### **Data Link :** [data\_transformation\_cleaning.xlsx](https://docs.google.com/spreadsheets/d/1wA3tUHNt7sh8nwoLEEsPL3GPt6OeOYvt/edit?usp=drive_link&ouid=116581504019999500009&rtpof=true&sd=true)

### **Data Cleaning Tasks**

1. **Remove Duplicates:**
   * Identify and remove any duplicate records in the SalesData and CustomerDetails sheets based on unique identifiers (e.g., OrderID for SalesData and CustomerID for CustomerDetails).
2. **Handle Missing Values:**
   * In the SalesData sheet:
     + Replace missing values in the Quantity and Price columns with appropriate default values (e.g., 0 for Quantity, average or median for Price).
     + Fill missing OrderDate with a placeholder like "Unknown" or the earliest available date.
     + Replace missing Country and SalesRep values with "Unknown."
   * In the CustomerDetails sheet:
     + Replace missing values in Name, Email, and Country columns with placeholders (e.g., "Unknown").
     + For missing PhoneNumber, use a placeholder like "N/A."
3. **Remove NULL and Empty Rows:**
   * Remove rows in the SalesData or CustomerDetails sheet where critical columns (like OrderID in SalesData or CustomerID in CustomerDetails) are completely empty.
4. **Trim Whitespace:**
   * Remove extra spaces from CustomerName, Product, SalesRep, and Name fields to ensure consistency.
5. **Data Type Corrections:**
   * Ensure that columns have the correct data types (e.g., Quantity as a whole number, Price as a decimal, OrderDate as a date).
6. **Standardize Text:**
   * Convert all text data (e.g., Country, CustomerName) to a consistent case (e.g., title case) to maintain uniformity.
7. **Validate Data:**
   * Check for invalid data (e.g., non-numeric characters in the Quantity column) and either correct or flag these entries.

### 

### 

### **Data Transformation Tasks**

1. **Column Splitting:**
   * Split OrderDate into separate columns for Year, Month, and Day for more granular analysis.
2. **Data Type Conversion:**
   * Convert OrderDate to a date data type for proper date-based operations.
   * Convert Quantity and Price columns to numeric data types if they are not already.
3. **Creating New Columns:**
   * Create a new column TotalValue in the SalesData sheet by multiplying Quantity by Price.
4. **Merge Queries:**
   * Merge SalesData with CustomerDetails based on CustomerName to enrich the sales data with customer information (e.g., Email and PhoneNumber).
5. **Conditional Columns:**
   * Create a new column OrderStatus in the SalesData sheet:
     + Mark orders as "Complete" if both OrderDate and SalesRep are not null.
     + Mark as "Incomplete" if any of these fields are missing.
6. **Replace Values:**
   * Standardize country names in both sheets (e.g., if there are variations like "U.S.A" and "USA," unify them to one format).
7. **Sort and Filter Data:**
   * Sort the SalesData based on OrderDate or TotalValue for chronological or revenue analysis.
   * Filter out rows that don’t meet certain criteria (e.g., remove orders where TotalValue is zero).
8. **Remove Unnecessary Columns:**
   * Remove columns that are not needed for analysis (e.g., PhoneNumber if customer communication analysis is not in scope).
9. **Group Data:**
   * Group sales data by Product or Country and aggregate relevant metrics (e.g., total Quantity and total TotalValue).
10. **Format Data:**
    * Apply appropriate number and date formats to the Price, Quantity, and OrderDate columns.

### 

### 

### 

### **Advanced Transformation Tasks**

1. **Custom Calculations:**
   * Calculate a column for DiscountedPrice if a discount strategy is to be applied (e.g., 10% off for orders over $1,000).
2. **Date Transformations:**
   * Add columns to classify OrderDate into quarters or financial periods for more detailed time-based analysis.

### **Tasks for : Filtering and Managing Parameters**

#### **Task 1: Filtering Rows Based on Conditions**

* Filter the data to include only relevant information, such as a specific date range, specific countries, or sales above a certain value.

#### **Task 2: Creating and Managing Parameters**

* Create parameters to allow for dynamic data filtering, such as setting a minimum order date or filtering by a specific country.
* Use these parameters to modify data transformations dynamically, enabling users to change their view of the data without directly editing the queries.

#### **Task 3: Using Parameters in Data Filtering**

* Apply these parameters to filter the SalesData table based on user-defined conditions, such as selecting a specific date range or filtering by country.

#### **Task 4: Testing and Modifying Parameters**

* Apply and test the filters using these parameters to ensure that data updates dynamically in response to changes.