****

**Business Decision Support Systems (MIS41040)**

**Tableau Assignment**

|  |  |
| --- | --- |
| **Student Number** | **Student Name** |
| **Leena Mary Varghese** | **23200412** |
| **Jeyaram Muthukrishnan** | **23200671** |
| **Shubhashree Saradavalli Ramesh** | **23200943** |

**Submission Date: 17-04-2024**

**Declaration: We declare that all material in this assignment is our work.**

**Navigating the Impact of COVID-19 on Air Travel: A Tableau Exploration**

**Problem Statement**

The COVID-19 pandemic in 2020 caused a period of confusion and uncertainty that presented the aviation industry with previously unheard-of challenges. Governments rushed to impose stringent regulations to stem the virus's spread as it swept over the globe. The aviation sector was severely and permanently harmed by a number of these measures, some of the most notable of which were travel restrictions, border closures, and nationwide lockdowns. The airline industry faced several challenges after these unprecedented disruptions, ranging from decreased passenger counts and income declines to operational challenges and strategic dilemmas.

**Access the Visualization here:**

<https://public.tableau.com/views/MIS41040Team33/Dashboard1?:language=en-GB&:sid=&:display_count=n&:origin=viz_share_link>

**Addressing Key Questions - Our project sets out to answer critical questions:**

1. **Impact on Airlines:** Our goal is to comprehend how the epidemic affected airlines throughout the globe. Examining how their daily operations, financial results, and crisis management strategies have changed are all part of this.
2. **Use of Aircraft:** We're looking at patterns in the use of aircraft during the epidemic and how these patterns differed between different airline firms and geographical areas.
3. **Passenger Behaviour:** Future plans will not be possible without an understanding of how passenger travel habits, preferences, and demand changed in response to safety concerns and constraints associated to the epidemic.

**Approach to Data Handling:**

Our analysis relies on the OpenSky Network dataset and involves several steps to ensure data integrity and reliability:

1. **Data Cleaning:** We take great care to eliminate errors like missing numbers, outliers, and inconsistencies from the dataset. This guarantees that our analysis is founded on trustworthy and accurate facts.
2. **Data Integration:** We include other datasets, such as airline and airport locations.
3. **Pattern Analysis:** Analysing patterns in missing data helps us understand possible causes for their recurrence, which helps us make well-informed judgements on imputation techniques and data management tactics.

**Data Cleaning Process:**

1. **Data Selection:**

Choosing the right time period for our research was a critical first step in our analysis of how the COVID-19 outbreak affected air travel. Since the epidemic was unusual and had a significant impact on air travel worldwide, we strategically decided to only look at data from 2020. The aim to record air travel at the height of the pandemic, free of any pre-existing patterns or post-pandemic recovery initiatives, was the driving force for this choice.

1. **Removal of Null Values:**

Dealing with null values was one of the first issues we ran across when cleaning the data, especially in the "Origin" column. We made the decision to exclude these null values from the dataset after giving it some thought. We made this decision because we realised that transit points—which are indicated by null values in the "Origin" column—did not add anything useful to our understanding of how the pandemic affected air travel. Our goal in removing these missing information was to make sure that our dataset was accurate and comprehensive.

1. **Exclusion of Military and Private Aircraft:**

We made the decision to exclude military and private aircraft from the dataset in order to keep the emphasis of our study on commercial air travel and its connection to the pandemic. Despite the possibility that these aircraft types were in use during the epidemic, their inclusion would have increased noise and masked the real effects of COVID-19 on commercial aviation. We aimed to improve the reliability and clarity of our study by eliminating these outliers, giving stakeholders valuable insights into the impact of the pandemic on the aviation sector.

1. **Removal of Outliers:**

Another important step in our data cleansing process was locating and dealing with outliers. We deliberately focused on anomalies, such private runoff flights, as they didn't significantly add to our understanding of how COVID-19 affected air traffic. Our goal was to increase the accuracy of our results and minimise noise by removing these unusual data points.

1. **Standardisation of Time Data:**

For analysis and visualisation to be made easier, time-related data has to be formatted consistently. In order to accomplish this, we combined the "day and time" column into a single "day" column, standardising its structure. Our goal in streamlining the dataset's temporal aspect was to make it easier for stakeholders to understand and expedite the analysis process. We made this choice because we wanted to make it easier to grasp and comprehend the data by presenting it clearly and coherently.

1. **Integration of External Data:**

We added external data sources with airline name information to enhance our dataset and offer more context for analysis. Our capacity to associate call signals with certain airlines was made possible by this integration, which improved the interpretability of our results. Our objective was to provide stakeholders with a more all-encompassing comprehension of the issues impacting air travel during the epidemic by integrating external data.

1. **Matching of IATA Codes:**

We matched airport names with their corresponding International Air Transport Association (IATA) codes after conducting a thorough data-cleaning procedure that prioritised consistency and clarity. The dataset's origin and destination locations were made easier to understand by this alignment. We sought to improve the accessibility and comprehensibility of our dataset for stakeholders by making sure that IATA codes matched airport names appropriately.

1. **Creation of Route Column:**

During the data cleansing process, another important choice was to augment the dataset with a new "route" column. We made it easier to examine aircraft paths and patterns during the epidemic by combining origin and destination data. Stakeholders were able to learn more about the historical development of aircraft routes and the geographic distribution of air transport thanks to this innovation.

1. **Removal of Altitude Data:**

We decided to exclude the altitude column from the dataset as it was not relevant to our examination of COVID-19's effects on air travel. Significant insights into the impact of the pandemic on aircraft frequency or passenger behaviour were not revealed by altitude data. We sought to simplify the dataset and concentrate on factors that were specifically pertinent to our study goals by removing this unnecessary material.

1. **Airport Name Integration:**

We combined airport names using origin data to improve our comprehension of passenger flow patterns. We were able to determine which airports were the busiest and examine trends in passenger flow during the epidemic thanks to this connection. We streamlined the analytic process and gave stakeholders useful information about airport activity throughout the pandemic era by grouping airport names together inside the dataset.

During the data cleaning process, we made a number of calculated choices with the goal of streamlining the dataset and obtaining useful information on how the COVID-19 epidemic affected commercial aviation. Every choice we made was carefully considered in light of how it would affect the precision, applicability, and clarity of our study.

**Setting Up Tableau:**

We started off by purchasing Tableau Desktop, a potent tool for data analysis and visualisation. We can now turn raw data into interactive dashboards using Tableau, giving people a more intuitive way to explore and comprehend large, complicated information.

**Visualisation Objectives:**

Visualisation Objectives: Our main objective was to develop an interactive decision support system (DSS) that went beyond conventional reporting by giving users access to visualisations. These visualisations are tools for decision-making in addition to being a way to display facts. In order to enable users to extract useful insights from the data, we included tools like filters and user-controlled parameters.

**Analysis in Tableau:**

We performed a thorough study of the aviation information in Tableau, concentrating on important metrics and trends that are essential for industry decision-making. An interactive representation of our findings is provided below.

1. **Route Map**: We visualised frequent flight routes, allowing users to filter by month and country of origin to analyse variations over time.
2. **Top 5 Airlines**: Through interactive charts, we showcased the performance of the top 5 airlines along with their percentage contribution enabling users to explore monthly variations in flight operations.
3. **Dips in Flights operated**: A dynamic visualisation displayed the dips in the number of flights operated across each month in the year 2020, offering insights into seasonal trends and overall activity levels.
4. **Top Routes**: By highlighting the busiest routes according to passenger volume, we made route optimisation and strategic planning easier.
5. **Total Routes per Month**: Through the visualisation of the monthly route operations, users may discern trends and variations in the growth or reduction of the network.
6. **Busiest Airport**: An interactive map pinpointed the busiest airports, allowing users to delve into passenger traffic volumes and operational trends.
7. **Total Number of Aircrafts Operated**: We provided information on the total number of aircraft operated using comprehensible charts, which aided in capacity planning and fleet management.
8. **Average Flight Distance**: The average flight distance in miles was visualised to offer information about fuel usage and route planning.

**Key Insights:**

Following a comprehensive Tableau analysis of the aviation dataset, several interesting discoveries were made that will be helpful to aviation sector decision-makers. Even though users may freely explore and modify their analysis using interactive filters, a few crucial findings hold true in all scenarios:

1. **American Airlines Dominance**: In spite of the difficulties caused by the COVID-19 epidemic, American Airlines was the best-performing airline all year long. This resilience demonstrates how strong its operating plans are and how much demand there is for its services.
2. **Passenger Behaviour**: Observations of passenger conduct throughout the pandemic era revealed important patterns in demand, preferences, and travel behaviour. There were observable trends in route preferences and travel frequency despite limitations and safety concerns, indicating the intricate interaction of elements impacting customer decision-making.
3. **Busiest Airport**: Dallas Fort Worth International: Despite the difficulties caused by the epidemic, our data showed that Dallas Fort Worth International Airport continued to be the busiest airport, managing a large amount of passenger traffic and flight operations. This emphasises its strategic significance as a pivotal aviation network hub.
4. **Frequent Flight Routes**: During the period under analysis, the domestic routes in the United States stood out among the frequently detected aeroplane routes. The popularity of this route indicates a high level of passenger demand and its strategic significance for both business and leisure tourists.
5. **Peak in Routes in March**: March marked the highest number of routes operated during this period, suggesting that March saw a peak in aviation activity. This may be a result of seasonal travel trends or particular occasions generating greater demand for air travel.
6. **Top Route**: Tokyo to Osaka: The route from Tokyo to Osaka was found to be the most popular overall based on our data, demonstrating a high volume of operating and passenger travel between these locations. For airlines that operate in these areas, this information may help with marketing and route planning.

**Conclusion:**

To sum up, the process of developing and executing the decision support system (DSS) for aviation analytics has been rewarding. We have gained essential knowledge about managing, analysing, and visualising data, which has led to the creation of an effective tool for well-informed decision-making. Our commitment to quality and teamwork have produced a strong DSS that gives stakeholders the confidence they need to successfully negotiate the challenges of the aviation sector.

**References:**

*1.Zenodo OpenSky Dataset (2022) OpenSky Network Dataset, [online] Available at:* [*https://zenodo.org/records/5815448*](https://zenodo.org/records/5815448)

*2.www.kaggle.com. (n.d.). Covid-19 Global Dataset. [online] Available at:* [*https://www.kaggle.com/datasets/josephassaker/covid19-global-dataset?resource=download*](https://www.kaggle.com/datasets/josephassaker/covid19-global-dataset?resource=download)

*3.World Bank. (n.d.). Global Airports Dataset. Retrieved from World Bank Data Catalog:* [*https://datacatalog.worldbank.org/search/dataset/0038117/Global-Airports*](https://datacatalog.worldbank.org/search/dataset/0038117/Global-Airports)

*4.Banerjee.S(n.d). Airline Dataset. Retrieved from Kaggle:*

[*https://www.kaggle.com/datasets/iamsouravbanerjee/airline-dataset*](https://www.kaggle.com/datasets/iamsouravbanerjee/airline-dataset)