SAS-PROJECT

CIS-5250

Consumer Airfare Report -1996 to 2020

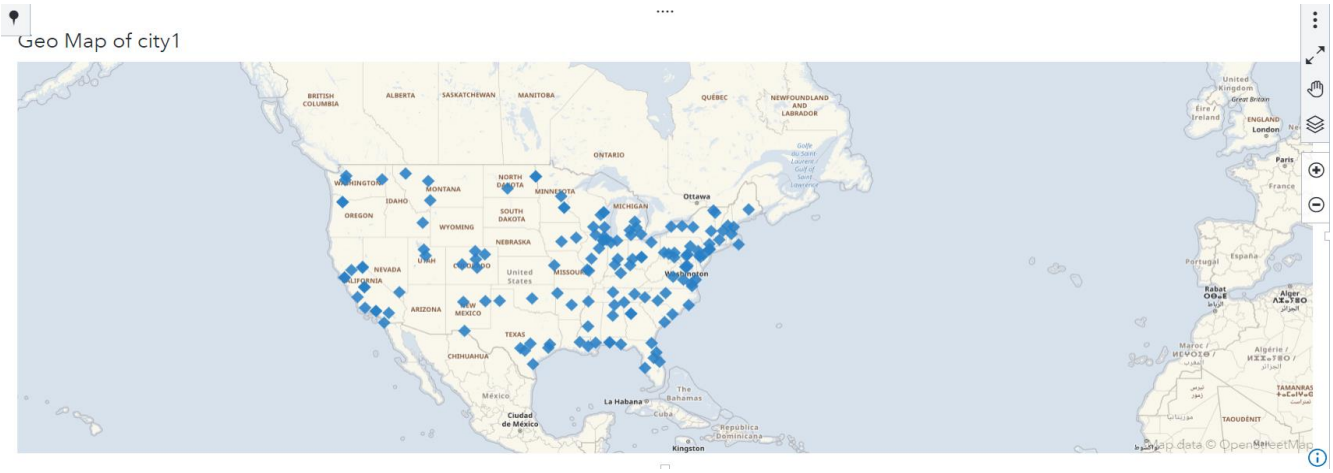
Submitted By: Pradnya Patil (401385189)

Section A:

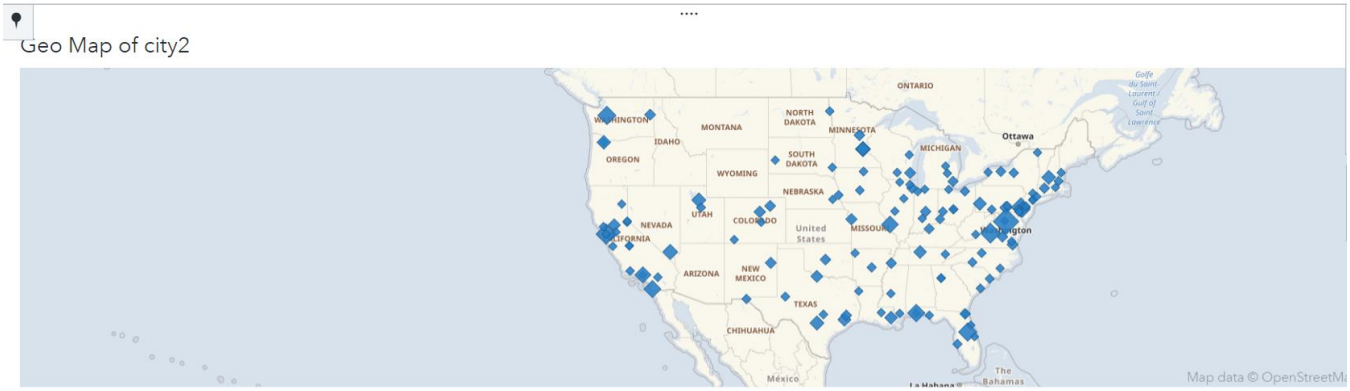
Introduction:

The aviation industry encompasses almost all aspects of air travel and the activities that help to facilitate it. The air travel consumer report is a yearly report which provides the information of Quarterly product of the Department of Transportation Office of Aviation Consumer Protection. The report provides detailed information of airfare range, number of passengers, mileage covered by air-carrier. The report is designed to assist consumers with information on the quality of services provided by the airlines. This is informative data that provides details of consumer airfare from the year 1996 to 2020 -four a quarterly basis. I found 'Domestic Airline Consumer Airfare Report' most informative as it provides details of changes in airfare range from the year 1996 to 2020. I will start the analysis by grouping the years like 1996 -2000, 2000-2005,2005-2010, 2010-2015 & 2015- 2020. For analysis, I have selected to and from destination cities. There will be two groups of cities created City1 & City2. Following are the details of Cities:

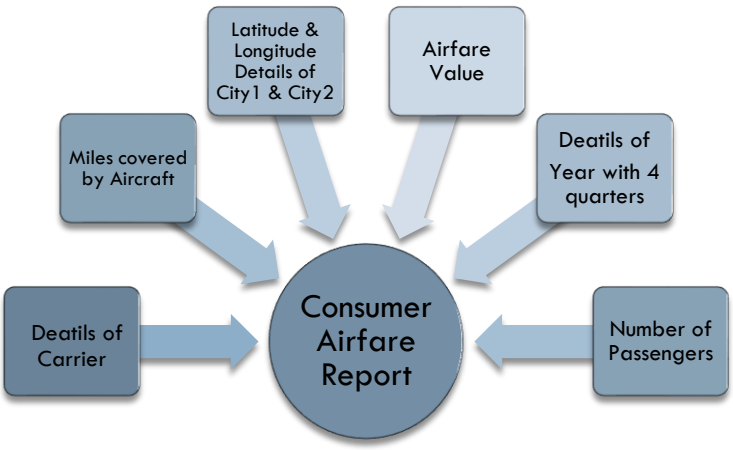
City 1: the latitude and longitude details are present in Consumer Airfare Report. Following are the cities covered under the city-1 group



City 2: The latitude and longitude details are present in Consumer Airfare Report. Following are the cities covered under the city-2 group



I will create a visual analysis by considering the important factors in the consumer airfare report:



The primary factor of analysis is the range of airfare and the number of passengers. The visualization will be created after a detailed analysis of the airfare range and number of passengers from the destination.

I am planning to analyze by using the following questions:

1. **What is the factor responsible for changes in airfare from 1996 to 2020? How COVID-19 situation affect airfare range?**
2. **What is the factor responsible for changes in the number of passengers from 1996 to 2020? What is the impact of COVID-19 on the aviation sector?**
3. **What is the average mileage covered by air-carrier?**
4. **Explain the range of airfare specifically fare_lg and fare_low value variation?**

5. **Which are top 5 air carrier with high airfare and milage covered?**

The analysis will include the categorization of years- Starting from 1996 to 2000 which is based on years, I will analyze the changes in airfare range and number of passengers traveling during this year. The analysis of the pre-pandemic situation from the year 2017 to 2019 and details of post-pandemic situation (the year 2019-2020) will be captured in the report. I am planning to include the impact of Covid -19 on the aviation sector in visual reports.

REFERENCE

(ACI), A. C. (2021, March 2021). *The impact of COVID-19 on the airport business and the path to recovery*. Retrieved from Advisory Bulletins: https://aci.aero/2021/03/25/the-impact-of-covid-19-on-the-airport-business-and-the-path-to-recovery/

Knaflic, C. N. (2015). *Storytelling with Data*. Hoboken, New Jersey: John Wiley & Sons, Inc.

Section B:

Data Description:

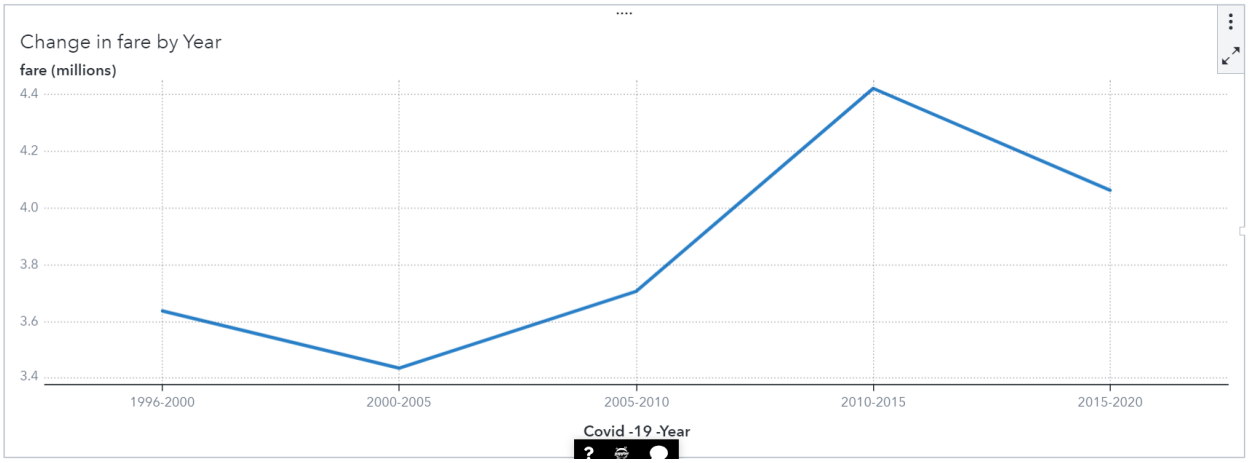
Field Name	Data Description	Data Type 1	Data Type 2
Year	Column represents the information of Year for which data is provided	double	N/A
Quarter	Data provided for four quarters	double	Quantitative (discrete data)
Citymarketid_1	Column represents the information of City 1 -makertid	double	Quantitative (discrete data)
Citymarketid_2	Column represents the information of City2-marketid	double	Quantitative (discrete data)
City1	Column represents the name of City1 -destination (from-destination)	char	Qualitative
City2	Column represents the name of City2- destination	char	Qualitative
nsmiles	Column represents the information of miles covered by aircraft	double	Quantitative (discrete data)
Passengers	Column represents the total number of passengers	char	Quantitative (discrete data)

Fare	Column represents the airfare value	double	Quantitative (discrete data)
Carriers	Column represents the name of carrier (aircraft)	char	Qualitative data
Large_ms	Columns represent the miles	double	Quantitative (discrete data)
Fare_lg	Column represents the high airfare range	double	Quantitative (discrete data)
City1_Latitude	Column represents the information of latitude of City1	double	Quantitative (Continuous Data)
City1_Longitude	Column represents the information of longitude of City1	double	Quantitative (Continuous Data)
City2_Latitude	Column represents the information of latitude of City2	double	Quantitative (Continuous Data)
City2_Longitude	Column represents the information of longitude of City2	double	Quantitative (Continuous Data)

Section C – Analysis & Visualization:

1. **What are factors responsible for changes in airfare ranging from 1996 to 2020? How COVID-19 situation affect airfare range?**

   Data Analysis: The analysis is generated from the information present in consumer airfare report from the year 1996 to 2020. The consumer airfare report provides the details of key factors which affect the changes in airfare range. As per the study of data from 1996 to 2020, the following are the observation estimated:

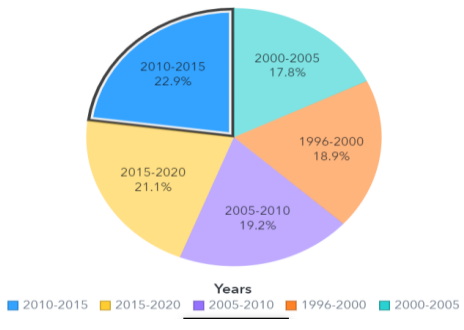


(Figure -a- Line Graph)

The above **line graph** provides details of changes in the rage of airfare from 1996 to 2020.

The following table represents the airfare value as per year:

Year Details	Airfare (Million)
1996-2000	36,37,494
2000-2005	34,36,066
2005-2010	37,06,649
2010-2015	44,20,190
2015-2020	40,61,678



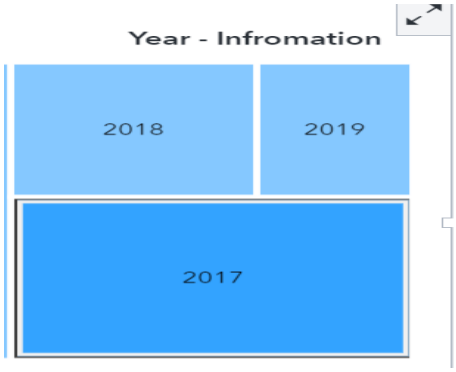
(Figure-b- Pie-Chart)

The airfare range data provided from 1996 to 2020, I created a group of 5 years to simplify the analysis of consumer airfare range. The major development and changes took place in the aviation sector in the year 1996 to 2000. hence the airfare range was high in these years. Due to innovation in the technology and aviation marketing field, the airfare range increased again in the year 2005 to 2010. Since then, the aviation sector is one of the leading sectors. In the year 2010 to 2015, the airfare range increase which leads to an increase in the profit rate. The year 2015-2020 analysis is conducted in two ways pre-pandemic analysis and pandemic analysis. The above **pie-chart (figure-b) represents** the percentage of increase in airfare value.

As per information provided in the table- the airfare range is highest between the year 2010-2015 (44,20,190 million) and lowest between the year 2000-2005 (34,36,066 million). As per reference of an Empirical Analysis of Airline Industry (NBER working Paper No. 14503). Changes in airfare range demand accounted for almost half of the 80% decline profit. In the year 2006, delays and the full flights had made passengers so adverse to connecting flights that adding a layover to a route could reduce the number of passengers on it by almost four-fifths. As a result, the average fare for connecting flights dropped by an estimated 12 percent, while the average fare of direct flights fell by only 4 percent. During this period, the average airline fare decreased by 4%. As per analysis, the average airfare rate decreased by 2% from 1996-2000 to 2000-2005. The airfare rate was slightly high in the year between 1996-2000 and reduced in 2000-2005.

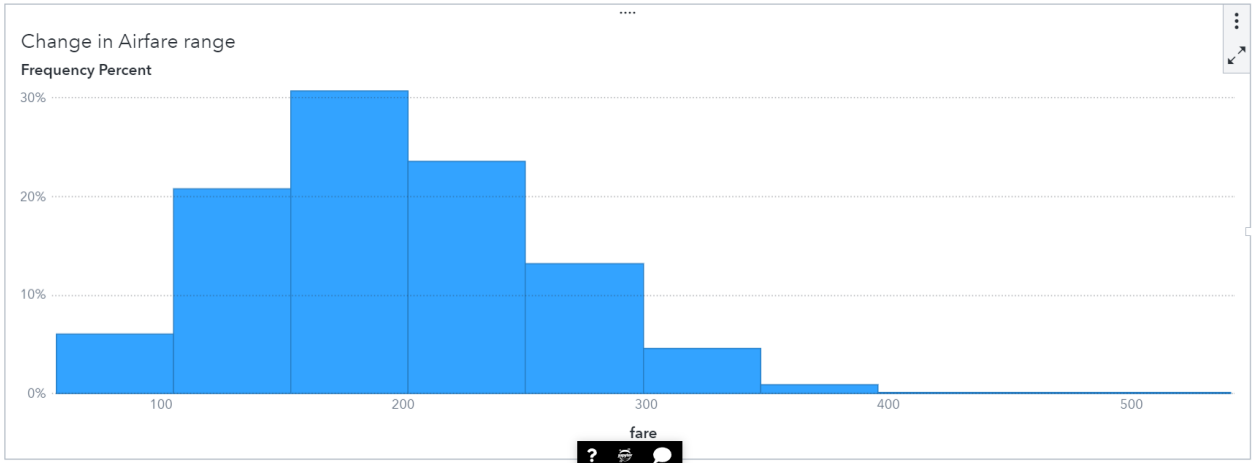
The following table represents airfare range from year 2017 to 2020 as per quarter:

Year	Q1	Q2	Q3	Q4
2017	3,99,054	2,23,099	2,12,185	2,15,942
2018	2,18,398	2,18,369	2,13,627	2,19,883
2019	2,19,027	2,22,753	2,16,737	2,20,458
2020	2,12,574	1,80,060	1,65,571	0



(Figure -c- Tree Map)

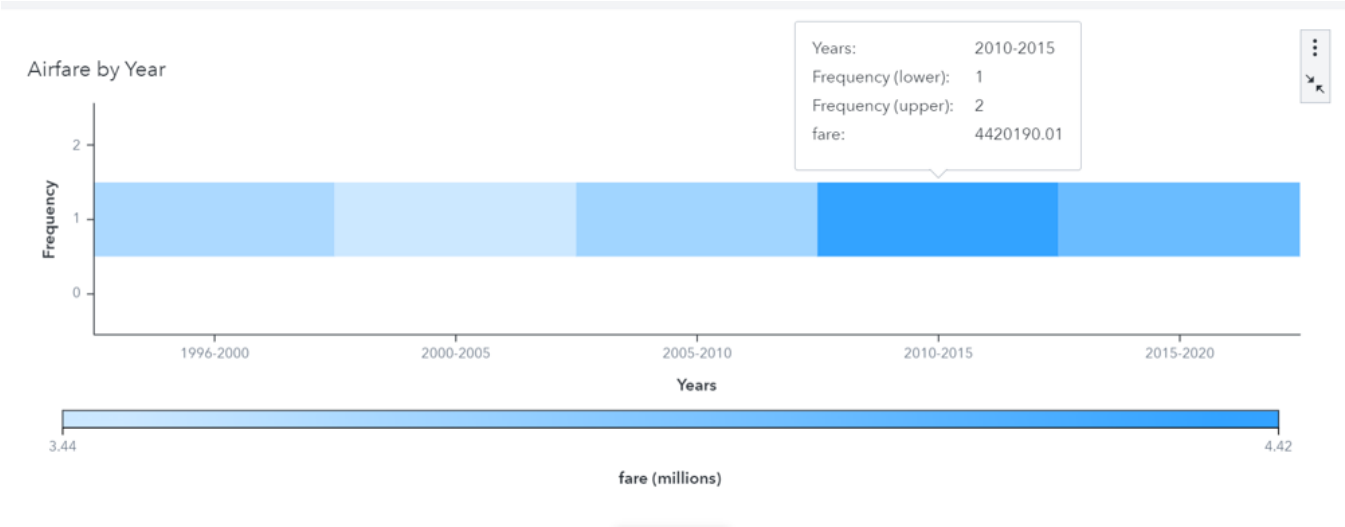
The frequency of percentage change in airfare rate is 30% and airfare change range is $200. The below histogram represents changes in airfare range.



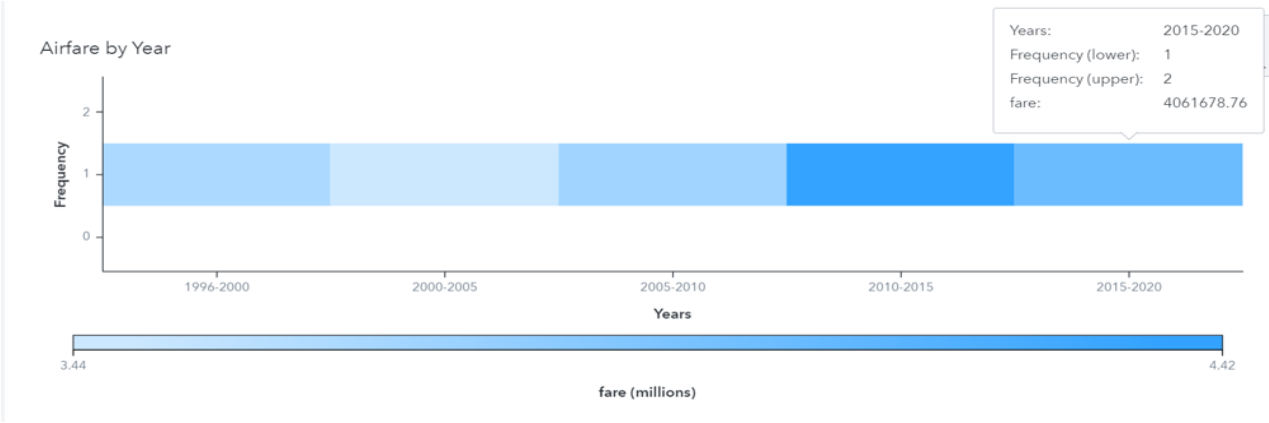
(Figure-d-Histogram Graph)

Conclusion: The COVID-19 situation leads to a global transportation crisis. The aviation sector was in high demanding till the year 2019. Due to the global lockdown situation, the airfare rate effected.

Following Heatmap represents airfare value:



(Figure-e-heatmap)



(Figure-f-heatmap)

Following is the range of change in airfare value:

Airfare by Year

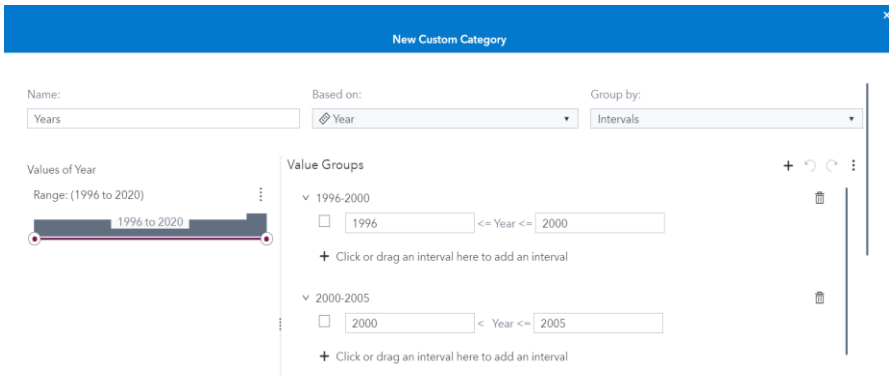
Years	Frequency (lower)	Frequency (upper)	fare
1996-2000	1	2	3637494.32
2000-2005	1	2	3436066.09
2005-2010	1	2	3706649.26
2010-2015	1	2	4420190.01
2015-2020	1	2	4061678.76

SAS Feature:

Following are the details of the graph used:

a. **Line graph**: I used a **line graph in figure-a to represent the changes in airfare value by year**. I used a line graph to represent the single series of data.

   - I created the new custom category –Year, which represents the group of five years (1996-2000).



   - Using a 5year group of data started my analysis of change in airfare rate. I found line graph is effective visualization to represent the change in airfare rate.
   - It is very easy to read the statistical analysis using line graph.

b. **Pie-Chart**: I used pie-chart (Figure -b) to represent the percentage changes in airfare value for group of years. Pie-chart is effective visualization to represent the exact statical notification. It is easy to understand the viewer and one can easily identify the percentage difference using pie-chart. I decide to use the SAS feature – Pie-Chart to represent the percentage change of airfare.

c. **TreeMap:** TreeMap is used to display hierarchical data using nested figures. I used treemap (figure-c) to represent the changes in airfare rate in specific years 2017 to 2020. I found hierarchical structure is effective visualization for representation airfare value from the year 2017 to 2020. The color combination of blue from dark to light represents the highest airfare range.

d. **Histogram:** Histogram is generally used for approximate representation of the distribution of numerical data. I used a histogram (figure-d) to represent the exact range of change in airfare value. I found every year range of change in airfare is nearly $200. ((ACI), 2021)

e. **Heat Map:** I used a heat map (figure- e & f) to represent the range of frequency changes in airfare value. The blue color shade- from dark to light represents the highest frequency to low- frequency value.

REFERENCES

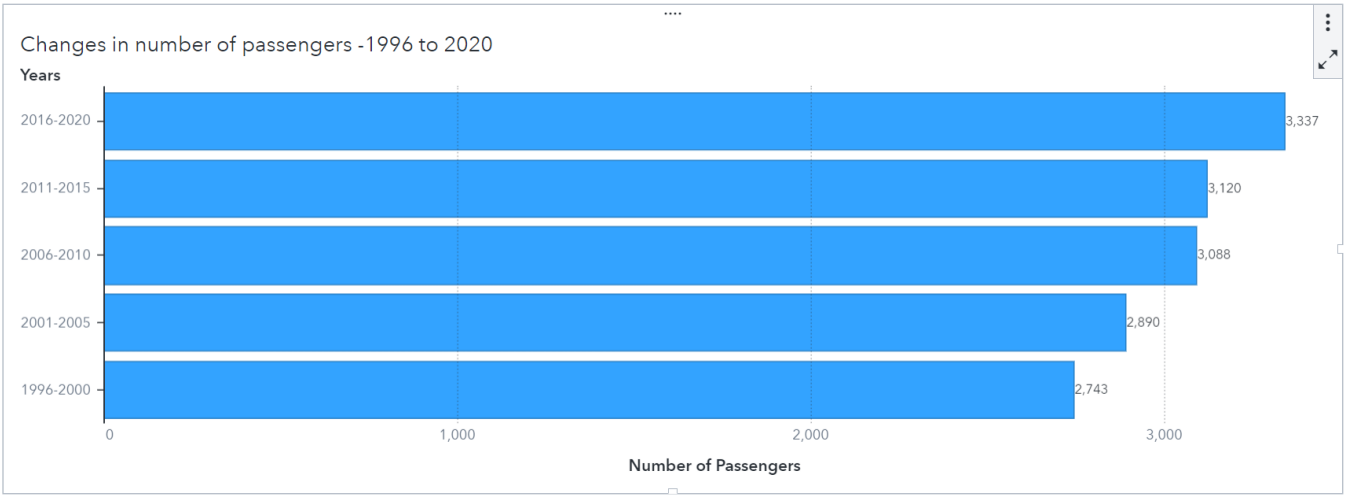
(ACI), A. C. (2021, March 2021). *The impact of COVID-19 on the airport business and the path to recovery*. Retrieved from Advisory Bulletins: https://aci.aero/2021/03/25/the-impact-of-covid-19-on-the-airport-business-and-the-path-to-recovery/

Knaflic, C. N. (2015). *Storytelling with Data*. Hoboken, New Jersey: John Wiley & Sons, Inc.

U.S. Department of Transportation. (2020, October 27). *Domestic Airline Consumer Airfare*. Retrieved from Transportation: https://www.transportation.gov/policy/aviation-policy/domestic-airline-consumer-airfare-report

Q2. What is the factor responsible for changes in the number of passengers from 1996 to 2020? What is the impact of COVID-19 on the aviation sector?

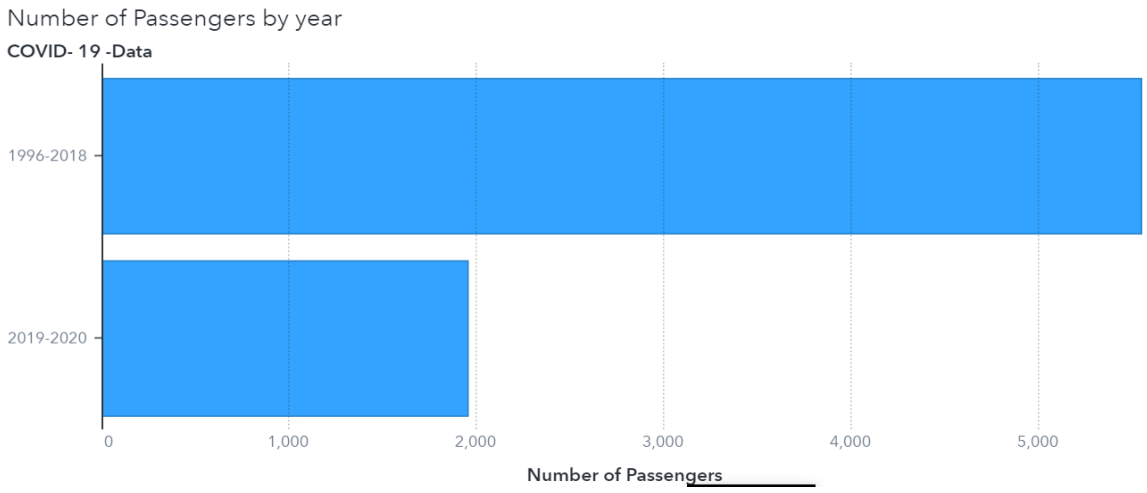
Data Analysis: The number of passengers is one of the important factors in consumer airfare report. The profit ratio will be calculated on basis of airfare range and number of passengers for that year. As per analysis, I came across the fact that the airfare range is inversely proportional to number of passengers. If the airfare value is low, then the number of passenger range will be high and if airfare range is high then number of passengers will below.



(Figure-g-Horizontal Bar Graph)

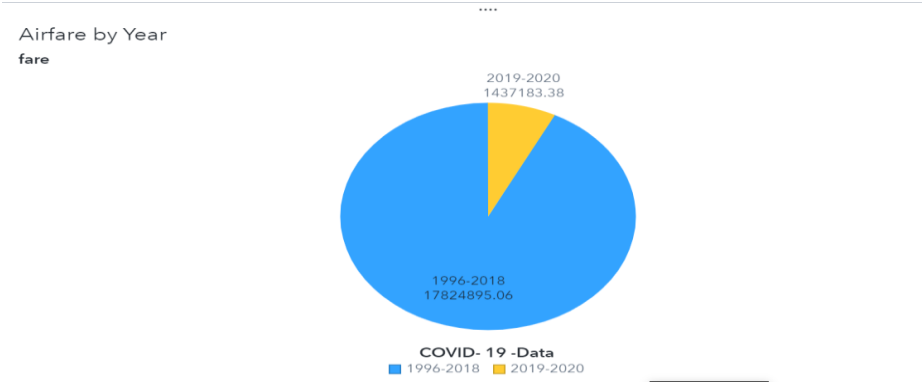
For the year 1996 to 2000, the airfare rate was high hence this directly impacted on the number of passengers. The aviation sector was growing from year 199 to 2000 hence the price range was high. Few numbers of passengers select airway travel due to high range of airfare.

For year 2001 to 2005, the airfare rate was slightly low compared to year 1996-2000, this leads to an increase in the number of passengers. For the year 2006 to 2010, leads to major changes in the aviation sector. Advance innovation and developed technologies lead to the introduction of new carriers and routes leads to reduce the time of traveling. The majority number of passengers preferred air transportation mode.



(Figure-h- Horizontal Bar Graph)

For year 2005 to 2018, there is a constant number of increases in passenger range and value of airfare. Year 2019 to 2020 leads to major changes in the aviation sector. Due to COVID-19 (pandemic situation), there are major changes in the number of passengers and range of airfare. Lockdown directly impacts on the number of passengers. The number of passengers gets decrease from quarter 3 of 2019 and the number came to 0 at quarter 4 of the year 2020. This pandemic situation directly impacts on a range of airfare and the number of passengers.



(Figure-i-Pie-Chart)

SAS Feature:

Following SAS Features are used for visual analysis:

g. **Horizontal Bar graph:** I used a horizontal bar graph (figure g, h) to represent the changes in the number of passenger values by year. I wanted to represent the number of passengers as per year, I consider the year group as a specific category hence to represent the categorical data, I preferred to use a Horizontal bar chart for visualization.

- In figure-h, I used horizontal bar graph to represent the number of passenger count of pre-pandemic (year 1996 to 2018) and pandemic year 2019 to 2020.
- The horizontal bar chart is easy for visualization and is used to define the categorical data.

i.Pie -Chart: Pie chart is visualization is used to represent the impact of COVID-19 situation on the crisis of global transportation. I select the data label as the number of passengers to represent count with year.

- For the representation of pandemic data, I have created a custom category as Covid-19 year and formed a specific group of the year 2019 & 2020.

REFERENCES

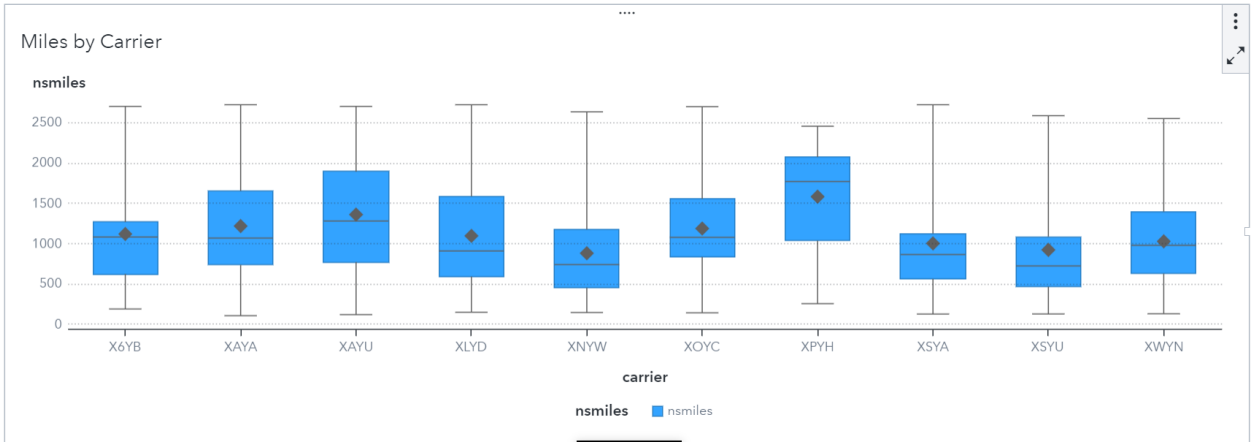
(ACI), A. C. (2021, March 2021). *The impact of COVID-19 on the airport business and the path to recovery*. Retrieved from Advisory Bulletins: https://aci.aero/2021/03/25/the-impact-of-covid-19-on-the-airport-business-and-the-path-to-recovery/

Knaflic, C. N. (2015). *Storytelling with Data*. Hoboken, New Jersey: John Wiley & Sons, Inc.

U.S. Department of Transportation. (2020, October 27). *Domestic Airline Consumer Airfare*. Retrieved from Transportation: https://www.transportation.gov/policy/aviation-policy/domestic-airline-consumer-airfare-report

Q3. What is the average mileage covered by air-carrier?

Analysis: The average mileage covered by the top 10 carriers are mentioned in the below box plot graph. The airline miles are known as frequent flyer miles. Carrier -X6YB, XAYA, XAYU, XLYD, XNYD, XOYC, XPYH, XSYU, XWYN are the top 10 carriers with the highest mileage range.



(Figure -j-Box plot)

As per information from U.S air traffic passengers-miles from 2007 to 2020, following are the results:

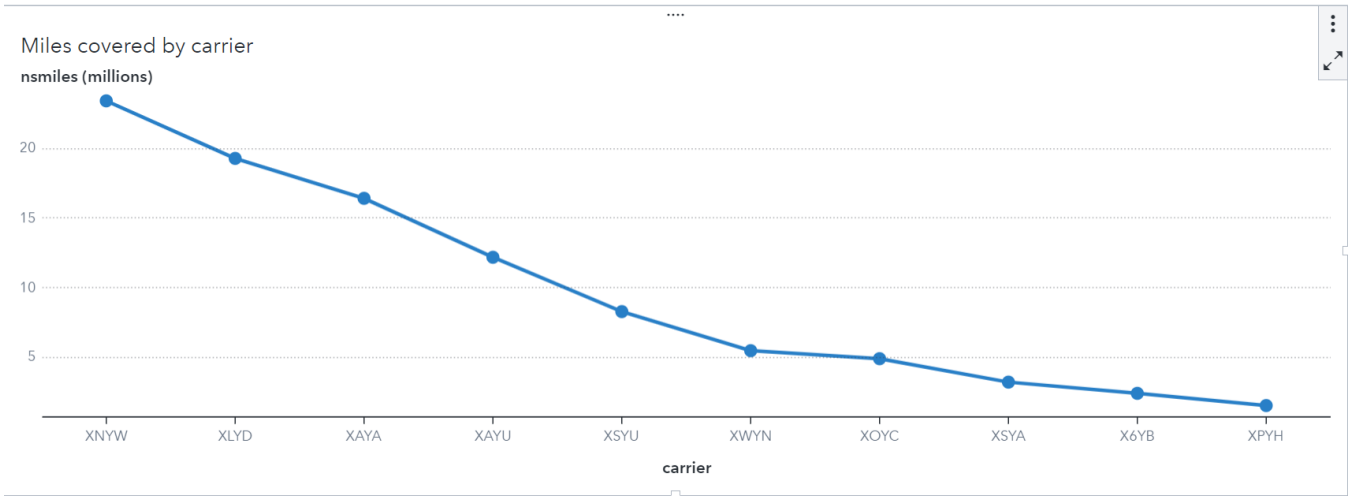
Years	Miles in Billions
2007	600
2008	580
2009	550
2010	570
2011	590
2012	580
2013	590
2014	600

2015	630
2016	650
2017	693
2018	730
2019	763
2020	306

For the year 2008 to 2013, the miles covered by the carrier is almost the same. The miles rate increased from the year 2015. **The year 2019 has the highest rate of miles but in the year 2020, due to corona virus pandemic, the distance covered is only 306 miles**. Following are details of the stop 10 carrier with the highest miles covered.

Miles by Carrier

carrier	Minimum	Lower Whisker	First Quartile	Average	Median	Third Quartile	Upper Whisker	Maximum	Std Dev	Count
X6YB	192	192	625	1121.6510976	1084	1271	2704	2704	672.09726167	2,141
XAYA	109	109	737	1220.1539605	1071	1657	2724	2724	622.3138737	13,458
XAYU	122	122	773	1362.4965332	1282	1894	2704	2704	706.51171483	8,942
XLYD	152	152	594	1098.470049	911	1589	2724	2724	645.03997211	17,562
XNYW	148	148	453	883.18587907	743	1180	2636	2636	538.6289638	26,528
XOYC	145	145	836	1189.7451982	1080	1559	2700	2700	546.362734	4,113
XPYH	258	258	1037	1584.4106029	1772	2075	2458	2458	579.79100689	962
XSYA	129	129	564	1004.9696115	867	1128.5	2724	2724	656.61782854	3,192
XSYU	130	130	468	924.99396581	726	1089	2588	2588	625.46038722	8,949
XWYN	133	133	629	1030.489919	981	1399	2553	2553	504.64302397	5,307



(Figure- k-Line graph)

SAS Feature:

Following SAS features are used for Visual Analysis:

j. **Box Plot**: I used box plot visualization (figure-j) to represent the miles covered by carrier. I found box plot visualization is an effective way to graphically depict a group of numerical data through their quartiles. I wanted to conduct an analysis basis on the carrier which covered the highest milage, I used rank functionality to filter the top 10 carrier with the highest milage.

k. **Line graph**: I used line graph(figure-k) for representation of top 10 carrier covered highest milage. I found line graph is effective visualization and one easily read the visual representation.

REFERENCES

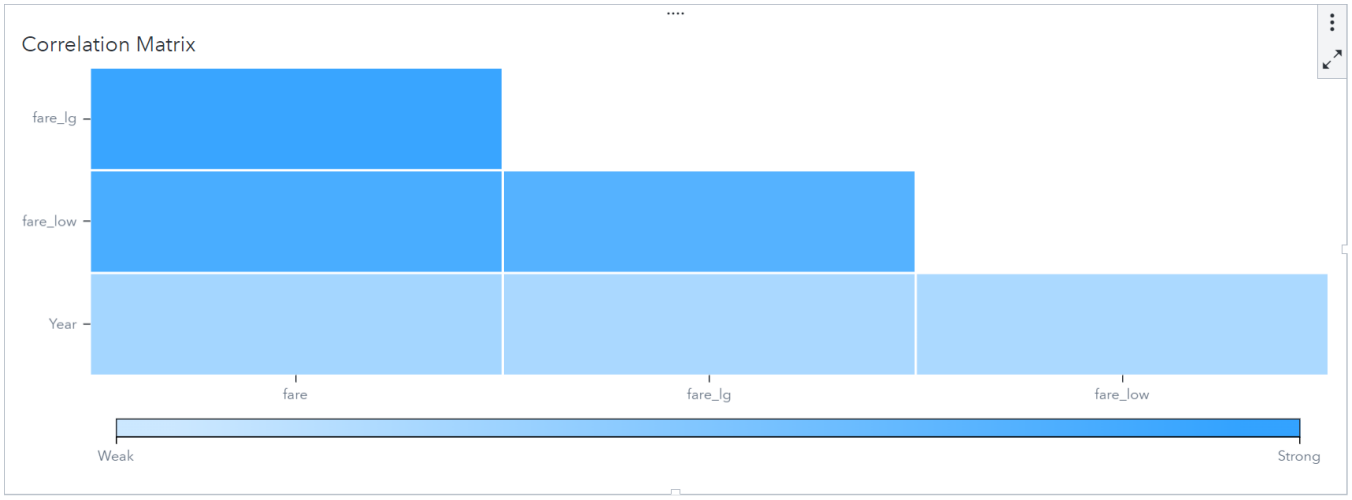
(ACI), A. C. (2021, March 2021). *The impact of COVID-19 on the airport business and the path to recovery*. Retrieved from Advisory Bulletins: https://aci.aero/2021/03/25/the-impact-of-covid-19-on-the-airport-business-and-the-path-to-recovery/

Knaflic, C. N. (2015). *Storytelling with Data*. Hoboken, New Jersey: John Wiley & Sons, Inc.

U.S. Department of Transportation. (2020, October 27). *Domestic Airline Consumer Airfare*. Retrieved from Transportation: https://www.transportation.gov/policy/aviation-policy/domestic-airline-consumer-airfare-report

Q4. Explain the range of airfare specifically fare_lg and fare_low value variation?

Data Analysis: The analysis is conducted from the data provided in the consumer airfare report. The below **correlation matrix** uses to represent the correlation between the normal airfare value, large airfare value and low airfare value for a specific year. This **measure is best used in a variable that demonstrates a linear relationship between factors**. This matrix depicts the correlation between all the possible pairs of values in the table. Here in this data visualization – correlation between range of airfare is represented. The actual airfare for a specific quarter and variation of airfare in the different quarter in the same year is specified by categorizing the high fare range and low fare range.



(Figure-l-Correlation Matrix)

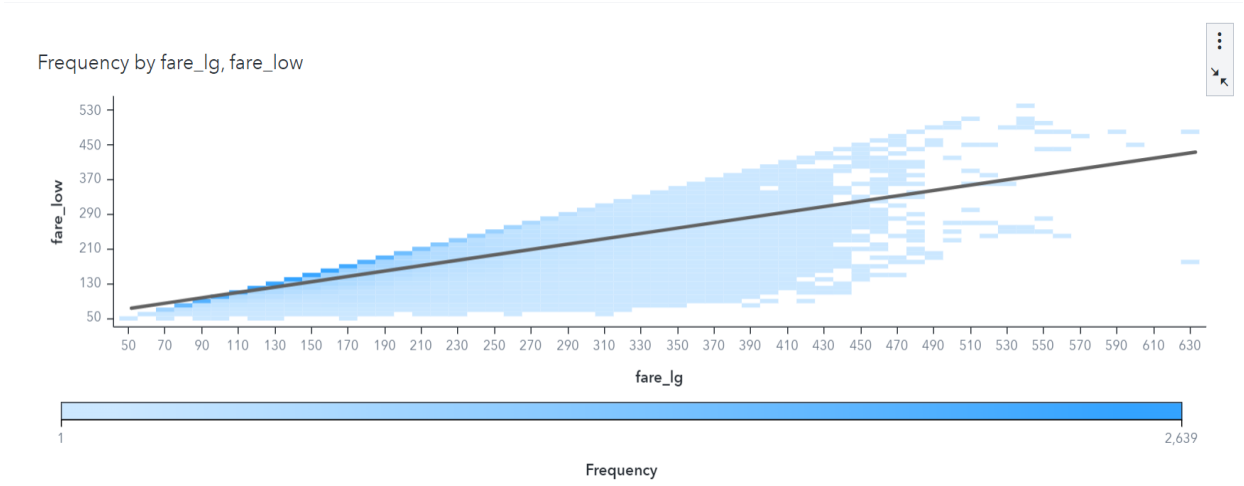
Following are the details of correlations:

Correlation Matrix

X Axis	Y Axis	Correlation
fare	Year	0.2583
fare	fare_low	0.8534
fare	fare_lg	0.9587
fare_lg	Year	0.2190
fare_lg	fare_low	0.7777
fare_low	Year	0.2138

Following forecast graph represents the frequency of changes in the range of airfare:

This forecasting graph represents the statistical analysis of changes in airfare values.



(Figure-m-forecasting graph)

Following are details of frequency of changes in airfare value:

fare_lg, fare_low | fare_lg, fare_low analysis

fare_lg (lower)	fare_lg (upper)	fare_low (lower)	fare_low (upper)	Frequency
45	55	45	55	10
65	75	45	55	2
85	95	45	55	1
95	105	45	55	2
115	125	45	55	2
125	135	45	55	1
165	175	45	55	1
55	65	55	65	121
65	75	55	65	26
75	85	55	65	28
85	95	55	65	32
95	105	55	65	11
105	115	55	65	9
115	125	55	65	14
125	135	55	65	4

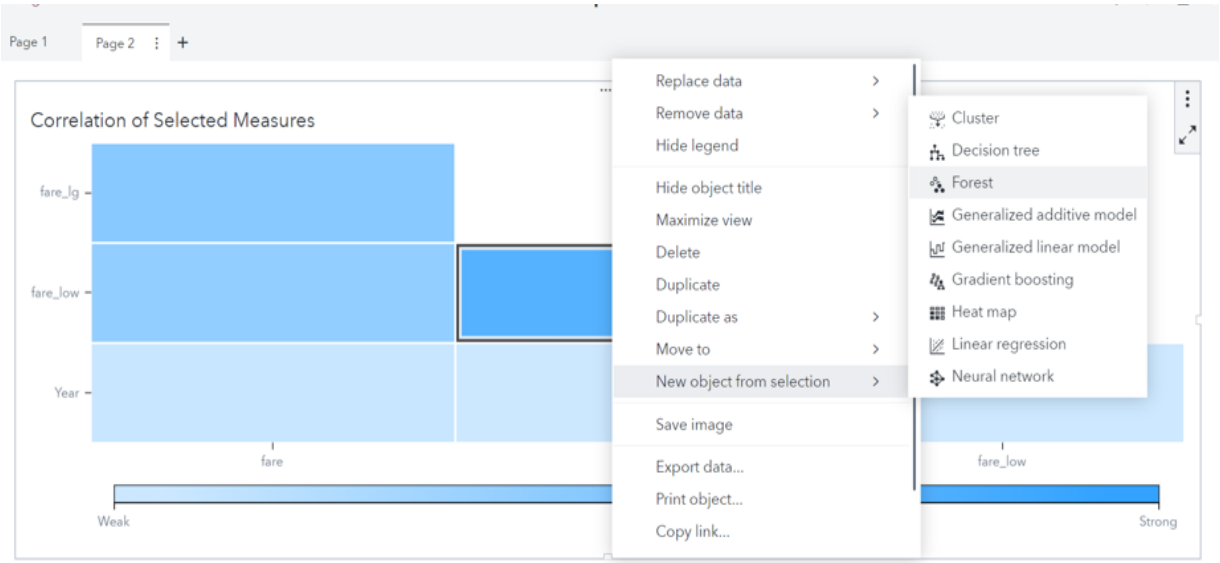
SAS Features:

Following SAS feature used for Visual Analysis:

l. **Correlation Matrix**: **Correlation matrix (figure-l) is use for represents the correlation of variables. I decided to use correlation matrix for analyzing the changes in airfare values**. I started analyzing exact changes of airfare range by comparing highest fare range and lowest fare

range in specific year. **I used SAS correlation matrix function and using new object selection- I have created foresting graph for analyzing the frequency change in airfare rate**.

Following is screenshot for reference:



REFERENCES

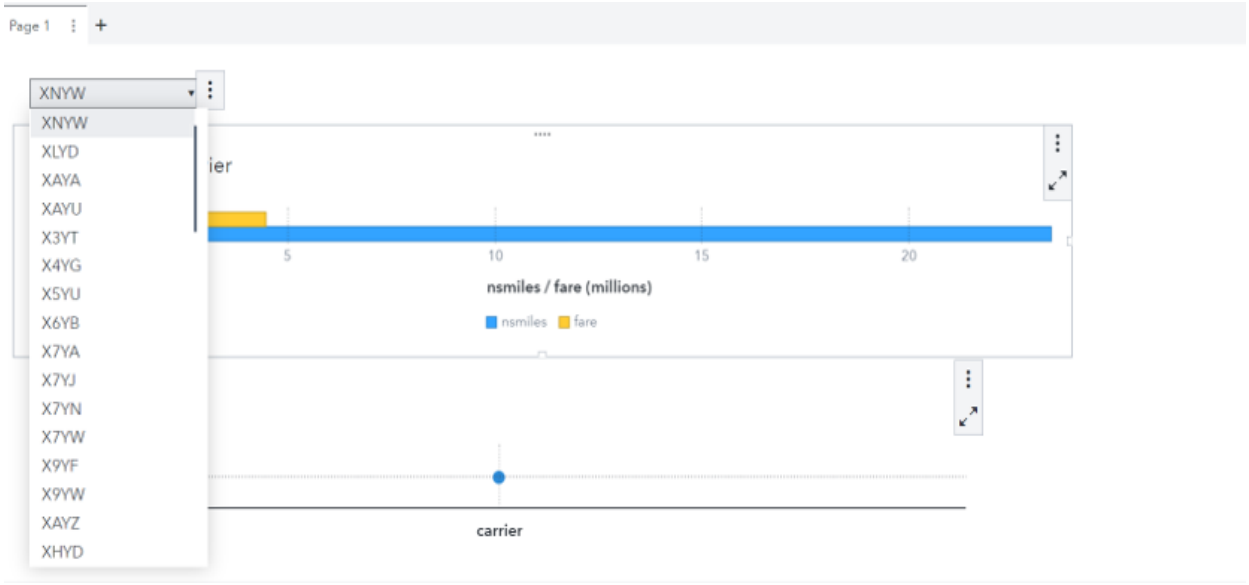
(ACI), A. C. (2021, March 2021). *The impact of COVID-19 on the airport business and the path to recovery*. Retrieved from Advisory Bulletins: https://aci.aero/2021/03/25/the-impact-of-covid-19-on-the-airport-business-and-the-path-to-recovery/

Knaflic, C. N. (2015). *Storytelling with Data*. Hoboken, New Jersey: John Wiley & Sons, Inc.

U.S. Department of Transportation. (2020, October 27). *Domestic Airline Consumer Airfare*. Retrieved from Transportation: https://www.transportation.gov/policy/aviation-policy/domestic-airline-consumer-airfare-report

Q5. Which are top 5 air carrier with high airfare and milage covered?

Data Analysis: The analysis is conducted by using the details of consumer airfare report. Analyzing the details of carrier visual representation is created. **The dropdown value feature is used for selection of carriers.**



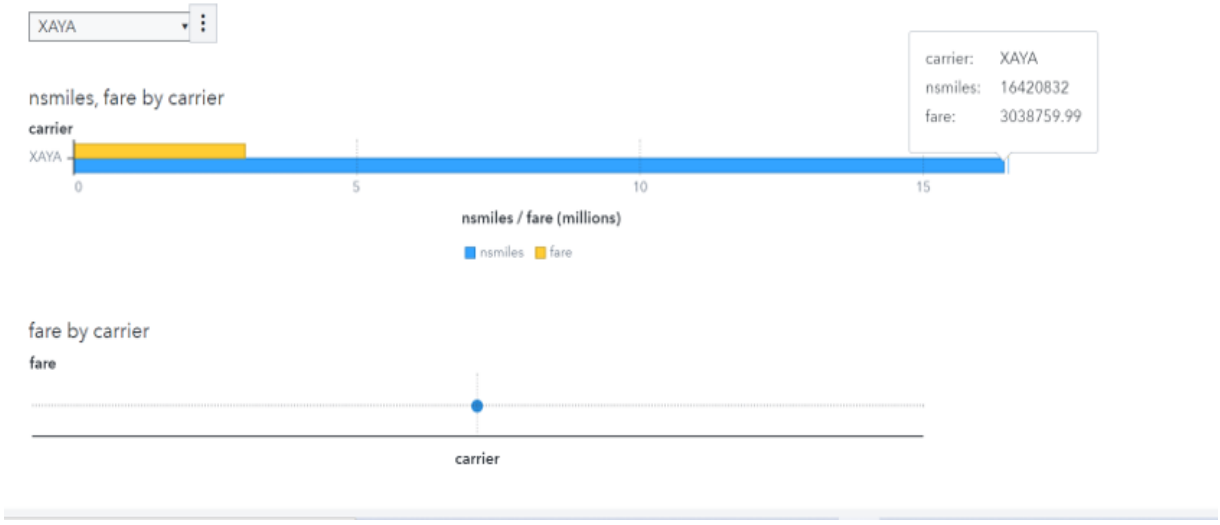
(Figure-n-dropdown feature)

Following is list of top 5 carrier with high airfare range and milage covered:

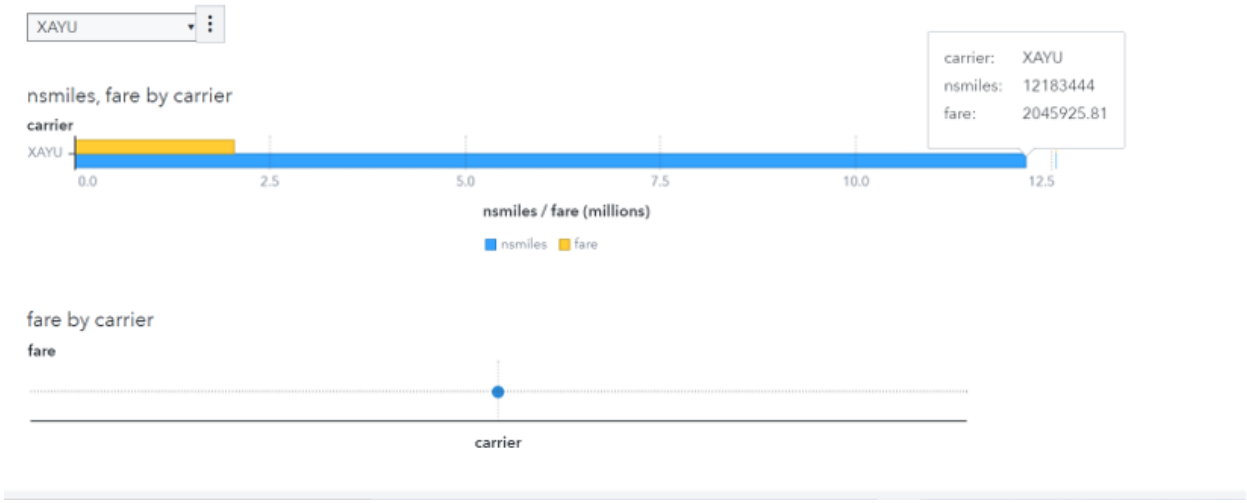
1. Carrier XNYW - As per analysis, the carrier XNYW covered highest mile – 23429155 and having good range of airfare.

2.Carrier XLYD: The carrier XLYD is second highest miles covered carrier and having good range of airfare value.

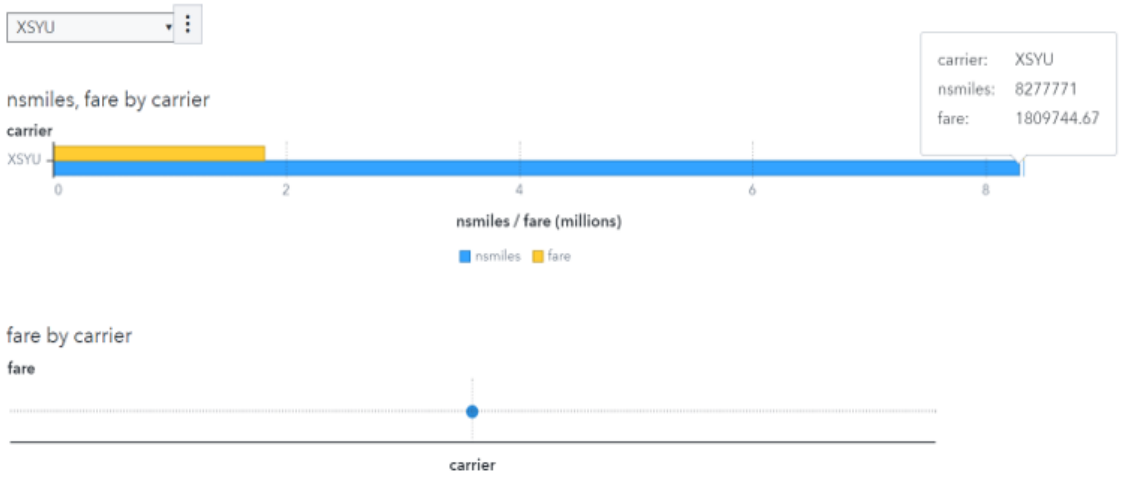
3.Carrier XAYA: The carrier XAYA – is third most popular carrier which covered highest miles.



4.Carrier XAYU: The milage covered by carrier XAYU is 12183444. It is forth highest milage covered carrier.



5.Carrier XSYU: The milage covered by carrier XSYU is 8277771. It is top 5th milage covered carrier.



SAS Feature:

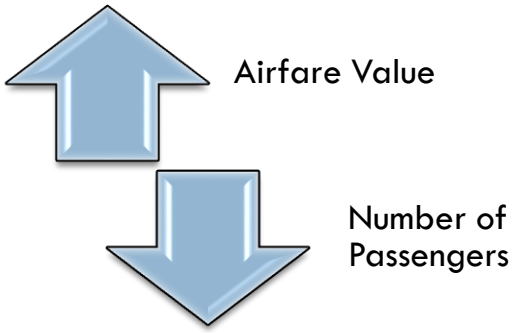
Following are SAS functionality used for Visual Analysis: **I preferred to use dropdown functionality for data analysis. The carrier's name values are used as dropdown value and**

after selection of specific carrier. I use dot plot and horizontal bar graph representation for mile covered and air fare value.

Section D:

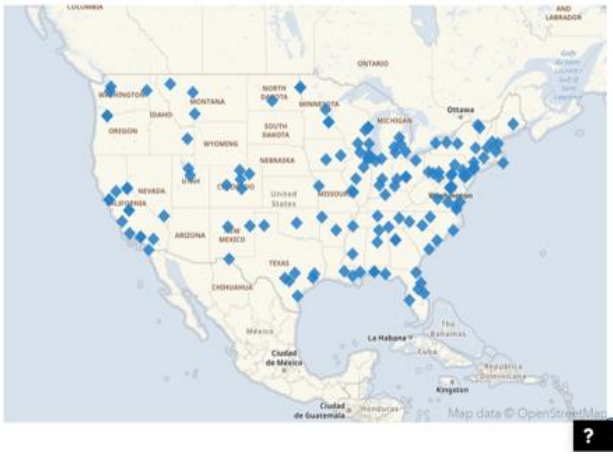
Dashboard & Storytelling:

The aviation industry is one of the well-known industries, and aviation marketing has some things in common with marketing in other industries. The target market can coincide with the luxury market, and the markets for high-end products and services. Aviation attracts people who love science, machinery and flying. The growth in the aviation business leads to establishing new routes and serving more passengers by procuring new aircraft, thereby adding more seats. **Many brands understand the importance of using storytelling in their aviation marketing strategy to connect with their customer on a deeper level**. The aviation sector is one of the top leading sectors hence the main factors are airfare, passengers and service provided in the carrier. **I think the airfare range is inversely proportional to the number of passengers. As the airfare rate goes high the number of passengers will reduce.**

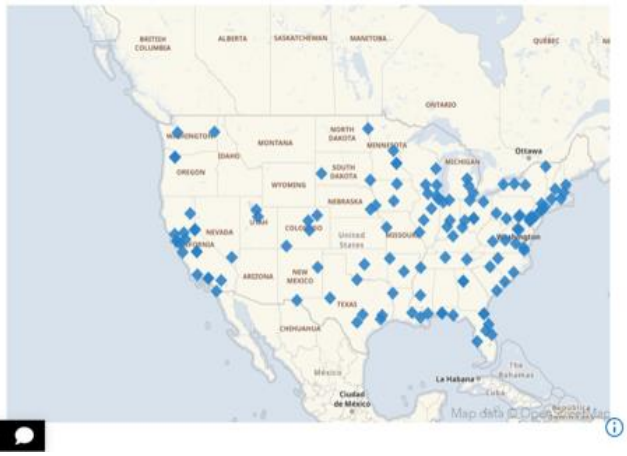


To provide the details information aviation market, the consumer airfare report provides information of four quarters in every year. **The four-quarter data from 1996 to 2020 are included in the report.** The data represents details of airfare value for specific travel destinations and the total number of passengers are traveled. For detailed analysis, data of domestic airlines are considered. The data of the following cities are provided.

Geo Map of city1

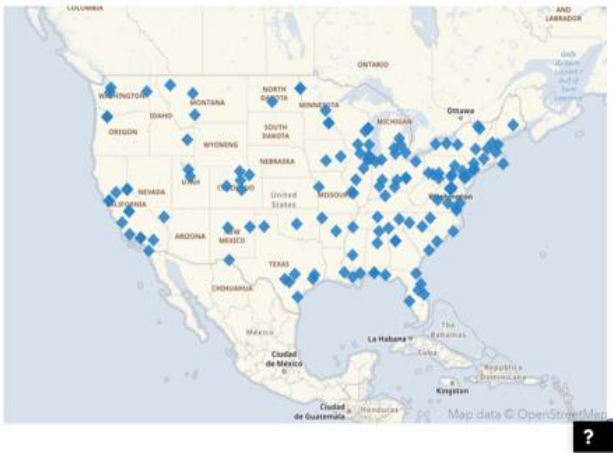


Geo Map of City 2

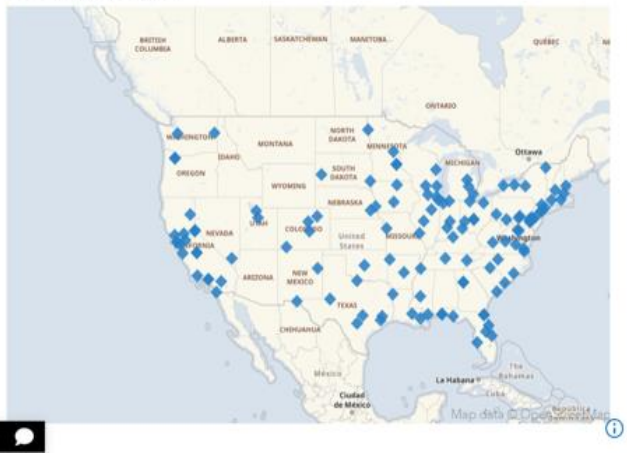


The Consumer Airfare report provides detail information on global transportation in the aviation sector. **The story narrates the information of a few air carriers**. This report is specifically for the domestic aviation sector. The data of the report is extracted a quarterly basis each year. The four-quarter data from 1996 to 2020 are included in the report. The data represents details of airfare value for specific travel destinations and the total number of passengers are traveled during that time. **The COVID-19 situation leads global transportation crisis which directly impacts the aviation sector.** The year 2019 to 2020 leads to a decrease in domestic airfare rate and the number of passengers. The evolution from 1996 to 2020 leads to many changes in the aviation sectors. Following is the Geo Map for two destination cities –

Geo Map of city1



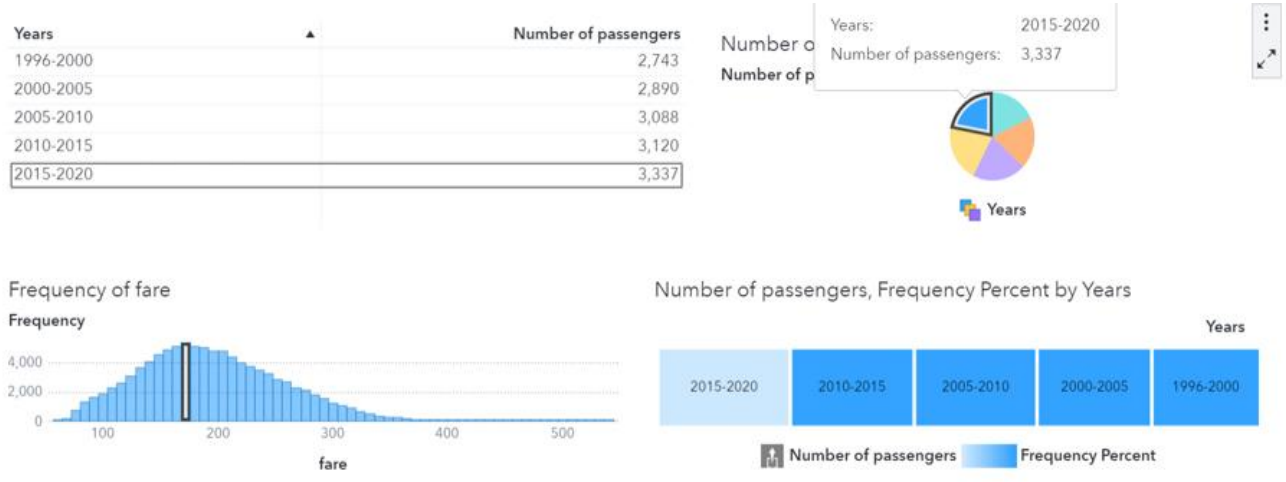
Geo Map of City 2



I came across the analysis, that from 1996 to 2000 prices that passengers paid for tickets declined slightly, even before adjusting for inflammation. The number of passengers is low because of the higher ticket range.

The year 2015 leads to great changes in airfare rate and the number of passengers. The profitability index was high during the year 2015 to 2019. The COVID-19 situation leads to global transportation crisis. The pandemic situation directly impacts on number of passenger range. Because of the lockdown situation, citizens are restricted to travel the domestic regions and international places. This leads to major loss to the aviation sector. The pre-pandemic situation in the year 2019- leads to slowly reduce the number of passengers but the airfare was high because of the limited number of transportations. Many countries close borders to avoid infection of COVID-19. **Third quarter in the year 2020 leads to major loss in all business sectors**. The following dashboard represents the how changes in number of passengers directly impact on airfare range.

Dashboard:

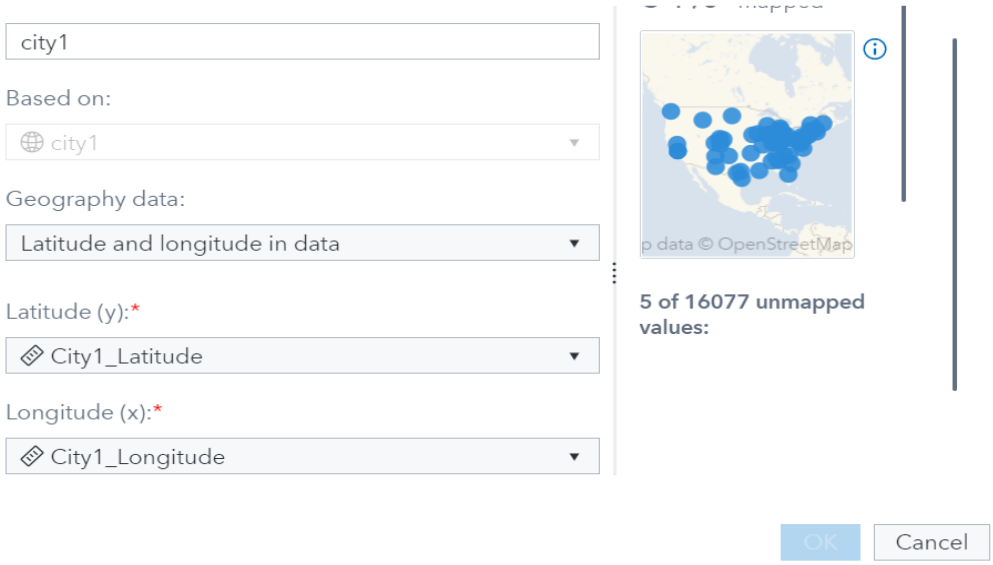


The dashboard helps to analyze the changes in the range of airfare value and range the number of passengers. The data is specific for two cities mentioned in the above Geo-Map. **The domestic consumer airfare report data is informative, and visualization helps me to analyze that how the impact on one**

factor directly effect on the range of other factors. The airfare, range of passenger and carrier details are key factors for an aviation consumer reports.

SAS Feature:

Geo Map: To represent the data of two cities, I used **Geo map**. The geo map is created using the cities latitude and longitude values.



REFERENCES

(ACI), A. C. (2021, March 2021). *The impact of COVID-19 on the airport business and the path to recovery.* Retrieved from Advisory Bulletins: https://aci.aero/2021/03/25/the-impact-of-covid-19-on-the-airport-business-and-the-path-to-recovery/

Knaflic, C. N. (2015). *Storytelling with Data*. Hoboken, New Jersey: John Wiley & Sons, Inc.

U.S. Department of Transportation. (2020, October 27). *Domestic Airline Consumer Airfare*. Retrieved from Transportation: https://www.transportation.gov/policy/aviation-policy/domestic-airline-consumer-airfare-report

Project Analysis:

- The consumer airfare report provides detailed information of the aviation sector.
- The report mainly focuses on the impact of change in airfare value on the number of passengers.
- The airfare range is based on the type of carrier, the capacity of carrier and milage rate covered by the carrier.
- The visual analysis helps to analyze the major changes between certain years into graphical format.
- The statistical representation is easily represented by using effective visuals. The impact of the COVID -19 situation on the aviation sector is represented by using effective visualization.
- The consumer airfare report is generated every four quarters in a year.