Projects

- Data Analytics and Information Dashboards for the dataset in the selected domain

- An Empowering Data-Driven Decisions with Tableau: A Comprehensive Analysis of dataset to Reveal Hidden Relationships

Author Name: Darssh Anand Vejayalakshme

Application: Tableau

Table of Contents

1. About the dataset

The dataset named "flights.csv" offers comprehensive information on the flights operated by an airport. This dataset contains comprehensive information on flight details, including the scheduled departure and arrival times, any delays encountered, the airline operating the flight, the flight number, the origin and destination of the trip, the duration of the flight, the distance covered, the specific hour and minute of the flight, and the precise date and time of the flight.

This data may be utilized for management research and strategic decision-making, offering valuable insights on flight performance and the effectiveness of placement businesses. This data can also be utilized by airport authorities in USA airports for the effective operation improvements.

Link to the dataset : https://www.kaggle.com/datasets/mahoora00135/flights/data

The data analysis in this dataset can serve as a foundation for the following analysis.

- Analysis on Low Season Months (February, March, October, November)
- Analysis of Departures (in Hours)
- Analysis on flight arrivals followed by average of arrival time in minutes and average distance in miles.
- Analysis on Total Distance travelled based on Origin Airport departures by Month
- Analysis of Distance travelled by Airlines on Total, Average, and Maximum distance covered
- Analysis on Total Distance covered by airlines based on different time frame.
- Analysis of total delay and arrivals based on origin airport and destination airport.
- Airtime Analysis in Minutes for the year 2013

Nr	Fields	Description
1	Id	A unique identifier for each flight record in the dataset.
2	Year	The year in which the flight took place (2013 in this dataset).
3	Month	The month in which the flight took place (1 to 12).
4	Day	The day of the month on which the flight took place (1 to 31).
5	Dep_time	The actual local departure time of the flight, in 24-hour format (hhmm).
6	Sched_dep_time	The scheduled local departure time of the flight, in 24-hour format (hhmm).
7	Dep_delay	The difference between the actual and scheduled departure times of the flight, in minutes. A positive value indicates a delayed departure, while a negative value indicates an early departure
8	Arr_time	The actual local arrival time of the flight, in 24-hour format (hhmm).
9	sched_arr_time	The scheduled local arrival time of the flight, in 24-hour format (hhmm).
10	arr_delay	The difference between the actual and scheduled arrival times of the flight, in minutes. A positive value indicates a delayed arrival, while a negative value indicates an early arrival.
11	carrier	The two-letter code of the airline carrier for the flight.
12	flight	The unique flight number of the flight.
13	tailnum	The unique identifier of the aircraft used for the flight.
14	origin	The three-letter code of the airport of origin for the flight.
15	dest	The three-letter code of the destination airport for the flight.
16	air_time	The duration of the flight, in minutes.
17	distance	The distance between the origin and destination airports, in miles.
18	hour	The hour component of the scheduled departure time, in local time
19	minute	The minute component of the scheduled departure time, in local time.
20	time_hour	The scheduled departure time of the flight, in local time and format (yyyy-mm-dd hh:mm:ss).
21	name	The name of the airline carrier for the flight.

Table 1 : Dataset fields and its description.

2. Analytics Task

Problem Statement: The provided dataset contains flight information, including details such as departure and arrival times, delays, carrier details, flight distance, and more. The primary goal of analysing this dataset is to understand flight performance and identify key patterns affecting flight delays.

Focus Groups: Airline / Carrier decision making authorities and Airport authorities in the United States may also make use of this data to enhance their operations more effectively.

Key Questions and Issues to Explore:

1. **Flight Delay Analysis / Patterns:**
    - Which flights or carriers experience the most frequent delays?
    - What are the primary factors contributing to delays (e.g., weather, carrier, origin, destination)?
    - How does the time of day or day of the week impact flight delays?
    - Are delays more common for specific routes or destinations?
2. **Carrier Performance:**
    - Which airlines have the best and worst on-time performance?
    - Are there specific routes where certain airlines consistently perform better or worse?
3. **Airport Performance:**
    - Which airports experience the most delays?
    - Is there a pattern of delays specific to certain origins or destinations?
4. **Flight Time Patterns:**
    - How do flight times vary across different routes?
    - Are there specific times or seasons when flights are generally faster or slower?
5. **Correlation Between Flight Distance and Delays:**
    - Is there a significant relationship between the distance of a flight and its likelihood of being delayed?
    - Do longer flights tend to have longer delayed?

Possible insights from these questions:

- **Airlines and Airports:** These insights help airlines and airports optimize their operations and improve customer satisfaction.
- **Passengers:** Passengers benefit from knowing which flights and airlines are more reliable, helping them make informed travel decisions.
- **Regulatory Authorities:** Understanding delays and patterns in the aviation industry aids in policy formulation and infrastructure planning.

Expected Outcomes:

- Identification of the main factors contributing to flight delays.
- An overview of which airlines and airports have the best and worst on-time performance.
- Actionable recommendations for airlines and airports to improve their efficiency and reduce delays.

We can now proceed with the analysis by diving into the data to extract actionable insights.

3. DASHBOARDS

3.1 Airtime Analysis in Minutes for the year 2013

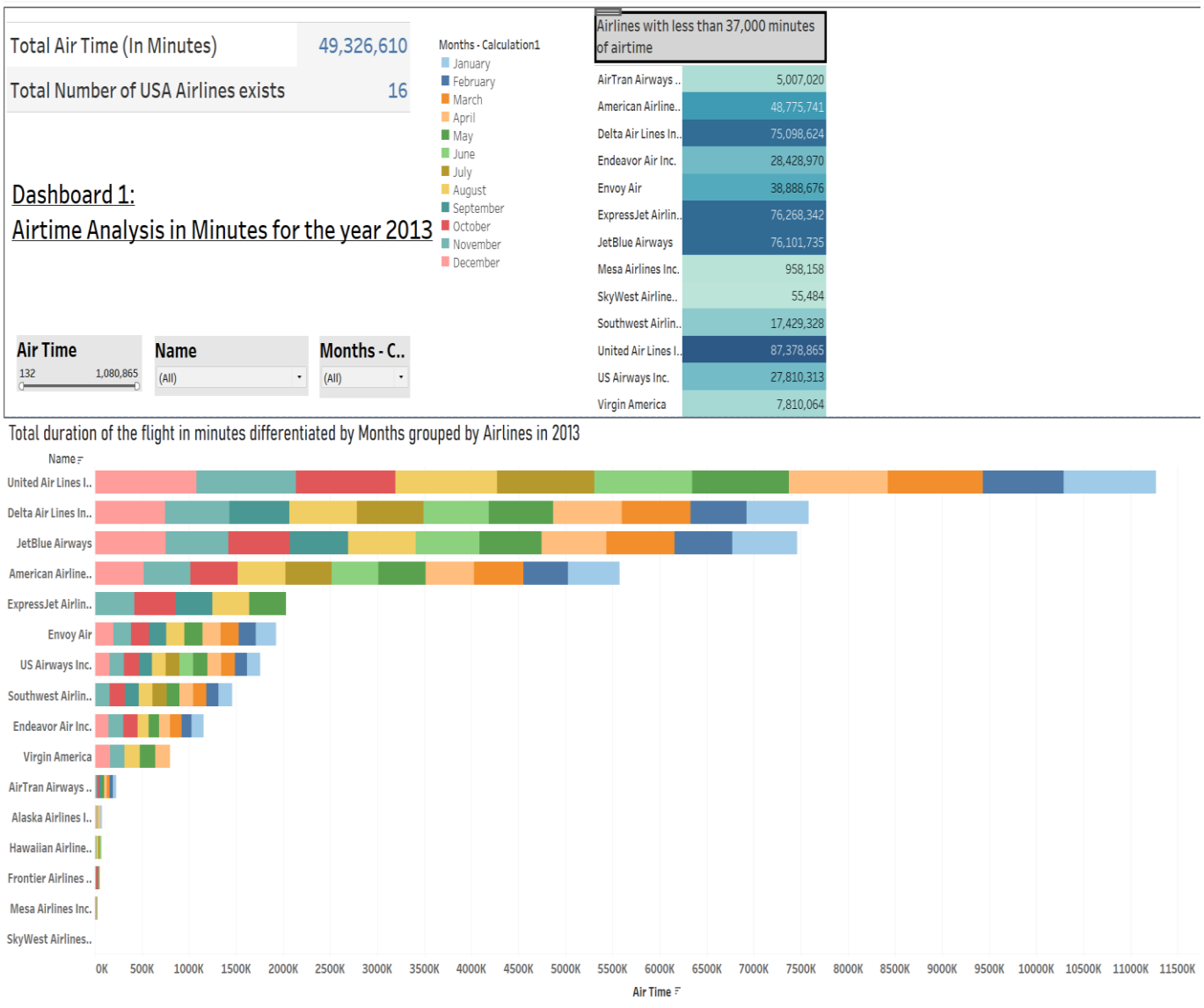


Fig 1. Dashboard 1 : Airtime Analysis in Minutes for the year 2013

1. Key Insights:

- **Airline Performance:**

United Airlines: Leads in total airtime, signifying extensive operations and route coverage.

Delta and American Airlines: Follow closely behind United Airlines in airtime, reflecting their extensive flight networks.

Smaller Airlines: Airlines like Mesa Airlines, SkyWest, and Frontier show significantly less airtime, highlighting their focus on regional routes or smaller fleets.

- **Monthly Trends:**

High Airtime Months: July and August exhibit higher airtime across most airlines, consistent with the summer travel season where flight demand typically increases.

Low Airtime Months: February shows the lowest airtime, likely due to the fewer days in the month and potential adverse weather affecting flight operations.

- **Airlines with Limited Airtime:**

The list of airlines with less than 37,000 minutes of airtime includes AirTran Airways, Mesa Airlines, and SkyWest Airlines, among others. These airlines have specialized operations focusing on specific regions or markets, resulting in fewer total flight hours.

2. Actionable Intelligence:

- **Operational Efficiency:** Major airlines can use these insights to optimize fleet usage and adjust scheduling to maintain high service levels during peak seasons.
- **Fleet Strategy:** Smaller airlines should assess opportunities for expansion based on these insights, possibly targeting underserved routes.
- **Resource Allocation:** Airports should allocate more resources during peak travel months (July and August) to handle the increased passenger traffic efficiently.
- **Operational Planning:** Airports can identify airlines with consistent flight operations and adjust resource planning accordingly.
- **Travel Planning:** Passengers can plan travel during off-peak months to avoid crowds and flight delays.
- **Airline Selection:** The data helps passengers identify airlines with consistent and extensive operations, which can aid in selecting reliable flights.

Summary: The dataset focuses on flights from the year 2013, providing detailed airtime information for different airlines. The dashboard visualizes airtime across various airlines, broken down by months and overall total minutes. This analysis offers stakeholders an overview of airline airtime patterns, helping them optimize operations, resource allocation, and travel plans based on the data-driven insights.

3.2 Analysis of total delay and arrivals based on origin airport and destination airport.

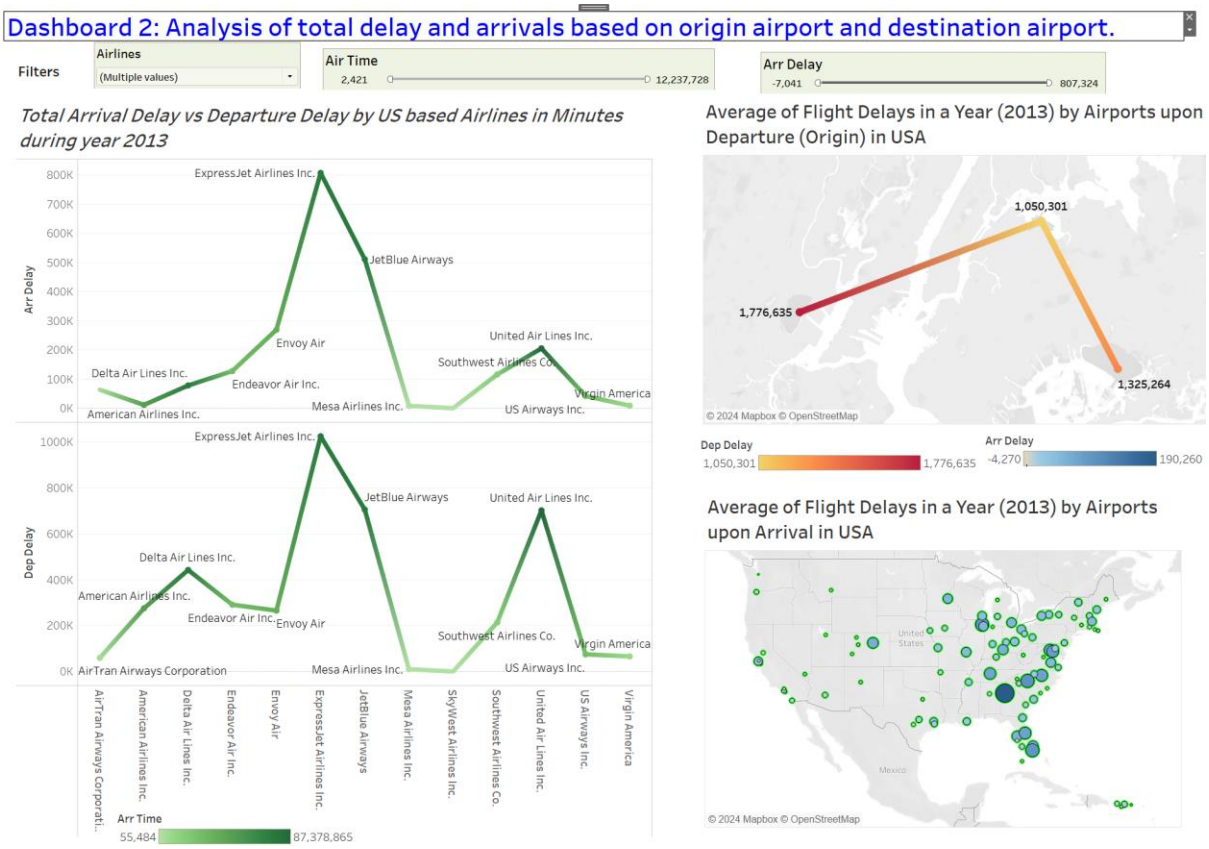


Fig 2: Dashboard 2: Analysis of total delay and arrivals based on origin airport and destination airport.

Key Insights

1. **Flight Delays by Airline:**
    - **ExpressJet Airlines:** Shows the highest average arrival delay, indicating potential operational issues or scheduling challenges.
    - **JetBlue Airways and United Airlines:** Also have notable arrival delays but relatively lower than ExpressJet.
    - **Mesa Airlines and Endeavor Air:** Although smaller airlines, they display significant departure delays, highlighting operational challenges.
2. **Flight Delays by Airports:**
    - **Departure Delays:** High departure delays are observed at specific airports, such as the one with over 1.7 million minutes of departure delay.

- **Arrival Delays:** Some airports experience significantly high arrival delays, with one airport exceeding 1.3 million minutes in 2013.

3. **Geographic Distribution of Delays:**

    - The geographical maps highlight clusters of delays across U.S. airports.
    - Airports in the Eastern and Central regions of the U.S. tend to have higher delays on average than those on the West Coast.

Actionable Insights

1. **Airlines:**

    - **Operational Efficiency:** Airlines with consistently high delays should investigate root causes and improve operations, including better scheduling and fleet management.
    - **Route Optimization:** Airlines may need to reconsider routes consistently associated with high delays and develop strategies to mitigate delays.

2. **Airports:**

    - **Resource Allocation:** Airports with high delays should enhance resource allocation to reduce congestion and improve flight punctuality.
    - **Infrastructure Planning:** Airports with recurrent delays should invest in infrastructure to address bottlenecks and improve operational efficiency.

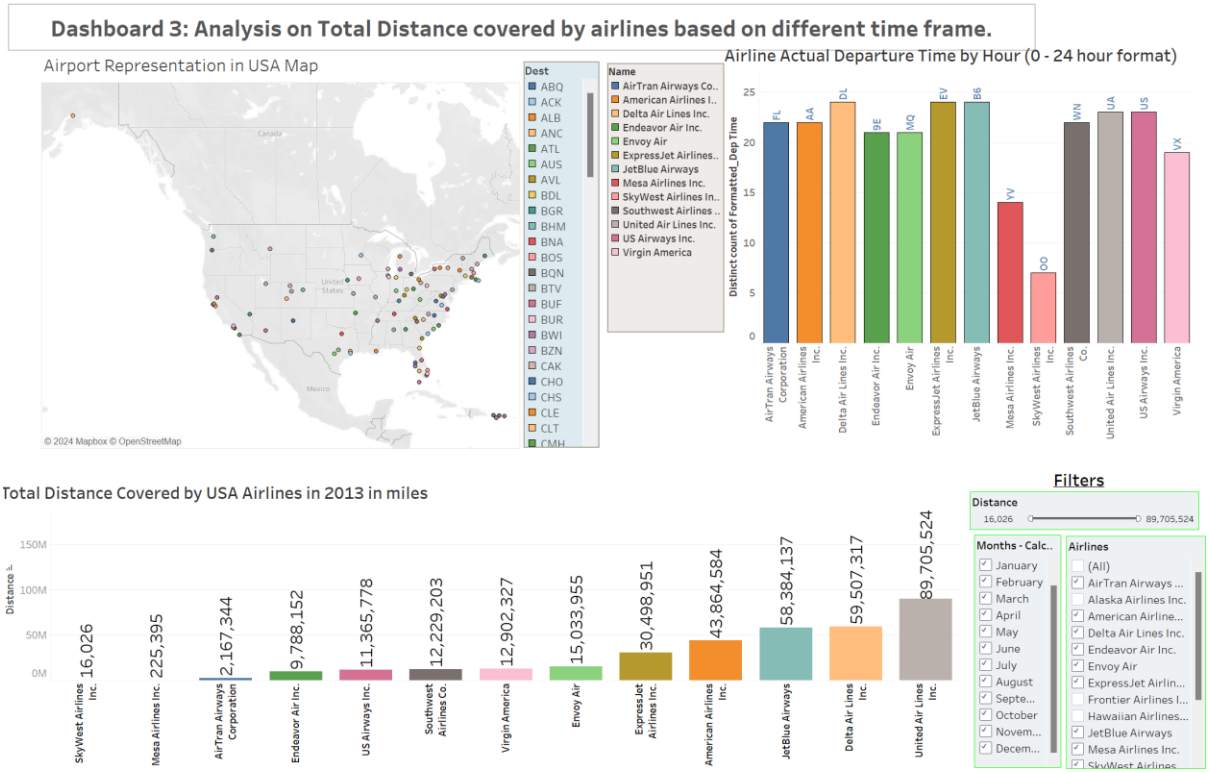
3. **Passengers:**

    - **Travel Planning:** Passengers can choose flights from airports with lower delays to minimize travel disruptions.
    - **Flexible Scheduling:** Travelers should allow flexibility in their schedules, particularly when flying to/from airports with a history of significant delays.

Summary

The dataset offers comprehensive insights into flight delays across different airlines and airports in the U.S. during 2013. The dashboard includes visualizations of average flight delays for both departure and arrival across various U.S. airports and airlines. The analysis provides a detailed overview of the delay trends across various airlines and airports, enabling stakeholders to identify operational inefficiencies and potential improvement areas.

3.3 Analysis on Total Distance covered by airlines based on different time frame.



Key Insights

1. **Total Distance Covered by Airlines:**

United Airlines: Leads by covering the highest total distance, over 89 million miles in 2013. This shows its extensive route network.

Delta Air Lines: Comes second in terms of distance covered, reflecting its vast global reach.

Smaller Airlines: Airlines like Mesa Airlines and Virgin America cover significantly fewer miles, indicating their focus on regional or niche markets.

2. **Airtime Analysis by Airline:**

United Airlines: Again, leads with the highest total airtime in minutes, confirming its extensive operations.

JetBlue Airways and **Delta Air Lines:** Follow closely behind, showing similar extensive operations.

3. **Monthly Airtime Distribution:**

Different months show variable airtime distributions across airlines, which might reflect seasonal travel patterns and airline schedules. Airlines like United, Delta, and American Airlines show more consistent airtime month over month compared to smaller airlines.

4. **Actual Departure Times by Airline:**

The graph on actual departure times suggests a varied pattern in departure schedules across different airlines. Airlines like AirTran Airways, American Airlines, and Delta Air Lines have a relatively higher count of flights departing throughout the day compared to other carriers, showcasing their widespread network.

Actionable Insights

1. **Airlines:**
    - **Operational Efficiency:** Major airlines can use these insights to optimize fleet usage and adjust scheduling to maximize efficiency in their flight operations.
    - **Fleet Strategy:** Smaller airlines can find growth opportunities by analysing peak periods and under-served routes.
2. **Airports:**
    - **Resource Allocation:** Airports can allocate more resources during peak travel months and times to handle the increased passenger traffic efficiently.
    - **Operational Planning:** Airports should identify airlines with consistent flight operations to optimize resource planning.
3. **Passengers:**
    - **Travel Planning:** Passengers can plan their travel during off-peak months to avoid crowds and potential delays.
    - **Airline Selection:** The data helps passengers identify airlines with consistent and extensive operations, helping them choose more reliable flights.

Summary

The dataset provides comprehensive data on flights across different airlines in the U.S. in 2013. The dashboard visualizes key metrics, such as airtime distribution, actual departure times, and total distance covered by each airline. Overall, this analysis offers a holistic view of airline operations across the U.S., providing valuable insights to optimize operations and enhance travel experience.

3.4 Analysis of Distance travelled by airlines on Total, Average, and Maximum distance covered.

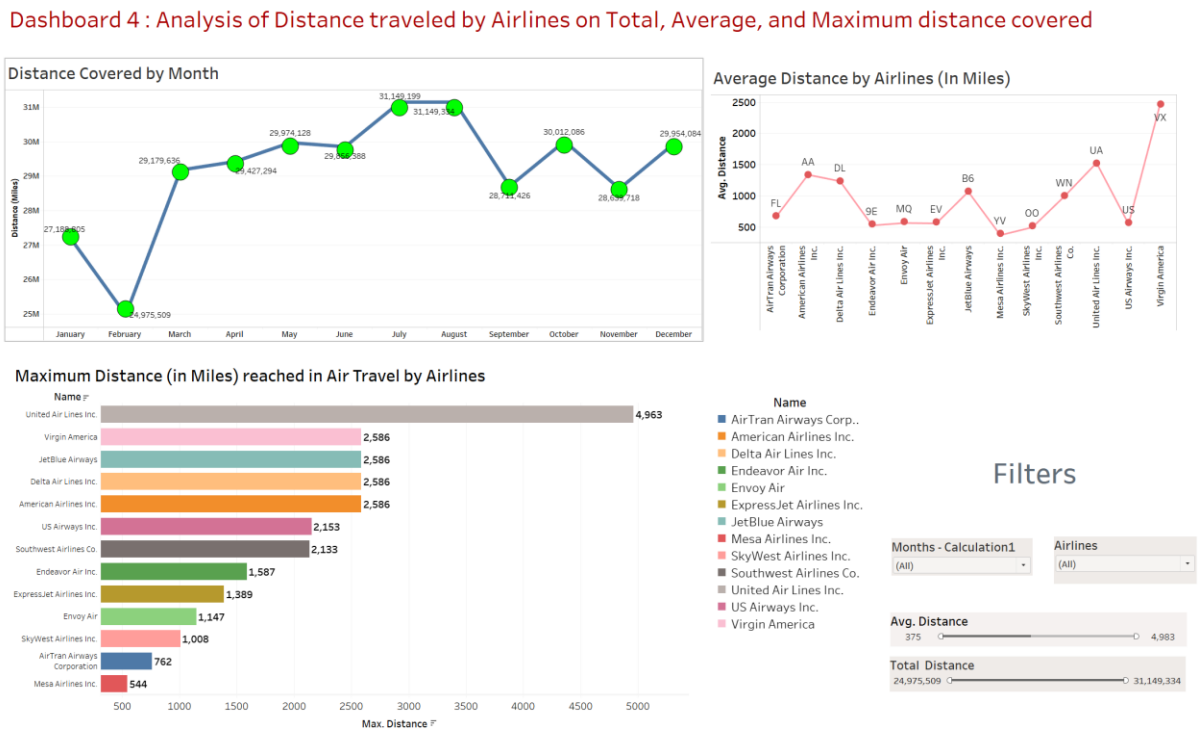


Fig 3: Dashboard 4 - Analysis of Distance travelled by airlines on Total, Average, and Maximum distance covered.

Key Insights

1. **Total Distance Covered by Month:**
    - **July and August:** These months have the highest total distances covered, exceeding 31 million miles, reflecting peak summer travel demand.
    - **February:** Has the lowest total distance covered at just under 25 million miles, likely due to its fewer days and potential weather disruptions.
2. **Average Distance by Airlines:**
    - **Virgin America:** Shows the highest average flight distance, indicating its focus on long-haul routes.
    - **US Airways and United Airlines:** Follow closely behind in terms of average distance, showcasing their extensive operations across longer routes.

3. **Maximum Distance in Air Travel by Airlines:**

    - **United Airlines:** Has the longest single flight distance, reaching nearly 5,000 miles, showing its reach across continents.
    - **Virgin America, JetBlue, Delta, and American Airlines:** Have a similar maximum distance, each reaching up to 2,586 miles, indicating their transcontinental operations.

Actionable Insights

1. **Airlines:**

    - **Route Optimization:** Airlines should assess which routes provide the most efficient use of their fleet and identify opportunities to expand their long-haul services.
    - **Fleet Planning:** Airlines with significant long-haul operations need to optimize their fleet to manage the high utilization rates efficiently.

2. **Airports:**

    - **Resource Allocation:** Airports need to allocate more resources during peak travel months (July and August) to handle the increased flight distances and traffic.
    - **Operational Planning:** Airports should enhance infrastructure and services to support airlines with significant long-haul operations.

3. **Passengers:**

    - **Travel Planning:** Passengers can expect more flight options during peak travel months and plan their trips accordingly.
    - **Route Selection:** Passengers looking for long-haul flights can identify airlines with extensive operations and reliable service.

Summary

The dataset provides an analysis of flight distances covered by various airlines in 2013. The dashboard visualizes total, average, and maximum distances across different airlines and months, providing comprehensive insights into flight operations. The analysis provides detailed insights into the flight distance patterns of different airlines, enabling stakeholders to optimize their operations and travel plans effectively.

3.5 Analysis on Total Distance travelled based on Origin Airport departures by Month

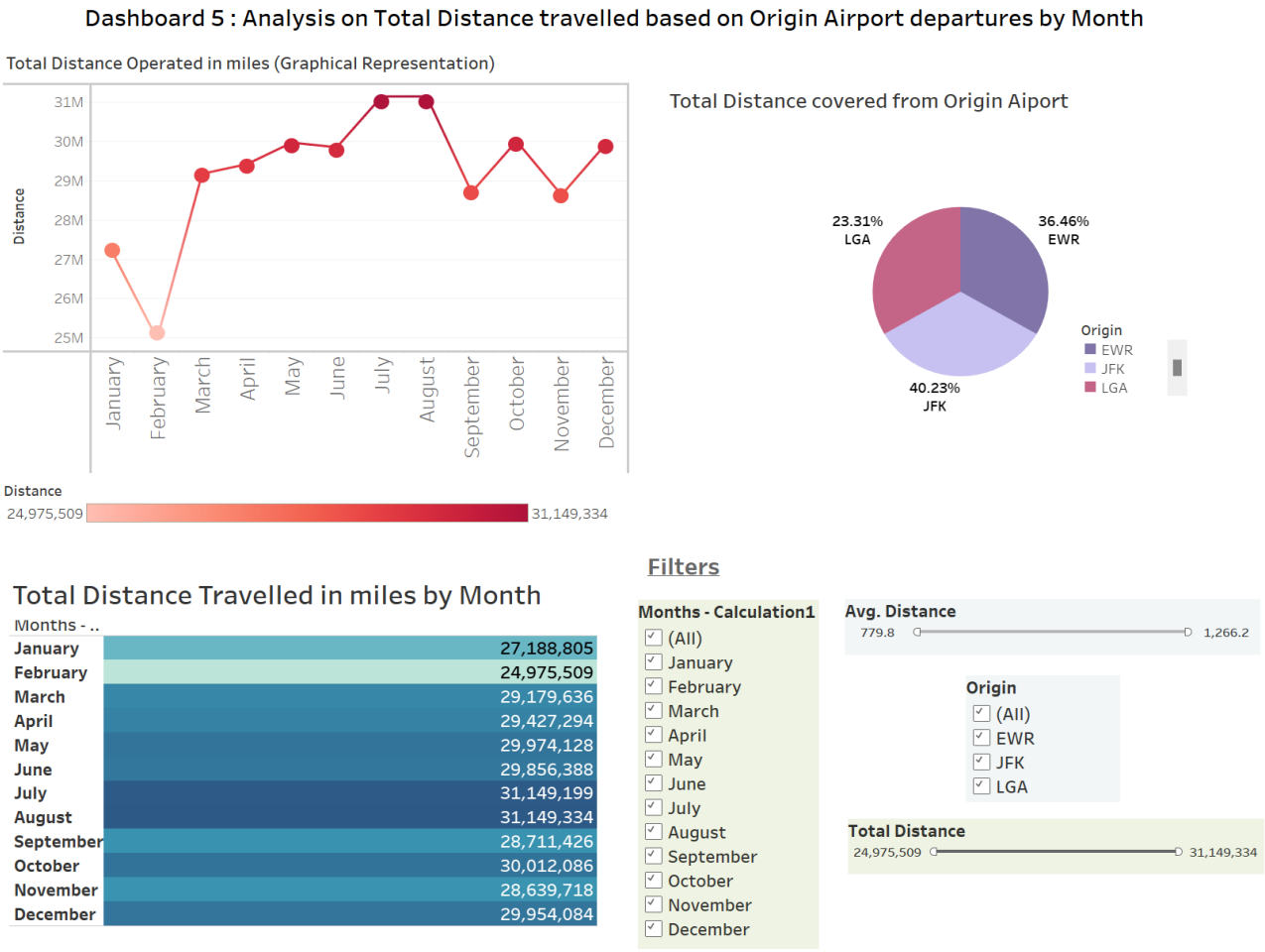


Fig 5: Dashboard 5 - Analysis on Total Distance travelled based on Origin Airport departures by Month

Key Insights

1. **Total Distance Travelled by Month:**

   - **July and August:** Represent the peak travel months, with each exceeding 31 million miles travelled, indicating the summer travel season's impact.
   - **February:** Has the lowest total distance covered, just under 25 million miles, attributed to fewer days and potential seasonal weather challenges.
   - **Month-to-Month Variability:** There is a noticeable drop in travel distance in September and October, suggesting a lull after the peak summer travel.

2. **Distance Distribution by Origin Airport:**

    - **JFK:** Accounts for the largest share of the total distance covered, with over 40% of the flights originating from this airport.
    - **EWR and LGA:** Newark (EWR) follows closely with around 36% of the total distance, while LaGuardia (LGA) contributes about 23%.

Actionable Insights

1. **Airlines:**

    - **Route Planning:** Airlines should optimize flight schedules around peak months (July and August) to maximize utilization of their fleet.
    - **Regional Focus:** Given JFK's dominance, airlines should consider adding more capacity to routes originating from this airport.

2. **Airports:**

    - **Resource Allocation:** Airports should prepare for the high traffic in summer and plan for peak travel months.
    - **Operational Planning:** Airports can use this data to streamline operations and manage flight schedules effectively.

3. **Passengers:**

    - **Travel Planning:** Passengers planning summer trips should book in advance to avoid congestion and delays.
    - **Airport Selection:** Travelers might prefer JFK due to its more extensive route network.

Summary

The dataset provides a detailed view of flight distances, focusing on total miles covered per month based on the origin airport. The visualization also highlights the distribution of miles travelled from specific origin airports (EWR, JFK, and LGA). This analysis offers valuable insights into how flight distances vary across months and origin airports. The data suggests specific times and routes where airlines and airports can focus their planning and resource allocation.

3.6 Analysis on Low Season Months (February, March, October, November)

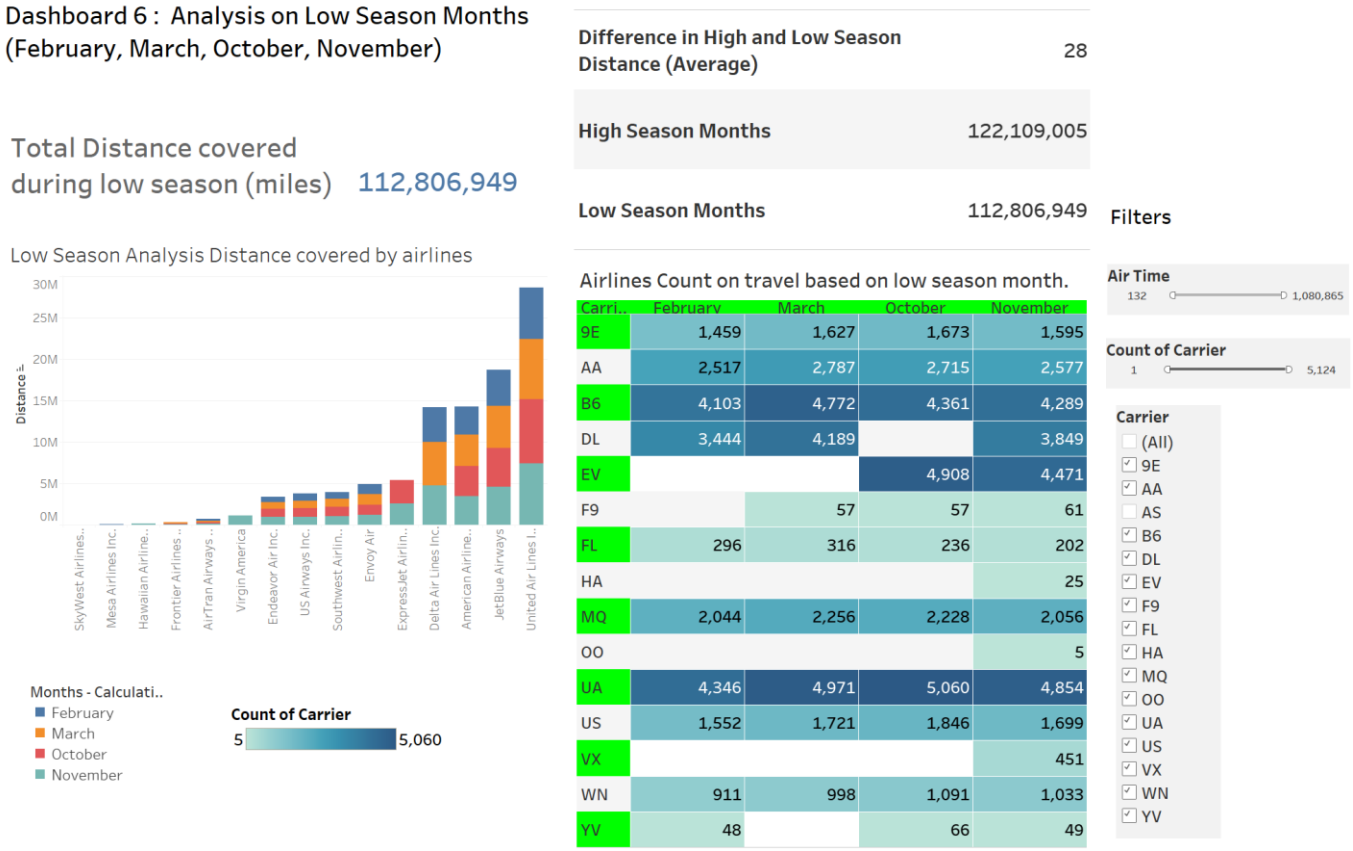


Fig 6: Dashboard 6 - Analysis on Low Season Months (February, March, October, November)

Key Insights

1. **Total Distance Covered in Low Season:**
    - A total of approximately 112.8 million miles were covered by all airlines during low-season months.
    - United Airlines, JetBlue Airways, and Delta Air Lines were the leading carriers in terms of miles flown during these periods.
2. **Difference in Distance Between High and Low Seasons:**
    - The difference between high-season and low-season distance is approximately 28 million miles, showing a significant drop in travel during the low season.
    - This reflects the lower travel demand in these periods compared to peak months like July and August.

3. **Airlines Count on Travel During Low Season:**

    - JetBlue Airways (B6) had the highest count of flights across the low-season months, indicating consistent operations despite the low season.
    - United Airlines (UA) also maintained a high number of flights, particularly in October.
    - Other airlines like Delta (DL), Envoy Air (EV), and American Airlines (AA) maintained a consistent flight count, though lower compared to peak months.

Actionable Insights

1. **Airlines:**

    - **Fleet Management:** Airlines can optimize their fleet usage by reducing the number of aircraft operating on certain routes during low-season months.
    - **Marketing Strategies:** To mitigate the reduction in travel, airlines can implement promotions and offers to attract more travellers during low seasons.

2. **Airports:**

    - **Operational Efficiency:** Airports can use these insights to adjust staffing and services during low-season months to improve cost-efficiency.
    - **Resource Allocation:** Resources can be reallocated to focus on high-demand services, potentially reducing operating costs.

3. **Passengers:**

    - **Travel Planning:** Passengers can leverage low season travel for more affordable flights and reduced congestion.
    - **Flight Options:** Travelers have a broader selection of flights from airlines like JetBlue and United during low seasons due to their consistent operations.

Summary:

The dataset focuses on analysing flight distances and airline operations during low-season months, namely February, March, October, and November. The dashboard shows total distance covered in miles, the difference between high and low season distances, and counts of flights by airlines. The analysis provides detailed insights into the travel patterns of airlines during low-season months, revealing valuable opportunities for optimization in fleet management, marketing, and operations.

3.7 Analysis of Flight Departures (in Hours)

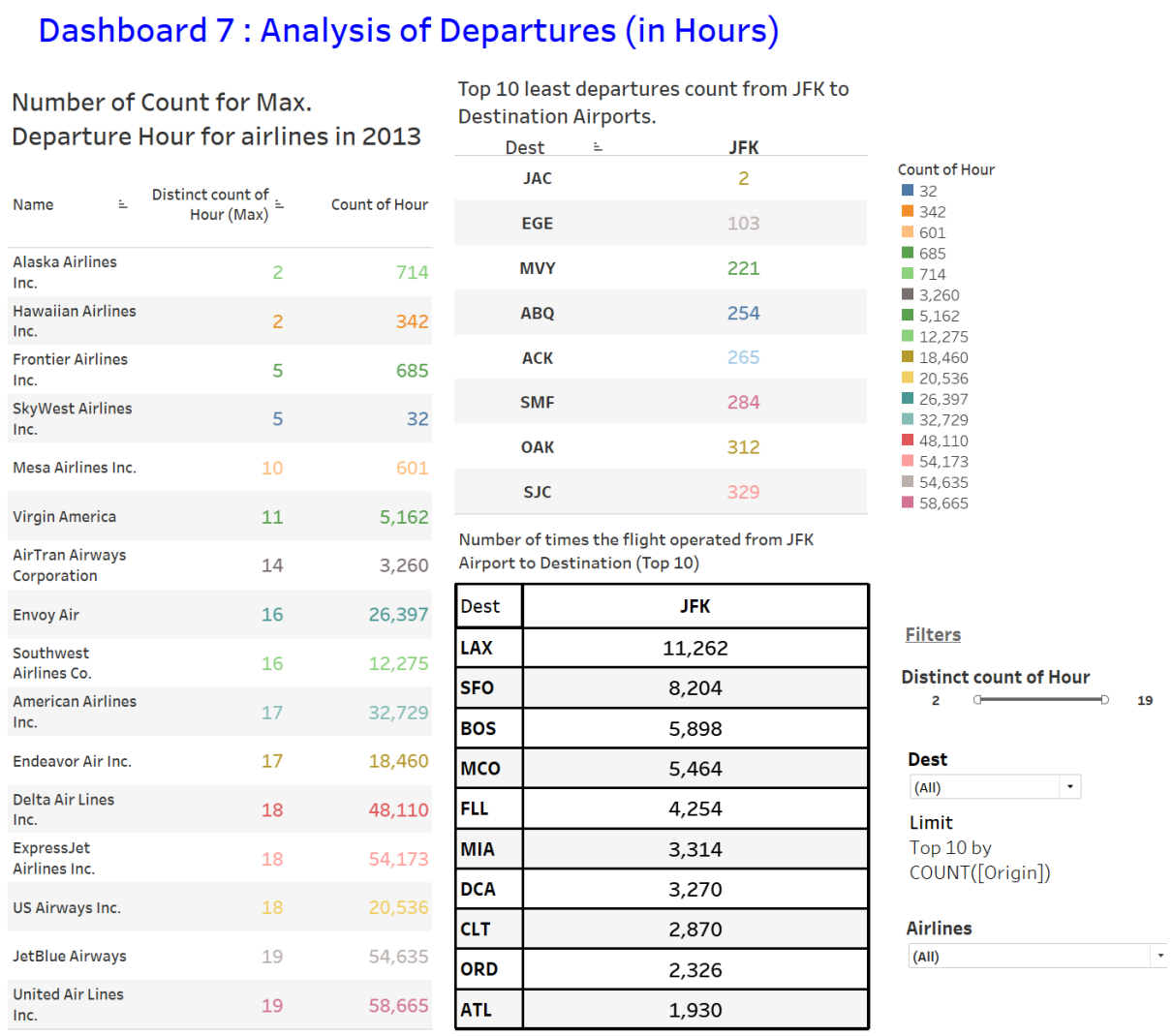


Fig 7 : Dashboard 7 - Analysis of Departure

Key Insights

1. **Departure Hour Analysis:**

   - **United Airlines and JetBlue Airways:** Both airlines have the highest distinct count of departure hours, suggesting extensive operations with frequent departures.
   - **Airlines with Fewer Departure Hours:** Airlines like Alaska, Hawaiian, and Frontier have fewer departure hours, reflecting limited or more focused flight schedules.

2. **Top 10 Least Departures from JFK:**

- **Airports with the Fewest Flights:** Airports such as Jackson Hole (JAC), Eagle (EGE), and Martha's Vineyard (MVY) have the fewest departures from JFK, which may indicate niche or seasonal routes.
- **Domestic Routes:** Most of these routes are domestic, suggesting that certain destinations are not served frequently from JFK.

3. **Top 10 Destinations from JFK:**

- **Popular Domestic Routes:** Los Angeles (LAX) and San Francisco (SFO) are the most frequently served destinations from JFK, highlighting their importance for business and tourism.
- **Key East Coast Destinations:** Boston (BOS) and Orlando (MCO) are also popular destinations from JFK, indicating high passenger demand.

Actionable Insights

1. **Airlines:**
    - **Route Expansion:** Airlines with fewer departure hours should consider expanding to popular routes, like LAX and SFO, to increase operations.
    - **Flight Frequency:** Airlines should adjust their flight frequency to align with passenger demand, particularly for high-traffic routes.
2. **Airports:**
    - **Resource Allocation:** JFK should allocate resources efficiently to handle the high traffic on routes to LAX, SFO, and other key destinations.
    - **Route Planning:** Airports with few flights from JFK should collaborate with airlines to enhance route connectivity.
3. **Passengers :** Passengers can plan travel more effectively by knowing which destinations have frequently flights and adjusting their schedules accordingly. They can also find opportunities to visit less frequently served destinations by planning trips around available flights.

Summary: The dataset provides an analysis of airline departures in terms of hours. It covers the count of maximum departure hours for various airlines, top 10 least departures from JFK to destination airports, and the top 10 destinations from JFK by number of flights. The analysis reveals key patterns in airline departure hours and popular flight routes, providing valuable insights into optimizing airline operations, airport resources, and passenger travel plans.

3.8 Analysis on flight arrivals followed by average of arrival time in minutes and average distance in miles.

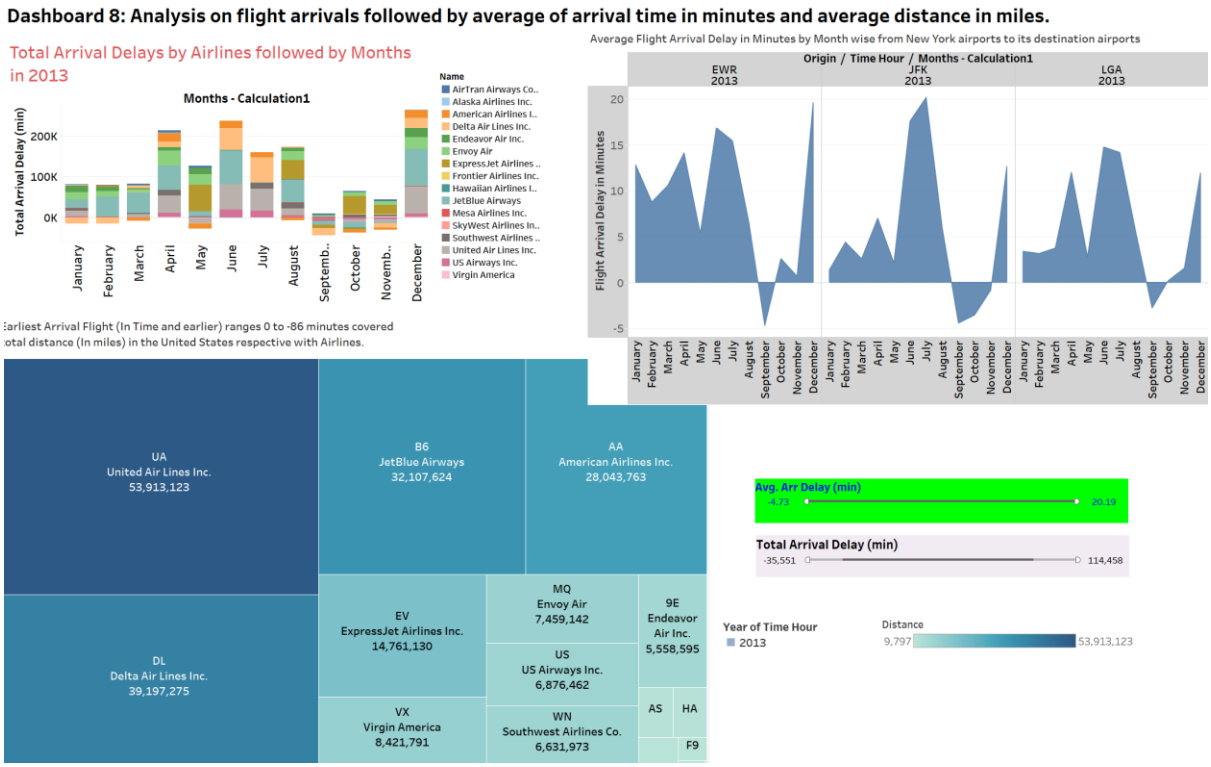


Fig 8 : Dashboard 8 - Analysis on flight arrivals followed by average of arrival time in minutes and average distance in miles.

Key Insights

1. **Total Arrival Delays by Airlines and Months:**
    - **United Airlines (UA):** Has the highest total arrival delay, with delays exceeding 50 million minutes, reflecting its extensive operations and potential for congestion.
    - **Delta Air Lines (DL) and JetBlue Airways (B6):** Also report significant delays, each exceeding 30 million minutes, suggesting operational challenges.
    - **Monthly Trends:** The summer months, particularly June and July, show notable spikes in arrival delays across most airlines, likely due to higher travel demand.
2. **Average Flight Arrival Delay:**
    - **Monthly Variations:** Significant variations in delay times across months, with peaks observed in June and September for airports like JFK and EWR.
    - **Airport Variations:** Each airport (EWR, JFK, LGA) exhibits different delay patterns, indicating specific operational or traffic challenges.

3. **Flight Distance and Arrival Delays:**

    - **United Airlines (UA):** Also leads in terms of flight distance, suggesting a strong correlation between long-distance flights and higher delays.
    - **JetBlue Airways (B6) and Delta Air Lines (DL):** Follow closely behind, reinforcing the relationship between flight distance and arrival delays.

Actionable Insights

1. **Airlines:**

    - **Operational Efficiency:** Airlines with high delays should evaluate their flight scheduling and operational strategies to reduce congestion and improve punctuality.
    - **Route Optimization:** Focusing on improving long-haul routes with significant delays can help improve overall arrival performance.

2. **Airports:**

    - **Resource Allocation:** Airports with significant delays should enhance resource allocation during peak months to manage congestion.
    - **Infrastructure Planning:** Investing in infrastructure improvements can alleviate bottlenecks and enhance operational efficiency.

3. **Passengers:**

    - **Travel Planning:** Passengers should anticipate potential delays during peak travel months and plan their trips accordingly.
    - **Flight Selection:** Selecting flights with a history of better on-time performance can minimize travel disruptions.

Summary

The dataset provides an analysis of flight arrival delays, highlighting total arrival delays by airlines, average delay times, and distance in miles. The visualizations focus on the year 2013, offering insights into the delay patterns across different months and airlines. This analysis reveals valuable insights into flight delays across different airlines and airports. The data suggests opportunities to optimize airline operations, improve airport infrastructure, and inform passenger travel plans.

4. An Empowering Data-Driven Decisions with Tableau: A Comprehensive Analysis of dataset to Reveal Hidden Relationships

Variables set for creating dashboard

The following sets of variables where there might be relationship between the variables and were exploring and understanding these relationships might lead to actionable insights. The variables are represented below followed by SQL queries to perform working on a dashboard.

- **Formatted_Dep_time** : INT([Dep Time] / 100)

The variable departure time is in integer format representing hours and minutes. Executing this query results in retrieving the departure time in hours.

- **Difference in High Season and Low Season (Average)** : AVG([High Season Months]) - AVG([Low Season Months])
- **High Season Months** : IF [Months - Calculation1] IN ("December", "June", "July", "August") THEN [Distance] ELSE 0 END
- **Low Season Months:** IF [Months - Calculation1] IN ("February", "March", "October", "November") THEN [Distance] ELSE 0 END
- **Months_Calculation1:** Using switch case, the variables in numbers were changed to string representing the months.

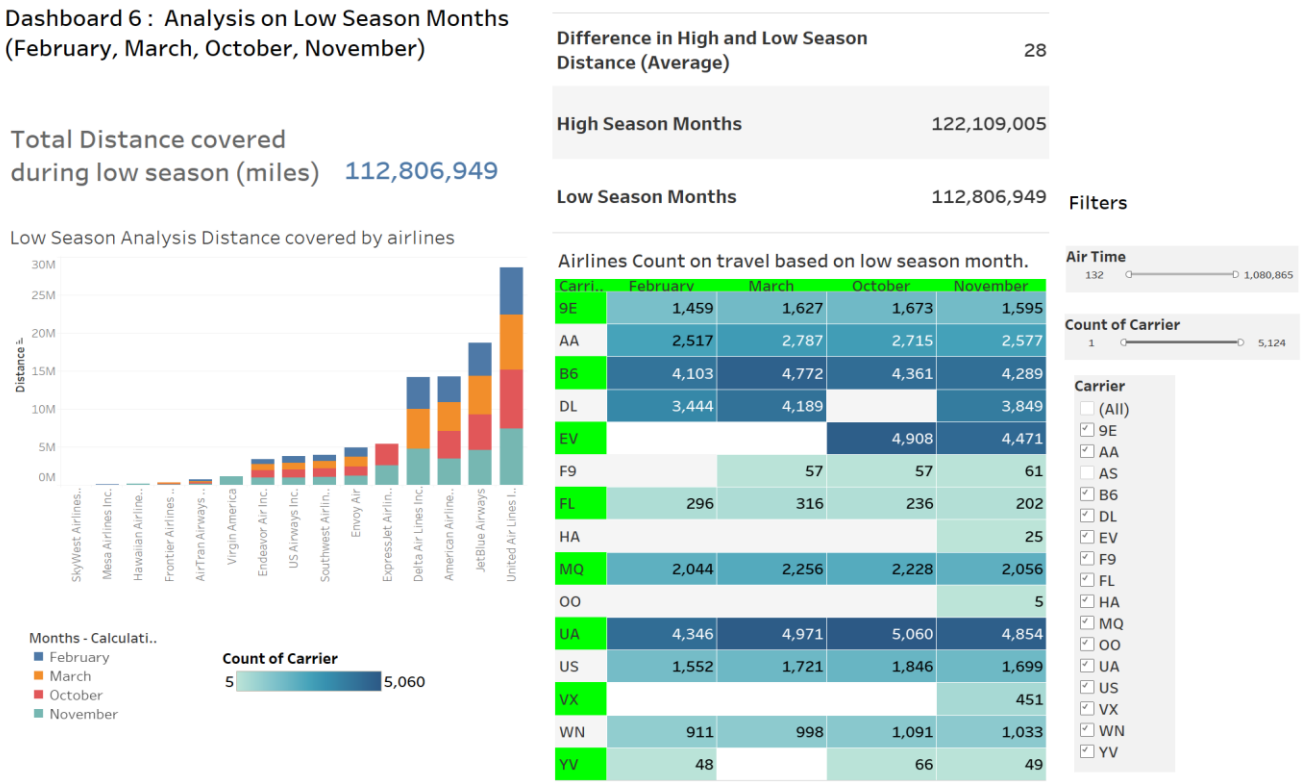


Fig 9: Dashboard - Analysis on Low season months (February, March, October, November)

Insights from Exploratory Data Analysis

Airlines with high airtime often experience significant delays, potentially due to the complexity of managing extensive networks. There is a notable increase in delays during specific months, particularly during summer. Certain airports experience higher delays than others, suggesting operational inefficiencies. Longer flight distances tend to be associated with higher delays, likely due to more complex logistics.

Possible Questions:

- What are the major operational challenges that airlines face in managing high airtime without incurring significant delays?
- What factors contribute to seasonal spikes in delays, and how can airlines anticipate and mitigate them?
- How can airports address infrastructure challenges to reduce delays, especially during peak travel times?
- How can airlines optimize long-haul flights to reduce delays?

Predictive and Prescriptive Analytics Approaches

Analysis Type	Purpose	Target Variables	Predictors
Regression Analysis	Used to quantify the impact of various factors on delays and predict future delays.	Flight delays	Airline, distance, seasonality, and airport.
Time-Series Analysis	To analyse trends over time and forecast future delays.	Monthly delays	Seasonal patterns, flight demand.
Classification Models	To categorize flights based on delay risk, allowing proactive management.	Delay risk levels.	Flight details, airline, airport.

Specific Target Audience for this dashboard : Airline operations teams, Airline planners, Airport management teams. Airline route planners. and airport authorities.

Purpose: Highlight delay patterns across different airlines to help improve operational efficiency. It shows monthly delay trends to prepare for peak travel times. Correlate flight distances with delays for optimized route planning.

Actionable Insights: Adjust scheduling to mitigate delays. Allocate more resources during peak seasons to assist customers demand. Plan capacity improvements to reduce delays. Improve flight schedules and optimize long-haul operations.

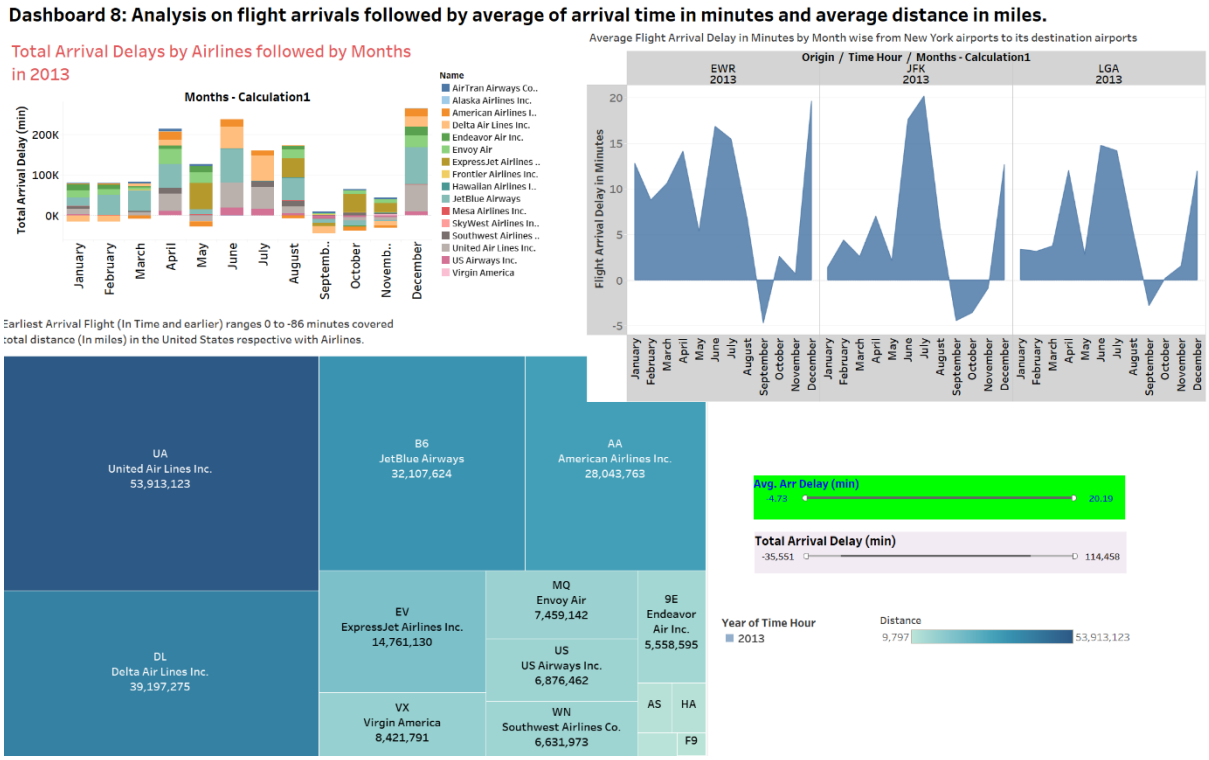


Fig 10 : Analysis on flight arrivals followed by average of arrival time in minutes and average distance in miles.

Dashboard Overview : The dashboard focuses on flight distances and delays in 2013, providing insights into average flight arrival delays, airtime distribution, and airline-specific delays.

Key Questions to Explore

1. **Airline-Specific Delay Patterns:** Which airlines have the highest delays, and what operational challenges contribute to these delays?
2. **Monthly Variations in Flight Delays:** How do delays vary across months, and what are the seasonal patterns affecting flight delays?
3. **Distance and Delay Relationship:** Is there a correlation between flight distance and delay, and how can airlines optimize routes based on this insight?
4. **Airport-Specific Delay Trends:** How do delays differ across airports, and what can be done to mitigate them?

Identified Relationships to Explore

1. **Airline and Delays:** Analysing specific airlines and their delay patterns can reveal operational inefficiencies.
2. **Month and Delays:** Understanding monthly variations can help airlines prepare for peak seasons.

3. **Distance and Delays:** Assessing the relationship between flight distance and delays can aid in route optimization.
4. **Airport and Delays:** Exploring delays across different airports can provide insights into infrastructure and capacity planning.

Predictive Data Analytics Approaches

1. **Regression Analysis:** To identify the correlation between flight distance and delays, helping airlines optimize routes.
2. **Time-Series Analysis:** For understanding seasonal patterns in delays to forecast future trends and prepare accordingly.
3. **Classification Models:** To categorize flights based on their delay risk, helping airlines proactively manage delays.

Target Audience: Airline operation managers, Airline planners, Airline route planners, airport authorities

Purpose: To identify patterns and causes of delays for specific airlines. To highlight monthly delay trends for proactive planning. To analyse delays by airport for infrastructure planning.

Actionable Insights: Identify operational bottlenecks and optimize scheduling. Plan capacity and resources based on historical trends. Improve airport efficiency by addressing capacity challenges.

Decision-Making Benefits

- **Enhanced Scheduling:** Airlines can optimize schedules based on historical delay patterns.
- **Resource Planning:** Airports and airlines can better allocate resources during peak seasons.
- **Route Optimization:** Airlines can adjust routes to minimize delays and improve efficiency.
- **Infrastructure Planning:** Airports can identify capacity issues and plan infrastructure improvements accordingly.

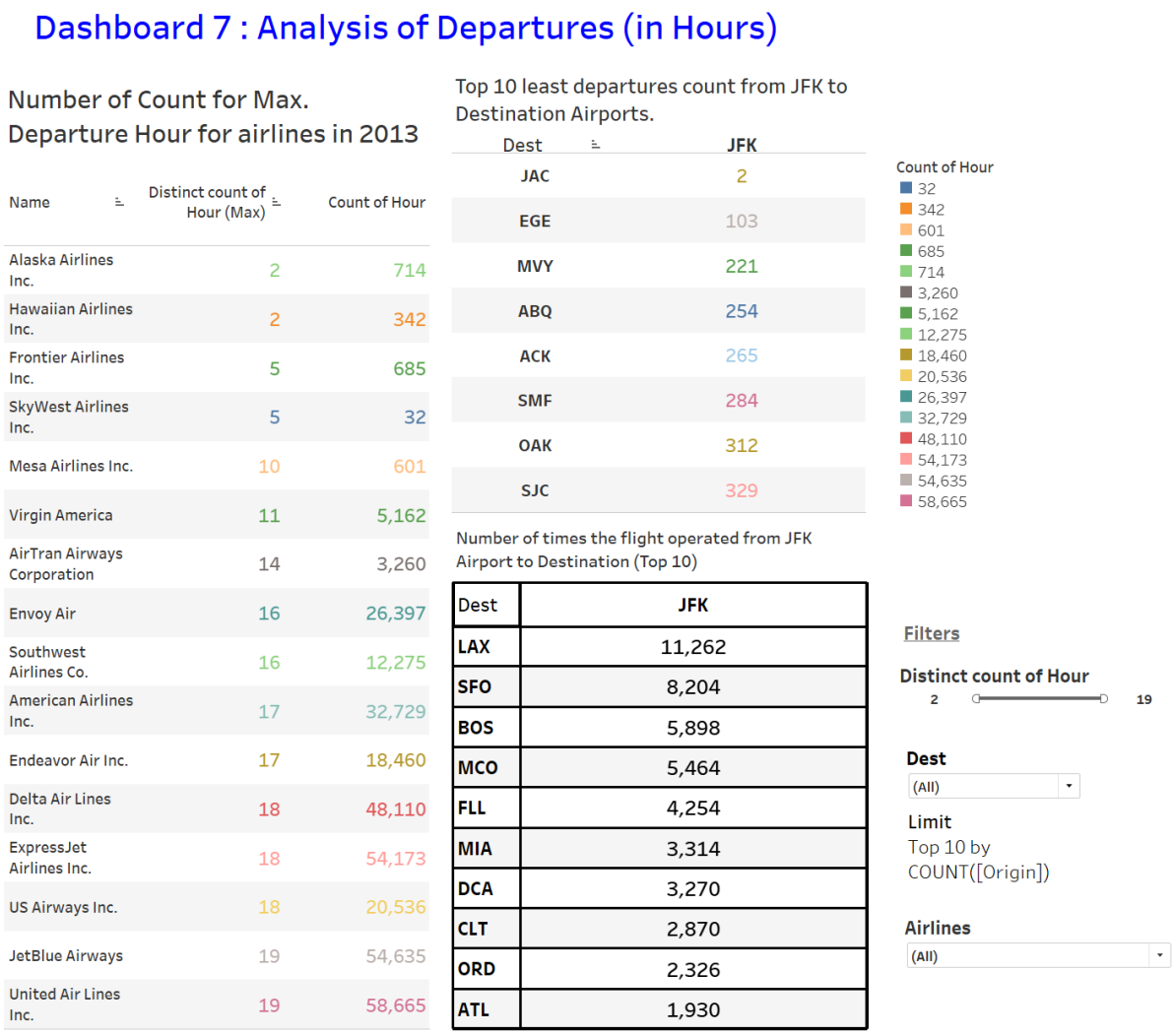


Fig 11 : Dashboard – Analysis of departures in hours

Insights from Exploratory Data Analysis

From the dashboard, various insights and questions emerge:

1. **Departure Hour Analysis:**
    - **Insight:** Airlines like United, JetBlue, and Delta have the highest distinct count of departure hours.
    - **Question:** How can these airlines optimize flight schedules to improve operational efficiency?
2. **Top Destinations from JFK:**
    - **Insight:** Los Angeles (LAX) and San Francisco (SFO) have the most flights from JFK, indicating high travel demand.
    - **Question:** How can airlines maximize revenue by optimizing capacity on these routes?

3. **Departure and Arrival Delays:**
    - **Insight:** Certain airlines consistently experience high departure and arrival delays, like ExpressJet and Envoy Air.
    - **Question:** What are the operational factors causing consistent delays for specific airlines?
4. **Least Departures from JFK:**
    - **Insight:** Destinations like Jackson Hole (JAC) and Eagle (EGE) have the fewest flights, suggesting limited demand.
    - **Question:** Should airlines continue servicing routes with low demand, or focus on more profitable routes?

Predictive and Prescriptive Analytics Approaches

1. **Regression Analysis:** Used to assess the impact of departure hours on flight delays.
    - **Target Variables:** Flight delays.
    - **Predictors:** Departure hour, airline, destination.
2. **Time-Series Analysis:** To analyse the demand for flights from JFK over time.
    - **Target Variables:** Flight demand.
    - **Predictors:** Month, year, destination.
3. **Classification Models:** To identify flights with high delay risk based on airline and destination.
    - **Target Variables:** Delay risk.
    - **Predictors:** Airline, destination, departure hour.

Decision-Making Benefits

- **Operational Efficiency:** Airlines can optimize flight schedules to improve punctuality.
- **Revenue Maximization:** Airlines can increase profitability by aligning capacity with demand.
- **Customer Satisfaction:** Airlines can improve customer satisfaction by minimizing delays and enhancing travel experience.

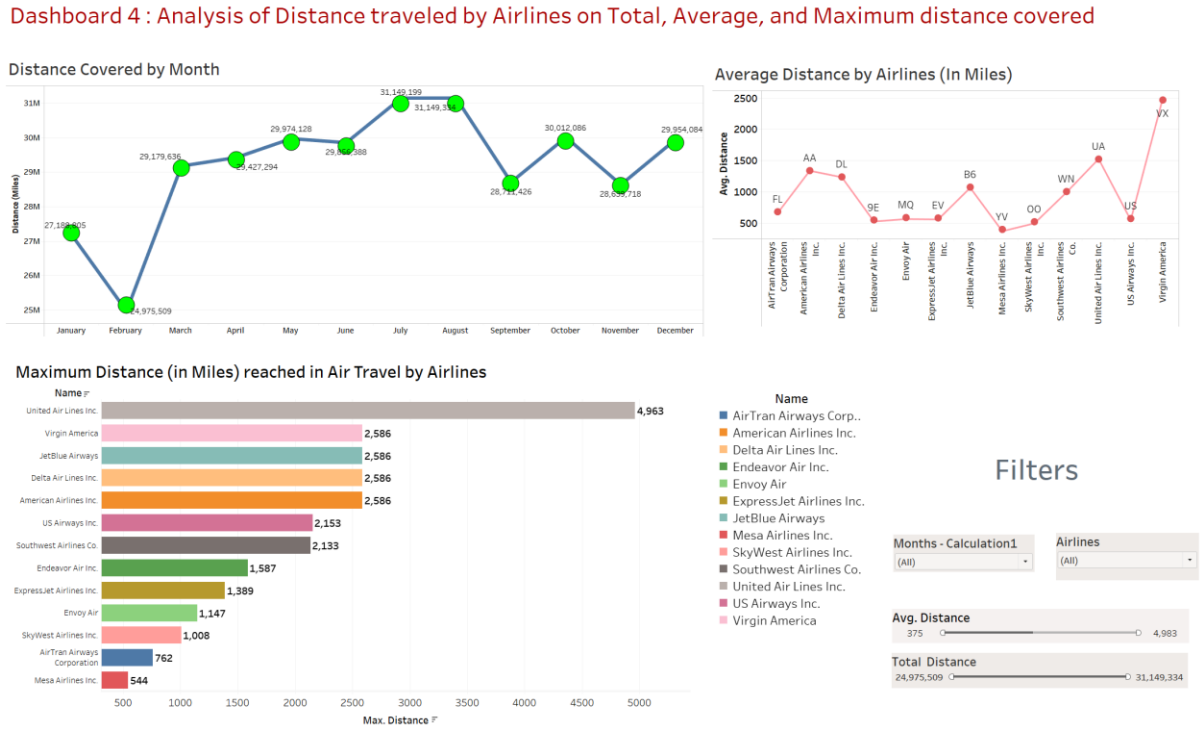


Fig 12 : Dashboard – Analysis of distance travelled by airlines on total, average and maximum distance covered.

Analysing the provided dashboard, we can identify key insights and questions to guide further analysis:

1. **Total Distance Covered by Airlines:**
    - **Insight:** United Airlines has the highest distance covered, suggesting its extensive network.
    - **Question:** How does network size affect delays and overall efficiency for United Airlines compared to other airlines?

2. **Seasonal Distance Trends:**
    - **Insight:** Travel peaks in the summer months of July and August, indicating high demand.
    - **Question:** How can airlines optimize their schedules to handle peak travel seasons efficiently?

3. **Average Distance by Airlines:**
    - **Insight:** Airlines like Virgin America focus on long-haul flights, resulting in high average distances.
    - **Question:** How do long-haul flights impact operational efficiency and customer satisfaction?

4. **Maximum Distance Reached:**
    - **Insight:** United Airlines reaches a maximum flight distance of nearly 5,000 miles.
    - **Question:** What challenges do long-haul flights pose, and how can airlines mitigate them?

Predictive and Prescriptive Analytics Approaches

1. **Regression Analysis:** To explore the impact of airline network size and flight distance on delays.
    - **Target Variables:** Flight delays.
    - **Predictors:** Airline network size, flight distance.
2. **Time-Series Analysis:** To analyse seasonal trends in flight distances and delays.
    - **Target Variables:** Monthly flight distances and delays.
    - **Predictors:** Time variables like month and year.
3. **Classification Models:** To identify flights at high risk of delays based on airline, distance, and season.
    - **Target Variables:** Delay risk levels.
    - **Predictors:** Airline, distance, season.

Audience: Airline operations managers, route planners, Airline planners and resource managers.

Purpose: Visualize airline-specific flight distances to optimize route planning, Assess the challenges of long-haul flights. Analyse the correlation between flight distances and delays. Highlight seasonal variations in flight distances.

Actionable Insights: Adjust flight schedules and network coverage, Optimize capacity and resources for peak seasons, Improve customer experience and optimize flight schedules, Improve route planning and scheduling.

- **Optimized Operations:** Airlines can better plan schedules and manage resources.
- **Improved Efficiency:** Airlines can identify inefficiencies in long-haul flights and address them.
- **Enhanced Customer Experience:** Airlines can improve customer satisfaction by minimizing delays.