# Data Cleaning with Pandas & NumPy

## 🧠 Objective

This project demonstrates basic data cleaning techniques using Python libraries `pandas` and `numpy`.

## 📂 Files

- `data/customer\_sales.csv`: Raw dataset with missing and duplicate values

- `cleaned\_data/cleaned\_customer\_sales.csv`: Cleaned and processed version (generated)

- `notebooks/data\_cleaning.py`: Python script for cleaning logic

## 🧹 Cleaning Steps

- Remove duplicates

- Handle missing values with mean/median/fill

- Convert `JoinDate` to datetime

- Create a new feature `SpenderCategory`

## 📦 Requirements

- Python 3.x

- pandas

- numpy

Install using:

```bash

pip install pandas numpy