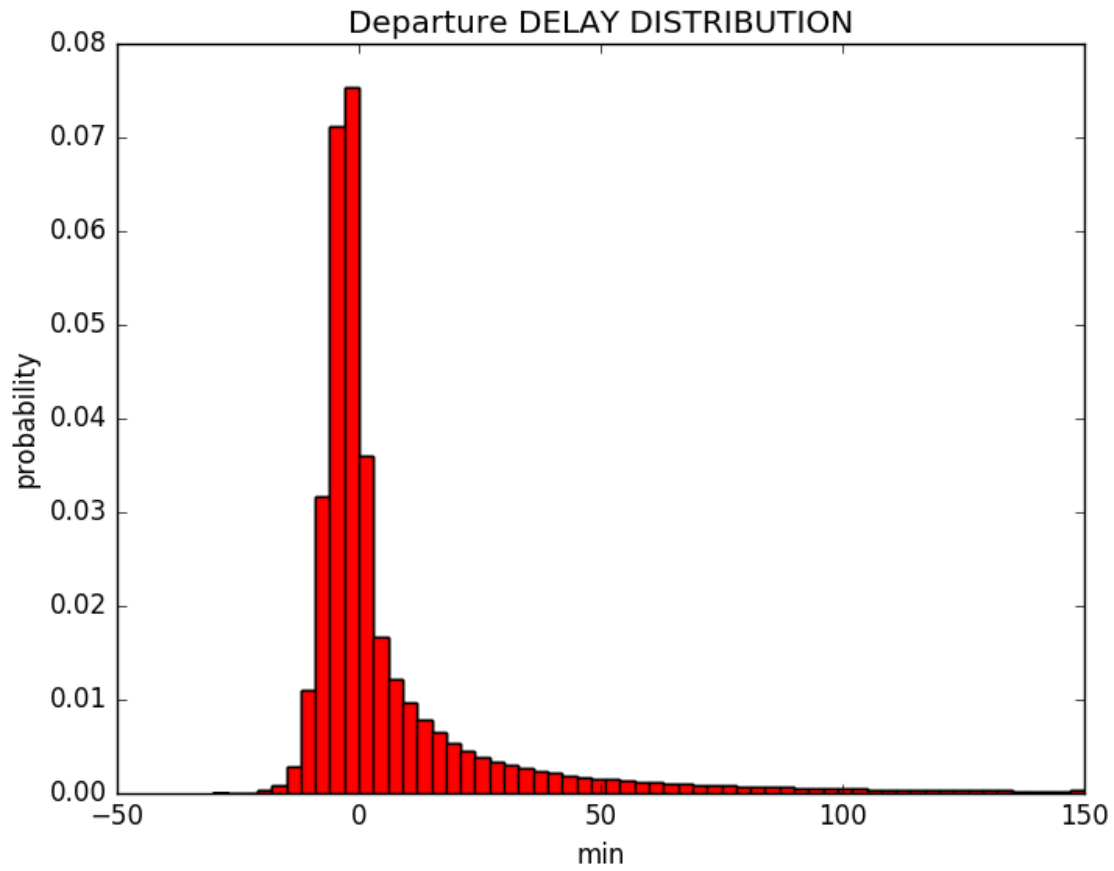


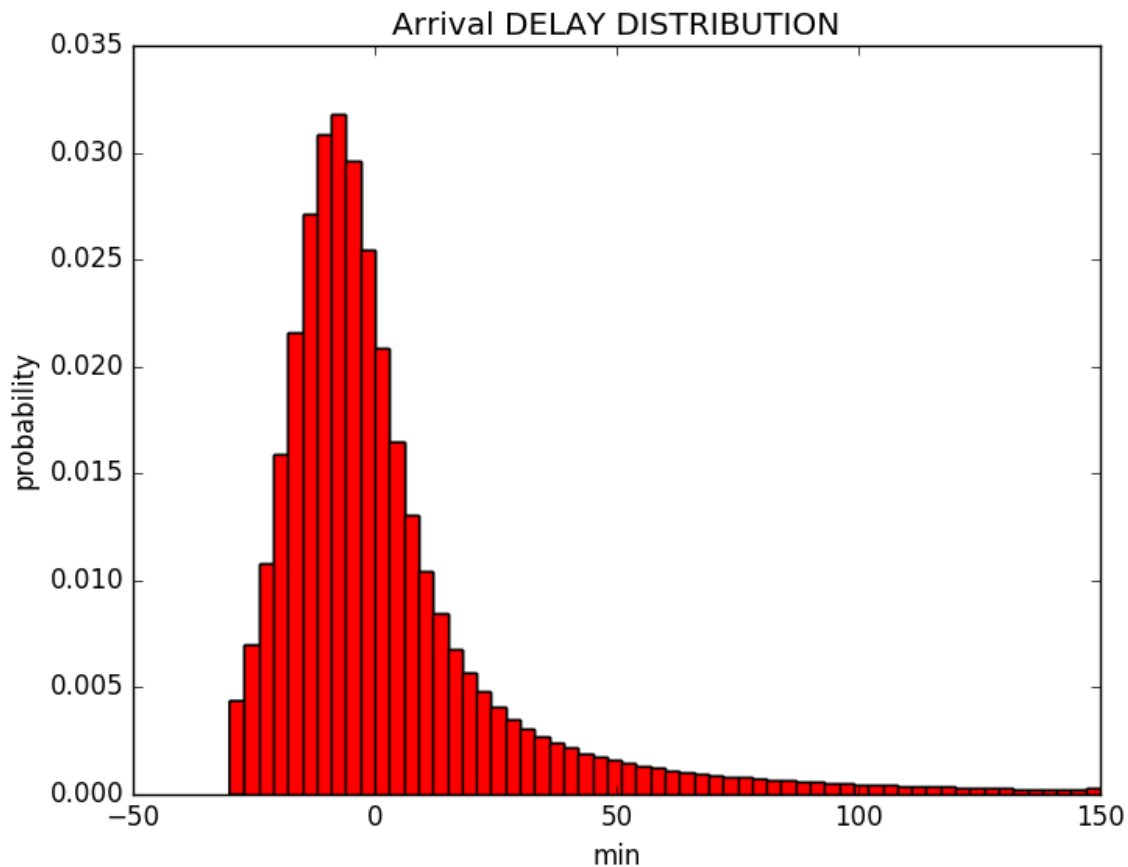
# Delay related insights from the flight data

Sonia Jaiswal  
13EC35003  
IIT KHARAGPUR

## Departure and Arrival Delay Distribution:

- First the departure and arrival delay distribution is plotted over the whole data.



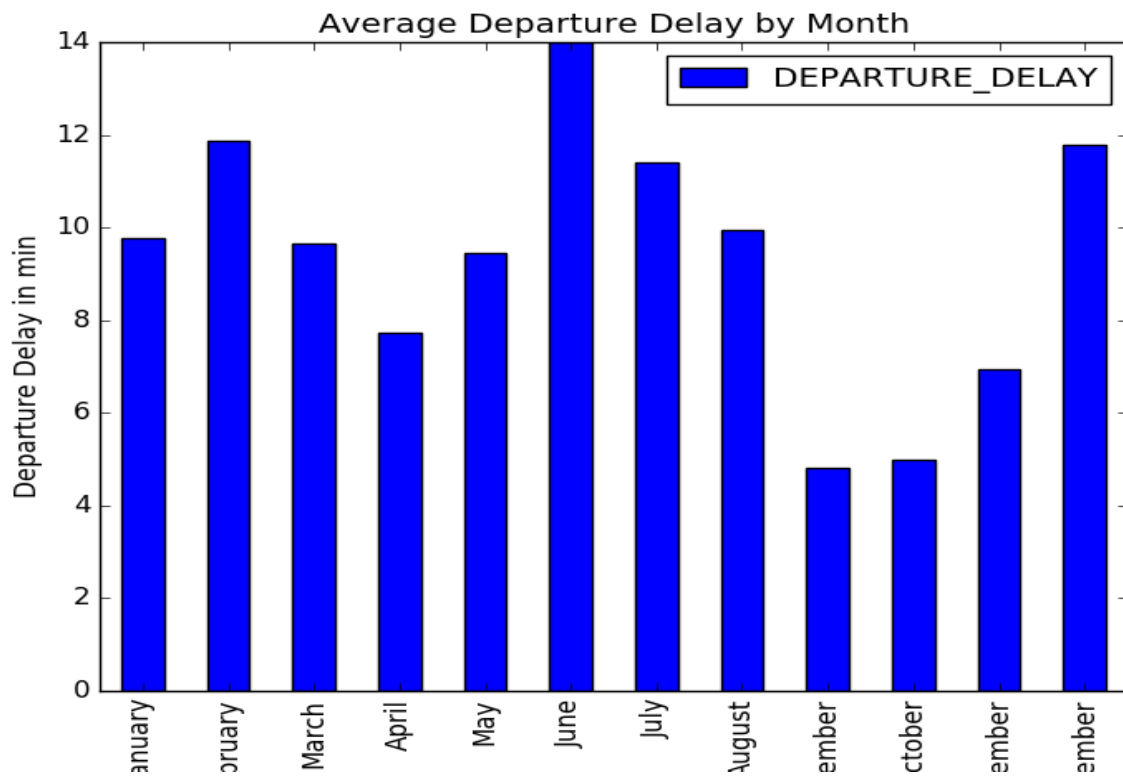


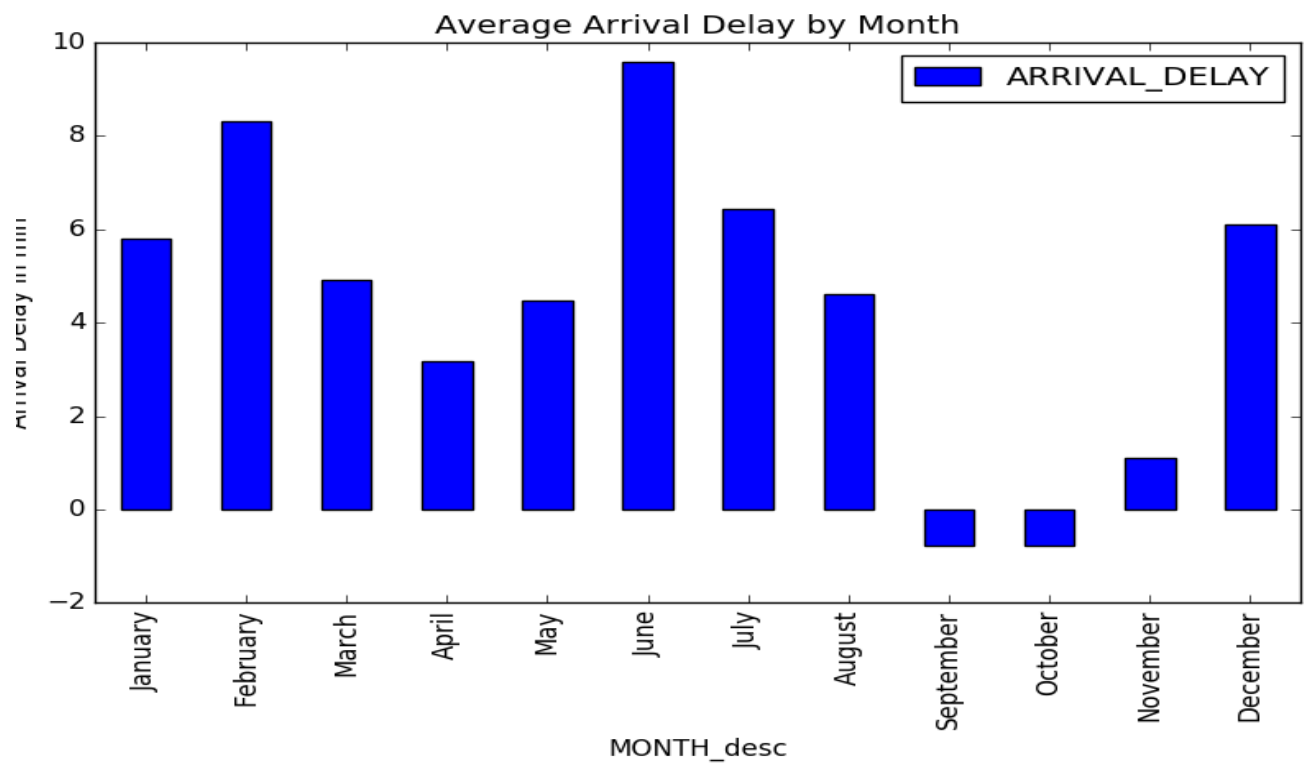
- We notice a much higher probability of short delays - actually negative, so we can consider them advances - for departure delays and a wider distribution (in minutes) for arrivals. Notice the long right-hand tails. Some flights are delayed for very long times, over two hours. On the other hand, the delays are centred just below zero.
- In both cases, the mode of the distribution is less than zero, meaning most of the flights leave from gate and arrive at gate even before the published schedule time of departure and arrival. As we will show below, the longer delays cancel out the shorter negative delays (advances), leading to average delays that are above zero.
- The x-axes for the two plots are to scale. As a result, we can see that the arrival delay distribution, compared with the departure delay distribution, leans toward left.
- A flight delay is defined as the time difference between the scheduled time of an event compared to the actual time of the event. Airlines usually put extra buffer time in a flight to ensure on-time arrival. Therefore, the distribution of the difference in the departure delay and arrival delay indicates that some departure delays are recovered during the flights due to the extra amount of time embedded in the flight time between two airports.

- In terms of analysis, it makes sense to consider separately departure and arrival delays, since the impacting factors may be different for each

### Average Departure and Arrival Delay by Month:

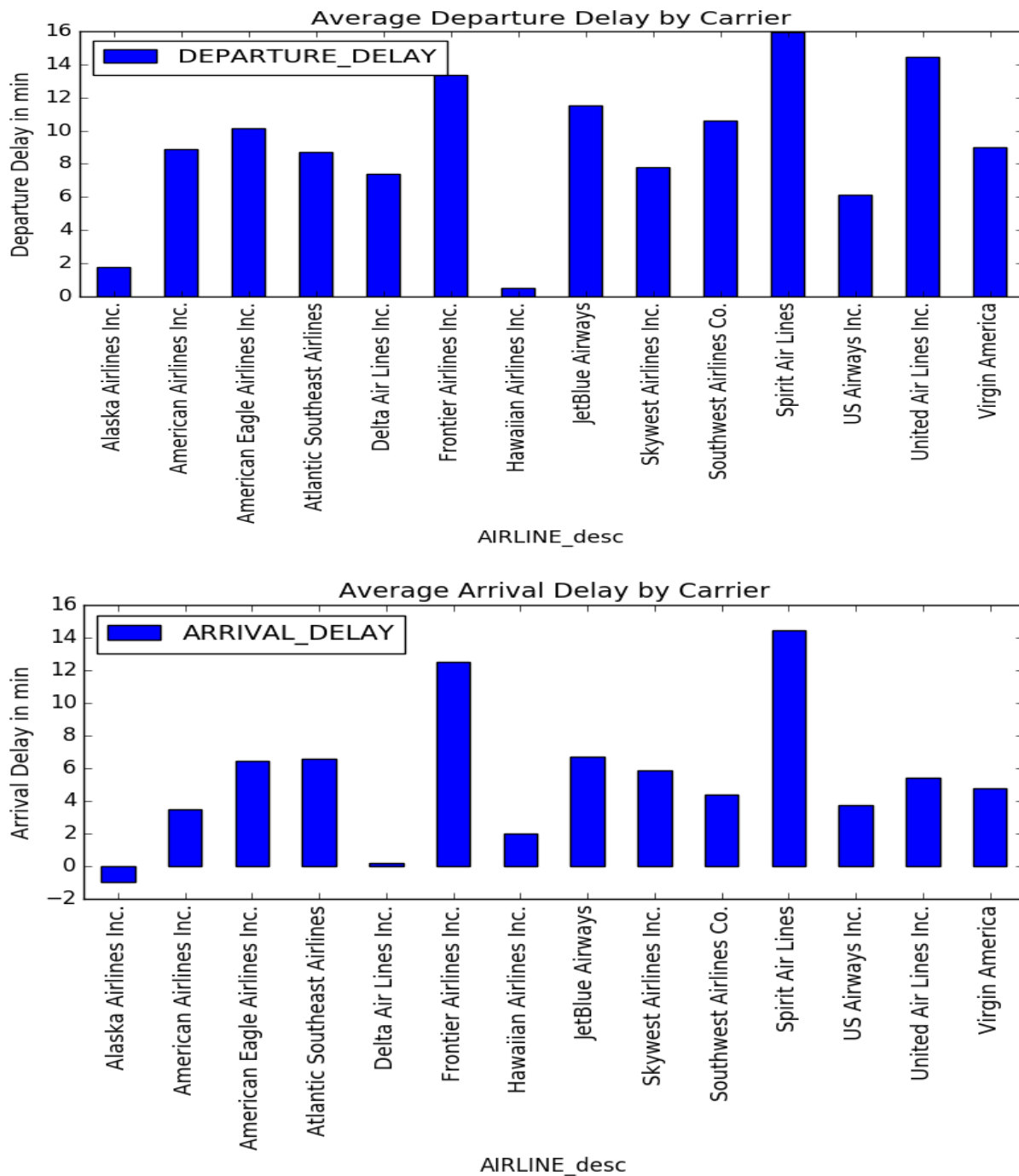
- It can be seen that the delay is maximum in June may be due to summer vacations. It's also high in December and January which is obvious due to weather.





## Average Departure and Arrival Delay by Carrier:

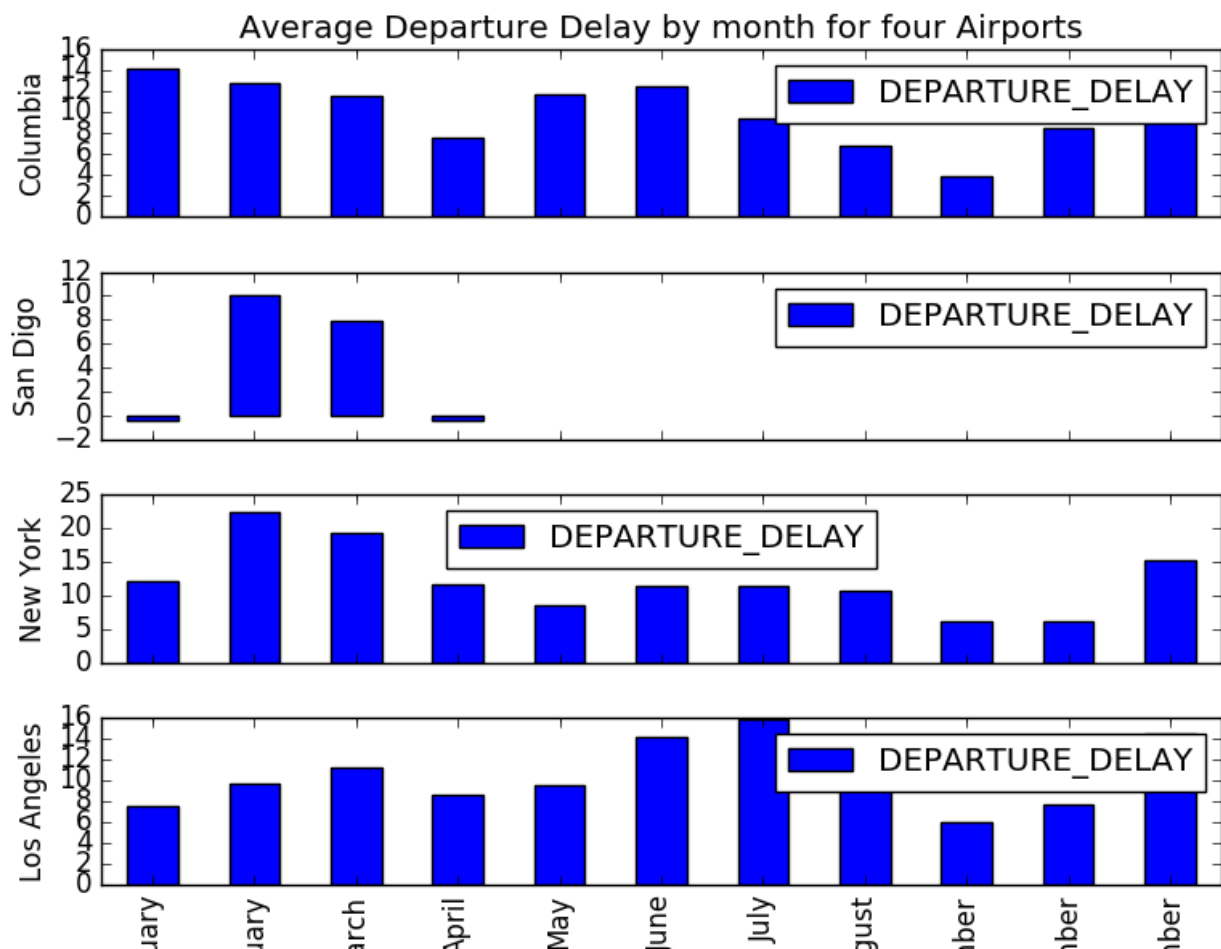
For flights with one of 20 unique carriers, average flight delays vary consider



Here in particular the analysis is affected by the number of flights for each carrier. Some carriers with a small number of flights, like Hawaiian Airlines (HA), have the lowest mean delays. This is to

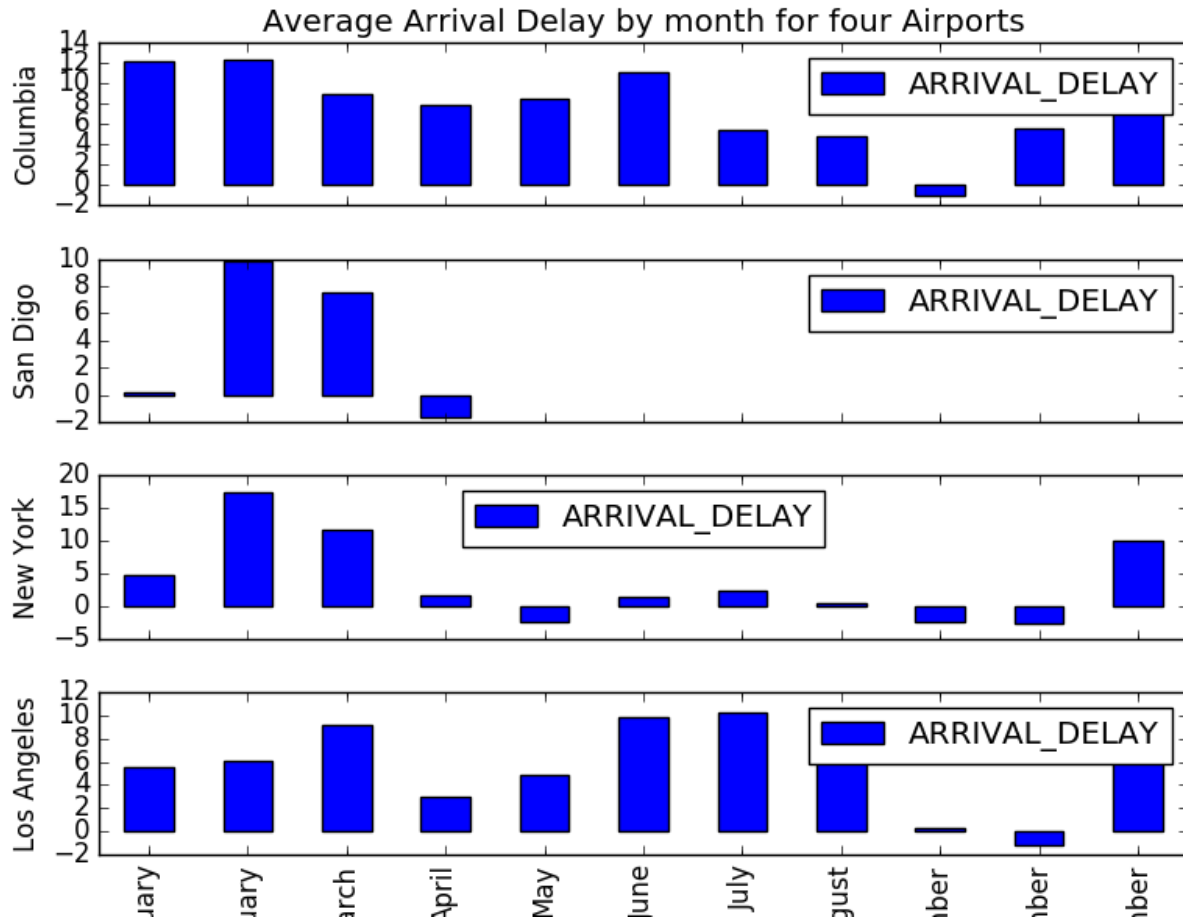
be expected. Hawaiian Airlines mostly operate flights between Hawaii and the continental United States. Since these flights are mostly 5-6 hours, they are usually exempt from Ground Delay Programs (GDPs) which occurs when an airport encounters inclement weather conditions. This means that no delay is imposed on these flights by the air traffic management system. Therefore, the average delays for these two carriers can be generally lower than other (scheduled) carriers, and can in fact be quite short.

### Average Departure Delay by month for four Airports:



We see differences between airports - Columbia, New York and Los Angeles are similar to the overall profile for mean delay at all airports, with higher delays in December and January and a midsummer bump. On the other hand, San Diego shows higher mean delays in the beginning of the year and lower or no delays in other months. New York has more delay in the springtime, with February, March and June with a higher mean delay than in December. In all locations, December is a month with higher than normal delays.

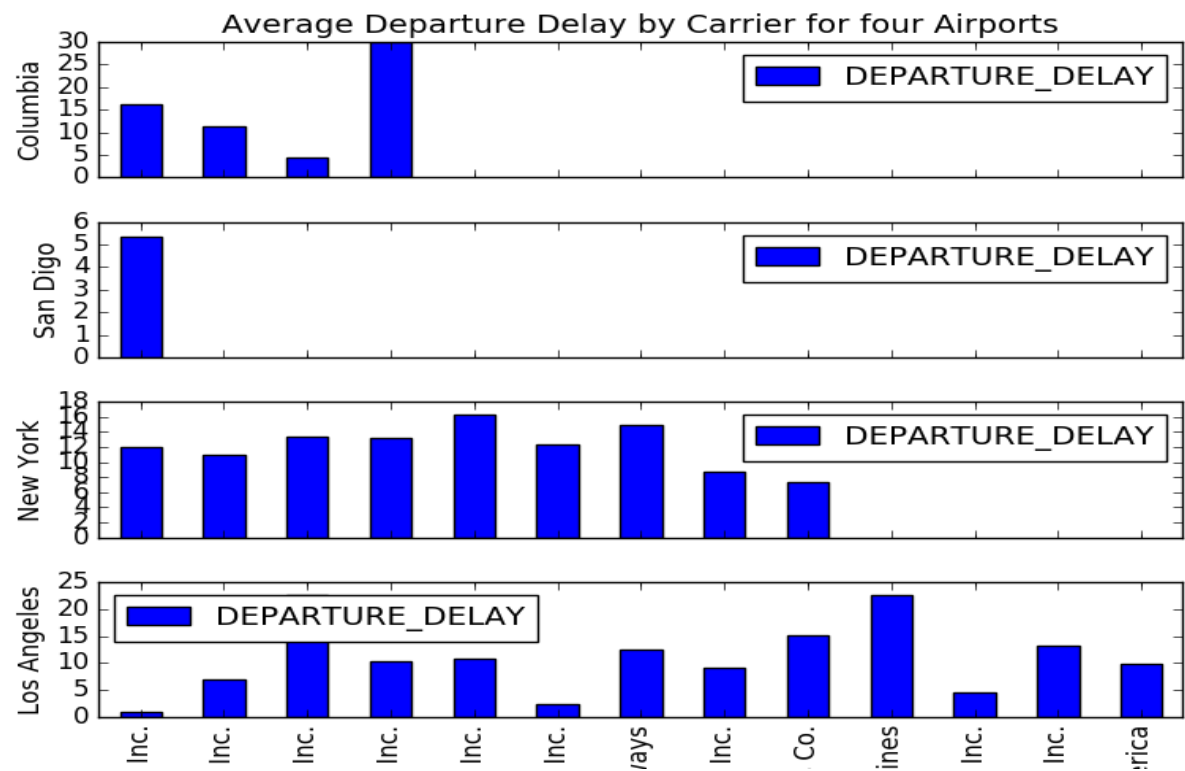
Average Arrival Delay by month for four airports:

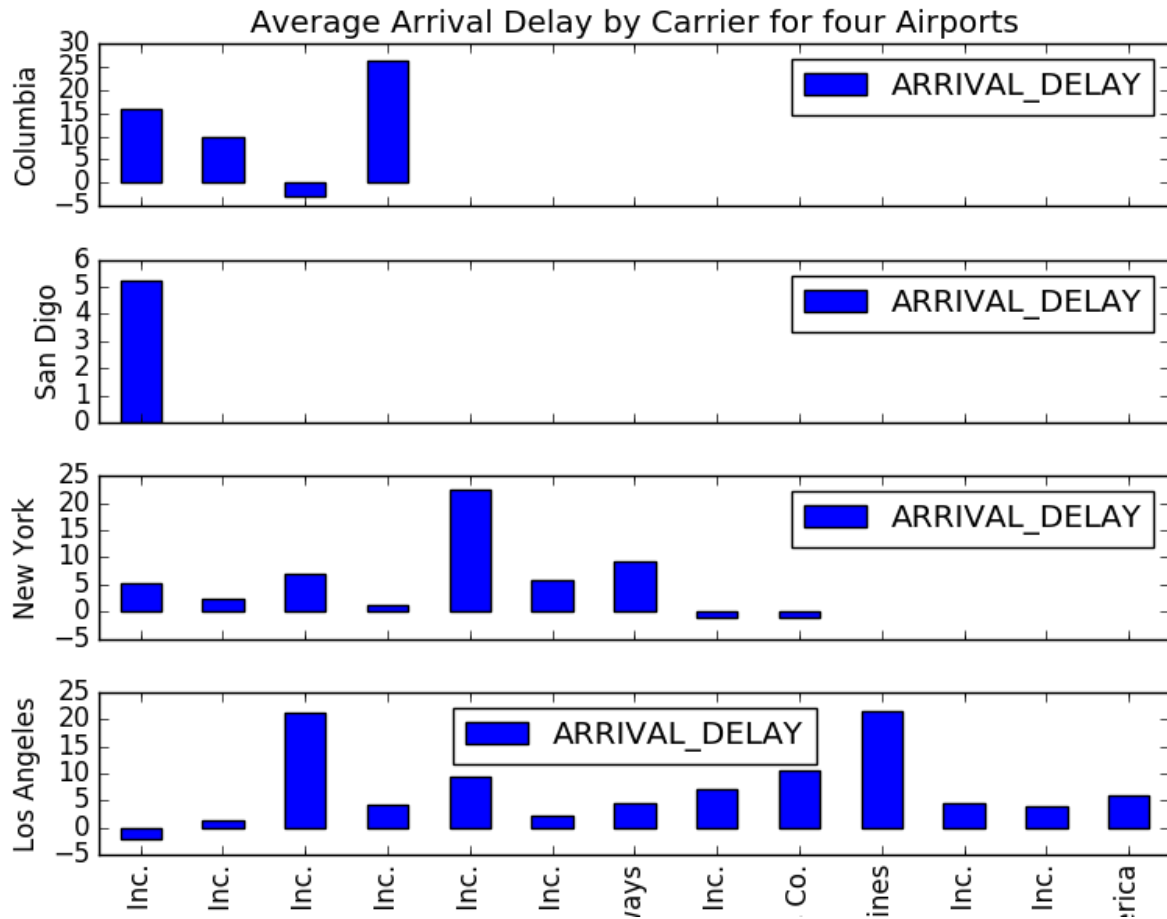


For arrival delays, we see two distinct peak months for Chicago Columbia and Los Angeles in February and June. The latter may be because of holiday travel at these typical vacation travelling periods. The former may also be entirely weather-related since these three months have the worst climatic conditions.



Average Departure and Arrival Delay by carrier for four Airports:

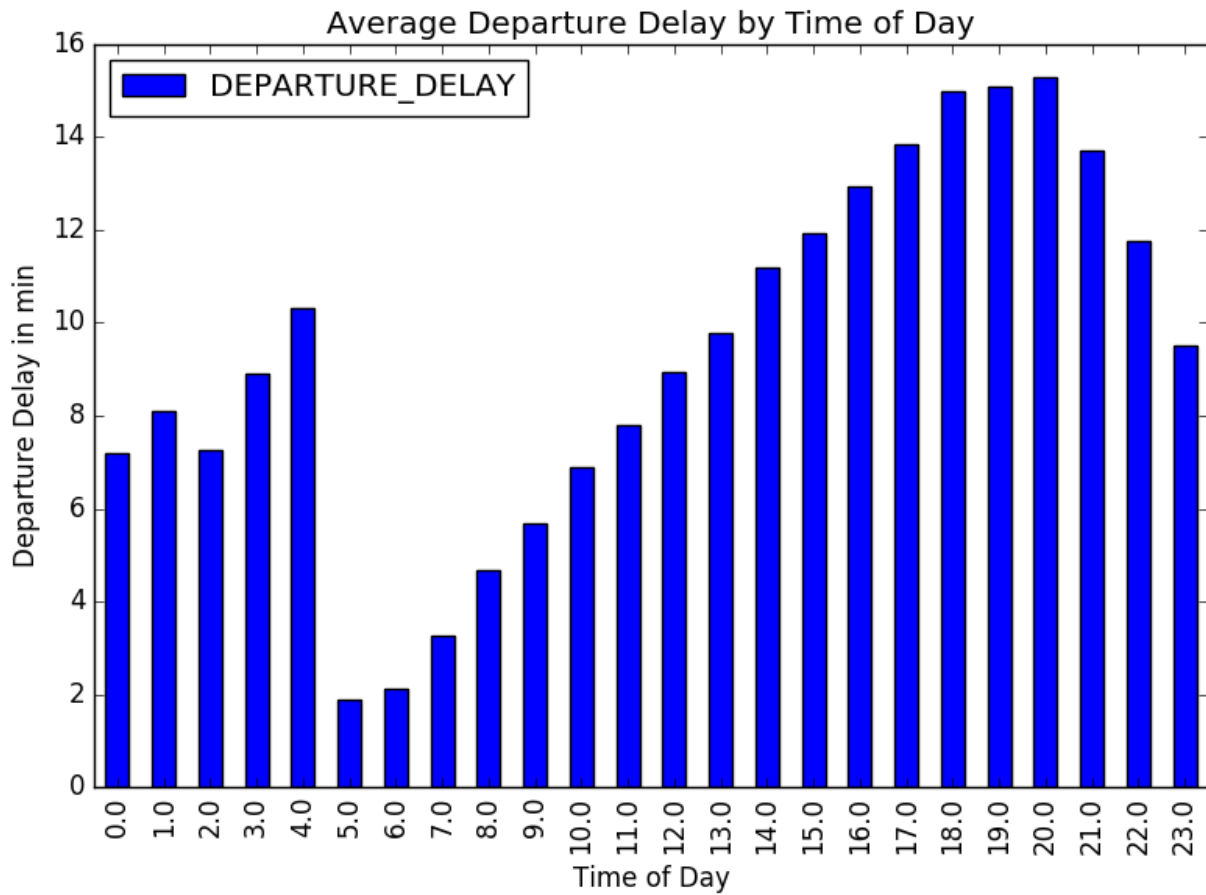


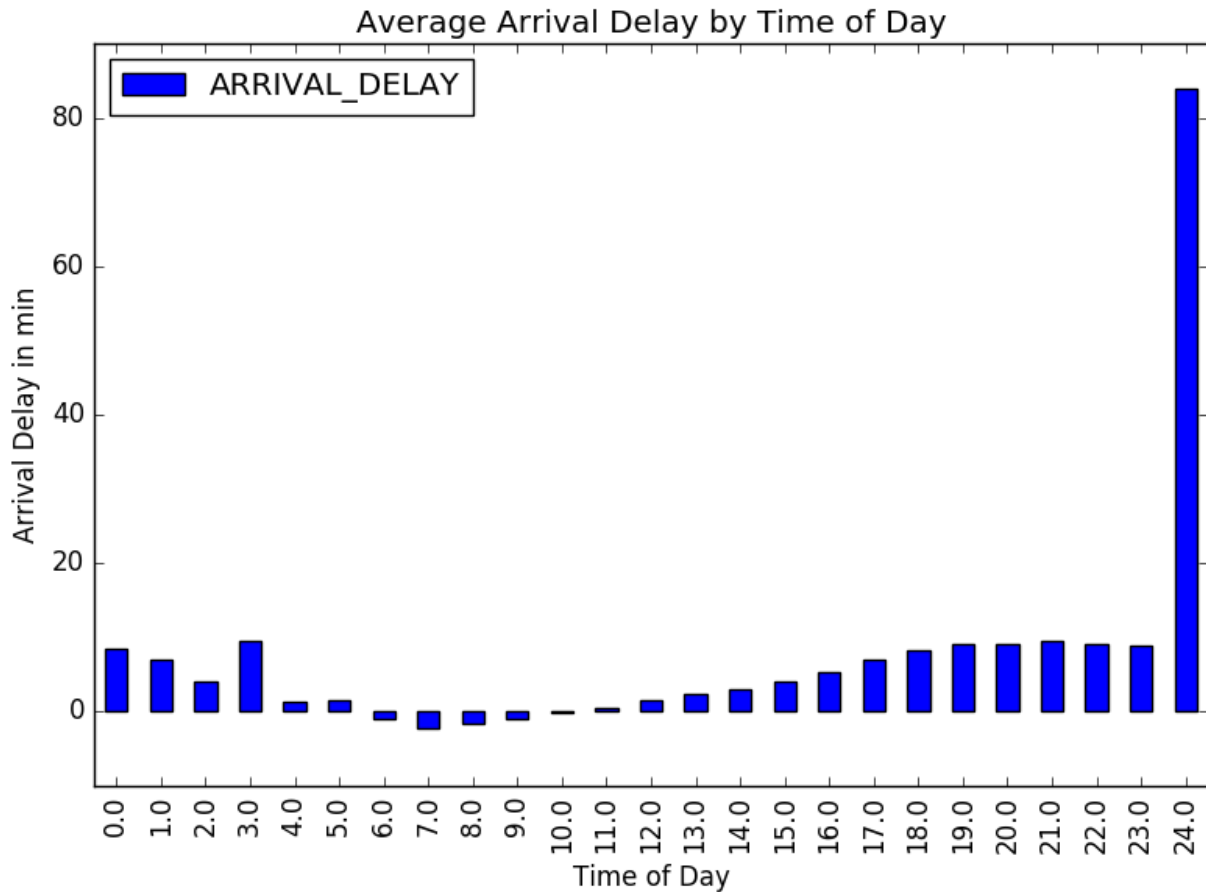


Here we see that certain carriers, the small and/or low cost ones, have the largest departure and arrival delays, such as Delta Airlines Mainstream carriers like North-Western and US Air have smaller delays. It is also apparent that the profile is more or less similar for both departure and arrival delays. San Diego has delay only on Alaska Airlines may be it has only some Airlines facilities and low mobility.

## Average Departure and Arrival Delay by Time of Day”

From the plots, we see a marked "V" shaped decline in delay with the lowest delays in early morning hours.

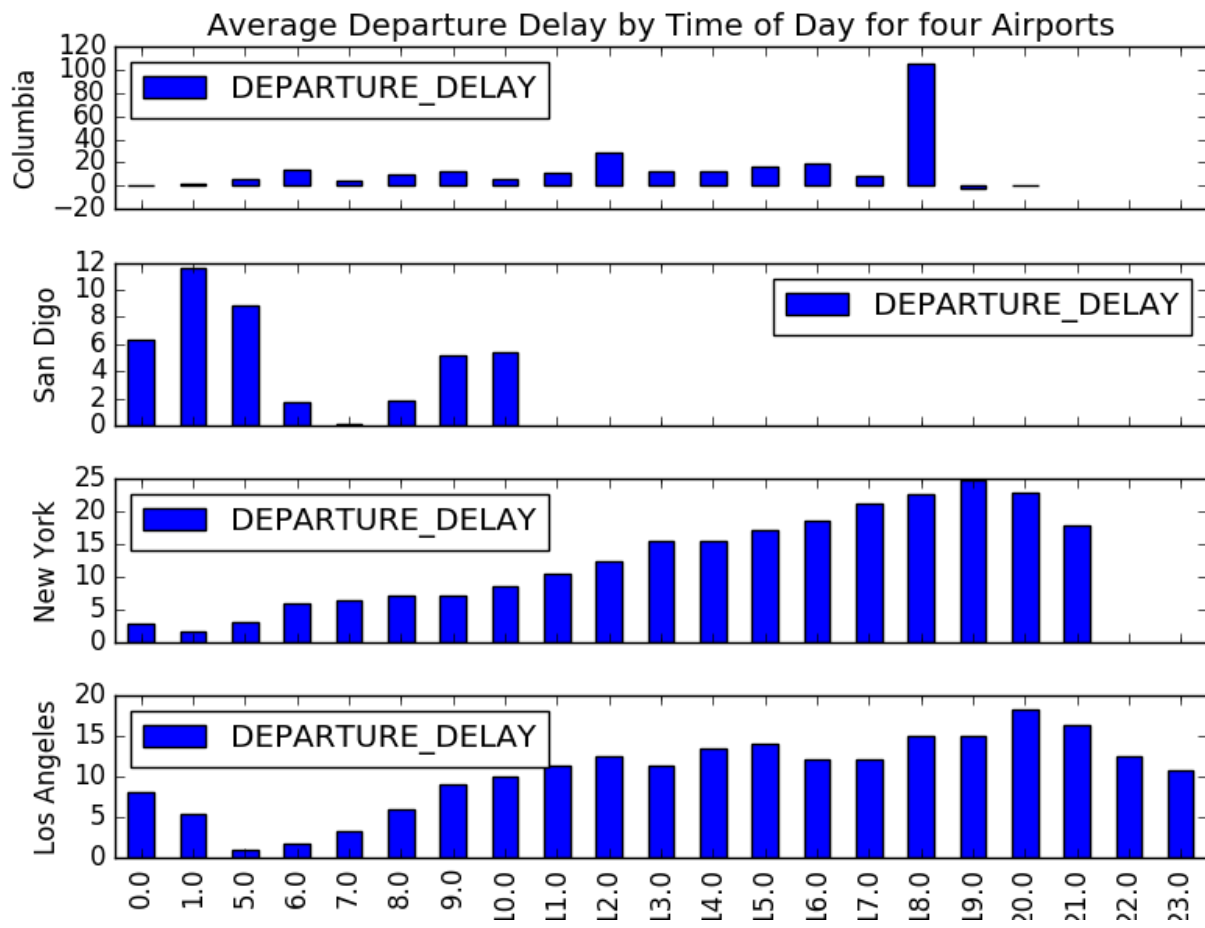




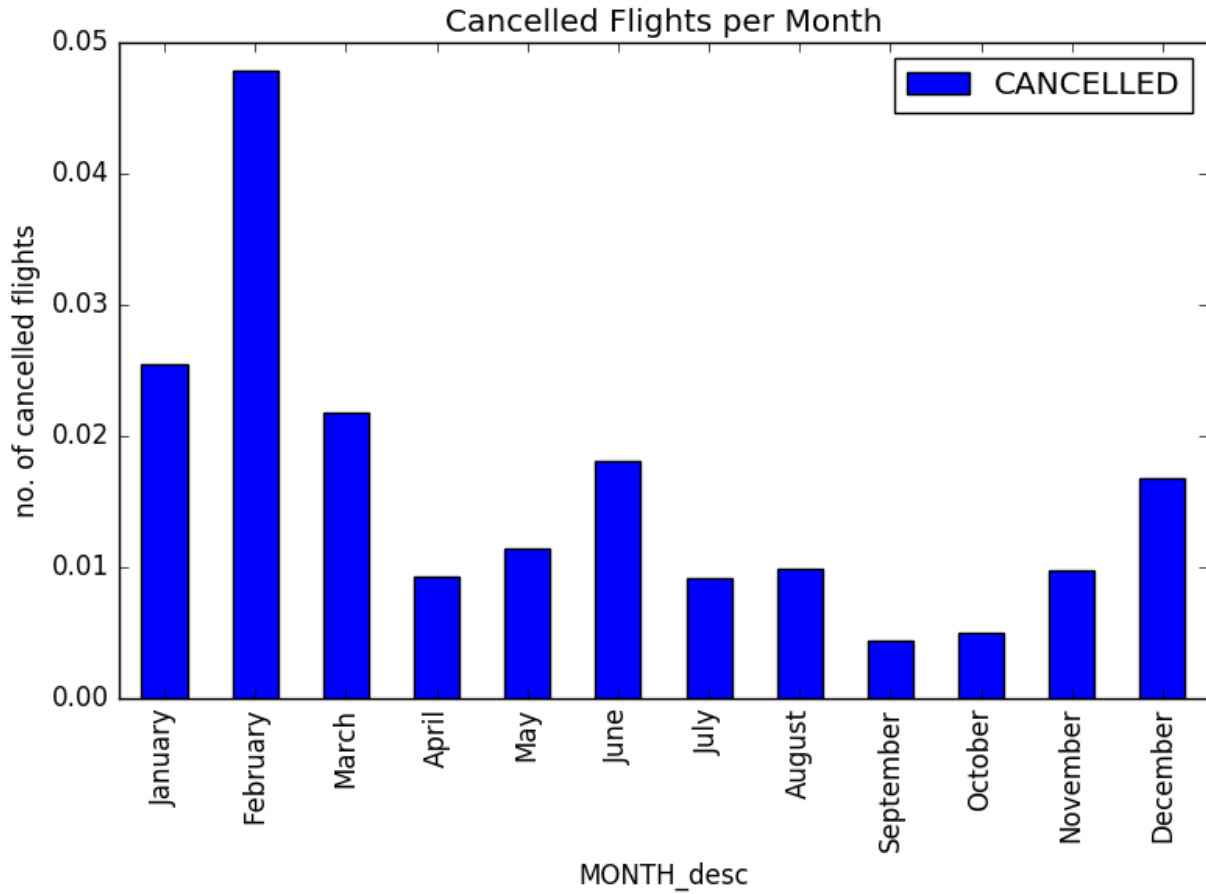
Both departure and arrival delays accumulate from the earlier morning hours reaching their peaks in the evening hours. For departure, the highest mean delay is during prime-time of 18:00 to 21:00, and for arrivals, it is slightly later (the average flight duration is a few hours) and peaks at around 24:00. The increasing of flight delay by the hours of the day is mainly caused by flight delay propagation. Although a flight is built with scheduled buffer time for unforeseeable flight delay during the flight operations, it is not sufficient to cover all types of delay. As a result, if a flight is delayed, the next flight has to wait for the late arrival flight to be ready before it can be operated. Hence, flight delays for both departure and arrival flights do increase over time. As a conclusion from the exploration, we see that the variation in mean delay by hour that we see implies that 'Hour of Day' should be a good predictor of flight delay.

## Average Departure Delay by time of day for four Airports:

Pattern is almost similar except in San Diego due to low mobility, delay is minimum during the evening time.

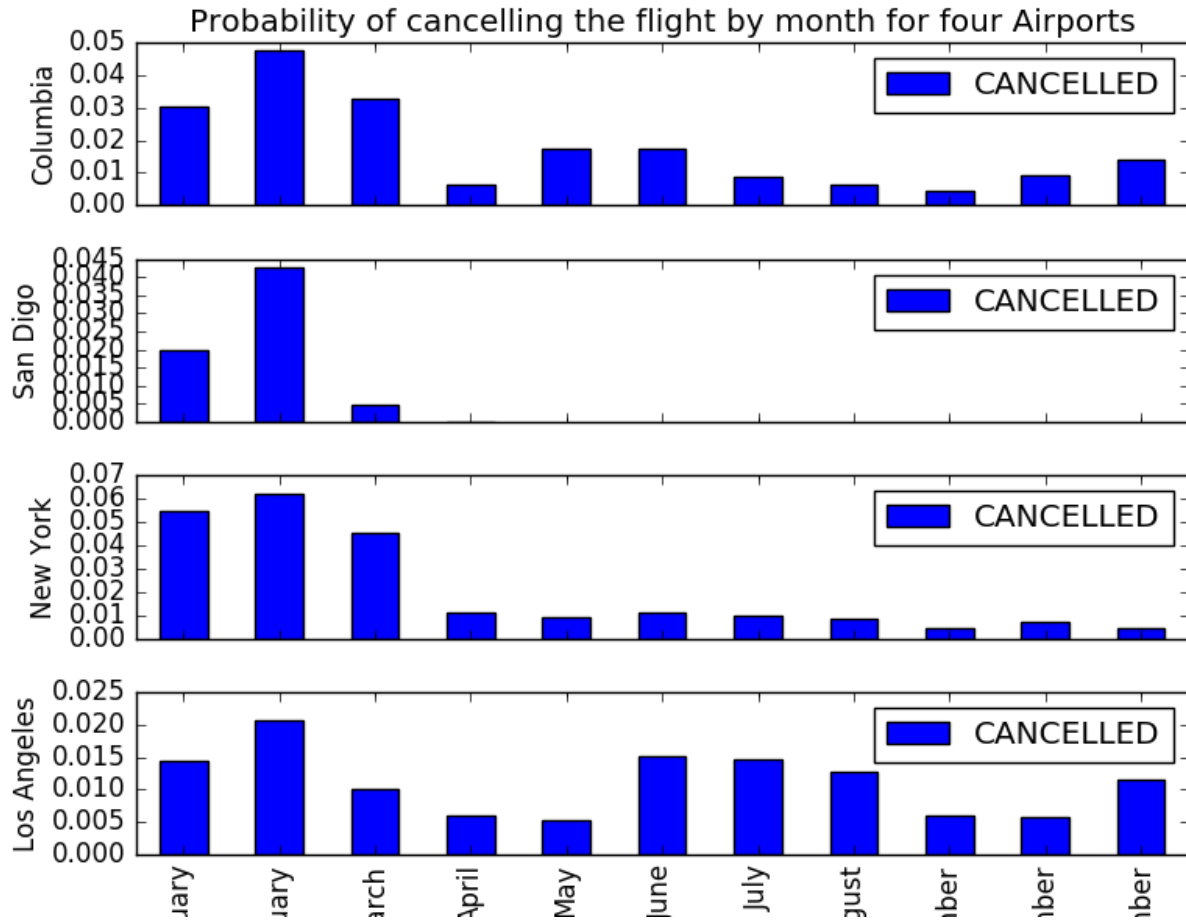


## Probability of Cancelling Flights per Month:



Probability of cancelling is highest in month of February due to weather condition then in January and December. After that in June due to vacation travelling thus causing higher mobility

Probability of cancelling the flight by month for four Airports:



For Los Angeles pattern is same as for all the data in New York and Columbia due to weather highest is in month of February In San Diego probability of cancellation is very high in February all other time due to low mobility low no. of flights hence low probability of cancellation.