**TEAM 7 – UConn – Atlas Air Worldwide**

This document helps you to understand how to navigate and use various files. All ipynb files needs to be opened in <https://colab.research.google.com/?utm_source=scs-index> (Recommended) and another option is Jupyter Notebook.If you are planning to use Jupyter Notebook on your own computer, you should download all the datasets into your local machine and then read datasets by updating path directly from there.

**OPIM-5770\_Team-7ModelingFileCode.sql**

This file contains the SQL code related to the Modeling part. We have worked on logic to join various Schedule files, Delay data, Flight Data files and finally sampled the data. Details of each column and its description is present in the next Modeling file.

**OPIM\_5770\_Team-7\_Attrition\_Modeling.ipynb**

This file consists of the modeling part of the file with all code written and comments added for better understanding. This needs to be opened on Google Colab for better view of data and understanding.

**OPIM-5770 Team-7 Scraping tweets\_for\_AltasPilots.ipynb**

This file contains the code for scraping tweets from the official account @AltasAirPilots. We successfully scraped 2153 tweets which are used for further analysis in the next Google Colab file.

**OPIM-5770 Team-7 Text\_analysis.ipynb**

This file is to gain more information from the pilots’ aspects, which enables us to come up with reasonable and suitable recommendations to solve the business problem. We have collected three different text documents from different places. The First one is from the tweet account @AtlasAirPilot. The second one is from the engagement survey provided by the sponsor already. The last one is from the Glassdoor account @Altas Air Worldwide. This needs to be opened in the Google Colab environment.

**Note**: We recommend to place the Final Analytics folder in personal Google Drive and then change the drive path to access these files(we mark where you can update the path in the Colab files) . This is because of the challenges in accessing a folder dynamically in Google Colab.

**OPIM\_5770\_Team7 Data used for text analysis**

This folder contains three datasets that was used for text analysis:

* all\_tweets.csv–2153 tweets combined into csv file
* Glassdoor Reviews.xlsx–Manually collected from Glassdoor website
* Engagement Survey 2021 - Scheduling EXAMPLES.xlxs–Shared by Sponsor

**OPIM5770\_Team-7\_airports\_geolocation.sql**

This file contains latitude and longitude information and airport codes. This enabled us to get Domestic and International trips in Atlas.

**OPIM5770\_Team-7 Atlas Air Pilots Different Metrics PostgreSql.sql**

This file contains different scripts used to generate aggregated data which we used for data explorations in Tableau

**DataExploration\_MidTermPresentation.twbx**

This tableau file contains the explorations related to published vs Actual schedules which we prepared for midterm presentation

**DataExploration\_FinalPresentation.twbx**

This tableau file contains the explorations related to published vs Actual schedules based on Trip No, fleet, based on midterm feedback