# Data Cleaning & Analysis Steps – Digital Wallet Transactions Project (Excel)

## Data Analyst Training Topics

1. Data Cleaning  
2. Pivot Table Creation  
3. Data Visualization

## Project Overview

Dataset Used: digital\_wallet\_transactions.xlsx (sample data from Kaggle)

Objective: To clean, standardize, and analyze digital payment transactions using Excel.

## 1. Data Cleaning Process

• Preserve Raw Data: Save the original file as 'digital\_wallet\_transactions\_raw.xlsx' and work on a duplicate copy.

• Inspect Data Quality: Verify record count, check for blanks, duplicates, and inconsistent formats.

• Adjust Column Widths: Home → Format → AutoFit Column Width / AutoFit Row Height.

• Clean Text Fields: Use Find & Replace to remove text within parentheses and standardize case using PROPER(), UPPER(), or LOWER() functions.

• Handle Whitespace & Spelling: Use TRIM(PROPER()) to fix unwanted spaces and capitalization.

• Fix Date and Time Columns: Use DATEVALUE() to convert text-based dates to Excel date format.

• Handle Missing Values: Ctrl + A → Find & Select → Go To Special → Blanks → Type 'NA' → Ctrl + Enter.

• Merge or Split Columns: Use TEXTJOIN() to combine, or Data → Text to Columns to split text data.

• Remove Duplicates: Data → Remove Duplicates (based on TRANSACTION\_ID).

• Validate Numeric Columns: Convert text to numbers using VALUE() and apply proper currency/number formatting.

## 2. Pivot Table Creation

Objective: Summarize and visualize transaction performance by category, merchant, or payment method.

Steps:

• Convert data to a table: Select data → Insert → Table → Check 'My Table Has Headers'.

• Insert Pivot Table: Insert → PivotTable → Choose tbl\_transactions → Place in new worksheet.

• Example Analyses:

• - Total Sales by Product Category (Rows: PRODUCT\_CATEGORY, Values: Sum of PRODUCT\_AMOUNT)

• - Transaction Count by Payment Method (Rows: PAYMENT\_METHOD, Values: Count of TRANSACTION\_ID)

• - Top 5 Merchants by Sales (Rows: MARCHANT\_NAME, Values: Sum of PRODUCT\_AMOUNT → Sort → Top 5)

• - Average Fees by Payment Method (Rows: PAYMENT\_METHOD, Values: Average of TRANSACTION\_FEES)

• Visualization: Insert PivotChart → choose chart type (Column, Pie, Line) for summary visuals.

## 3. Best Practices

• Always refresh PivotTables after updates (Alt + F5).

• Rename fields for clarity and consistent naming.

• Use conditional formatting for quick insights.

• Insert Slicers for easy data filtering (Insert → Slicer).

## 4. Deliverables

|  |  |
| --- | --- |
| File Name | Description |
| digital\_wallet\_transactions\_raw.xlsx | Original dataset |
| digital\_wallet\_transactions\_cleaned.xlsx | Cleaned and formatted dataset |
| Digital\_Wallet\_Pivot\_Report.xlsx | Pivot tables and charts |
| Sales\_Performance\_Dashboard.xlsx | Visual dashboard summary |

## 5. Key Learnings

• Importance of preserving raw data.

• Using Excel formulas for text cleanup and standardization.

• Replacing blanks and converting formats for consistency.

• Creating professional PivotTables and dashboards.