Milestone 4 – Week 6 Report

Cancellation and Seasonal Analysis

Project Title: AirFly Insights: Data Visualization and Analysis of Airline Operations
Intern Name: Sarthak Mokal
Organization: Infosys – Internship Program (Data Analytics & Visualization)
Milestone: 4 – Week 6

---

1. Introduction

The focus of **Week 6** was to extend the analysis from route- and airport-level delay behavior (Week 5) to a **cancellation- and seasonal-level study**.
This milestone aimed to examine the **patterns, reasons, and seasonality** behind flight cancellations across U.S. airlines and to understand how factors such as **weather, security, air-traffic system (NAS)**, and **carrier issues** contributed to disruptions.

By integrating temporal and environmental factors, the analysis helped uncover **when** and **why** flights are most frequently cancelled, offering operational insights into **carrier reliability**, **seasonal impacts**, and **holiday-period performance**.

---

2. Objectives

The main objectives for Week 6 were:

- To summarize the **overall cancellation statistics** and identify the major causes.
- To analyze **monthly cancellation trends** and highlight peak months of disruption.
- To visualize **airline-wise and month-wise cancellation distribution**.
- To evaluate the **impact of different delay causes** (Carrier, NAS, Weather, Security).
- To compare **seasonal (Winter vs Non-Winter)** and **holiday vs non-holiday** cancellation behavior.

---

3. Dataset Summary after Preprocessing

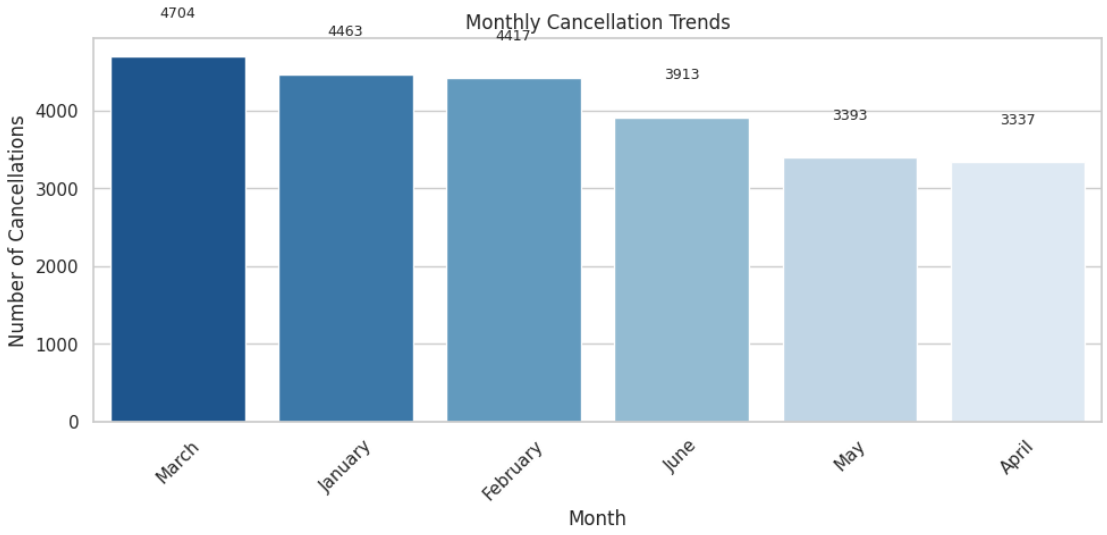
Metric	Value
Unique Airlines	12
Unique Routes	3377
Cancelled Flights	24,227
Cancellation Reasons	Carrier (12,161), NAS (5,951), Weather (4,940), Security (175)

Majority of cancellations were due to **Carrier issues**, followed by **NAS** and **Weather** factors. Security-related cancellations were minimal.

---

4. Visual Analysis and Findings

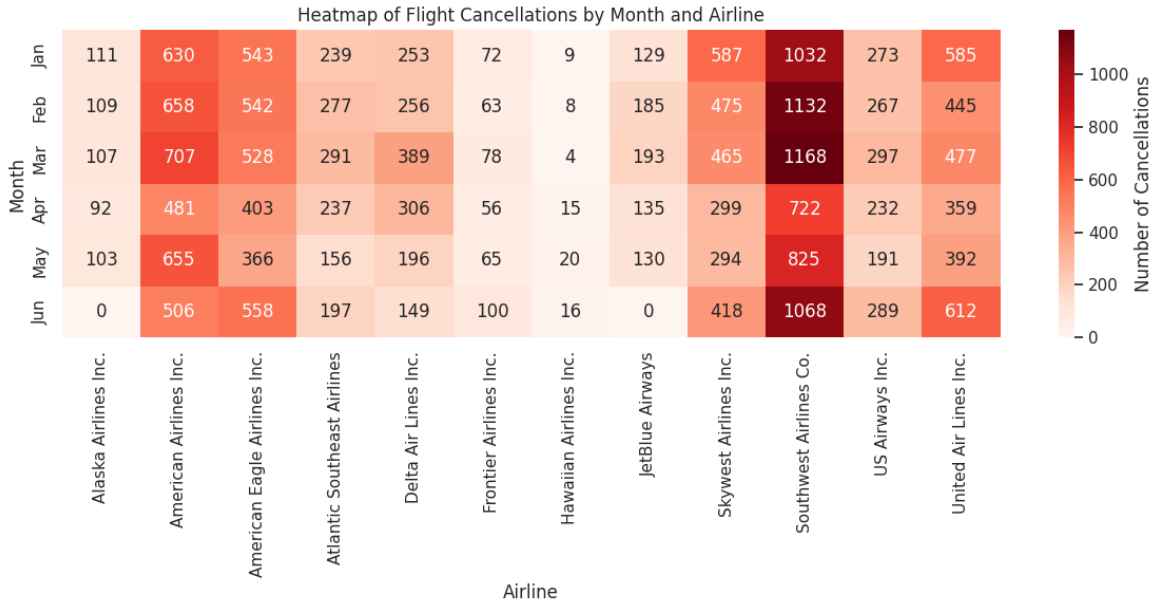
4.1 Monthly Trends



A bar-chart visualization of monthly cancellations showed:

- **March** had the **highest** number of cancellations (~4,700 flights).
- **January** (4,463) and **February** (4,417) also showed high values, indicating **early-year seasonal disruptions** (winter storms and congestion).
- Cancellations steadily declined through **April to June**, reflecting operational normalization.

4.2 Airline and Month Heatmap

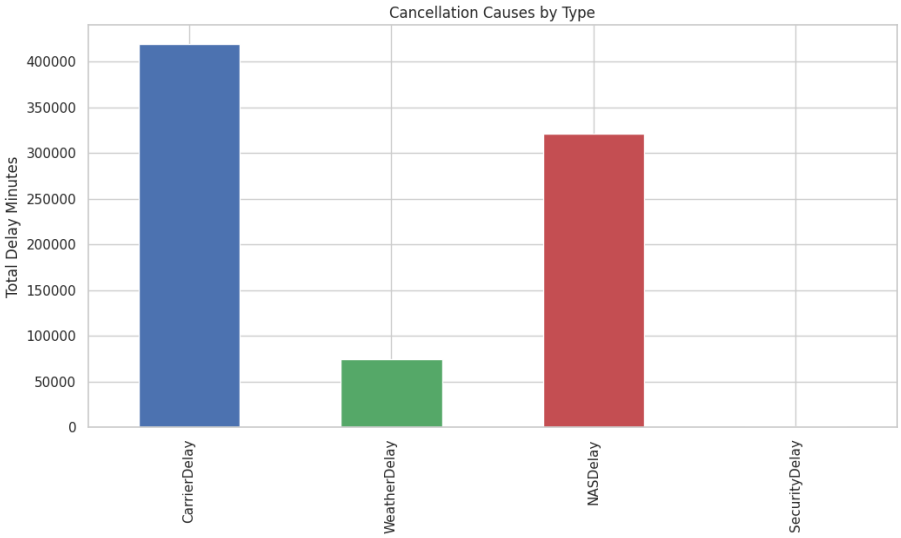


A heatmap cross-tabulating airline vs month revealed:

- **Southwest Airlines Co.** and **American Airlines Inc.** consistently reported the highest cancellation counts across all months.
- **JetBlue and US Airways Inc.** had moderate cancellations, while smaller carriers (Frontier, Hawaiian) showed minimal disruptions.
- March again showed intensified cancellation activity across multiple airlines.

Insight: Larger network carriers exhibit higher cancellation volume but not necessarily higher rates, due to greater flight frequency.

4.3 Cancellation Causes by Type

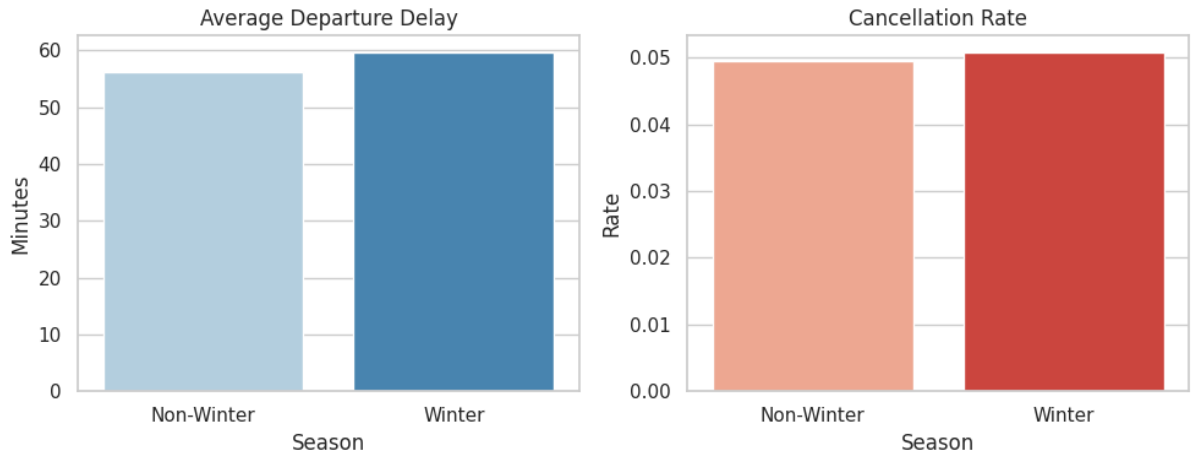


A categorical bar graph comparing total delay minutes showed:

- **Carrier Delays** ≈ 420 K minutes → primary contributor.
- **NAS Delays** ≈ 320 K minutes → air-traffic system congestion.
- **Weather Delays** ≈ 75 K minutes → significant but less frequent.
- **Security Delays** negligible.

Interpretation: Most cancellations originate from **carrier-related operational issues** and air-traffic system inefficiencies rather than external weather.

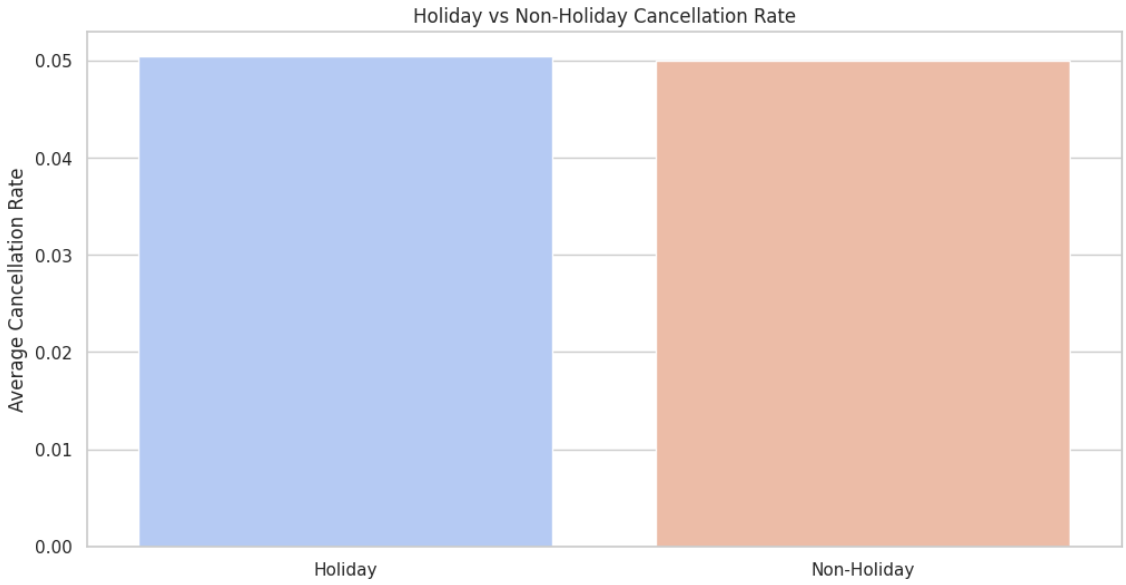
4.4 Seasonal Analysis (Winter vs Non-Winter)



Season	Avg Departure Delay	Cancellation Rate
Non-Winter	≈ 56 minutes	0.049
Winter	≈ 59 minutes	0.051

- Winter season shows **slightly higher average delays and cancellation rate**.
- Reflects typical winter disruptions like storms and de-icing delays.

4.5 Holiday vs Non-Holiday Analysis



Bar chart comparison revealed **similar cancellation rates** (~5%) for both periods.

Insight: Despite high traffic during holidays, airlines maintain consistent performance due to advanced planning and resource allocation.

---

5. Key Insights

- **Carrier issues** remain the top reason for cancellations (> 50%).
- **March** marked the highest cancellation volume.
- **Southwest and American Airlines** account for the largest share of cancellations.
- **Winter months** experience slightly higher delay and cancellation rates.
- **Holiday season** operations are handled efficiently with no significant spikes.

---

6. Conclusion

Week 6 provided a comprehensive understanding of **flight cancellation behavior** across time, airlines, and causes.
While weather and security events play a role, the dominant impact comes from **carrier and air-traffic (NAS)** related delays.

The findings emphasize the importance of **proactive carrier maintenance and airspace management** during peak winter months to minimize disruptions.