**2024W -T2 BAM 2053 – DATA VISULAIZATION 01**

**Analyzing Flight delays and Cancellation in the US (January 2019 – August 2023)**

**Team Members**

* **Hashmeet Kaur: C0913169**
* **Deepinder Kaur: C0905697**
* **Mohammed Saaduddin: C0911305**
* **Omer Mohd: C0910940**

**Introduction**

Air travel is a vital kind of transportation that links individuals and companies worldwide. On the other hand, both passengers and airlines may experience severe inconvenience and interruptions because of aircraft delays and cancellations. It is essential to comprehend the trends and causes of these delays to increase customer satisfaction and operational effectiveness.   
  
We examine airline flight delay and cancellation data from January 2019 to August 2023 in this study, which comes from the Bureau of Transportation Statistics of the U.S. Department of Transportation. The dataset offers insightful information about flight paths, event time frames, and causes of delays and cancellations.

**Business Problem**

The aviation sector deals with several issues pertaining to flight cancellations and delays, which influence travelers, airlines, and the trip experience in general. The following are a few of the major business issues this analysis addresses:

* Operational Efficiency: To guarantee effective operations, save expenses, and increase revenue, airlines strive to maintain ideal flight schedules. Airlines can optimize their crew scheduling, flight planning, and resource allocation by having a better understanding of the causes that lead to delays and cancellations.
* Customer Satisfaction: Unfavorable experiences and passenger unhappiness can result from flight delays and cancellations. In order to keep consumers and retain a positive brand image, airlines work hard to deliver a dependable and enjoyable travel experience. Enhancing customer satisfaction can be achieved through identifying prevalent concerns and enhancing communication during disruptions.
* Financial Impact: Airlines bear heavy costs because of delays and cancellations, including passenger compensation, rebooking fees, and possible revenue loss. Airlines can find economical ways to lessen these financial effects by evaluating the data.
* Safety and Compliance: In the aviation sector, maintaining safe and legal operations is crucial. By comprehending the causes of delays and cancellations, airlines can immediately address safety problems and comply with regulatory obligations.

**Purpose and Objective**

This analysis's goal is to examine the large dataset of airline flight delay and cancellation data from January 2019 to August 2023 that the Bureau of Transportation Statistics of the U.S. Department of Transportation has made available. The following are the analysis's goals:

* Determine Trends: Examine trends and patterns in the number of canceled and delayed flights over the given time frame.
* Recognize Causes: Look at the primary causes and contributing elements of flight delays and cancellations.
* Evaluate the effects of cancellations and delays on travelers, airlines, and the industry.
* Forecast prediction: graphs for predicting and controlling future delays and cancellations are created by using historical data.
* Suggestions: Offer suggestions to airlines and industry participants regarding tactics to minimize delays, boost operational effectiveness, and elevate the traveler experience.

**Key Stakeholders**

* Airlines: Since they are immediately impacted by flight delays and cancellations, airlines are the main stakeholders. They can enhance customer service, optimize schedules, and improve maintenance procedures by comprehending the issues.
* Passengers: Travel plans, connections, and general pleasure are all negatively impacted by delays and cancellations for passengers. Better communication and compensation practices for affected tourists may result from the analysis's insights.
* Airports: Airport management is a critical function of airports. Airports can optimize runway usage, gate assignments, and ground handling procedures by comprehending delay trends.
* Government Organizations: To increase flight efficiency and safety, government organizations like the Federal Aviation Administration (FAA) use data to create rules and policies.
* Travel Industry: To better serve clients and manage expectations, travel companies, internet booking platforms, and associated businesses can gain from knowledge into flight delays.

**Project Strategy**

1. **Dataset and Collection**

Source: Bureau of Transportation Statistics, US Department of Transportation  
Duration: January 2019 through August 2023   
  
Flight paths (origin, destination) are variables.   
Event time ranges (minutes, local time)   
Reasons for delays and cancellations with limited attributions   
Details of the flight: airline, time of arrival, time of departure, delay of arrival and departure, taxi in and out, wheels on and off, canceled, diverted, airtime, distance, weather delays, National Airspace System delays, security delays, carrier delays, aircraft delays, on-time flights, and airport delays.

1. **Extract, Transform, Load, or ETL Process**
2. Extract

* Click the given link to download the dataset: Data on Airline Flight Cancellations and Delays.
* Compile monthly data subsets between January 2019 and August 2023.

1. Change

* To address errors, missing values, and inconsistencies, clean up the data.
* Convert data types (such as datetime conversion) if needed for precise analysis.
* Combine datasets by year to produce an all-inclusive dataset.

1. Fill up

* For analysis, save the cleaned dataset in an appropriate format (such as a CSV file or SQL database).

1. **Entity-Relationship Diagram (ERD) and Normalization**
2. **Entities:**

* Flights
* Airlines
* Airports
* Delay Reasons

1. **Relationships:**

* Flights have relationships with Airlines (via airline code), Airports (via origin and destination codes).
* Delay Reasons are associated with Flights (delay by weather, NAS, security, carrier, aircraft)

1. **Normalization:**

* Normalize the dataset to reduce redundancy and improve data integrity
* Break down into atomic tables like Flights, Airlines, Airports, Delay Reasons.

1. **Data Validation:**

* Check for Completeness: Ensure all required fields are present (e.g., flight number, origin, destination)
* Check for Accuracy: Verify data against known standards or external sources.
* Check for Consistency: Ensure data is consistent across different fields (e.g., arrival time after departure time)
* Check for Validity: Validate data types and ranges (e.g., dates, numeric values)
* Check for Integrity: Ensure relationships between entities are maintained (foreign key constraints)

1. **Key Performance Indicators (KPIs):**

* On-Time Performance: Percentage of flights that arrive or depart on time
* Average Delay Time: Average delay time for flights due to weather, NAS, security, carrier, aircraft
* Cancellation Rate: Percentage of flights that are cancelled.
* Diversion Rate: Percentage of flights that are diverted to another airport.
* Top Delay Reasons: Analysis of the most common reasons for flight delays
* Route Analysis: Busiest routes based on number of flights and delays.
* Airline Performance: Comparison of on-time performance and delays among different airlines
* Airport Performance: Analysis of delays and cancellations at different airports

**Analysis Questions**

1. What are the most common reasons for flight delays and cancellations?
2. How do flight delays vary by airlines and routes?
3. Are there specific times of the day or year when delays are more prevalent?
4. What trends can be observed in flight delays and cancellations from January 2019 to August 2023?
5. Are there any patterns indicating which airlines or routes are more prone to delays?
6. How do weather conditions impact flight delays and cancellations?
7. Can we identify any systematic issues contributing to delays that airlines can address?

By answering these queries, this analysis seeks to give stakeholders a thorough grasp of the variables affecting flight delays and cancellations, empowering them to decide wisely and take preventative action to improve the aviation industry's operational effectiveness and passenger experience.

**Methodology**

* Data Collection: The Bureau of Transportation Statistics of the U.S. Department of Transportation provided the dataset, which covered the period from January 2019 to August 2023.
* Data Preparation: Excel and Python were used to combine, clean, and modify the data. Variables were chosen for analysis, and missing values were dealt with.
* Exploratory Data Analysis (EDA): To comprehend the distribution of delays and cancellations, descriptive statistics, trend analysis, and visualizations were carried out.
* Route analysis is the process of determining which routes have the most delays and cancellations, both in terms of frequency and duration.
* Analyzing temporal trends, such as hourly, daily, weekly, and seasonal fluctuations in cancellations and delays, is known as temporal analysis.
* Causes of Disruptions: Examining the main causes of weather-related, mechanical, air traffic, etc. aircraft delays and cancellations.
* Impact assessment: Examining how delays and cancellations affect passengers and airlines financially and operationally.
* Recommendations: Offering practical suggestions for airports, airlines, and regulatory agencies to enhance operations and lessen disturbances in light of the results.

The purpose of this research is to give stakeholders in the aviation sector useful insights into the intricacies of flight delays and cancellations, along with data-driven suggestions to improve operational effectiveness and customer experience.

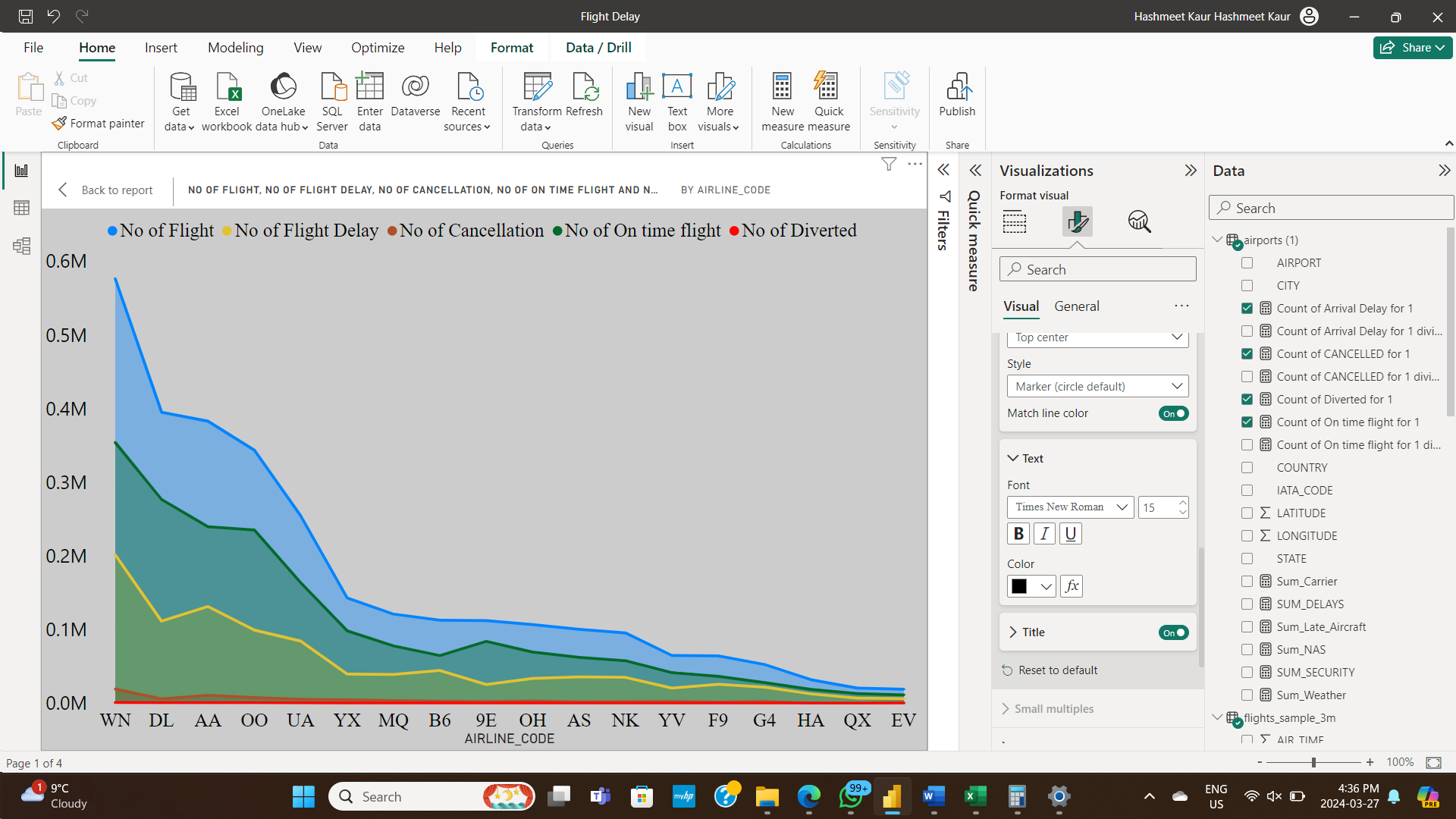
**Analysis and Data visualization**

* Airline Performance Comparison: Compare the performance of major airlines by looking at their cancellation and delay rates.

A screenshot of a computer

Description automatically generated

* Flight analysis by airlines through area chart.

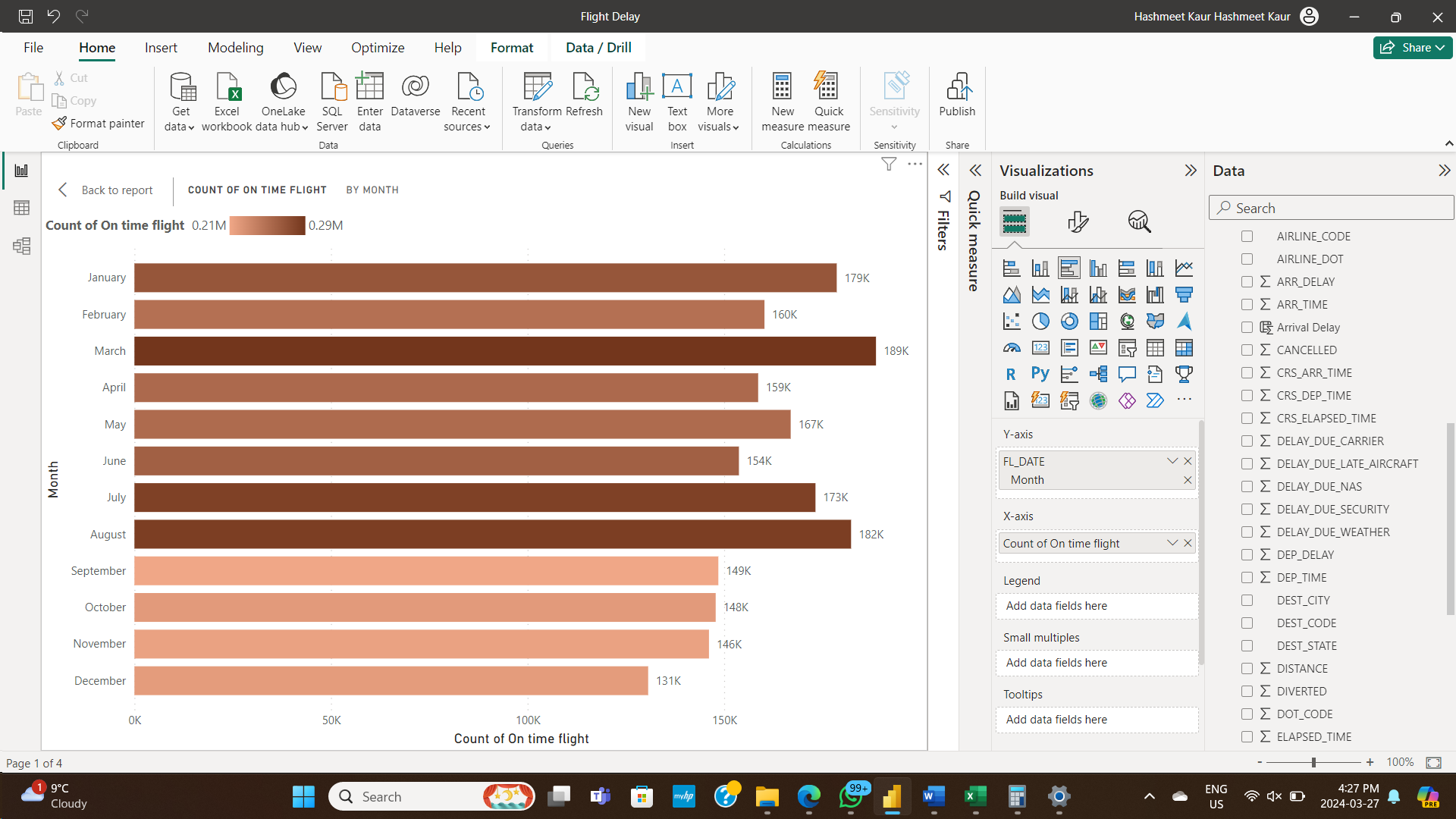


* Time Analysis: A line graph showing how cancellations and delays vary with the year.

A screenshot of a computer

Description automatically generated

* Seasonal Trends: A bar graph illustrating the on-time flight each season.



**Key Outcomes**

* Determining the main causes of flight cancellations and delays.
* Insights into the trends and patterns in the punctuality of different airlines and itineraries.
* Critical performance indicators, including seasonal changes, average delay durations, and cancellation rates, are visualised.
* Awareness of the effects of outside variables like the weather and traffic at airports.
* Suggestions for airlines on how to better communicate, schedule flights more efficiently, and handle delays.

**Potential Benefits to Stakeholders**

* Airlines have the potential to increase customer satisfaction, cut costs, and optimize operational efficiency.
* Travelers can plan ahead for possible interruptions, reduce annoyances, and make well-informed travel selections.
* Airport managers can arrange its facilities and resources to best suit passenger volumes and flight schedules.
* Regulators have the authority to impose rules and guidelines to enhance passenger safety and on-time performance.

**Recommendations**

* Take preventative action to deal with common flight delay factors, such as crew availability or maintenance problems.
* Increase passenger communication amid disruptions by giving them access to real-time information and alternate routes.
* Optimize flight scheduling by considering peak travel periods and historical data.
* Improved cooperation between airports and air traffic control can help reduce traffic jams and delays.
* Invest in cutting-edge weather forecasting equipment to lessen the effects of unfavorable weather.

**Conclusion**

Analyzing flight delay and cancellation data from airlines offers important insights into the difficulties the aviation sector faces. Airlines may boost customer happiness, optimize operations, and improve on-time performance by implementing tactics based on an understanding of the underlying causes of disruptions and trend identification. The purpose of this report is to provide the aviation industry with a solid basis for informed decision-making and ongoing improvement.

**References**

* U.S. Department of Transportation, Bureau of Transportation Statistics: <https://www.transtats.bts.gov>
* DOT's monthly Air Travel Consumer Report: <https://www.transportation.gov/airconsumer>
* Dataset Source: https://www.transtats.bts.gov