Data Wrangling Report

1. Introduction This report outlines the data wrangling process applied to the dataset 'Supermarket Sales.' The primary goal was to clean, transform, and prepare the data for analysis.

2. Data Loading The dataset was loaded into a Pandas DataFrame from a CSV file.

3. Handling Missing Values

- Checked for missing values in all columns.
- Visualized missing values using a heatmap.
- No significant missing data was found, so no imputation was needed.

4. Handling Duplicates

- Identified and counted duplicate rows.
- No duplicates were found, so no action was required.

5. Data Cleaning and Transformation

- Renamed the column 'cogs' to 'cost of goods sold' for clarity.
- Dropped unnecessary columns: 'Branch' and 'Invoice ID.'
- Ensured data types were correctly assigned to relevant columns.

6. Summary The dataset was cleaned and prepared for further analysis. The next step involves extracting business insights from the processed data.