

Data Wrangling Report

Overview of the Data :

The dataset contains supermarket sales records, including branch details, product categories, prices, payment methods, and customer ratings. The goal is to clean and prepare the data for analysis.

Data Cleaning Process :

1. Handling Missing Values :

- Checked for missing values using `isnull().sum()`.
- Removed missing values with `dropna()`.
- Verified that no missing values remained after cleaning.

2. Removing Duplicates :

- Used `drop_duplicates()` to eliminate duplicate records.
- Confirmed that all duplicates were removed with `duplicated().sum()`.

3. Converting Data Types :

- Transformed the Date column into datetime format for easier time-based analysis.
- Converted categorical values, like Payment Method, into numerical codes using `.astype("category").cat.codes`.

4. Normalizing Data :

- Scaled the Total Sales column between 0 and 1 to ensure consistency in analysis.

5. Saving Cleaned Data:

- Exported the cleaned dataset as cleaned_data.csv for further analysis.