**ETL Homework The BeardedGingers**

We started by looking at possibly doing a study between poverty in seniors whpo live in Chicago over a 5 to 10 year period to see if there were poverty clusters. Decided against it because it was too depressing.

Next we thought about a study of “Did you get the Holiday you Wanted?” We choose 6 main holidays: Thanksgiving, Christmas, New Years Eve, July 4th, Memorial and Labor Day weekends. We looked at when airlines forecast when travelers would travel compared with the number of flights cancelled during the time period for each holiday. We downloaded data from the FFA website for cancellations by holiday date and created a CVS sheet of dates but there were less than 350 rows of data so we thought that was to small of a dataset.

Next we looked at Kaggle again and decided to download data from kickstater and the world bank to see if there was any correlation between successful projects and the countries income per capita.

We first took the csv file from worldbank which was a 131 kb file and imported it into pandas to clean it up which resulted in a file that was only 24 kb. We took the kickstarter data which was about 50kb and imported it into pandas and cleaned up the data resulting in a file size of 24kb. The files were outputted and imported into postgres. We created the tables in postgres and imported the data into the tables. We then started writing queries to summarize the information we collected.

We had a problem turning our world bank data frame into just three columns so we used excel to rearrange the data into a table we could utilize for the project.

We chose postgres because we like sql better than Mongodb. We joined the tables on the Country field in each table and made a junction table so the data could be joined between the world bank and kickstarter data sets. We made an ERD to show the database relationships.

We created a number of queries to review the data and answer our question. And we discovered we couldn’t make a correlation between successful kickstarter campaigns and the world bank data. We learned we should have taken each countries population into consideration because it showed for example the United States had an average of 13,590 successful projects with an average capita of 46,499 compared to Luxemburg which has an average of 6.3 kickstarter projects and an average per capita 57,654 dollars.

In short, per capita has no relationship to the number of successful kickstarter projects.