ETL\_Project: what we did it.

Extract: SQL Query & Schemata

Our first task after acquiring a file called life expectancy.csv, was to join it to the other 2 existing csv files worldhappinessreport2021.csv and happinessalcoholconsumption.csv.

In pgAdmin4 we created our tables and then performed a 3-way join. We used country as the key, a column they all had in common. After joining we altered the query to only pull certain columns needed to pull the data to create our database in python.

Transform: File Clean up in Python & Creating Database for plotting

Once we had our table created we brought it into Python saving it as a Pandas dataframe.

Load:

Using existing code we created a scatter plot with our new data. Now seeing a relationship between Happiness Score and life expectancy. Saving the scatter plot to the images file.