

Load the dataset into Python using Pandas.

analysis.

Identify important fields that will be useful for analysis.

Subtask 2: Download and Explore Dataset

Find and Download an E-Commerce Dataset

· Visit Kaggle and search for "E-commerce sales data".

**★ How You Can Perform This Task?** 

**Quantity, and Revenue.** 

Load the Dataset in Python

Explore the Dataset Structure

Identify Missing Values and Duplicates

Identify Key Columns for Analysis

behavior.

region.

exploration.

**Tasks** 

Document Observations

• Explore the structure of the dataset (columns, data types, missing values, etc.).

This step ensures that you understand the dataset before proceeding with cleaning and

Select a dataset that includes fields like Order ID, Date, Customer ID, Product,

· Download the dataset in CSV format and store it in a working directory.

· Display the first few rows to get an overview of the dataset structure.

Check the total number of rows and columns to understand dataset size.

• Examine if key fields like Order Date, Revenue, and Customer ID are present.

• Kote down and in transactions that ormating (e.g., incorrect date formats).

• Focus on sales and revenue-related columns that will be used in trend analysis.

Note any data quality issues that need to be addressed in the cleaning process.

Pro Tip: If the dataset is too large, consider using a sample of the data to speed up initial

Look for customer segmentation features like purchase frequency and spending

Note down any geographical information if available, to analyze sales distribution by

Identify different data types (e.g., numerical, categorical, date fields).

Scan the dataset for missing or null values in important columns.

· Write a short summary of key insights about the dataset.

Save the dataset in a structured format for further analysis.

Downloaded an e-commerce dataset from Kaggle.

Identified important columns for sales and customer analysis.

Explored dataset structure and key fields.

Checked for missing values and duplicates.

Documented key observations for data cleaning.

Open the CSV file using a data analysis tool like Pandas in Python or Excel.

Preprocessing

☐ Subtask 4: Submission

Step 3: Customer

**Analysis** 

3 Segmentation Using RFM

Step 4: Sales Forecasting

**Using Time Series Analysis** 

Step 5: Business Insights &

Recommendations

2 Step 2: Sales Trend Analysis >