

Data Cleaning Summary (Excel)

1. Handled Missing Values

- Filled missing **Age** with the **median value (39 years)**.
- Filled missing **Purchase_Amount** with the **median value** (calculated in Excel).

2. Removed Duplicates

- Used Excel's Remove Duplicates tool and deleted 3 duplicate rows.

3. Standardized Text

- Corrected spelling errors in the Name column (e.g., Maleria Lopez → Maria Lopez, John SMaleith → John Smith).
- Standardized Gender values (M/F → Male/Female).
- Standardized Country names (us → USA, india → India, etc.).

4. Fixed Date Formats

- Converted all Join_Date values into a single format: dd-mm-yyyy.

5. Renamed Columns

- Changed column headers to clean format: customer_id, name, age, gender, join_date, country, purchase_amount.

6. Verified Data Types

- Age and Purchase_Amount → Numeric.
- Join_Date → Date format.

Final Output:

The dataset is complete, consistent, and duplicate-free. It is ready for further analysis and visualization.