

Commercial Flight Delay Analysis

ECE 180 Python for Data Analysis

Group 4

Will Chen

Sam Vineyard

Jinhao Tang

Zexi Chen

Siyi Wang





Background and Motivation

- **Purpose:**
 - Flight delay is frustrating
 - Help commercial flight passengers in the United States (US) choose flights with minimal delays
 - Provide delay statistics so that people can plan things ahead.



Dataset and Data Reduction Methods

- **Source of Data:**
 - Department of Transportation (DOT)
- **Type of Data:**
 - CSV for each month
- **Data Viability:**
 - Reliable
 - Large quantity (5,000,000 flight entries)
 - Features with variability (location, time, airlines, etc)
 - Diverse Features
- **Data Limitations:**
 - 1.2% missing value
 - Only used data from 2016



Dataset and Data Reduction Methods

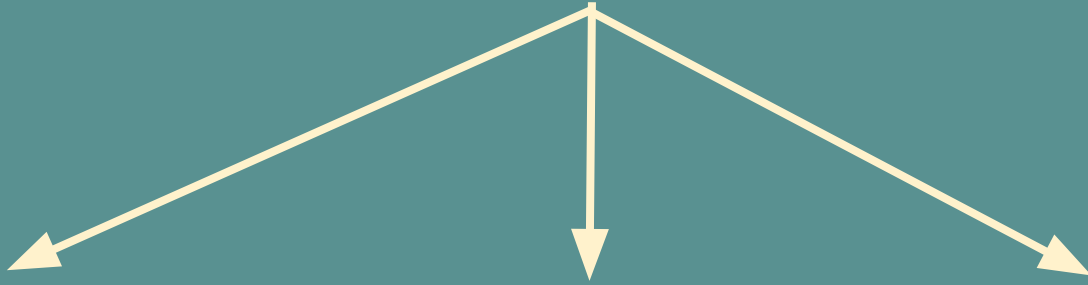
- **Data Relevancy:**
 - Airline
 - Time
 - Airport
 - Delay Duration
 - Delay Cause
 - Distance
- **Data Cleaning:**
 - Remove duplicate features (eg. state code and state name)
 - Remove useless features (eg. wheel on time)
 - Remove entries with missing features

Delay Rates

of Delayed Flights / # of Total Flights

Delayed Flight = 15 min delay

Delay Rates



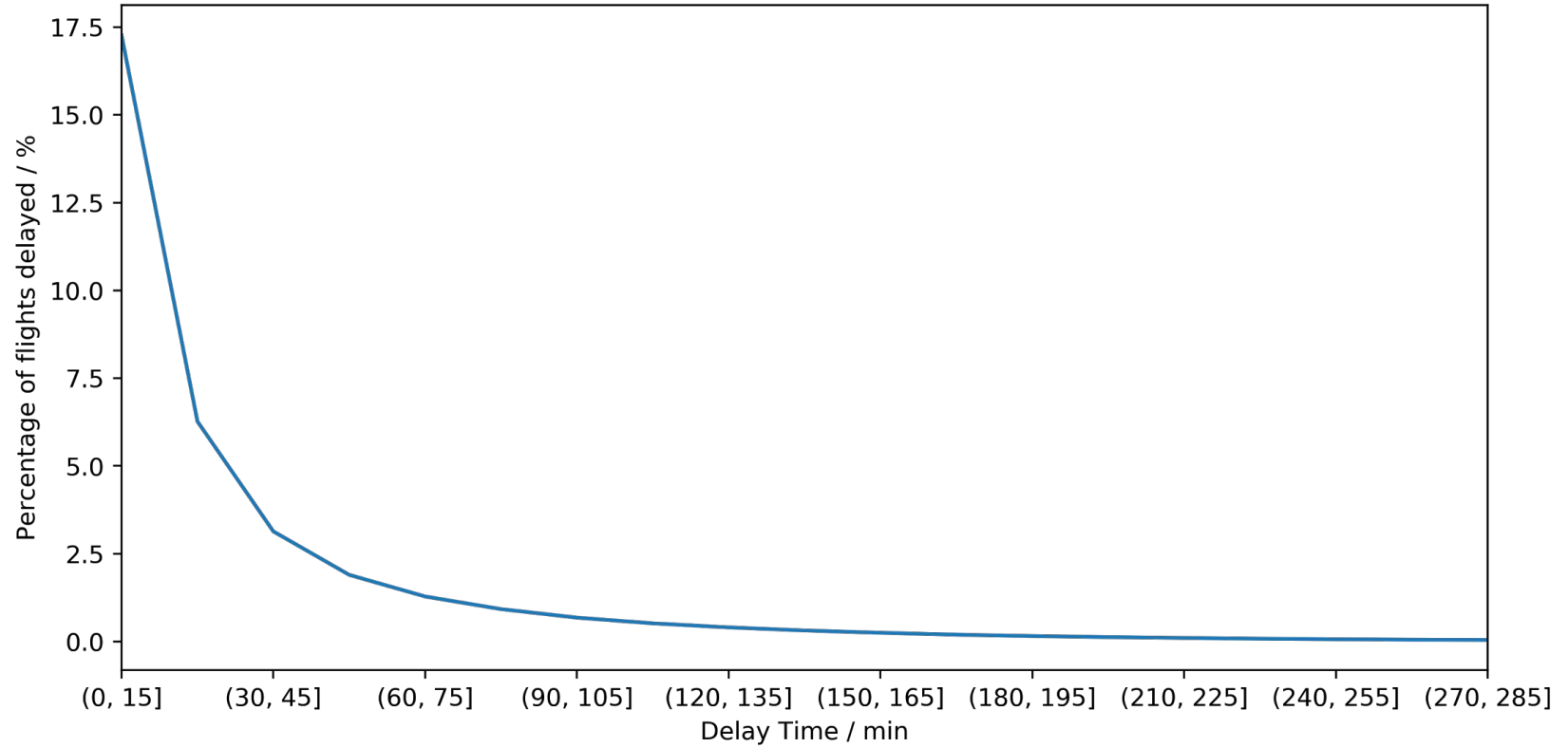
Airline

Time

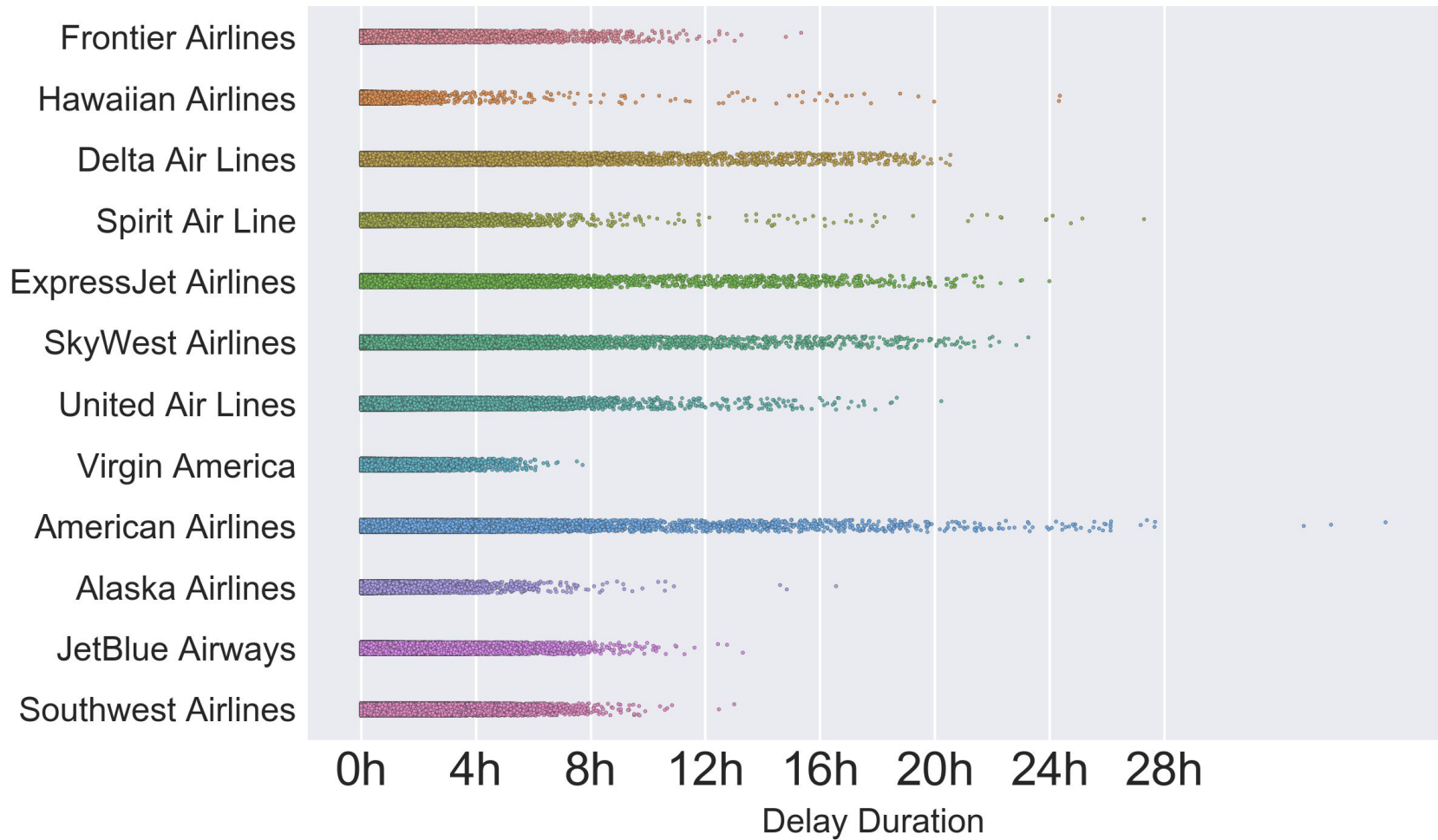
Location

Delay Duration Distribution

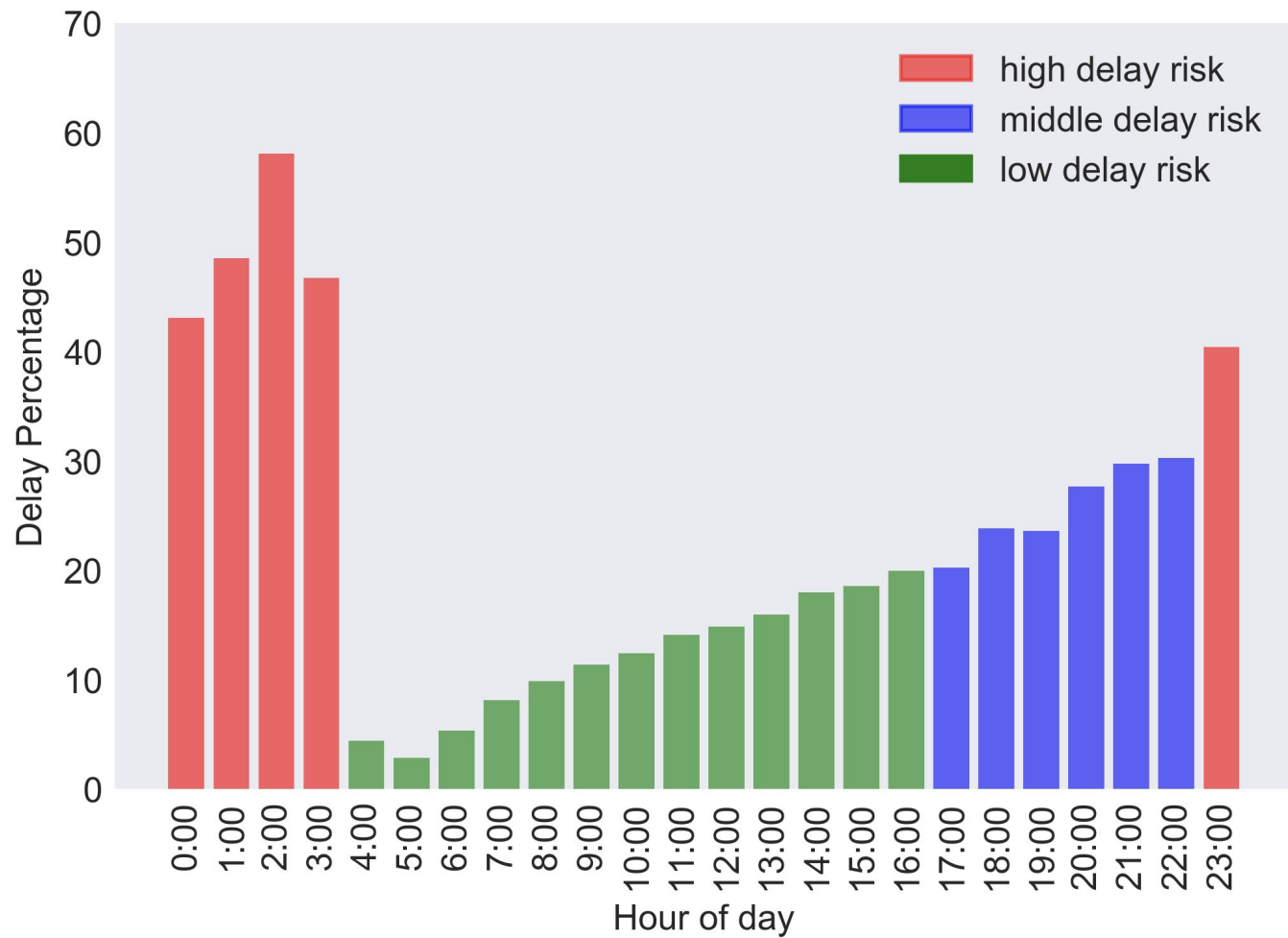
Duration of delays



Delay Durations vs Major Airline

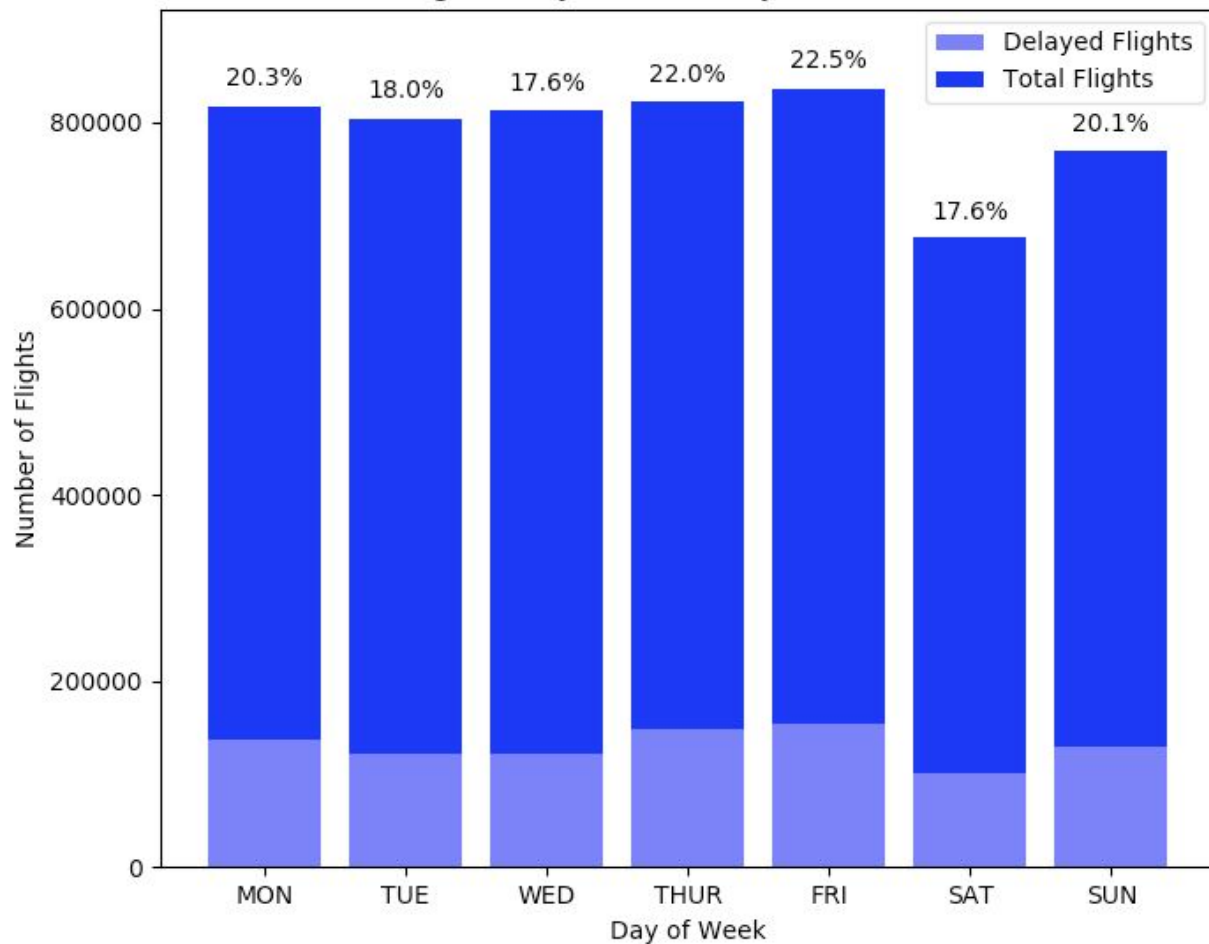


Delay Rate vs Hour of Day

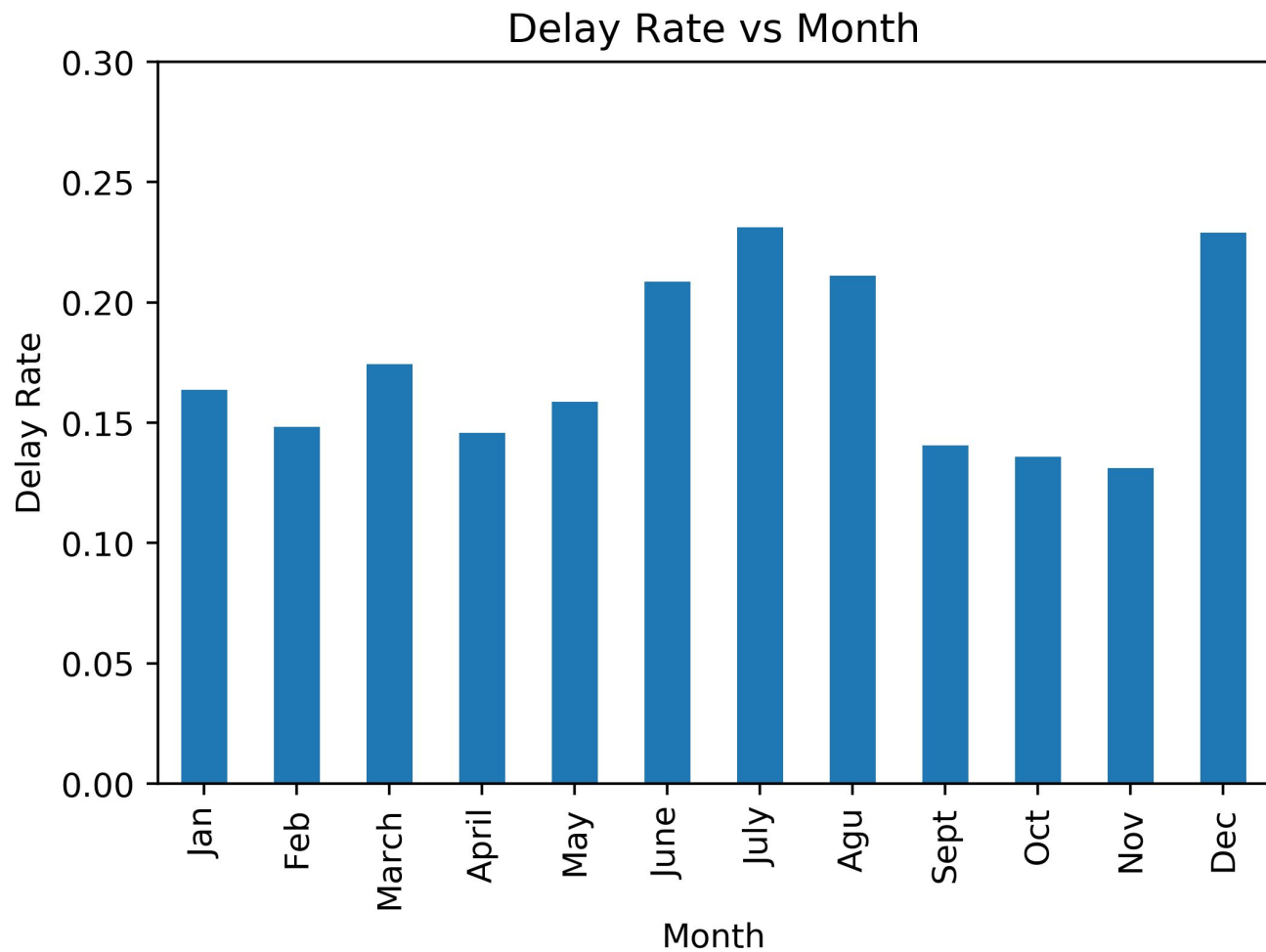


Delay Rate vs Day of Week

US Flight Delay Rates vs Day of Week in 2016

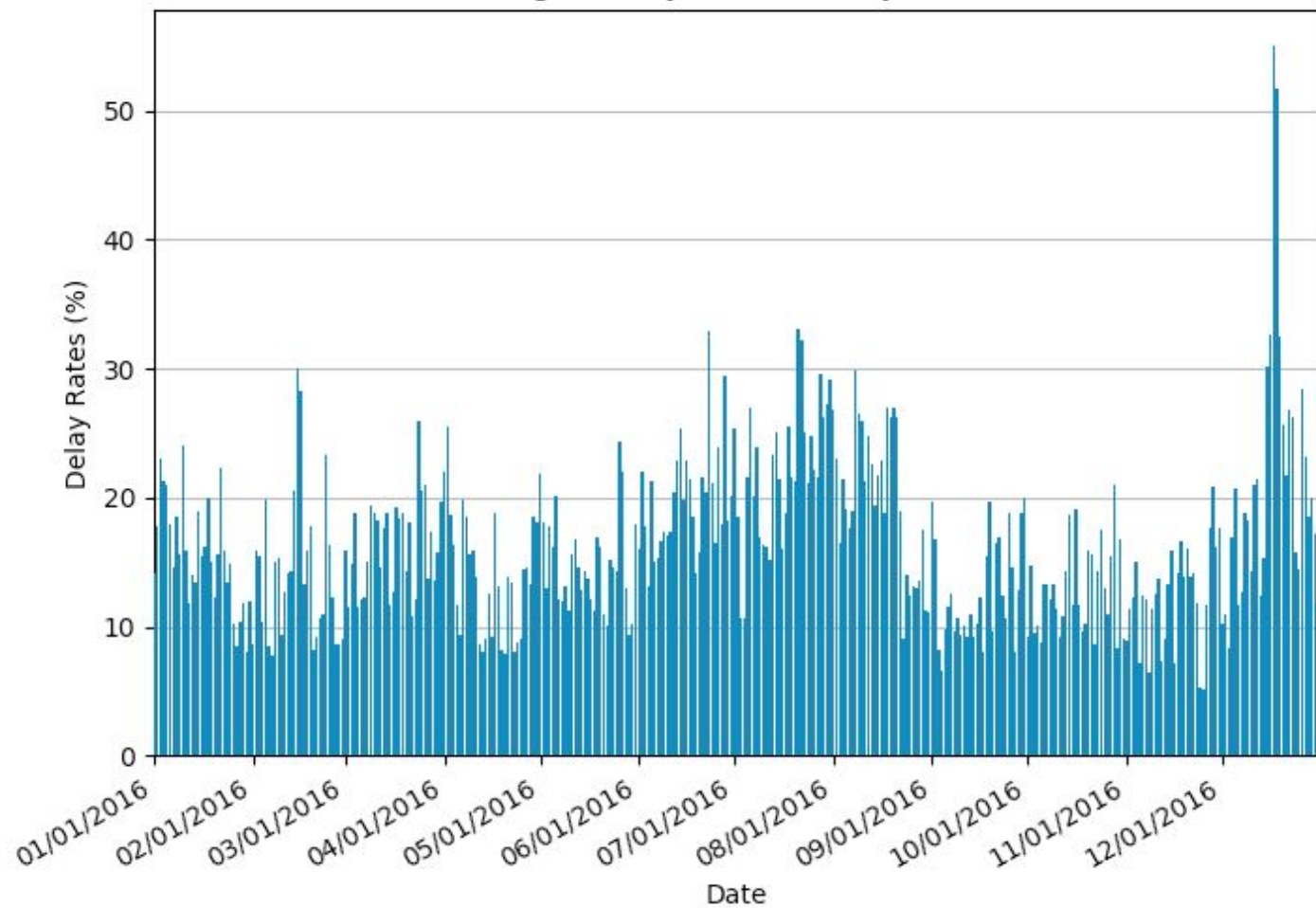


Delay Rate vs Month



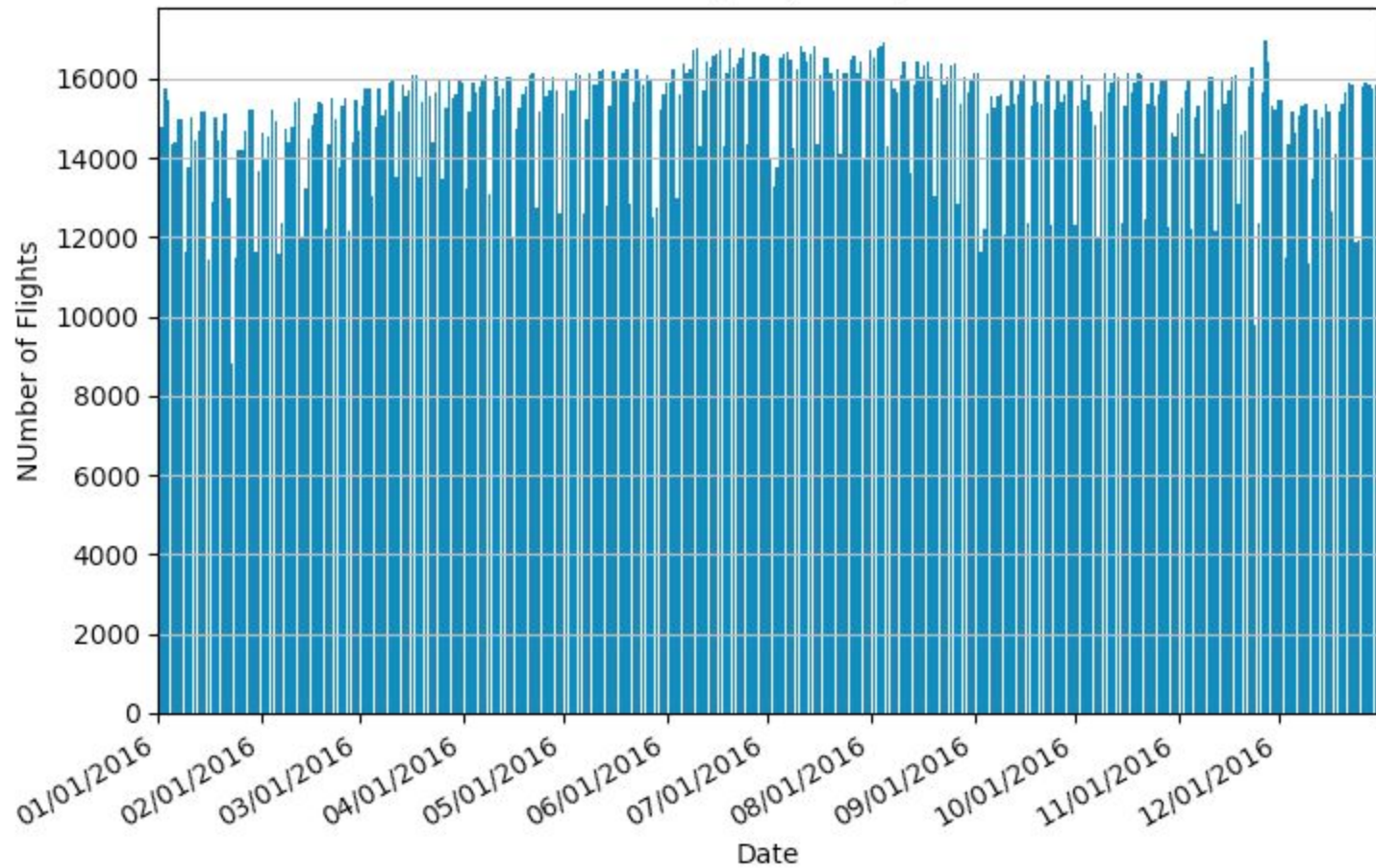
Delay Rate vs Day of Year

US Flight Delay Rates vs Day of 2016



Number of Flights vs Day of Year

Number of Flights per Day in 2016

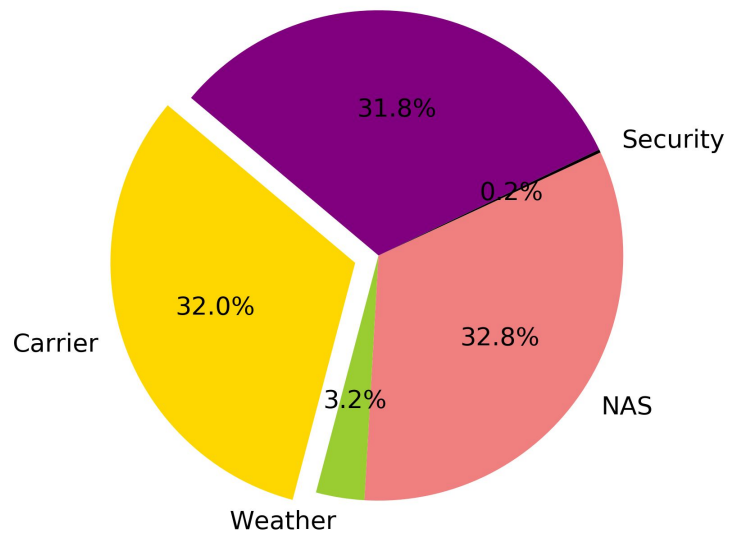


Delay Rate and Duration vs Causes

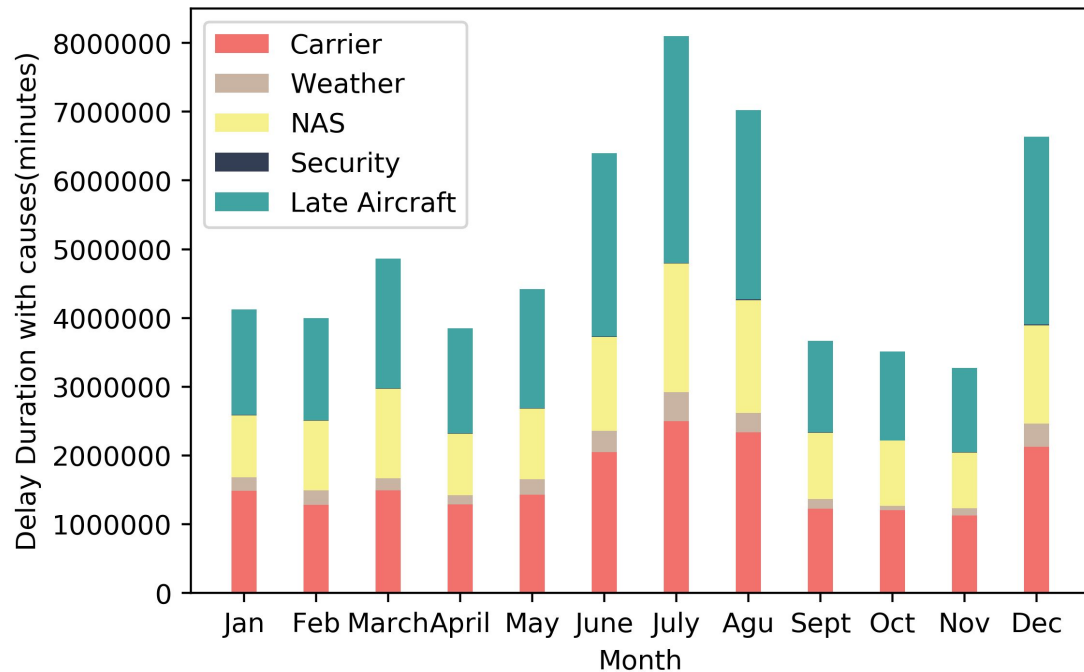


Delay Rate vs Causes

Late aircraft



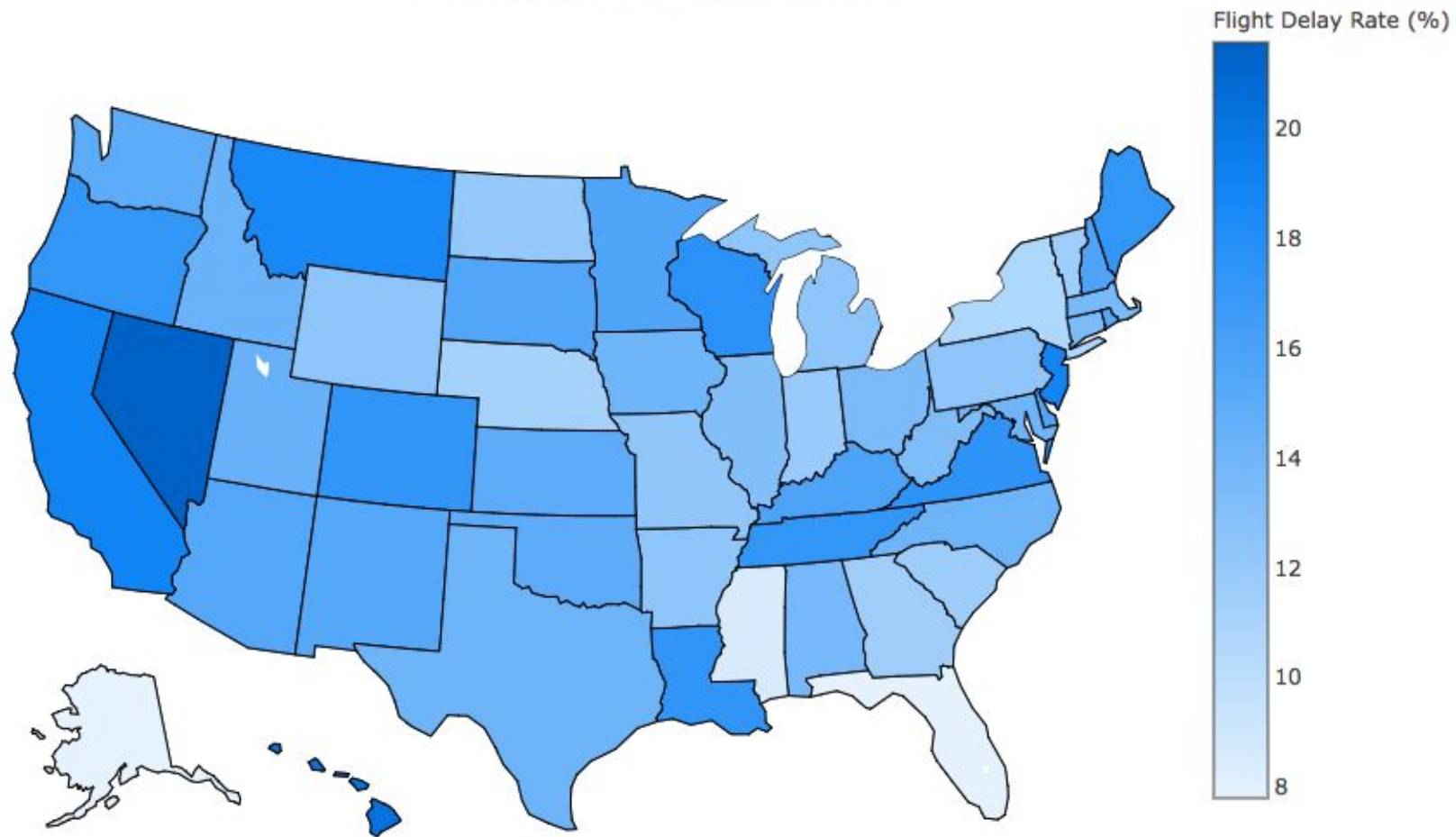
Delay Duration with causes vs Month



Delay Rate vs State

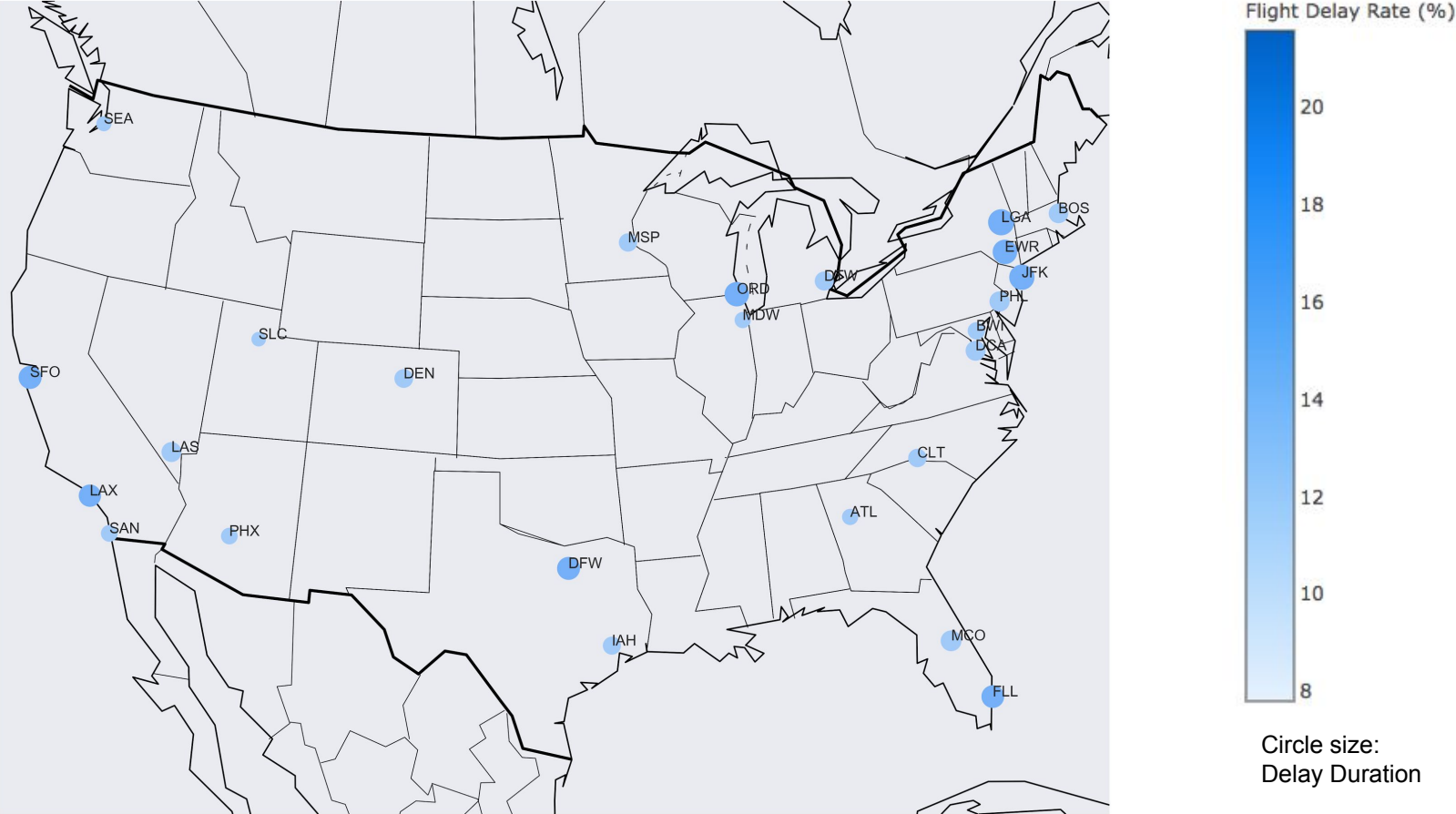


2016 Flight Delay Rates vs State



Delay Rate vs Origin Airport

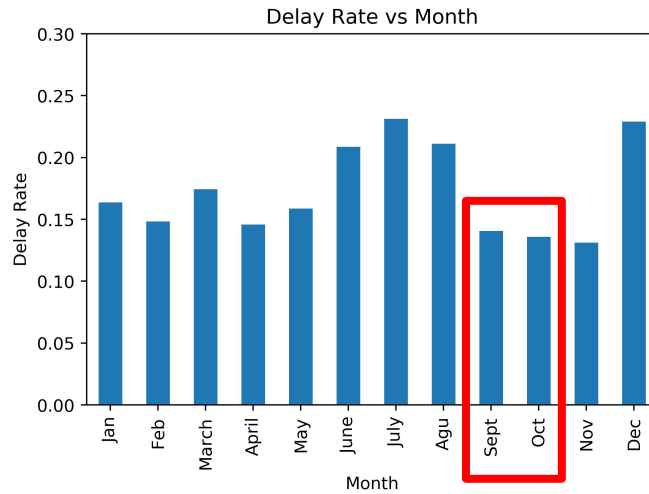
Flight Delay Rate and Duration vs Airport



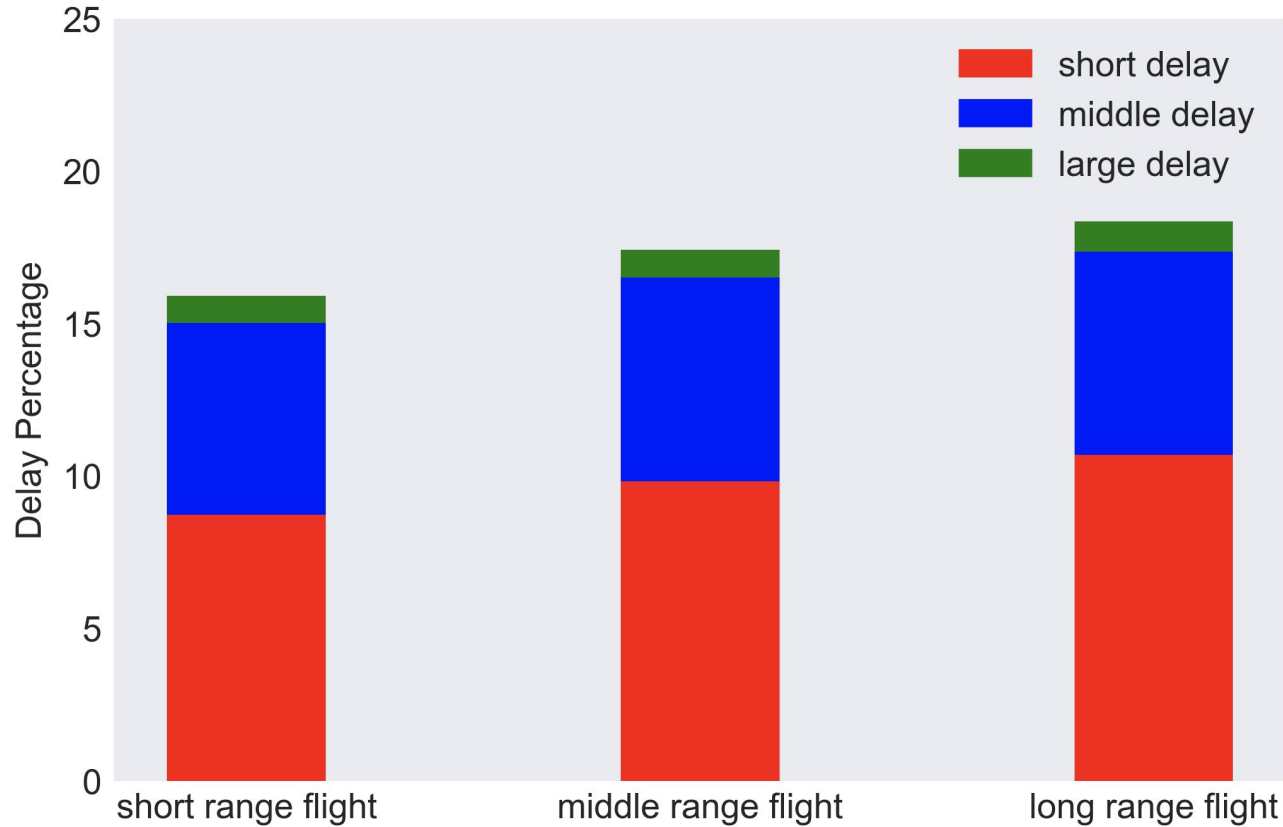
Example: (SAN) San Diego Airport

Airline and Month Filters for Varying Destination Airports

- **Specific** departing airport would be useful
- **Different** delay rate and duration from the general case
- Example of delay rate vs month - September and October



Delay Rate vs Flight Distance





Conclusion

- Meaningful results for various aspects of flights
- Useful information for passengers to use
- For future work, more in-depth analysis
 - Cause of delay
 - More data from previous years