US 2023 Civil flights, delays, meteo & aircrafts

CIS9650 Programming for Analytics
Prof. Jason Tseng

Zicklin School of Business - Baruch College

Moksha Shah, Nikita Gautam, Tashfiqul Haq, Yashasvi Bhati

May 11, 2024

INTRODUCTION

Source of Data: Kaggle [Dataset Link](#)
Project Goal: To analyze US flights in 2023, with a focus on canceled flights, using visualization and detailed data analysis to reveal valuable insights.

Key Findings:
- Boeing is the industry's dominant manufacturer with a 46% market share.
- Southwest Airlines Co. is the largest airline in the United States in terms of the number of flights operated.
- June and July had the highest flight cancellation rates, indicating a correlation with the volume of scheduled flights.
- The most canceled route was BOS To LGA
- 89.7% of canceled flights were small-haul
- Afternoons experienced the highest flight cancellation rate at 31.1%

Why we picked this Data set:
- Our team of 4 was driven by shared experiences of flight disruptions, as everyone has encountered canceled or delayed flights at least once in their lives.

Information about your cleaned data (who can use this data?)
- The cleaned dataset, focused on US flights in 2023, particularly on canceled flights, is valuable for aviation analysts, airlines, researchers, and policymakers aiming to understand and address flight cancellation patterns.

Challenges Faced:
- Understanding flight cancellation and delay patterns.
- Dealing with primarily categorical data.

Team Tasks:
- Data Collection and Cleaning: Yashasvi Bhati, Nikita Gautam
- Analysis, Coding, Visualization: Moksha, Nikita Gautam, Yashasvi Bhati, Tashfiqul Haq
- Report and Presentation: Moksha, Nikita Gautam, Yashasvi Bhati, Tashfiqul Haq

Topics Researched:
- Reviewed existing research on flight delays and cancellations to understand common patterns and trends.

Expansion Possibilities:
- Investigate the specific reasons for flight delays, such as weather, air traffic control issues, mechanical problems, or crew scheduling.

Database:
- Saved the raw CSV data to SQLite database to make the lookup easier for the further analysis