Q1

You are provided with data with 10-mins, 60-mins and 1-day resolution (Merge.csv)

Please merge them into a pandas Dataframe with 2-hours resolution in between 7:00 – 17:00 only as index.

Please take the average of the 10-mins and 60-mins resolution prices during the 2-hours window and forward fill the 1-day resolution prices in the 2-hours window.

Q2

You are provided with a daily energy consumption data from 2016 to date (Consumption.csv).

Please create a Pandas DataFrame with to show the consumption of each year. The expected format is to have the year number as column name and mm-dd as index.

Please also create a seasonal plot showing 5-years (2016-2020) range (**shaded**) & average (**dashed line**), and year 2021 (**line**) & 2022 (**line**).

Please comment on your observation on the plot

Q3

You are asked to clean the number(s) in a string format. Currently, all numbers in our report are marked with thousands-separator and grouped together in a string separated by ONE white space.

Example 1: "1,247 4,568,123"

In addition, there are some whitespaces mistakenly added among the 3 digits after the thousands-separator.

Example 2: "1,2 47 4,56 8,1 23"

You are asked to write a function to remove the whitespace between the 3 digits after the thousands-separator but keep the whitespace separator (i.e. the whitespace between two numbers).

So, after the operation, the output of the above two examples should be:

Example 1: "1,247 4,568,123"

Example 2: "1,247 4,568,123"

Q4

You are provided with sample trades data in a data file (trades.csv). It contains a collection of trade data during 5 working days with volume and price. You are asked to process the sample data into a data set for creating the candlestick plots, with open, high, low & close prices (OHLC) and total volume for each product and contract, over a time interval.

You are expected to write a function to create a dataframe contains OHLC and trading volume with (begin, end, product(s), freq) as input. It should be able take freq input, e.g., 15MIN, 1H, 1D.

Expected Result – A python model file detailing process with comments:

Note:

- The product “Emission - Venue A” and “Emission - Venue B” are the same product trading in two different venues, please combine them when queried
- If product list has more than one contract, please generate OHLC and volume data for each contract
- Please limit output within trading hours, i.e., 7:00 – 17:00, except when freq >= 1D
- Please do not use OHLC resampler