# **Sens Automation Script**

## **Overview**

This script processes microbiome data, assigns tags, and generates summaries based on predefined conditions. It merges multiple data sources, performs calculations, and produces output files with meaningful insights.

## **Requirements**

* **Python version**: 3.6 or later
* **Dependencies**:
  + pandas
  + os
  + sys

Install missing dependencies using:

pip install pandas

## **Usage**

Run the script using the following command:

python3 Sens\_automation.py <IOM\_file> <Reference\_file.csv> <Text\_for\_tags.csv> <Tags\_table.csv>

### **Arguments:**

1. **<IOM\_file>**: CSV file containing microbiome data with a column starting with "IOM".
2. **<Reference\_file.csv>**: CSV file with reference data.
3. **<Text\_for\_tags.csv>**: CSV file mapping tags to text descriptions.
4. **<Tags\_table.csv>**: CSV file mapping tags to labels.

## **Input File Structure**

### **IOM File**

* Must contain an "IOM" column (e.g., IOM025FX210924).
* Includes rows of microbiome data for analysis.

### **Reference File**

* Must contain the following columns:
  + #OTU ID
  + Bacteria Tag 1 in Report
  + Positive Impact

### **Text for Tags File**

* Maps tags to corresponding text descriptions.
* Should include columns List Of Tags and tag-specific mappings (e.g., Tag1, Tag2, etc.).

### **Tags Table File**

* Maps tags to specific labels.
* Should include columns Label and tag-specific mappings (e.g., Tag1, Tag2, etc.).

## **Output Files**

1. **Intermediate File**:
   * Name: <IOM\_file>\_SENS.csv
   * Contains merged data between the IOM file and the reference file.
2. **Final Output File**:
   * Name: <IOM\_file>\_SENS\_Final\_Output.csv
   * Contains a pivot table with assigned tags, levels, and text descriptions.

## **Script Workflow**

### **1. Input Validation**

* Checks for the required number of command-line arguments.
* Confirms the presence of a column starting with "IOM" in the input file.

### **2. Data Processing**

#### **Top 20 Selection:**

* Sorts the IOM file by the "IOM" column in descending order.
* Selects the top 20 rows for further processing.

#### **Merging Data:**

* Merges the top 20 rows with the reference file using the #OTU ID column.

#### **Pivot Table Creation:**

* Groups the merged data by Labels.
* Calculates the sum of the "IOM" column for each label.

### **3. Tag Assignment**

#### **Functions Used:**

1. **find\_tag(score)**:  
   * Determines the tag and level based on the input score.
   * **Logic**:
     + Score > 20: Tag1, High
     + 10 < Score <= 20: Tag2, Medium
     + 0 < Score <= 10: Tag3, Low
     + Score <= 0: Absent, Absent
2. **assign\_tags(df, tags\_tab, text\_tags\_db)**:  
   * Assigns tags, levels, and text descriptions based on label matching with the tags and text files.

### **4. Output Generation**

* Saves intermediate and final results as CSV files in the specified output directory.

## **File Paths**

* **Base Directory**: /Users/rajathtalpady/Desktop/Iom\_Bioworks/Sens\_Automation
  + **Input Directory**: input
  + **Output Directory**: output

The script automatically creates the output directory if it does not exist.

## **Example Command**

python3 Sens\_automation.py IOM025FX210924.csv Reference.csv Text\_for\_tags.csv Tags\_table\_sens.csv

## **Error Handling**

* If no column starting with "IOM" is found, the script exits with an error message.
* Input validation ensures the required files and columns are present.

## **Future Improvements**

* Add support for customizable score thresholds and tag logic.
* Implement logging for better debugging.
* Add visualization of the pivot table output.

## **Contact**

For issues or questions, please contact: **Rajath Talpady**

* Email: rajathtalpady@example.com
* GitHub: [github.com/rajathtalpady](https://github.com/rajathtalpady)