**SMART Question**

*Which U.S. airports consistently experience the highest cumulative delays across different delay causes (Carrier, Weather, NAS, Security, Late Aircraft) for both origin and destination flights, and how does this ranking vary on a monthly or seasonal basis over the observed period?*

This question is crafted to generate actionable insights by focusing on identifying not just the airports with the highest delays but also the primary causes and their seasonal patterns. Answering this question can guide targeted interventions and resources to reduce delays and enhance overall travel efficiency.

**Exploratory Data Analysis (EDA)**

The dataset for this analysis comprises records of flight delays across U.S. airports. Each entry includes details such as origin and destination airports, delay causes, and delay duration in minutes. Preprocessing and EDA were crucial steps in transforming this data into meaningful insights.

**Graphical Representation**

To visualize the insights obtained through EDA, three primary graphs were generated:

1. Frequency of Delays at Top 10 Origin and Destination Airports
2. Heatmap of Delay Causes at Top 10 Airports
3. Seasonal Delay Patterns at Top Airports

**Frequency of Delays at Top 10 Origin and Destination Airports**

To identify the airports with the highest delay frequency, we aggregated the data by origin and destination airports and calculated the total count of delays for each. Using the ggplot2 library, a bar chart was created to visualize these counts, with separate bars for origin and destination delays.

A graph of a number of airports

Description automatically generated with medium confidence

**Code:**

A computer screen shot of a computer code

Description automatically generated

**Findings:** The bar chart illustrated that Chicago O'Hare International Airport (ORD), Dallas/Fort Worth International Airport (DFW), and Hartsfield-Jackson Atlanta International Airport (ATL) consistently had the highest delay frequencies. This visualization allows us to see which airports contribute most to delays, offering insight into which locations could benefit most from improvement initiatives. These airports serve as significant hubs, which likely increases the volume and frequency of delays due to their high operational demands.

**Heatmap of Delay Causes at Top 10 Airports**

To explore which specific delay causes are most prevalent at the top 10 airports, a heatmap was generated. We first isolated the delay causes for each airport and then aggregated them to calculate the total delay minutes per cause for each airport. The causes included Carrier delays, Weather delays, NAS (National Airspace System) delays, Security delays, and Late Aircraft delays.

A graph of a flight

Description automatically generated with medium confidence

**Code:**

A screenshot of a computer program

Description automatically generated

**Findings:** The heatmap highlighted Late Aircraft as the leading cause of delays across most airports, followed by NAS delays. This suggests that delays in arriving aircraft often lead to further delays in departing flights. Weather and Carrier delays were less frequent, and Security delays had minimal impact. The use of colour intensity in the heatmap provided a clear view of where each delay type is concentrated, making it easier to pinpoint which airports are heavily affected by specific delay causes.

**Seasonal Delay Patterns at Top Airports**

Lastly, to understand seasonal patterns, a bar chart was created to display the cumulative delay minutes across different seasons (Winter, Spring, Summer, and Fall) for each of the top 10 airports. This analysis helps in determining if certain seasons contribute more to delays, which could be attributed to seasonal weather changes, travel demand fluctuations, or operational factors.

A graph of a number of flights

Description automatically generated

**Code:**

A screenshot of a computer code

Description automatically generated

**Findings:** The seasonal bar chart revealed that winter had the highest delay minutes, followed by spring. Winter delays are expected due to adverse weather conditions that impact flight operations. This seasonal analysis underscores the importance of season-specific delay mitigation strategies, such as deploying additional resources or preparing contingency plans during peak delay seasons.

**Conclusion:**

This analysis provides a data-driven perspective on the primary factors contributing to delays at major U.S. airports. Key findings include:

Top Airports by Delay Frequency: Chicago O'Hare (ORD), Dallas/Fort Worth (DFW), and Hartsfield-Jackson Atlanta (ATL) were identified as the airports with the highest frequency of delays for both origin and destination flights. As major hubs, these airports face increased operational demands, which may contribute to their higher delay frequencies.

Primary Delay Causes: Late Aircraft delays were the predominant cause, indicating that delays propagate through interconnected flight schedules. NAS delays were also a significant contributor, emphasizing the importance of efficient airspace management to reduce delays. Carrier and Weather delays had a lower impact, while Security delays were minimal. These insights highlight the critical need for efficient management of flight schedules and airspace operations to reduce delay propagation.

Seasonal Impact on Delays: Winter was observed to have the highest total delay minutes, followed by spring, likely due to inclement weather and increased travel demand during these seasons. This suggests that airports should be better prepared for winter and spring seasons by allocating additional resources and implementing contingency plans.

**Implications of the Analysis:**

By understanding which airports, delay causes, and seasons contribute the most to flight delays, airline and airport management teams can implement targeted strategies to mitigate delays. For instance, during high-delay seasons like winter, airports could allocate additional ground staff, deploy resources for de-icing and snow removal, or adjust flight schedules to accommodate potential delays. Additionally, since Late Aircraft delays are a major issue, airlines could prioritize maintaining buffer times between connecting flights to reduce the knock-on effects of delays.

Overall, this analysis provides valuable insights into optimizing airport and airline operations, particularly for high-traffic hubs. Addressing these delays requires coordinated efforts from airport management, airlines, and regulatory bodies to improve efficiency and minimize the inconvenience caused to passengers.