Aviation Analytics

Team Members: Deepankit Shukla (1211213182), Dhruv Arora (1211076409), Swati Kanchan (1211213195)

1. Project Idea

As per the Bureau of Transportation Statistics, in 2008, there were 113.2 million air passengers who took a domestic US flight. This number soared in 2016; 895.5 million boarded a domestic flight. As evident, number of people travelling by an airplane are increasing every year. We chose this problem, because of the sheer number of people which can benefit from our analysis and make informed decisions.

In this project, we build an airline recommendation system which works on the principle of comparing various US flight carriers based on their on-time performance during a calendar year (Jan ‘11-Nov ‘15). We also make use of tweets, posted regarding airline companies, during the year 2015, and apply sentiment analysis. Datasets for both problems have been downloaded by us.

We plan to analyze: on-time arrival performance of airline carriers, correlation between delay (weather and security) and airport, airline preferred by people as per the tweets, finding the distribution of airline carriers in different regions, travel patterns during various months across the country by examining four years of data.

1. Data

Data is mainly taken from [Bureau of transportation statistics](http://www.transtats.bts.gov/OT_Delay/OT_DelayCause1.asp?pn=1) for flight information(Jan ‘11-Nov ‘15) and [Kaggle](https://www.kaggle.com/crowdflower/twitter-airline-sentiment) (there is a single dataset) for tweets. Below are **some** of the features which we plan to use:



1. Project Team

For workload distribution, please see the attached image.

1. Related Work

