

Flight Delays & Cancellations Analysis

Kefan Liao, Michael Gattone, Fay Muhammad, and Chris Scarlata

Introduction

The Covid-19 pandemic disrupted industries across the world in 2020. Our analysis explores the pandemic's impact on the U.S. airline industry, focusing on domestic flight delays and cancellations. We hypothesize that the pandemic significantly contributed to an increase in these forms of flight disruptions. To explore this, we analyzed data from the U.S. Department of Transportation from 2018 to 2022 to understand what occurred prior to and following the pandemic. We explore four main questions over this period:

1. Which airline(s) dominated the U.S. domestic flight market? Why?
2. What are the trends in domestic flight cancellations? What events contributed to spikes in flight cancellations?
3. Did the pandemic impact the frequency or duration of flight delays?
4. Which airlines experienced frequent delays and cancellations?

This paper is divided into four sections, each of which investigates these questions to better understand how the pandemic affected domestic air travel.

1. Which airline(s) dominated the U.S. domestic flight market? Why?

We used a clustered column graph to compare the flight activity of different airlines operating in the US across a five-year span between 2018 and 2022. In Figure 1, the y-axis represents the volume of domestic flights, and the x-axis is labeled with the names of the different airlines. Each airline shows a cluster of five (5) columns representing the total number of flights by year.

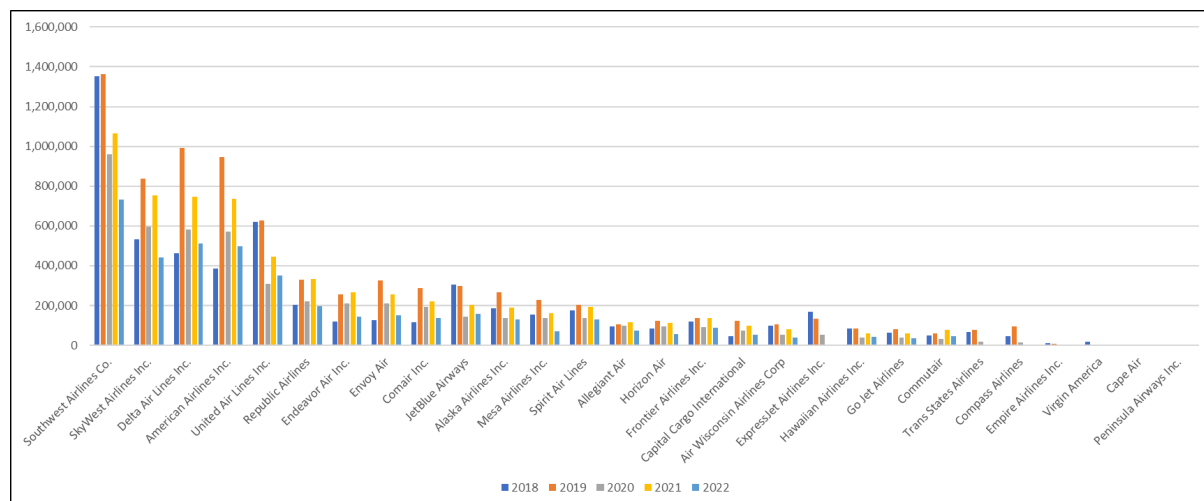


Figure 1. Annual Comparison of Total Domestic Flights by Airline (2018 - 2022): Total number of domestic flights operated by airlines in the U.S. from 2018 to 2022.

Figure 2 displays the cumulative count of flights per airline over the five-year period, allowing for an easy visual comparison of flight volume across different airlines.

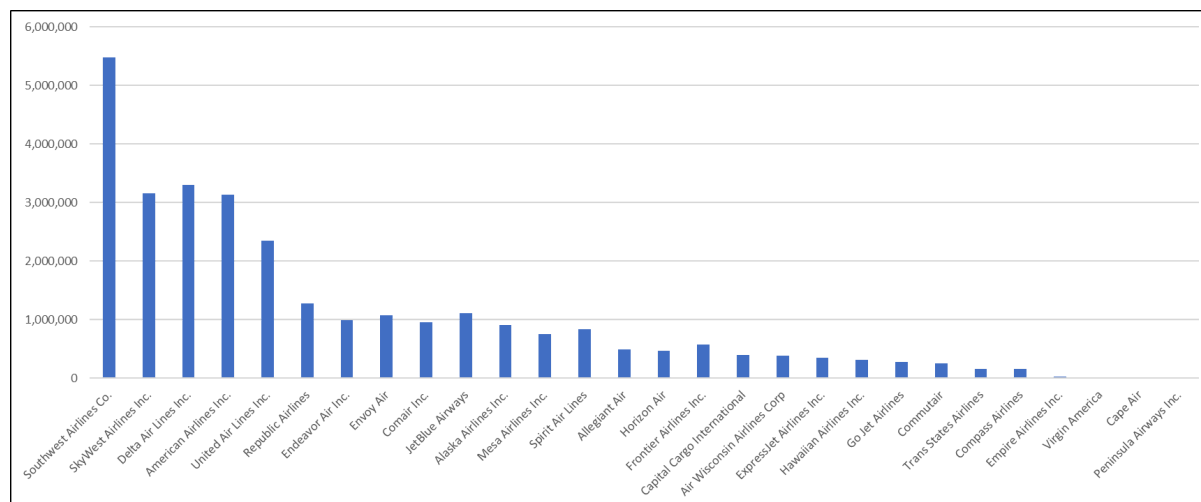


Figure 2. Total Domestic Flights per Airline (2018 - 2022): Total number of domestic flights operated by airlines in the U.S. from 2018 to 2022.

According to the graphs, the top five airlines ranked from highest to lowest in terms of flight volume both before and after the pandemic are Southwest, SkyWest, Delta, American, and United. These airlines operated many flights in 2018 and 2019 and saw an increase in flight activity in the latter year particularly. However, in 2020 there was a significant decline in the number of domestic flights. While airlines saw an uptick in flight volume the following year, from 2020 to 2021, the graph shows that the number of flights post-pandemic did not reach pre-pandemic levels at the end of 2021. It's important to note that the dataset for 2022 is incomplete, only including data from the months of January to July of that year. Therefore, because of this limitation, the visualization may not accurately reflect a true decline in flight volume from 2021 to 2022.

Based on our research on Southwest, we found that the airline's focus on domestic markets and low-cost business model, coupled with its route network, are the primary reasons why they dominate the domestic air travel market in the US. They continued to dominate through 2020 because they were the slowest airline to react and reduce capacity during the pandemic.

2. What are the trends in domestic flight cancellations? What events contributed to spikes in flight cancellations?

We used a cluster column graph to compare the total non-cancelled and canceled domestic flights in the U.S. on an annual basis and a heatmap to visualize the frequency of domestic flight cancellations on a monthly-daily basis.

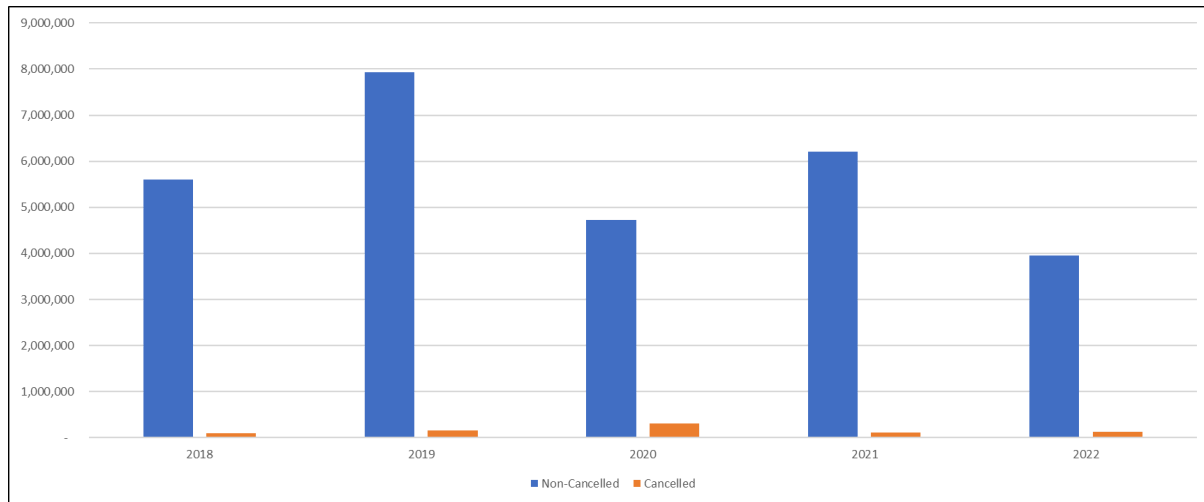


Figure 3. Annual Comparison of Non-Cancelled and Canceled Domestic Flights (2018 - 2022): This graph presents an annual comparison of the total non-cancelled and canceled domestic flights in the U.S.

In Figure 3, the volume of flight cancellations appears to be consistent both before and after the pandemic. However, there's a noticeable spike in flight cancellations in 2020. This increase is likely due to the travel bans and other Covid-19 restrictions that were put into place at the time.

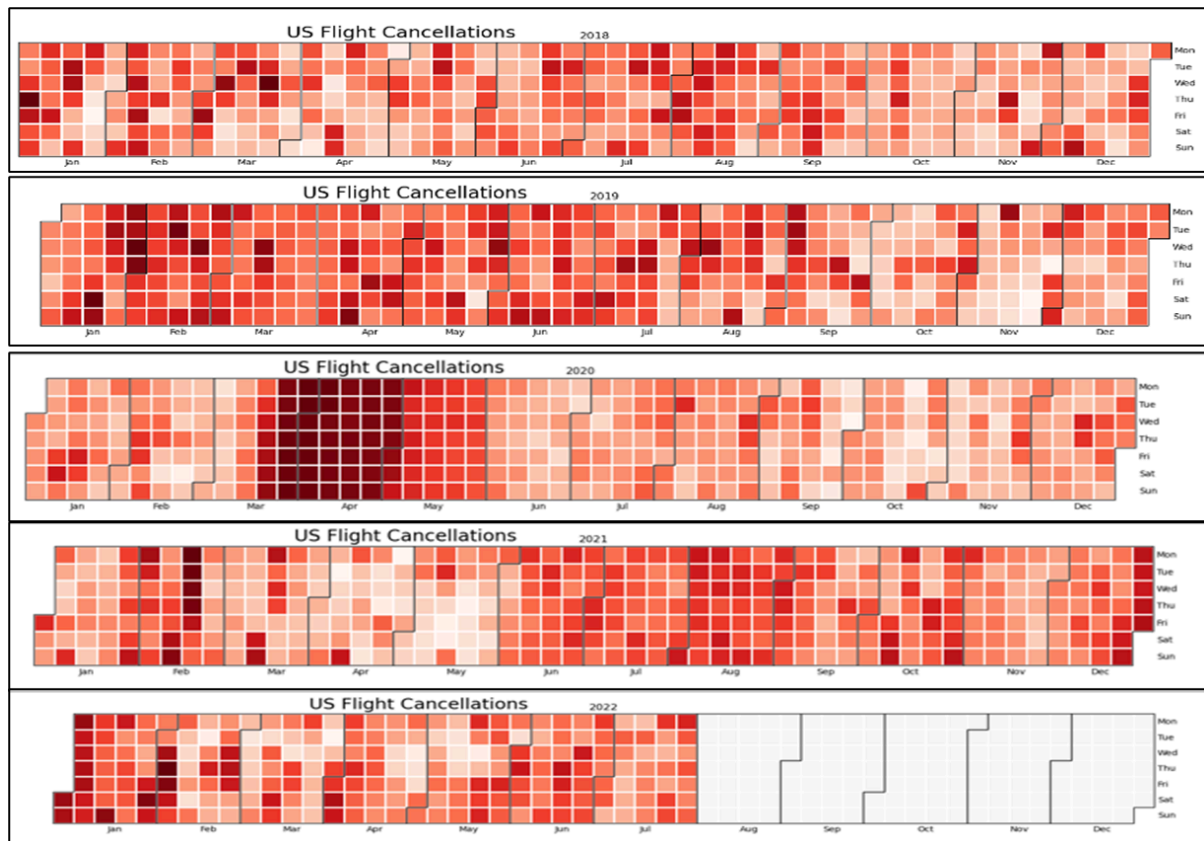


Figure 4. Seasonal Trends in Flight Cancellations (2018 - 2022): This heat map depicts the frequency of domestic flight cancellations in the U.S. from January to December 2018 through 2022. Darker shades of red indicate a high number of cancellations, whereas lighter shades reflect low cancellation numbers.

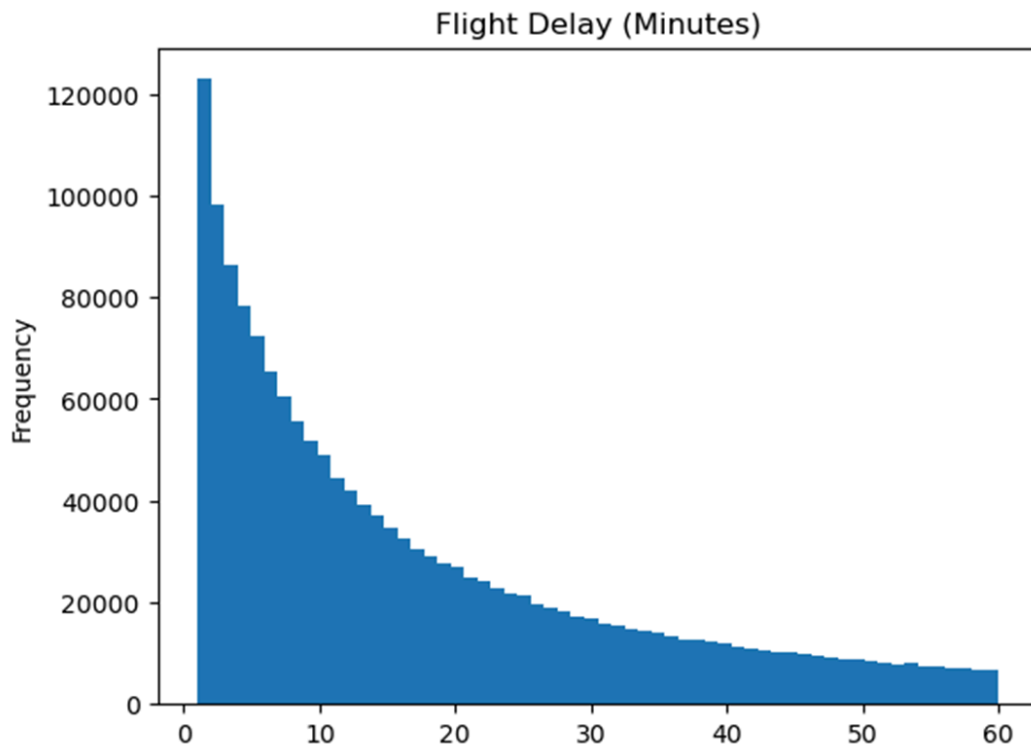
In Figure 4, the heatmap offers a more granular view of flight cancellations for each month of the year between 2018 and 2022. It's clear that the heightened occurrence of flight cancellations took place in the middle of March 2020, and lessened in the end of May. This coincides with the timeline of the WHO declaring Covid-19 a global pandemic.

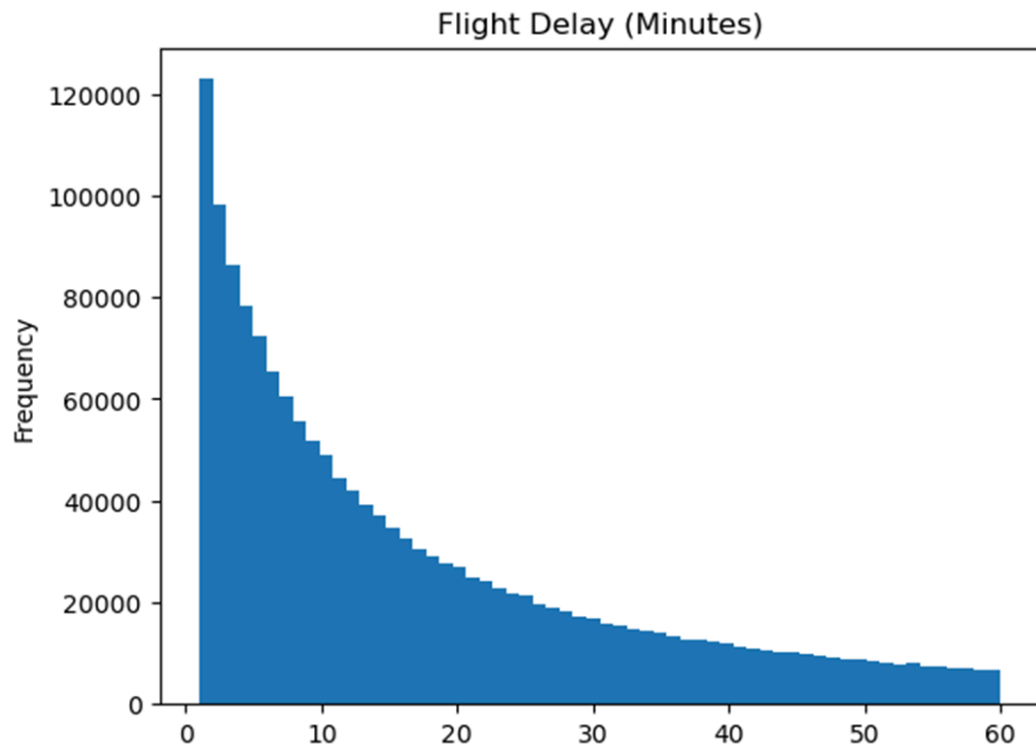
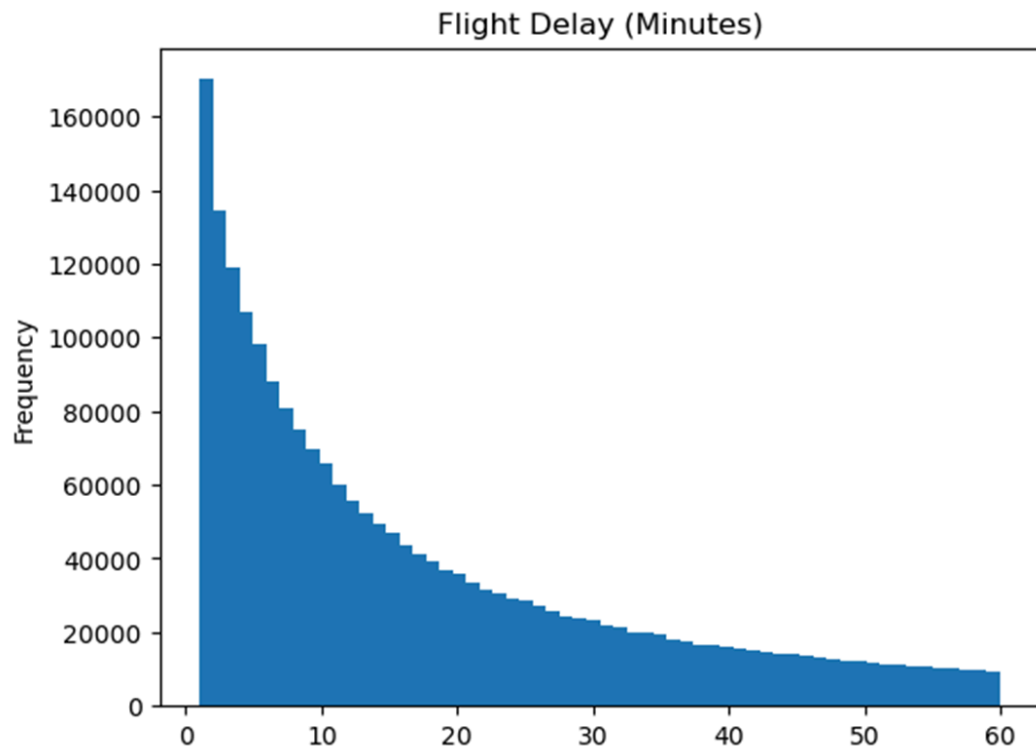
The primary reason more flights were canceled in 2020 compared to 2021 is the onset of the COVID-19 pandemic. As 2021 progressed, vaccination efforts began, and countries implemented measures to control the spread of the virus. Many airlines adjusted their operations, implemented safety protocols, and adapted to the evolving circumstances. The demand for air travel gradually increased, and the need for widespread flight cancellations diminished in comparison to the more severe disruptions experienced in 2020. The world experienced widespread lockdowns, travel restrictions, and a significant decrease in passenger demand due to the rapid spread of the virus. These factors led to a sharp decline in air travel, causing airlines to cancel flights in response to reduced demand, financial challenges, and logistical difficulties. Although the aviation industry continued to face challenges, the situation generally improved compared to the previous year.

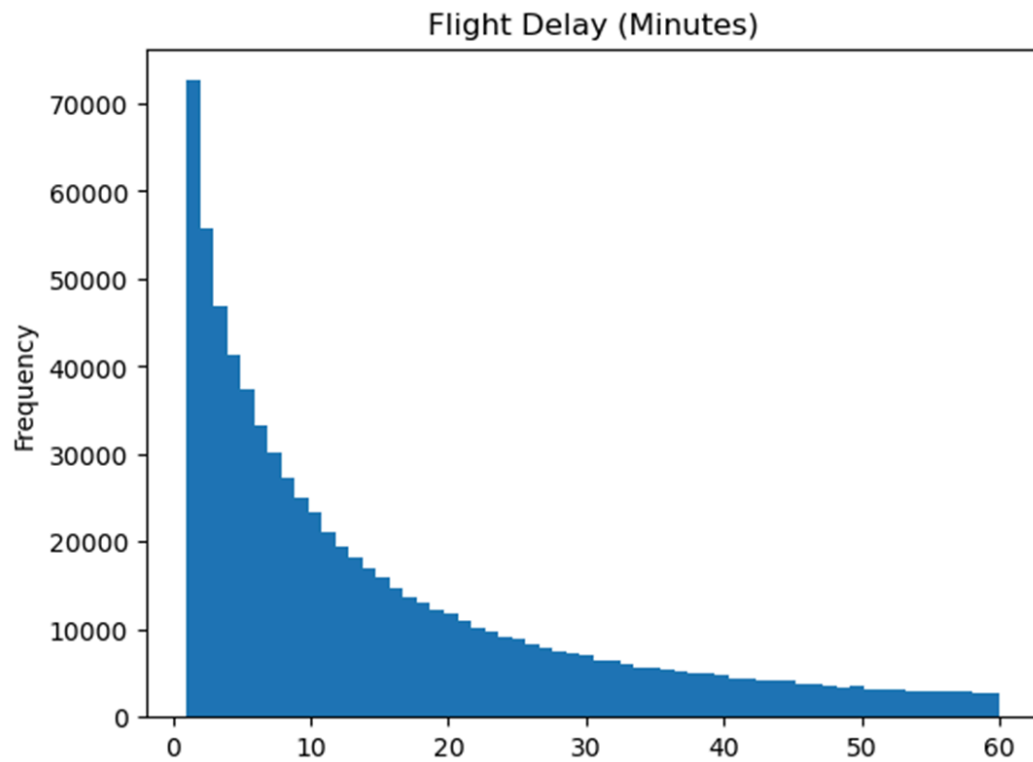
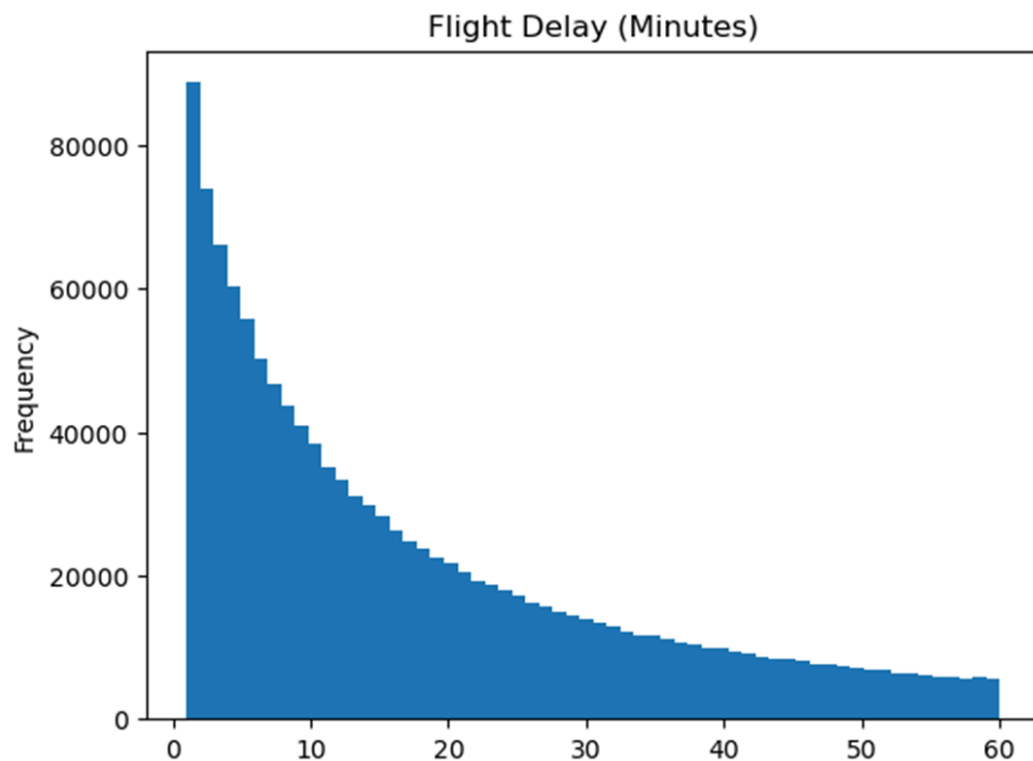
3. Did the pandemic impact the frequency or duration of flight delays?

Based on the histograms in Figures 5A-5E, it appears that the pandemic did not significantly alter the existing patterns of flight delays.

2018



20192020

20212022

Figures 5A-5E. Distribution of Flight Delays (2018 - 2022): The set of histograms above shows the frequency of domestic flight delays ranging from 0 to 60 minutes in 10-minute intervals.

The histograms for each year display a similar pattern, where the distribution sharply slopes downward before leveling off. This indicates that across the five-year period analyzed, the most common occurrence of delays lasted 0 to 10 minutes. As the delay durations increase, the frequency of delay occurrences also falls and levels off as it approaches the 60-minute mark.

In short, there is a consistent pattern across all years indicating the long-duration delays are significantly less frequent than short-term delays despite the COVID shock.

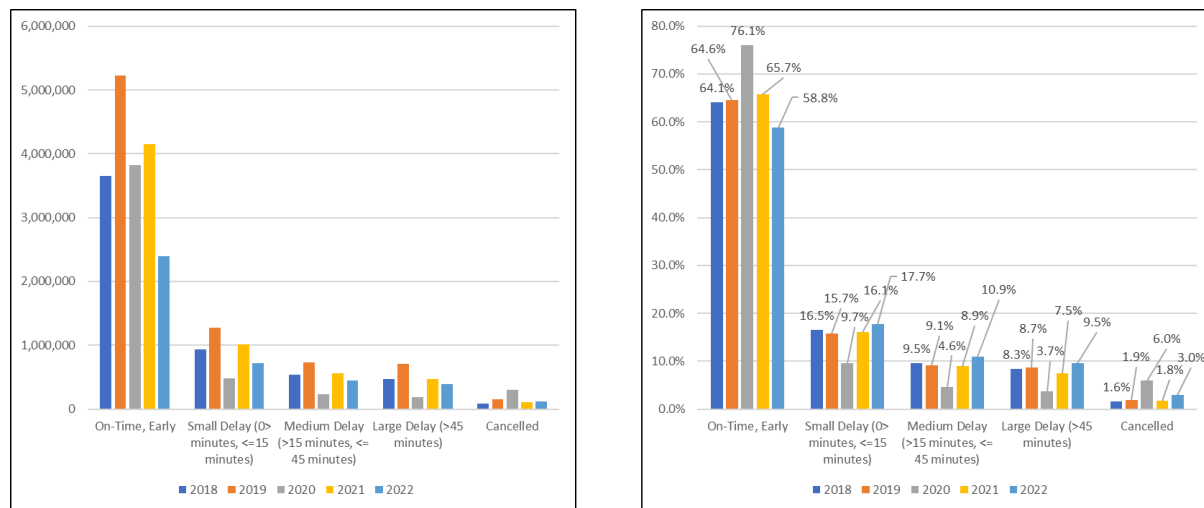


Figure 6A-6B. Distribution of Flight Disruption by Category (2018 - 2022): These clustered bar graphs show the frequency of flight delays in terms of raw count (left) and percentage (right) categorized into five (5) groups: On-time/Early, Small Delay (1-15 minutes), Medium Delay (16-45 minutes), Large Delay (>45 minutes), and Canceled.

In Figures 6A-6B, we can see that year-on-year, the large majority of flights operated on-schedule over a five-year period. However, in 2020 there is a noticeable increase in flight cancellation occurrences compared to the years prior and after the pandemic. While it's highly probable that the high cancellation rate was driven by the Covid-19 pandemic, more statistical analysis is needed to confirm the relationship at a significant level.

The heatmap tables in Figures 7A to 7D shows the percentage of flights in each month that fall under the five flight disruption categories: On-time/Early, Small Delay (0-15 minutes), Medium Delay (16-45 minutes), Large Delay (>45 minutes), and Canceled. We observe that cancellations are not very common both pre- and post-pandemic, with cancellation rates occurring less than 6% of the time in each month. However, flight cancellation rates were an anomaly in 2020 accounting for 17% and 41% of flights in March and April 2020, respectively. However, the average Flight cancellation rate from May 2020 through December 2020 is about 1% which is mainly due to the fall of demand and airline capacity cut.

2018 Delay Groups by Month

| Delay Group | On Time Early | Small Delay | Medium_Delay | Large_Delay | Cancelled |
|-------------|---------------|-------------|--------------|-------------|-----------|
| Month | | | | | |
| 1 | 65.0% | 15.2% | 9.1% | 8.0% | 2.8% |
| 2 | 60.5% | 18.4% | 11.1% | 8.1% | 1.9% |
| 3 | 62.0% | 18.2% | 10.1% | 7.1% | 2.6% |
| 4 | 65.8% | 16.4% | 9.2% | 7.6% | 1.0% |
| 5 | 60.5% | 18.0% | 10.7% | 9.6% | 1.2% |
| 6 | 57.1% | 19.1% | 11.5% | 10.8% | 1.4% |
| 7 | 58.5% | 18.1% | 11.1% | 10.7% | 1.7% |
| 8 | 59.2% | 16.5% | 10.4% | 11.6% | 2.3% |
| 9 | 70.2% | 13.5% | 7.6% | 7.1% | 1.5% |
| 10 | 69.2% | 15.3% | 8.2% | 6.5% | 0.78% |
| 11 | 64.9% | 16.7% | 9.4% | 7.8% | 1.2% |
| 12 | 65.1% | 16.6% | 9.5% | 7.6% | 1.2% |

2019 Delay Groups by Month

| Delay Group | On Time Early | Small Delay | Medium Delay | Large Delay | Cancelled |
|-------------|---------------|-------------|--------------|-------------|-----------|
| Month | | | | | |
| 1 | 65.5% | 14.9% | 8.6% | 7.9% | 3.1% |
| 2 | 60.1% | 16.3% | 10.4% | 10.0% | 3.1% |
| 3 | 65.7% | 16.1% | 8.7% | 7.5% | 2.0% |
| 4 | 65.9% | 14.8% | 8.5% | 8.5% | 2.4% |
| 5 | 63.3% | 15.9% | 9.4% | 9.4% | 2.0% |
| 6 | 58.1% | 17.1% | 10.9% | 11.8% | 2.1% |
| 7 | 62.1% | 16.0% | 9.5% | 10.4% | 2.1% |
| 8 | 63.2% | 15.7% | 9.5% | 9.8% | 1.8% |
| 9 | 71.8% | 13.4% | 6.8% | 6.3% | 1.7% |
| 10 | 68.5% | 15.2% | 8.3% | 7.1% | 0.89% |
| 11 | 70.3% | 15.1% | 7.7% | 6.1% | 0.85% |
| 12 | 61.1% | 18.0% | 10.6% | 9.2% | 1.06% |

2020 Delay Groups by Month

| Delay Group | On Time Early | Small_Delay | Medium_Delay | Large_Delay | Cancelled |
|-------------|---------------|-------------|--------------|-------------|-----------|
| Month | | | | | |
| 1 | 71.3% | 14.2% | 7.3% | 5.9% | 1.3% |
| 2 | 71.4% | 13.6% | 7.2% | 6.8% | 1.0% |
| 3 | 67.8% | 7.8% | 4.2% | 3.4% | 16.9% |
| 4 | 52.9% | 3.0% | 1.4% | 1.3% | 41.3% |
| 5 | 82.5% | 6.7% | 2.5% | 1.9% | 6.4% |
| 6 | 83.8% | 9.6% | 3.6% | 2.5% | 0.43% |
| 7 | 83.8% | 8.3% | 3.8% | 3.3% | 0.79% |
| 8 | 84.3% | 8.2% | 3.6% | 2.8% | 1.1% |
| 9 | 85.3% | 8.4% | 3.2% | 2.4% | 0.73% |
| 10 | 83.4% | 9.3% | 3.8% | 2.9% | 0.53% |
| 11 | 85.2% | 8.3% | 3.4% | 2.6% | 0.54% |
| 12 | 77.8% | 11.7% | 5.5% | 3.8% | 1.07% |

2021 Delay Groups by Month

| Delay Group | On Time Early | Small Delay | Medium Delay | Large_Delay | Cancelled |
|-------------|---------------|-------------|--------------|-------------|-----------|
| Month | | | | | |
| 1 | 81.6% | 9.2% | 4.4% | 3.7% | 1.1% |
| 2 | 72.0% | 11.5% | 5.8% | 5.0% | 5.8% |
| 3 | 76.6% | 12.6% | 5.6% | 3.9% | 1.3% |
| 4 | 76.0% | 13.4% | 5.8% | 4.2% | 0.54% |
| 5 | 70.7% | 15.9% | 7.3% | 5.7% | 0.45% |
| 6 | 57.0% | 18.5% | 11.6% | 11.4% | 1.6% |
| 7 | 55.1% | 19.3% | 12.1% | 11.8% | 1.7% |
| 8 | 57.9% | 17.5% | 11.1% | 10.4% | 3.1% |
| 9 | 69.7% | 15.3% | 7.7% | 5.9% | 1.4% |
| 10 | 62.5% | 17.5% | 9.9% | 8.0% | 2.1% |
| 11 | 65.8% | 18.4% | 9.2% | 6.0% | 0.62% |
| 12 | 57.2% | 19.1% | 11.9% | 9.5% | 2.4% |

2022 Delay Groups by Month

| Delay Group | On Time_Early | Small_Delay | Medium_Delay | Large_Delay | Cancelled |
|-------------|---------------|-------------|--------------|-------------|-----------|
| Month | | | | | |
| 1 | 61.1% | 14.8% | 9.2% | 8.6% | 6.3% |
| 2 | 60.2% | 17.2% | 10.0% | 8.2% | 4.5% |
| 3 | 59.8% | 18.4% | 11.0% | 9.2% | 1.5% |
| 4 | 58.8% | 18.1% | 11.2% | 9.7% | 2.3% |
| 5 | 59.2% | 18.9% | 11.0% | 8.9% | 2.0% |
| 6 | 55.2% | 18.7% | 12.1% | 10.9% | 3.1% |
| 7 | 57.8% | 17.7% | 11.8% | 10.9% | 1.8% |

Figure 7A - 7E. Frequency of Flight Disruption Types by Month (2018 - 2022): With the exception of Figure 7E., these heat maps depict the frequency of domestic flight delays and cancellations in the U.S. from January to December 2018 through 2021. Darker shades of blue indicate a higher frequency of delays, and/or cancellations. The 2022 dataset trunks through January through July.

The charts above support our hypothesis that COVID-19 pandemic significantly contributed to an increase in domestic flight cancellations in 2020. However, the data also show that On-Time Performance (OTP) improves. COVID-19 caused a severe decrease in airline traffic across all routes. This created a situation of excess capacity at airports which releases the stress and congestion at airports.

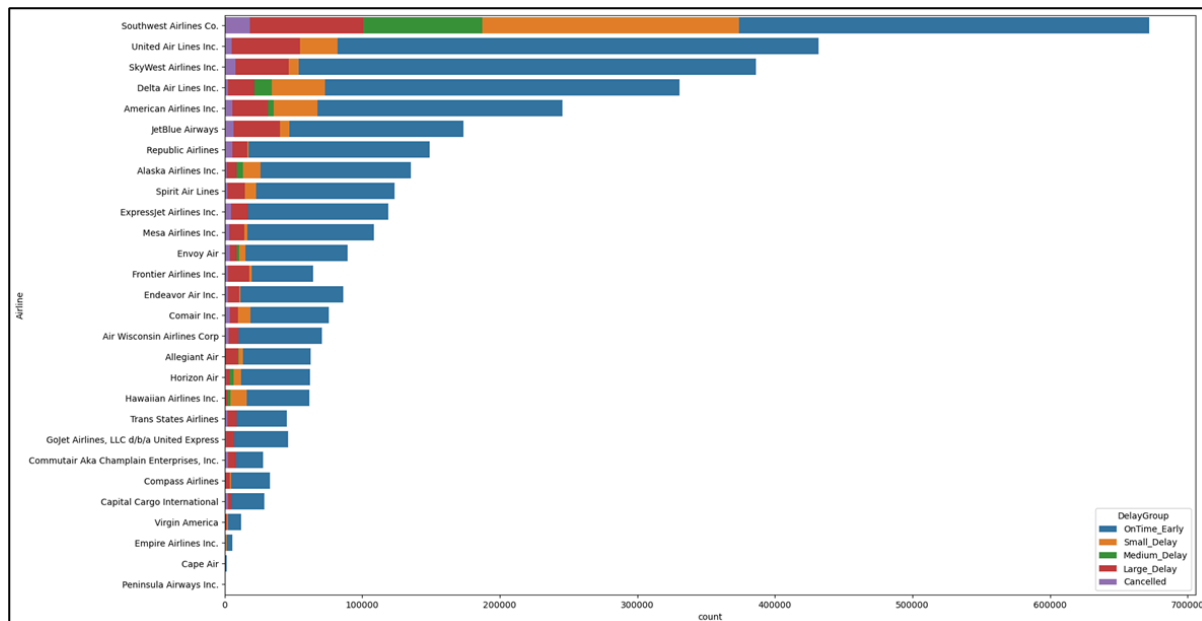
4. Which airlines experienced frequent delays and cancellations?

The top five airlines, except for Southwest, managed to operate a significant majority of their flights on-time or early both pre- and post-pandemic from 2018-2019 and 2021-2022. Southwest, however, the leading domestic airline over this period, had a lower proportion of flights operating on time or early. Instead, they show a pattern of operating more than half of their flights with delays and cancellations both pre- and post-pandemic.

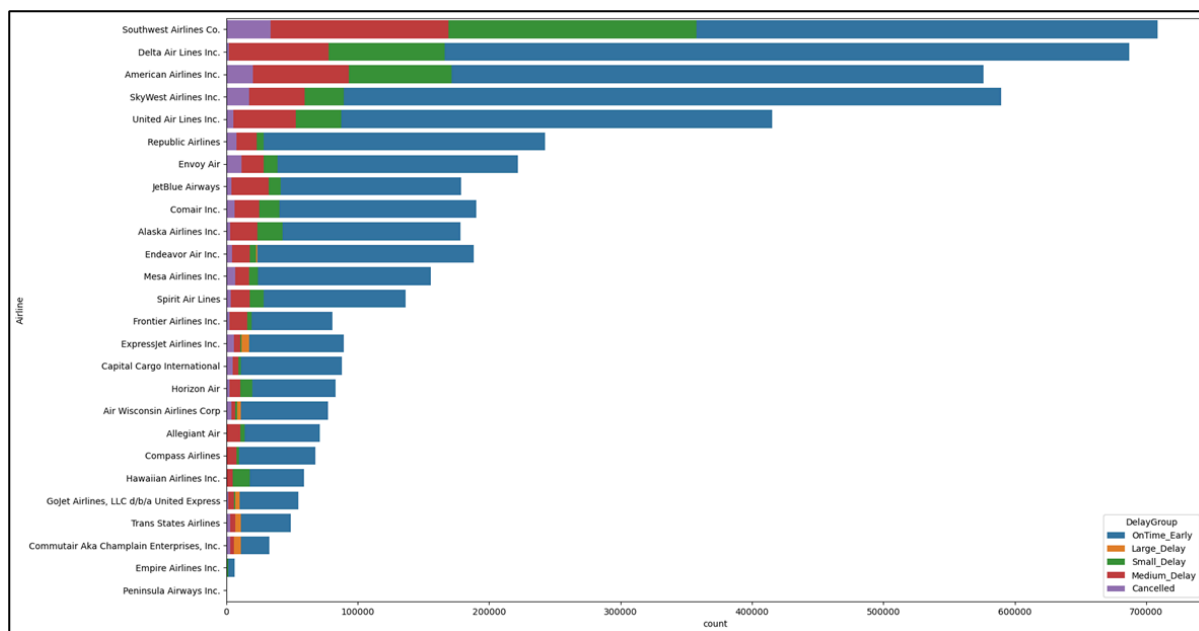
Figures 8A-8E show that in 2018, 2021, and 2022, small and large delays were the most prevalent types of flight disruptions. In 2019, we see small and medium delays being more common, while in 2020, small delays and cancellations were the dominant type of flight disruption.

Relative to the total number of flights operated by each airline in 2020, most flights operated on time and early during the pandemic year. With fewer people traveling during the pandemic, airlines experienced less air traffic, which likely contributed to fewer delays. While delays were minimized, airlines saw higher rates of flight cancellations in 2020 relative to pre- and post-pandemic years.

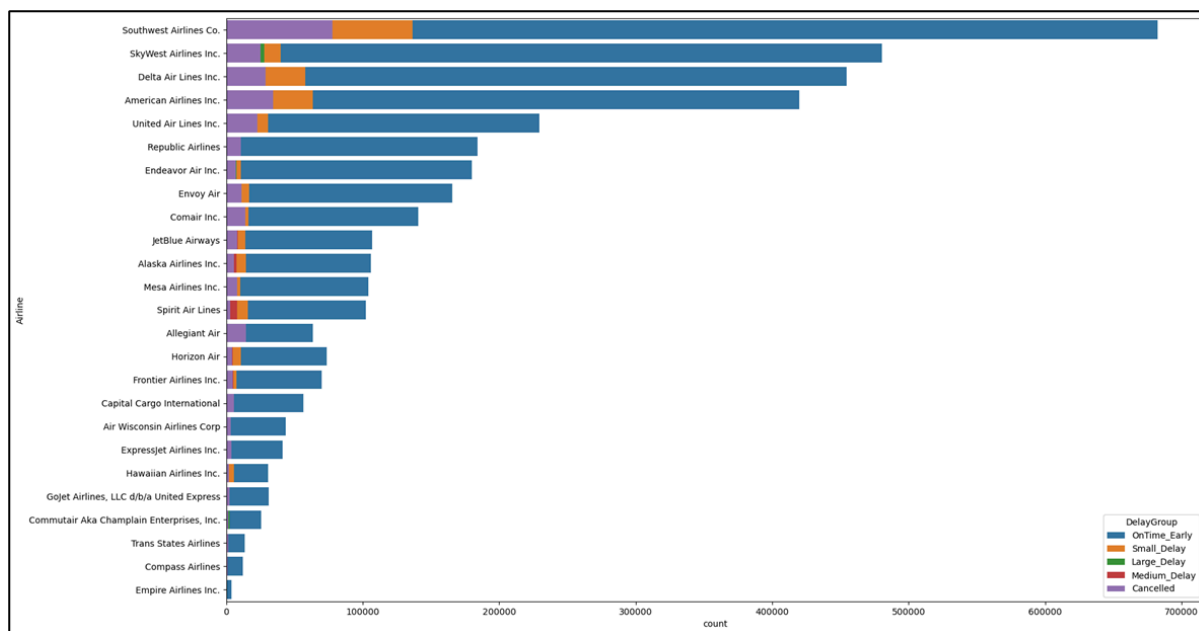
2018



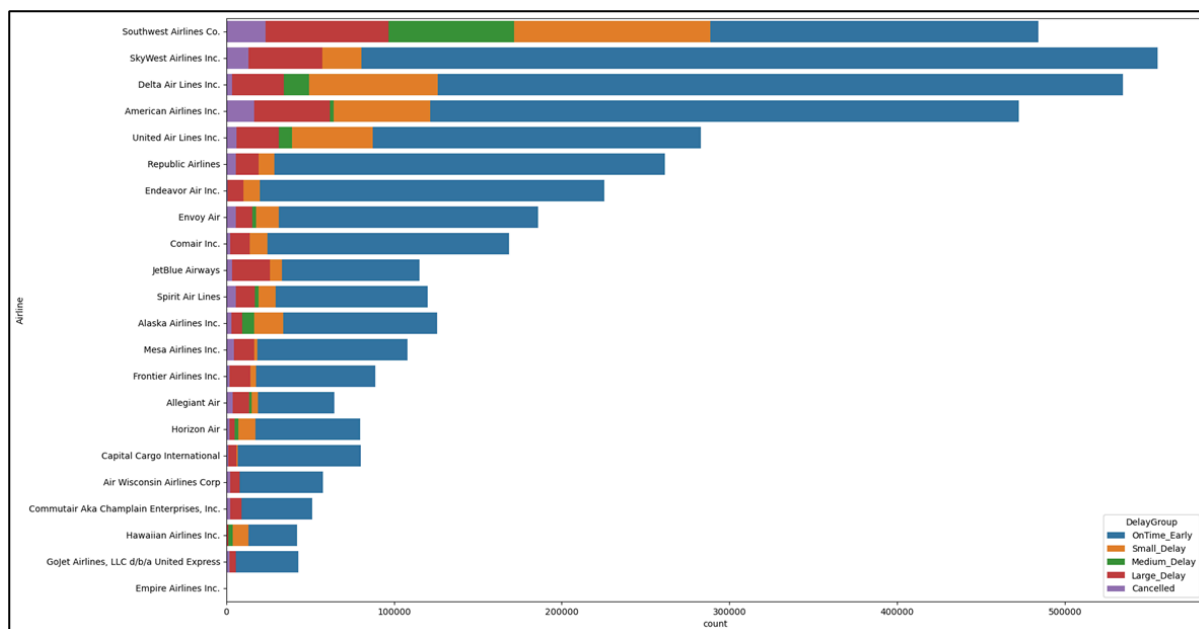
2019



2020



2021



2022

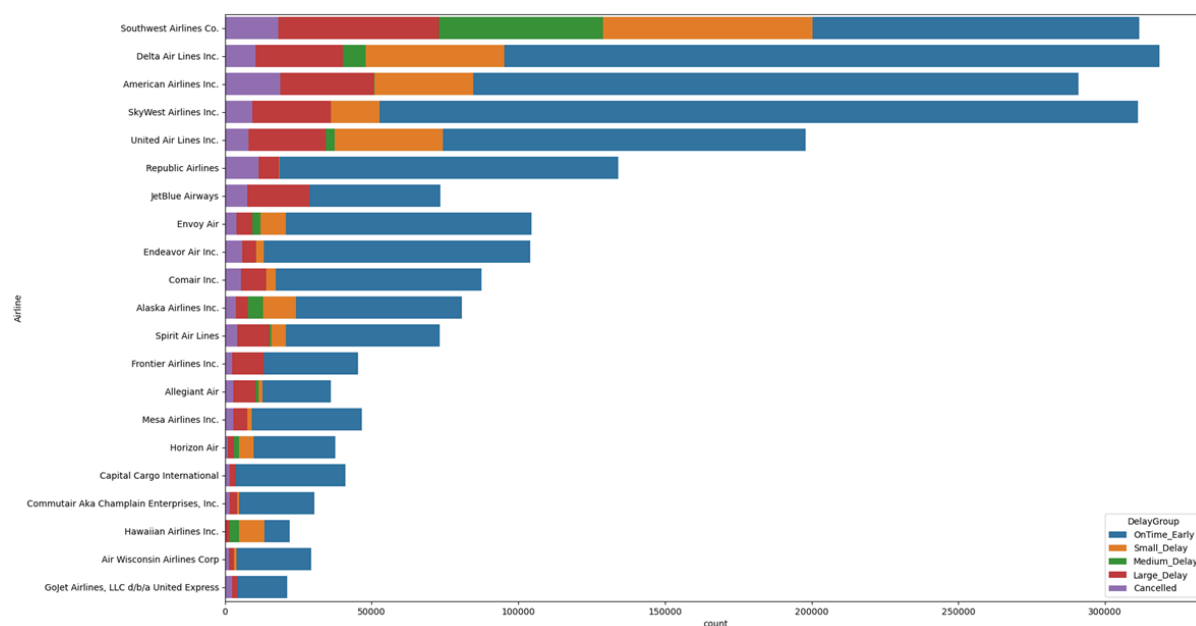


Figure 8A-8E. Flight Disruptions by Airline (2018 - 2022): These stacked bar graphs show the number of flights by various airlines for each year between 2018 and 2022. The segments within the bar are color-coded to depict the following categories: On-time/Early, Small Delay (0-15 minutes), Medium Delay (16-45 minutes), Large Delay (>45 minutes), and Cancelled.

Conclusion & Further Research

Based on our analysis, we conclude that our initial hypothesis is partially supported by the data. The data indicates that Covid-19 contributed to more domestic flight cancellations. However, the data doesn't back our hypothesis that the pandemic led to more flight delays. On the contrary, we observed fewer delays during the pandemic. With fewer people traveling during the pandemic, airlines experienced less air traffic and congestion, which likely helped reduce flight delays in 2020.

For further research, we could incorporate geographical data and analyze changes in passenger preferences in terms of flight routes before and after the pandemic. There's also the opportunity to study how local government Covid-19 restrictions affected passenger preference and flight routes.