Pinal Mistry, Giancarlo Giudice, Evan Duffy

Carleton Bootcamp

Project 1

August 31, 2023

Project Proposal

We have all been in trapped in airport purgatory: not knowing when your flight will depart, if it will depart even, and being at the mercy of the announcements blaring in the heavens. What we are asking today is if the is a way for airlines and/or airports to better predict potential delays in flights and, if so, if there is a better way to deliver these updates to flyers than tv boards and announcements. For example, VIA rail trains have a feature that can update train passengers if there are delays as well as provide the new arrival time and they do so by email or by text message. We propose that by analyzing a data set of more than half a million domestic flights in the US last year that there is indeed a way to predict incoming delays to flights. As such, we will be working with airline departure data in the hopes of answering this question.

We are refining our guiding premise into the following questions:

1. What are the biggest factors in determining flight delays?
2. Is there a subset of the data (Aircraft age, aircraft type, time in air, departure/arrival airport, etc.) that provides a true correlation to flight delays?
3. Does the number of delays early on in the year cause airlines to receive less business in the latter part of the year?

As might have been surmised, our group will be working in the airlines industry to better answer this question, and we will be working with a dataset that provides flight information for domestic flights in the US over the whole year of 2022. We found our dataset on Kaggle and are confident that it is robust enough, with over half a million rows and 40 columns of data, to conduct a thorough analysis. The dataset can be found here: <https://www.kaggle.com/datasets/jl8771/2022-us-airlines-domestic-departure-data?select=CompleteData.csv>

Below is our task list and schedule:

|  |  |  |
| --- | --- | --- |
| Task | Due by | Date completed |
| Download data | 8/31/2023 | 8/31/31 |
| Create and Submit Proposal Doc | 8/31/2023 |  |
| Put into main branch | 8/31/2023 |  |
| Complete data cleanup (Giancarlo) | 9/3/2023 |  |
| Complete analysis of data | 9/5/2023 |  |
| Complete Visualizations (Pinal) | 9/7/2023 |  |
| Complete Analysis and Conclusions | 9/9/2023 |  |
| Prepare slide deck | 9/11/2023 |  |
| Upload files to github | 9/13/2023 |  |
| Submit Project | 9/14/2023 |  |