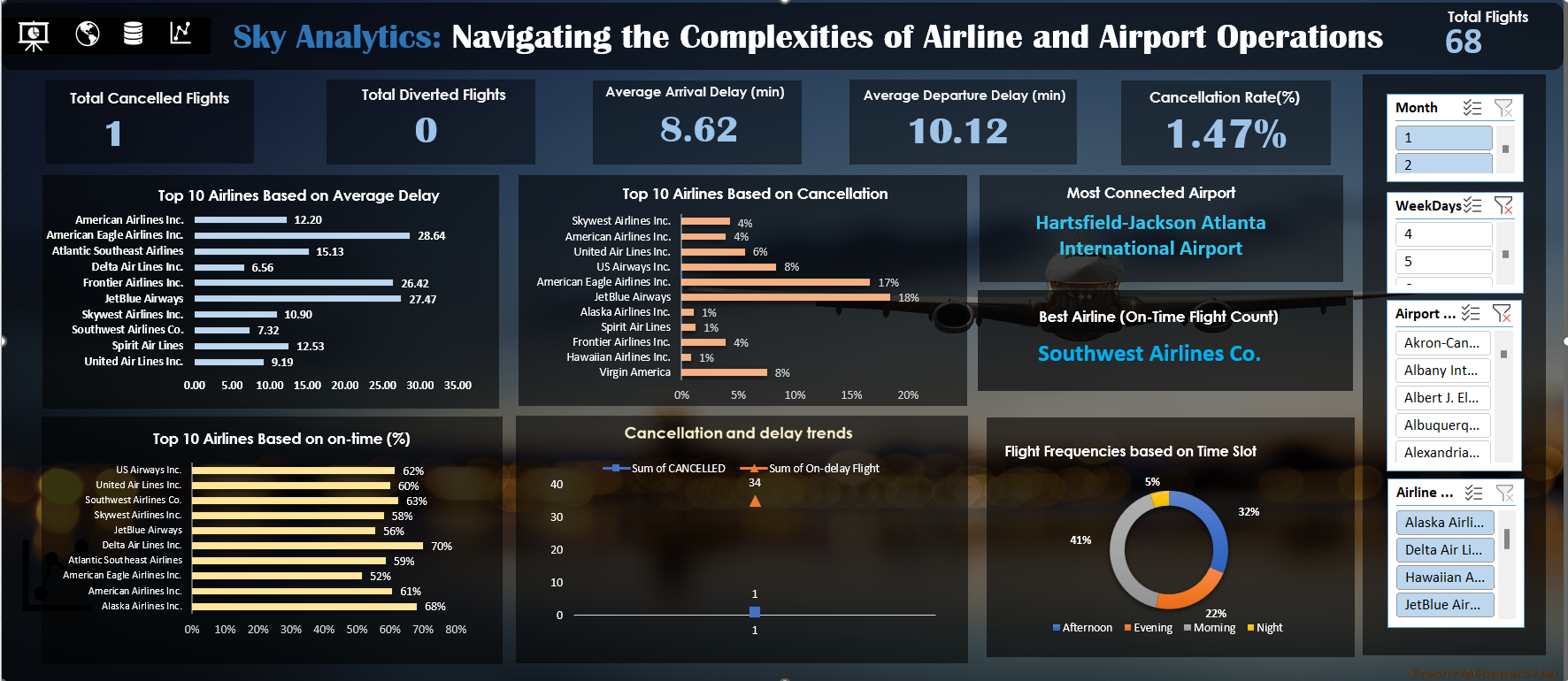
**Travel & Hospitality**

**Sky Analytics: Navigating the Complexities of Airline and Airport Operation:**

SkyNet Analysis Inc. is a leading consultancy firm specializing in aviation analytics. With the aviation industry's rapid expansion and the increasing complexity of global air travel, SkyNet plays a critical role in providing data-driven insights to airlines, airports, and regulatory bodies. The company has access to extensive datasets that cover a wide range of information, including flight schedules, delays, airline operations, and airport traffic details. These datasets offer a unique opportunity to explore and understand the multifaceted nature of the aviation industry, from operational efficiency and customer satisfaction to logistical challenges and environmental impact.

**Objective:**

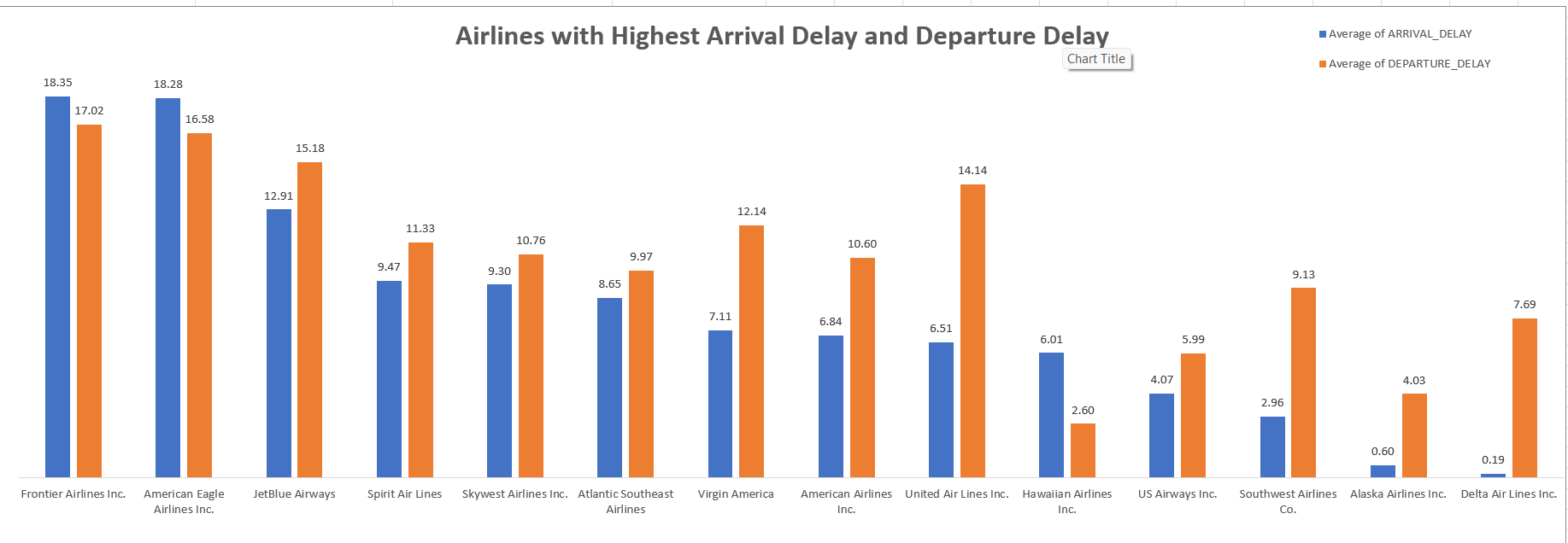
The primary objective of this case study, titled "Sky Analytics: Navigating the Complexities of Airline and Airport Operations," is to deeply analyze and interpret the extensive datasets encompassing flights, airlines, and airports - namely "flights.csv", "airlines.csv", and "airports.csv". The analysis aims to uncover critical insights into flight operations, delay patterns, airline efficiency, and airport traffic dynamics. By exploring these datasets, the study seeks to identify key factors influencing operational efficiency, understand the intricacies of flight scheduling and delays, and evaluate the performance metrics of airlines and airports. The ultimate goal is to provide strategic recommendations to enhance operational effectiveness, improve customer experiences in air travel, and contribute to the overall advancement of the aviation industry's standards and practices.



**1.Flight Delays Analysis**: Determine the average flight delay per airline. What are the top 3 airlines with the highest average delays?

Based on the analysis of flight delays, the top 3 airlines with the highest average delays are:

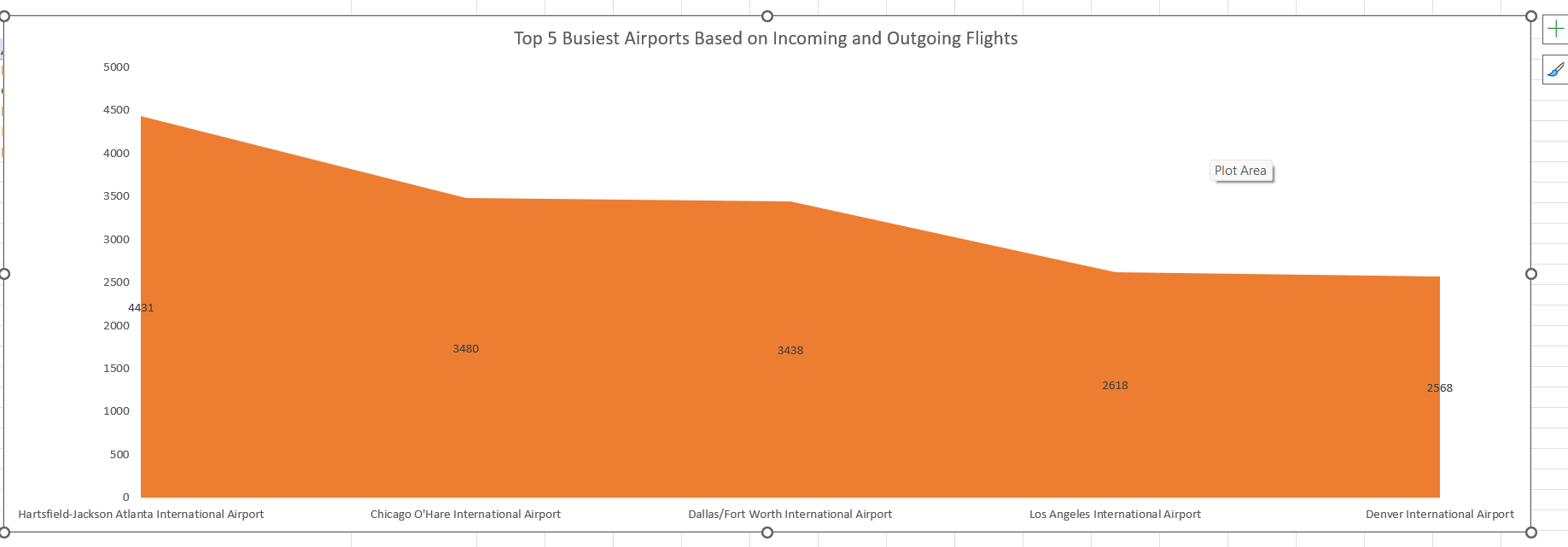
* 1. **Frontier Airlines Inc.** with an average arrival delay of 18.35 minutes and an average departure delay of 17.02 minutes.
  2. **American Eagle Airlines Inc.** with an average arrival delay of 18.28 minutes and an average departure delay of 16.58 minutes.
  3. **JetBlue Airways** with an average arrival delay of 12.91 minutes and an average departure delay of 15.18 minutes.

These airlines have been identified as having the highest average delays in both arrival and departure based on the data analyzed and represented in the chart.

2.**Airport Traffic Volume**: Identify the top 5 busiest airports based on the number of incoming and outgoing flights.

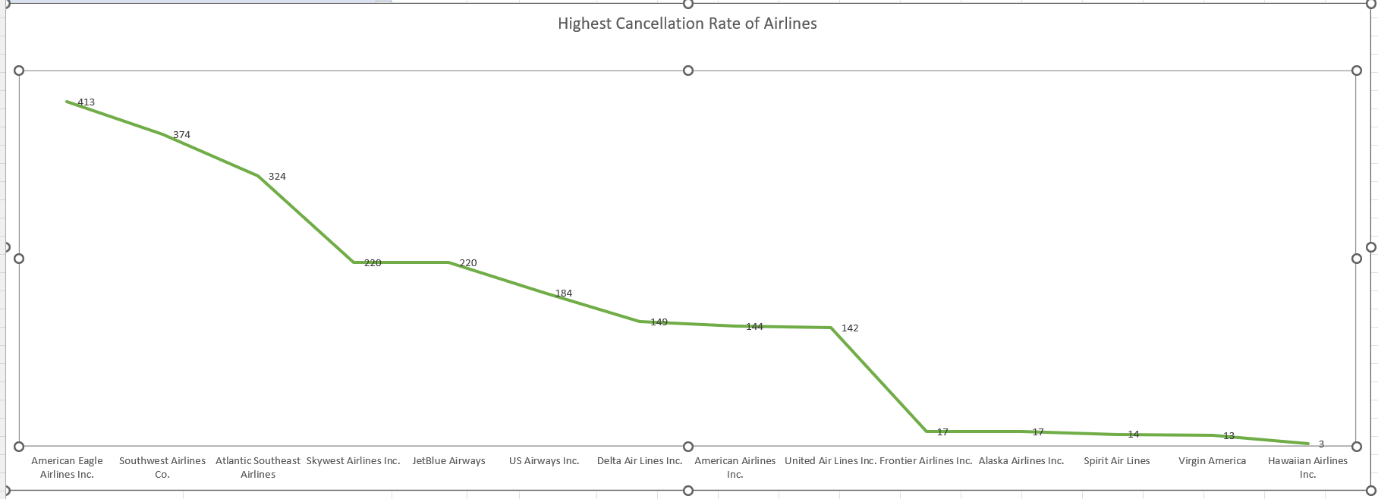
The top 5 busiest airports based on the number of incoming and outgoing flights are:

* 1. **Hartsfield-Jackson Atlanta International Airport** with 4,431 flights.
  2. **Chicago O'Hare International Airport** with 3,480 flights.
  3. **Dallas/Fort Worth International Airport** with 3,438 flights.
  4. **Los Angeles International Airport** with 2,618 flights.
  5. **Denver International Airport** with 2,568 flights.



**3. Flight Cancellation Insights**: Analyze the flight cancellations: Which airline has the highest cancellation rate, and what are the most common reasons for cancellations?

The airline with the highest number of cancelled flights is **American Eagle Airlines Inc.** with 413 cancellations.



**4.Seasonal Variations in Flight Operations**: Examine if there are seasonal patterns in flight operations. Are certain months more prone to delays or cancellations?

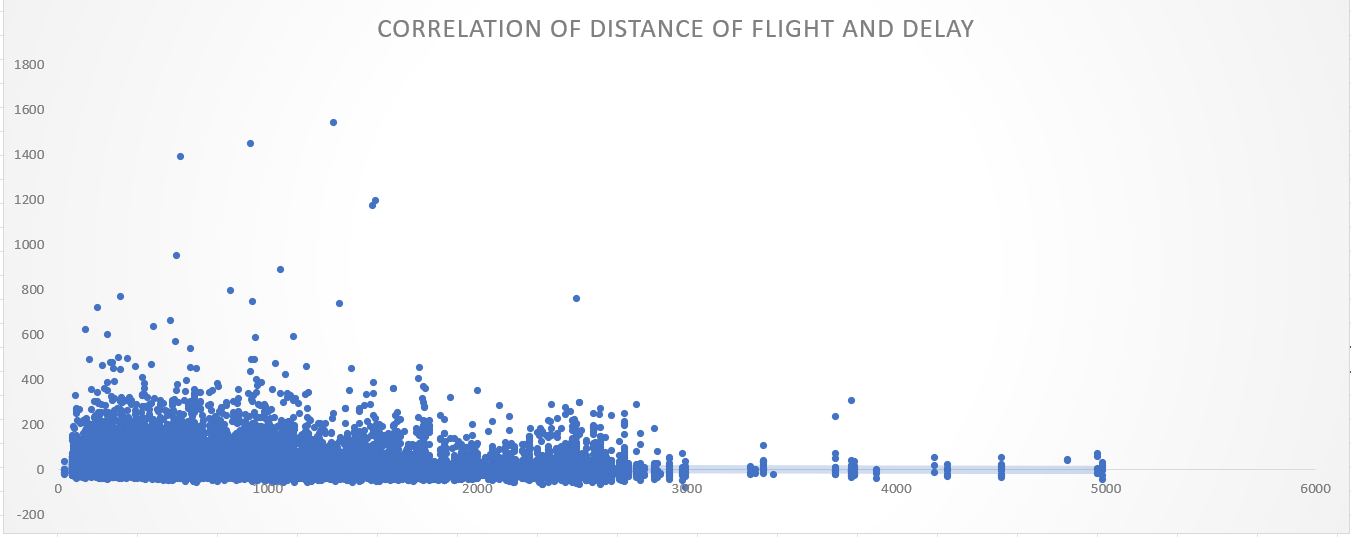
**Month 2** appears to be more prone to cancellations, with the highest count among the provided data.

**Month 1** has a higher number of departure delays, indicating potential seasonal factors affecting flight operations.

| **Month** | **Count of CANCELLED** | **Count of DEPARTURE\_DELAY** |
| --- | --- | --- |
| 2 | 1,161 | 23 |
| 1 | 1,073 | 29 |
| **Grand Total** | **2,234** | **52** |

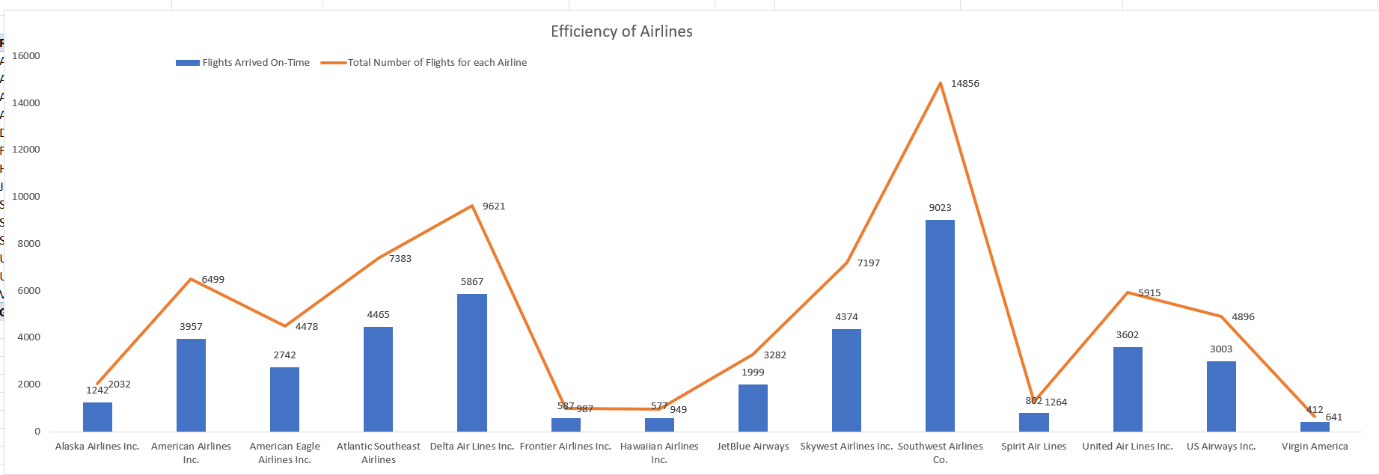
5. **Correlation between Distance and Delays**: Investigate if there's a correlation between the distance of the flight and the length of delays. Use scatter plots for visualization.

The correlation coefficient of -0.03253 suggests that there is almost no significant linear relationship between flight distance and delay time. This means that other variables likely have a greater impact on flight delays, and distance is not a major factor influencing delays

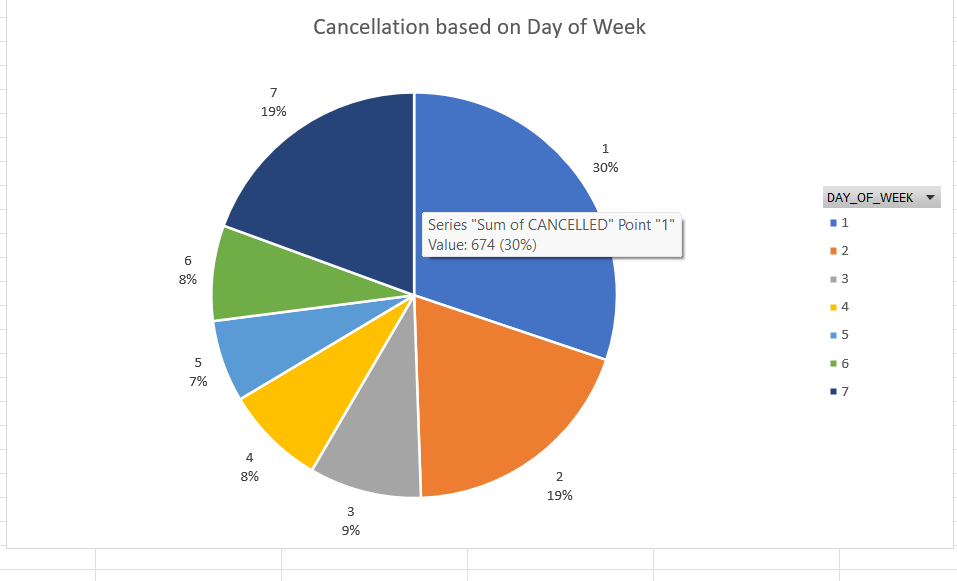


**Efficiency of Airlines**: Calculate the on-time performance (percentage of flights that are not delayed) for each airline. Rank them based on this metric.





**8.Impact of Day of Week on Flight Operations**: Assess how flight operations (delays, cancellations) vary by the day of the week.

flight operations vary significantly by the day of the week. Mondays have the highest cancellation rates, while Wednesdays and Thursdays show the best on-time performance. Fridays are the most efficient overall. Saturdays face higher cancellations, and Sundays maintain strong on-time performance. Recognizing these patterns helps airlines optimize operations and allows passengers to plan better for smoother travel experiences.

9.**Analysis of Airport Connectivity**: Which airports serve as the most significant hubs in terms of connectivity (most destinations served)?

The top five airports serving as the most significant hubs based on connectivity are:

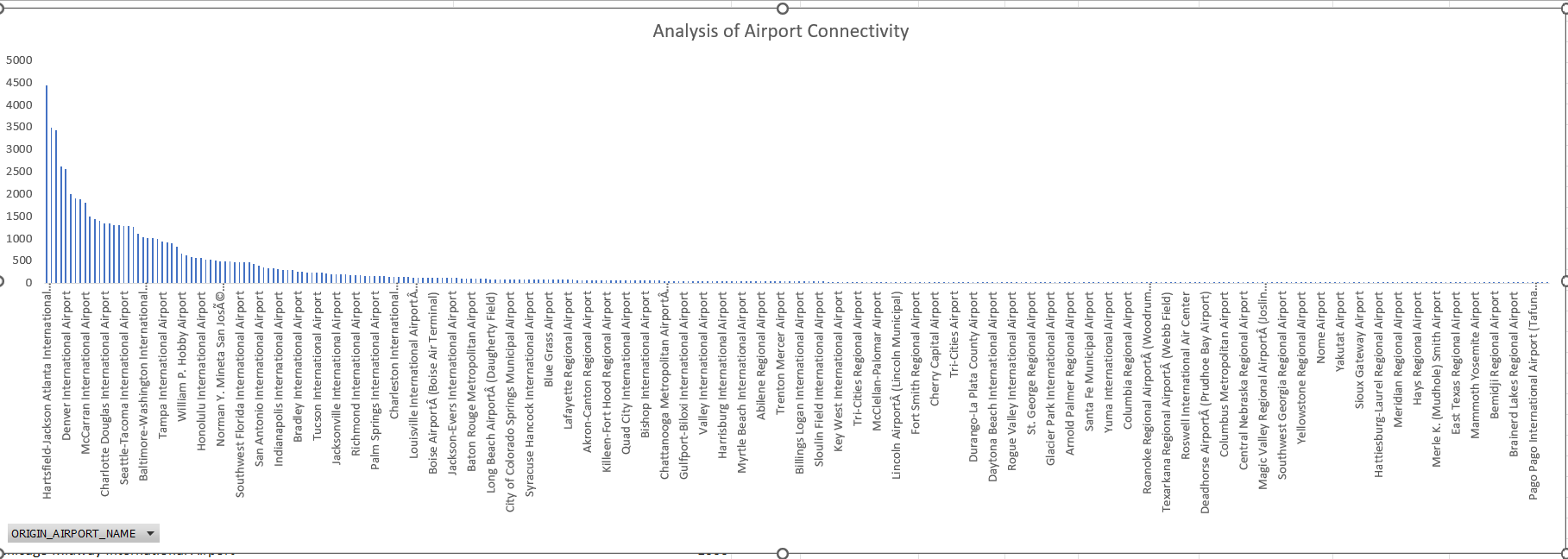
Hartsfield-Jackson Atlanta International Airport leads with 4431 destinations, making it the most connected airport.

Chicago O'Hare International Airport follows with 3480 destinations, reflecting its role as a major transportation hub.

Dallas/Fort Worth International Airport serves 3438 destinations, cementing its status as a key connectivity point.

Los Angeles International Airport with 2618 destinations, highlights its importance, particularly for international and transcontinental flights.

Denver International Airport rounds out the top five with 2568 destinations, emphasizing its strategic location and connectivity.

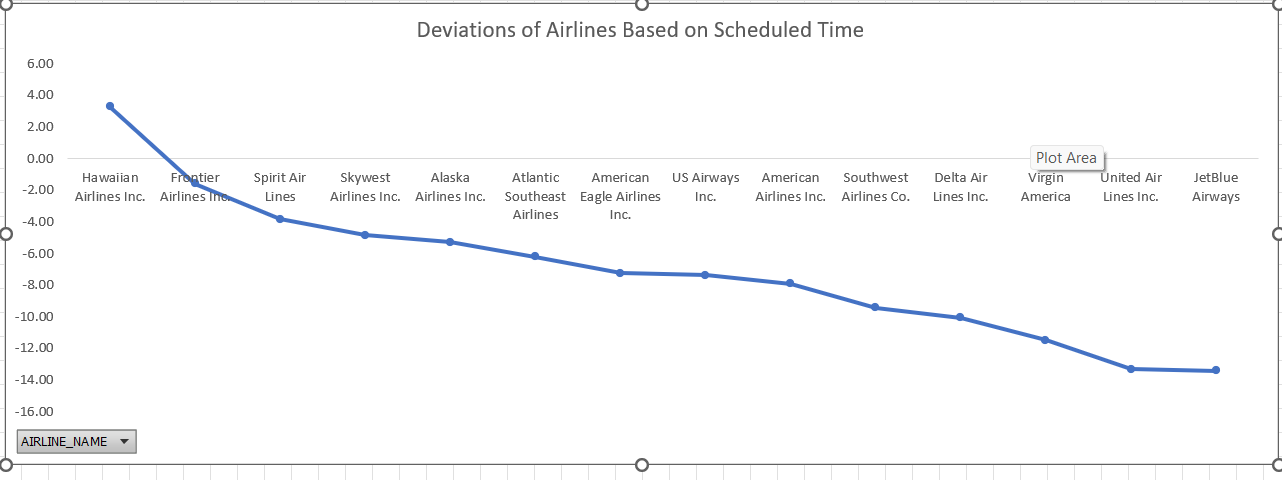


**10.Flight Duration Accuracy**: Compare the scheduled flight duration versus the actual flight duration. Which airlines have the most and least deviation?

Most Accurate (Least Deviation):Hawaiian Airlines Inc. stands out with an average deviation of 3.29 minutes, indicating it often arrives slightly later than scheduled but very close to the planned times.

Least Accurate (Most Deviation):JetBlue Airways has the largest average deviation of -13.40 minutes, meaning it frequently arrives earlier than scheduled by a significant margin.

Hawaiian Airlines shows the highest accuracy in terms of adhering to the scheduled flight durations, while JetBlue Airways exhibits the least accuracy, often arriving much earlier than scheduled.

11.**Airline Fleet Utilization**: Based on the tail numbers, determine which airline has the highest number of flights per aircraft, indicating fleet utilization.

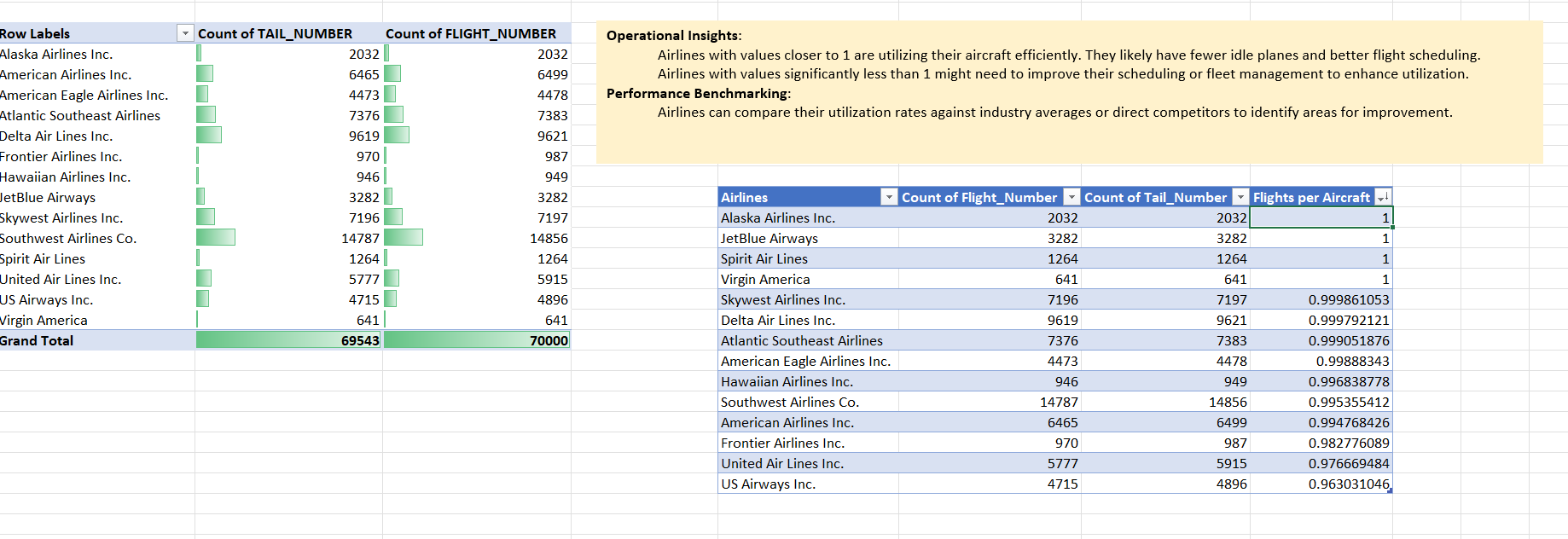
**Operational Insights**:

Airlines with values closer to 1 are utilizing their aircraft efficiently. They likely have fewer idle planes and better flight scheduling.

Airlines with values significantly less than 1 might need to improve their scheduling or fleet management to enhance utilization.

**Performance Benchmarking**:

Airlines can compare their utilization rates against industry averages or direct competitors to identify areas for improvement.



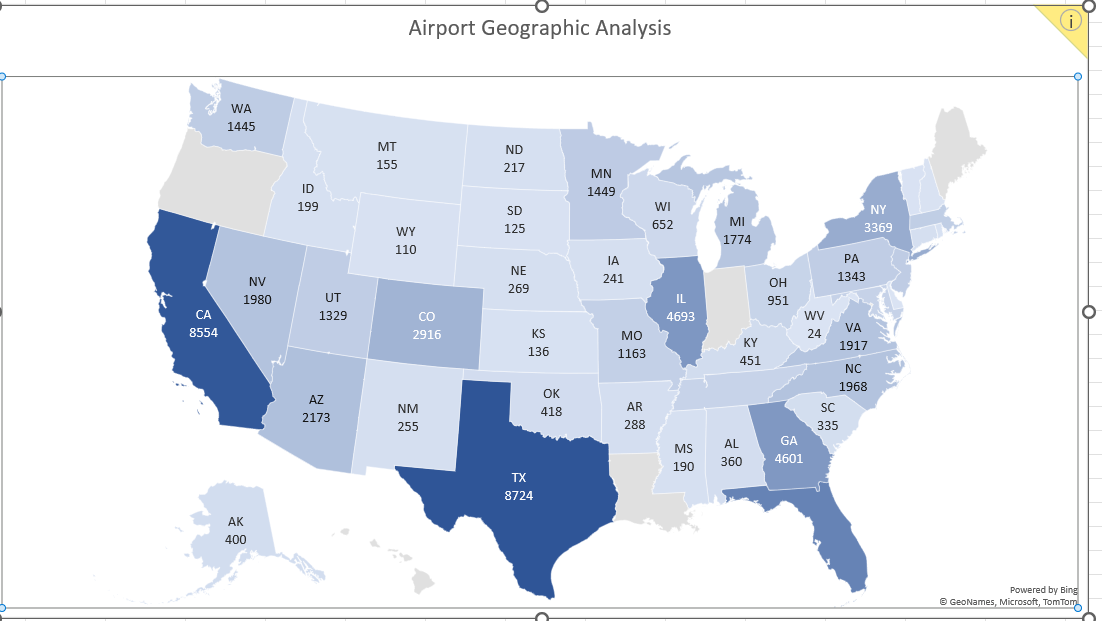
1. **Airport Geographical Analysis**: Using latitude and longitude data, analyze the geographical distribution of airports. Which states or regions have the highest concentration of airports?

1.Texas (TX) has the highest concentration of airports with a total of 8,724.

2.California (CA) follows closely with 8,554 airports, reflecting its vast size and significant air traffic needs.

3.Florida (FL) ranks third with 5,909 airports, indicative of its status as a major tourist destination.

4.Illinois (IL) and Georgia (GA) round out the top five with 4,693 and 4,601 airports, respectively

These states have the most extensive airport networks, supporting both domestic and international travel, and playing crucial roles in the country's transportation infrastructure

**13.Delayed Flights and Delay Types Analysis**: For flights that are delayed, break down the delay types (airline, weather, security, etc.) and analyze their proportions.

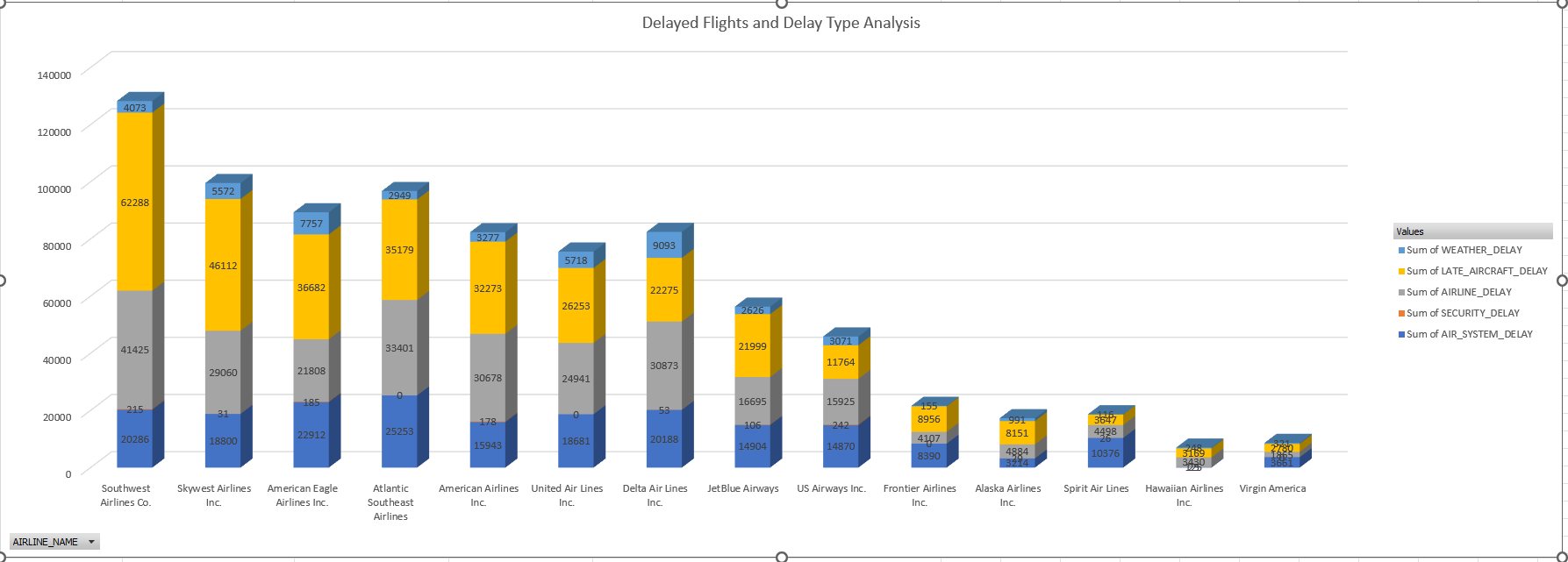
Airline Delays: Among the listed delay types, airline delays and late aircraft delays are the most prevalent contributors to flight delays, accounting for a significant proportion of delays across airlines.

Weather Delays: Weather delays also play a notable role, particularly in regions prone to adverse weather conditions. While weather delays contribute to delays, they are relatively lower compared to airline and late aircraft delays.

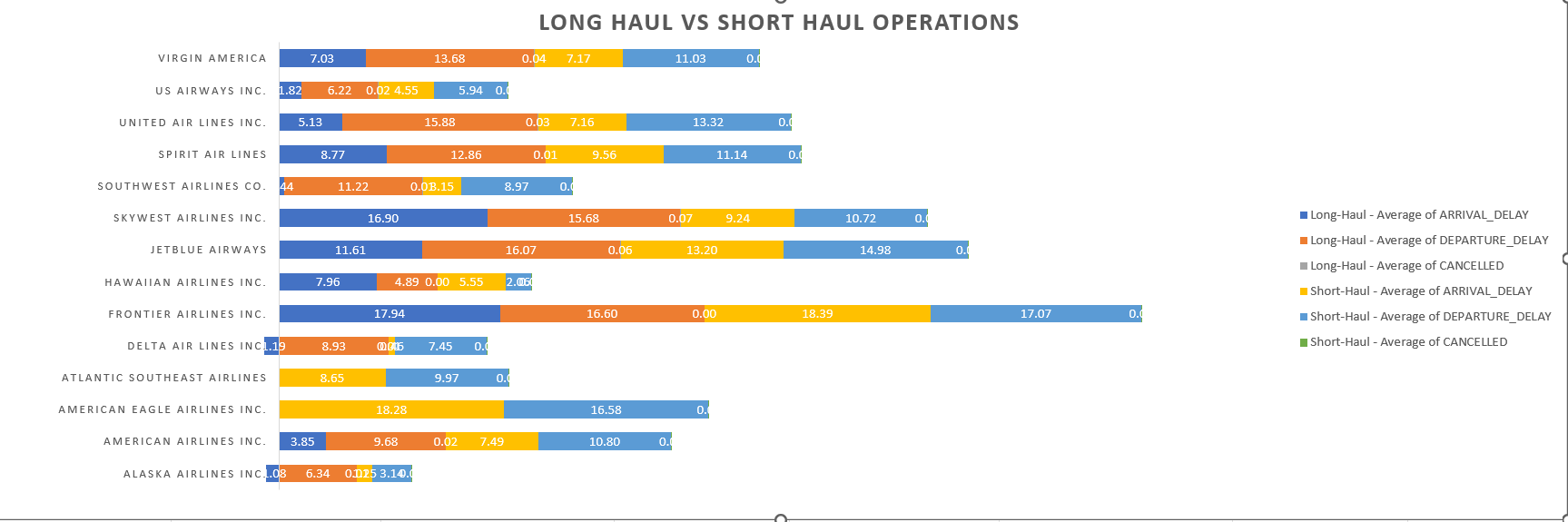
Air System Delays: Air system delays, which include delays attributed to air traffic control and airspace congestion, are another contributing factor to flight delays but to a lesser extent compared to airline and late aircraft delays.

Security Delays: Security delays have a minor contribution to overall flight delays, indicating that security-related issues have less impact on flight operations compared to other delay types.

airline and late aircraft delays are the primary contributors to flight delays, followed by weather delays and air system delays. Understanding the proportions of these delay types helps airlines identify areas for improvement and implement strategies to minimize delays and enhance overall operational efficiency.



14. **Long-Haul vs Short-Haul Operations**: Compare the operational metrics (delays, cancellations) between long-haul and short-haul flights for different airlines.

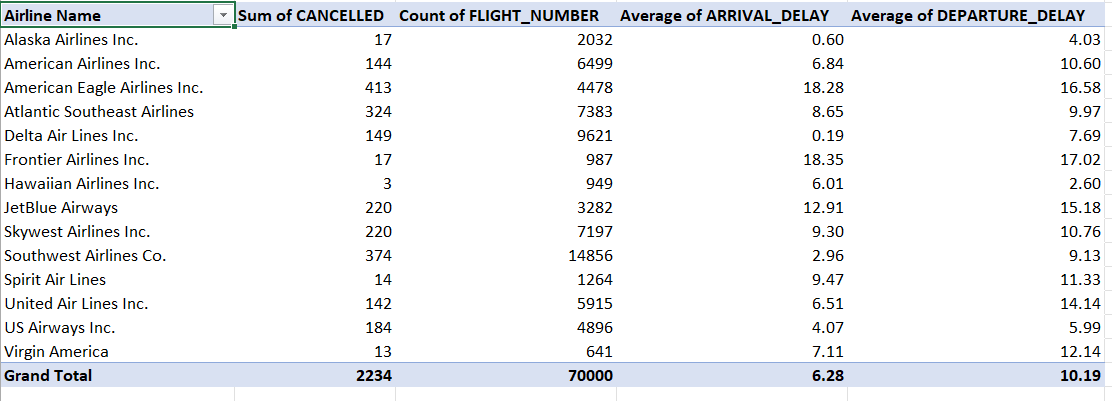


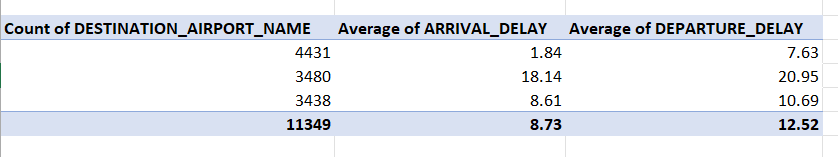
**15.Pivot Analysis of Flights Data**: Use pivot tables to summarize key operational metrics (like average delay, number of flights, cancellations) by airline and airport.

Delay Variation: Long-haul flights generally experience lower average delays compared to short-haul flights, except for specific airlines like Frontier and Skywest.

Low Cancellation Rates: Overall, cancellation rates for both long-haul and short-haul flights are relatively low across airlines, typically less than 0.05.

Performance Disparities: There are significant differences in delay metrics and cancellation rates among airlines. Hawaiian Airlines stands out for its consistently lower delays and cancellation rates, indicating superior operational efficiency.

Operational Focus: Airlines with higher delay metrics may need to prioritize improving operational processes to enhance reliability and customer satisfaction, while those with lower metrics can capitalize on their strengths to maintain high service standards.

16. **Data Integration for Comprehensive Insights**: Merge data from the "airlines.csv" and "flights.csv" to provide enhanced insights, such as correlating airline names with operational metrics. Analyze the merged data to determine the overall on-time performance of each airline, considering both arrival and departure delays.

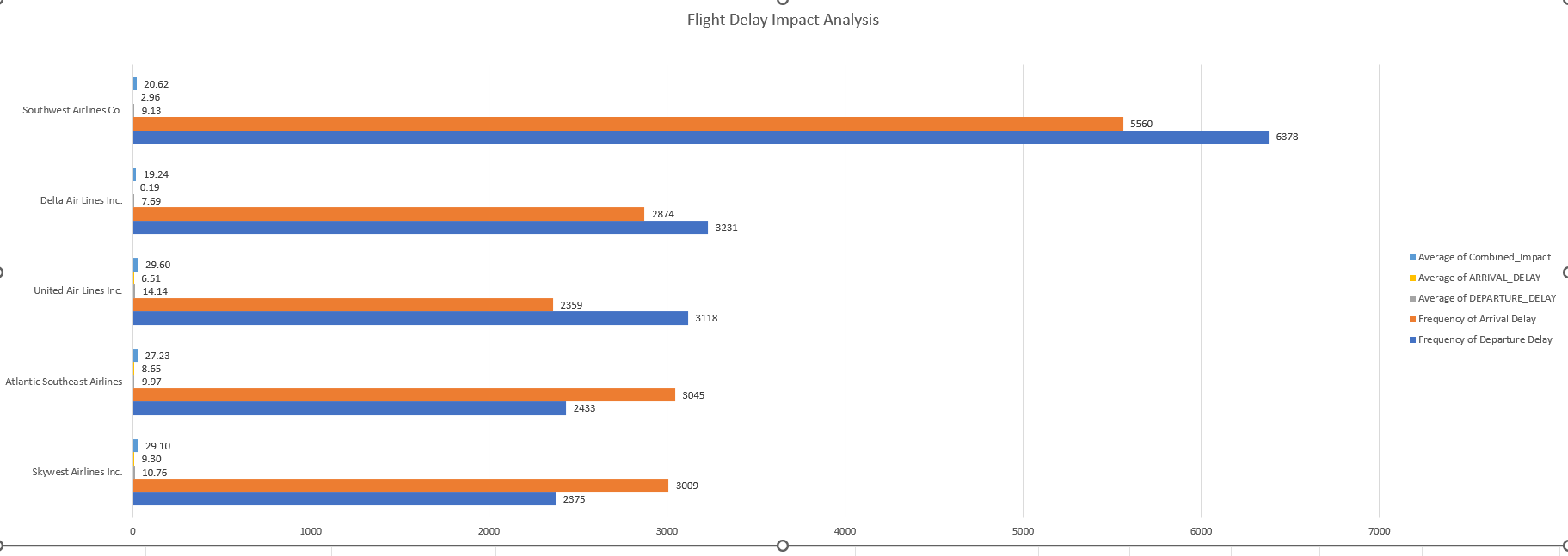
**17.Flight Delay Impact Analysis**: Calculate the cumulative impact of delays for each airline. Consider both the frequency of delays and the average delay time. How do these factors combine to affect overall airline performance?

High Cumulative Impact: Airlines like Southwest Airlines Co., United Air Lines Inc., and American Eagle Airlines Inc. exhibit a high cumulative impact of delays, indicating significant operational challenges and potential customer dissatisfaction.

Combined Impact Factor: JetBlue Airways and Frontier Airlines Inc. also show considerable combined impact factors due to higher average delay times, highlighting areas for improvement in managing delays.

Lower Impact Airlines: Hawaiian Airlines Inc. and Virgin America have relatively lower cumulative impacts, suggesting better overall performance in managing delays.

Operational Focus: Airlines with high cumulative impacts should focus on strategies to reduce both the frequency and duration of delays to improve overall performance and customer satisfaction. This could involve optimizing scheduling, enhancing communication and coordination, and investing in technology for better delay management.



**18.Optimal Flight Path Efficiency Assessment**: Using the distance data from "flights.csv" and geographical coordinates from "airports.csv", calculate the efficiency of various flight paths. Determine if there are significant differences in the efficiency of flights (measured as a ratio of actual flight time to the shortest possible time based on distance) for different airlines or types of aircraft (identified by tail number).

**Interpretation of Efficiency Ratios**:

**Efficiency Ratio**: This is calculated as the actual flight time divided by the shortest possible time to fly the given distance.

**Ratio ≈ 1**:

If the ratio is close to 1, it means the flight time is very close to the shortest possible time. This indicates high efficiency.

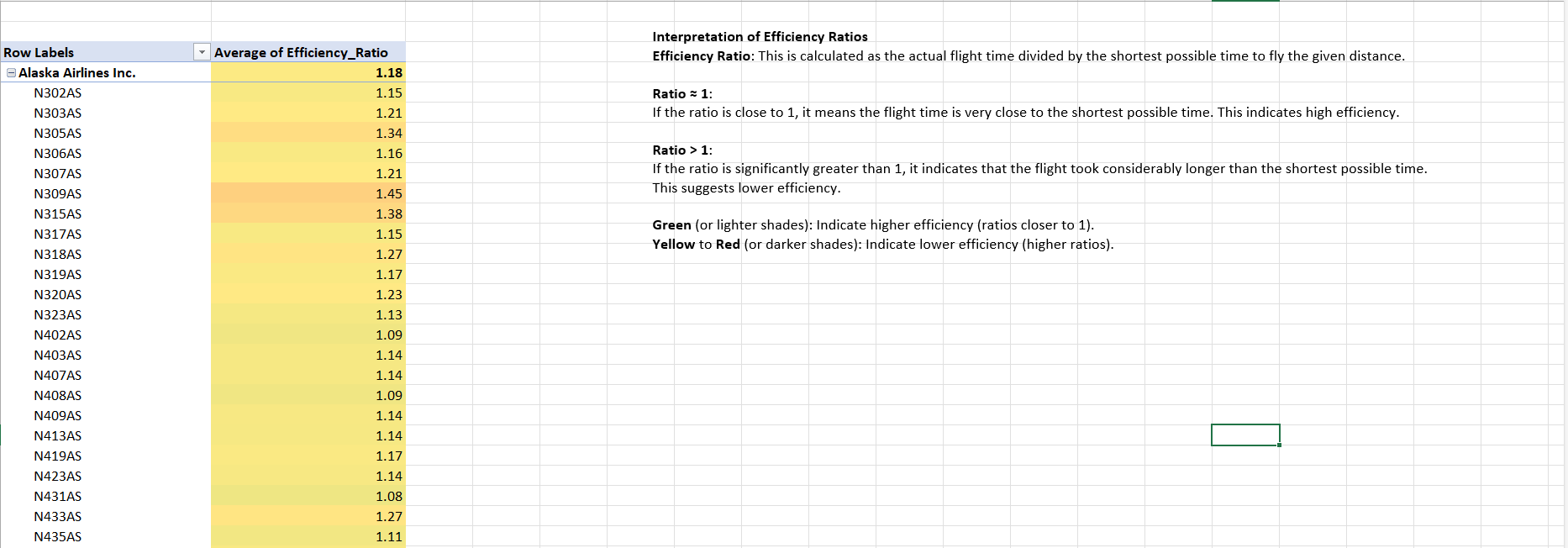
**Ratio > 1**:

If the ratio is significantly greater than 1, it indicates that the flight took considerably longer than the shortest possible time.

This suggests lower efficiency.

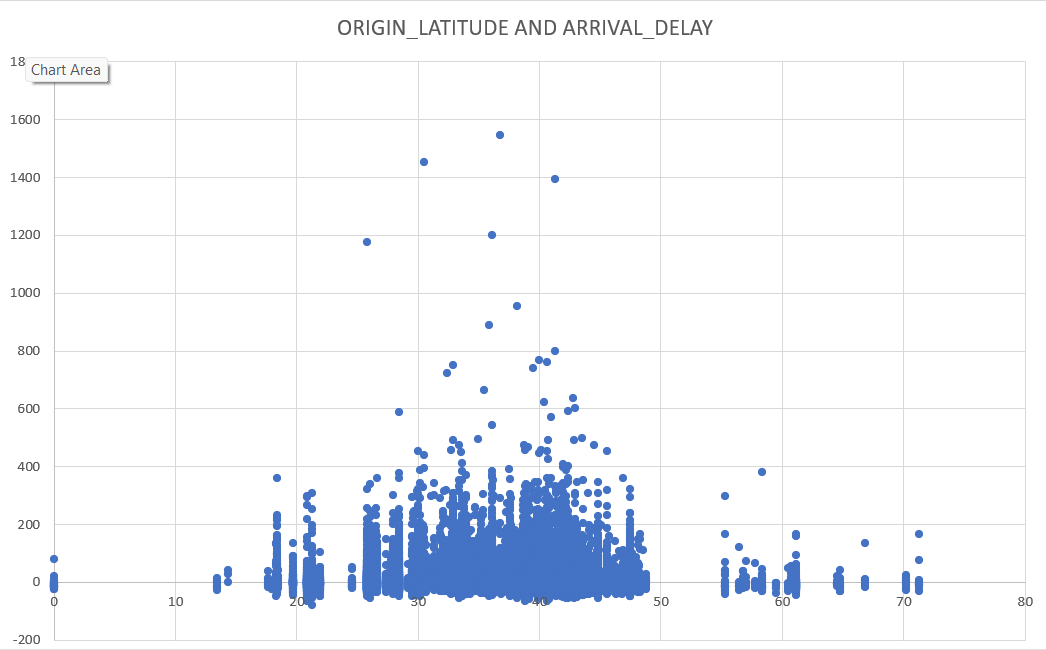
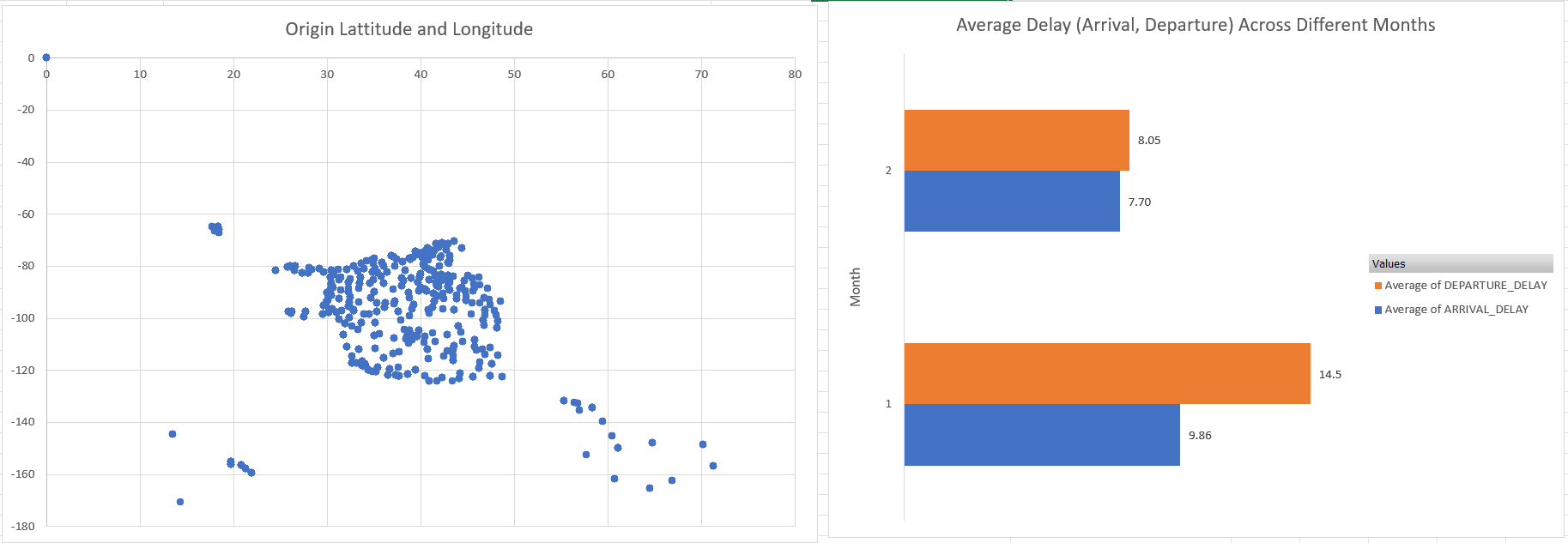
**Green** (or lighter shades): Indicate higher efficiency (ratios closer to 1).

**Yellow** to **Red** (or darker shades): Indicate lower efficiency (higher ratios).



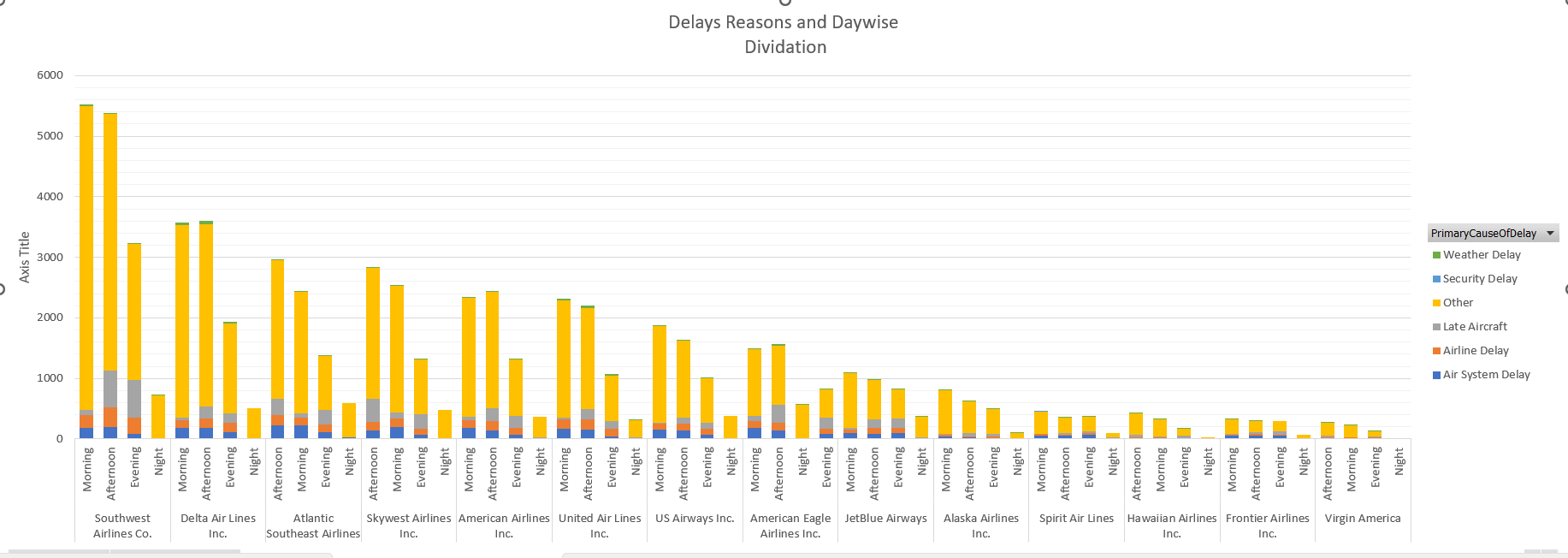
19.**Airport Performance and Environmental Factors Correlation**: Investigate if there's a correlation between airport performance (in terms of delays) and environmental factors like location (latitude and longitude from "airports.csv") and time of year (seasonal weather conditions). Are certain airports more prone to delays due to their geographical location and the associated weather patterns?

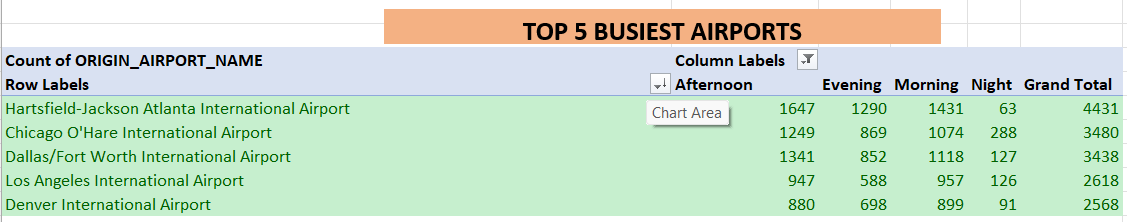
The correlations between geographical coordinates (latitude and longitude) of airports and flight delays (departure or arrival) are close to zero, indicating no significant linear relationship. Similarly, there's no significant linear relationship between flight distance and delays. These findings suggest that factors other than geographical location or distance have a greater influence on flight delays.

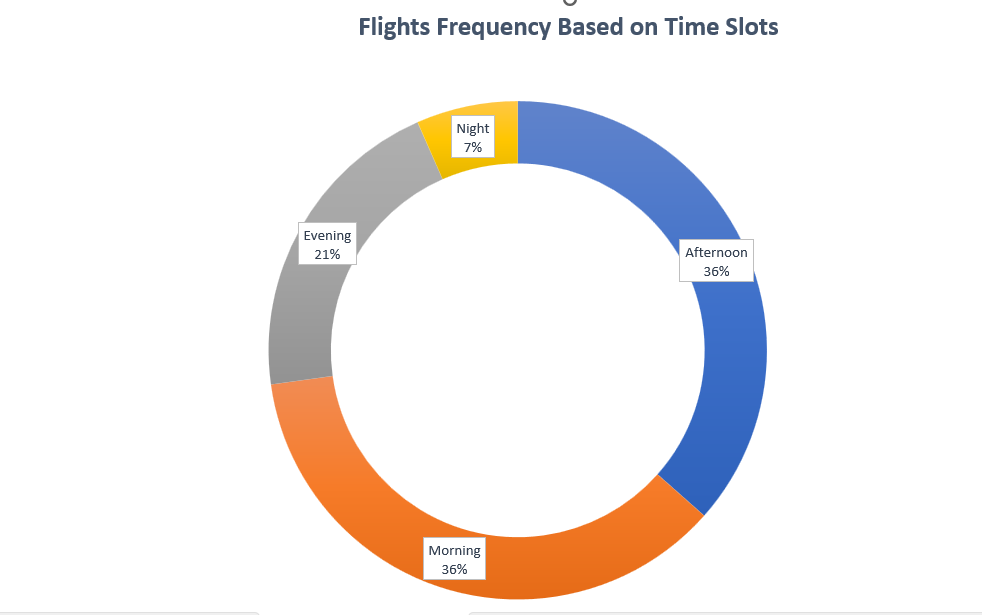


1. **Complex Delay Cause Analysis**: Use nested functions to analyze the primary cause of delays for each airline. Determine if the predominant cause of delay (like airline delay, weather delay, security delay) varies by airline and time of day.

there is no predominant cause of delay across airlines and times of day. Each airline experiences various delay causes, including air system delay, airline delay, late aircraft, security delay, and weather delay. The frequency of delay causes varies by airline and time of day, suggesting that there is no single factor consistently responsible for delays across the industry



1. **Analysis of Flight Frequency and Peak Hours**: Analyze the flight frequencies to determine the peak operating hours for major airports. (Use a combination of Excel functions to categorize flights into different time slots (e.g., morning, afternoon, evening, night) and calculate the number of flights in each slot for the top 5 busiest airports.)



Morning: Approximately 33.06% of flights occur between 6:00 AM and 11:59 AM.

Afternoon: Approximately 36.51% of flights occur between 12:00 PM and 5:59 PM.

Evening: Approximately 25.70% of flights occur between 6:00 PM and 11:59 PM.

Night: Approximately 4.73% of flights occur between 12:00 AM and 5:59 AM.