|  |  |
| --- | --- |
| **Title** | Analysis of Airline Data |
| **Name** | Kyi Thin Nu |
| **Tools** | Excel, Power Pivot and Power Query |
| **Total Working Hours** | 16 Hours |
| **Timeline** | 22/Jun/2024 – 20/July/2024 |

# **Objective**

Purpose:

* To explore the airline data nature and how the analysis could be done based on the airline data to offer insights into functioning and efficiency of the aviation industry.

Motivation:

* To understand the analysis process for the specific dataset (airline dataset in this project) using Excel and Power Pivot.

## **Data Source**

* Dataset downloaded from Kaggle: <https://www.kaggle.com/datasets/iamsouravbanerjee/airline-dataset/data>.
* This is a Secondary Dataset.

## **Data Cleaning Process**

1. Check the data types: Are the features identified with correct data types?
2. Check the null values.
3. Remove the columns that are not necessary in this analysis project.
   1. Remove Customer First and Last Name
4. Add the age group according to the age columns:

Here are the rules for defining the age group in this project:

|  |  |
| --- | --- |
| Age | Age Group |
| 0-2 | Infant |
| 3-12 | Child |
| 13-20 | AdultJunior |
| 21-35 | Adult |
| 36-60 | AdultSenior |
| 61-90 | Senior |
| Above 90 | GoldenAge |

1. For Data modelling: I need to add a new sheet that contains country and airport code (country + code) // skip in this project
2. Data Cleaning >> the arrival airport = (0, -) values
3. Create a new column >> **Route by combining Departure Airport and Arrival Airport (done)**
4. **Create a new column >> Flights by combining Route and Departure Date (done)**

## **Dashboard Functionality**

1. Overview Session (Key Metric and KPIs)
   1. Total Passengers: Display the total number of unique passengers.
   2. Total Flight
      1. Merge (airport country code + arrival airport + departure status) >> total (unique flights)
2. Passenger Demographic Analysis
   1. Gender Distribution: What is the percentage of Men and Women distribution?
   2. Age group distribution for airline: Which age groups are the most popular in the airline?
   3. Nationality Distribution: Top 10 nationalities among the passengers.
3. Flight Routes Analysis**:** Analysis of flights by departure and arrival airports.
4. The most popular routes in the airline for the year 2022? (top 10 most departure and visited country)
   * The most popular route with time series: in which month which route is popular?
5. **Popular Arrival Airports**
   * Which continent is the most visited?
6. **Popular Departure Airports**
   * the most common departure airports
7. how many routes for each day? (extra)
8. total passengers in each single flight
9. Flight Status Analysis: What is the flight status throughout the year?
   1. **Flight Status Overview:** Pie chart showing the distribution of flight statuses (on-time, delayed, cancelled).
      1. **Filter with time (Monthly, Quarterly)**
      2. Which month is the most happening for each flight status?
   2. **Flight Status by Pilot:** Performance of flights categorized by pilot names (on-time, delayed, cancelled).
      1. To get the pilot column back from the dataset
   3. Flight Status by Airport (Airport Performance): Which airport is the most on time, delayed, cancelled?
10. Trend Analysis
11. **Flights Over Time:** Line chart showing the number of flights over the months of 2022.
12. **Passenger Trends Over Time:** Line chart showing the number of passengers over the months of 2022.
13. Geographical Analysis
    1. **Flight Routes Map:** Interactive map showing flight routes from departure to arrival airports.
    2. **Passenger Nationality Map:** Map showing the distribution of passenger nationalities.

## **Metrics and Charts**

1. Overview Session (Key metrics and KPI)

|  |  |
| --- | --- |
| Chart Title | Total Passengers |
| Chart Type | Card |
| Dimensions | Number of unique passengers |
| Metrics | unique passengers count |

|  |  |
| --- | --- |
| Chart Title | Total Flight (not sure to proceed) |
| Chart Type | Card |
| Dimensions | Number of unique flight count |
| Metrics | unique flight count (1 unique route = by taking a unique data from airport name, departure date, arrival airport) |

1. Passenger Demographic Analysis

|  |  |
| --- | --- |
| Chart Title | **Gender Distribution** |
| Chart Type | Bar Chart / Pie Chart |
| Dimensions | Gender (Male, Female) |
| Metrics | Count the number of PassengerID (by Genger) |

|  |  |
| --- | --- |
| Chart Title | **Age Group Distribution** |
| Chart Type | Bar Chart / Pie Chart |
| Dimensions | Age Groups |
| Metrics | Count the number of PassengersID (by Age Groups) |

|  |  |
| --- | --- |
| Chart Title | **Nationalities Distribution** |
| Chart Type | Bar Chart / Pie Chart |
| Dimensions | Nationality |
| Metrics | The number of passengers by Nationality |

1. Flight Route Analysis

|  |  |
| --- | --- |
| Chart Title | **Popular Routes** |
| Chart Type | Bar Chart |
| Dimensions | Route |
| Metrics | The top 10 routes (based on frequence or count) |

|  |  |
| --- | --- |
| Chart Title | **Popular Arrival Airports** |
| Chart Type | Bar Chart |
| Dimensions | Arrival Airports |
| Metrics | The most common arrival airports (based on frequence or count) |

|  |  |
| --- | --- |
| Chart Title | **Popular Departure Airports** |
| Chart Type | Bar Chart |
| Dimensions | Departure Airports |
| Metrics | The most common departure airports (based on frequence or count) |

1. Flight Status Analysis

|  |  |
| --- | --- |
| Chart Title | **Flight Status Overview** |
| Chart Type | Pie Chart |
| Dimensions | Flight status (on-time, delayed, cancelled) |
| Metrics | Distribution of flight status (on-time, delayed, cancelled) |

|  |  |
| --- | --- |
| Chart Title | **Flight Status by Pilots** |
| Chart Type | Stacked Bar Chart |
| Dimensions | Pilot, Flight status (on-time, delayed, cancelled) |
| Metrics | Performance of flights categorized by pilot names (on-time, delayed, cancelled). |

|  |  |
| --- | --- |
| Chart Title | **Flight Status by Airport** (Airport Performance) |
| Chart Type | Stacked Bar Chart |
| Dimensions | Airport, Flight status (on-time, delayed, cancelled) |
| Metrics | Performance of flights categorized by airport names (on-time, delayed, cancelled). |

1. Trend Analysis

|  |  |
| --- | --- |
| Chart Title | **Flights Over Time** |
| Chart Type | Line Chart |
| Dimensions | Number of Flights, Months |
| Metrics | the number of flights over the months of 2022. |

|  |  |
| --- | --- |
| Chart Title | **Routes Trends Over Time** |
| Chart Type | Line Chart |
| Dimensions | Number of Routes |
| Metrics | the number of routes over the months of 2022. |

1. Geographical Analysis

|  |  |
| --- | --- |
| Chart Title | **Flight Routes Map** |
| Chart Type | Interactive map |
| Dimensions | Departure and Arrival Airports |
| Metrics | flight routes from departure to arrival airports. |

|  |  |
| --- | --- |
| Chart Title | **Passenger Nationality Map** |
| Chart Type | Interactive map |
| Dimensions | Nationalities |
| Metrics | the distribution of passenger nationalities. |

## **Dashboard Mockup**

A screenshot of a computer

Description automatically generated

# Final Report

Note: Some data are irrelevant to construct the analysis, therefore the final report is a little bit different.

A close-up of a computer screen

Description automatically generated

## **References**

<https://github.com/shwezin-coder/Analysis-Project-Planning-Template>

[Search Vizzes | Tableau Public](https://public.tableau.com/app/search/vizzes/airline)

<https://public.tableau.com/app/profile/andy.kriebel/viz/AirlineDelays_2/AirlineDelays>

https://public.tableau.com/app/profile/madhu.samudrala/viz/Airline\_14/AirlineDelaySummary