**Documentation 1**

**Partner Sustenance 1(source Partner side)**

**Documentation for Data Processing Module**

**Overview**

This module provides functionality to clean, filter, and process two pandas DataFrames based on a specified threshold for the 'gsv' column. It includes functions to load CSV files, filter data, compare results, and merge duplicate entries. The results are saved to a CSV file.

**Dependencies**

* **pandas**: This module requires the pandas library for data manipulation and analysis.

**Functions**

1. **clean\_and\_load\_csv(file\_path)**

Reads a CSV file, strips whitespace from column names, and returns the DataFrame.

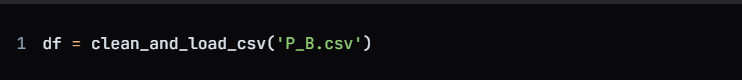
**Parameters:**

* **file\_path** (**str**): The path to the CSV file to be loaded.

**Returns:**

* **pd.DataFrame**: A cleaned DataFrame with stripped column names.

**Example:**



1. **filter\_gsv(df1, df2, threshold=10000)**

Filters two DataFrames based on the 'gsv' column.

**Parameters:**

* **df1** (**pd.DataFrame**): The first DataFrame to be filtered.
* **df2** (**pd.DataFrame**): The second DataFrame to be filtered.
* **threshold** (**int**, optional): The threshold value for filtering the 'gsv' column. Default is 10,000.

**Raises:**

* **KeyError**: If either DataFrame does not contain a 'gsv' column.

**Returns:**

* **tuple**: A tuple containing the filtered DataFrames.

**Example:**



1. **filter\_and\_display(df5, source\_partner)**

Filters **df5** to get the top 10 rows for each source country based on 'gsv'.

**Parameters:**

* **df5** (**pd.DataFrame**): The DataFrame to be filtered.
* **source\_partner** (**str**): The source partner to be added to the DataFrame.

**Raises:**

* **KeyError**: If required columns do not exist in **df5**.

**Returns:**

* **pd.DataFrame**: The top rows for each source country.

**Example:**



1. **compare\_and\_filter(top\_rows, filtered\_df1)**

Compares **top\_rows** with **filtered\_df1** and keeps only matching rows.

**Parameters:**

* **top\_rows** (**pd.DataFrame**): The DataFrame containing top rows to be compared.
* **filtered\_df1** (**pd.DataFrame**): The DataFrame to compare against.

**Raises:**

* **KeyError**: If required columns do not exist in **filtered\_df1**.

**Returns:**

* **pd.DataFrame**: The filtered top rows.

**Example:**



1. **merge\_duplicate\_rows(final\_results)**

Merges duplicate rows based on **source\_partner**, **source\_country**, and **destination\_country**.

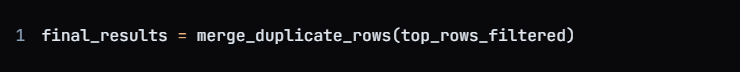
**Parameters:**

* **final\_results** (**pd.DataFrame**): The DataFrame containing results to be merged.

**Returns:**

* **pd.DataFrame**: The merged results with summed 'gsv' values.

**Example:**



1. **add\_regional\_director(merged\_results, r\_d\_file\_path)**

Adds a 'regional\_director' column to **merged\_results** based on matches in **R\_D.csv**.

**Parameters:**

* **merged\_results** (**pd.DataFrame**): The DataFrame to which the regional director will be added.
* **r\_d\_file\_path** (**str**): The path to the **R\_D.csv** file.

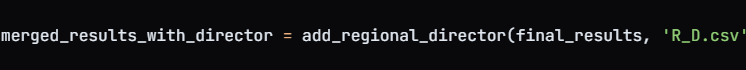
**Raises:**

* **KeyError**: If required columns do not exist in **R\_D.csv**.

**Returns:**

* **pd.DataFrame**: The merged DataFrame with the regional director information.

**Example:**



**Main Execution Block**

The main execution block of the module performs the following steps:

1. Cleans and loads the first DataFrame.
2. Cleans and loads the second DataFrame.
3. Filters both DataFrames based on the 'gsv' column.
4. Iterates through a list of source partners to filter, compare, and merge results.
5. Saves the results to a CSV file.

**Example Usage:**

A computer screen shot of a program code

Description automatically generated

**Notes**

* Ensure that the CSV files used in the module are correctly formatted and contain the necessary columns.
* Ensure that the source\_partners is filled manually with all the partner names.
* The module assumes that the 'gsv' column contains numeric values or values that can be coerced to numeric.
* The results are saved to a CSV file named **Mer.csv**, which can be modified as needed.