Drivers of Flight Disruptions at Ronald Reagan Washington National Airport (DCA)

Georgetown University
DSAN 5100 Probabilistic Modeling and Statistical Computing
Instructor: Prof. Amineh Zadbood, Prof. Purna Gamage
Author: Adam He, Zizheng Wang, Ken Qin
December, 14th, 2025; Fall 2025

Table of Contents

1. Abstract

Flight delays and cancellations are common at airports operating near capacity. Dense flight schedules and limited runway space leave little flexibility when disruptions occur. Our research focuses on Ronald Reagan Washington National Airport (DCA) and examines the key factors contributing to flight disruptions. The analysis considers three key elements: weather conditions, holiday travel, and differences in airline operations. Flight-level data from the U.S. Department of Transportation were combined with weather observations from the National Oceanic and Atmospheric Administration and a structured holiday calendar. The dataset includes all flights operating at DCA from 2023 to 2024.

We combined descriptive analysis, statistical tests, and time-series models to compare cancellation rates, flight delay patterns, and changes over time under different conditions. Out of those factors, weather plays the most important role in disrupting flights. Rain is linked to higher cancellation rates, while snow causes more severe impacts, including widespread cancellations and longer flight delays. Flights during holidays face an increased risk of cancellation, but average departure delays remain similar to non-holiday. We conclude that airlines manage holiday congestion mainly by cancelling flights rather than allowing delays. Performance also differs across airlines, reflecting variations in scheduling practices and operational recovery strategies. Overall, flight disruptions at DCA result from the combined effects of environmental constraints and airline decision-making.

2. Keywords

Flight Disruptions; Flight Delays; Flight Cancellations; Airport Operations; Weather Effects; Holiday Effects; Airline Performance; Time-Series Analysis; Constrained Airports; Ronald Reagan Washington National Airport; DCA; Washington, DC

3. Introduction

3.1 Background and Motivation

In modern-day aviation, flight disruptions are an everyday issue, causing significant delays and cancellations and affecting various travelers within the U.S. Delays occur when a flight departs or arrives later than the scheduled time, which can be caused by many different factors, and cancellations occur when a planned flight does not operate. While small delays are often expected, frequent flight disruptions might dramatically reduce the reliability of air travel. Issues of flight delays or cancellations are more visible at busier airports, such as Reagan National Airport (DCA), where capacity is limited. In such situations, minor problems can lead to large amounts of flight delays or cancellations in a short period.

Flight delays and cancellations are not just a one-flight matter, as their impacts can extend beyond individual flights. When the regular schedule is disrupted, airlines may face greater operating pressures, and airports may experience flight congestion. Additionally, passengers may face longer travel times, leading to uncertain travel plans. Knowing the effects of flight disruptions, improving flight reliability is the key challenge for modern aviation systems.

3.2 Complexity of Flight Disruptions

Normally, flight disruptions have multiple causes rather than a single cause. For example, bad environmental conditions can limit safe operations, increased travel demand can strain available capacity, and airline scheduling decisions can increase system vulnerability. When these factors occur simultaneously, disruptions are more likely. Thus, this complexity makes it difficult to explain flight disruptions using simple reasoning.

The interaction between environment, demand, and airline operations is especially important at constrained airports. These airports operate with limited runway space and strict scheduling windows. There is little room for recovery when problems occur. Ronald Reagan Washington National Airport is a clear example of such a setting. Studying disruptions at this type of airport helps reveal how system pressure affects reliability.

3.3 Existing Understanding and Gaps

In general, when people think about flight delays and flight cancellations, weather conditions are frequently blamed, holidays are associated with increased risk of disruptions, and airline reliability is judged based on personal experience. While these explanations may seem reasonable, they are often based on limited observations of each individual and thus do not fully reflect the complexity of airline operations, especially at a busy and complex airport such as Ronald Reagan Washington National Airport (DCA). Personal experiences of individuals capture only a small portion of overall flight behavior and may be strongly influenced by chance. As a result, distinguishing real patterns from coincidence is difficult without a systematic analysis. In this research, we will use a data-driven approach to understand the true factors contributing to flight disruptions.

3.4 Research Focus and Scope

The goal of our research is to better understand the key factors driving flight disruptions at DCA by analyzing and identifying the broad patterns that explain their occurrence. Our analysis aims to clarify which factors play the most significant roles in shaping reliability and in causing flight disruptions. Due to the many factors causing flight disruptions, our study will focus on three complementary perspectives: environmental conditions, seasonal and holiday effects, and airline operations. Together, our research will provide a structured framework for studying disruption behavior without relying on a single explanation.

3.5 Practical Relevance

The result of our research, understanding flight disruptions, could have practical value for many different groups. Passengers could benefit from clearer expectations and more reliable travel planning; airlines can use these insights to improve flight scheduling; the DCA planners can better manage congestion; policymakers can design rules that improve stability. These benefits together are important for the long-term performance of the air transportation system around DCA.

---

4. Analysis and Statistical Methods

4.1 Data Sources

Our research integrated three different datasets, with two publicly available datasets from the U.S. government, to examine factors influencing flight disruptions at the Ronald Reagan Washington National Airport, also known as the DCA. Each of the following data sources was selected for its reliability and consistency.

4.1.1 U.S. Department of Transportation’s Bureau of Transportation Statistics

The flight data were obtained from the U.S. Department of Transportation’s Bureau of Transportation Statistics and consist of detailed flight-level operational records from 2023 to 2024. This dataset includes information on flight dates, airlines, flight numbers, origins, destinations, scheduled and actual departure and arrival times, and number indicators for delays and cancellations. It also contains the lengths of departure and arrival delays in minutes, along with the associated delay or cancellation type.

4.1.2 National Oceanic and Atmospheric Administration (NOAA)

The DCA weather data were obtained from the National Oceanic and Atmospheric Administration's Local Climatological Data Service. The dataset contains random time-stamped observations recorded at the airport weather station, including measures of temperature, humidity, visibility, wind conditions, and atmospheric pressure. The most important field in this data for our research is the description of sky conditions and current weather.

4.1.3 Contingency Table of Holiday Status

For the holiday table, we've decided to use LLM tools to generate a structured calendar table that classifies each date as either a holiday or a non-holiday. In addition to the holiday flag, we generated columns for the holiday type, holiday name, number of days to the next holiday, day of the week, and weekend flag. Various potential helpful columns that might benefit our research. The simplicity of such a holiday table supports clear interpretation and consistent integration with flight and weather data.

4.1.4 Data Cleaning Techniques

For the flight data cleaning process, since the data can only be downloaded by month, we first merged monthly CSV files into a single table and filtered the data to retain only flight data relevant to the study airport, including various irrelevant identifiers and redundant fields. Next, we standardized the raw date and time fields by converting flight dates from text data into proper date formats and transforming scheduled and actual times from HHMM values into usable datetime representations. For the blank values in the delay column, we manually added zeros. For columns with truly missing information, such as flight duration for cancelled flights, null values were preserved. Next, we cleaned various indicators to ensure consistent interpretation across the dataset and standardized categorical fields, such as airline carriers.

For the NOAA hourly weather data, we first converted the original timestamps to a uniform datetime format and floored them to whole hours to ensure temporal alignment, so we could create relationships with the other datasets using whole hours. Then, the relevant variables were retained, and meteorological fields stored as character strings were cleaned by removing non-numeric flags. Next, we've converted blanks and invalid entries to missing values and cast variables such as temperature, pressure, humidity, visibility, and wind to numeric types. Last, we converted the encoded weather and sky-condition fields into interpretable categorical variables, treating blank weather as no significant weather and blank sky conditions as unknown.

4.2 Weather Effect

4.2.1 Research Question: Weather Effects on Flight Cancellation Rates

In this section, we will conduct some statistical analysis to determine wether there's any relationship between the Flight cancellation rates and the adverse weather, in other words, we want to examine whether adverse weather conditions significantly increase the probability of flight cancellations. This analysis will give us some insights about how environment or weather factors constrain the airport's operations.

4.2.1.1 Statistical Methods

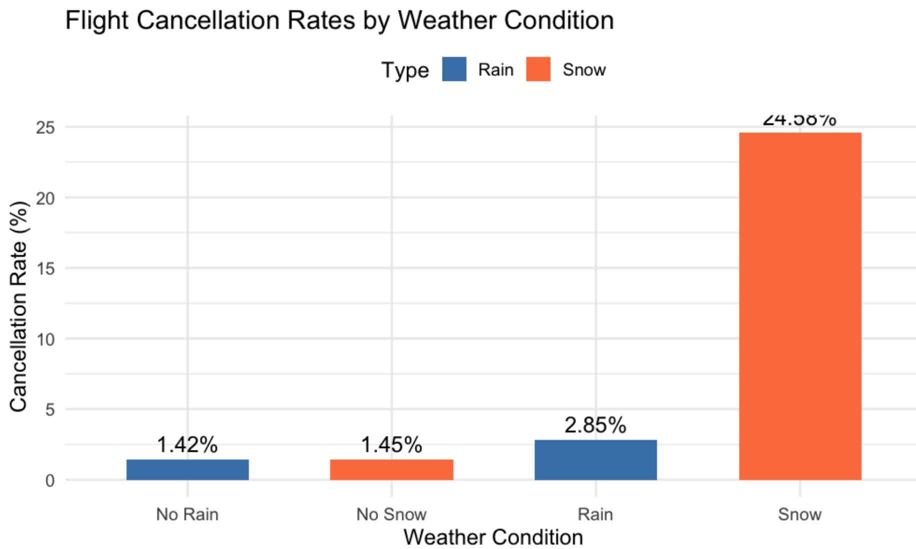
In order to assess the relationship between the weather conditions and the airport's operations, we decided to use the chi-square test which can be an ideal method to examine whether two factors are independent or not. We will use the cancellation status which is canceled or not and weather conditions which are adverse conditions or not. In this case, we have two different kinds of adverse weather which are snowy weather and rainy weather. Therefore, we conducted separate analyses for these two different weathers to identify condition-specific effects. The Chi-square test evaluates whether observed cancellation frequencies differ from those expected under the assumption that weather and cancellations are independent. A statistically significant result would indicate that weather systematically affects cancellation probability.

4.2.1.2 Research Result

The Chi-square analysis for rain versus no rain produced a test statistic of $\chi^2 = 258.36$ with 1 degree of freedom and a p-value nearly equal to zero ($p < 2.2e-16$). The cancellation rate during rainy conditions was 2.85%, compared to 1.45% during normal conditions. This represents a 1.4 percentage point increase in cancellation probability which is nearly doubling the rate. Then we can reject the null hypothesis that the rain condition and the cancellation rate are independent. The analysis for snow versus no snow revealed even more dramatic relationships. The Chi-square test yielded a statistic of $\chi^2 = 2{,}895.6$ with 1 degree of freedom and an extremely small p-value ($p < 2.2e-16$). The cancellation rate during snowy conditions was 24.58%, compared to 1.45% during non-snowy conditions. This represents a 23.13 percentage point increase, indicating that snow raises cancellation risk by approximately 17 times relative to baseline. We can also reject the null hypothesis based on these findings that the snowy weather and the cancellation rate are independent of each other.

4.2.1.3 Interpretation

The results give us a lot of statistical evidence that adverse weather significantly increases flight cancellation probability at DCA including the rainy weather and the snowy weather. Rain approximately doubles cancellation risk, reflecting moderate operational constraints during wet conditions. However, snow produces a dramatically more severe impact over the airport operation system, with nearly one in four flights being canceled during snowy weather. These findings highlight the vulnerability of airport operations to winter weather events and suggest that snow represents the most disruptive weather for DCA flight reliability. This can be intuitively interpreted as well that during the snowy weather, the airport needs to handle lots of complex conditions like the icy runway or visibility issues, which are the major issues affecting the planes.



4.2.2 Research Question: Weather Effects on Severe Delay Rates

In this section, we will discuss the relationship between the adverse weather and severe delay rates. We will have some tests on deciding if the adverse weather will increase the probability of severe delay. In this case, we define the severe delay as a departure delay of 15 minutes or longer, based on the FAA's on-time performance standard that considers flights departing within 15 minutes of schedule as on-time. Unlike the cancellation analysis, this question focuses on flights that do operate, examining whether weather conditions systematically degrade on-time departure performance.

4.2.2.1 Statistical Methods

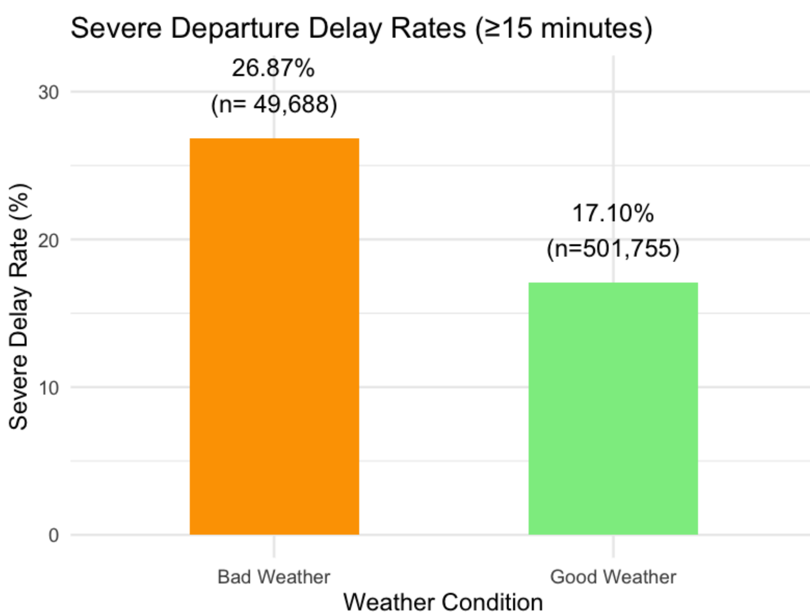
To compare severe delay rates between adverse weather and good weather conditions, we conducted a two-proportion z-test. This test is appropriate for comparing proportions across two independent groups and evaluates whether the observed difference in severe delay rates is statistically significant. The analysis was restricted to non-canceled flights to avoid confounding cancellations with delay behavior. We calculated the proportion of flights experiencing delays of 15 minutes or more under bad weather conditions versus good weather conditions and tested the null hypothesis that these proportions are equal.

4.2.2.2 Research Result

The two-proportion z-test produced a Chi-square statistic of 1323.7 with 1 degree of freedom and a p-value effectively equal to zero ($p < 2.2e\text{-}16$). Among non-canceled flights, the severe delay rate under bad weather conditions was 26.87%, compared to 17.10% under good weather conditions. This represents a 9.77 percentage point increase in absolute terms and a 57.1% relative increase in severe delay probability. In total, 14966 out of 55726 flights departing in bad weather experienced severe delays, compared to 85711 out of 501155 flights departing in good weather. Based on the results we can reject the null hypothesis that the adverse weather do not increase the probability of severe delays.

4.2.2.3 Interpretation

The results demonstrate that the adverse weather not only increases cancellation risk but also significantly affects the on-time performance of flights that actually operate. The 57% relative increase in severe delay probability indicates that weather conditions impose substantial operational delay even when flights are not canceled. This finding suggests that passengers flying during adverse weather should anticipate longer delays, and airlines may face greater challenges maintaining schedule integrity.



4.2.3 Research Question: Flight Cancellation pattern along with time series analysis

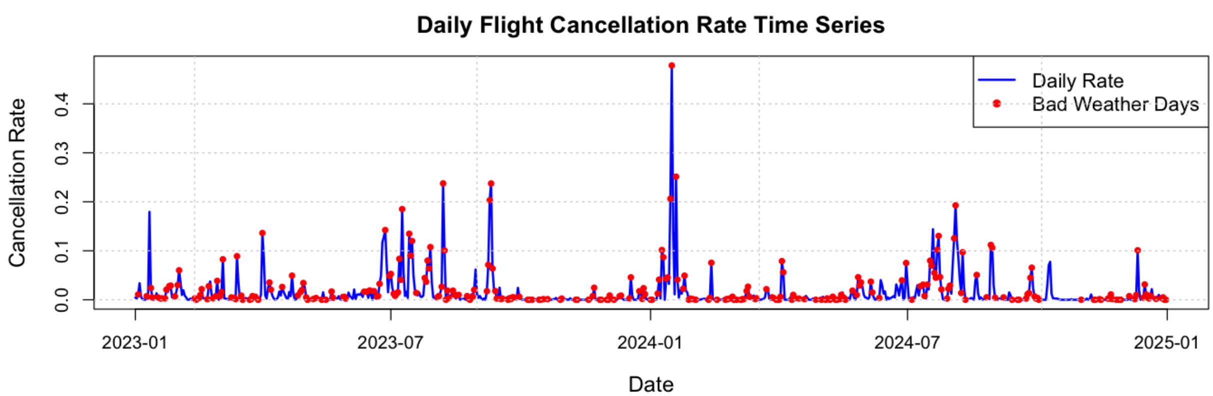
The third research question of the environment section, we examine whether daily cancellation rates exhibit systematic or what we called patterned spikes during weather events and whether these patterns can be modeled using time-series analysis. Rather than analyzing individual flights, this question aggregates data to the daily level to identify temporal dynamics and weather related disruption patterns over time.

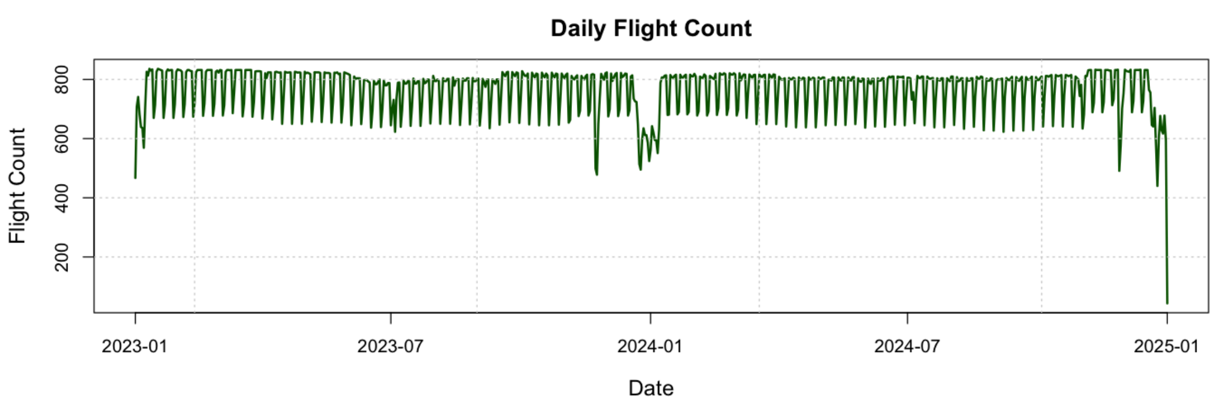
4.2.3.1 Statistical Methods

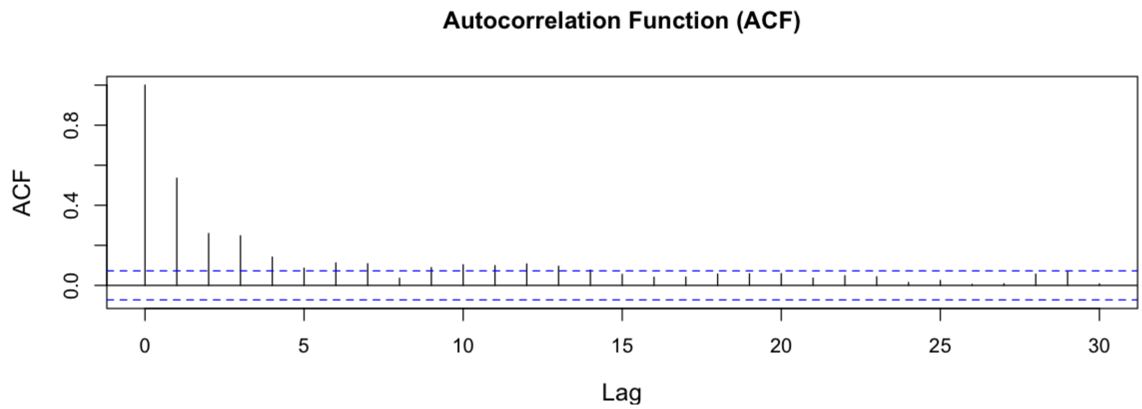
We applied ARIMA time-series models to analyze daily cancellation rates. Daily data were constructed by aggregating flights by date and calculating the proportion of flights canceled each day, along with indicators of whether adverse weather occurred that day. A two-sample t-test was first conducted to compare average cancellation rates on bad weather days versus good weather days. And then we fit our ARIMA model to the daily cancellation rate time series to assess autocorrelation and patterns. The model specification was selected using the auto.arima function, which evaluates candidate models based on the Akaike Information Criterion (AIC). Stationarity was assessed using the Augmented Dickey-Fuller (ADF) test, and autocorrelation structure was examined using ACF and PACF plots.

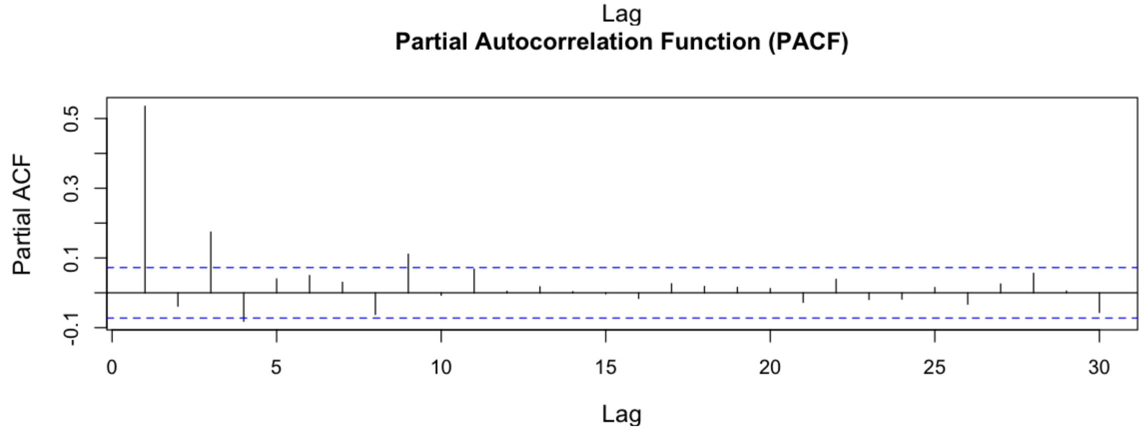
4.2.3.2 Research Result

The daily aggregation produced 730 days from January 2023 to December 2024. Of these, 218 days (29.9%) experienced adverse weather conditions, while 512 days (70.1%) had good weather. The average daily cancellation rate on bad weather days was 2.65%, compared to 0.82% on good weather days. A Welch two-sample t-test yielded a t-statistic of 9.91 with 318 degrees of freedom and a p-value effectively equal to zero ($p < 2.2e-16$), providing strong evidence that cancellation rates are significantly higher on adverse weather days. The model achieved an AIC of -3,053.84 and BIC of -3,021.67, which indicates a strong fit over the data.









4.2.3.3 Interpretation

The ARIMA time-series analysis reveals that daily cancellation rates exhibit both short-term temporal persistence and weekly seasonality. The three autoregressive terms indicate that cancellation rates on a given day depend on rates from the previous three days, reflecting multi-day weather systems and operational recovery delays. The positive AR coefficients (AR(1)=0.091, AR(2)=0.123, AR(3)=0.146) suggest mild positive autocorrelation, meaning elevated cancellation rates tend to persist across several days rather than be affected immediately. The MA(1)=0.494 captures the delayed effects of random shocks like random weather events, which influence cancellation rates beyond the day of occurrence. The SAR(1)=0.088 with weekly period reveals weekly patterns in cancellation behavior. We assume that this kind of pattern is likely reflecting operational scheduling patterns and day-of-week effects like Monday has its own behavior in terms of operational scheduling etc. Visualization of the time series clearly shows cancellation spikes aligned with bad weather days, confirming that weather events are a primary driver of temporal variation.

4.3 Seasonal (Holiday) Effects

4.3.1 Research Question: Holiday Effects on Flight Cancellation Rates

In this study, we will determine whether there is a true difference in mean flight cancellation rates between holiday and non-holiday periods at Ronald Reagan Washington National Airport (DCA). With the result of this question, we are able to see whether holidays introduce additional seasonal stress, leading to a higher likelihood of cancellations that are not caused by daily variations.

4.3.1.1 Statistical Methods

To examine the true mean difference between holiday status and flight cancellation, we've decided to use a Chi-Square Test of Independence. By applying the Chi-Square Test of Independence, we will be able to assess whether flight cancellations occur independently of whether flights occur on a holiday or a non-holiday. The test is appropriate for identifying statistically significant differences in categorical outcomes across two distinct time periods.

4.3.1.2 Research Result

After running the Chi-Square test, it yielded a test statistic of 47.52 with a p-value of 5.45×10^{-12}. The cancellation rate during holidays was 2.23%, compared to 1.52% during non-holiday periods. In absolute terms, 328 cancellations occurred out of 14,714 holiday flights, and 8,275 cancellations occurred out of 545,332 non-holiday flights. Since the p-value is well below the 0.05 significance threshold, we reject the null hypothesis of independence. The Chi-Square analysis provides strong evidence of a relationship between holiday status and flight cancellations. Cancellation rates during holiday periods are significantly higher than those during non-holiday periods.

4.3.1.3 Interpretation

The result above indicates that holiday periods are associated with increased cancellation risk at DCA. The explanation is that increased travel demand and reduced scheduling flexibility during holidays are likely to constrain DCA's ability to recover from disruptions. When operational issues arise, airlines may be more inclined to cancel flights rather than attempt a recovery. As a result, holidays represent periods of heightened vulnerability in airport operations.

4.3.2 Research Question: Holiday Effects on Average Departure Delays

In this second seasonal research question, we will investigate whether average departure delays differ between holiday and non-holiday periods. Unlike the last research question, this question will focus on delay duration rather than cancellation frequency and evaluate whether holiday congestion leads to longer delays for flights that do operate.

4.3.2.1 Statistical Methods

To compare the average departure delays between holiday and non-holiday flights, we've decided to conduct a Welch two-sample t-test. We are using the Welch two-sample t-test because it is suitable when group sizes differ, and the variances may differ across groups. Also, this test evaluates whether the mean departure delay differs significantly between the two time periods.

4.3.2.2 Research Result

The results from the Welch two-sample t-test indicate that there is no statistically significant difference in average departure delays between holiday and non-holiday flights. The test produced a t-statistic of -0.46 with 7,552 degrees of freedom and a p-value of 0.6445. The mean departure delay for holiday flights is 10.48 minutes, and for non-holiday flights is 10.78 minutes. The 95% confidence interval for the difference in mean delays ranged from -1.56 to 0.97 minutes. Since the confidence interval includes zero and the p-value is greater than the significance threshold of 0.05, we failed to reject the null hypothesis.

4.3.2.3 Interpretation

Based on the t-test results above, holiday periods do not increase the duration of departure delays at DCA. Combining the results from the first research question on holidays and differences in cancellation rates, we conclude that while cancellations increase during holidays, flight delays tend to remain at the same level as on non-holiday days. This conclusion is consistent with the previous interpretation that airlines manage holiday congestion through cancellations rather than delaying flights.

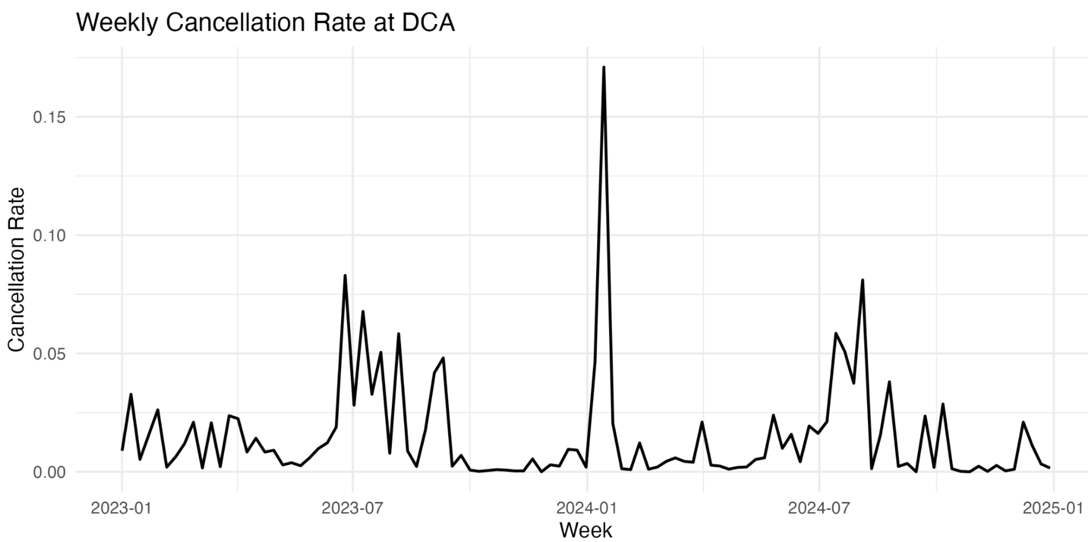
4.3.3 Research Question: Seasonal Patterns in Flight Disruptions Around Holidays

For the last seasonal research question, we will examine whether flight disruptions have recurring seasonal patterns around the periods of major holidays. Instead of focusing on isolated holiday effects, we will evaluate whether disruptions follow consistent and repeatable temporal structures aligned with the holiday calendar.

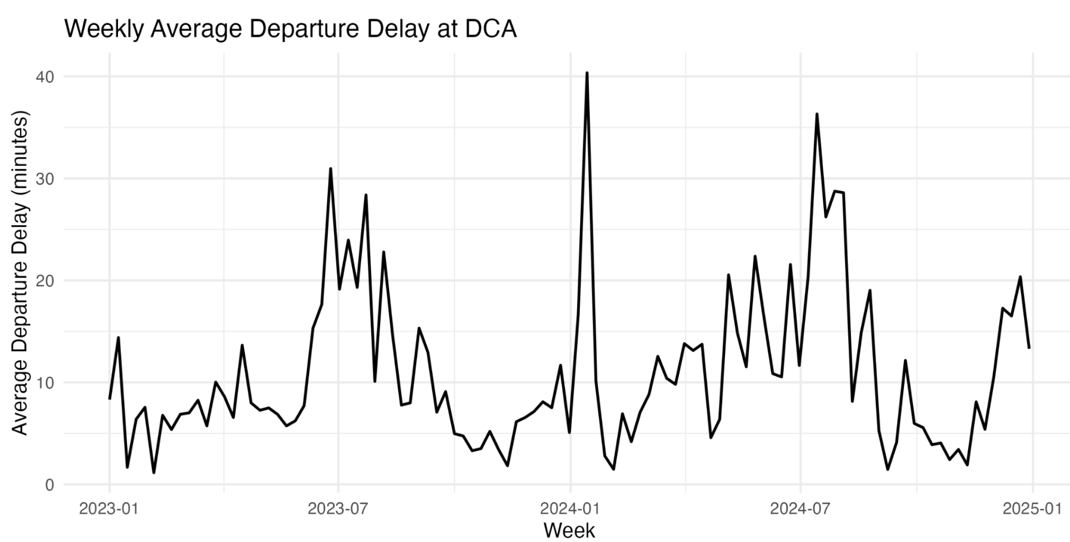
4.3.3.1 Statistical Methods

For this research question, the ARIMA time-series modeling was used to analyze temporal patterns in flight disruptions. By using this method, we can determine whether current disruption levels depend on past values and whether seasonal behavior is present. In total, we will conduct two separate analyses, one for cancellations and one for departure delays. We will examine if there are any differences in their underlying temporal dynamics.

4.3.3.2 Research Result



The ARIMA time series of weekly cancellation rates contains spikes and short-term fluctuations instead of a stable seasonal pattern. The auto.arima function from R selected an ARIMA(1,0,0) model with a non-zero mean, meaning that the present cancellation rate depends on the value of the previous week. With this model, the estimated AR(1) coefficient is 0.28, and the estimated mean cancellation rate is around 1.5%, indicating a mild positive autocorrelation. This suggests that increases or decreases in cancellation rates one week tend to be followed by similar movements the next week, though the relationship is not strong. The model fit statistics AIC and BIC are around -493.77 and -485.81, respectively, indicating a well-fitting specification. Combined with the residual diagnostics, which show low residual autocorrelation, we conclude that short-term persistence is present, but the cancellation rate does not follow a strong seasonal pattern, meaning future rates are better predicted by recent activity rather than by repeating yearly cycles.



The ARIMA time series of weekly average departure delay shows higher volatility than the cancellation rates, and it exhibits a gradual upward trend. The auto.arima function from R selected an ARIMA(0,0,1)(0,1,0)[52] model with drift, indicating a weak moving-average dependence and a weak long-term trend component. The model returned an estimated MA(1) coefficient of 0.26, while the drift term suggests a slight upward trend in average delays over the long term. The AIC and BIC are approximately 377.25 and 383.16, respectively, which support this specification. Although some clustering of higher delays is visible during peak seasons, the time series lacks consistent, repeating spikes aligned with major holidays, suggesting a weak seasonal structure.

4.3.3.3 Interpretation

From the two ARIMA time-series results and visualization above, we are able to conclude that holiday effects on flight disruptions at DCA do not exhibit seasonal patterns. Both the cancellation rate and departure delays showed several spikes, but none were recurring cycles that aligned with the holiday calendar. The cancellation series showed only mild weekly persistence, suggesting that increased cancellation risk reflects recent operational conditions rather than predictable annual patterns. The departure delay series shows a weak annual pattern and a slight long-term upward trend, but similarly, no recurring holiday patterns. With these results and conclusions above, it implies that holidays might be a temporary stress amplifier within the system instead of a driver of enduring seasonal dynamics.

4.4 Airline Operational Effects

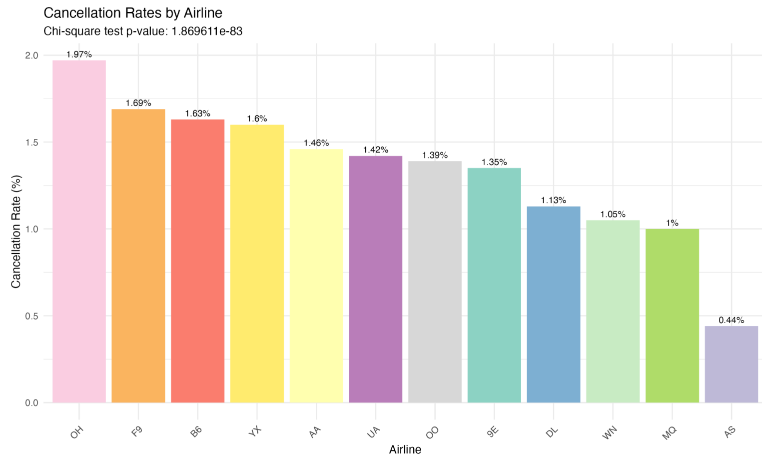
4.4.1 Research Question: Differences in Airline Cancellation Rates

The first research question examines whether major airlines differ in their flight cancellation rates and whether cancellation behavior is associated with airline carriers. Specifically, the goal is to determine if cancellations occur independently of airline choice or if certain carriers are systematically more or less likely to cancel flights.

4.4.1.1 Statistical Methods

To address this question, a Chi-square test of independence was conducted using a contingency table that cross-classified airline carriers and flight cancellation status (canceled versus not canceled). This nonparametric test is well-suited for categorical outcomes and evaluates whether observed cancellation frequencies differ from those expected under the assumption of independence. The test compares the observed counts of canceled flights for each airline to the counts that would be expected if all airlines had the same cancellation rate.

4.4.1.2 Research Result



The Chi-square analysis produced a test statistic of $\chi^2 = 421.26$ with 11 degrees of freedom, yielding a p-value effectively equal to zero. This extremely small p-value indicates that the observed differences in cancellation frequencies across airlines are far larger than would be expected due to random sampling variation alone.

4.4.1.3 Interpretation

Based on these results, the null hypothesis of equal cancellation rates across airlines is decisively rejected. There is overwhelming statistical evidence that airline carrier and cancellation status are not independent, meaning that airline choice significantly affects the likelihood of a flight being canceled. From a practical perspective, this finding suggests that travelers, airlines, and policymakers should consider carrier-specific operational practices when evaluating reliability, as cancellations are not evenly distributed across airlines.

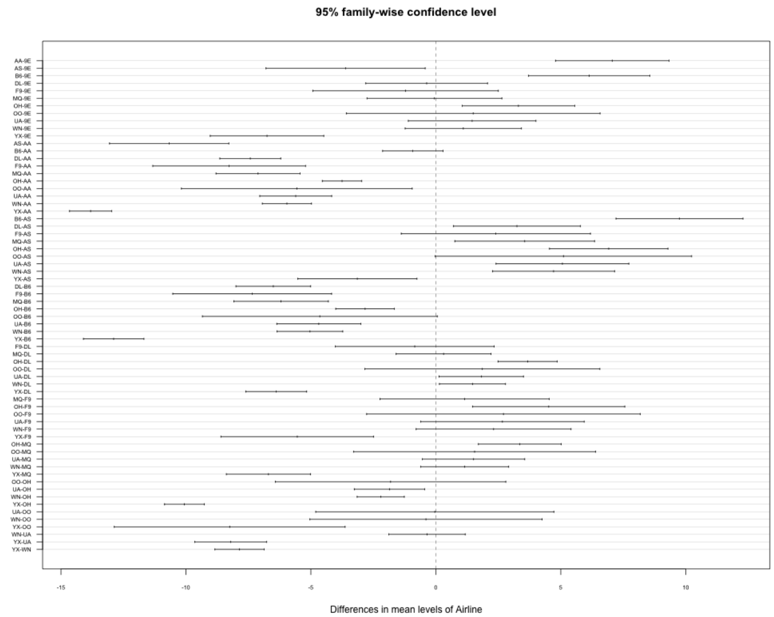
4.4.2 Research Question: Differences in Mean Arrival Delay Across Airlines

The second research question investigates whether major airlines differ in their average arrival delays, focusing on whether differences in central tendency across carriers reflect meaningful variation in operational performance. The objective is to determine whether observed differences in mean delays are statistically significant or simply the result of random variation.

4.4.2.1 Statistical Methods

A one-way Analysis of Variance (ANOVA) was employed to compare mean arrival delays across multiple airlines simultaneously. ANOVA decomposes total variation in arrival delays into variation attributable to differences between airlines and variation within airlines. Because ANOVA alone only indicates whether at least one group differs, post-hoc pairwise comparisons were subsequently used to identify which specific airlines exhibited significantly better or worse delay performance.

4.4.2.2 Research Result



The ANOVA results led to rejection of the null hypothesis of equal mean arrival delays. The ratio of between-airline variance to within-airline variance was approximately 441, indicating that differences in average delay between airlines are dramatically larger than would be expected due to random fluctuation. Among the carriers analyzed, American Airlines exhibited the highest mean arrival delay of 12.77 minutes, with a 95% confidence interval ranging from 7.58 to 18.95 minutes, highlighting both elevated delay levels and notable variability.

4.4.2.3 Interpretation

These findings demonstrate that airline delay performance varies significantly across carriers and that the magnitude of these differences is substantial rather than trivial. The large variance ratio underscores that airline-specific operational factors play a major role in determining arrival delays. Follow-up comparisons further reveal that while some airlines consistently perform better in terms of on-time arrivals, others, such as American Airlines, experience systematically worse delay outcomes.

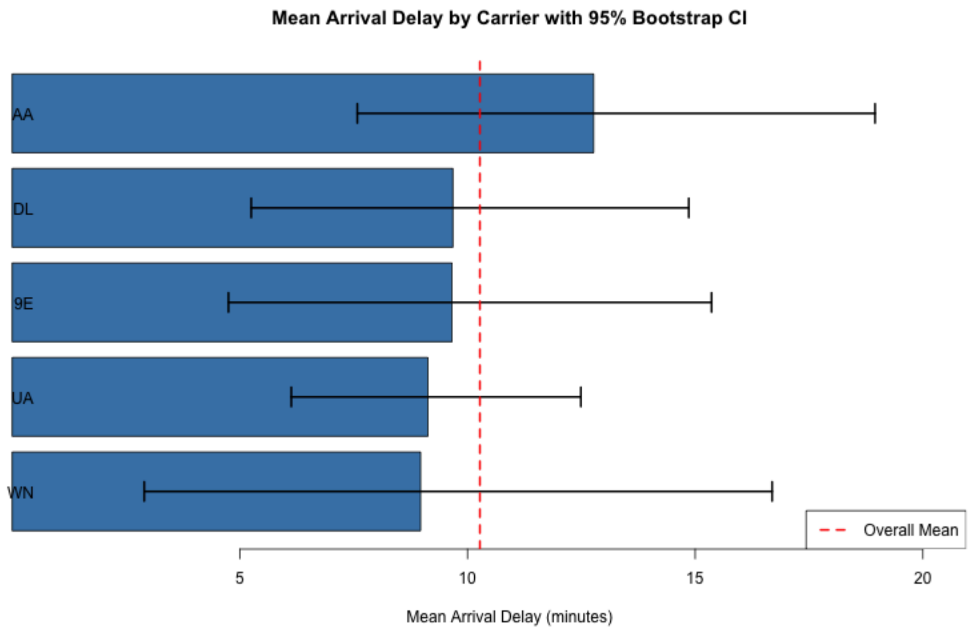
4.4.3 Research Question: Airline Delay Performance Bootstrap Resampling

The third research question explores whether bootstrap resampling techniques can be used to obtain robust confidence intervals for airline delay performance. The goal is to assess not only point estimates of mean delay but also the stability and uncertainty surrounding those estimates, particularly when traditional parametric assumptions may be questionable.

4.4.3.1 Statistical Methods

A nonparametric bootstrap approach was applied to arrival delay data by repeatedly resampling the observed data with replacement and recalculating mean delays for each airline. This method does not rely on assumptions of normality and provides an empirical distribution of the mean delay, from which confidence intervals can be constructed.

4.4.3.2 Research Results



The bootstrap analysis produced confidence intervals that clearly differentiated airline performance. American Airlines maintained a mean delay of approximately 12.77 minutes, with confidence intervals that did not overlap with those of some competing carriers. Additionally, variation in confidence interval widths across airlines revealed differences in performance stability, with some airlines exhibiting narrow intervals (indicating consistent delays) and others displaying wider intervals (indicating greater variability).

4.4.3.3 Interpretation

Bootstrap resampling offers a robust and informative perspective on airline delay performance by explicitly quantifying uncertainty. The results reinforce earlier findings that American Airlines experiences poorer delay performance while also demonstrating greater inconsistency relative to other carriers. More broadly, the bootstrap approach highlights the value of resampling methods for assessing operational reliability, as they provide insights into both average performance and variability across airlines.

---

5. Results

5.1 Environmental Effects on Flight Disruptions

This section presents our statistical findings on how adverse weather like rainy weather and snowy weather affect flight disruptions at DCA. We examined three complementary dimensions of weather impact: the likelihood of cancellations under different weather conditions, the probability of severe delays among operating flights, and the temporal dynamics of cancellation rates during weather events. Across all three analyses, the results consistently demonstrate that weather represents a major determinant of operational reliability and that its effects extend beyond isolated incidents to create sustained disruption patterns.

5.1.1 Weather and Cancellation Likelihood

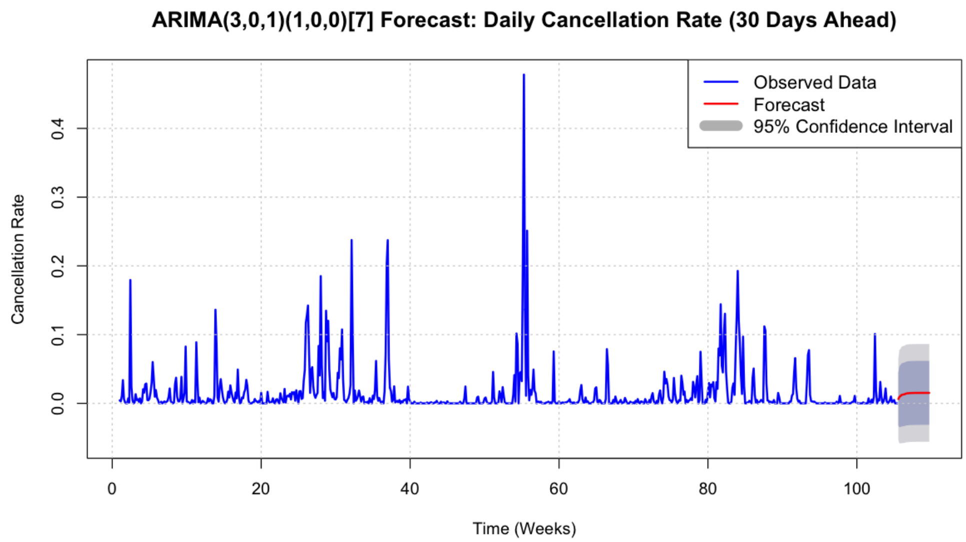
Weather conditions exhibit a powerful and statistically significant association with flight cancellation probability, with the magnitude of impact varying substantially by weather type. Rain increases the cancellation rate from a baseline of 1.45% to 2.85%, representing a near-doubling of cancellation risk. While this effect is statistically significant and operationally meaningful, it is modest compared to the impact of snow. During snowy conditions, the cancellation rate rises dramatically to 24.58%, meaning that nearly one in four flights is canceled when snow occurs. This increase relative to good weather conditions identifies snow as the most disruptive weather factor at DCA. The consistency and magnitude of these effects across thousands of flights indicate that weather-related cancellations are not isolated events but rather systematic responses to environmental constraints that limit safe and efficient airport operations.

5.1.2 Weather and Severe Delays

Beyond cancellations, adverse weather significantly causes the on-time performance of flights that do operate to perform bad. Among non-canceled flights, the probability of experiencing a severe departure delay increases from 17.10% to 26.87%. This increase represents a 57% relative increase in severe delay risk. The p-value also gives us a significant result that confirms that this effect is robust. The finding that more than one in four flights experiences severe delays during adverse weather underscores the dual nature of weather impact: not only more flights are canceled, but also those flights that do not get canceled are facing severe delays. This pattern suggests that weather degrades multiple aspects of airport operational capacity, including de-icing and air traffic control throughput.

5.1.3 Time-Series Evidence of Weather Impact

The time-series analysis of daily cancellation rates provides evidence that weather affects both short-term persistence and weekly seasonality. On days with adverse weather, the average cancellation rate is 2.65%, compared to 0.82%. Visualization of the daily cancellation rate time series reveals clear spikes aligned with weather events, and the ARIMA(3,0,1)(1,0,0)[7] model successfully captures both the short-term persistence and weekly patterns in cancellation behavior. The positive autoregressive coefficients indicate positive autocorrelation, which means that the rise of cancellation rates tend to persist across multiple days. The moving-average component suggests that weather shocks have delayed effects, as operational recovery requires time even after weather conditions improve. The SAR(1)=0.088 captures recurring weekly patterns, revealing that cancellation risk varies systematically by day of week, likely due to operational scheduling practices and demand fluctuations. These findings reveal the complexity of managing disruptions in a constrained operational environment like DCA.



5.2 Seasonal (Holiday) Effects on Flight Disruptions

In this section, we've looked at how flight disruptions at Ronald Reagan Washington National Airport (DCA) differ between holiday and non-holiday periods. In total, we've focused on three different aspects, including flight cancellation frequency, flight delays length, and holiday-related seasonal patterns. The results of these aspects indicate that holidays are temporary stress points instead of seasonal drivers of disruption. Holiday periods increase the likelihood of cancellations, but they do not create lasting or recurring patterns in flight delays or cancellations over time.

5.2.1 Holiday vs Non-Holiday Cancellations

Based on our hypothesis testing, flight cancellations at DCA occurred more frequently during holiday periods than on regular days. During the holiday periods, about 2.23% of flights were cancelled, compared to 1.52% on non-holiday days. 328 flights were cancelled during holiday periods, and 8,275 cancellations occurred on non-holiday days. The pattern of cancellation rate is consistent across all holiday periods, even though the overall percentage difference seems pretty small. The higher cancellation rate during holiday periods indicates that flights scheduled during holidays face a noticeably greater risk of being cancelled.

5.2.2 Holiday Effects on Delay Duration

Although the previous result indicates that cancellation rates are higher during holidays, the flight delay length remains similar across holiday and non-holiday periods. On average, departing flight delays on holidays are approximately 10.5 minutes, compared to 10.8 minutes on non-holiday days. This difference in average flight delays between these two periods is too small to be statistically significant. With this result, we conclude that holiday travel does not lead to longer flight delays. Instead, congestion during holidays appears to be managed through cancellations rather than by allowing delays to grow longer.

5.2.3 Seasonal Time-Series Patterns

When we applied ARIMA time-series models to flight disruptions, holiday effects did not form a clear or repeating seasonal pattern. Both cancellation rates and delays showed several spikes, but these spikes do not occur consistently around the same holidays each year; instead, they might be caused by weather during that holiday period. However, cancellation rates do tend to reflect short-term conditions, meaning that higher cancellation levels in one week are sometimes followed by slightly higher levels the next week. These changes do not repeat on a predictable annual cycle tied to the holiday calendar. Just like cancellation rates, departure delays fluctuate over time and show a mild long-term upward trend, but they do not consistently increase during holiday periods.

5.3 Airline Operational Effects on Flight Disruptions

Across all three research questions, the analysis provides strong and consistent evidence that airline operational practices play a significant role in flight disruptions, including both cancellations and delays. Statistical tests reveal that variation in flight outcomes is not random but is systematically associated with airline carrier. Together, the Chi-square analysis, ANOVA results, and bootstrap resampling demonstrate that airline choice meaningfully influences both the likelihood of flight cancellation and the magnitude and stability of arrival delays. These findings confirm that airline-level operational differences are a key determinant of passenger travel reliability.

5.3.1 Airline Differences in Cancellation Rates

The Chi-square test of independence provides overwhelming evidence that cancellation rates differ significantly across major airlines. The extremely large test statistic ($\chi^2 = 421.26$, df = 11) and near-zero p-value indicate that cancellation behavior is strongly associated with airline carriers rather than occurring uniformly across airlines. This result suggests that some airlines are systematically more prone to cancel flights than others, highlighting the importance of carrier-specific operational policies and capacity management when evaluating flight reliability.

5.3.2 Airline Differences in Delay Performance

The analysis of mean arrival delays reveals substantial and statistically significant differences in delay performance across airlines. The one-way ANOVA indicates that between-airline variation in arrival delays is dramatically larger than within-airline random variation, underscoring the practical significance of these differences. American Airlines, in particular, exhibits the highest mean arrival delay, reflecting comparatively poorer on-time performance. These findings demonstrate that arrival delays are not evenly distributed across carriers and that airline-specific operational efficiency plays a critical role in determining passenger delay experiences.

5.3.3 Bootstrap Confidence Intervals for Airline Delay Performance

Bootstrap resampling further strengthens the analysis by providing robust confidence intervals for airline delay performance that quantify both average delays and uncertainty. The bootstrap results confirm that American Airlines' delay performance is not only worse on average but also more variable than that of some competing carriers, as evidenced by wider confidence intervals. More broadly, the bootstrap approach highlights meaningful differences in performance stability across airlines, offering a nuanced understanding of operational reliability beyond simple point estimates. This resampling-based perspective reinforces the conclusion that airline operational effects on delays are both statistically significant and practically relevant.

---

6. Conclusions

6.1 Key Findings

In this research, we analyzed flight disruption factors at the Ronald Reagan Washington National Airport (DCA), including environmental conditions, airline operational practices, and holiday or seasonal factors. Among all the factors we researched, weather has the strongest and most immediate impact on flight delays and cancellations. Weather conditions that are not ideal for flights significantly increase the risk of cancellation, with the flight cancellation rate doubling on days of rain and 17 times on days of snow. The results of the statistical methods show that weather impacts on flight delays or cancellations are greater than those associated with holidays or airline operational factors.

For our research on holidays, we found that flights scheduled during holidays are more likely to be cancelled than those on non-holiday days. Although cancellations are more likely, flight departure delays tend to remain similar. This result might indicate that during holidays, when operational pressure at DCA is higher, airlines tend to manage pressure by cancelling flights rather than allowing delays.

For airline operational factors, we conclude that the difference in cancellation rates and delay rates across carriers is statistically significant. The results show that some carriers experienced higher cancellation frequencies and longer delays than others. This might be because of differences in airline scheduling density, operational practices, and recovery approaches. Once disruptions occur, airline decisions influence whether flights are cancelled or delayed and how quickly operations recover. When considered alongside the effects of weather and holidays, these findings show that environmental conditions constrain overall system performance, while airline operations determine how disruption risk is distributed across flights and passengers.

6.2 Implications for Stakeholders

The findings of our research could have clear implications for passengers, airlines, and airport operations at Ronald Reagan Washington National Airport (DCA). For passengers, weather conditions pose the greatest risk to travel reliability. Adverse weather significantly increases the likelihood of flight cancellations, with passengers facing nearly double the cancellation risk during rain and an extremely elevated risk during snow events. Holidays also increase cancellation risk, though flights that operate on holidays experience delays similar to those on non-holiday days. These results suggest that passengers should plan for greater uncertainty during bad weather and holiday periods by allowing additional flexibility, closely monitoring forecasts, and considering airline reliability when booking flights.

For airlines operating at DCA, our results highlight the importance of proactive operational planning under adverse weather or peak-demand conditions. Severe weather appears to overwhelm recovery capacity, causing an increase in cancellations rather than longer delays. During holidays, airlines similarly manage congestion by cancelling flights instead of extending delays, reflecting limited flexibility at this constrained airport. Also, large differences in cancellation and delay rates across airlines indicate that internal scheduling and recovery strategies play a major role in determining reliability. More conservative scheduling and stronger weather-contingent planning may help to improve performance during periods with higher risks.

For airport operations, our research indicates that disruptions at DCA are primarily driven by short-term operational stress instead of seasonal patterns. Days with bad weather events have the greatest impact on flights, and holidays serve as temporary stress points. This suggests that airport planners may benefit more from flexible response strategies than from fixed seasonal planning. Enhanced coordination during bad weather could help reduce the severity of flight disruptions and improve passenger experiences.

6.3 Limitations

During the process of our research, we experienced several limitations related to data availability and the scope of analysis. Our analysis relied on publicly available flight and weather datasets, which did not contain any detailed operational factors such as airline crew availability, aircraft rotations, or gate assignments. These unobserved factors above have a very high chance of influencing cancellation and flight delays, but cannot be directly measured in this study. Also, for each of the holidays, we've treated levels of travel demand and operational intensity as the same. We did not distinguish between different holiday seasons.

We also had several limitations related to modeling assumptions. In our analysis, we assumed that observed relationships between flight disruptions and their drivers are relatively stable over time, even though, in reality, airline operations, staffing levels, and regulatory environments may significantly influence flight disruptions. Time-based patterns reflect average behavior and may not fully capture rare but extreme disruption events. As a result, while the findings provide strong evidence of broad patterns, they should not be interpreted as precise predictions for individual flights.

---

7. References

7.1 Data Sources

U.S. Department of Transportation (DOT), Bureau of Transportation Statistics (BTS). (n.d.). *Airline On-Time Performance / On-Time Arrival Data (Reporting Carrier On-Time Performance).* Transtats. https://www.transtats.bts.gov

National Oceanic and Atmospheric Administration (NOAA), National Centers for Environmental Information (NCEI). (n.d.). *Local Climatological Data (LCD).* https://www.ncei.noaa.gov/access/search/data-search/local-climatological-data (Weather station used: Washington/Reagan National (DCA))

U.S. Office of Personnel Management (OPM). (n.d.). *Federal Holidays.* https://www.opm.gov/policy-data-oversight/pay-leave/federal-holidays/

7.2 Methodology Texts and Documentation

Agresti, A. (2019). *An Introduction to Categorical Data Analysis* (3rd ed.). Wiley.

Montgomery, D. C. (2020). *Design and Analysis of Experiments* (10th ed.). Wiley.

Efron, B., & Tibshirani, R. J. (1994). *An Introduction to the Bootstrap.* Chapman & Hall/CRC.

Hyndman, R. J., & Athanasopoulos, G. (2021). *Forecasting: Principles and Practice* (3rd ed.). OTexts. https://otexts.com/fpp3/

NIST/SEMATECH. (n.d.). *e-Handbook of Statistical Methods.* https://www.itl.nist.gov/div898/handbook/

7.3 Software and Packages (R)

R Core Team. (n.d.). *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria. https://www.r-project.org

Wickham, H., et al. (2019). *Welcome to the tidyverse. Journal of Open Source Software, 4*(43), 1686. https://doi.org/10.21105/joss.01686

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis.* Springer. https://ggplot2.tidyverse.org

Grolemund, G., & Wickham, H. (2011). *Dates and Times Made Easy with lubridate. Journal of Statistical Software, 40*(3), 1–25. https://www.jstatsoft.org/article/view/v040i03

Robinson, D., & Hayes, A. (n.d.). *broom: Convert Statistical Objects into Tidy Tibbles.* CRAN. https://CRAN.R-project.org/package=broom

Kassambara, A. (n.d.). *rstatix: Pipe-Friendly Framework for Basic Statistical Tests.* CRAN. https://CRAN.R-project.org/package=rstatix

Fox, J., & Weisberg, S. (2019). *An R Companion to Applied Regression* (3rd ed.). Sage. https://socialsciences.mcmaster.ca/jfox/Books/Companion/

Canty, A., & Ripley, B. (n.d.). *boot: Bootstrap R (S-Plus) Functions.* CRAN. https://CRAN.R-project.org/package=boot

Hyndman, R. J., & Khandakar, Y. (2008). *Automatic time series forecasting: The forecast package for R. Journal of Statistical Software, 27*(3), 1–22. https://www.jstatsoft.org/article/view/v027i03

Trapletti, A., & Hornik, K. (n.d.). *tseries: Time Series Analysis and Computational Finance.* CRAN. https://CRAN.R-project.org/package=tseries

---

8. Appendix

Figure A1. Overview Dashboard: Monthly Flight Activity, Delays, and Cancellations at DCA

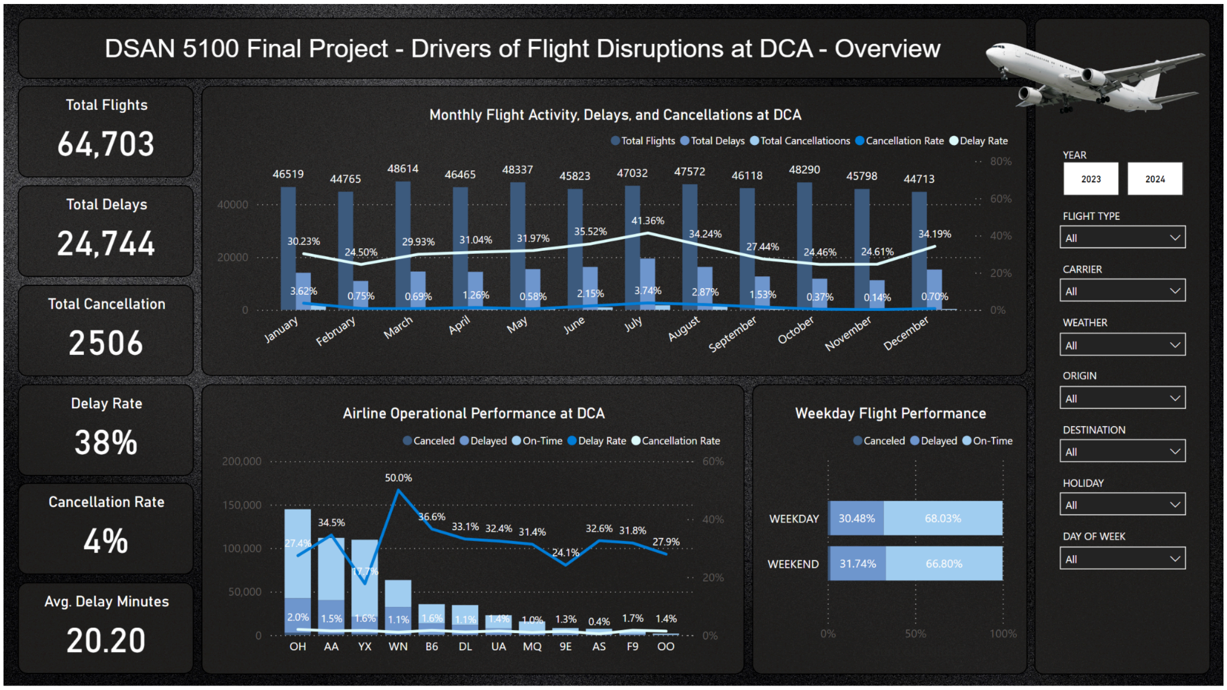


Figure A2. Overview Dashboard: Monthly Flight Activity, Delays, and Cancellations at DCA (Departing Flights, 2023–2024)

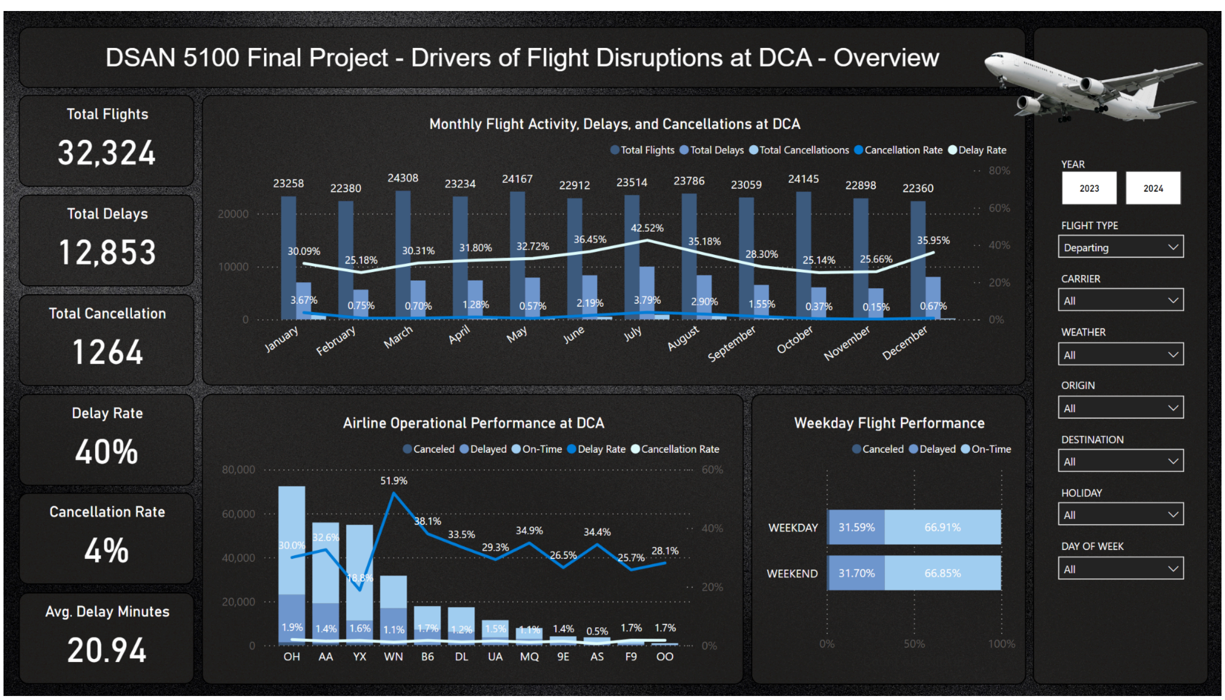


Figure A3. Environmental Dashboard: Weather Effects on Flight Activity, Delays, and Cancellations at DCA (Departing Flights, 2023)

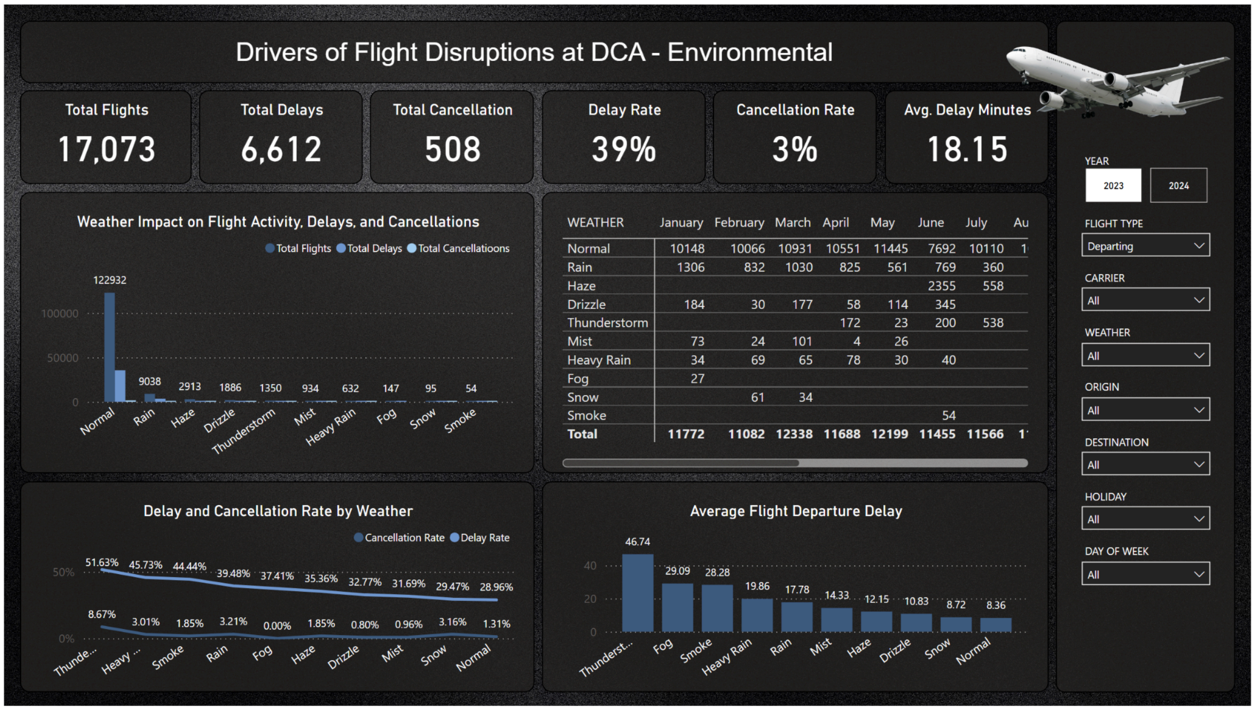


Figure A4. Environmental Dashboard: Weather Effects on Flight Activity, Delays, and Cancellations at DCA (Departing Flights, 2024)

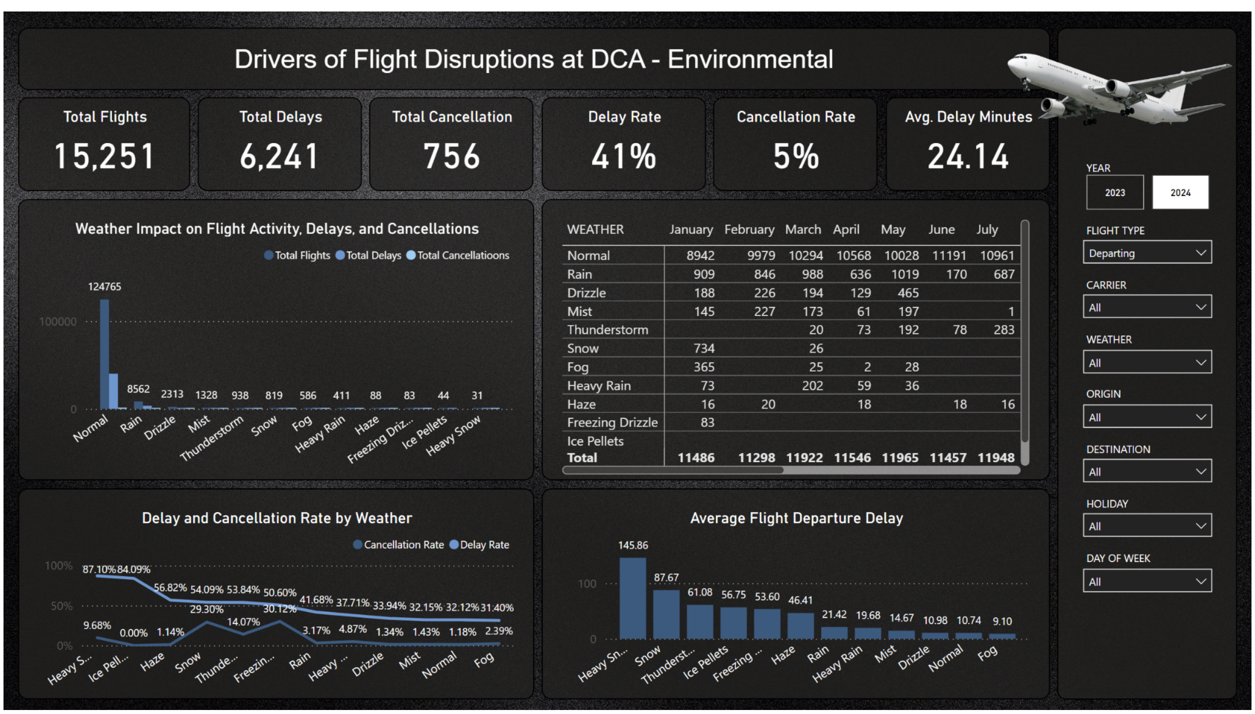


Figure A5. Environmental Dashboard: Weather Effects on Flight Activity, Delays, and Cancellations at DCA (Departing Flights, 2023–2024)

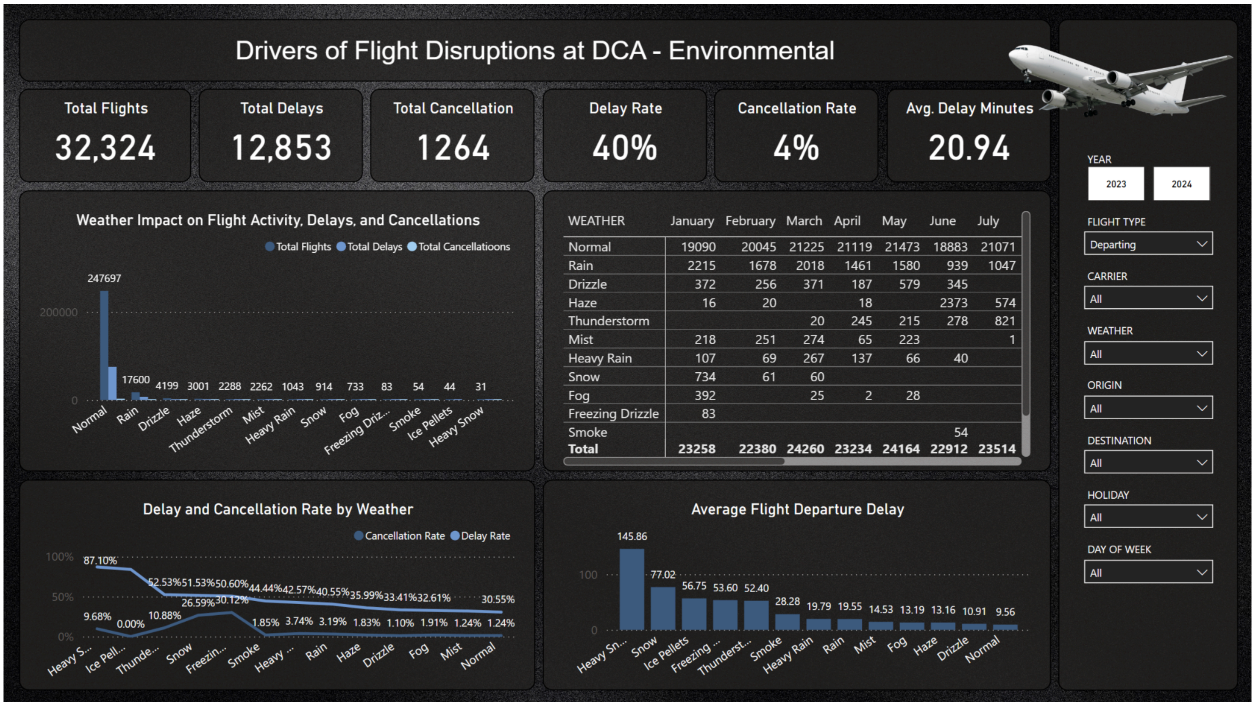


Figure A6. Environmental Dashboard: Weather Effects on Flight Activity, Delays, and Cancellations at DCA (Arriving Flights, 2023–2024)

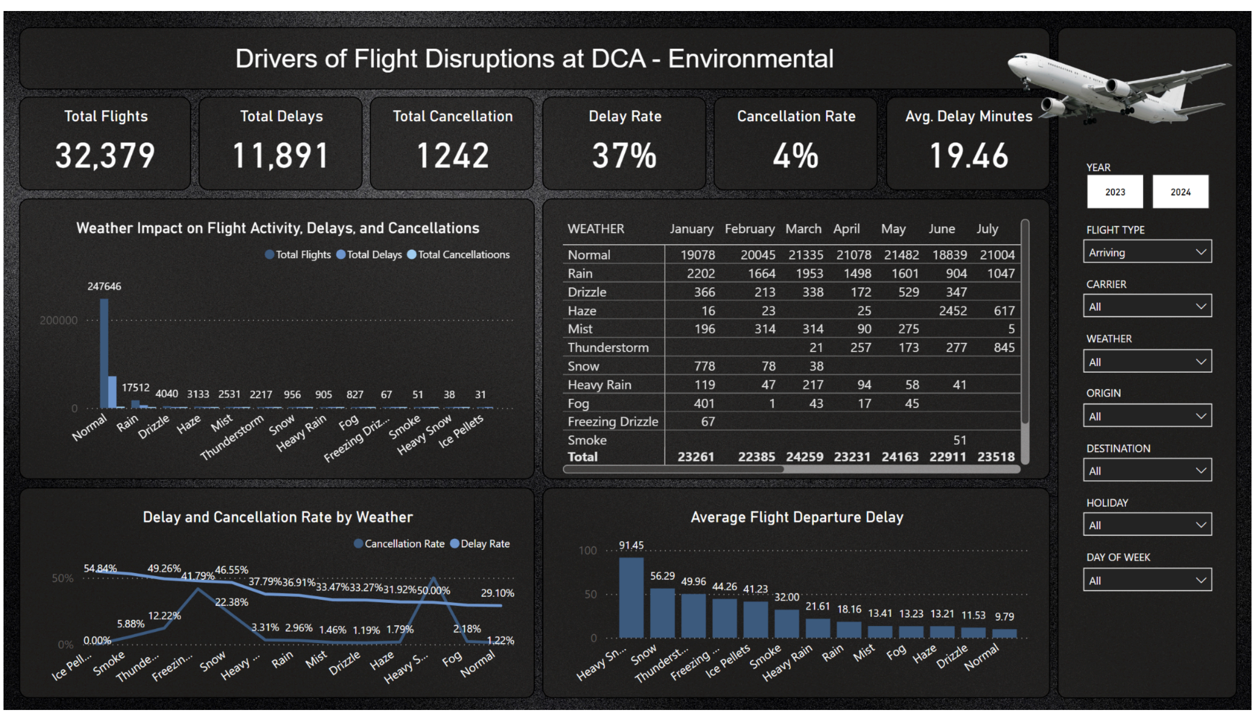


Figure A7. Seasonal Dashboard: Holiday and Monthly Patterns in Flight Delays and Cancellations at DCA (All Flights, 2023–2024)

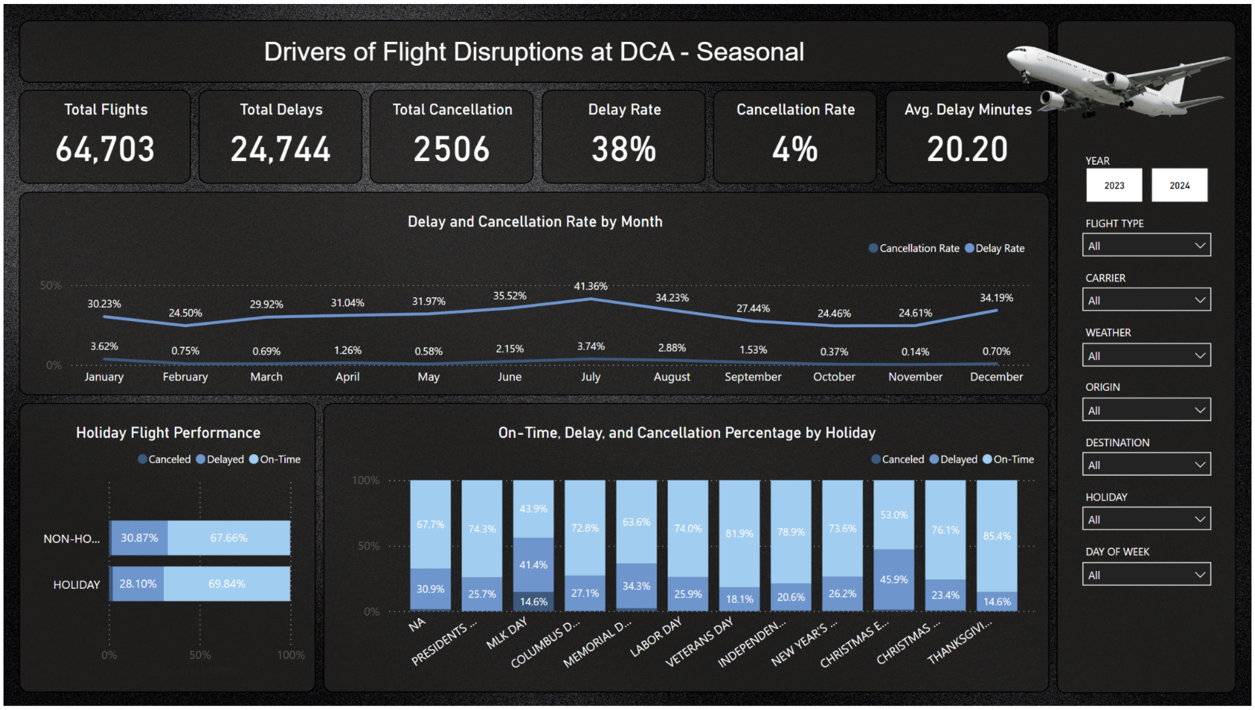


Figure A8. Airline Operational Dashboard: Airline-Level Differences in Delays and Cancellations at DCA (All Flights, 2023)

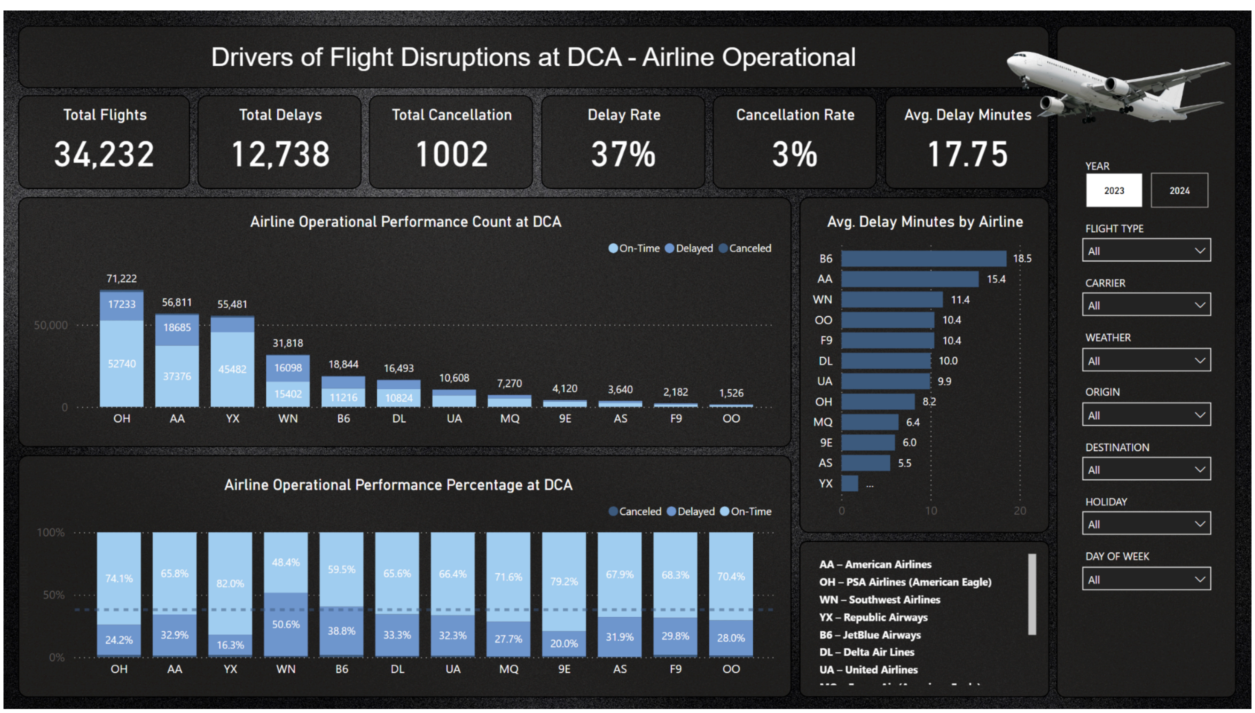


Figure A9. Airline Operational Dashboard: Airline-Level Differences in Delays and Cancellations at DCA (All Flights, 2024)