

Data Analyst Internship – Task 1

Data Cleaning and Preprocessing Summary Report

Objective:

To clean and prepare a raw dataset by handling missing values, duplicates, inconsistent formats, and incorrect data types using Microsoft Excel. **Dataset Used:**

Pizza Sales Dataset **Tool Used:**

Microsoft Excel **Data Cleaning Steps Performed:**

1. Removed duplicate records using Excel's "Remove Duplicates" feature.
2. Identified and handled missing values by filling blanks with "Not Available" or appropriate values.
3. Standardized text data by converting all column headers and text entries to lowercase.
4. Removed unnecessary spaces using the TRIM function.
5. Converted date columns into a consistent format (dd-mm-yyyy).
6. Verified numeric columns such as "Quantity" and "Total" to ensure proper data types.
7. Ensured consistency and accuracy across all dataset fields for further analysis.

Outcome:

The Pizza Sales dataset is now clean, consistent, and ready for data analysis or visualization tasks. All missing, duplicate, and formatting issues have been successfully resolved. **Prepared By:**

K. Bharath

Data Analyst Intern