

Data Visualisation Assignment

by Ben Roberts

Submission date: 08-Jan-2024 12:49AM (UTC+0000)

Submission ID: 220872463

File name: 41744_Ben_Roberts_Data_Visualisation_Assignment_531139_310757898.pdf (2.42M)

Word count: 2240

Character count: 13613

Assignment 1 - Critiquing and Creating Data Visualisations

Ben Roberts (2205382)

Data Analysis & Visualisation (ACCB4002)

University of Wales Trinity Saint David – School of Applied Computing

Table of Contents

Introduction.....	3
Task 1 – Import all the Data.....	4
Task 2 – Data Preparation.....	6
Task 3 – Inspect the relationship of the data in the "Data Model" tab.....	9
Task 4 – Analyse the data using visualisations.....	10
Task 5 – Add a user-interactive filter.....	13
Task 6 - Top 5 Airlines by total cost.....	16
Task 7 – Count of airlines in different countries.....	17
Task 8 – Analysing Busiest Port.....	18
Task 9 – Analysing Freight Out.....	19
Task 10 - Formatting Report.....	20
Conclusion.....	21

Table of Figures

Figure 1: Importing aircraft datasets.....	4
Figure 2: All datasets imported.....	4
Figure 3: Aircraft dataset with added columns.....	5
Figure 4: Marking the first row of the airline table as a header.....	6
Figure 5: Replacing null with 0.....	6
Figure 6: All instances of null have been replaced with 0.....	6
Figure 7: Changing the data type to a fixed decimal number.....	7
Figure 8: The transformed dataset with fixed decimal numbers.....	7
Figure 9: Splitting columns with PowerBI's transform data tool.....	8
Figure 10: View of the dataset with the "airline" column split by leftmost delimiter.....	8
Figure 11: View of the relationships within the dataset.....	9
Figure 12: Full-screen view of the bar chart with the data included.....	10
Figure 13: Full screen view of the stacked area chart showing the current total cost of all airlines.....	10
Figure 14: Full screen view of a line graph showing the yearly trend of increases and decreases in freight.....	11
Figure 15: The same visualisation filtered to only display freight in information for Fiji.....	11
Figure 16: Normal sized view of the line chart without filtering.....	11
Figure 17: Card-type visualisation containing the sum of the total number of orders.....	12
Figure 18: Early version of the visualisation with a filter.....	13
Figure 19: Explicit filters set to only show data for Airbus A300 and Airbus 310 aircraft.....	14
Figure 20: Current version of the filter which is a drop-down box and has a select all option.....	15
Figure 21: Small view of a treemap visualisation showing the top five most expensive airlines.....	16
Figure 22: The above visualisation in fullscreen.....	16
Figure 23: Map visualisation of count of airlines in different countries.....	17
Figure 24: DAX function to find the sum of all in and out of every port.....	18
Figure 25: Full screen view of the matrix visualisation with the percentage grand total visible.....	18
Figure 26: Function generated using PowerBI's quick measures option.....	19
Figure 27: Full screen view of the gauge chart.....	19
Figure 28: Fully completed visualisation with the title "Executive Summary Report".....	20
Figure 29: The above visualisation with a filter active.....	20

Introduction

Delta Fast is a company that manufactures and assembles commercial aircraft. It is relatively new, having been in business for only two years. This report has been commissioned by Delta Fast because there is a budgeting review coming up and the board of directors are planning to decide on which areas of investment to focus on in the upcoming financial year.

The intention of report is to determine the total number of orders for specific types of aircraft and analyse the total cost of said aircraft, calculate the top five airlines that could potentially produce the highest possible revenue, and analyse the total input and output of freight and passengers of specific airports to achieve greater insight into which is the most profitable.

This data visualisation will be utilising PowerBI because it allows the user to quickly produce graphical visualisations of the information and can produce a wide variety of graphs and charts.

Task 1 – Import all the Data

The first step of the visualisation process is to import the dataset. In this case, each of the datasets are stored as a comma separated values (CSV) file and must be imported separately.

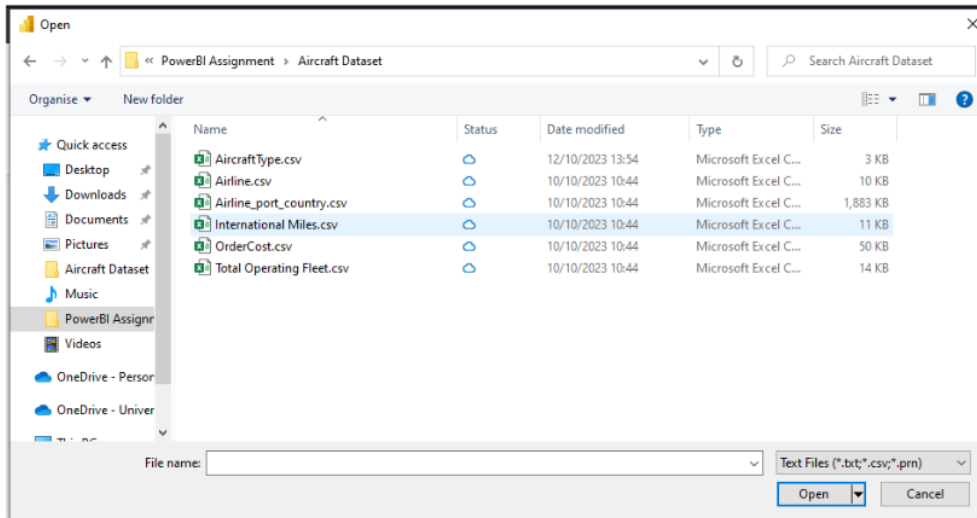


Figure 1: Importing aircraft datasets

Upon importing the data, it was added to the data menu. This allows it to be dragged-and-dropped into a visualisation.

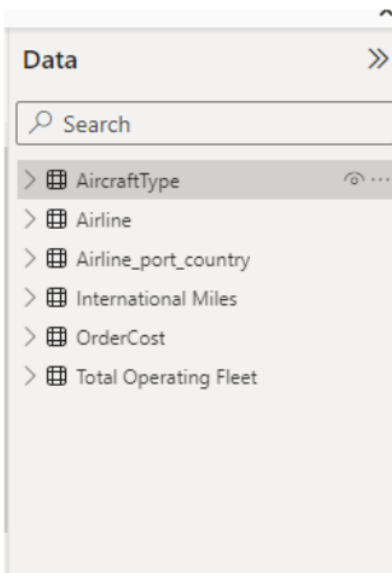


Figure 2: All datasets imported

The screenshot shows a data visualization tool interface. The main table displays aircraft data with columns for Aircraft Type, Aircraft ID, and Column1. The table is titled 'Table: AircraftType (101 rows) Column: Column2 (1 distinct values)'. The right sidebar shows a search bar and a list of columns, including AircraftType, Column1, Column2, and various date and time columns.

Aircraft Type	Aircraft ID	Column1
Airbus A320	7001	
Airbus A320	7002	
Airbus A321	7003	
Airbus A321	7004	
Airbus A321	7005	
Airbus A321	7006	
Airbus A321	7007	
Airbus A321	7008	
Airbus A321	7009	
Airbus A321	7010	
Airbus A321	7011	
Airbus A321	7012	
Airbus A321	7013	
Airbus A321	7014	
Airbus A321	7015	
Airbus A321	7016	
Airbus A321	7017	
Airbus A321	7018	
Airbus A321	7019	
Airbus A321	7020	
Airbus A321	7021	
Airbus A321	7022	
Airbus A321	7023	
Airbus A321	7024	
Airbus A321	7025	
Airbus A321	7026	
Airbus A321	7027	
Airbus A321	7028	
Airbus A321	7029	
Airbus A321	7030	
Airbus A321	7031	
Airbus A321	7032	
Airbus A321	7033	
Airbus A321	7034	

Figure 3: Aircraft dataset with added columns

After importing the dataset, the next step was to add more blank columns to the Aircraft dataset. These columns ultimately went unused, but they could be included is future-proofing so that more data could quickly be added to the visualisation.

Task 2 – Data Preparation

The first row of some of the CSV files were marked as headers during the import process by using the transform data tool. However, select datasets were imported as-is and had to be set as headers using the transform tool.

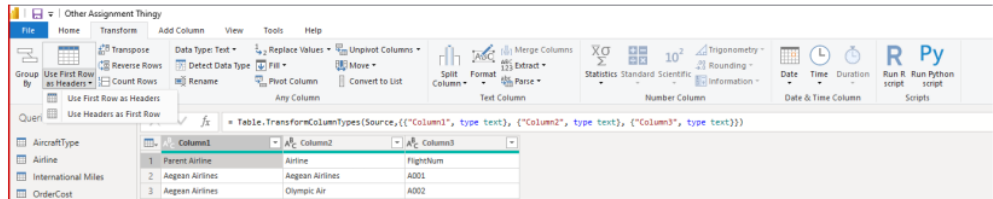


Figure 4: Marking the first row of the airline table as a header

The next step was then to use the data transformation utility built into PowerBI to alter the datasets in a way that makes them easier to visualise.

To make the data visualisations able to display null values, the first step was to change all instances of “null” to “0”.

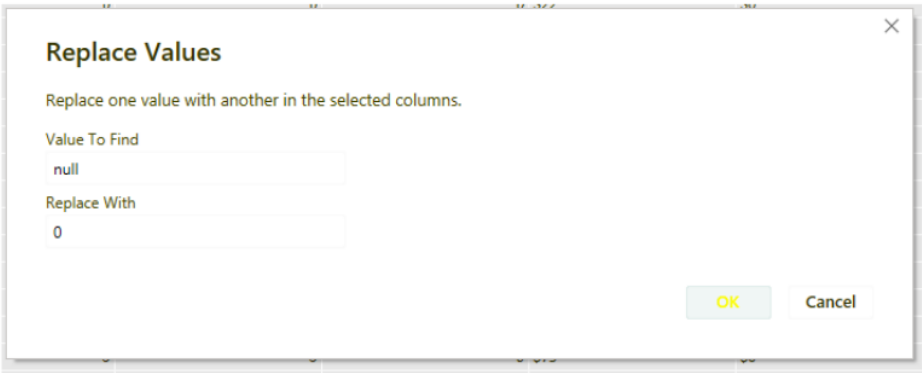


Figure 5: Replacing null with 0

1.2 Customer	1.2 Flight	1.2 Route	1.2 Total	1.2 Order	1.2 Unit Cost	1.2 Total Cost (Current)	1.2 Average Age	1.2 FlightNum
1	0	0	0	0	0 \$50	\$50	11.6 A001	0 A001
2	0	0	0	0	0 \$90	\$0	0 A002	0 A002
3	38	0	3	41	0 \$98	\$3,724	7.5 A001	0 A002
4	0	0	0	0	0 \$98	\$0	0 A002	0 A002
5	8	0	0	8	0 \$115	\$919	10.3 A001	0 A002
6	2	0	0	2	0 \$22	\$44	2.4 A002	0 A002
7	0	0	0	0	0 \$22	\$0	0 A002	0 A002
8	0	0	17	17	0 \$74	\$0	0 A001	0 A001
9	0	0	0	0	0 \$14	\$0	0 A001	0 A001
10	12	0	7	19	0 \$11	\$376	12.6 A002	0 A002
11	0	0	14	14	0 \$75	\$0	0 A003	0 A003
12	38	0	5	31	0 \$90	\$2,190	12.7 A004	0 A003
13	0	0	15	15	0 \$90	\$0	0 A003	0 A003
14	30	0	0	20	0 \$90	\$896	11 A005	0 A005
15	71	2	3	75	21 \$98	\$6,958	5.7 A005	0 A005
16	5	0	8	13	0 \$98	\$490	15.2 A004	0 A004
17	31	0	8	39	0 \$115	\$3,162	3.6 A003	0 A003
18	22	0	0	22	0 \$240	\$5,280	5.9 A003	0 A003
19	0	0	0	0	22 \$112	\$0	0 A003	0 A003
20	0	0	3	3	0 \$75	\$0	0 A003	0 A003
21	0	0	2	2	0 \$75	\$0	0 A004	0 A004
22	0	0	0	0	0 \$28	\$0	0 A004	0 A004
23	20	1	18	17	28 \$74	\$1,480	1.6 A003	0 A003
24	12	0	0	12	0 \$74	\$888	2.2 A004	0 A004
25	17	1	5	23	0 \$74	\$1,258	9.6 A004	0 A004
26	0	0	5	5	0 \$74	\$0	0 A005	0 A005
27	7	2	0	9	2 \$116	\$2,211	16.7 A004	0 A003
28	0	0	13	13	0 \$187	\$0	0 A003	0 A003
29	0	0	9	9	0 \$187	\$0	0 A004	0 A004
30	15	0	2	17	1 \$295	\$4,418	2.5 A003	0 A003
31	6	5	1	12	5 \$295	\$1,767	17.7 A004	0 A003
32	0	0	0	0	22 \$285	\$0	0 A003	0 A003
33	8	2	1	12	0 \$11	\$290	14.8 A005	0 A004
34	0	0	1	1	0 \$45	\$0	0 A004	0 A004

Figure 6: All instances of null have been replaced with 0

The dataset included the numeric values for total cost stored as text, which meant they could not be used for the visualisation and would not work when used in calculation. To convert it, all special characters had to be removed via the replace values tool and the data type needed to be changed to a fixed decimal number.

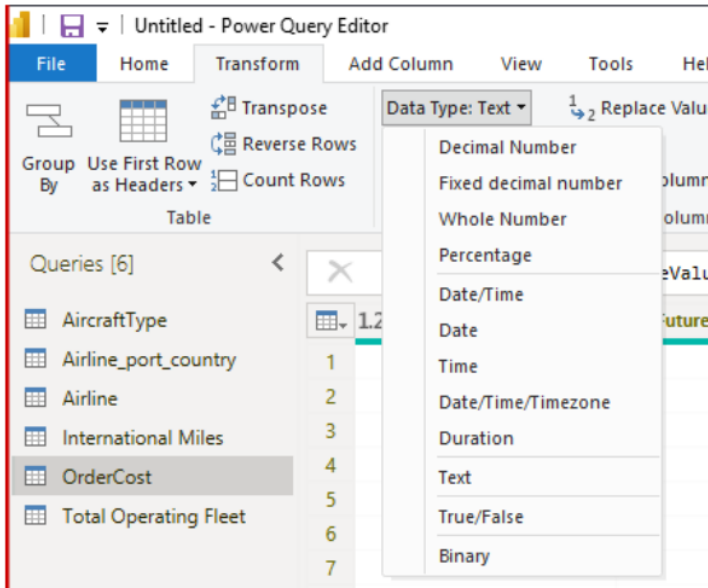


Figure 7: Changing the data type to a fixed decimal number

The screenshot shows the Power Query Editor with a table of data. The table has columns for 'AircraftType', 'Airline_port_country', 'Airline', 'International Miles', 'OrderCost', and 'Total Operating Fleet'. The 'Total Cost' column is highlighted, showing numeric values. The table is sorted by 'Total Cost' in descending order. The 'OrderCost' column is also highlighted, showing numeric values. The 'Total Operating Fleet' column is highlighted, showing numeric values. The table is displayed in a grid format with rows and columns. The 'Total Cost' column is the first column, followed by 'OrderCost', 'Total Operating Fleet', and then the other columns. The values in the 'Total Cost' column are numeric, indicating that the data type has been successfully changed from text to a fixed decimal number.

Figure 8: The transformed dataset with fixed decimal numbers

The airline column was also split, although no practical reasons for splitting the column were found and it also broke some of the relationships within the dataset. The column was later re-merged as a result, which repaired the data set.

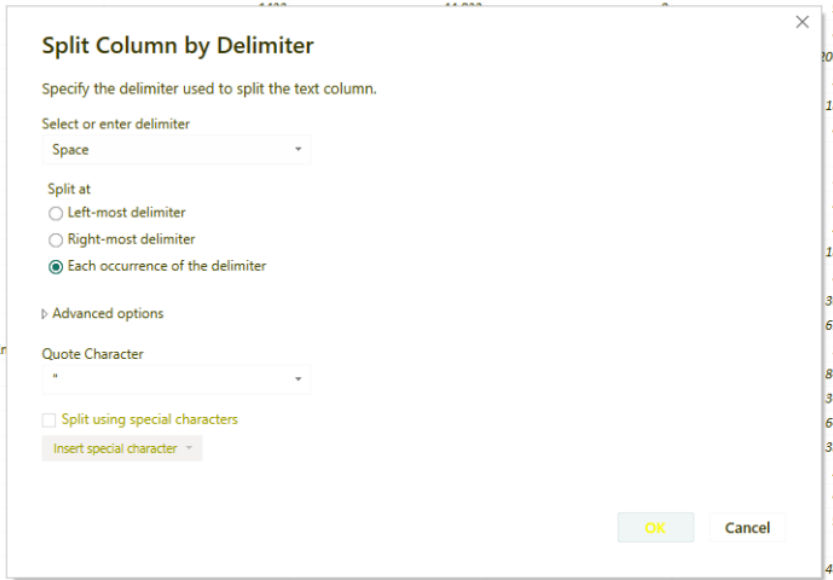


Figure 9: Splitting columns with PowerBI's transform data tool

[illegible]

Figure 10: View of the dataset with the "airline" column split by leftmost delimiter

After that, the data transformation tool was closed and the modified data was loaded into PowerBI.

Task 3 – Inspect the relationship of the data in the “Data Model” tab

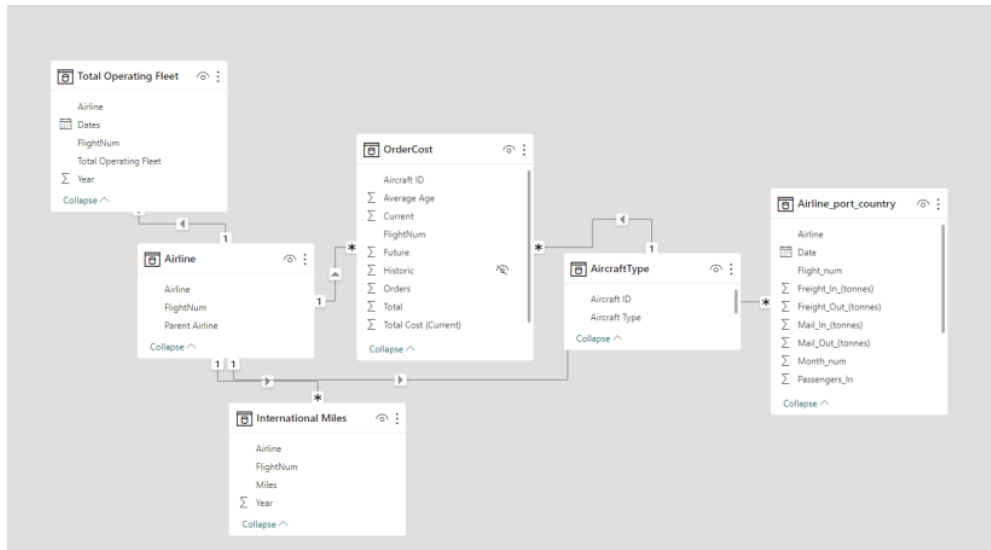


Figure 11: View of the relationships within the dataset

This is a view of the entities and relationships within the dataset. The relationships are represented by lines with symbols on each end to show whether it is a one-to-many relationship or a one-to-one relationship. By viewing the model tab, it can be noted that all of the relationships within the dataset are one-to-many.

Task 4 – Analyse the data using visualisations

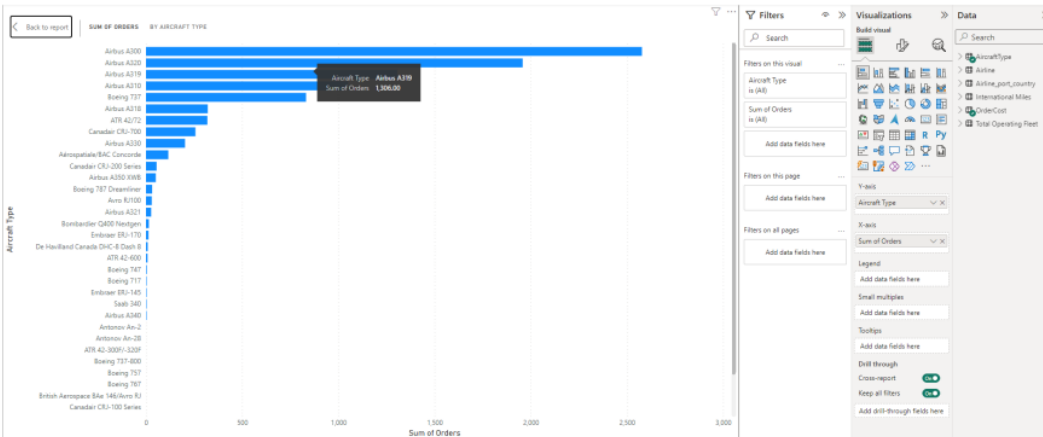


Figure 12: Full-screen view of the bar chart with the data included

This is a horizontal bar chart showing the bulk of orders for each type of aircraft. It was created by adding a bar chart element to the PowerBI visualisation and dragging-and-dropping the “Aircraft Type” and “Orders” values; the latter of which was automatically set as a sum (hence why it is labelled “Sum of Orders”).

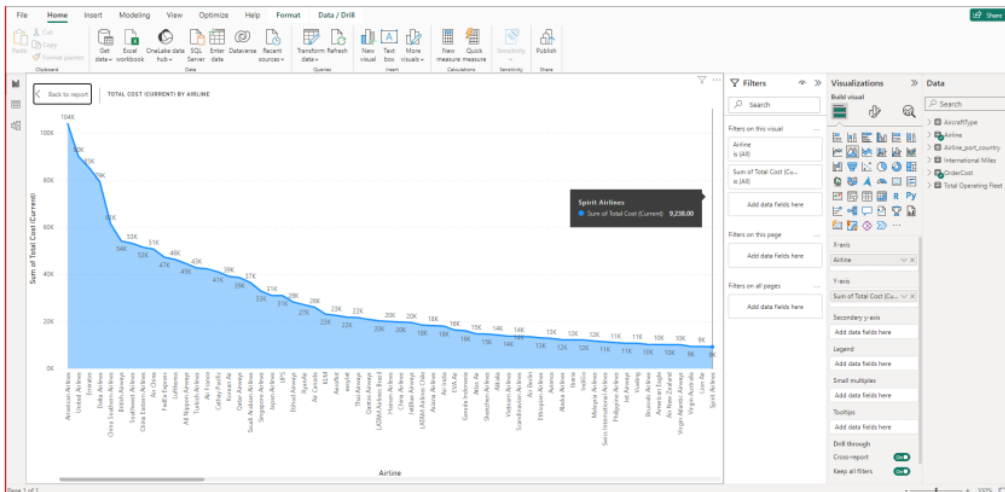


Figure 13: Full screen view of the stacked area chart showing the current total cost of all airlines

The next step was to create a stacked area chart to show which airlines are incurring the highest total cost. It was created similarly to the bar chart visualisation, but using a stacked area chart element and the “Airline” and “Total Cost” values. At \$103,892, American Airlines has the highest total cost.

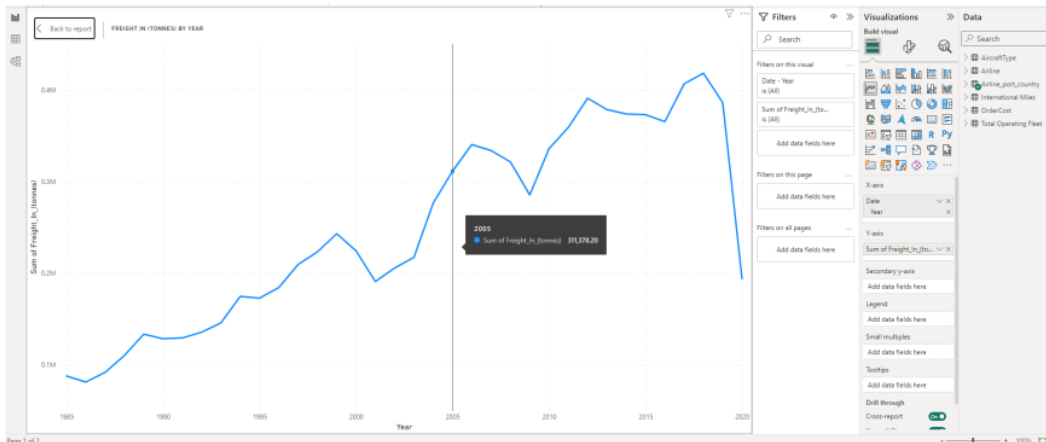


Figure 14: Full screen view of a line graph showing the yearly trend of increases and decreases in freight

A line graph was used to show the yearly trend of freight in information. In this instance, the calculated value "Sum of Freight_In_(tonnes)" was used for a numeric value by dragging it into the Y axis. The date value also had to be filtered to explicitly only display the year, which was done by unticking everything except the year in the date hierarchy.

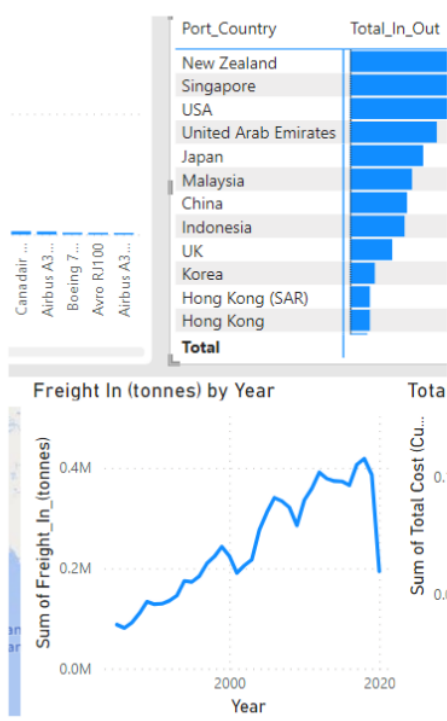


Figure 16: Normal sized view of the line chart without filtering

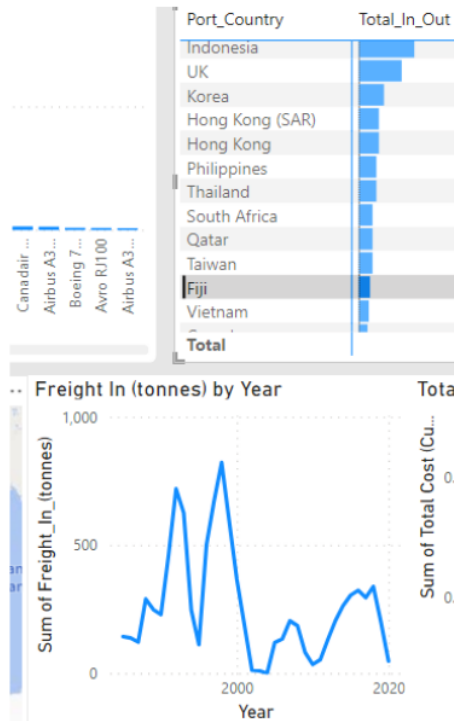


Figure 15: The same visualisation filtered to only display freight in information for Fiji

It is also possible to filter the visualisation as a whole by selecting a specific value on another visualisation. In this instance, Fiji has been selected on the matrix visualisation (created later) to only show freight in information for that country.



Figure 17: Card-type visualisation containing the sum of the total number of orders

A card visualisation was created to show the sum of the total number of orders at a glance. Its settings were changed so that it would have rounded corners and display a 30px rounded border.

Task 5 – Add a user-interactive filter

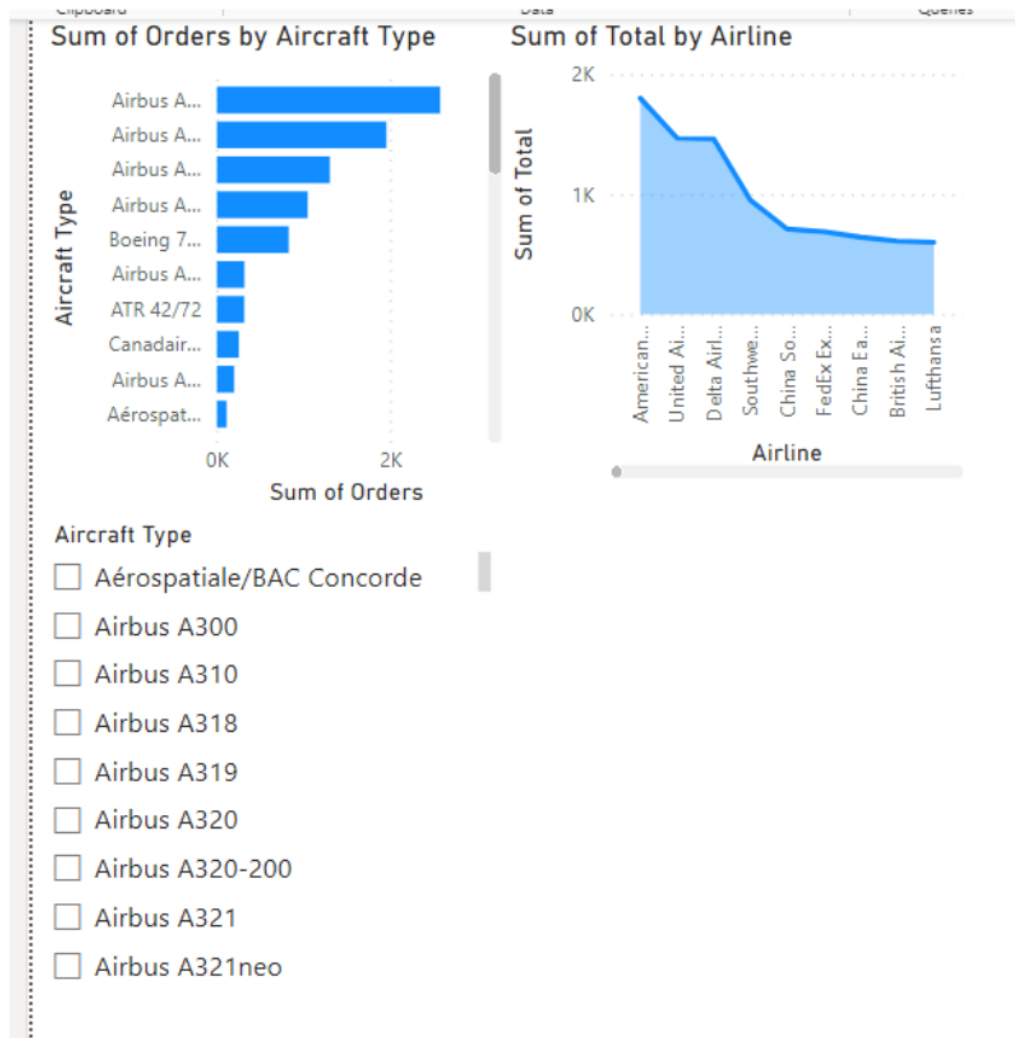


Figure 18: Early version of the visualisation with a filter

A slicer was added to allow the user to filter by aircraft type. While this can be done by clicking on any aircraft name in any visualisation, the slicer is an easier and more obvious way to filter specific values.

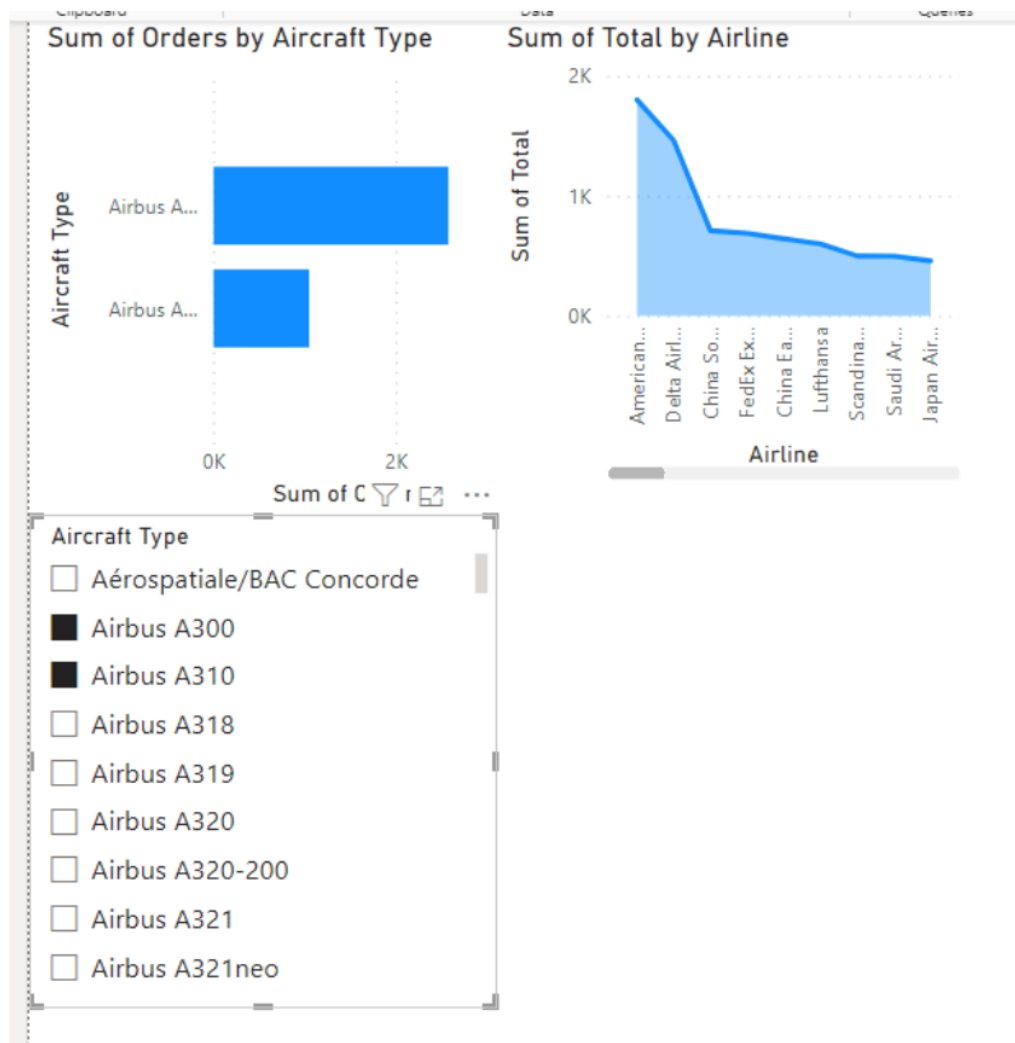


Figure 19: Explicit filters set to only show data for Airbus A300 and Airbus 310 aircraft

In this instance, the filter was set to display only values related to the Airbus A300 and Airbus A310 aircraft. This adjusts the values shown on the other visualisations to only utilise data related to the selected aircraft.

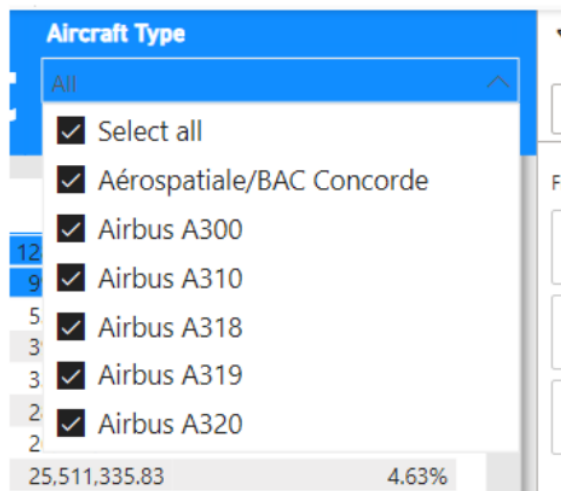


Figure 20: Current version of the filter which is a drop-down box and has a select all option

Finally, the visualisation was changed into a dropdown box and given a "select all" button which is active by default. It was also modified to use more conventional tick boxes; before doing so it required holding down the control key whilst selecting an aircraft type in order to select multiple filters.

Task 6 - Top 5 Airlines by total cost

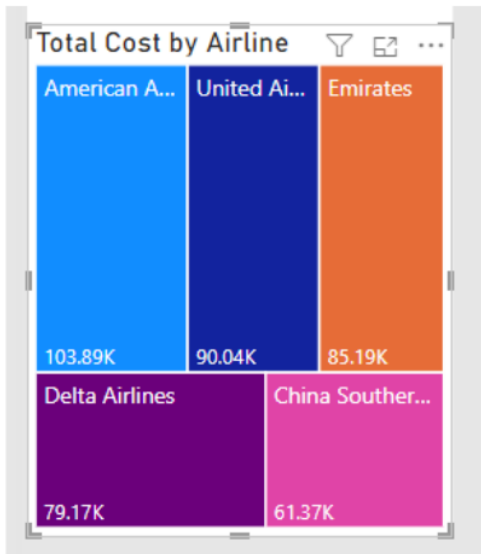


Figure 21: Small view of a treemap visualisation showing the top five most expensive airlines

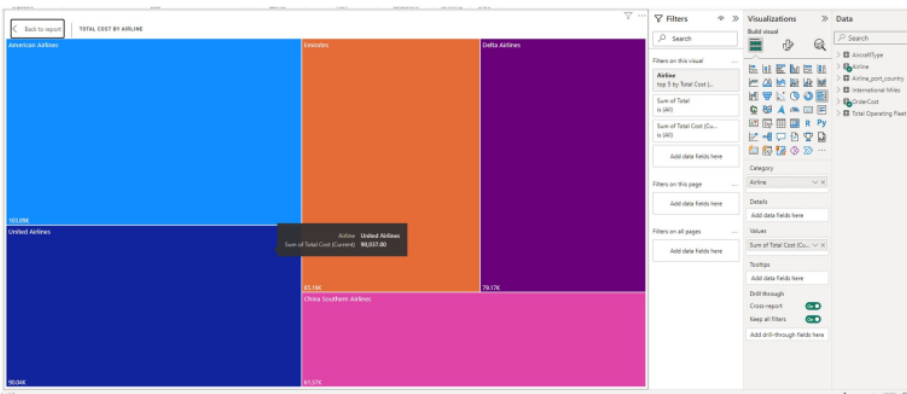


Figure 22: The above visualisation in fullscreen

A treemap visualisation was added to display only the top 5 airlines by total cost. This was done by adding a filter that would only return the top 5 values with the highest total cost. In order from the top to the bottom, the most expensive airlines were American Airlines, United Airlines, Emirates, Delta Airlines and China Southern Airlines.

Task 7 – Count of airlines in different countries

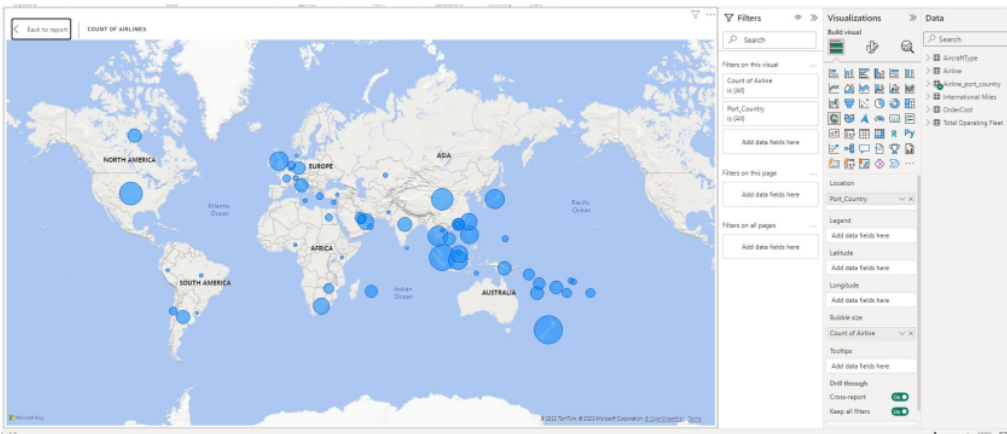


Figure 23: Map visualisation of count of airlines in different countries

This specific visualisation makes use of geographic data from the Port_Country value to display larger circles depending on which countries have more airlines in them. For security reasons, PowerBI cannot display map data by default so this feature had to be enabled in the settings.

The airline value had to be set to a count in order to get the exact number. The largest circle is New Zealand, which has 2533 airlines.

Task 8 – Analysing Busiest Port

```
1 Total_In_Out =
2 SUM('Airline_port_country'[Freight_In(tonnes)]) + SUM('Airline_port_country'[Freight_Out(tonnes)]) + SUM('Airline_port_country'[Mail_In(tonnes)]) + SUM('Airline_port_country'[Mail_Out(tonnes)]) + SUM
('Airline_port_country'[Passengers_In]) + SUM('Airline_port_country'[Passengers_Out])
```

Figure 24: DAX function to find the sum of all in and out of every port

To calculate the total in and out of every port, a calculated measure custom DAX function had to be created to add up every In and Out value.

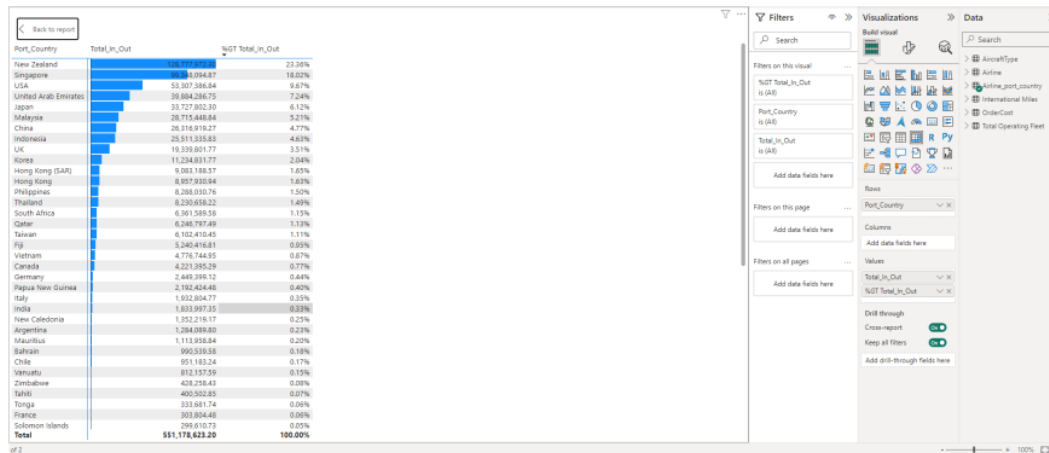


Figure 25: Full screen view of the matrix visualisation with the percentage grand total visible

Those values were then added to a matrix visualisation. The same value was added to the visualisation twice, but the second one was modified using a PowerBI setting to display as a percentage of grant total. The Total_In_Out value was also set to display a meter as well.

Task 9 – Analysing Freight Out



Figure 26: Function generated using PowerBI's quick measures option

Another DAX function was created to analyse the variance of freight out for every country. PowerBI includes a selection of common functions in the quick functions menu, so the quick function “variance by category” was used to generate this DAX function.

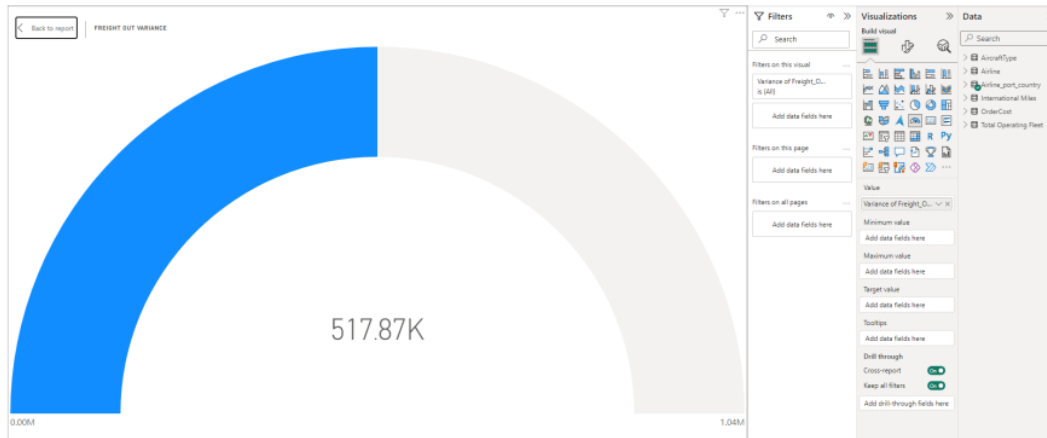


Figure 27: Full screen view of the gauge chart

The resulting variance function was added to a gauge chart visualisation. The title of that visualisation was changed to say Freight Out variance.

Task 10 - Formatting Report

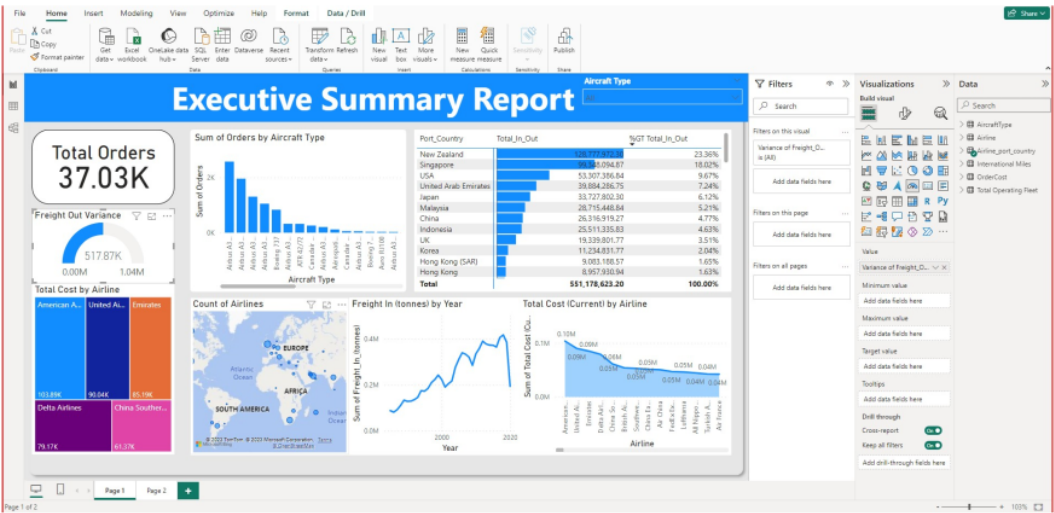


Figure 28: Fully completed visualisation with the title "Executive Summary Report"

The final step was to organise the visualisation. In this instance a title bar was added to the top of the page that reads "Executive Summary Report". The interactive filter was placed at the top of the page and can be used to filter the entire visualisation (see below). The map, line and stacked area charts were also grouped together in the visualisation.

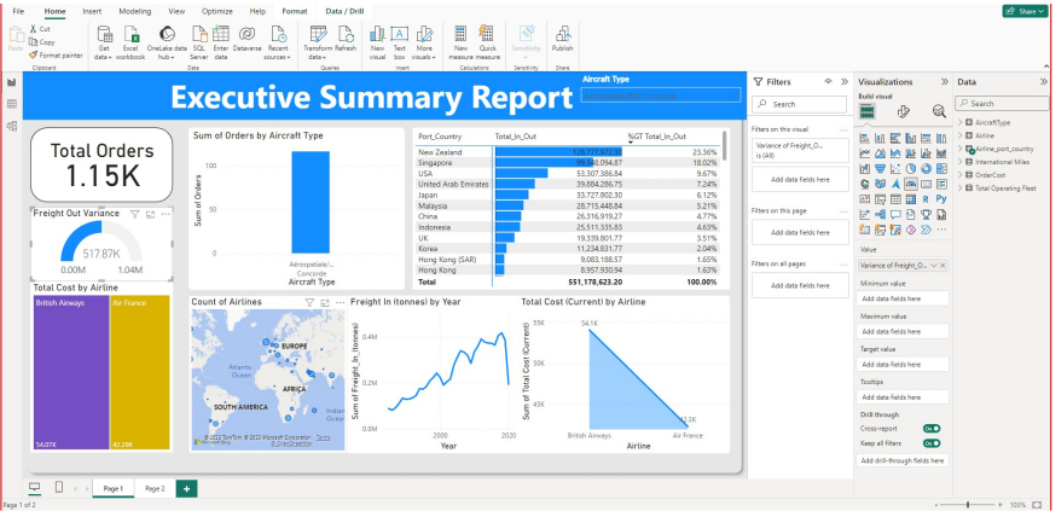


Figure 29: The above visualisation with a filter active

Conclusion

In conclusion, the completed data visualisation shows an at-a-glance view of the data related to aircraft sales and usage as well as a working filter to sort through specific aircraft types as necessary. This visualisation should hopefully make it easier for the board of executives to decide what to areas allocate the budget for the next fiscal year into.

There were some notable occurrences in the visualisation that could be used for analysis of the dataset, for example the global freight in value experienced a large drop during 2020 (presumably due to the COVID-19 pandemic) and that the Airbus A300 was the most commonly ordered aircraft.

Potential improvements to the visualisation could include a more explicit airline filter rather than just using the matrix chart as a filter.

Data Visualisation Assignment

GRADEMARK REPORT

FINAL GRADE

GENERAL COMMENTS

94/100

Task 1 :	10
Task 2 :	4
Task 3 :	5
Task 4 :	18
Task 5 :	10
Task 6 :	4
Task 7 :	5
Task 8 :	14
Task 9 :	14
Task 10 :	10

Hi Ben,

Good start.

A very clear introduction about the scenario explaining the need for a PowerBI dashboard

Here are the comments for your tasks:

Task 1:

All the necessary tables were imported to the Power BI application successfully.

Task 2:

Data preparation was done excellently making all the necessary changes before loading the data. All the screenshots after incorporating the necessary changes enrich the document however the screenshot after changing the first row as header is missing.

Task 3:

explanation of the relationships between the tables is commendable.

Task 4:

Good and clear visuals have been created. However, all the subsections could have been labelled for ease of evaluation. Analysis of stacked area charts focuses on highlighting the values related to American Airlines. Similar in-depth analyses could have been provided for the values derived from other charts as well.

Task 5 and 6:

step-by-step process of including the slicer and creating a tree map has been clearly and effectively presented.

Task 7:

The count of airlines was distinctly differentiated among countries. The screenshots taken to highlight the creation of the required visual in the spotlight greatly aided in better understanding.

Task 8 and 9:

Right DAX expression was created in a measure to be utilized in matrix visual. however, the final goal was to analyze the busiest airport, and textual information about the analysis could have been provided.

Task 10:

The report was very well formatted with clear differentiation between all the visuals

Overall, A well presented document with clear introduction and conclusion.

I hope these comments will be of help.

PAGE 1

PAGE 2

PAGE 3

PAGE 4

PAGE 5

PAGE 6

PAGE 7

PAGE 8

PAGE 9

PAGE 10

PAGE 11

PAGE 12

PAGE 13
