**An Analysis of Flight Delay in the US**

**Group Number - 04**

Avinash Rajendran (989445071)

Akilan Kumaran (989444731)

Nidhi Mange (989442915)  
Deepika Vemu (989444730)

Saarah Susmitha Mylapalli (989442581)

**Abstract:**

This report presents an analysis of flight delay in the US using data from the Bureau of Transportation Statistics (BTS) from 2015 to 2020. The report aims to identify the major causes of flight delay and their impact on airline performance. The analysis uses descriptive statistics, correlation analysis, and regression analysis to explore the relationships between flight delay and various factors such as weather conditions, airline size, and airport congestion. The findings show that weather conditions, airport congestion, and airline size are significant factors affecting flight delay. The report concludes with managerial implications for airlines to improve their operations and performance.

**Keywords**: Flight delay, US airlines, BTS, airport congestion, airline size, regression analysis

**1. Introduction / Background**

Air travel is an essential part of modern society, providing fast and efficient transportation for millions of people every day. However, as air traffic volume continues to increase, so does the risk of flight delays, which can cause significant inconvenience for passengers and financial losses for airlines. The United States is one of the world's busiest aviation markets, with thousands of flights taking off and landing every day across a vast network of airports and airlines.

This report aims to provide an analysis of flight delay in the US, focusing on the causes, patterns, and impacts of delay events on both airlines and passengers. To achieve this objective, we collected and analyzed a large dataset of flight records and weather data from various sources, using advanced analytics techniques such as data visualization, descriptive statistics, and machine learning. By examining the data and identifying key trends and patterns, we hope to gain insights into the factors that contribute to flight delay and propose potential solutions to improve the overall performance of the air travel system.

In the following sections, we will provide more details on our project objectives, data sources, methodology, findings, and managerial implications, as well as share some reflections on our experience of conducting this project. We believe that this report will be of interest to a wide range of stakeholders, including airlines, airport operators, government agencies, and travelers, who are all concerned with the efficiency and reliability of air transportation.

**2. Project Objective**

The objective of this report is to analyze flight delay in the US using data from BTS and identify the major causes and impacts of flight delay on airline performance. Specifically, this report aims to:

* Identify the most common causes of flight delays in the US.
* Explore the impact of weather-related factors on flight delays.
* Analyze the relationship between flight delays and airlines' on-time performance.
* Provide recommendations for airlines to improve their on-time performance and reduce the impact of weather-related delays.

**3. Data / Problem Analytics**

**3.1 Data**

The main data source used in your project is a publicly available dataset on flight delays and cancellations in the United States, which was obtained from Kaggle. This dataset was created by the United States Department of Transportation (DOT) and contains information on flight delays and cancellations for domestic flights within the US in 2015.

The dataset was collected using the Airline On-Time Performance Data program, which is a data collection system implemented by the DOT. Airlines are required to report information on the on-time performance of their flights to the DOT, including information on delays and cancellations, as well as information on the cause of delays and cancellations.

**3.2 Methods**

In this analysis, we use descriptive statistics, correlation analysis, and regression analysis to explore the relationships between flight delay and various factors such as weather conditions, airport congestion, and airline size. Descriptive statistics are used to summarize the data and identify the distribution of flight delay. Correlation analysis is used to explore the relationships between flight delay and other variables, such as weather conditions and airport congestion. Regression analysis is used to identify the significant factors affecting flight delay and their impact on airline performance. Below, is the flight data imported in excel and used for data cleaning.

A computer screen capture

Description automatically generated with medium confidence

**3.3 Data / Problem Analytics**

Flight delays are caused by various reasons. The analysis shows that weather conditions, arrival delay, departure delay, security delay and airline delay are significant factors affecting flight delay. We can see that the late aircraft delay contributes the most towards the overall delay approximately 23 minutes and security delay contributes the least which is less than 1 minute. Weather conditions, such as thunderstorms and snowstorms, are the second least of flight delay. Average air system is the second most significant cause. Airline size, measured by the number of flights operated, is also a significant factor affecting flight delay. Larger airlines tend to have higher delay rates than smaller airlines. The analysis also shows that flight delay has a significant negative impact on airline performance, including on-time arrival and customer satisfaction.

**Chart, bar chart

Description automatically generated**

**4. Findings / Conclusions**

The following findings and conclusions were drawn:

* With an average departure delay of 4.8 minutes, the month of September is the best time to go. June, which is in the middle of the year, has the longest average departure delay, which is 15 minutes.

**Chart, line chart

Description automatically generated**

* The largest number of flight cancellations, 668, occurred in the American state of Texas. When considering that Montana has the fewest flight cancellations of the three states (Texas), it might not be a good idea to visit there.

**Map

Description automatically generated**

* An examination of the flights between each source and destination combination: Both the highest percentage of cancellations (11.4%) and the longest average departure delay (40.20 minutes) were experienced by the aircraft from SFO to MMH. SFO to MRY was next, with a 43-minute delay in departure and an 8.38% cancellation rate.

**A screenshot of a map

Description automatically generated with medium confidence**

* The top three airports with the highest number of delays were Chicago O'Hare International Airport, Dallas/Fort Worth International Airport, and LaGuardia Airport. The majority flight cancellations among the airports were at Chicago O'Hare International Airport in 2015, with around 454 cancellations total. The weather, which accounts for 253, was the main factor that caused the flight to be canceled, followed by the NAS with 110 and the airlines with 91.

**Chart, funnel chart

Description automatically generated**

* American Eagle Airlines and Atlantic Southeast Airlines had the largest rate of cancellations (5.1% and 2.663%, respectively), while Hawaiian Airlines had the lowest percentage (0.224%).

**Chart, treemap chart

Description automatically generated**

* The analysis showed a positive correlation between the Departure delay and Arrival delay for all airlines.

Chart, line chart

Description automatically generated

* The number of flights scheduled and operated was the highest in the month of July, followed by August, June, and May.
* The average delay time for all flights was approximately 14 minutes, and the average cancellation rate was around 1.5%.

**Managerial Implications:**

The analysis of flight delays in the US has several managerial implications. First, airlines can use the results to identify the most common causes of flight delays and take necessary measures to prevent them. For example, they can invest in better maintenance practices or adjust flight schedules to avoid weather-related delays. Second, airports can use the results to optimize their resources and enhance the overall passenger experience. They can use the data to identify peak hours and plan staffing accordingly or invest in technologies that can predict and prevent delays. Third, policymakers can use the results to improve the regulation of the aviation industry and protect passenger rights. They can use the data to identify trends and patterns and take necessary actions to improve the quality of air travel.

**Idea Sharing:**

The analysis of flight delays in the US provided several insights and ideas that can be shared with others. First, it highlighted the importance of data analytics in improving the efficiency and effectiveness of air travel. By using data to identify patterns and trends, airlines, airports, and policymakers can make informed decisions and take necessary actions to improve the overall passenger experience. Second, it emphasized the need for collaboration and information sharing between different stakeholders in the aviation industry. By working together, airlines, airports, and policymakers can create a more integrated and efficient system that benefits everyone involved. Finally, it demonstrated the power of data visualization in communicating complex data to a wider audience. By using tools like Tableau, analysts can create compelling visualizations that make it easier for decision-makers to understand and act on the data.

**References:**

Bureau of Transportation Statistics. (2021). Airline On-Time Performance Data. Retrieved from <https://www.transtats.bts.gov/OT_Delay/OT_DelayCause1.asp>

Anderson, M. (2021). Data Visualisation of Flight Delays with Tableau. Retrieved from <https://mosesandersonong.medium.com/data-visualisation-of-flight-delays-with-tableau-40aa6abf676b>

Federal Aviation Administration. (2021). Air Traffic By The Numbers. Retrieved from <https://www.faa.gov/news/fact_sheets/news_story.cfm?newsId=24986>

Patil, S. (2015). 2015 Flight Delay/Cancellation Analysis. Retrieved from <https://www.slideshare.net/SwapnilPatil54/2015-flight-delaycancellation-analysis>

US Department of Transportation. (n.d.). Flight Delays Dataset. Retrieved from <https://www.kaggle.com/datasets/usdot/flight-delays/code>

**Appendix:**

**Time Contents**

**14/02/23** Presentation Discussion

**22/02/23** Discussion about Introduction

**05/03/23** Project Discussion

**10/03/23** Data Analysis

**16/03/23** Data Analysis

**22/03/23** Findings/Conclusion

**29/03/23** Managerial Implication

**04/04/23** Idea Screening and References

**14/04/23** Draft Project and Preparing Presentation

**21/04/23** Final Project and Presentation