Umar Hassan & Abby Lemon

ETL – Project

March 7, 2020

Target & Starbucks (menu) / Location Data

**EXTRACT:**

Starbucks Data Source:

<https://www.kaggle.com/starbucks/store-locations>

This data was formatted as a CSV

**TRANSFORM:**

I loaded the data into Python and worked to identify columns of interest. We decided to keep to drop international locations to have some coordination with Target location data. Furthermore, we decided to drop coordinates column in the interest of preventing data cluster as we already have individual addresses of the different store locations. We also had to change the postcode data type to integer from text to avoid connection conflict with Postgres SQL.

We proceeded to create a SQL database with similar columns to stage it for connection to the Python file.

**LOAD:**

Upon the completion of the data cleaning we finalized our code and created the connection to Postgres SQL. We choose this method because it allowed us to create a database with the data of interest for further analysis.