Domestic Airline Delays: a regional analysis on the impact of the COVID-19 pandemic on flight delays

Natalia Mosquera, Wyatt Tomko, Andrew Heneghan, Kyle Knox

Introduction

Background

Air travel has become an integral part of our lives. The punctuality of domestic flights is paramount for both travelers and the aviation industry. Understanding the intricacies of flight delays is a practical necessity to enhance operational efficiency and ensure consumer satisfaction. We aimed to investigate whether there is an association between the region and year of travel. Additionally, we analyzed whether the COVID-19 pandemic affected regional delays.

Research Goal

The goal of this analysis was to determine which domestic flights experience the most delays across the United States. To achieve this, we conducted a comprehensive examination of airline data provided by the Bureau of Transportation spanning from 2018 to 2019 and 2022 to 2023. The data was grouped into two datasets. The first included pre-pandemic data from 2018 to 2019 and the second included post-pandemic data from 2022 to 2023. The years 2020 and 2021 were excluded from the analysis due to the global pandemic. This exclusion serves as a dual purpose, to preserve the data integrity and enable a comparative analysis of pre-pandemic and post-pandemic flight performances.

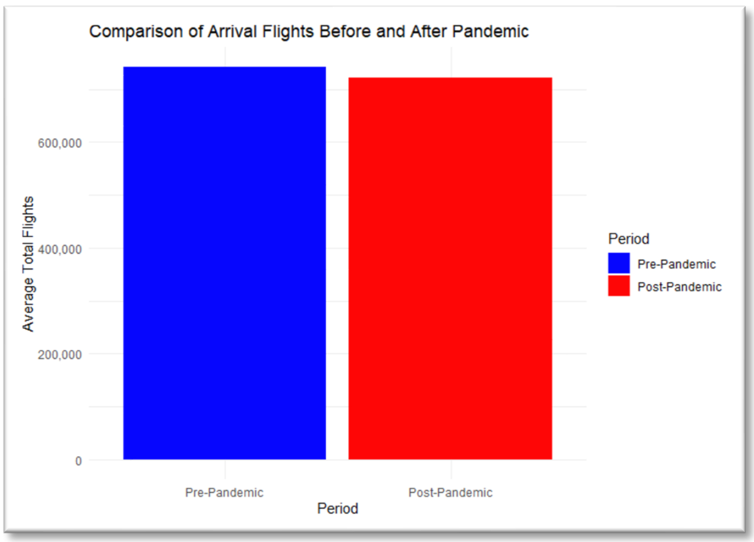
Methods

The study hinges on a crafted sampling methodology, drawing from a population of domestic flights operated by major domestic airline carriers, as defined by the Bureau of Transportation. The regional flights documented in the datasets travelled to seven distinguished territories: Northwestern, North Central, Northeastern, Southwestern, South Central, and Southeastern states, Alaska and Hawaii. In the stratification of the data, we assigned weights to each city based on the total number of flights to each of the regions. This stratification ensured that our final sample is not only representative of the flight population but also accounted for the varying number of flights across different regions.

Findings

The objective of our analysis was to see if there was a significant difference in domestic flight delays because of the COVID-19 pandemic. It was found that flights to the southern part of the United States have witnessed significantly more delays compared to those flights to the northern part of the nation, notably after the COVID-19 pandemic. The findings found in the analysis transcend statistical analysis, honing consumer confidence, operational efficiency, and cost implications within the aviation world. The exploration of flight volume and delays unveils insights into operational dynamics that impose enhancing efficiency and mitigating disruptions. The cost implications of flight delays exert financial pressures on airlines through escalated fuel costs, potential compensations to customers, and an increase in employee compensations.

Visual Exploration of the Data



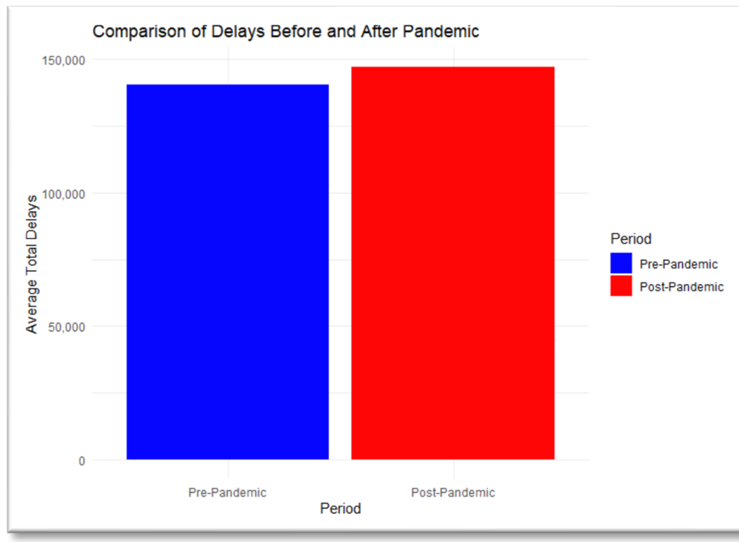


Figure 1: Illustrates the number of arrival flights before and after the pandemic. The distributions are similar, the figure shows pre-pandemic had more flights compared to post-pandemic.

Figure 2: Illustrates the number of delayed flights before and after the pandemic. The distributions are similar, but the figure shows that post-pandemic has more delays compared to pre-pandemic

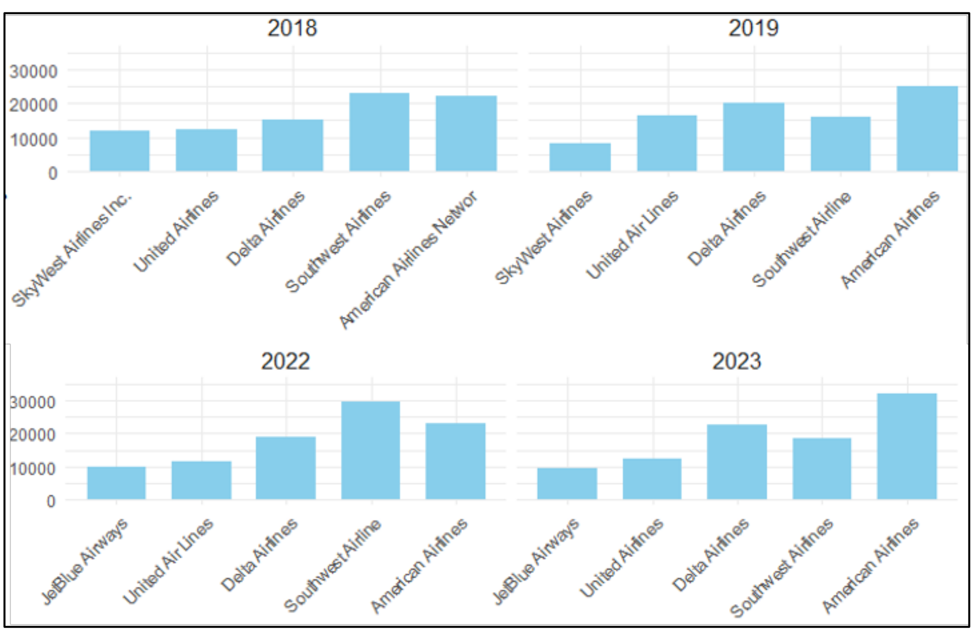


Figure 3: Shows the top 5 airlines with the most flights during the four years of this study. United Airlines, Delta Airlines, Southwest Airlines, and American Airlines are leaders during all four years. These airlines fly across all regions of the United States, leading to more flights.

Visual Exploration of the Data

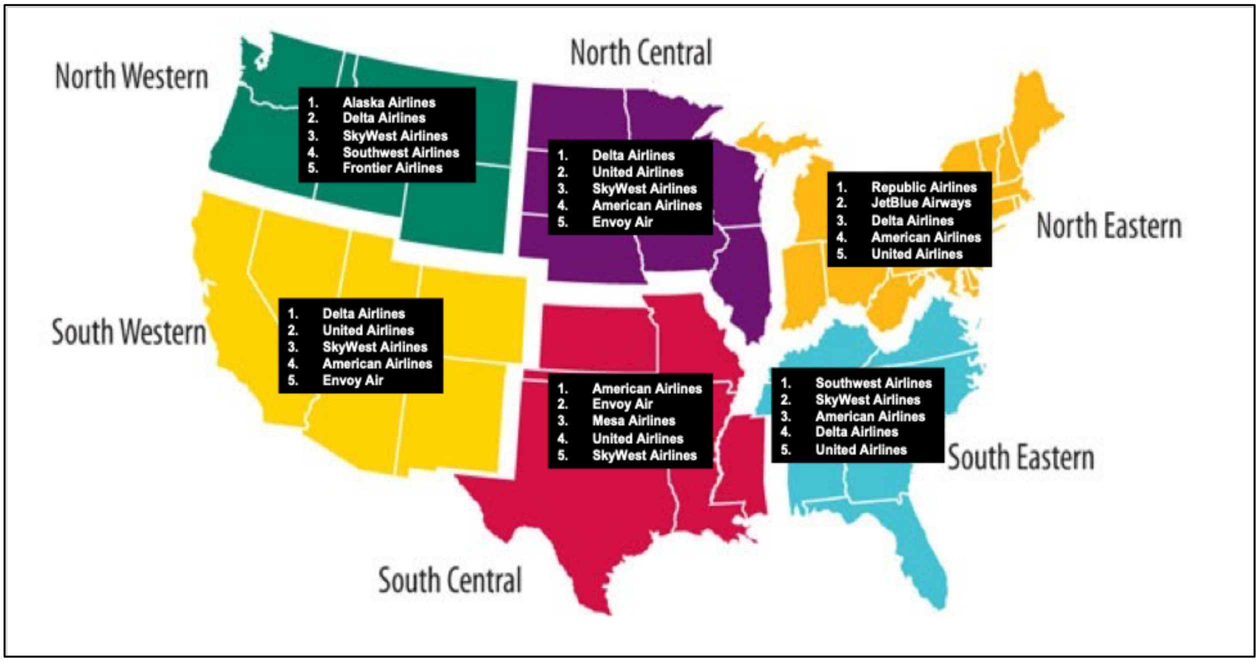


Figure 4: Illustrates the top 5 airlines with the most flights in each region of the United States. Delta Airlines and American Airlines are leading airlines in almost all regions. Furthermore, emphasizing them as the top airlines in the United States, further confirming the findings in Figure 3.

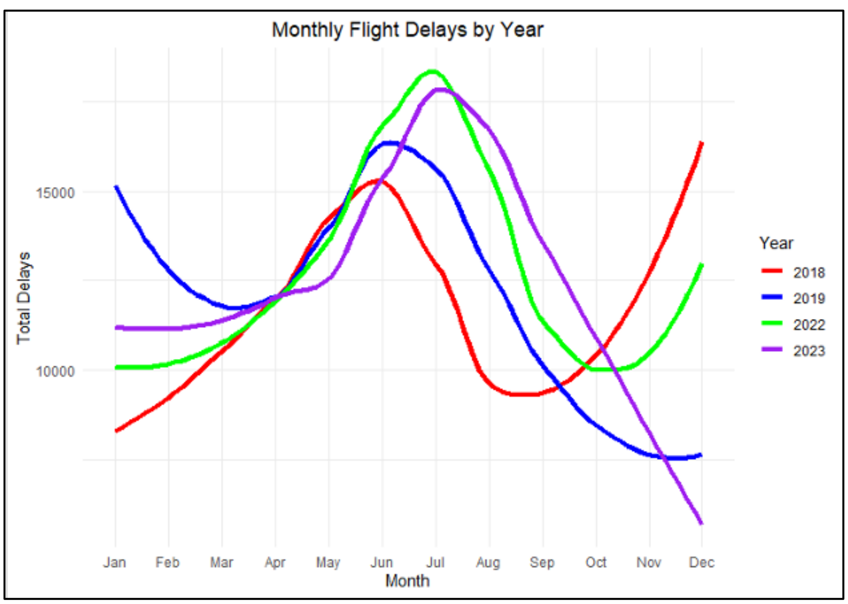


Figure 5: Illustrates the distribution of the number of flights by each year in the study. The peak months in all years are from the months May through July. The density of flights rises again towards the end of year near the holidays. However, in 2023 the flights did not increase at the end of the year.

Visual Exploration of the Data

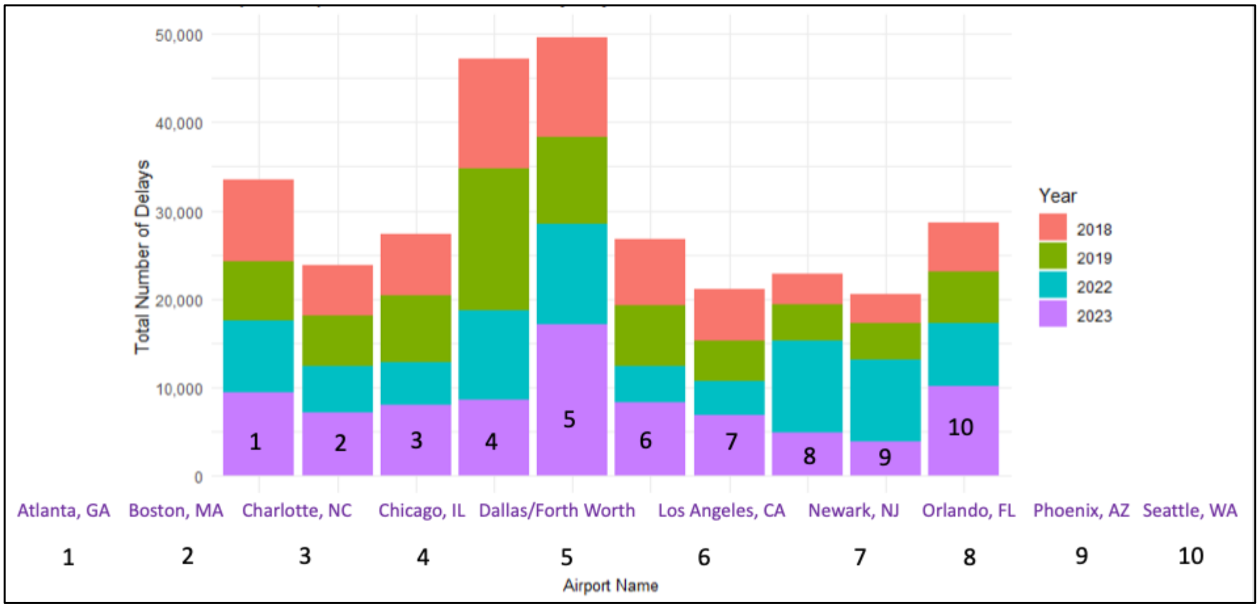


Figure 6: Illustrates the airports with the most delays. Dallas/Fort Worth, Chicago, and Atlanta have the most delays. These airports are also three of the busiest airports in the United States.

Figure 7: Illustrates the delays for each region separated by year. The Northeast region has the most delays in 2018 and 2019 (pre-pandemic), whereas the Southeast region has the most delays for 2022 and 2023 (post-pandemic).

Methods

Defining the Data

The data used in this project originates from the U.S. Department of Transportation's (DOT) Bureau of Transportation Statistics (BTS). It is used to track on-time performance of domestic flights operated by large air carriers. Only 2018-2019 (pre-COVID) and 2022-2023 (post-COVID) flight data was used to get the sample for the analyses. The years 2020 and 2021 are excluded since there were less people flying these years due to the COVID-19 pandemic. This led to less flights and delays and could thus bias the results. Also, only plane observations that arrived at airports that were classified as "major" by the Bureau of Transportation Statistics were used for creating the population we sampled from.

The main outcome variable for our analyses was a proportion variable we created, called prop_del. This is the proportion of delayed flights (arr_del15) over total arrived flights (arr_flights) for a certain airline at a certain airport. Another important variable we created was region, which was our main stratification variable for sampling. It describes a section of the U.S. based off location of a major domestic airport where each sampled plane landed. Also, while we did use year in which an airplane was observed to subset the data, we further categorized this variable by creating a new variable with two levels called newYear. For flights that were sampled pre-COVID (2018 or 2019), the value of newYear was 0, while for flights sampled post-COVID (2022-2023), newYear was 1. This was used to further stratify the data. Other variables used for sampling include the name of the airline an observed plane belongs to and the month each airplane was observed since data was collected between January through December for each year.

Sampling Methods

In our statistical analysis of domestic flight delays across the United States, we employed a stratified random sampling technique to ensure that the sample was representative of various geographical regions. The dataset contained information for the total flights operated by various airlines, categorized into six distinct regions: North Central, Northeast, Southeast, South Central, Southwest, and Northwest. Initially, the goal of the stratification was to capture the potentially different dynamics of flight operations and delays in these regions, as well as different factors that lead to delays such as weather, security, aircraft issues, among others that vary significantly across different parts of the country. Ultimately, stratification was used to determine whether regulatory policies implemented due to the COVID-19 pandemic led to flight delays.

The total sample size of 2,800 was distributed equally among the four years: 2018, 2019, 2022, and 2023, with each year receiving a sample size of 700 flights. Each year, flights were further categorized based on regional stratification so that the sample was representative of the population. This approach allowed us to observe any temporal changes in flight delay patterns, potentially influenced by changes in airline operations, regulatory policies, or external factors like economic shifts caused by the COVID-19 pandemic. Each regional sample was chosen to reflect the proportion of total flights operated in that region, ensuring that more heavily trafficked regions like the Northeast and Southeast had correspondingly higher representation in the sample compared to less trafficked regions such as the Northwest. For instance, in 2023, the Northeast had a sample size of 803 flights compared to the Northwest's 225, mirroring their real-world flight operation volumes. It is important to note that the flight values depended on each distinct airline. This methodological choice enhances the accuracy and reliability of our analysis, allowing for a nuanced understanding of regional differences in flight delay patterns across the United States.

Statistical Methods

To analyze the effect of the difference of flight delays pre and post covid, two statistical methods were employed for the investigation. First, to look at an overall pre and post covid analysis, the sample proportion plus 95% confidence interval for stratified samples was calculated to look at potential significant differences. This analysis used samples of 1400 for pre and post covid. Secondly, we used a two-way ANOVA model to investigate our research question. The model will be using the categorical variables of region and newyear. The two-way model will also look at the overall pre and post covid flight delays. The model is also used to investigate regional differences and pre- and post-covid differences for the individual regions. For these analyses, a significance level of 0.05 was used to establish any differences. As for the 95% confidence intervals, we rounded the conventional z-score of 1.96 to a z-score of 2.

Diagnostics for this two-way ANOVA appear to check out as well. The six region groups by year type (pre- or post-COVID) were sampled randomly and independently. The condition of homogeneity among variances is also met since the Levene's Test for the interaction ($F=1.43$, $p = 0.1520$) gives evidence to conclude no signs of inequality among variances for the different region group means by year type. As for normality, a Shapiro-Wilk test could not be done due to the large sample size. The Kolmogorov-Smirnov ($D = 0.076798$, $p < 0.01$) does not give evidence to conclude the region group means by year type are normally distributed. We attempted to use a box-cox transformation on the data to fix this, but it did not work. Therefore, we just chose to be cautious with interpreting the results.

Results

Proportion and Confidence Interval

Looking into the sample proportion and 95% confidence interval calculations, we found that the proportion of pre-COVID flight delays was about 0.192 with a confidence interval of 0.171 to 0.213. The post-COVID flight delay proportion was 0.205 with a confidence interval of 0.183 to 0.227. Those results are also shown in Table 1 below.

Time Period	Proportion	Lower Bound	Upper Bound
Pre Covid	0.192	0.171	0.213
Post Covid	0.205	0.183	0.227

Table 1. Results of the Sample Proportion Calculation with 95% Confidence Intervals

The results above show that there is not a significant difference in the proportion of flight delays pre- and post-COVID. This shows that in the United States, that flight delays have not changed because of the COVID-19 pandemic.

ANOVA Model

A two-way ANOVA model was used to investigate the other questions of our study. When creating the model, we found that region had a significant effect on delays, with a p-value of 0.0063. The created variable newYear was also found significant with a p-value of < 0.0001 as well as the interaction of region and newYear which also had a p-value of < 0.0001. This model shows that region and which side of the pandemic had a significant effect in predicting the proportion of delays. The interaction also shows that the effect of region on the proportion of delayed flights differs depending on which side of the COVID-19 pandemic the observation is on.

Variable	P-value
Region	0.0063
NewYear	<0.0001
Region*NewYear	<0.0001

Table 2. Main Effects of the Two-Way ANOVA Model and their Significance

Overall Differences

The first question of interest is looking into the pre- and post-COVID differences with the ANOVA model. This is the same investigation we did with the proportions and confidence intervals. With the ANOVA model, we found that there is a significant difference (p-value <0.0001) in the proportion of pre- and post-COVID flight delays. The results show that there is an increase of 0.027 or about 2.7% of flight delays post-COVID when compared to pre-COVID. Pre-COVID has a proportion of 0.202 with a 95% confidence interval of 0.197 to 0.209 while post-COVID has a proportion of 0.229 with a 95% confidence interval of 0.222 to 0.236. Below is the table with the proportions and 95% confidence intervals.

Time Period	Proportion	Lower Bound	Upper Bound
Pre Covid	0.202	0.197	0.209
Post Covid	0.229	0.223	0.236

Table 3. Results of the Two-Way ANOVA Model for Proportion of Delayed Flights

There is a significant difference in Table 3 above, as the confidence intervals for pre-COVID and post-COVID do not overlap. Each one is distinctly outside each other's confidence bounds, showing a significant difference between the two groups.

Regional Differences

When looking at each region versus each other, the two-way ANOVA model found that the only significant difference was the North Central Region, which includes airports in Chicago, Detroit and Minneapolis for example, have significantly less delays on arrival flights compared to those in other regions besides the Northwest. The North Central region had a mean proportion of flight delays of 0.199, while the next closest region was the Northwest at 0.212, this is a difference of 1.3%. However, we did find that any other regions over the four-year study had a significant difference in delays of arriving flight.

Pre-COVID and Post-COVID Regional Differences

Since we are investigating overall differences in pre- and post-COVID flight delays, it only makes sense to investigate each region's pre and post differences. The trend that was observed is that airports in the northern U.S., did not have significant differences in pre- and post-COVID flight delays. In the North Central and Northeastern U.S., proportion of flight delays were virtually equal. In the Northwest U.S., flight delays increase by 0.028, however this was not considered significant (p-value = 0.0544). In the Southern U.S., the proportion of flight delays increased significantly in all three regions. Each region went up somewhere between 0.032 and 0.049 post covid compared to pre covid. This shows that in the Southern U.S., the COVID-19 pandemic may have had an impact on the proportion of flight delays. Below is the table of the differences in flight delays by region. Differences are calculated post-COVID minus pre-COVID.

Region	Difference	P-value
North Central	0.004	0.7049
Northeast	-0.002	0.8124
Northwest	0.028	0.0544
South Central	0.322	0.0080
Southeast	0.049	< 0.0001
Southwest	0.049	< 0.0001

Table 4. Regional Differences of Proportion of Flight Delays Post Covid Minus Pre Covid

Discussion

Our analysis investigated two different techniques for studying pre- and post-COVID proportion of flight delays to see if the pandemic has had a potential effect on delays or if the flight industry is back to normal. Our first investigation investigated the sample proportion and confidence interval calculations. These calculations resulted in a nonsignificant difference between pre- and post-COVID for flight delays, indicating that nothing has changed due to the pandemic. However, when using the ANOVA model to calculate the proportion of delayed flights, the model indicated that there was a significant difference in flight delays pre- and post-COVID. The model showed that there was a significant increase in delays. This could be attributed to the fact that the model is calculating mean proportion of delays while the first calculation found the true proportion based on weights of the strata. From this analysis, we recommend that there has been an uptick in the proportion of flight delays, but they are not significant.

When using the ANOVA model to investigate certain regions, our analysis found that the North Central U.S. has the smallest proportion of delayed flights compared to other regions in the U.S. The only region that it was not different from was the Northwest. North Central had the smallest proportion of delays pre- and post-COVID. We can conclude that the North Central region is the best area to fly in the U.S. The Northeast and Southeast have the largest proportion of delays and thus are the worst areas to fly to in the U.S. Even though there is not a significant difference with those regions compared to the others besides the North Central, we can conclude that they are the worst regions for arriving flights.

Lastly, investigating the differences in pre- and post-COVID for each region, we found that the northern regions of the U.S. did not have a significant difference in proportion of flight delays. In the North Central and Northeast regions, there were virtually no difference in the proportion of the delays. In the Northwest region, there was a slight increase, however this increase was determined to be nonsignificant. On the opposite end, the Southwest and Southeast had the largest and most significant increase in the proportion of flight delays. Each region had an increase of 4.9% in the proportion of delays. Overall, the whole Southern U.S. had significant increases in delays. From these findings we can conclude that flights to the Southern U.S. have become more delayed compared those in the North after the COVID-19 pandemic. The Southern U.S. airport had a 3-5% increase in delays while in the North, they stayed virtually the same.

Avenues for Further Research

In the pursuit of continuing our research to better understand flight delays across the United States, there are several avenues offering insights and opportunities for exploration. One avenue entails an analysis of the types of delays experienced in aviation, including carrier delays, security delay, weather delay, and mechanical delay. For instance, we can explore how seasonal factors, major events like conferences and sporting events, and peak travel periods affect flight delays. We could also analyze how certain weather patterns and types affect the likelihood of a delay for different airlines or domestic regions. Plus, airport security problems, such as evacuations, bad screening equipment, and threats to cybersecurity, can lead to flight delays, and the amount can vary by region. By further investigating such avenues, researchers can understand the underlining factors contributing to delays and target strategies for mitigation. This can even be paired with comparative analysis of airlines, where we compare the delay performance and operational characteristics of different airlines within the same region.

Additionally, delving into the economic impact of flight delays imposes a whole new idea of how detrimental delays can be. This study would entail quantifying the financial repercussions on various stakeholders such as airlines, passengers, airports, and local economies. By quantifying the costs, researchers can provide actionable insights into the economic landscape of aur travel. We can also investigate how flight delays can sometimes lead to extra expenses for passengers, such as being forced to stay at hotels or requiring extra food, that the customers or airlines pay for.

The development of predictive modeling for delay prediction would be another promising avenue of continued research. By using historical data alongside variables like weather forecasts, air traffic, and maintenance schedules, researchers can construct models capable of predicting delays based off potential disruptions. These models could help with enhancing operational efficiency and passenger satisfaction. This satisfaction from passengers can be used to analyze how certain amounts and patterns of delays for different airlines can affect how frequently an airline is chosen for use by fliers. These, in turn, can work hand in hand with analysis of customer reviews, feedback surveys, social media sentiment related to flight delays.

References

Airline On-Time Statistics and Delay Causes. (n.d.). *Bureau of Transportation Statistics*. Retrieved from https://www.transtats.bts.gov/OT_Delay/OT_DelayCause1.asp

Berman H.B., "*Sample Size Calculator*", [online] Available at: https://stattrek.com/survey-sampling/sample-size-calculator URL [Accessed Date: 4/29/2024].