Programming Scope for New Programmers

Notes

- Use Python with pandas and openpyxl for Excel handling
- Handle missing data carefully
- Ensure no duplicates remain after merging
- Validate file formats and column names before processing

Assignment Scope: Data Merge and Validation for Substation Components

Objective

You are given:

- Two .csv files of two different substations (SUB1, SUB2)
- One .xlsx files of the list of components used in this TLS (Transmission Line Segment) (e.g. SUB1-SUB2)

Your task is to merge, validate, and update data based on this information.

1. **Load CSV Files**
   a. Load both station .csv files separately

2. **Merge the CSV Files**
   a. Combine the two data frames into one.
   b. Save the merged file as SUB1-SUB2 xxkV.csv

3. **Cross-reference with TLS (.xlsx) File**
   a. TLS components.
   b. Match components in the combined .csv file with those in the TLS using:
      i. Primary match: OID column
      ii. If OID is missing or does not match, use:
         1. Substation Name
         2. Component Description
         3. Additional Info (if applicable)

4. **Highlight Mismatches**

    a. Identify components in the combined file that do not match the TLS file.

    b. Add a new column called Mismatch:

        i. Mark as Yes if there's a mismatch

        ii. Mark as No if it's valid

5. **Update Combined File**

    a. Add column "Type of Change" into the new combined excel file.

    b. For any mismatched entries, update the incorrect or missing fields using the TLS excel data only if the component exists in the TLS excel file

    c. Save this new file as: Sub1-Sub2 xxkV_updated.csv

6. **Generate Summary Report**

    a. Create a .csv file listing all changes made, including:

        i. OID

        ii. Column(s) updated

        iii. Old Value

        iv. New Value

    b. Save this file as: SUB1-SUB2 xxkV_summary_report.csv

Final Submissions

1. SUB1-SUB2 xxkV.csv - Original combined data

2. SUB1-SUB2 xxkV_highlighted.csv - Combined file with mismatch column

3. SUB1-SUB2 xxkV_updated.csv - Final updated file with TLS corrections

4. SUB1-SUB2 xxkV_summary_report.csv - Changes made

Put all these files into the Final Folder