Aviation Accident Analysis: Understanding Aviation Accident Severity Through Data to Inform Preventative Actions.

Name: Teresia Ndung'u

LinkedIn: https://www.linkedin.com/in/teresia-ndung-u-8b9308a4/

Date: 30th June 2025

Teresia Ndungu
& LinkedIn

Date

Project Overview

This project analyzes aviation accident data of 1962 to 2023 with an aim of uncovering key factors influencing injury severity during aviation accidents and suggest safety improvements.

This project will explore exploring variables such as weather conditions, aircraft characteristics, and flight purpose, with the aim of deriving actionable insights that will support aviation safety improvements and help prevent future incidents.

Business Understanding

Why This Project Matters

On January 23, 2024, a Cessna 172N crashed near Weston, Florida during a personal cross-country flight operated by a flight school, tragically killing both commercial pilots on board. Aviation safety is critical, as aircraft accidents, while rare, often result in severe outcomes including death and financial losses.

Identifying factors that influence accident severity such as the weather conditions, aircraft type, aircraft engines and flight purpose can greatly help prevent future incidents. Insights from this analysis support improvements in regulatory policies, aircraft standards and pilot’s training programs.

Project Statement and Goal.

Problem Statement:

Aviation accidents have varying outcomes. Some outcomes are fatal, severe while others are non fatal. Understanding what factors contribute to more severe cases can help reduce fatalities and improve flight safety.”

Goal:

Generate insights to support policy decisions and preventative actions in aviation safety.

Data Understanding

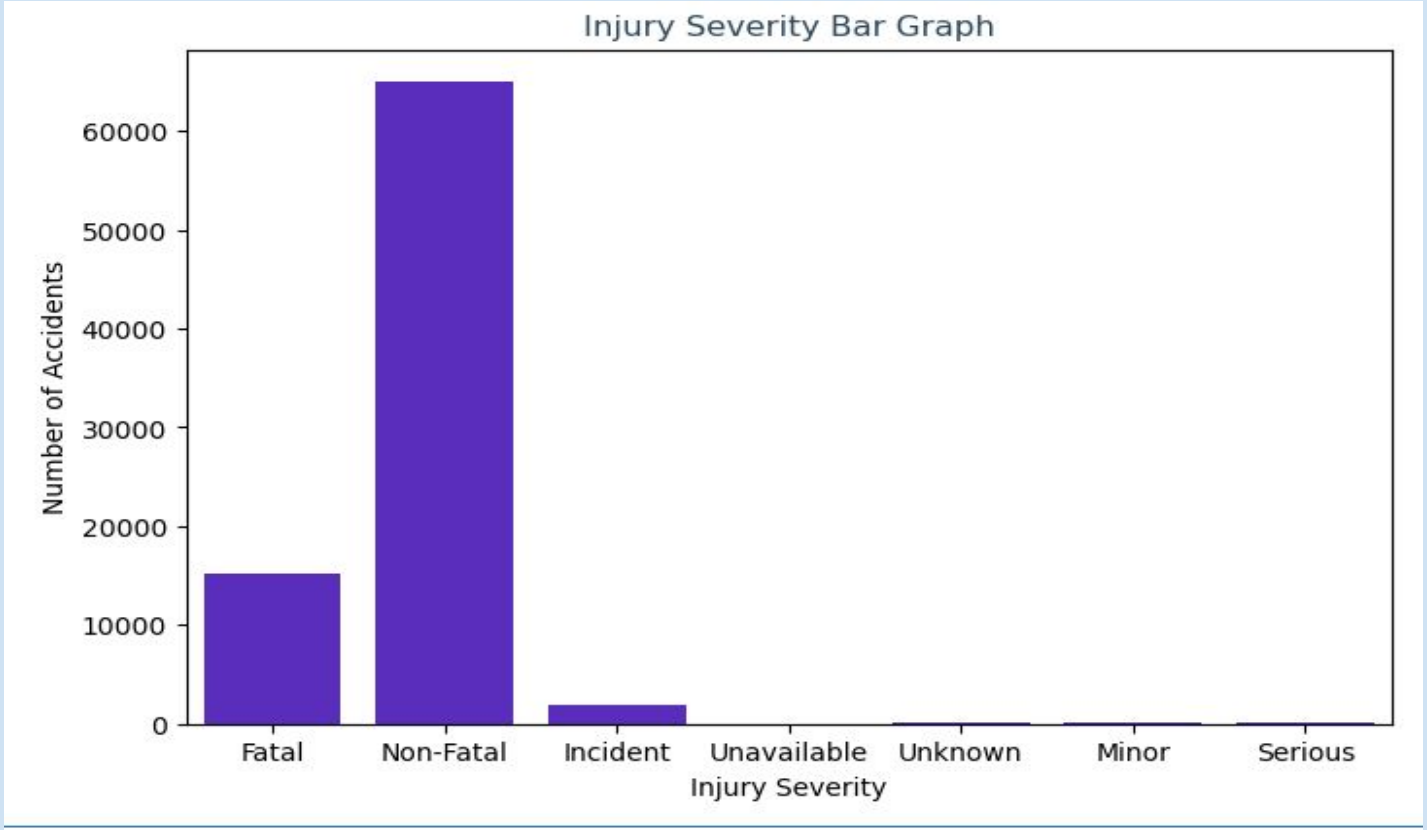
The dataset is from the National Transportation Safety Board that includes aviation accident data from 1962 to 2023 about civil aviation accidents and selected incidents in the United States and international waters. The dataset has 88889 rows and 31 columns.

To improve accuracy and consistency, the dataset underwent standardization steps such as capitalization of categorical values, merging similar categories and creation of cleaned columns such as “injury_severity_cleaned” for clearer analysis

Key variables include injury severity, weather, conditions, engine type, aircraft make/model, purpose of flight.

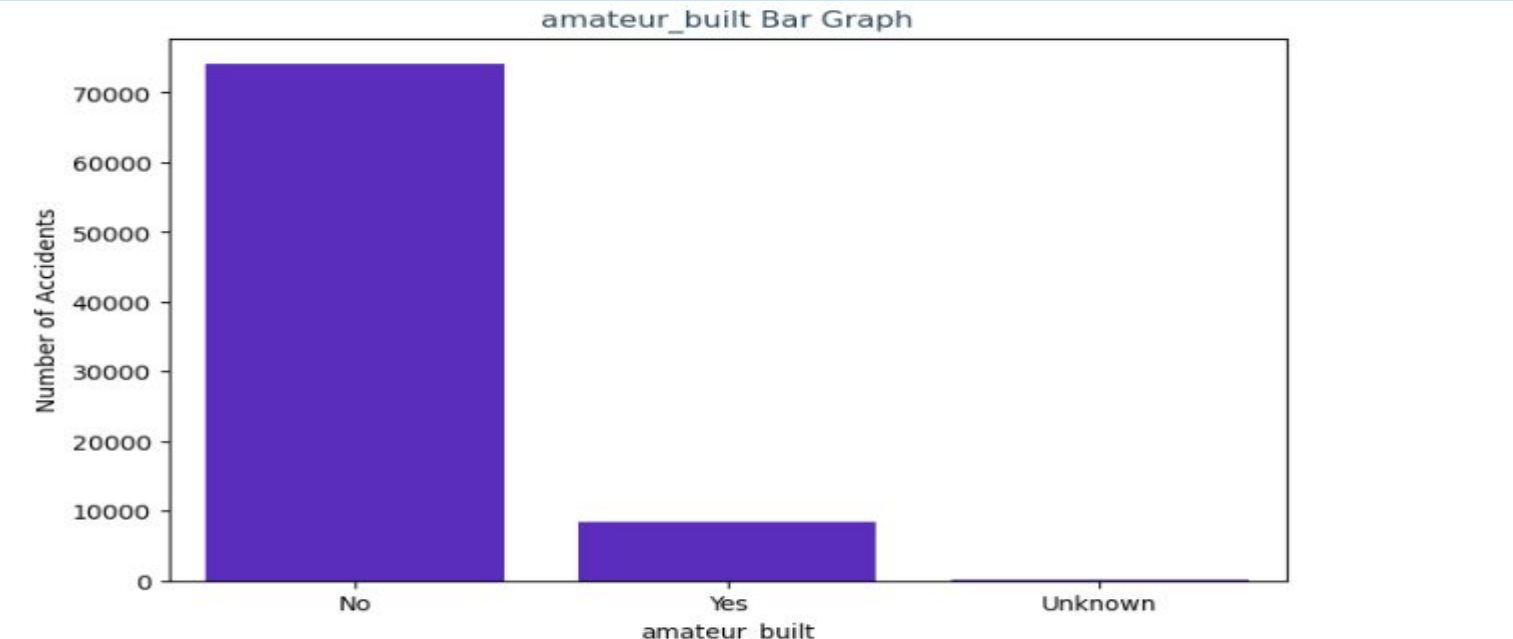
Univariate Analysis;Injury Severity Bar Graph

The output shows that non-fatal accidents were the majority (76%), followed by fatal accidents which had some accidents.



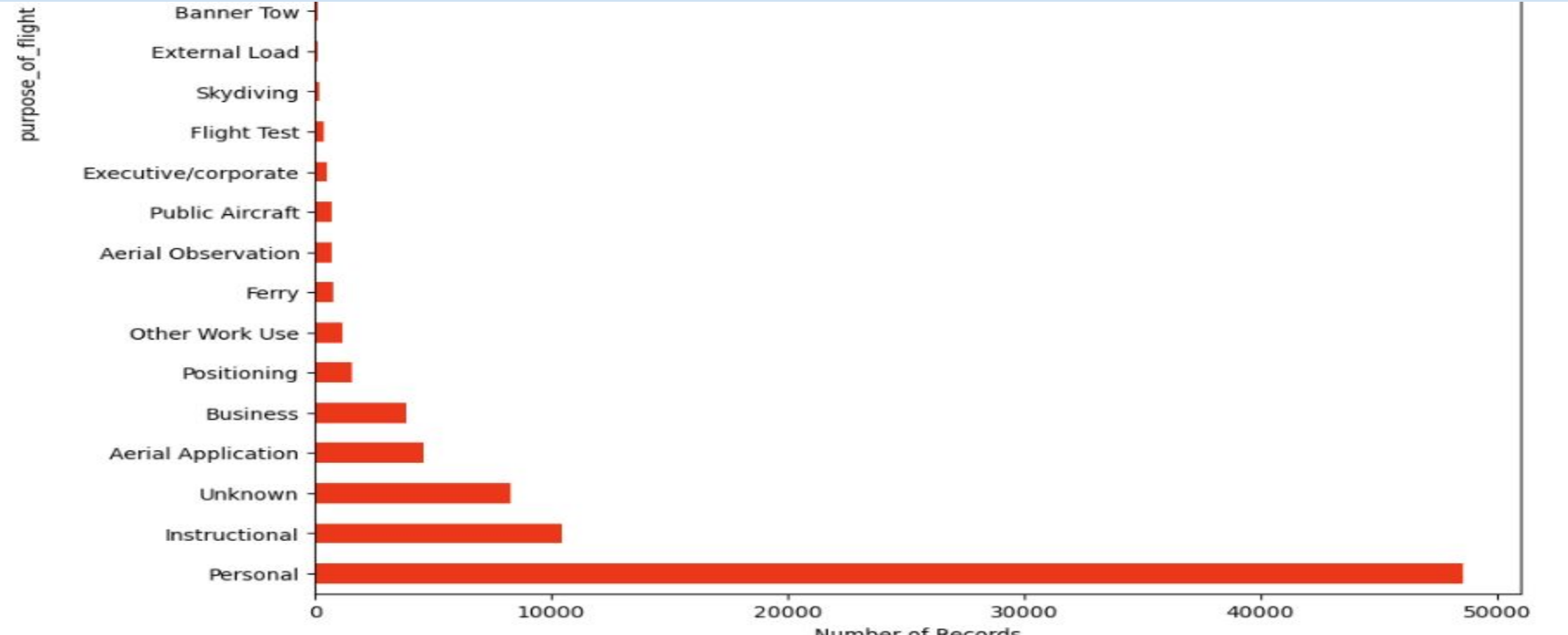
Univariate Analysis- Amateur Built Bar Graph.

The plot shows that the aircrafts involved in an accident and were built by an amateur were 74139, while those that were built by a non-amateur were 8319 while 24 aircrafts were from unknown amateur.



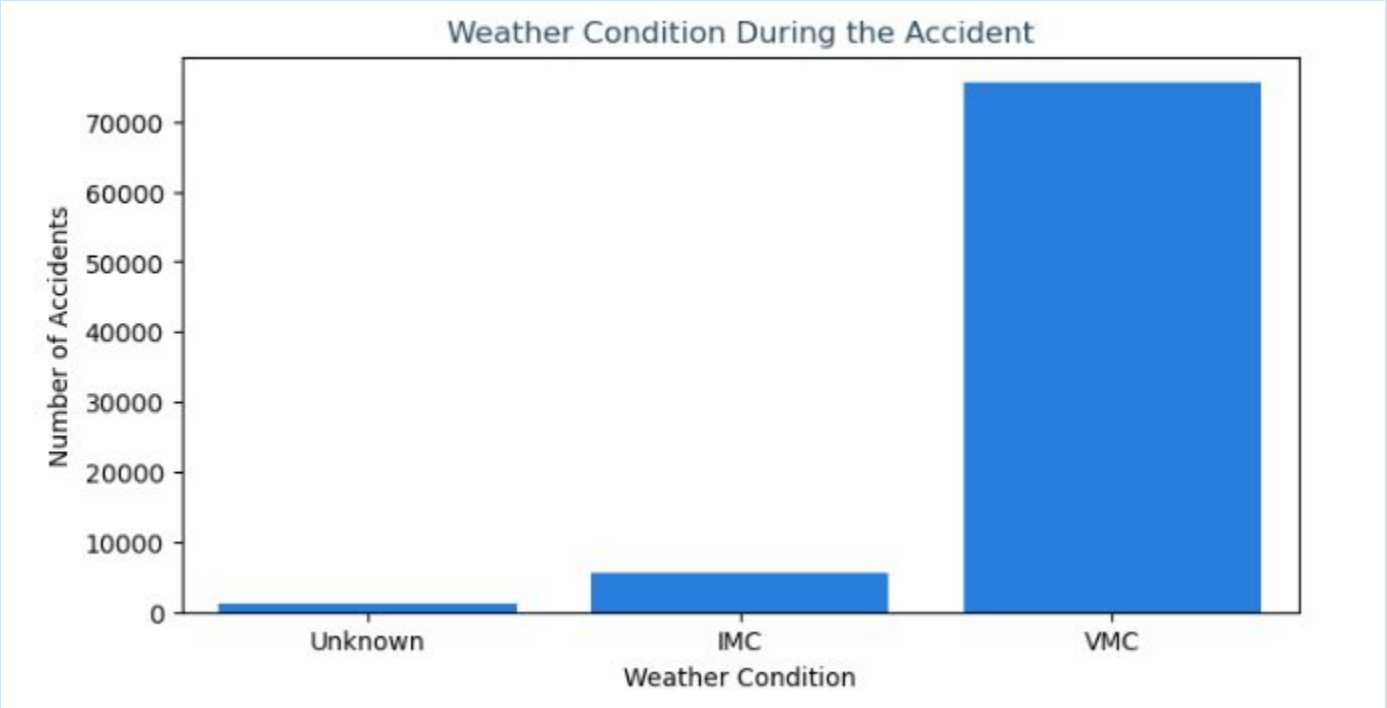
Univariate Analysis;Purpose of the Flight.

The majority of aviation accidents occurred during personal flights (48,582), highlighting higher risks in private or recreational flying. Instructional (10,438) and unknown purpose (8,279) flights followed, indicating training-related risks. Operational uses like aerial application and business flights also contributed significantly. Less frequent activities such as skydiving and firefighting showed lower counts but may involve unique hazards. Public aircraft flights were rare but present.



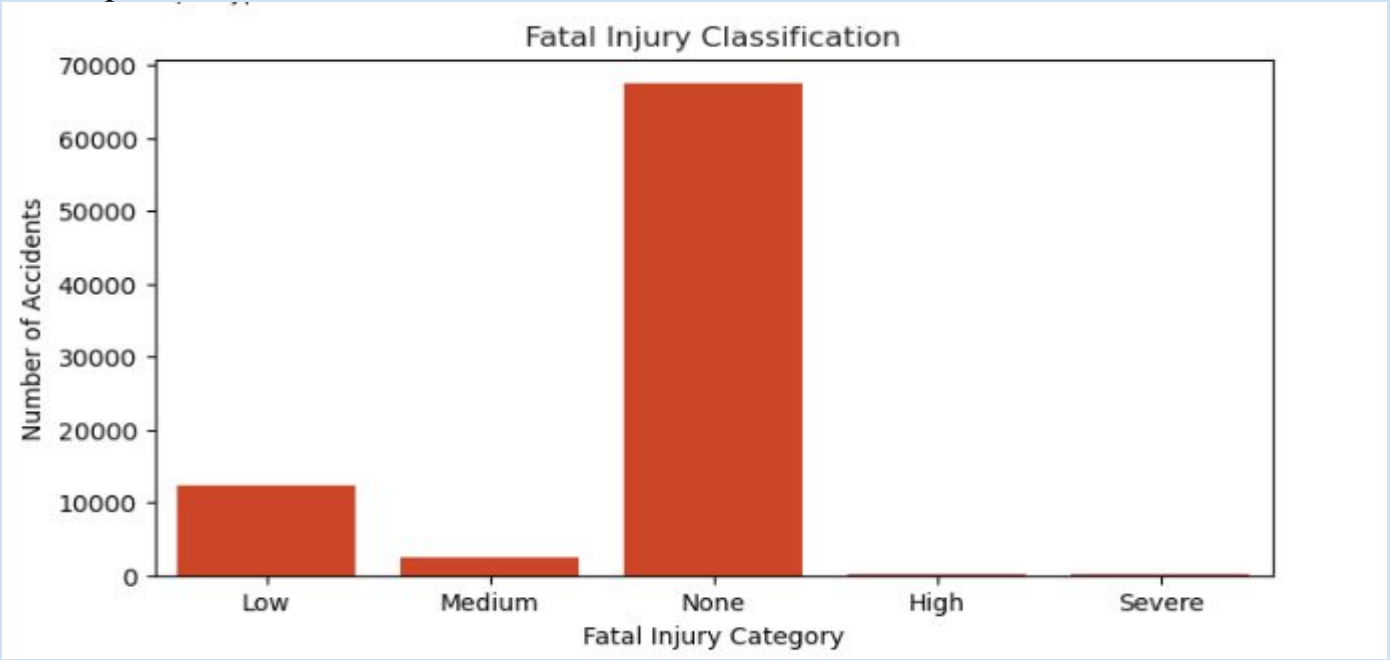
Univariate Analysis;Weather Condition During the Accident

Visual Meteorological Conditions (VMC) refer to clear weather, which is exceptionally safe for flying.Instrument Meteorological Conditions (IMC) refer to poor weather conditions, which mostly entail fog, clouds, and rain. This kind of weather is not suitable for flying. The Unknown category refers to the unspecified or missing weather data due to incomplete reporting.The bar chart shows that most of the accidents occured when the weather was clear and safe for flying.



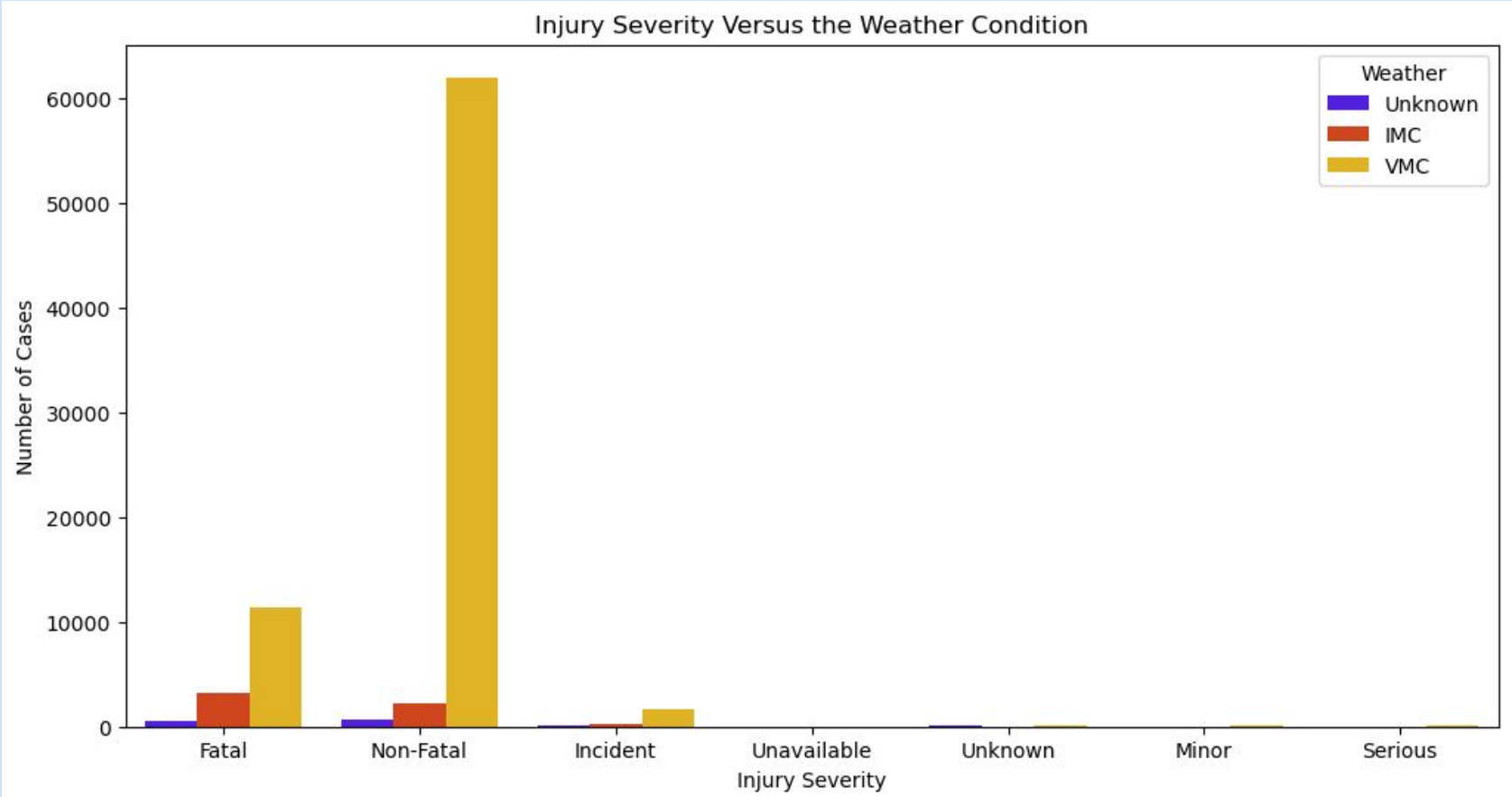
Univariate Analysis: Total Fatal Injuries Classification

The above output shows that 76% of the accidents had no fatalities. Based on these results, we can conclude that most of the aviation accidents are non-fatal. 14% of these accidents were classified as low, while 2.7% were classified as medium. High and severe accidents were rare and they constituted 0.4%.Accidents do occur, and most of them do not lead to loss of life. However, there is a need to investigate the small fraction of cases with high fatal consequences.



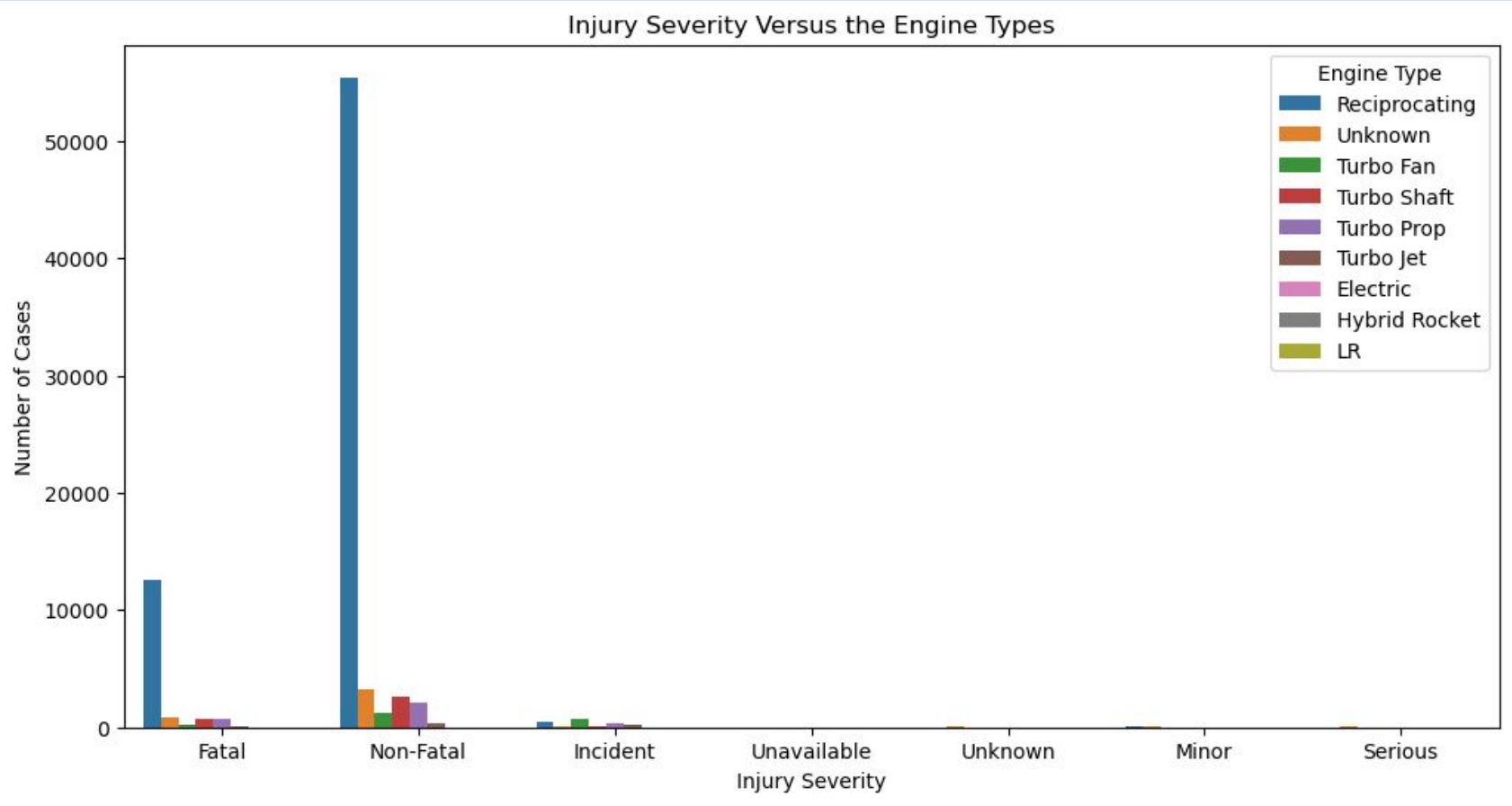
Bivariate Analysis: Injury Severity Versus Weather Condition

The above plot of injury severity versus the weather conditions has revealed that the Visual Meteorological Conditions (VMC), which refers to clear weather and most preferred for flying, has the highest number of non-fatal accidents.



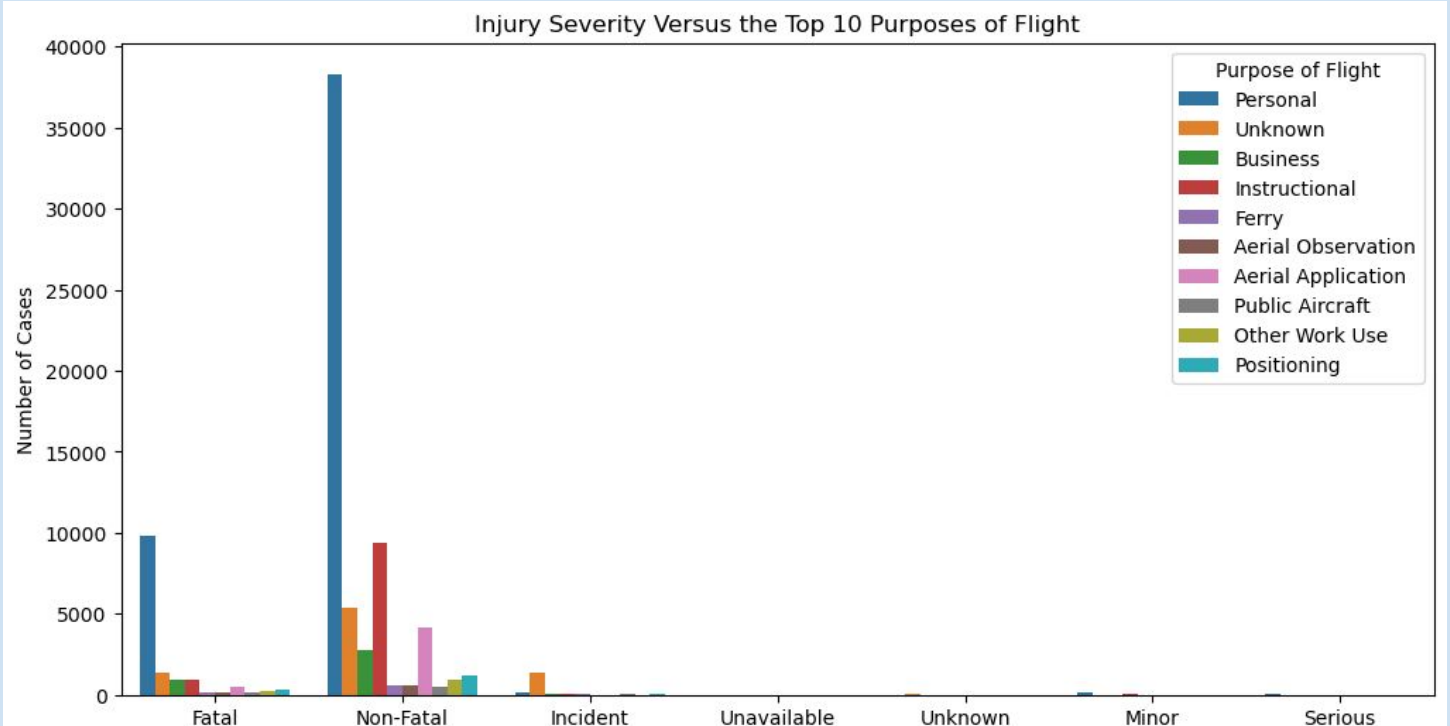
Bivariate Analysis: Injury Severity Versus Engine Types

The analysis shows that reciprocating engines were the most common in both fatal and non-fatal accidents. They were followed by turbojet, turboprop, and turbo fan engines, in that order of frequency.



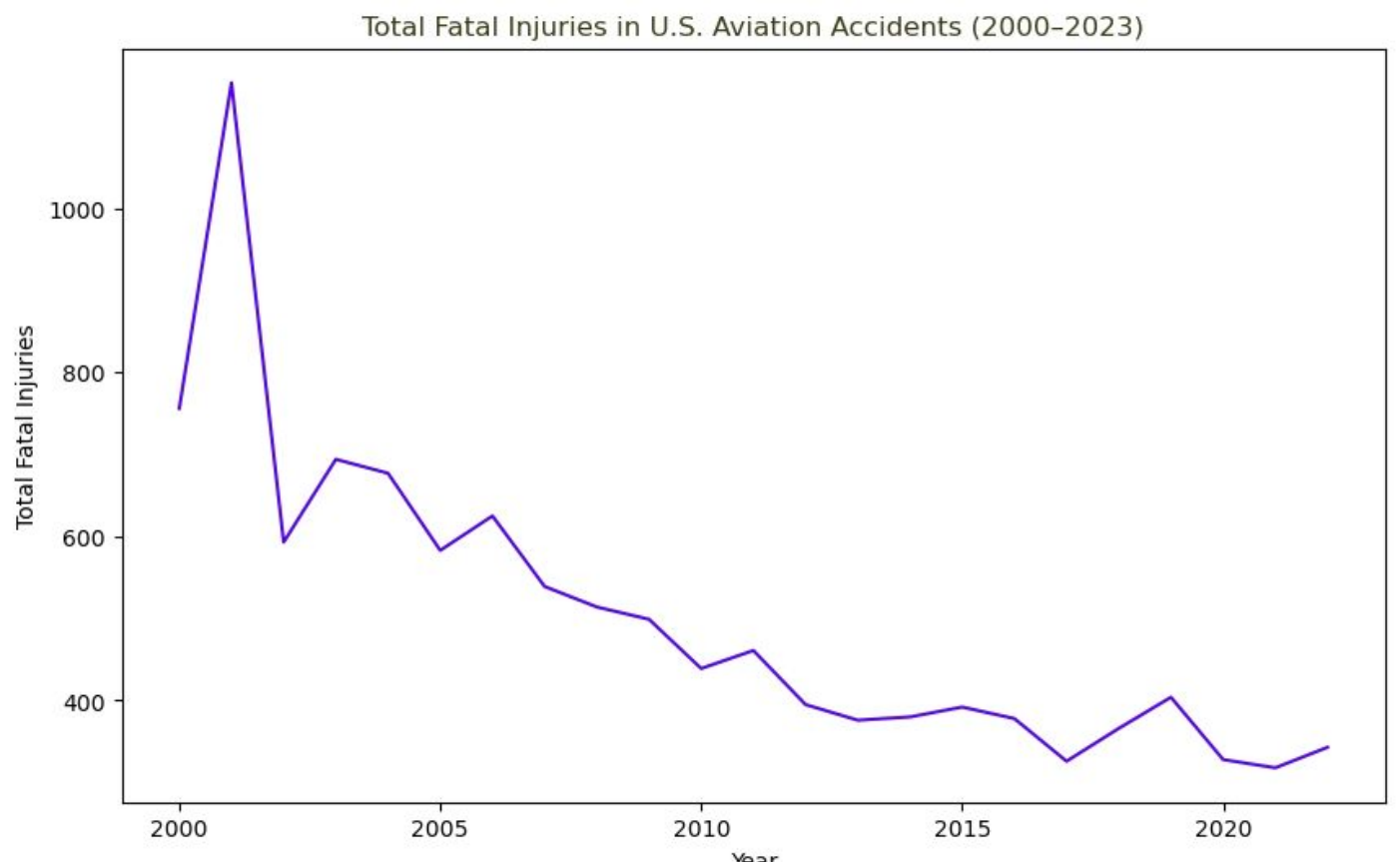
Bivariate Analysis: Injury Severity Versus Top 10 Purposes of Flight.

The bar plot shows that personal flights recorded the highest number of accidents for the fatal and non-fatal injury severity. This shows that personal use of aviation poses a greater risk. Notably, instructional flights had a high number of non-fatal injuries severity. Instructional flights are learning environments hence, the injury severity is expected to be low.Further, aerial aircraft had a notably high number of non-fatal accidents and very low fatal accidents. Aerial aircraft pilots are highly trained and experienced. The pilots are able to overcome dangerous situations safely. Additionally, the aerial aircraft designs are well-rugged, which significantly reduces the injury severity during crashes.



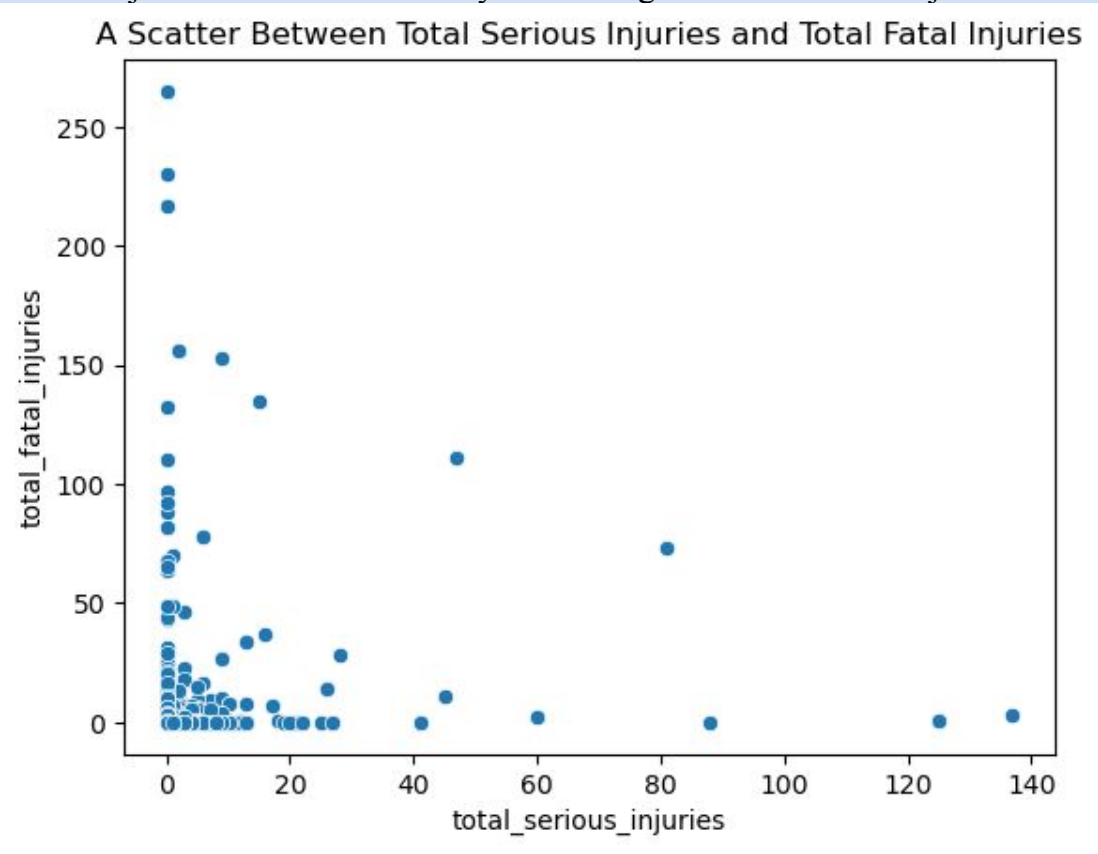
Bivariate Analysis: Total Fatal Injuries in U.S Aviation Accidents(2000-2023)

The line graph of the total fatal injuries of aviation accidents in the US, between 2000 and 2023 have declined steadily over time. The line graph shows that there was a notable high fatalies in 2000. There is a notable steady decline in thee fatal injuries after year 2001 which means that there are great improvements in aviation safety regulations, technology, and training.



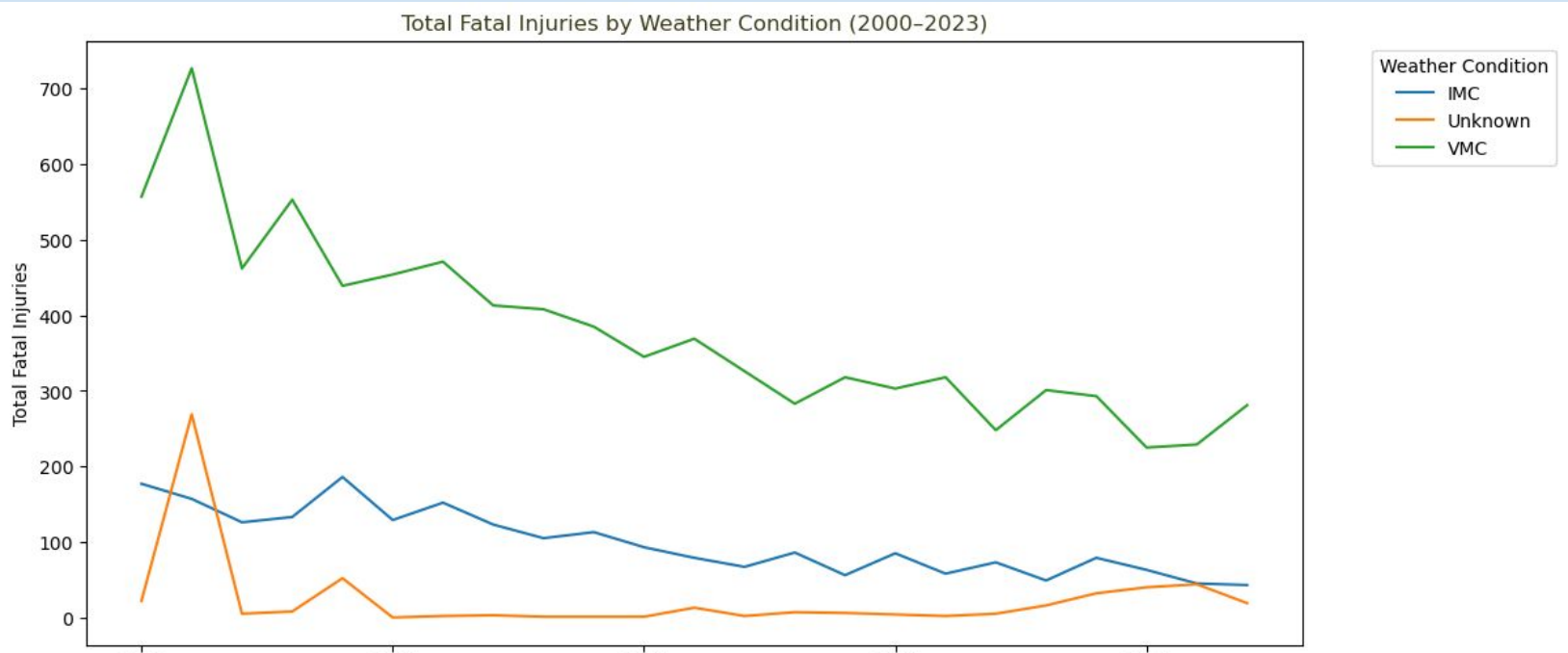
Bivariate Analysis: Total Fatal Injuries in U.S Aviation Accidents(2000-2023)

The above scatterplot shows that there is no clear linear relationship between serious injuries and fatal injuries since the data points are quite spread. Therefore, there is no strong correlation between the number of serious injuries and fatal injuries. A high number of serious injuries does not necessarily mean a high number of fatal injuries.

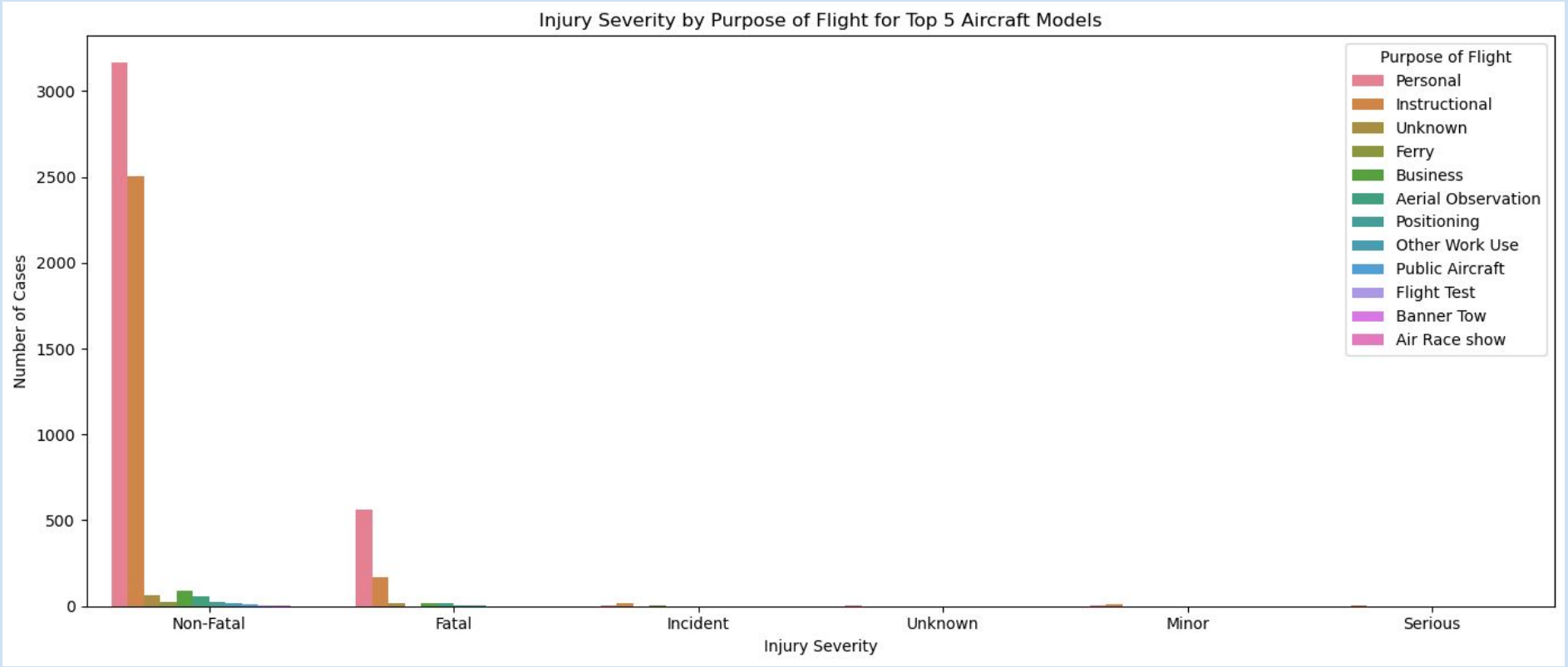


Multivariate Analysis: Total Fatal Injuries by Weather Conditions by Year(2000-2023)

The above line graph illustrates the total number of fatal injuries by weather condition between 2000 and 2023.Even though Visual Meteorological Conditions (VMC) is considered the most suitable weather for flying, it had the highest number of fatal injuries in 2000 and 2023 compared to Instrument Meteorological Conditions (IMC), which had a lower total number of fatal injuries. IMC weather conditions are very unfavorable for pilots because of low visibility.Interestingly, I expected the IMC to have the highest number of fatal injuries per year. However, according to our data, most of the accidents occur during the VMC when there is a high volume of flights, which increases accidents exposure. Additionally, IMC flights are operated by experienced pilots who observe strict flying protocols.



Multivariate Analysis: Injury Severity by Purpose of Flight for Top 5 Aircraft Models.



Conclusion

The analysis revealed that while non-fatal accidents are the majority, factors such as weather conditions, aircraft characteristics (like engine type and amateur-built status), and flight purpose significantly influence injury outcomes. Most accidents occur in Visual Meteorological Conditions (VMC), despite VMC being considered safe, indicating other underlying factors. Aircraft with reciprocating engines are most commonly involved in both fatal and non-fatal accidents, suggesting higher risks in general aviation. Personal flights recorded the highest number of fatal and non-fatal accidents, posing a greater risk than other flight purposes. Overall, fatal injuries in U.S. aviation accidents have seen a steady decline between 2000 and 2023, suggesting improvements in safety regulations, technology, and training.The presentation concludes that generating these insights supports policy decisions and preventative actions in aviation safety.

Business Recommendations

1. Improve Safety Protocols in Favorable Weather because Accidents still occur under Visual Meteorological Conditions (VMC). It will be better if reinforcement training is done even in “good” weather.
2. Implement more frequent safety audits and pre-flight inspections for privately flown aircraft, which account for the majority of accidents.
3. Track engine types involved in severe accidents and tighten maintenance and certification standards accordingly.
4. Focus safety messaging on common accident scenarios in personal and instructional flights.
5. Establish stricter protocol adherence even when conditions appear safe to avoid complacency-related errors.

Next Steps

1. Incorporate data on pilot experience, flight hours, and certification levels (if available) to better understand human factors influencing accident severity.
2. Analyze maintenance history, air traffic control communication, or altitude profiles to identify deeper systemic causes.
3. Collaborate with flight schools, aircraft manufacturers, and private pilot associations to implement tailored safety improvements.
4. Allow stakeholders to filter by weather, engine type, location, etc. to support real-time risk monitoring and intervention planning errors.

Thank You

Teresia Ndung’u

LinkedIn:https://www.linkedin.com/in/teresia-ndung-u-8b9308a

GitHub link:https://github.com/MissTerryStar