CSX415 Project : Airline Delays and Cancellation Analysis

Ajey Patil, ajeypatil@gmail.com (mailto:ajeypatil@gmail.com)

04/15/2018

1 Objective

The Objective of this project is to analyze Flight Delay and Cancellations for US Airlines and determine best airline for least flight delays and cancellations.

2 Overview

When multiple airlines are available in metro area, then residents have choice of airline. An analysis of airlines performance in terms of flight delay and cancellation is required for passengers to choose an airline with least cancelations and delays.

3 Data Sources

The U.S. Department of Transportation's (DOT) Bureau of Transportation Statistics tracks the on-time performance of domestic flights operated by large air carriers. Summary information on the number of on-time, delayed, canceled, and diverted flights is published in DOT's monthly Air Travel Consumer Report and in this dataset of 2015 flight delays and cancellations. https://www.kaggle.com/usdot/flight-delays/data (https://www.kaggle.com/usdot/flight-delays/data)

4 Stake Holders

Airline Passengers: this analysis will help them choose airline depending on airport performance of flight delays and cancellations

Airport management: this report will help them determine systemic and process related issues affecting flight delays and cancellations

Airline management: this report will help them determine best performing airports in terms of flight delays and canceellations and their reasons in scheduling and flight route planning.

Regulatory boards: this analysis will assist in identifying consistenly low performing airlines in terms of flight cancellations and delays and their reasons

5 Deliverables

The deliverables for this project will include a report of all the findings, r code, all of which will be published on GitHub.

Specific questions to be explored

```
What is co-relation between airline and day of week of flight that got cancelled ?

What is co-relation between airline and day of week of flight that got delayed ?

What is co-relation between delay time and airline and day of week ?

Categorize the delay reason for flights for airline and numbers of delays per delay r
eason

Based on number of delays per delay reason and number of delays per day of week,
predict best airline for minimum flight delay and cancellations.
```

6 Success Criteria

Success criteria of project is ability of R-code to generate report for deliverables for data from different years of observations of data.

7 Risks

Availability of data in required format for different years of observations of data

8 Deployment Phases

Initial Analysis Phase: R-code for preliminary analysis of data

Initial Prediction Phase: R-code for preliminary prediction of deliverables

Final Model Deployment Phase: R-code to generate reports for yearly data

9 Costs

To be estimated