**Part 5- Integration document**

**1. Importing Data into Excel**

To import data into Excel from a database or other data sources, follow these steps:

Open Excel:

Launch Microsoft Excel and open a new workbook.

Import Data from a File:

Click on the Data tab in the Excel ribbon.

Select Get Data > From File > From Workbook.

Browse to the location where the data file (e.g., CleanWaterAndSanitation.xlsx) is stored.

Select the file and click Import.

Select the Sheets to Import:

A Navigator window will appear, showing all the available sheets in the Excel file.

Select the sheets you need (e.g., Communities, Water\_Sources, Water\_Quality\_Tests, Projects, Project\_Funding, and Beneficiaries).

Click Load to import the data into separate sheets in your workbook.

Verify Data Import:

Once the data is loaded, verify that all tables are imported correctly by checking each sheet.

Ensure that all columns are correctly aligned and that there is no missing or misaligned data.

**2. Ensuring Data Consistency**

Maintaining data consistency is crucial for accurate analysis and reporting. Here are the steps to ensure consistency in your Excel workbook:

Check Data Types:

Make sure that each column in your data tables has the correct data type (e.g., numbers, dates, text).

Format columns as needed (e.g., format dates as Date, numbers as Number, and text as Text).

Remove Duplicates:

For each table, ensure there are no duplicate records that could skew the analysis.

Use Excel's Remove Duplicates feature by selecting the relevant columns and clicking Data > Remove Duplicates.

Validate Data Relationships:

Ensure that the relationships between tables are consistent. For example, each Community\_ID in the Projects table should match a valid Community\_ID in the Communities table.

Cross-reference related data points across tables (e.g., ensure that the Source\_ID in Water\_Quality\_Tests corresponds to an existing Source\_ID in Water\_Sources).

Data Cleaning:

Remove any unnecessary blank rows or columns.

Address any missing values by filling them in or using appropriate data imputation techniques.

Create Named Ranges (Optional):

For ease of reference, you can create named ranges for key data tables. Select a table and use the Formulas tab to define a named range.

Set Up Data Validation Rules (Optional):

If there are specific criteria that data must meet (e.g., Potability Status should only be "Safe" or "Unsafe"), set up data validation rules to ensure that future data entries meet these criteria.