**ETL Project Write Up:**

**EXTRACT**

For our ETL project, we decided to gather data on Beer Styles and Beer Reviews.

To do this, we gathered two data sets, one that was already in a CSV file from Kaggle and then we accessed the open API on brewerydb.com. The CSV file found on Kaggle contained the reviews of the different beer styles, in fact it contained over 1.5 millions of beer reviews. The review criteria were the date of review, overall review, aroma, appearance, profile name, palate, and taste. The open API on brewerydb.com was in JSON format and contained data concerning the beer names, the maximum and minimum international bitterness units (IBUs), the maximum and minimum alcohol by volume (ABV), the maximum and minimum standard reference method (SRM), the minimum original gravity (og), and the minimum and maximum final gravity (fg).

**TRANSFORM**

As we transformed our data the first thing we did was normalize the data. We focused on this because each data source had the beer styles, however, they did not necessarily match. It was important for us to match the beer styles from each data source with each other before loading any data, as that was one of the primary ways we were joining our tables. Our next step was building the tables that we were going to run queries on in PostgresSql. We made five different tables (Beer\_Styles, Beer\_Reviews, Brewery\_Table, User\_Table and Beer\_Table). We linked these tables with different Primary and Foreign Keys.

**LOAD**

As our group kept working on transforming the data, we also decided to perfect the loading of the data before we finished the transformation. For this, we created dummy data frames that matched the SQL tables we created and wen through the entire load process. This made it easier to load the correct data from our various data sources. After cleaning and transforming our data, we combined the code we used to transform the data and the dummy code we used to load the dummy data. Following that, all of our data was loaded into PostgresSql.