**Project 1: Aircraft Crashes**

****

[**Aircraft Crashes**](https://images.app.goo.gl/wyoacizRxNgTqJU99)

### **Team Members**

| **Flight Personnel** | **Role** |
| --- | --- |
| Kimberly Bonilla | Cabin Crew |
| Diana Tarasovets | Cabin Crew |
| Maero Lutta | Cabin Crew |
| Edgardo Perez Santiago | First Day Pilot, no training |

**Background**

Aviation safety is paramount in the transportation industry, and understanding the historical trends and factors contributing to aviation crashes is essential for improving safety measures. This project delves into aviation crash data dating back to 1908, examining variables such as fatalities, locations, and aircraft operators. By leveraging datasets containing comprehensive crash details and city information, the project aims to uncover patterns and insights that can shed light on the causes and circumstances surrounding aviation accidents over time. Through rigorous analysis of this data, stakeholders in the aviation industry, regulatory bodies, and safety authorities can gain valuable insights to enhance safety protocols, mitigate risks, and prevent future accidents, ultimately contributing to the advancement of aviation safety worldwide.

**Motivation**

Aviation accidents have far-reaching implications, prompting the need for comprehensive analysis to enhance safety protocols and regulations. This project is driven by the imperative to understand the multifaceted factors influencing aviation accidents. By meticulously analyzing crash data spanning over a century, the project seeks to uncover patterns and insights crucial for stakeholders in the aviation industry, regulatory bodies, and safety authorities. Through this analysis, the project aims to provide actionable insights that can inform the development of proactive safety measures, regulatory reforms, and training programs. Ultimately, the project's findings have the potential to contribute significantly to the advancement of aviation safety, reducing the occurrence of accidents and ensuring the well-being of passengers, crew members, and individuals on the ground.

### **Questions to answer**

1. What are the aircraft types and operators associated with the highest number of fatalities

1. When and where did aviation crashes occur, and are there any discernible patterns or trends in terms of time and location?
2. Is there a correlation between the number of passengers on board an aircraft and the fatalities resulting from aviation crashes?

### **Tools/Modules that were used**

* Python
* Pandas
* Matplotlib
* NumPy
* SciPy

### **Datasets used**

The project utilizes two main datasets:

1. *Airplane Crashes and Fatalities Since 1908:*

This dataset contains detailed information about aviation crashes, including the date, time, location, operator, aircraft type, number of passengers aboard, fatalities, and a summary of the incident. It spans from 1908 to the present day, providing a comprehensive historical perspective on aviation accidents. The dataset was obtained from <https://www.kaggle.com/datasets/cgurkan/airplane-crash-data-since-1908>

1. *World Cities:*

This dataset provides information about cities worldwide, including their names, geographical coordinates (latitude and longitude), country, population, and administrative details. It serves as a reference for identifying the locations of aviation crashes and analyzing the geographical distribution of incidents. This dataset was obtained from the bootcamp instructor.

Both datasets are stored in the output\_data directory and are crucial for conducting the analysis and generating insights into aviation safety and accident patterns.

### **Task Breakdown**

Student 1: Kim/ Edgardo - Collect/Clean the data.

Student 2: Maero - Dashboards

Student 3: Diana - Visualizations

Student 4: Edgardo, Kim, Diana, Maero - Analysis & Presentation

**A Summary Highlight of the Analysis**

**Analysis and Conclusion**

**What are the aircraft types and operators associated with the highest number of fatalities in aviation crashes?**

The analysis revealed that certain aircraft types and operators have been consistently associated with the highest number of fatalities in aviation crashes. Among the aircraft types, the Douglas DC-3 emerged as the aircraft with the highest fatalities, followed by Antonov AN-26, Douglas DC-6B, Douglas C-47, and McDonnell Douglas DC-9-32. These aircraft types have been involved in numerous accidents resulting in substantial loss of life over the years.

In terms of operators, Aeroflot, the national airline of Russia, topped the list with the highest number of fatalities. Military organizations, particularly the U.S. Air Force, also featured prominently, indicating that military aviation accidents contribute significantly to the overall fatalities in aviation. Other major commercial operators such as Air France, American Airlines, and Pan American World Airways were also among the operators with the highest fatalities.

The findings underscore the importance of understanding the historical performance and safety records of specific aircraft types and operators in ensuring aviation safety. Identifying trends and patterns associated with high-fatality aircraft types and operators can inform regulatory authorities, airlines, and aviation stakeholders in implementing targeted safety measures and improvements to mitigate the risks associated with these aircraft and operators.

**When and where did aviation crashes occur, and are there any discernible patterns or trends in terms of time and location?**

The analysis revealed several patterns and trends regarding when and where aviation crashes occur. In terms of time, crashes were more frequent during certain months, with December, September, and August experiencing the highest number of incidents. This suggests a potential seasonal trend in aviation accidents, although further investigation would be needed to ascertain the underlying causes. Additionally, crashes occurred more frequently during certain decades, with the 1970s and 1980s seeing a higher incidence of accidents compared to other decades.

Regarding location, the United States emerged as the country with the highest number of flight crashes, followed by the Philippines, Colombia, and Russia. This distribution reflects the global nature of aviation accidents and the prevalence of crashes in regions with significant air traffic. Furthermore, specific routes, such as Tenerife - Las Palmas and Tokyo - Osaka, were associated with a disproportionately high number of fatalities, highlighting potential safety concerns on these routes.

Overall, the analysis suggests that aviation accidents exhibit both temporal and spatial patterns, with certain months, decades, countries, and routes experiencing a higher frequency of crashes. Understanding these patterns can inform efforts to improve aviation safety through targeted interventions, such as enhanced monitoring, infrastructure improvements, and regulatory measures in high-risk areas and during periods of heightened risk.

**Is there a correlation between the number of passengers on board an aircraft and the fatalities resulting from aviation crashes?**

The analysis examined the correlation between the number of passengers on board an aircraft and the resulting fatalities from aviation crashes. A linear regression analysis revealed a positive correlation between these two variables, with a regression equation of y = 0.58x + 4.0 and an R-squared value of 0.57. This indicates that, on average, an increase in the number of passengers on board is associated with a corresponding increase in the number of fatalities resulting from aviation crashes.

The findings suggest that the number of passengers plays a significant role in determining the severity of aviation accidents. This correlation underscores the importance of passenger safety measures and emergency response protocols in mitigating the impact of crashes. Additionally, it highlights the need for airlines and aviation authorities to prioritize safety considerations, such as aircraft capacity limits, evacuation procedures, and emergency preparedness, to minimize the risk of fatalities in the event of an accident. Overall, the analysis provides valuable insights into the relationship between passenger numbers and fatalities in aviation crashes, informing strategies to enhance aviation safety and mitigate the consequences of accidents.

**Limitations**

* Data accessibility: Access to comprehensive aviation crash data can be restricted due to confidentiality concerns, national security issues, or proprietary information held by airlines
* We encountered challenges within our dataset, particularly with entries from certain states in the United States lacking country designation. This lack of clarity significantly impeded our ability to accurately depict the global distribution of aviation incidents on a world map within the stipulated time of the project.

**Probable Next Steps / Recommendations**

1. **Data Refinement**: Allocate time to gather more comprehensive and diverse datasets related to aviation incidents, including historical crash data, weather patterns, aircraft performance metrics, pilot records, and air traffic data. Enhance data cleaning and preprocessing techniques to ensure high quality and consistency.
2. **Feature Engineering**: Invest efforts into feature engineering to extract meaningful features from raw data. Explore advanced techniques such as dimensionality reduction, time-series analysis, and anomaly detection to uncover hidden patterns and insights that can improve predictive performance.