### 

When considering the choice of graphs to display, there were three major questions that we hoped to answer. First, has air travel become more dangerous over time? Next, how does the safety of air travel compare to other modes of transportation? Finally, how have airline revenues trended over time?

INTRODUCTION:

Airline safety has consistently been a primary concern for airlines. To evaluate aviation safety and fatalities data, datasets from the Aviation Safety Network (airline safety) and airplane crashes from Kaggle were utilized. This assessment was conducted to obtain metrics on airline safety.

### Design Choices

I chose power BI as the tool to create dashboard for DSC 640 project visualizations. I have used color palette from venngage.com and avoided colors as per suggestion.

### Main Dataset

1. [Airline Safety](https://github.com/fivethirtyeight/data/tree/master/airline-safety), Aviation Safety Network - airline-safety.csv downloaded from Github.
2. [Accidents and Fatalities Per Year](https://docs.google.com/spreadsheets/d/1SDp7p1y6m7N5xD5_fpOkYOrJvd68V7iy6etXy2cetb8/edit#gid=1448957446), Accident and fatalities data as provided in project task.

### Supplemental Dataset

1. [Passengers transported](https://www.transtats.bts.gov/Data_Elements.aspx?Qn6n=F) per year downloaded from Bureau of Transportation Statistics.

Downloaded the data excel and filter total yearly number of passengers before saving the file. This was to avoid the multiple rows for each year which made it easier to create visualizations in Power BI.

1. [Airlines Operating Revenue](https://www.transtats.bts.gov/Data_Elements_Financial.aspx?Qn6n=K) (In Thousands of Dollars) from Bureau of Transportation Statistics.

Downloaded the data excel and filter total yearly revenue per year before saving the file. This was to avoid the multiple rows for each year which made it easier to create visualizations in Power BI.

### Dashboard

A close-up of a graph

Description automatically generated

The above dashboard on Airline Fatalities was generated using Power BI. The primary goal of the dashboard is to provide insights to the management on airline fatality and incident statistics and give recommendations on problem areas.

Since the safety of passengers was so integral to this study, many of our graphs are directly involved in showing that, overall, air travel has become safer over these years. It was important to put air travel into perspective, and simply comparing the number of fatalities year over year from one method to the other doesn’t give us the entire story. We compared the number of fatalities to hundred million miles transported, which let us compare the relative safety of car travel to air travel. In this case, a line graph with both metrics on it made the most sense. It was important to maintain the same axis for both, so as to not sow confusion. On average, we found that air travel fatalities per hundred million miles were 8.75 times lower than the corresponding value for car travel.

Next, we wanted to compare overall incidents over time. A grouped bar chart was chosen so we could easily compare the number of incidents for 2000-2014 to the 1985-1999 totals, and also the fatal incidents over both time periods. We were able to see that both the number of incidents and the number of fatal incidents both trended downwards. These incidents are becoming less prevalent, and when they do happen, they tend to be less fatal than they were in the past.

We wanted to see whether incidents were limited to a single geographical region, or a specific carrier, which led to the development of the next two metrics. We found that, of the 56 carriers we had information for, 7 of them were responsible for nearly 70% of incidents. A pie chart was selected to show just how much of the incidents were attributed to these high-risk carriers. Only one was a US-based carrier.

Our last graph seeks to answer how airline revenues have been affected. As you can see, since 1995 that trend has been mostly positive. You can see the downturns in 2002 (post 9/11) and 2009 (2009 Recession), but generally within 2 years that downward trend has corrected itself. We’ll see quite a spike here due to COVID-19, but in general the market has recovered.

Put together, these metrics paint the picture that generally air travel has become safer, not more dangerous, over time, especially compared to comparable transportation methods. Not only that, but revenues from flights are increasing steadily as well. It’s clear to see that as it stands, more people are comfortable flying and it remains one of the safest methods of transportation we have.

**The Scenario/ Business Problem**

Due to recent unfortunate airline crashes, the media has been promoting statistics stating air is no longer a safe way to travel. The news and media outlets have been bombarding the public with reports and figures about the trends of airline safety and that things are not looking good. What was previously thought as the safest way to travel, especially when compared to automobiles, is now being presented as one of the most dangerous to the public.

**Objective**

Explore available data to present facts /trends to present airline’s side of the story.

Investigate the data that is available to present facts or trends to present the airline's perspective.

**Design Tool**

I will be using Tableau with excel data source to present various facts through visualization

**Data Source**

To proceed with my study, I have referred “**Aviation Safety**” as core data source which is found in below link:

<https://aviation-safety.net/statistics/>

I also used supplemental data to look at existing flight statistics in US. Supplemental source is available in following link:

<https://www.transtats.bts.gov/Data_Elements.aspx?Data=1>

1. Airline Incidents Per Airline: This visualization presents data from the Airline Safety dataset, showcasing the total number of incidents reported. By examining the stacked bar chart, Aeroflot\* was responsible for most incidents.

Recommendation: It would be beneficial to conduct a thorough analysis of the incidents and identify potential safety measures to prevent similar incidents from occurring again in the future.

### Donut Chart

I have divided the period between two groups, 1985-1999 and 2000-2014, and then compared the Air Traffic Fatalities between them. This shows, how fatalities have reduced almost to half during 2000-2014 as compared to 1985-1999. Hence, showing how the air traffic has become more safer over the period of time.

### Line Chart – Passenger & Revenue

This chart is showing the Passenger count as well as Revenue as two different lines over the same period range. This shows how the passengers and revenues have shown to grow over the period and air travel has become more and more popular.

1. Line Chart – Fatalities per Year: This line chart depicts the number of fatalities per year, allowing us to observe trends over time. By analyzing the spikes in fatalities, we can identify potential reasons for the increase. We can see that the number of Fatalities started trending down in 1999, falling by 40.10% (389) in 10 years.﻿ The number of Fatalities dropped from 970 to 581 during its steepest decline between 1999 and 2009. Thanks to technological advancements and increased safety measures in air travel, the number of fatalities has significantly decreased.

All the above graphs prove that the media is portraying it in the wrong direction to the audience, and airline travel is still safer.

ETHICAL IMPLICATIONS:

* + - 1. Ensuring data accuracy is crucial, especially when the data is obtained from a public website rather than an airline-specific or government one. It is difficult to ascertain the accuracy of such data. We can see a discrepancy between the Airline Safety data and the airline crash data provided on Kaggle, which brings into question the reliability of the data source.
      2. It's worth noting that the crash reports dataset only goes up until 2014, which is a decade ago. It would improve the accuracy of the analysis if more recent data were available.

### References

**Data References:**

*Death Rate per Year* Retrieved from Bureau of Aircraft Accidents Archives: <http://www.baaa-acro.com/statistics/death-rate-per-year?page=0>

*Airline Safety*. Retrieved from Github - fivethirtyeight: <https://github.com/fivethirtyeight/data/tree/master/airline-safety>

*Operating Revenue (In Thousands of Dollars)* Retrieved from Bureau of Transportation Statistics: <https://www.transtats.bts.gov/Data_Elements_Financial.aspx?Data=7>

*Passengers All U.S. Carriers - All Airports*. Retrieved from Bureau of Transportation Statistics: <https://www.transtats.bts.gov/Data_Elements.aspx?Data=1>

**Color Palette Reference:**

*How to Use Color Blind Friendly Palettes to Make Your Charts Accessible*. Retrieved from Venngage: <https://venngage.com/blog/color-blind-friendly-palette/#3>

**Summary:**

As a data analyst at an airline company, I have been tasked with investigating media claims that air travel is no longer a safe way to travel. To help internal stakeholders combat this I created a high-level dashboard containing key information related to Airline safety and the negative publicity’s effect on our company. To begin with, the dashboard contains a comparison of fatality rate for Airline and Automobiles based on Billion vehicle miles traveled. The next visualizations show the declining trend in airline fatalities by year in a line chart and the total number of automobile fatalities per year in a bar chart. I have also included a pie chart indicating the number of fatalities by airline company for the years 2000-2014. Next, I have included a line plot showing the declining trend in total airline accidents and the trend in Operating Revenue/Net Profit by year.

I have been given the responsibility of looking into media reports that flying is no longer a safe mode of transportation as a data analyst for an airline corporation. I developed a high-level dashboard with important data pertaining to airline safety to assist internal stakeholders in addressing this.

**Presentation Plan:**

To begin my presentation, I would provide context on the media narrative around air travel before discussing where our company falls in the industry for number of fatalities. Understanding this may help our team determine if we want to shift focus to a single competitor if they are responsible for many fatalities with our own stats being low this could be an opportunity to use that to poach customers. Going forward we can examine the declining trends in fatalities and total accidents per year for airlines. This can be a launching point for discussion comparing the airline industry fatality rate and total fatalities per year to the automobile industry for talking points for the marketing team to use against the media narrative. Lastly, we can review the operating revenue and net profit per year to examine if we believe the media narrative has been impacting our bottom line.

**Ethical Considerations**

As with any project, we must ensure that our data is validated to ensure there are no misrepresentations or bias especially as this data is related to the safety of our airline and potentially life/death decisions for consumers.

Reference:

[Fatality Analysis Reporting System (FARS) | NHTSA](https://www.nhtsa.gov/research-data/fatality-analysis-reporting-system-fars)

[Impact | Airlines For America](https://www.airlines.org/impact/)

[Accidents and fatalities per year - Google Sheets](https://docs.google.com/spreadsheets/d/1SDp7p1y6m7N5xD5_fpOkYOrJvd68V7iy6etXy2cetb8/edit#gid=1448957446)

[data/airline-safety at master · fivethirtyeight/data (github.com)](https://github.com/fivethirtyeight/data/tree/master/airline-safety)

# Dashboard Summary

The focus for this milestone was to create a dashboard showing the facts about airline safety. Negative publicity within the media has recently portrayed airline travel as an unsafe method of transportation. As a Data Analyst for the airline organization, this dashboard will help summarize some of the main findings after review of four primary datasets. The audience for this dashboard is my internal peers within the airline organization (mainly the Data Science Management team). This dashboard will also help decision makers identify potential action to take for future improvements.

PowerBI was the tool chosen to generate the dashboard. The four datasets utilized for the initial analysis were from Github (Airline Safety), Aviation Safety Network, Bureau of Aircraft Accidents Archives, and Kaggle (The Statistics of Airplane Accidents). Importing the data from these sources was the first step. A link to the website for Bureau of Aircraft Accidents was used, otherwise a CSV file was generated (and cleaned) for the other three sources. Additional details about these sources can be found through the respective links within the Reference section. The visualizations and style of the dashboard were strategically created. There were eight visualizations chosen for the dashboard. These visualizations included a Slicer (by year), Card (showing airline accident count), Line and Stacked Column Chart, Map, 100% Stacked Column Chart, Pie Chart, Clustered Column Chart, and finally a normal Column Chart. The Slicer was implemented into the dashboard to allow users to review the data over specific years. The Card shows the number of airline accidents over the selected year range. This could be revised in future versions of the dashboard to show a ratio of the total outgoing flights compared to the number of accidents. The Line and Stacked Column Chart shows the trend of airline accidents over the years. This visualization shows how the number of accidents has drastically declined from 1946 until 2021. This chart also includes a five-year moving average to reiterate the trends over time. The Map shows the location of fatalities (and quantity based on bubble size) for the years with the highest number of fatalities. The 100% Stacked Column Chart compares the number of airline accidents to the number of fatalities over a particular period (2011-2021). The thought behind this visualization was that not all accidents may have resulted in fatalities. However, it does appear to show that an increase in accidents tends to result in additional fatalities. The Pie Chart shows a part-to-whole relationship of types aircrafts involved with the fatal accidents. The categories available from the data where Helicopters, Gliders, Gyroplanes, and Ultralight Aircrafts. The Clustered Column Chart shows the top ten airlines sorted by the number of incidents (between 2000-2014). This visualization also shows the number of fatal accidents associated with each airline. Lastly, the normal Column Chart represents the top ten airlines sorted by number of fatalities (between 2000-2014). The color scheme chosen for this dashboard was colorblind safe.

There are many summary points to conclude from this dashboard. In addition, there are future improvements and actions that can be taken within the organization. Improvement recommendations for the dashboard are: collect data for the most recent years to show the latest trends, work to manage the relationships between the different datasets used in the dashboard, and improve the visualization card showing a ratio of accident “opportunities” compared to actual “accidents”. Improvement recommendations within the organization are: continue driving the reduction of airline accidents over time (maybe set a goal to reduce by certain % each year), focus on helping airlines with the highest number of incidents reduce this number, focus on understanding airlines with the highest fatalities over time (understand why and then implement corrective actions), and lastly look into aircraft safety improvements to show how the type of aircraft is now safer over time. From an ethical standpoint, the facts need to be shown that airline accidents still occur periodically, however there needs to be communication showing improvements over the years. In addition, all of the data used for the dashboard can be found in the original PowerBI file for transparency.