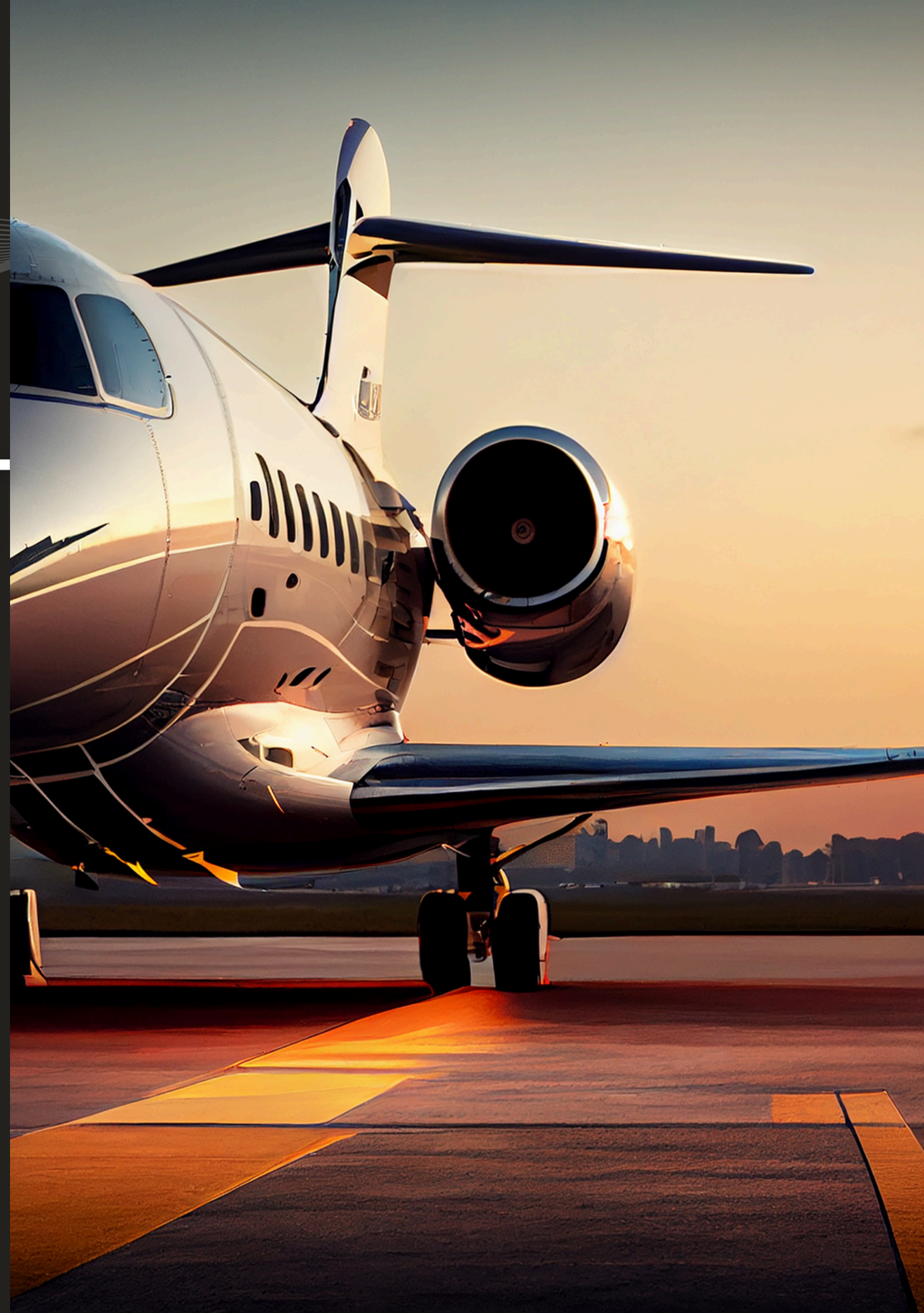
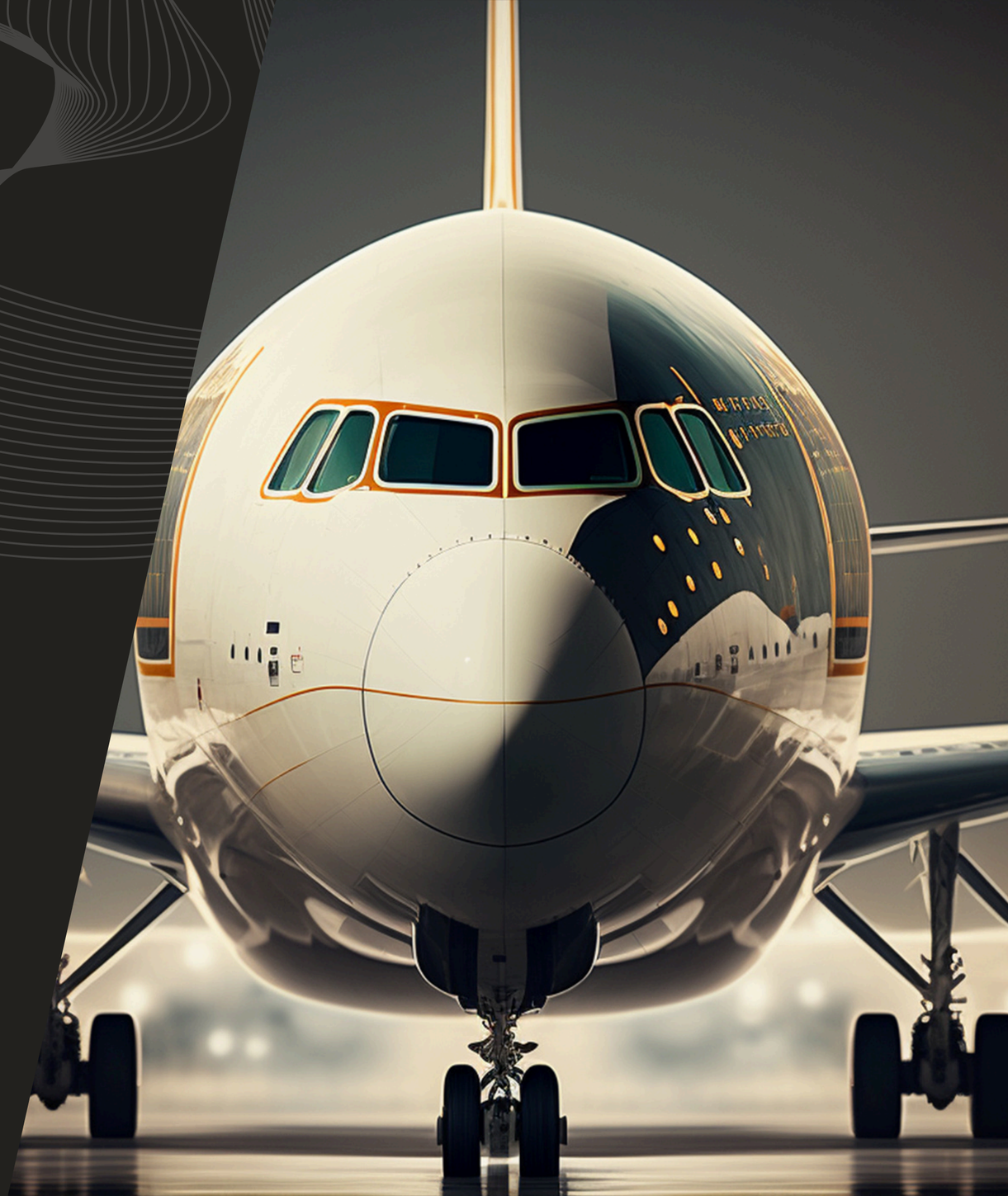
AVIATION AIRCRAFT PROJECT

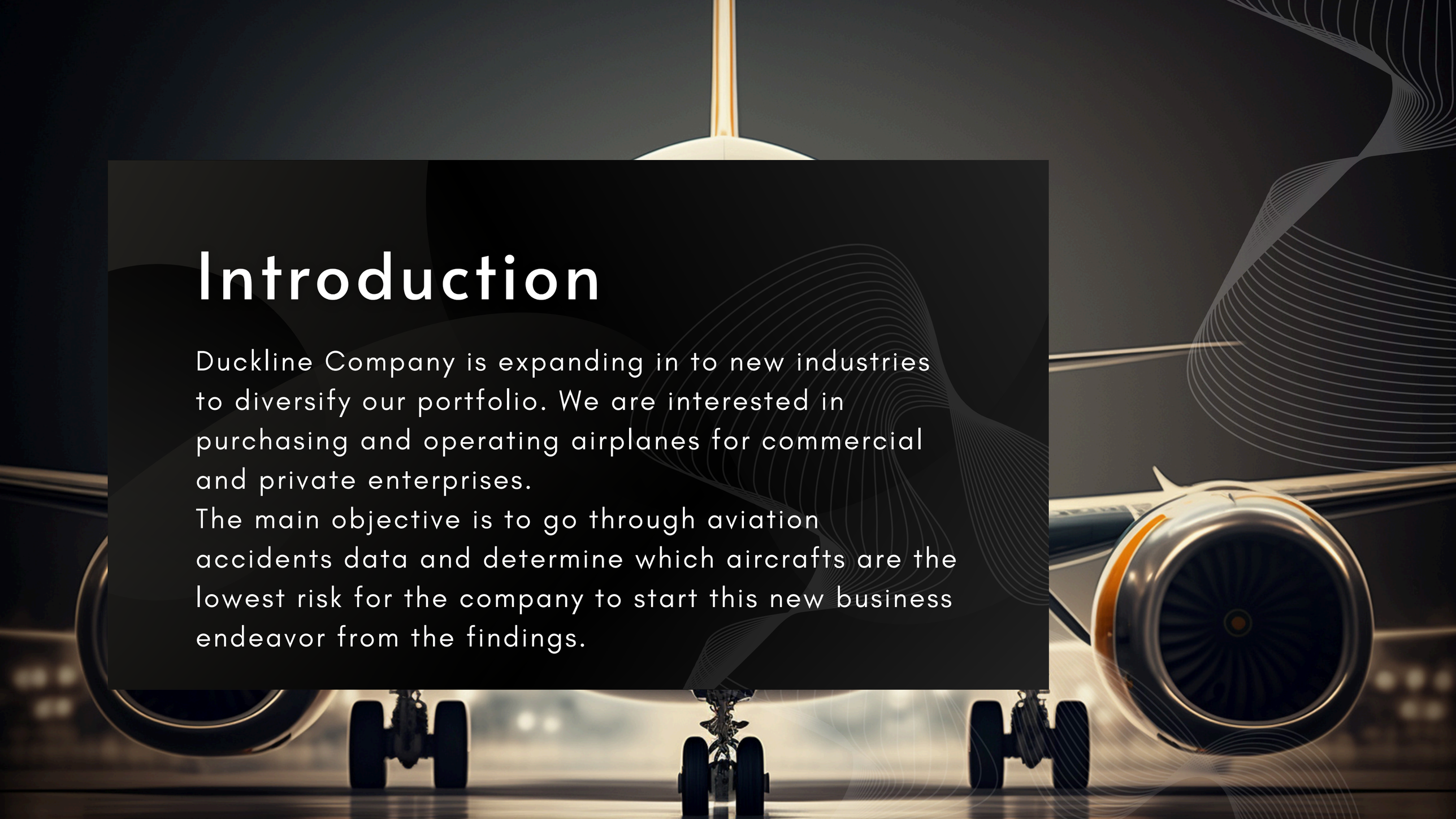
MAUREEN IMANENE



Overview

- About Us 01
- Overview 02
- Introduction 03
- Objectives 04
- About the Data 05
- Data Limitations 06
- Data cleaning 07
- Private aircraft Recommendation 08
- Commercial aircraft Recommendation 09



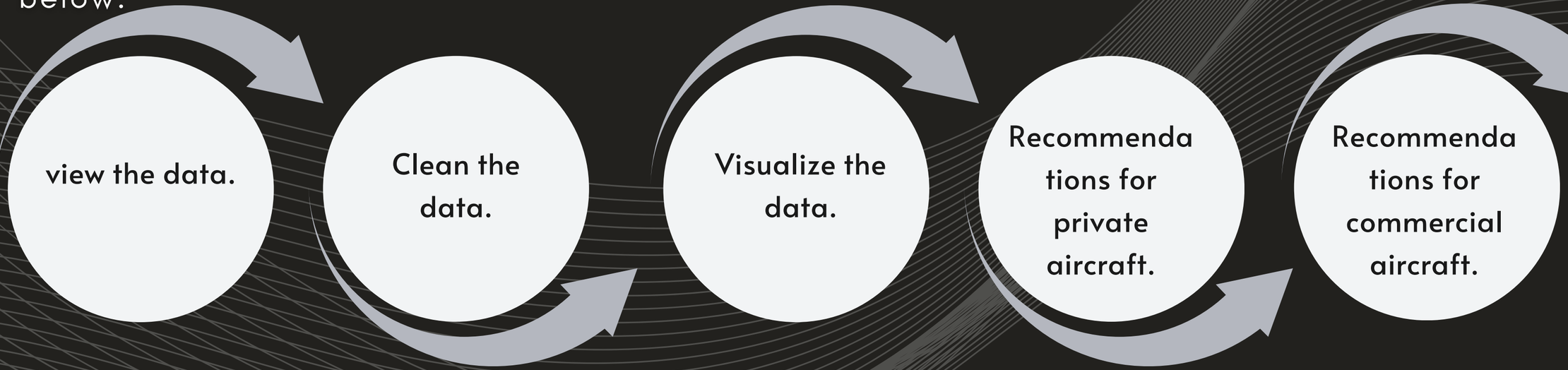


Introduction

Duckline Company is expanding in to new industries to diversify our portfolio. We are interested in purchasing and operating airplanes for commercial and private enterprises.
The main objective is to go through aviation accidents data and determine which aircrafts are the lowest risk for the company to start this new business endeavor from the findings.

Objectives

The objective of this project was to give recommendations for aircrafts that pose the lowest risk for the company in regards to accidents.
In-order to realise this objective, i took the path illustrated below.



About the Data

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 was used for this project.
This data set had 88889 rows and 31 columns.

Data Limitations

- VERY LARGE DATASET - The dataset was a very large one (This is supposed to be a good problem) however in our case we had:
- NON-RELEVANT COLUMNS - There were very many columns that weren't relevant to us in answering the question at hand . for example; Latitude, Longitude, Airport_Code, Amateur_Built, Schedule, Air_Carrier, Aircraft_Category, Far_Description, Report_Status etc.
- MANY NULL VALUES -There were very many null values present , in some columns more than half of the data was made up of null values.
- NAMING INCONSISTENCIES -Inconsistencies with naming/labeling of the air-craft makes & models , where one exact make is written in several different ways using commas or hyphens. Thus they are interpreted as different/distinct makes .

Data Cleaning

To clean the data I :

- **REMOVED NON-RELEVANT COLUMNS**:

I removed columns with data that was not assisting me to answer the question at hand.

- **MISSING VALUES:**

I checked for missing values and dealt with them on a case by case basis by removing some and replacing others with None.

- S**TANDARDIZATION:**

I standardized irregular make & models names so that each make was represented in a uniform manner for easier and more accurate interpretation.

Private Aircraft Recommendation

Based on the findings from the Aviation accident dataset I recommend the CESSNA aircraft make For the PRIVATE DIVISION.

This is because;

- Cessna had the LOWEST FATALITY RATE of all the makes.
- Cessna had one of the highest mean of UNINJURED PASSENGERS.
- Cessna has a large collection of different aircraft models to choosefrom.

Commercial Aircraft Recommendation

Based on the findings from the Aviation accident dataset I recommend the AIRBUS aircraft make for the COMMERCIAL DIVISION.

This is because;

- Airbus had the Highest mean of UNIJURED PASSENGERS of all the makes.
- Airbus had the Lowest FATALITY RATE mean.
- Airbus has a very wide selection of 89 different modelsto choose from;a300,a310etc.

Our team

This project was done by a team of one.

Maureen Nkatha Imanene
Data Science Student at The Moringa School.

Maureen Imanene
Data Scientist

Thank You

Contact Us

maureen.imanene@student.moringaschool.com

www.duckline.com

Gerente General