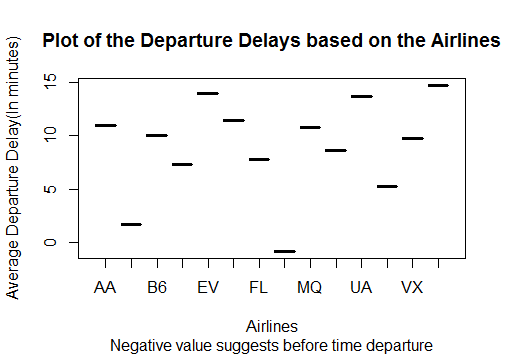
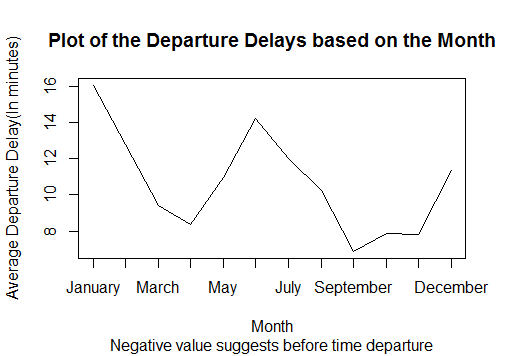
**Project Report**

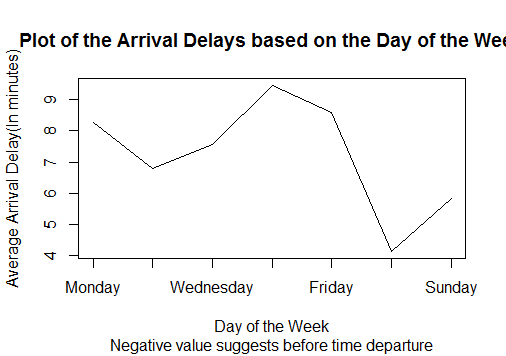
Air travel is increasingly becoming the most common mode of transport especially between large cities. With so many people dependent on air travel, delays are a very important factor. Being able to compare this delay based on the various airlines, origins, destinations, etc. can prove to be of great value. Hence the motivation to analyse and produce sensible results from data which can help people make better choices.

I have collected data for over a million flights for 14 major airlines during the year 2014 in the US. In order to be able to analyse the data, I have created various functions to perform this analysis. Each function takes into account various aspects of the data such as the airline, day of the week, day of the month, origin airport, destination airport, etc., and produces different results. The main algorithm that I used relies on concept of averages. The sum of the total delay for each airline for example is initially calculated. This sum is then then divided by the total number of records of data available for that airline. This gives the average delay for that airline. Also, the delay can be chosen to be departure delay or arrival delay.

The results that were obtained from the functions are then plotted to give a comparison between the results. To check if these results were correct, I had a look at the data manually and also performed individual checks for different sets of data.







Above are some of the plots that were obtained based on the analysis of the data by considering the different factors. Thus it can be inferred that Hawaiian Airlines has the least delay and Southwest Airlines has the most. Also, January has the most delay and September has the least. Finally, Thursday has the most delay and Saturday has the least.

Website for Data: <http://www.transtats.bts.gov/DL_SelectFields.asp?Table_ID=236>