Program Specification: Transformation Logic

**LLM:** GPT 5

**Context:**

I am an AI Prompt Engineer who is doing contract work for a company called Agentic AI Services. One of our customers is an energy brokerage company called Kilowatt Partners Inc. They act as a ‘middle man’ between Retail Energy Providers (abbreviated as REPs) and management companies. Each management company may have multiple accounts and each account may have multiple meters to measure energy consumption. Each meter is identified with an ESID and each account is currently identified with one manager.

The management company assigns managers to accounts. Accounts can be registered to bodies like Neighborhoods, HOA Boards, or Housing Development. Currently there is one manager per account. This is creating a problem that when the manager leaves or transitions to another account, the associated account is no longer uniquely identifiable.

Currently all of these relationships are managed by excel spreadsheets and Microsoft access. The goal is migrate off of access and move over to using a database schema so that the backend can utitlize supabase. The front end of the web application will be built using a React framework with Vite as a building tool.

**Modified**:

I am an AI Prompt Engineer who will be translating input excel spreadsheets from one format to another. Here are the requirements for the transformation.

1. Read in the input excel file. We will produce a destination transformation dataframe. The dataframe will have 17 columns.
2. Column A will be a counter that increments by 1 based on the previous row’s counter value.
3. Column B will be today’s date.
4. Column C will be the start date. This value is gathered from Column J of the input file.
5. Column D will be the zone – the value for zone will be the word that proceeds the word zone in column E of the input file. This value will either be NORTH, SOUTH, WEST, If Houston then COAST, otherwise NA.
6. Column E will be the load – that value will be HIGH, MED, LOW or NA. This information will be extracted from column E of the input file.
7. Column F will be the Retail Energy Provider (i.e. REP1). For the current input file this value will always be HUDSON.
8. Column G will be the Term. This value will be extracted from Column D of the input file. The only values we are concerned about are 12, 24, 36, 48, and 60.
9. Column H will be the minimum megawatt hours (abbreviated as Min\_MWh). This value will always be hard coded at zero for the current input file.
10. Column I will be the maximum megawatt hours (abbreviated as Max\_MWh). This value will always be hard coded at 1000 for the current input file.
11. Column J will be the Daily\_No\_Ruc. This value will be 1000\* Column H of the input file.
12. Column K will always be $0.00
13. Column L will be the sum of Columns J and Column K
14. Column M will always be $0.00
15. Column N will always be $0.00
16. Column O will always be $0.00
17. Column P will always be $0.00
18. Column Q will always be 5

Read in the Master Data file named “Master-Table” from directory “C:\Users\lcthu\Downloads\kilwatt-data-automation\2-copy-reformat”. Column A is the counter, our dataframe should start at 1 + the last entry in this master table file. Renumber the first column in the data frame starting with last number in the master file. Append the dataframe to the master file and write the data file out to the destination directory and name the new file master-file-updated. Do not modify the master file in the destination directory.

Please take the excel reader.py file and add the ability to download the master table file from the sharepoint similar to how excel processor.py does this operation

If the file to be downloaded already exists in the destination file, rename the file in the directory by appending the current date and time to the file name. Then download the file from the sharepoint.

Given the current background, please create a new python script that will download the master table file called "DAILY PRICING - new" and the input file called "HudsonMatrixPrices08272025020701PM.xlsm" from sharepoint as before. However, this time, please download the files into new\_files directory in the root of this project.

Yes, please chain the next step to transform the Hudson file and write the updated master copy into a separate output (without modifying the original).

No need to verify the master-file-updated. For next steps please add the ability to upload the master-file-updated to the original sharepoint link.

Have excel\_reader.py read in the input excel file and unhide the tab named 'Matrix Table' and process only that tab.

Template Formula Translations for Master Table Columns from Input File

Context: In between the input file that

**Column D** (Master): Zone Formula and Logic

Formula

=IF(C3="North","NORTH",IF(C3="West","WEST",IF(C3="South","SOUTH",IF(C3="Houston","COAST","NA"))))

Logic

if e2 == "North":

result = "NORTH"

elif e2 == "West":

result = "WEST"

elif e2 == "South":

result = "SOUTH"

elif e2 == "Houston":

result = "COAST"

else:

result = "NA"

Python Dictionary

mapping = {

"North": "NORTH",

"West": "WEST",

"South": "SOUTH",

"Houston": "COAST"

}

result = mapping.get(c2, "NA")

**Column G**: Term code formula translation =IF(O2=1,6,IF(O2=2,12,IF(O2=3,18,IF(O2=4,24,IF(O2=5,30,IF(O2=6,36,IF(O2=7,48,IF(O2=8,60))))))))

**Column J**: Daily\_No\_Ruc formula translation

=IFNA(S3\*R3,0)

In the formula above, column R is always hard coded at 10.

The formula for column S is below and it has a custom format:

=IF(Q3=6,E3,IF(Q3=12,F3,IF(Q3=18,G3,IF(Q3=24,H3,IF(Q3=30,I3,IF(Q3=36,J3,IF(Q3=48,K3,L3)))))))