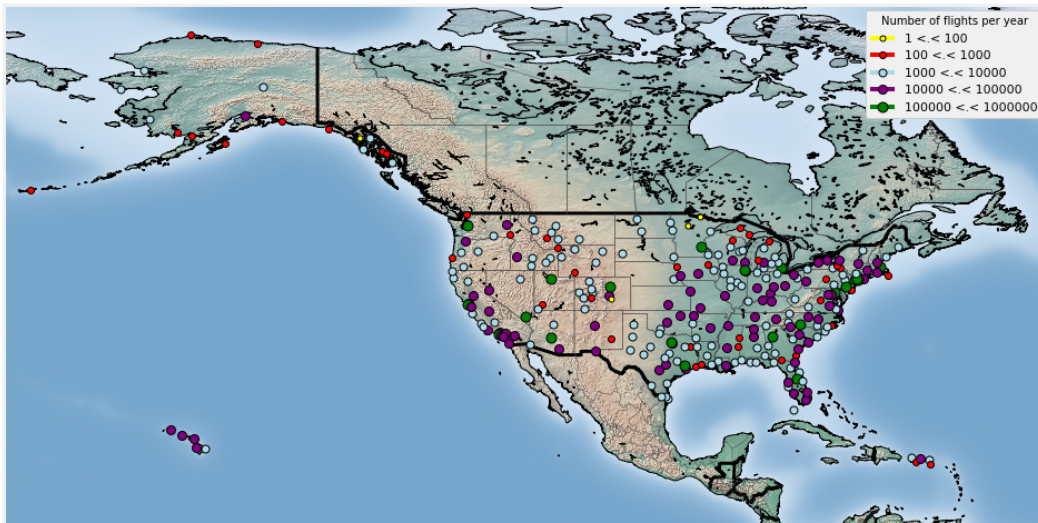


# INFSCI 2310 Final Report

## Network Analysis of Flights Between US Cities

Figure1: 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.

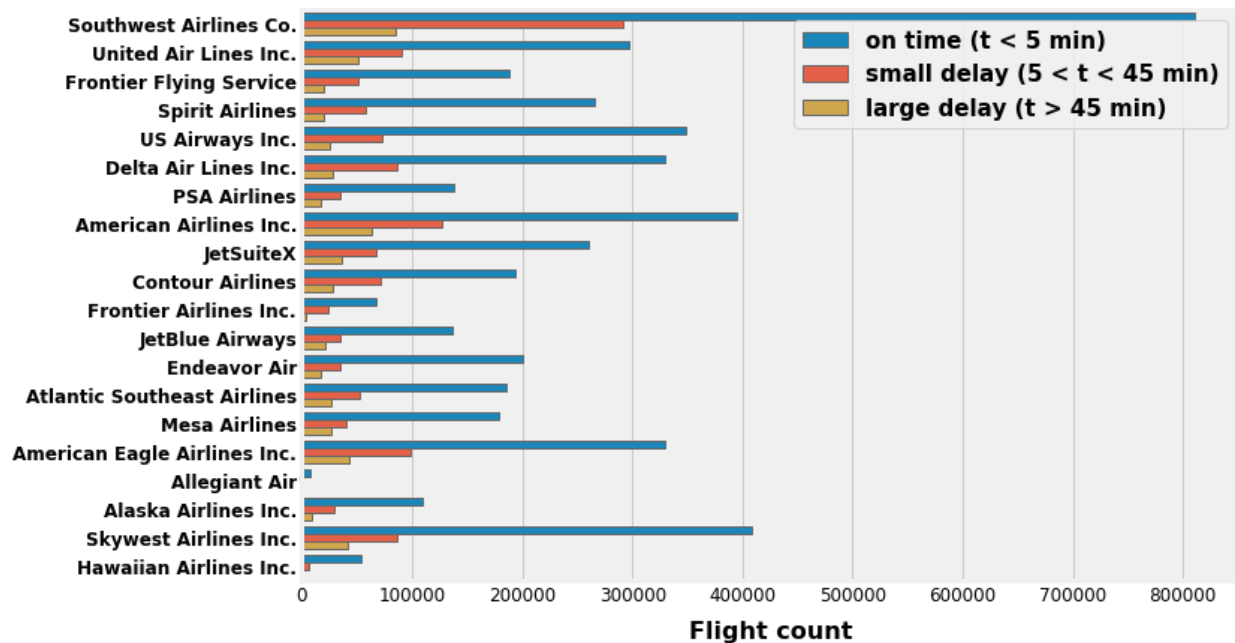
### Findings Text Introducing Highlights of the Produced Figure in Bulletin Points:

- ❖ 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:

<https://www.kaggle.com/datasets/vikalpdongre/us-flights-data-2008>, provided through Kaggle.

Figure3: Flight Count



### Legend Explained:

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

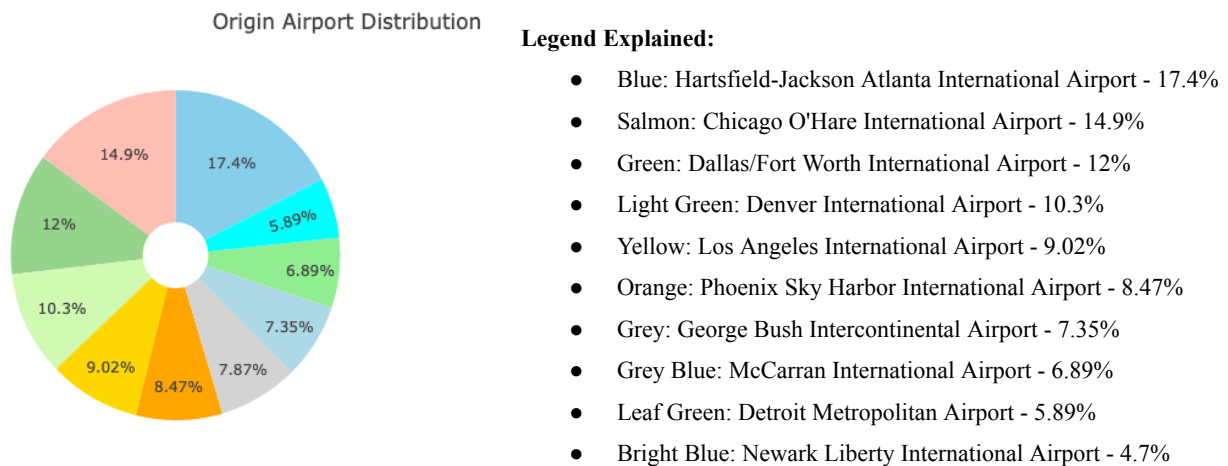
### Findings Text Introducing Highlights of the Produced Figure in Bulletin Points:

- ❖ 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>

Figure3: Origin Airport Distribution



### 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>