**Analysis Report**

**Extractions**

1. Download csv and use pandas to read in Jupyter notebook as a dataframe.
2. Use pandas to scrape table from the web to read into Jupyter notebook as a dataframe.

**Transformation**

1. Transformation in excel
   1. Drop columns
2. Transformation using pandas
   1. Dropping columns
   2. Renaming columns
   3. Resetting index
   4. Dropping zero and null values
   5. Changing shape of dataframe (sliced dataframe into separate tables and merged back together, similar to pivoting)
   6. Inner join of two tables from separate data sources to match country and region
   7. Groupby to transform data from country basis to region basis
   8. Remove rows from the supply utilisation dataframe that were deemed unnecessary based on the element info
3. Some of the daily prices in the csv contains commas, and as such, a comma separator needs to be noted when utilising read\_csc from panda (refer to code notebook attached separately)
4. the daily price dataframe is then updated with year code (extract year from Date column) and month code (extract month/year from Date column) to allow grouping for yearly price and monthly price analysis.

**Load**

1. Create schema for postgres using ERD
2. Establish connection to postgres database
3. Load dataframe from Jupyter notebook