey reduced accident cases thus improved functionality

Top 10 Makes and number of accidents by each

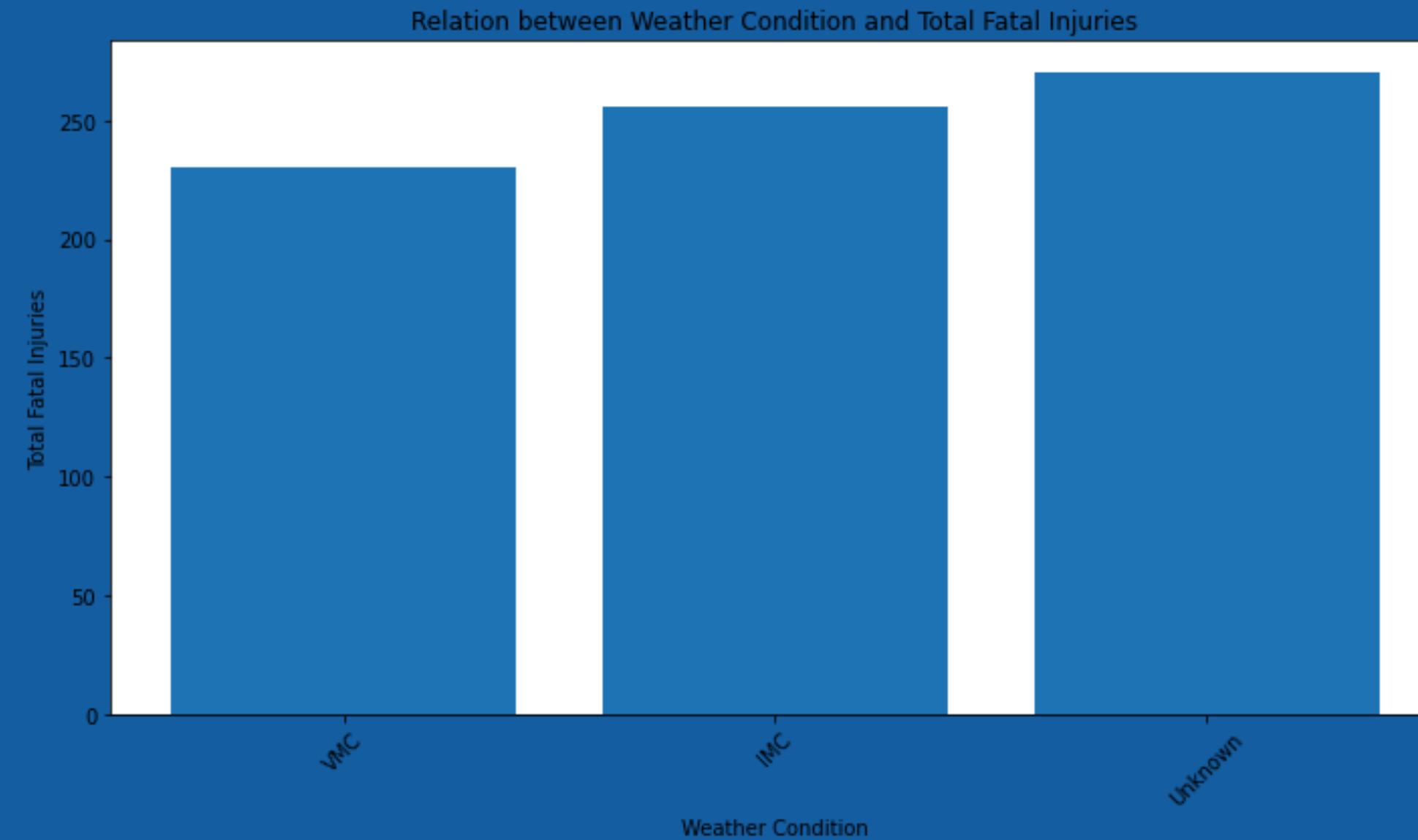


WHICH TYPE OF AIRCRAFT WHOULD THE COMPANY INVEST IN(AMATUER BUILT)?



Recommendation: According to the analysis, aircrafts which are not amateur built seem to have a high injury severity compared to those with amateur built. So the company should focus on airlines which are amateur built for reduced accidents.

WHICH IS THE BEST WEATHER CONDITION FOR AIRLINES TO PASS THROUGH?



Recommendation: The VMC indicates the best weather condition with most uninjured cases therefore sustainable for business operation. Pilots are advised to fly with the VMC condition.

Recommendations

1. In the Aviation industry, there has been a decreasing trend in the number of accidents as more aircrafts are designed. This indicates that the Aviation industry is an effective focus and it is worth investing as accidents are expected to reduce in the future thus the sustainability of the industry
2. The aircraft should invest in public Aircraft because it has recorded the least number of accidents and incidents thus has a low accident risk compared to the other purposes.
3. According to the analysis, aircrafts which are not amateur-built seem to have a high injury severity compared to those with amateur built. So the company should focus on airlines which are amateur built for reduced accidents.
4. While the visualization offers the leading companies in the number of accidents, the information might be limited to time constraints. The above visualization limits the company from investing in the above makes since they have a trend of causing accidents. However, the recommendation is not reliable with the current industry. The company should invest in the Boeing Make as they convey reduced accident cases thus improved functionality
5. The VMC indicates the best weather condition with most uninjured cases therefore sustainable for business operation. Pilots are advised to fly with the VMC condition.

NEXT STEPS

MACHINE
LEARNING FOR
EFFECTIVE
AIRCRAFT
DECISION MAKING
PROCESS

Thank You Note



Immaculate Kithei Kyalo

Data Scientist

Moringa School

Connect with us.



+254715614795



[https://www.linkedin.com/in/
mary-kyalo-bb2a67188/](https://www.linkedin.com/in/mary-kyalo-bb2a67188/)

