

FIRST SEMESTER PROJECT WORK

SUBJECT – PROGRAMMING FOR DATA SCIENCE

AUTHOR – SAWAN AICH

SHASHWATA MONDAL

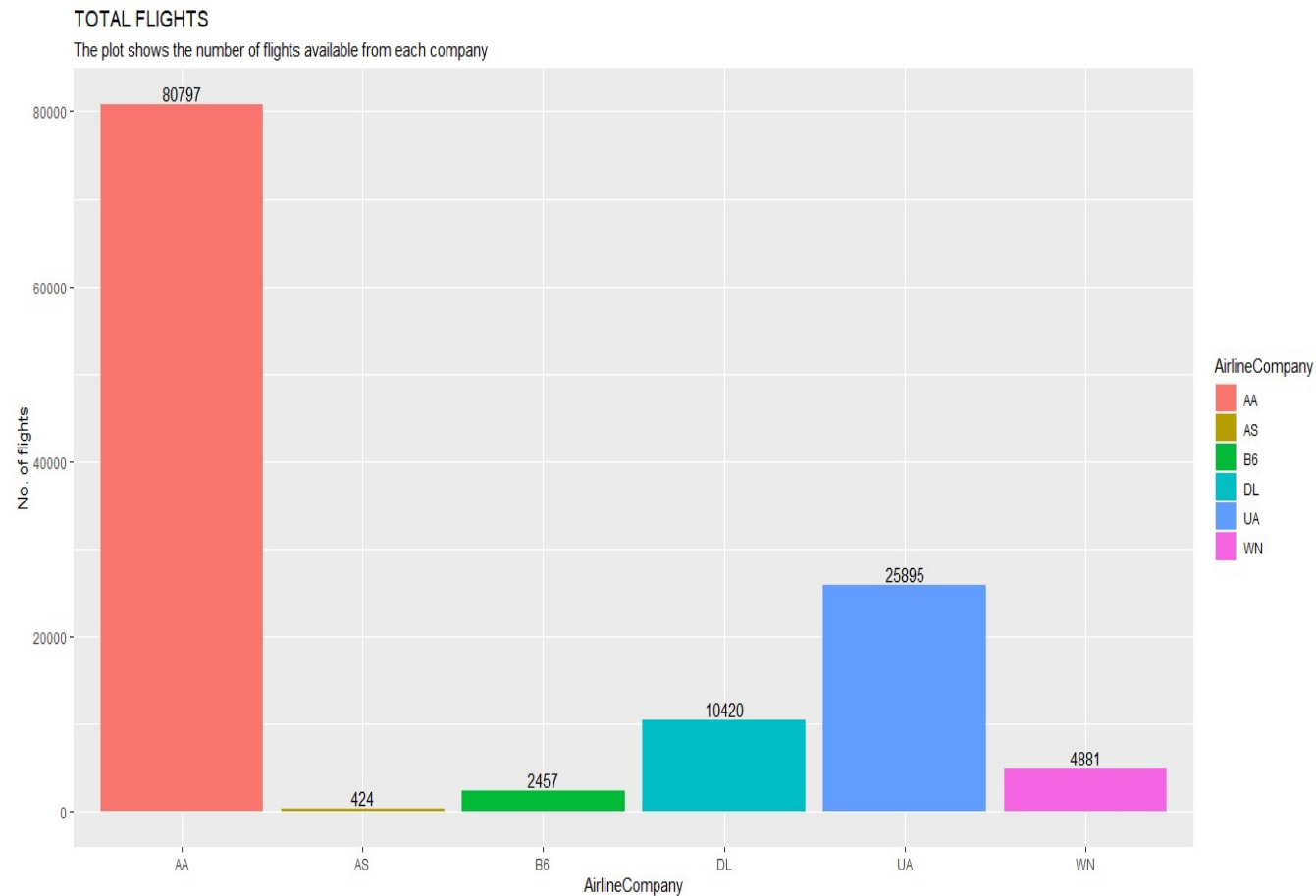
SOURAV BHATTACHARJEE

PROJECT OBJECTIVE

*TO FIND OUT THE MOST ECONOMICAL FLIGHT
TO TRAVEL FROM ANY ORIGIN*

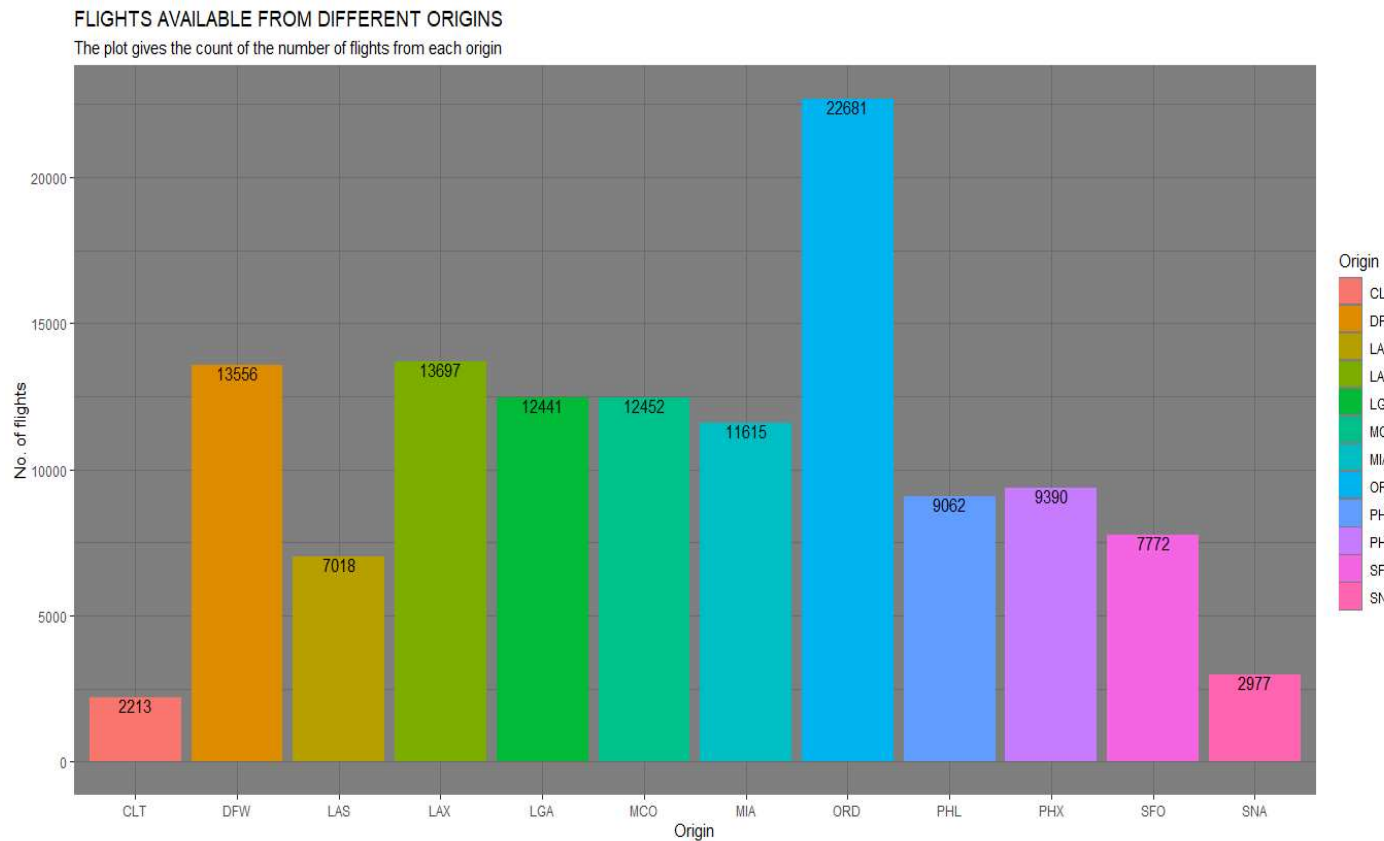
How many flights are available to travel?

The plot shows all the flights that are available to travel from six different companies. The plot also shows that American Airlines Inc. (AA) has the greatest number of flights available whereas Alaska Airlines Inc. (AS) has the least number of flights available.



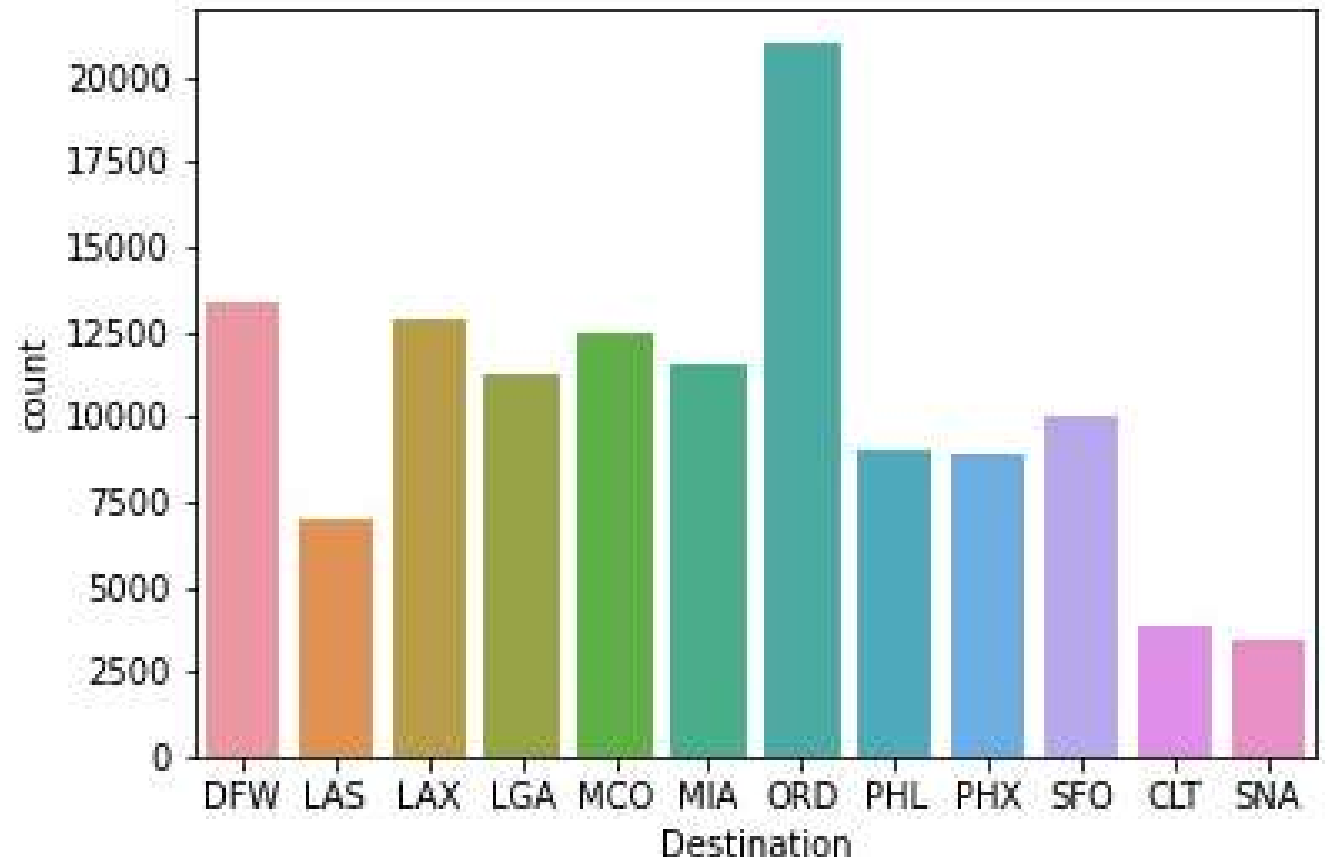
How many flights are available from each origin?

This plot shows that the origin ORD (Illinois) has the greatest number of flights leaving for various destinations. It is also evident that CLT (North Carolina) has the least number of flights going to different destinations. However, the flights are of which companies cannot be specifically identified here, so we need few other plots to analyze the data further.



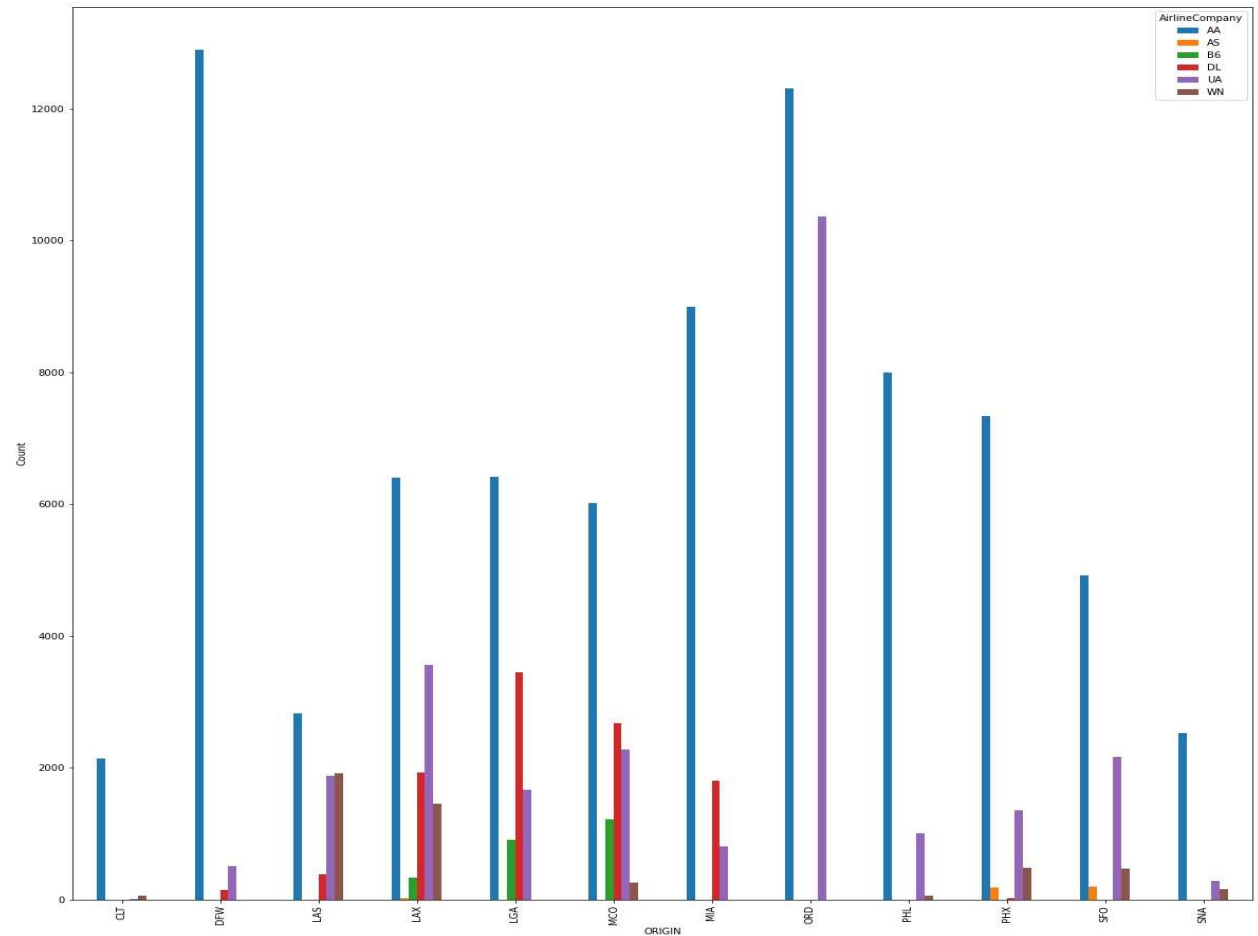
How many flights are available to every destination?

Here we see that ORD (Illinois) has the maximum number of flights reaching the place. In other words, most of the flights reach this place. However, we will be seeing something really interesting in an upcoming plot which will change our views about this place's flight availability.



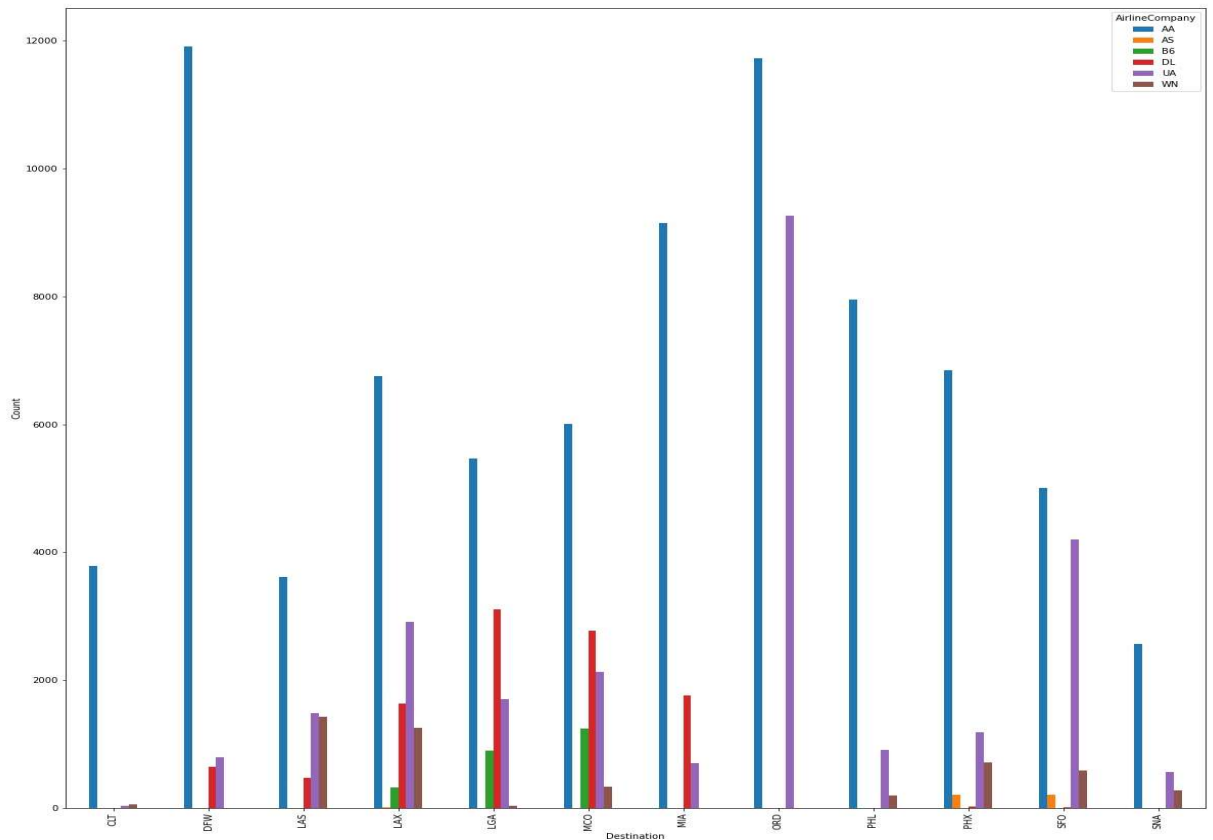
Number of flights coming from every origin point based on airline company.

In this plot, we come to know that majority of the American Airlines flights fly from DFW (Texas). ORD (Illionois) consists of only American Airlines and United Air Lines flight services available. LGA (New York) has a proper balance of all the flight services, Alaska Airlines being the only one not available there.



Number of flights coming to every destination point based on airline company.

Here, the trend is quite similar to that of the origin point of view. Here, we see that ORD (Illinois) has the highest number of flights reaching there, but the companies are limited to American Airlines and United Air Lines. So, though flight availability is high, the choices are way too less.



SB1

Heatmap showing the number of flights between 12 US cities. The color scale ranges from 0 (dark blue) to 3500 (dark red). The highest number of flights is between LGA and MIA, exceeding 3500.

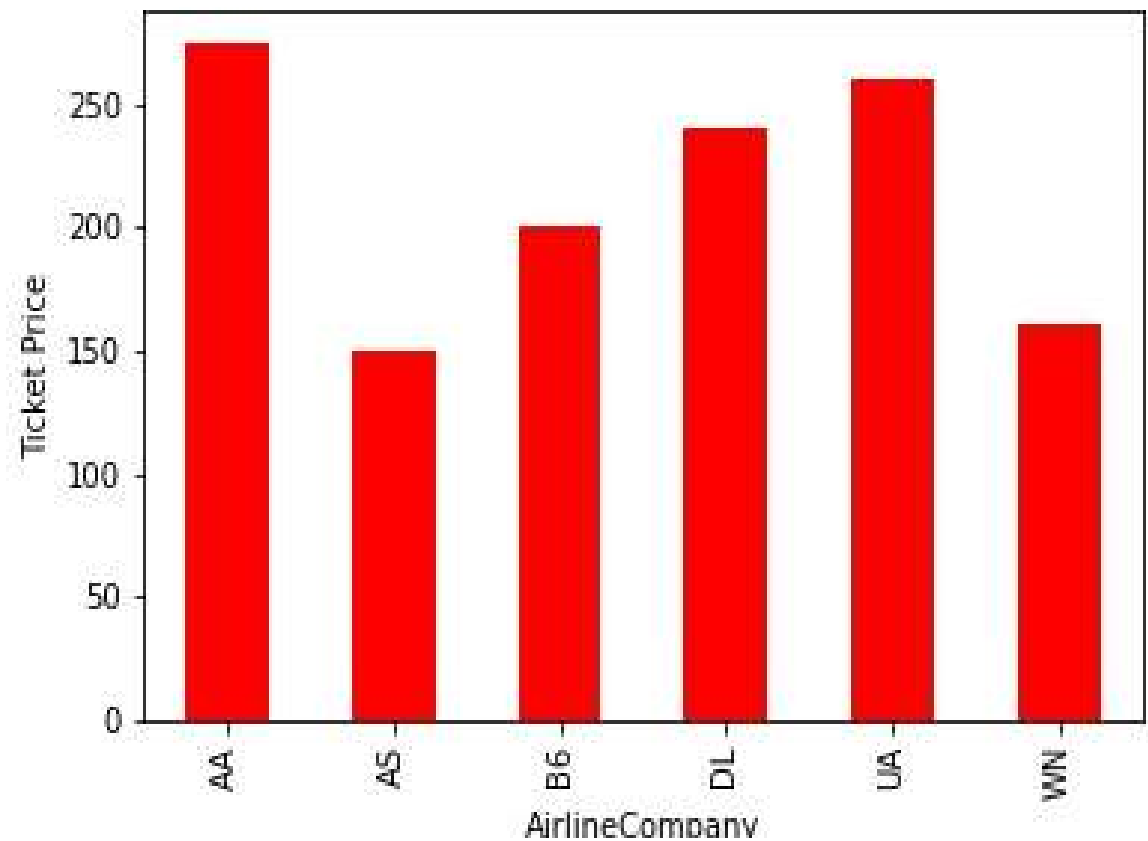
Slide 8

SB1

Sourav Bhattacharjee, 09-02-2021

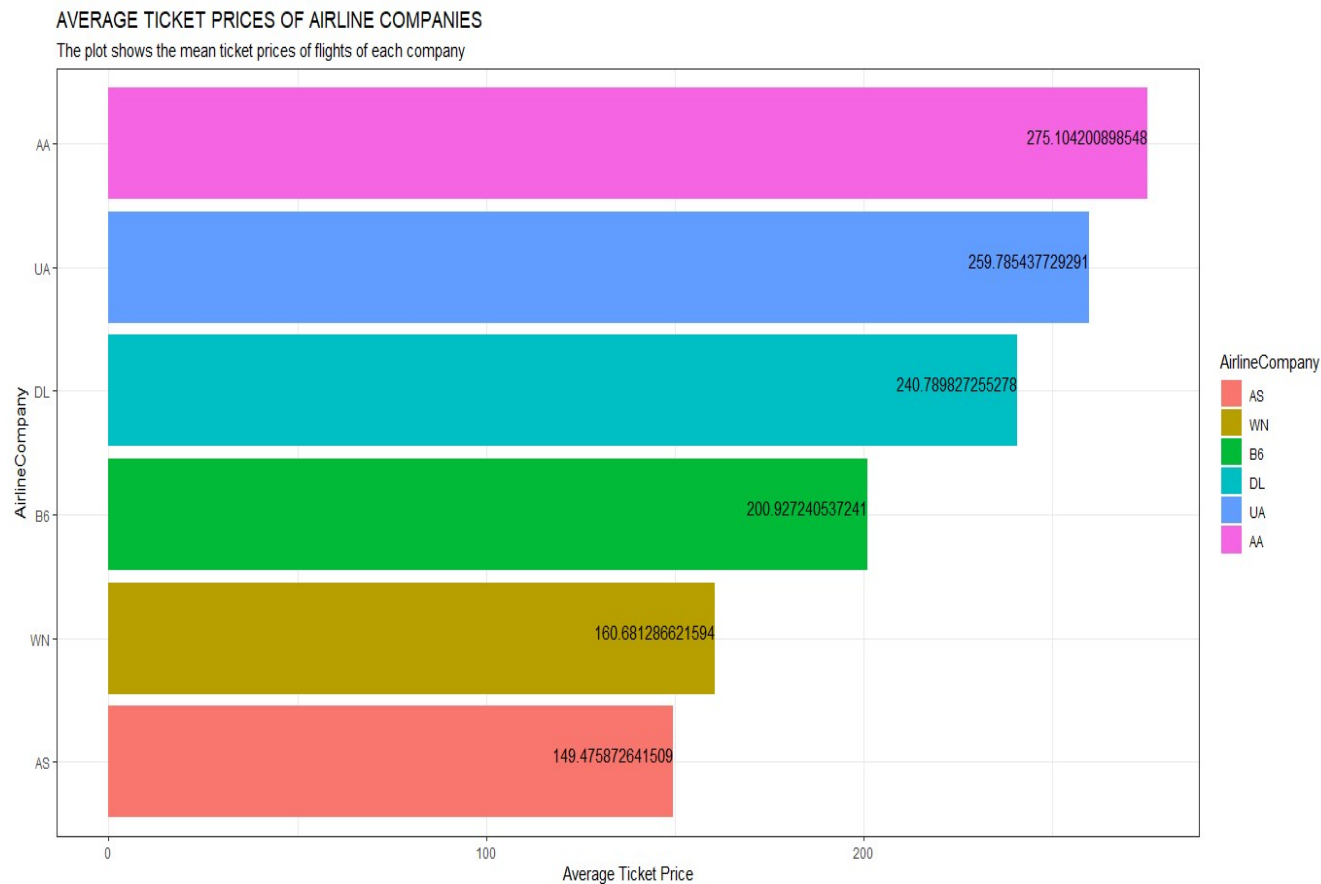
Average price ticket of every airlines company.

The barplot shows that the average price of the flight tickets are highest for AA airlines and lowest for AS airlines. So on an average traveling with AS is much cheaper compared to AA.



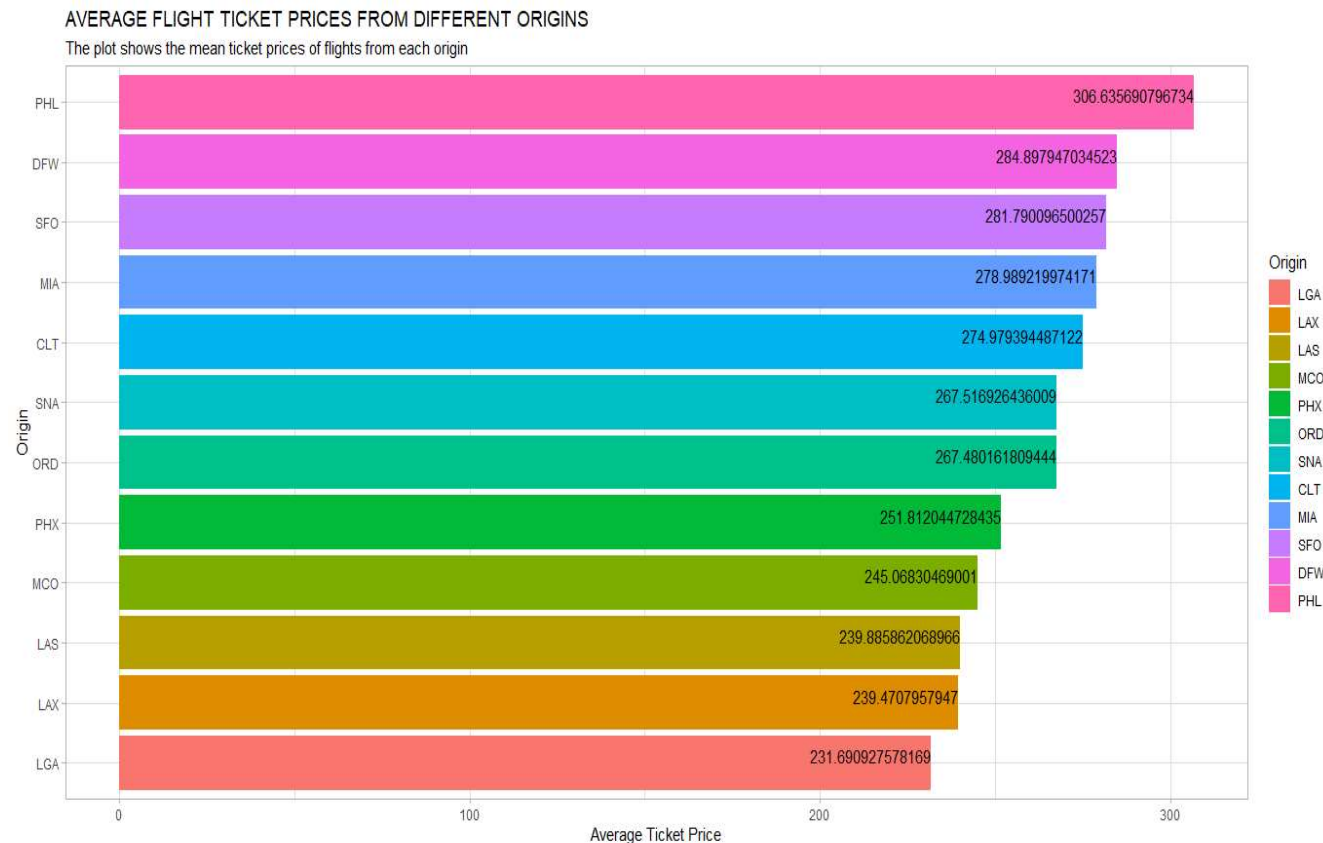
Which airline company is the most economical of all?

This plot shows that AS (Alaska Airlines Inc.) is the most economical flight to avail out of all. However, in a previous graph we have found out that AS is the least available flight. So, we come to know that even if AS is not available everywhere, it is best to board the flights of that company wherever it is available. Also, AA (American Airlines Inc.) is the most abundant of all, but it is also very expensive. So, boarders should choose flights of any other company if available for that particular origin – destination combination.



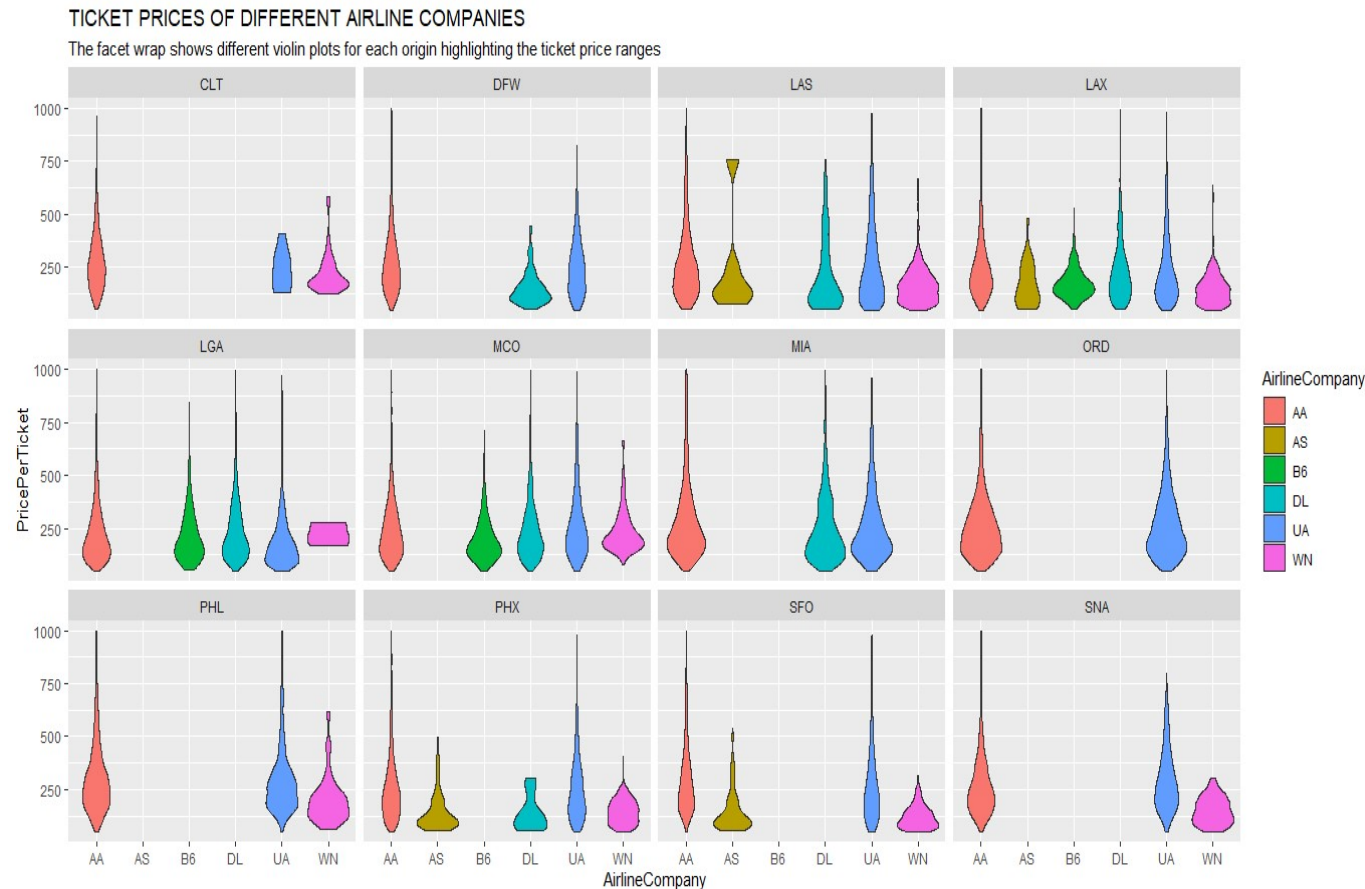
Which place gives the cheapest flight service?

This plot shows that from the origin LGA (New York), we will get the cheapest of all flights to some possible destinations (or maybe to all destinations). So, if flights from any particular origin seems to be quite expensive for someone, then they can travel to some nearby areas where the flight tickets are at a cheaper rate, provided they get their desired flight from that different origin. In other words, if someone is at a place which is near the top of the plot, then he can come to a place near to the bottom of the plot to avail his flight, provided he gets the flight for his desired destination.



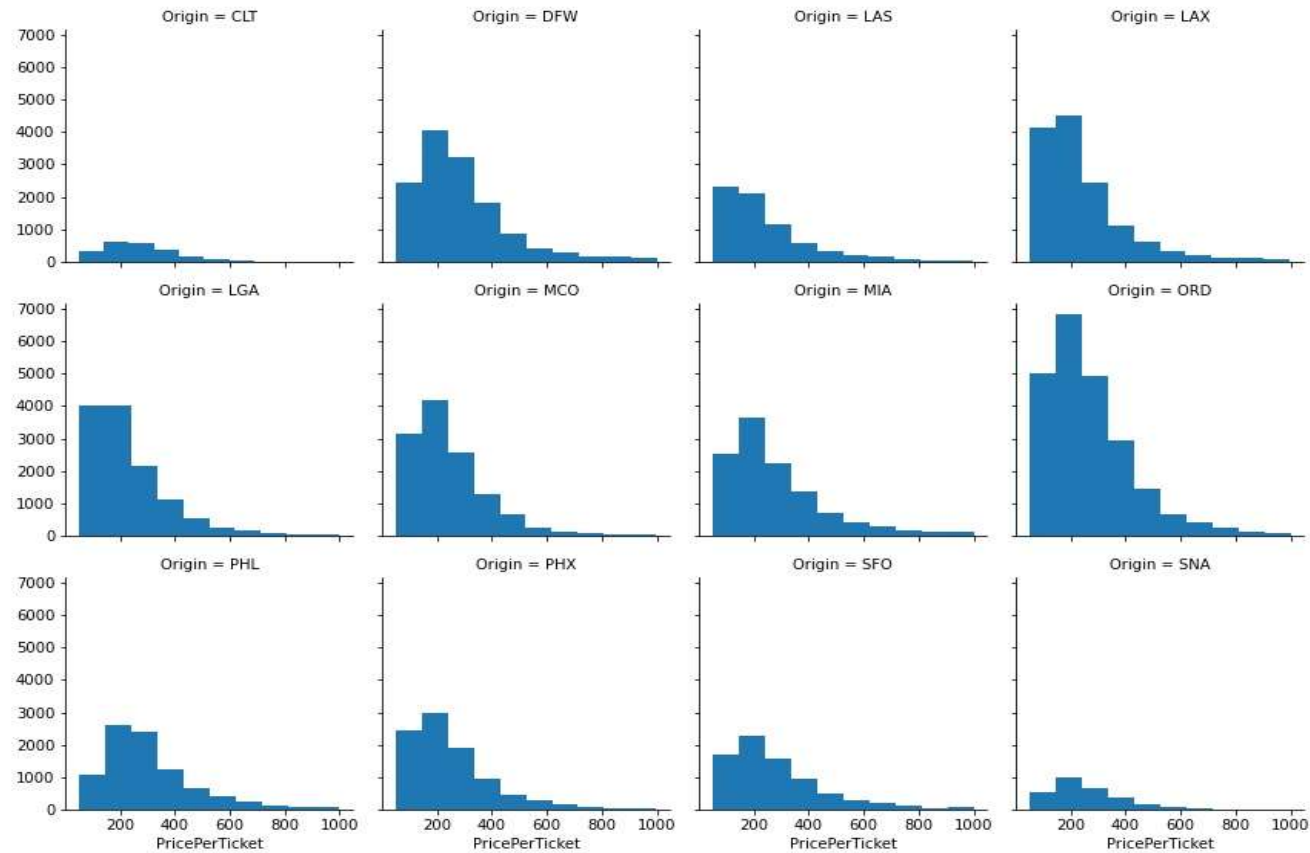
Find out the ticket price ranges of different flight companies from each origin.

The violin plots in each of the facet shows the availability as well the ticket price ranges of various flight companies for each origin. While we come to know that flights of each company fly from the origin LAX (California), ORD (Illinois) has just two options of flights to avail, even if maximum flights are available from ORD. The plot also shows the majority of the ticket prices is around 200 to 300, with few exceptions. The graph also confirms our previous prediction that flights of AS (Alaska Airlines Inc.) are less in number and are available only from few origins such as LAS, PHX, etc. However, the flat base of violin plots for AS clearly indicates its low-ticket price value.



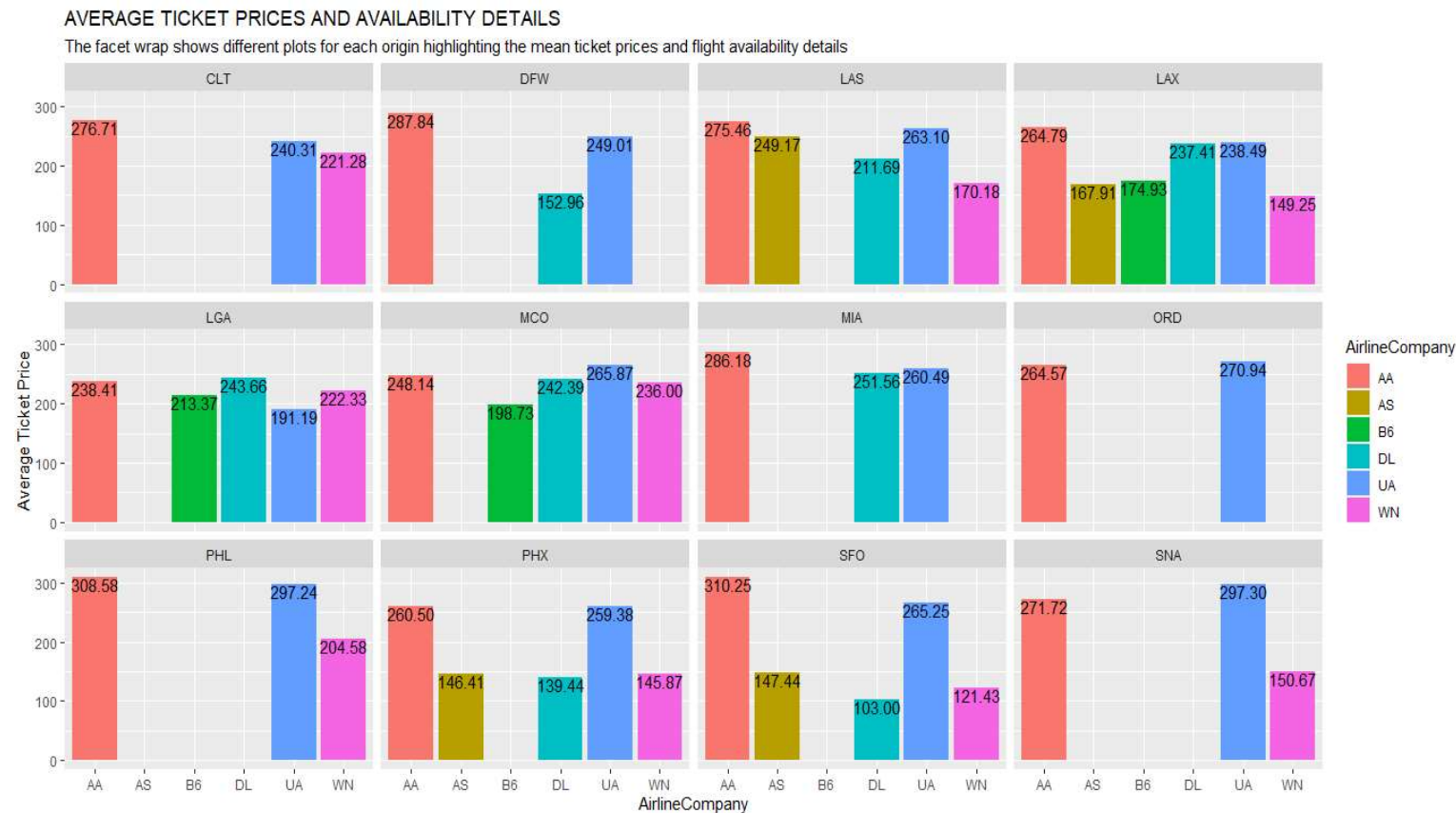
Origin wise ticket price variety.

The histograms here depicts the variation in the ticket prices of various locations. While ORD (Illinois) shows the highest variation, CLT (North Carolina) shows the least variation.



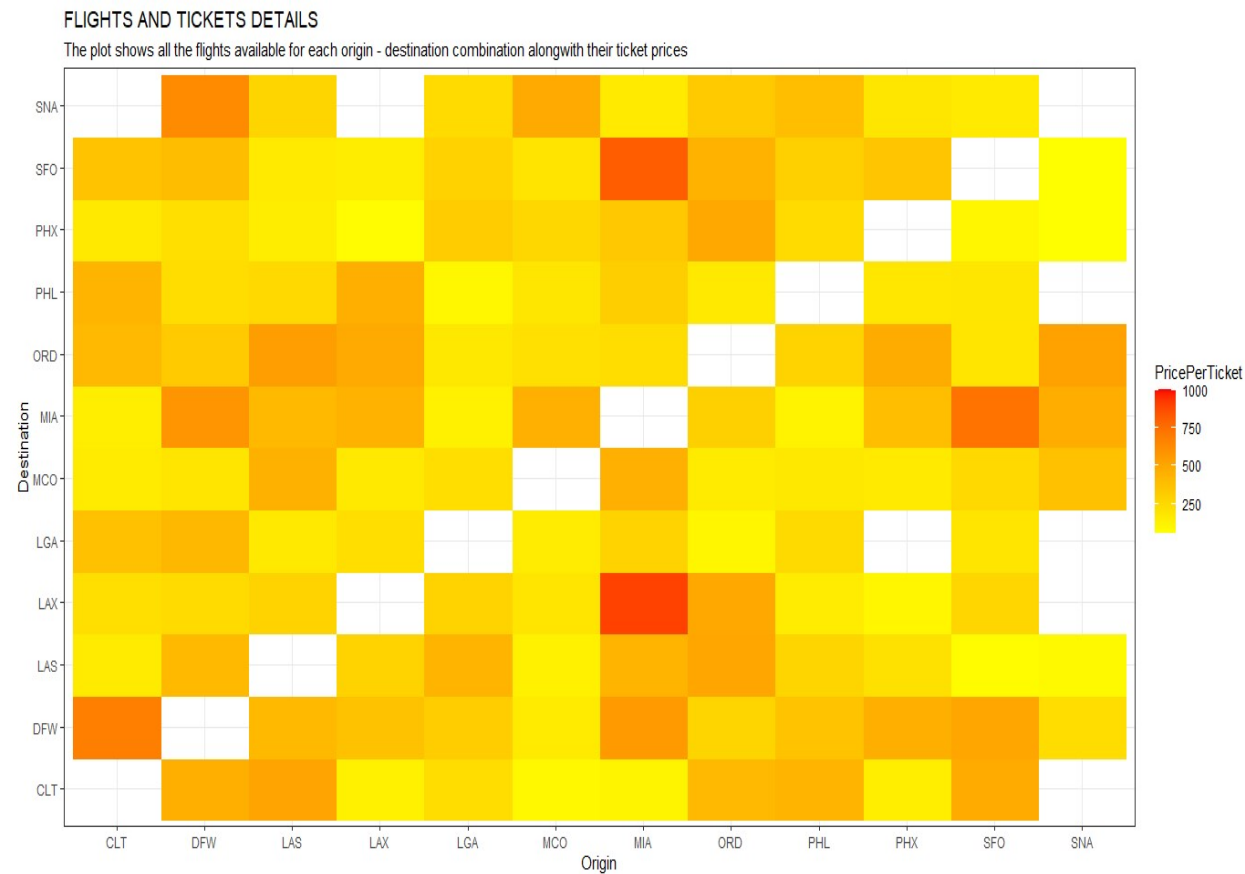
Find out the best flight options for someone for each origin.

Here, we find out the details of all the available flights from each origin as well as its mean ticket prices. Lower the ticket price, more economical it is for someone to avail it. This plot clearly confirms our prediction that the mean ticket price is around 200 to 300, with few exceptions.



Give the ticket price details for each origin-destination combination.

Here we perform a raster plot to get the ticket price details for each origin – destination combination. The right diagonal is blank as there are no flights for the same origin and destination. Few other boxes are also left blank indicating that there are no flights available for that origin – destination combination. Here, more reddish the color of the box, more is the ticket price for that particular origin – destination combination, or it is possible that more flights are available from that particular origin to that destination, thus making the sum of ticket prices high. Hence, this plot is useful as well as confusing at times, which we would definitely try to improve at times to come.



The most economical flight for each origin-destination combination.

With all the red lines, we understand that American Airlines is the most available flight service out here. LAX (California) has the most variety of flights available, with CLT and SNA having the least number of available flights. The ticket prices are fairly distributed, with few exceptions.



ACKNOWLEDGEMENT

In the end we would like to express our special thanks of gratitude to Mr Sudeep Mallick sir for giving us the opportunity to work on the visualization project. Developing the project has helped us to learn the languages immensely. Thanks for all the guidance and support without which this project would have been impossible.

Sawan Aich
Shashwata Mondal
Sourav Bhattacharjee