Data Clean Up

In order to have Michelle create great really good looking charts and graphs with Tableau, I had to make our data readable to everyone, not just people within the airline industry or our group.

To do this, I used Jupyter Notebooks, Pandas, and Microsoft Excel.

I loaded our original 2009 flight data into pandas data frame using JP. I dropped a few columns that have no needed value for Tableau, such as cancellation reason, or length of flight in miles and time.

Once the columns were dropped, i ran a for loop for determine all the airlines that we had data for.

We picked Skywest Airlines to do our analysis on. Once that was decided, I ran another loop to determine the origin and destination airports, which generated a list of all the airports that I needed. The airports were written in IATA codes, such as PHX or SFO.

After the list was generated, I used a code that would take the IATA code, match with the actual airport name, and create a new column with the airport name. The names of the airports were doing with the help of Google.

I did this process for both the origin and destination airports.

Once all this was done, I exported a CSV just to back up my data, and also backed it up to Github.

I opened the CSV in Excel and added two new columns, both for the IATA codes for the destination and origin in order to merge another CSV.

The other CSV file contained data on longitudes and latitudes for the airports. That CSV was cleaned in the same manner as our 2009 data. It was cleaned to just have the IATA code and the coordinates. This was done twice. Once for the origin and once for the destination.

Once everything was merged together into one data frame, it was exported again as a CSV, and opened in Excel for some final clean up.

It excel, I deleted some duplicate columns made within Pandas, converted military time into standard time, converted the dates (since they were in multiple formats) into one common format, and fixed some punctuation errors from adding the airport names in Pandas.

Everything everything was cleaned up and looked perfect, it was handed over to my team to create the dashboard portion.