**Author: Venkat Jagadeesh Jampani**

**Course: DSC-640 Data Presentation and Visualization**

**Assignment: Term Project Task 1 (Week 3&4)**

**Date: APR 9th, 2023**

**Airline Safety – Project Task-1**

## **Goal:**

The major goal of activity is to tell a story to the audience with visualization of different metrics to show airline is still one of the safest ways to travel and the fear of numbers accidents /fatalities are shown negatively in the media, but how in reality Air Travel is the safest way of travel.

## **Data Sources:**

1. **Primary Data:** Airline Safety
   1. **Filename:** Prim\_airline-safety.csv

This dataset contains the data for 56 airlines that were in the global top 100 as of Dec 2012 and which having operated continuously since 1985.

|  |  |
| --- | --- |
| **Field** | **Definition** |
| Airline | Airline (asterisk indicates that regional subsidiaries are included) |
| avail\_seat\_km\_per\_week | Available seat kilometers flown every week |
| incidents\_85\_99 | Total number of incidents, 1985–1999 |
| fatal\_accidents\_85\_99 | Total number of fatal accidents, 1985–1999 |
| fatalities\_85\_99 | Total number of fatalities, 1985–1999 |
| incidents\_00\_14 | Total number of incidents, 2000–2014 |
| fatal\_accidents\_00\_14 | Total number of fatal accidents, 2000–2014 |
| fatalities\_00\_14 | Total number of fatalities, 2000–2014 |

1. **Primary Data:** Accidents and Fatalities Per Year
   1. **File Name :** 
      1. Prim\_Accidents\_Fatalities\_Year.xls
      2. Sec\_Accidents\_Rate\_Per\_Year.xls
      3. Sec\_Airplane\_Crashes\_and\_Fatalities\_Since\_1908.csv
      4. auto-fatalities.xls
      5. Auto\_Fatalities\_1994\_2019.xlsx

This dataset contains the data on accidents and fatalities per year from 1946 to 2021. The dataset is accidents and fatalities from the above listed files are merged into one dataset.

|  |  |
| --- | --- |
| **Field** | **Definition** |
| Year | Year of the incident |
| Accidents\_Excl\_SBH | Accidents excluding suicide, sabotage, hijacking |
| Fatalities\_Excl\_SBH | Fatalities excluding suicide, sabotage, hijacking |
| Total\_Accidents | Total accidents including suicide, sabotage, hijacking |
| Total\_Fatalities | Total fatalities including suicide, sabotage, hijacking |
| Passenger\_Accidents | Passenger accidents including suicide, sabotage, hijacking |
| Passenger\_Fatalities | Passenger fatalities including suicide, sabotage, hijacking |
| Passenger\_Cargo\_Accidents | Passenger + cargo suicide, sabotage, hijacking |
| Passenger\_Cargo\_Fatalities | Passenger + cargo suicide, sabotage, hijacking |

1. **Supplemental Data:** Accidents and Death Rate per Year

This dataset contains total accidents and fatality rate from 1918 to 2022.

Reference: <http://www.baaa-acro.com/statistics>

|  |  |
| --- | --- |
| **Field** | **Definition** |
| Rank | Rank |
| Year | Year |
| Nb of Crashes | Number of Crashes |
| Fatalities | Fatalities count |

1. **Supplemental Data:** Total system passenger revenue and enplaned passenger

This dataset contains the consolidated airplane total operating revenue information from the year 1977 to 2021. Extracted another dataset containing the information about the consolidated airplane capacity information from the year 1950 to 2021

File Name: Financials\_Full Data\_data.csv

Reference: <http://web.mit.edu/airlinedata/www/Traffic&Capacity.html>

|  |  |
| --- | --- |
| **Field** | **Definition** |
| Year | Year |
| Cargo RTMs (mil) | Cargo RPM indicator |
| Departures (000) | Number of Departures |
| Load Factor | Load Factor |
| Passenger Enplanements (000) - Scheduled | Passenger Enplanements |
| Scheduled ASMs (000) | Aircraft ASM indicator |
| Scheduled RPMs (000) | Aircraft RPM indicator |

## **Why did you choose the visualizations you did? What were your findings?**

I have selected the data for 1. Primary safety data. 2. Consolidated Fatalities data. 3. Total Revenue of the Airlines over a period of time.

The selected datasets have been used to build the following charts in Tableau to expose the truth behind the negative news on air travel and the trends in each data selected. My primary goal is to show the trend of the airline crashes and usages to show that the airline travel is much safer today compared to early days.

1. **Total Air Accidents and Fatalities by Year: Line Chart**

This line chart is to show the total number of air incidents and fatalities occurred between the years 1920 and 2022. This is plotted between year and number of crashes and fatalities. The intention of this chart to show the trend is decreasing for number of crashes and fatalities as year passed on. The number of incidents and fatalities are high during 1940’s and 1970’s and gradually decreases for increase in year. This might be due to incorporating various safety measures which makes airline travel safer.

The trend tells us in the last 2 decades from 2000 to 2020, the crashes and fatalities have come down significantly.

1. **Passenger Flight Accidents and Fatalities Trend: Bar Chart/Line Chart**

This plot is also trend similar to previous chart. Here, the chart is plotted between year and passenger flight accidents and fatalities compared to previous which is plotted for total number of accidents and fatalities. I have used bar chart to show the accident trend while line chart is used to show fatalities trend. This chart also shows that trend is decreasing for increase in year similar to the earlier chart.

Both the accidents and fatalities incidents have come down increasingly in the last 2 decades. In the year 2020, the count has come down to single digits (both accidents and fatalities).

1. **Incidents by Airlines from 1985 to 1999 and 2000 to 2014 : Donut charts**

The next 2 plots are the donut charts plotted for number of incidents by airlines for the year ranges 1985 to 1999 and 2000 to 2014. I have chosen only top 15 airline companies by incidents to show in the charts. This is to show the total number of incidents have been reduced between 2000 and 2014 compared to those occurred between 1985 and 1999. In addition, we could see modest correlation between the charts. One of outlier we could see is Russia’s Aeroflot which had an extraordinarily high number of reported incidents in ’85-’99, with 28.15% of the total incidents. But many of these incidents are attempted hijackings around the time of the breakup of the Soviet Union. Contrastingly it has relatively low number in the ’00-’14 period. Another observation is Pakistan International Airlines and Ethiopian Airlines have had a persistently high rate of incidents.

The statistical inference from this chart is, some airlines are slightly safer to fly than others.

1. **Total Operating Revenue by Year: Bar Chart**

This bar chart is plotted for total operating revenue of airlines against year (for the time frame 1977 to 2021). We see the trend for operating revenue increases for each year except 2020 and 2021. The decrease in operating revenue for 2020 and 2021 are due to Covid-19. So, from this chart, we could understand that people prefer to travel in airplanes irrespective of few accidents reported. In for the rest of the years, the operating revenue is increasing year by year, that implies, more flights are operated over year by year, which further implies that more people prefer travel by air as it is safe.

1. **Number of Passengers on boarded : Area Graph**

This chart is similar to the previous plot of revenue. This area graph plotted between passenger usage and year shows the people’s preference for air travel increase for each year from 1950. The decrease in interest for 2020 and 2021 is due to Covid-19 situation. Each year the trend is more people are showing interest in travelling that shows the confidence in the general public’s preference in air travel, which shows it safe.

1. **Auto Deaths vs Airlines Deaths : Bar Chart**

This visual use supplemental data of auto fatalities in the USA for the same two periods (’85 - ’99, ’00 - ‘14). The airlines fatalities are filtered for US airlines and compared to the auto numbers. Although there are decreasing trend of incidents in both travels, auto casualties are consistently far greater than air travel. If the facts in the data are to be considered as truth, then this visual clearly shows that just by considering number of fatalities, travelling by road is far more dangerous than travelling by air. Though the number of incidents by road decreases over the period of time, but still the details due to air accidents are very minimal when compared to auto. This clearly shows that the it’s very safe to travel by air compare to Road.

## **How do you plan to present to your internal team?**

Will schedule a meeting with my internal team and walk through the visualizations created in Tableau with above observations. Also, I will prepare and share a PowerPoint presentation with all the charts and their observations and I will collect the feedback and will implement all the reasonable inputs from the internal team before presenting to the management.

## **Ethically what do you need to consider?**

Following are some ethical considerations I have considered in this research.

* Assessing only relevant information required for this analysis.
* Validating the data.
* Communicating the results appropriately.
* Collecting the required data that is relevant to the goal.
* Collecting the facts, whenever there is an issues with the trend. Eg : Covid, soviet union issue etc.
* Selecting the right charts that clearly present the goals.