**Data Cleaning & Data Exploration Coding**

**Note: This section is for the “2022\_Airline\_Data.ipynb” colab file**

Prerequisites:

* Combined\_Flights\_2022.csv
  + file can be found in github

Now that you have the csv file saved on your computer or drive and **imported panda, numpy, os** you can use the function:

“air= pd.read\_csv(‘’)”

**Note: you must name connect the file to ‘air’ variable for the rest of the code to work**

Now you can just run all the code cells from top to bottom, as long as you name the csv file air all the subsets and lines of code should run.

**Data Cleaning & Predictive Modelling**

**Note: This section is for the “Predictive\_Modelling\_2021\_2022.ipynb” colab file**

Prerequisites:

* final\_forecasting.csv
  + File can be found in github

The final\_forecasting.csv file was created by isolating the “flight date” and cleaned “cancelled” column and stacking the 2018, 2019, 2021, and 2022 flight data.