Flight Delays using EDA

Done by: Venkata Sumanth Teja Bitragunta

Introduction

Flight delay is inevitable, and it plays an important role in both profits and loss of the airlines. An accurate estimation of flight delay is critical for airlines because the results can be applied to increase customer satisfaction and incomes of airline agencies.

Dataset is taken from the Kaggle and the Dataset consists of more than 0.5 million rows and 29 columns. I have taken three datasets which consists of the information of Airports, Airlines and Flights. I have done Data Visualization to get the clear view of the data

Data Set

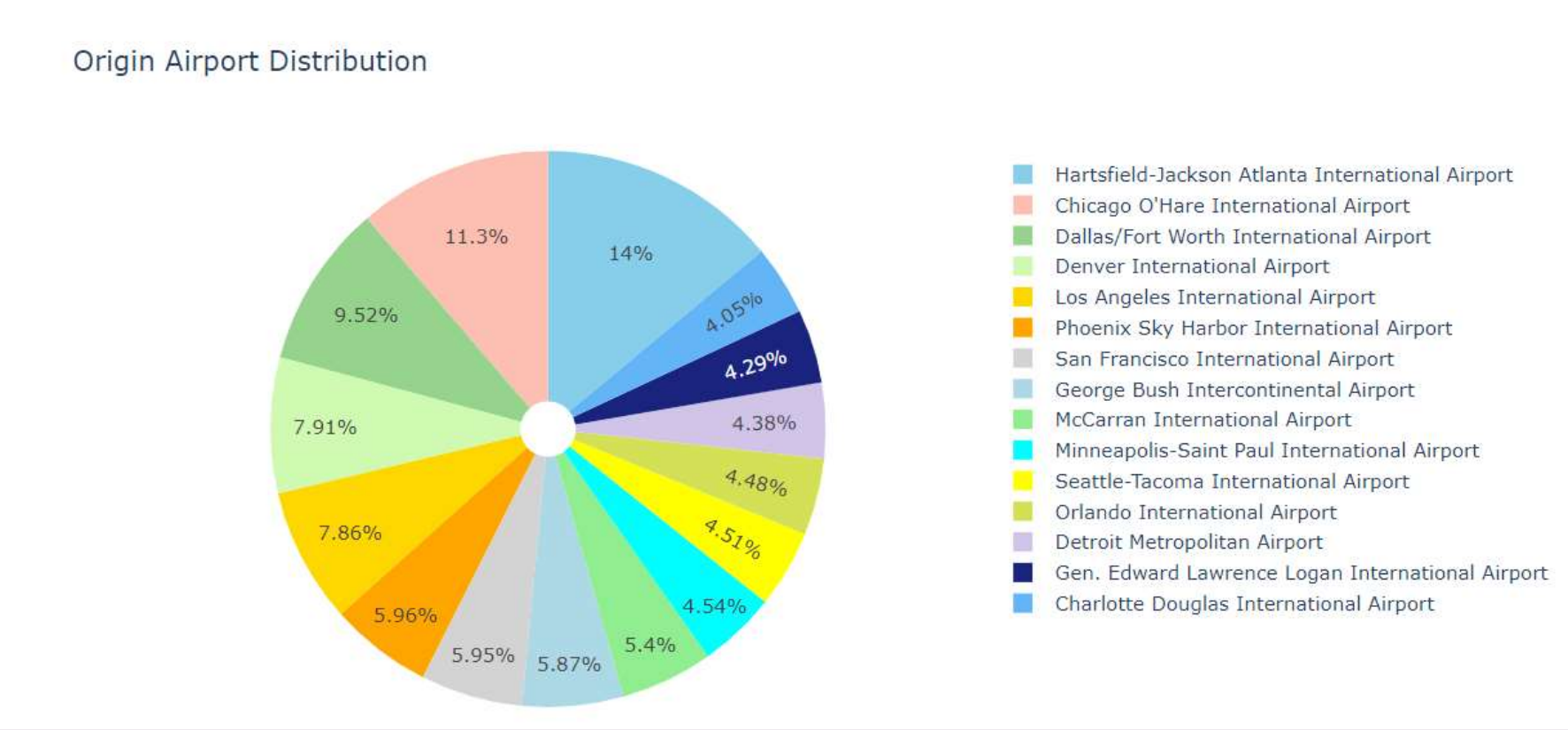
Column Name and its Datatype

```
In [5]: flights.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 5819079 entries, 0 to 5819078
Data columns (total 31 columns):
 #   Column               Dtype
---  ------               -----
 0   YEAR                 int64
 1   MONTH                int64
 2   DAY                  int64
 3   DAY_OF_WEEK          int64
 4   AIRLINE              object
 5   FLIGHT_NUMBER        int64
 6   TAIL_NUMBER          object
 7   ORIGIN_AIRPORT       object
 8   DESTINATION_AIRPORT  object
 9   SCHEDULED_DEPARTURE  int64
 10  DEPARTURE_TIME       float64
 11  DEPARTURE_DELAY      float64
 12  TAXI_OUT             float64
 13  WHEELS_OFF           float64
 14  SCHEDULED_TIME       float64
 15  ELAPSED_TIME         float64
 16  AIR_TIME             float64
 17  DISTANCE             int64
 18  WHEELS_ON            float64
 19  TAXI_IN              float64
 20  SCHEDULED_ARRIVAL    int64
 21  ARRIVAL_TIME         float64
 22  ARRIVAL_DELAY        float64
 23  DIVERTED             int64
 24  CANCELLED            int64
 25  CANCELLATION_REASON  object
 26  AIR_SYSTEM_DELAY     float64
 27  SECURITY_DELAY       float64
 28  AIRLINE_DELAY        float64
 29  LATE_AIRCRAFT_DELAY  float64
```

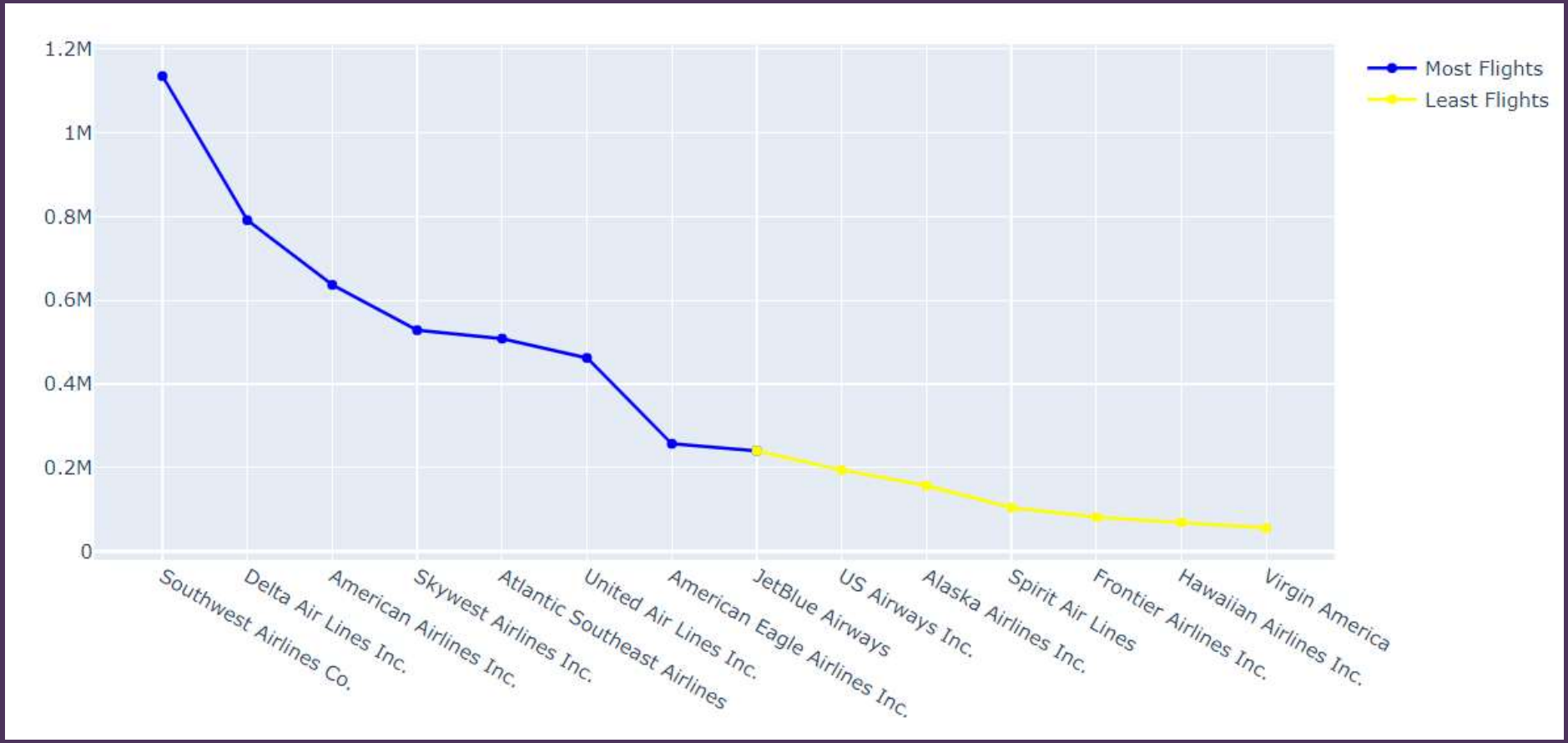
Airports with most Flights

- From this pie chart we can see that Atlanta airport is having the greatest number of flights

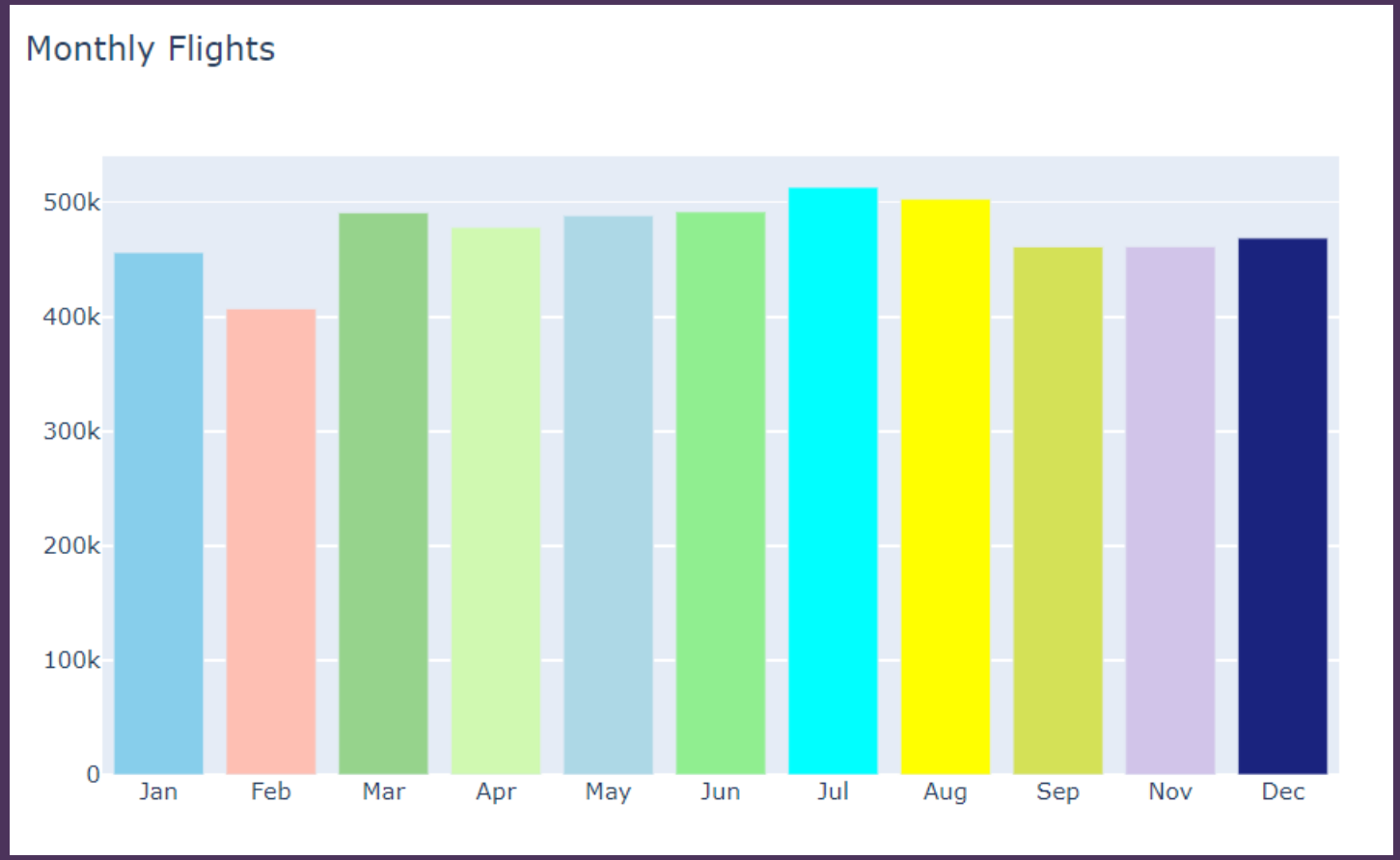


Cities with most Flights

Airline Distribution

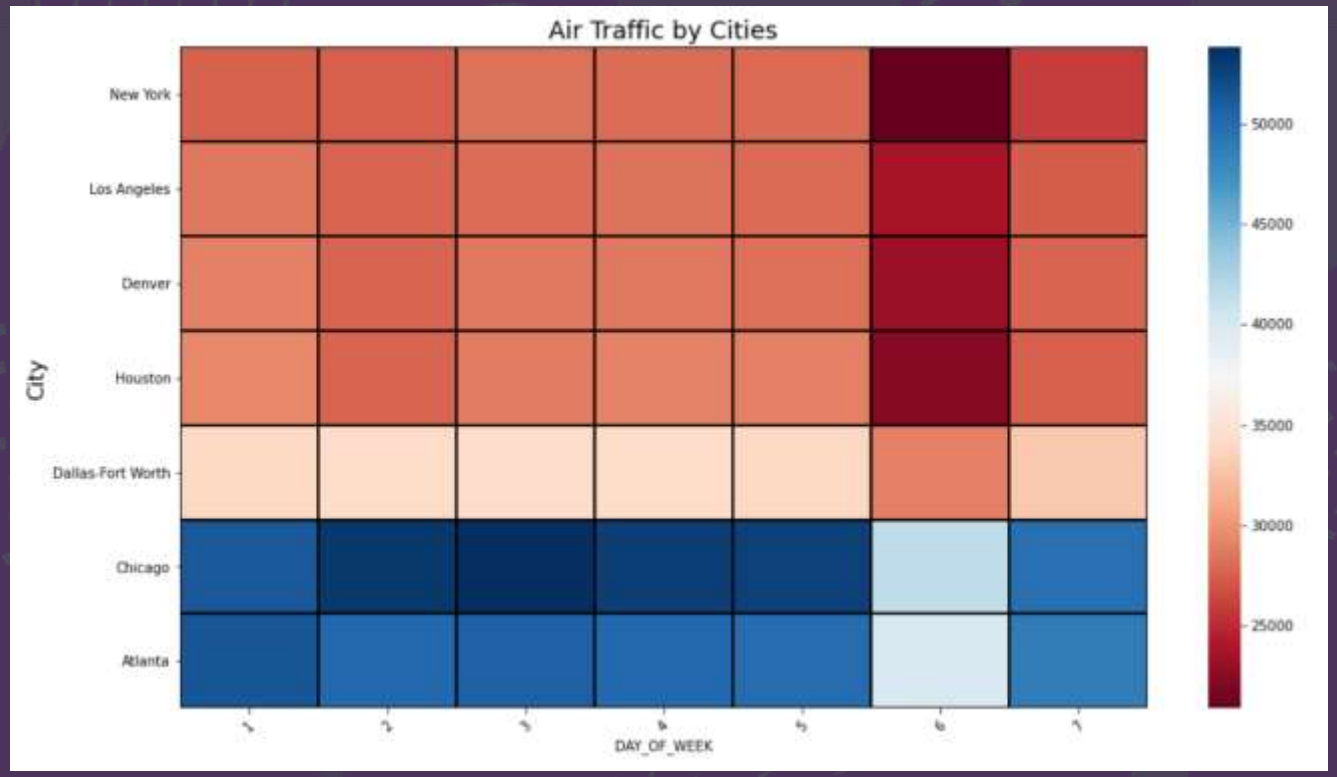


Number of flights per month

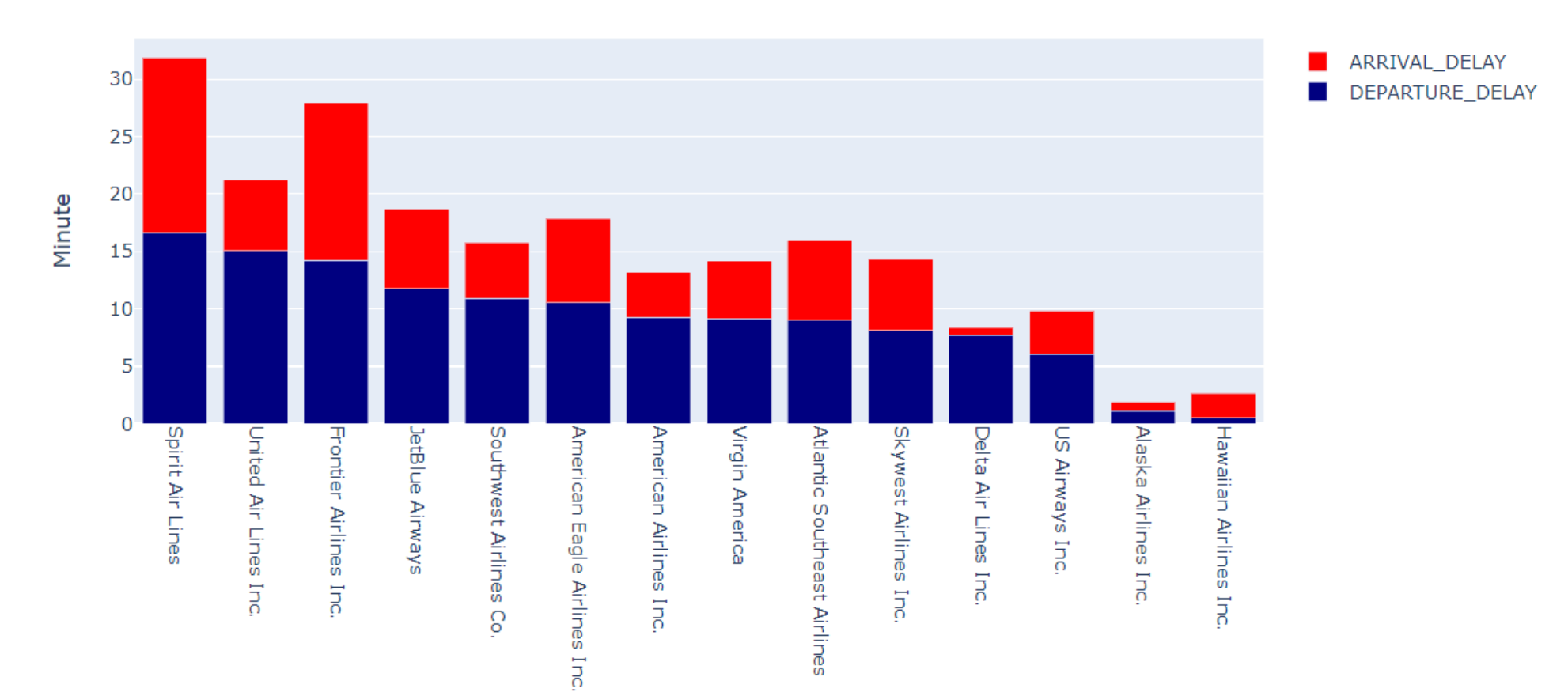


Air traffic by Cities

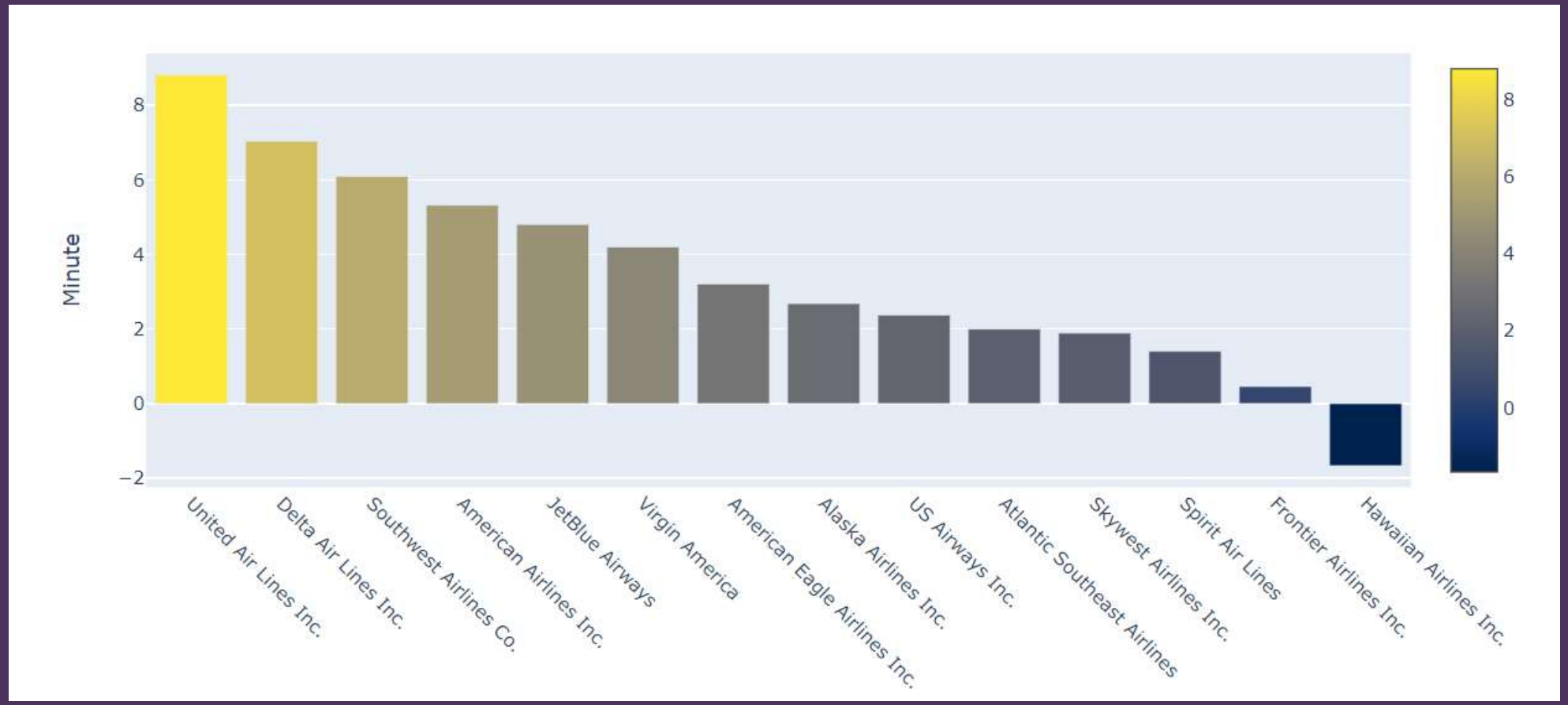
It is the Choropleth graph to identify the air traffic in the cities in a week.



Mean Arrival Delay and Departure delay



Mean(Departure-Arrival) Delay of Airlines



THANK YOU