**Documentation: Importing and Visualizing Data in Excel**

**Objective:** This documentation outlines the process I used to import data from a SQL database into Excel, ensure data consistency, and visualize the data using PivotTables and charts. The data includes patient records, hospital information, treatments, diseases, and regions.

**1. Preparation**

Before importing data, I ensured a clear understanding of the database structure and relationships between tables. The data includes:

**- Patients**

**- Hospitals**

**- Treatments**

**- Diseases**

**- Regions**

**- PatientDiseases**

**2. Export Data from SQL Database**

**A. Export Queries**

I executed the following SQL queries to export data:

**1. Patients Table**

SELECT \* FROM Patients;

**2. Hospitals Table**

SELECT \* FROM Hospitals;

**3. Treatments Table**

SELECT \* FROM Treatments;

**4. Diseases Table**

SELECT \* FROM Diseases;

**5. Regions Table**

SELECT \* FROM Regions;

**6. PatientDiseases Table**

SELECT \* FROM PatientDiseases;

**B. Export Results to CSV**

- **Using MySQL Workbench:**

1. I ran each SQL query.

2. I right-clicked on the result grid and selected **Export Resultset** > **Export as CSV**.

3. I saved each CSV file with a descriptive name (e.g., `Patients.csv`, `Hospitals.csv`).

- **Using SQL Server Management Studio:**

1. I ran each SQL query.

2. I right-clicked on the result grid and selected **Save Results As** > **CSV**.

3. I saved each CSV file with a descriptive name.

**3. Import Data into Excel**

**A. Open Excel and Create a New Workbook**

**B. Import CSV Files**

**1. Patients Table**

1. I went to the **Data** tab.

2. I clicked **Get Data** > **From Text/CSV**.

3. I selected `Patients.csv`.

4. I clicked \*\***Import**\*\* and then \*\***Load**\*\*.

5. I renamed the sheet to **Patients**.

**2. Repeat for Each CSV File**

- I imported `Hospitals.csv`, `Treatments.csv`, `Diseases.csv`, `Regions.csv`, and `PatientDiseases.csv` following the same steps.

- I renamed each sheet accordingly.

**4. Ensure Data Consistency**

**A. Verify Data Types**

- **Patients Table**:

- I ensured `PatientID` is numeric, `Name` is text, `Age` is numeric, `Gender` is text, and `RegionID` is numeric.

- **Hospitals Table:**

- I ensured `HospitalID` is numeric, `HospitalName` is text, and `RegionID` is numeric.

B. Check for Missing or Inconsistent Data

- I used **Conditional Formatting** to highlight any cells with missing or inconsistent data.

1. I selected the entire sheet.

2. I went to **Hom**e > **Conditional Formatting** > **Highlight Cell Rules** > **Blanks**.

3. I reviewed and corrected any inconsistencies.

**C. Ensure Referential Integrity**

- **Example Check for `PatientDiseases`:**

1. I ensured `PatientID` values exist in the `Patients` table.

2. I ensured `DiseaseID` values exist in the `Diseases` table.

3. I used **VLOOKUP** to verify consistency.

**D. Check for Duplicate Entries**

1. **Patients Table**:

- I went to the **Data** tab.

- I clicked **Remove Duplicates** and ensured no duplicate `PatientID` values.

2. **Repeat for Other Table** as needed.

**5. Consolidate Data if Necessary**

**A. Create a Consolidation Sheet**

1. I created a new sheet named **ConsolidatedData**.

2. I used **Data Consolidation**:

1. I went to **Data** > **Consolidate**.

2. I chose the consolidation function (e.g., **Sum**, **Average**).

3. I added ranges from each sheet and clicked **OK**.

**B. Prepare for Analysis**

- I ensured the consolidated data was structured consistently with clear headers.

**6. Create Visualizations**

**A. Generate PivotTables**

1**. Patients Analysis**

1. I selected the **Patients** sheet.

2. I went to **Insert > PivotTable**.

3. I dragged `RegionID` to Rows and `PatientID` to Values to count patients per region.

4. I placed the PivotTable on a new sheet named **Data**.

**2. Repeat for Other Sheets** to create PivotTables as needed (e.g., for hospitals, treatments).

**B. Create Charts**

**1. Disease Distribution**

1. I selected the PivotTable from **PatientDiseases.**

2. I went to **Insert > Pie Chart.**

3. I customized the chart with appropriate labels and titles.

**2. Treatment Analysis**

1. I selected the PivotTable from **Treatments.**

2. I went to **Insert > Bar Chart.**

3. I customized the chart with titles and data labels.

**7. Documenting the Process**

**A. Create a Step-by-Step Guide**

1. I documented each step in the import and visualization process.

2. I included screenshots or examples for clarity.

**Maintain a Log of Issues and Resolutions**

1. I recorded any issues encountered, such as data type mismatches or missing values.

2. I documented how each issue was resolved.

**Prepare a Summary Report**

1. I summarized the final data structure, transformations, and visualizations.

2. I provided recommendations for future data imports and consistency checks.

**8. Review and Finalize**

**A. Cross-Check Data**

1. I reviewed the final Excel workbook for accuracy.

2. I performed a final cross-check with the original SQL data.

**B. Save and Share**

1. I saved the Excel workbook with a descriptive name (e.g., `Malaria\_Data\_Analysis.xlsx`).

2. I shared the document with stakeholders or team members, along with the documentation.