**Extract**

The purpose of this analysis is determine what factors correlate to the overall happiness of a country. The two data sources used were both CSV files extracted from Kaggle: Countries of the World and World Happiness Report.

**Transform:**

The issues we faced during the transformation process in this analysis were the following:

1. There were more countries listed in the Countries of the World dataset than the World Happiness Report dataset.
2. Some columns from the datasets were ambiguous and their purpose could not be determined by simply looking at the datasets.
3. In the Countries of the World dataset, there were spaces after each value in the Country column so they would not join properly with the Country column in the World Happiness Report dataset
4. After fixing the issue in #3 the spaces were removed from the middle of the Country column values if the country name included more than one word. For example, “United States” became “UnitedStates”. This led to those values transforming into Null values in the table we created via joining.

To resolve the issues we ran we first created an ID column for both datasets to act as a Primary key which enabled us to properly join the datasets using the Country column. We then removed the ambiguous columns for simplicity since they were not necessary for our analysis. There turned out to be two spacing issues with the Country column in the Countries of the world dataset. We fixed the first issue by stripping off the space at the end of each value, and we fixed the second issue by inserting commas between each name of the values that consisted of multiple names.

**Load:**

We created a relational database using MySQL so we could save the new updated tables and transport them into new schemas with MySQL Workbench.