**Extraction:**

Use data.world or kaggle to get two different CSV files.

<https://www.kaggle.com/fernandol/countries-of-the-world#countries%20of%20the%20world.csv>

(countries population and GDP)

<https://data.world/ndsouza/alcohol-v-life-expectancy> (countries Alcohol v Life Expectancy)

**Transformation:**

Drop all columns from country\_population.csv except for “Country”, “Population”, “GDP”

Rename columns

Drop rows with null values

Join the csv files on their respective “Country” columns

**Loading:**

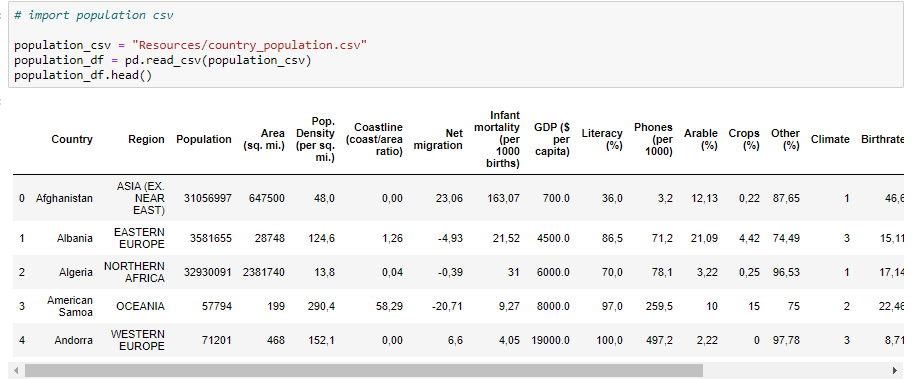
-Final database to load into will be Postgres using SQLalchemy ‘create engine’

-confirmed table names for ETL\_project db.

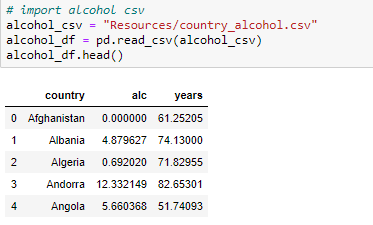
-Quick DBD to visualize the table.([https://app.quickdatabasediagrams.com](https://app.quickdatabasediagrams.com/#/))

**Finding Data: (kaggle and data.world)**

1. <https://www.kaggle.com/fernandol/countries-of-the-world#countries%20of%20the%20world.csv>

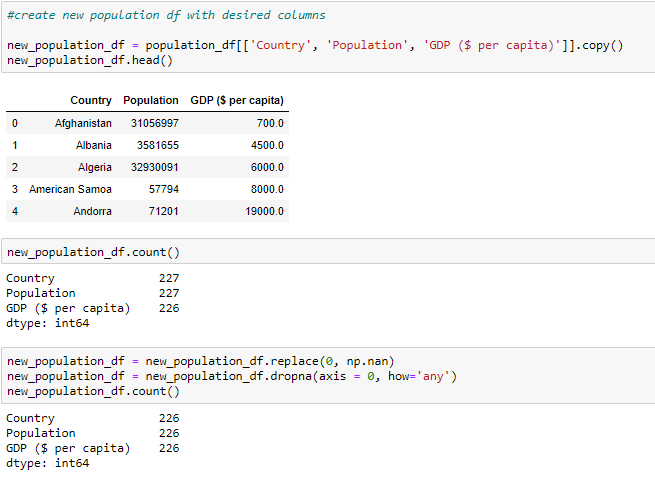


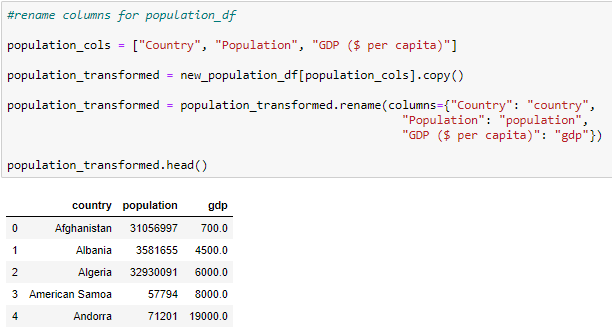
1. <https://data.world/ndsouza/alcohol-v-life-expectancy>



**Data Cleaning**

1. country\_population.csv
   1. From country\_population.csv, drop all columns from country\_population.csv except for “Country”, “Population”, “GDP”
   2. Drop value = 0
   3. Rename the columns

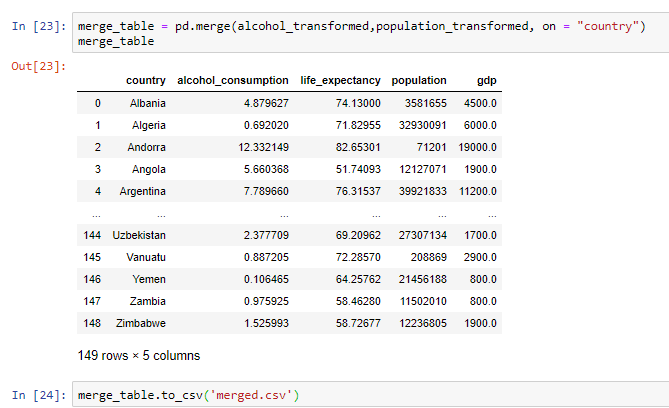




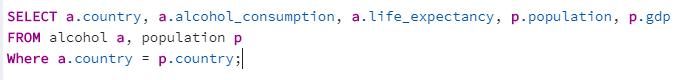
1. country\_alcohol.csv
   1. From country\_alcohol.csv, rename column
   2. Drop value = 0

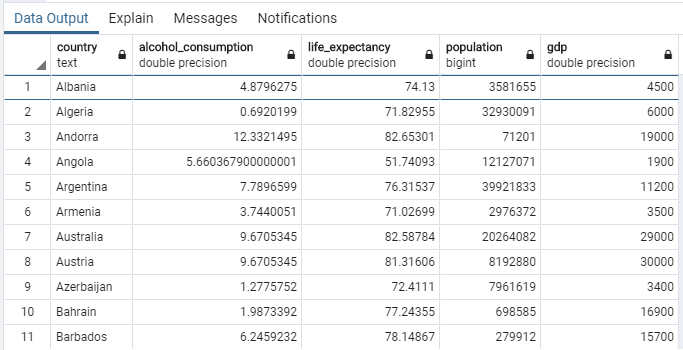


**Merging Data by using pandas**

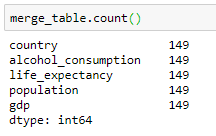
****

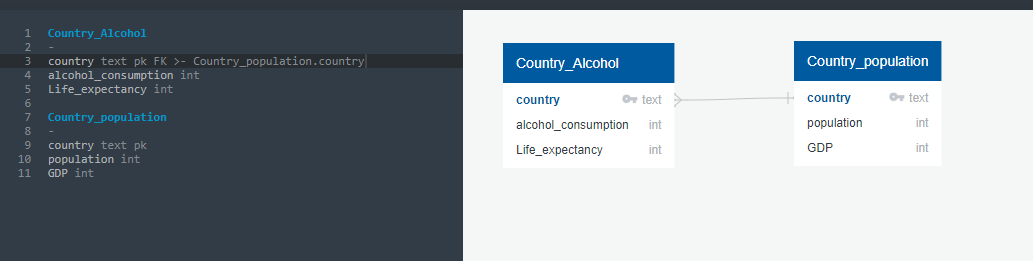
**Merging Data by using SQL**

****

****

**Only 149 counties left after Data cleaning and merging.**

****

**Quick DBD**