# P03 Instructor & Reference Guide – Data Cleaning & Excel Dashboard

## Theory Brief

Real‑world data often contains inconsistencies, duplicates, and missing entries. Data cleaning is essential to ensure accurate analyses. Excel Tables provide structured formatting with built‑in filters and formulas, which simplify dashboard creation and reporting.

## Worked Example

Below is a snapshot of the first few rows of the synthetic dataset and summary statistics:

Date Product Region Sales Quantity  
2025-01-01 Tablet West 50792.45 17.0  
2025-01-02 Laptop South 72066.25 13.0  
2025-01-03 Laptop South 45394.43 3.0  
2025-01-04 Laptop East 61204.76 10.0  
2025-01-05 Accessory East 36131.75 2.0

### Basic Statistics

* Sales\_mean: 49928.71
* Sales\_median: 49639.56
* Sales\_mode: 36131.75
* Sales\_var: 143470248.75
* Sales\_std: 11977.91
* Quantity\_mean: 10.07
* Quantity\_median: 9.00
* Quantity\_mode: 19.00
* Quantity\_var: 34.73
* Quantity\_std: 5.89

### Correlation Matrix

Sales Quantity  
Sales 1.000000 0.012947  
Quantity 0.012947 1.000000

### Visualisations

An example plot is saved in the results folder as P03\_plot1.png.

## Evaluation Rubric

* Effectiveness of duplicate and missing value handling (30%)- Proper conversion to Excel Table and use of table features (30%)- Quality and clarity of the dashboard (30%)- Documentation and reasoning (10%)

## Common Pitfalls

Overlooking hidden duplicates, improper treatment of missing values (e.g., deleting too many rows), or cluttered dashboard design.