Coffee Shop Sales Data Cleansing (Python)

1. Project Goal

The goal of the project was to prepare a raw CSV file containing coffee shop transactions for further analysis. The project aimed to demonstrate the ability to clean, validate, and prepare data for business analysis.

1. Data Cleansing Process Steps

-Importing the Required Libraries

-Using the Pandas Library to Load a CSV File

-Backing Up the File

-Learning Basic Data Information

-Replacing Errors in Columns as Null Values

-Converting Columns to the Appropriate Type

-Checking for Null Values

-Converting Error Values to NaN

-Creating Temporary Columns to Later Replace in Regular Columns

-Completing Calculated Values in Temporary Columns

-Removing Temporary Columns

-Checking for Duplicates

-Detecting Potential Outliers

-Completing Missing Values

-Saving the .csv File

3. Skills Demonstrated in the Project

Python & Pandas: Loading Data, Manipulating DataFrames Type conversion.

Data Cleaning: Handling missing data, duplicates, outliers, and normalizing text data.

Data Validation: Validating numeric columns and checking value consistency (Total Spent).

Problem Solving: Analyzing and resolving common problems in sales data.

Portfolio-ready Skills: Preparing data for further analysis and visualization in BI.

4. Result

A cleansed and consistent data set, ready for analysis: e.g., sales reports by product, payment method, and location.