INFSCI 2310 Final Report

Network Analysis of Flights Between US Cities

Number of flights per year

1 << 100

100 << 1000

1000 << 10000

10000 << 100000

100000 << 100000

Figure 1: Flight Network Visualization

Legend Explained:

- Yellow dots: Airports with less than 100 flights per year.
- Red dots: Airports with 100 to 1,000 flights per year.
- Blue dots: Airports with 1,000 to 10,000 flights per year.
- Purple dots: Airports with 10,000 to 100,000 flights per year.
- Green dots: Airports with 100,000 to 1,000,000 flights per year.

Findings Text Introducing Highlights of the Produced Figure in Bulletin Points:

- The majority of flight activity is concentrated in specific regions, suggesting major travel hubs.
- ❖ There is a clear trend of increased flights in coastal areas, indicating higher travel demand.
- ❖ A significant number of airports handle a very high volume of traffic, emphasizing the need for efficient air traffic control in these areas.

Data Gathered from:

 $\underline{https://www.kaggle.com/datasets/vikalpdongre/us-flights-data-2008}, provided through Kaggle.$

Southwest Airlines Co. on time (t < 5 min) United Air Lines Inc. small delay (5 < t < 45 min) Frontier Flying Service large delay (t > 45 min) Spirit Airlines US Airways Inc. Delta Air Lines Inc. PSA Airlines American Airlines Inc. letSuiteX Contour Airlines Frontier Airlines Inc. JetBlue Airways Endeavor Air Atlantic Southeast Airlines Mesa Airlines American Eagle Airlines Inc. Allegiant Air Alaska Airlines Inc. Skywest Airlines Inc. Hawaiian Airlines Inc. 100000 200000 300000 400000 500000 600000 700000 800000 Flight count

Figure3: Flight Count

Legend Explained:

- Blue bar: Flights on time ($t \le 5$ minutes)
- Orange bar: Flights with a small delay (5 minutes < t < 45 minutes)
- Gray bar: Flights with large delay (t > 45 minutes)

Findings Text Introducing Highlights of the Produced Figure in Bulletin Points:

- Southwest Airlines Co. has the highest number of on-time flights, which reflects their operational efficiency.
- ❖ While United Air Lines Inc. and Frontier Airlines Inc. show a significant number of flights, they also have a notable proportion of small and large delays.
- ❖ Smaller airlines like Allegiant Air and Hawaiian Airlines Inc. have a relatively high on-time performance, suggesting efficient punctuality management despite fewer flights.
- ❖ The overall flight count indicates that while larger airlines handle a greater volume of traffic, they also tend to have a higher occurrence of delays.

Data Gathered from:

https://www.kaggle.com/datasets/vikalpdongre/us-flights-data-2008

Figure 3: Origin Airport Distribution

Origin Airport Distribution

14.9% 17.4% 12% 5.89% 6.89% 10.3% 7.35% 9.02% 8.47%

Legend Explained:

- Blue: Hartsfield-Jackson Atlanta International Airport 17.4%
- Salmon: Chicago O'Hare International Airport 14.9%
- Green: Dallas/Fort Worth International Airport 12%
- Light Green: Denver International Airport 10.3%
- Yellow: Los Angeles International Airport 9.02%
- Orange: Phoenix Sky Harbor International Airport 8.47%
- Grey: George Bush Intercontinental Airport 7.35%
- Grey Blue: McCarran International Airport 6.89%
- Leaf Green: Detroit Metropolitan Airport 5.89%
- Bright Blue: Newark Liberty International Airport 4.7%

Findings Text Introducing Highlights of the Produced Figure in Bulletin Points:

- Collectively, the top three airports constitute nearly 45% of the origin flights, indicating a high concentration of traffic at these hubs.
- ❖ The distribution shows that a small number of airports dominate as starting points for US flights, which may impact national air traffic patterns and resource allocation.

Data and Method Text Describing the Data and Method Used in This Process:

- ➤ Analytical Tools: Utilized Python for data cleaning, manipulation, and visualization, employing libraries such as Pandas for dataset operations, Matplotlib for creating charts, and geospatial libraries for mapping.
- Methodology: Statistical methods were used to analyze flight distributions and punctuality across US airports and airlines. The visualizations are designed to clearly depict the scale and traits of the flight data.

Significance Statement on Why the Presented Figures are Important:

- ★ The geospatial distribution map underscores the spatial dynamics of flight frequencies, highlighting infrastructural and logistical hotspots.
- ★ The pie chart of origin airport distribution offers a clear view of the central hubs of US air traffic, which is pivotal for strategic decisions in capacity planning and route management.
- ★ The airline punctuality bar chart reflects the reliability of different carriers, essential for consumer choice and airline competitive strategy.

GitHub Link: https://github.com/QuinceyNiu/InfoViz2023Fall