Flight Delays Data Analysis

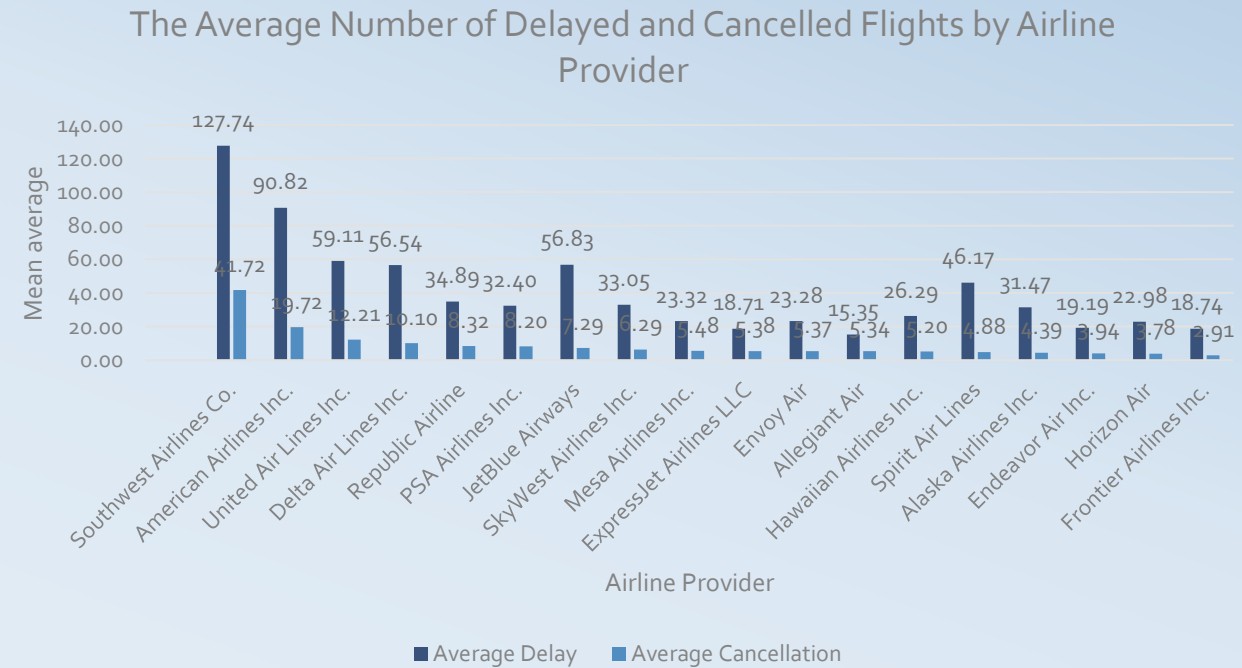Using Python and Google Colab

Brief

- The object of this exercise was to create a way to visualise and analysis airline flight delays to answer the following questions:

  1. What Airlines have the most delayed flights?

  2. Which Airline is the most affected by both delayed and cancelled flights?

  3. What are the main causes of flight cancellations and delays for the company Southwest Airlines Co.?

  4. What are the main cause of flight cancellations and delays among the different airlines?

  5. Did COVID-19 pandemic have an effect on flight delays and cancelations?

- Google Colab was used to create the code to perform analytic research to answer the questions. Google excel was use to take generated analysis to create visual displays to create visualisation to be used in Powerpoint.
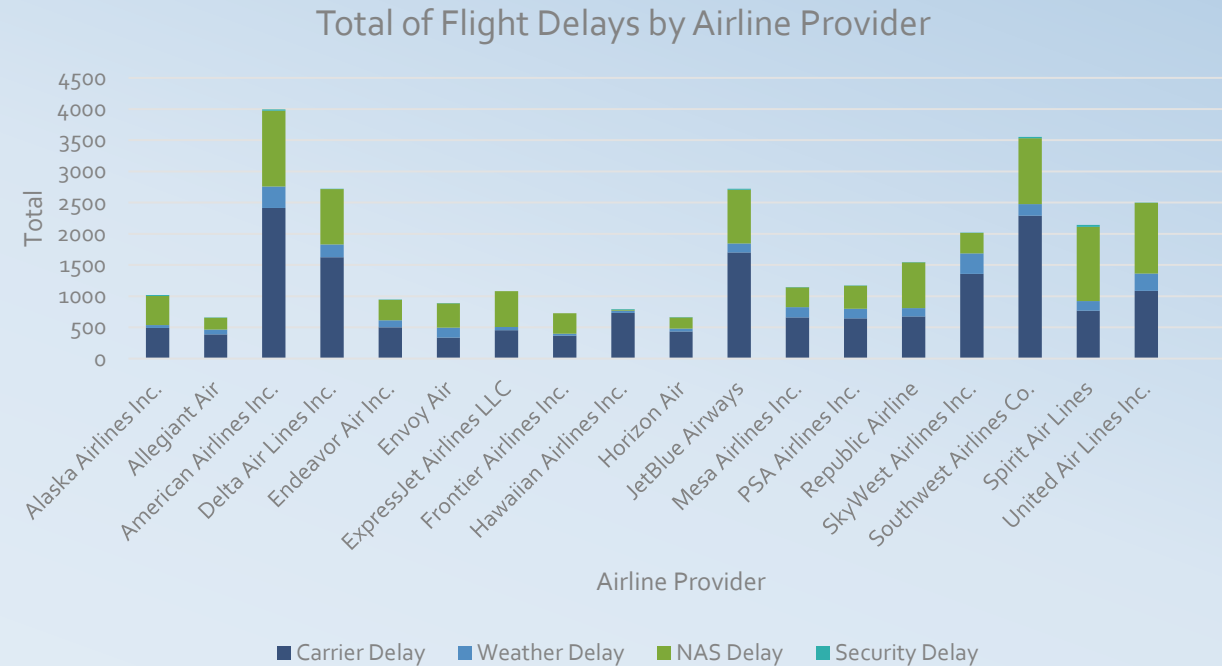
Analysis

- The mean average of the number of delayed flights per airline is visualised using the dark blue columns in the graph, located to the right. This shows that Southwest Airlines Co. had the most delayed flights at a mean average of 128 flights being delayed. Southwest Airlines Co. had 34% more delays than United Air Lines Inc, the next most delayed airline with a mean average of 91.

- Again, the mean average of the number of flights cancelled was also calculated and is displayed using the lighter blue columns in the graph located on the right. This again shows that Southwest Airlines Co. was the airline service most effected by the number of cancelled flights, with an average of 42. This number is considerably higher than the next most affected airline, American Airlines Inc, which had an average delay of 20. This indicates that Southwest Airlines Co. were 72% more likely to cancel their flights compared to American Airlines Inc.
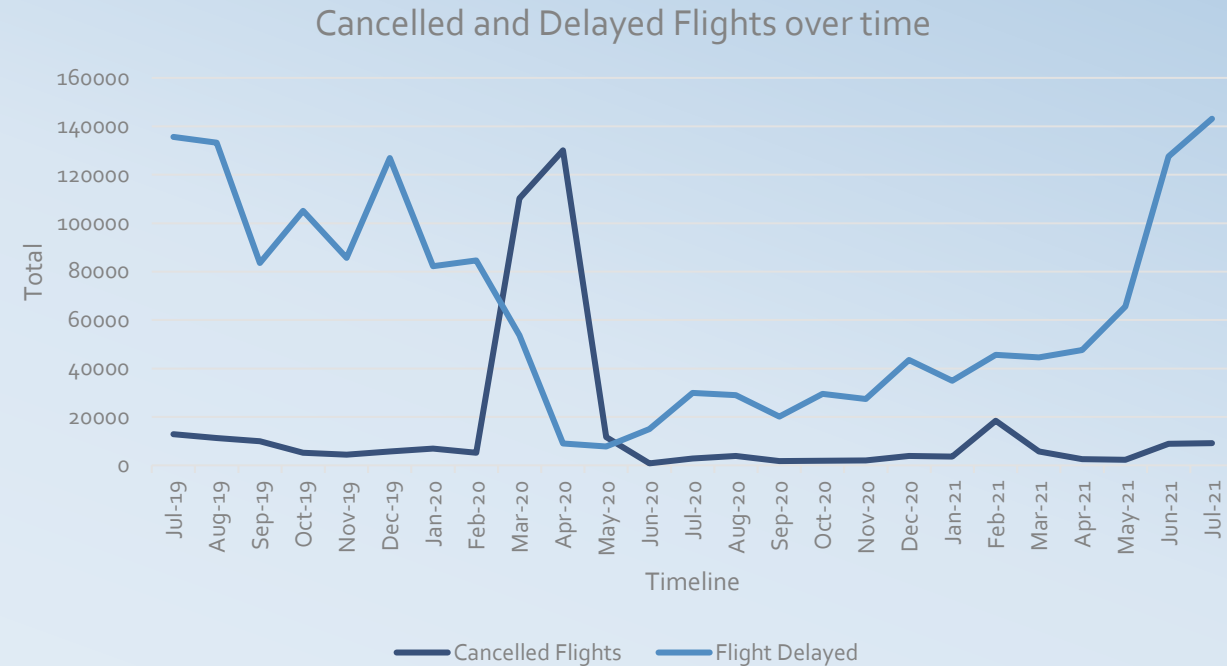

The Average Number of Delayed and Cancelled Flights by Airline Provider

Analysis

- To understand what were the main causes for flight delays anaylsis was conducted to look at the total number of delays caused by the weather, security, the National Airspace System (NAS) and the Carrier themselves.

- Looking at the data from all airline carriers, the main cause of delays is as a result of the carrier themselves, which accounts for 56% of the reasons for a delay. NAS is the next biggest cause for delays, resulting in 35% of delays. Weather accounts for 9% of delays . Security was a minimal factor in delays, accounting for less than 1%.

- Specially, looking at Southwest Airlines Co. The main reason flight were delayed was due to the company themselves, with 64% of the delays being result of the carrier, the NAS accounted for 30% of the delays, whilst the weather (5%) and security (1%) impacted minimally as the reason for delays. Whilst Southwest Airlines Co. did not have the largest total number of delays (American Airlines Inc. had the highest total of delays), it was one of only four airline carriers to feature in the top ten for the highest number of delays for each of the four (Carrier, Weather, NAS and Security delays) categories. The other carriers were American Airlines Inc., Delta Air Lines Inc. and JetBlue Airways.

Total of Flight Delays by Airline Provider

Analysis

- Looking at the graph on the right there is a clean indication that Covid-19 had a clear impact on flight delays and cancellations.

- The average number of flights delayed during the period from the beginning of the dataset until the onset of Covid-19 (July 2019 until December 2019) showed that on average 670,118 planes were delayed. However, during the period of Covid (January 2020 until June 2020) this average dramatically decreased to 42,067. During the post-Covid period (July 2020-July 2020) delays did increase, but were on average less (at 58,898 planes delayed) than pre-Covid levels. However there has been a steady increase in the delay of planes in the in the post-Covid period, where it gradually increased until reaching 135882 delays for the month of July 2021, which sees a return to a similar level recorded for July 2019 (135,630).

- For the cancellation of this was fairly consistent with an average number of cancellations being 7,715 for the period of July2019 to February 2020. However there was a disproportional large increase in the number of flight being cancelled for March and April 2020, which saw an average of 120,133 flights cancelled during this period. This quickly returned to within normal range from May 2020 and in fact there has been a slight improvement, with an average of 5,288 flights between cancelled during the period of May 2020-July 2021.

Cancelled and Delayed Flights over time

Conclusions

1. ## What Airlines have the most delayed flights?

   A. Southwest Airlines Co. has a higher number of flights delayed by quite a wide margin.

2. ## Which Airline is the most affected by both delayed and cancelled flights?

   A. Southwest Airlines Co. is most affected by both delayed and cancelled fights, being the lead for both categories by quite a wide margin.

3. ## What are the main causes of flight cancellations and delays for the company Southwest Airlines Co.?

   A. The two biggest causes of cancellations and delays are the courier causing delays and delays caused by National Airspace System (NAS).

4. ## What are the main cause of flight cancellations and delays among the different airlines?

   A. The main cause of flight cancellation and delays is a result of the airline themselves, followed by NAS then weather and finally security.

5. ## Did COVID-19 pandemic have an effect on flight delays and cancelations?

   A. Covid did play a significant factor in causing an increase of flight cancelations between March and April 2020, before returning to a slight improved normal. Flight delays plummeted during the period of July 2020-July 2020 before slowly increasing to reaching pre-Covid levels.