Question8:

8. Scenario: You are a data scientist working for a company that sells products online. You have

been tasked with analyzing the sales data for the past month. The data is stored in a Pandas data

frame.

Question: How would you find the top 5 products that have been sold the most in the past month?

Answer:

import pandas as pd

df = pd.read\_csv(r"D:\datasets\question3.csv")

df.columns = df.columns.str.strip()

df['Date'] = pd.to\_datetime(df['Date'], format="%d-%m-%Y", errors='coerce')

cutoff\_date = pd.Timestamp.today() - pd.Timedelta(days=30)

recent\_sales = df[df['Date'] >= cutoff\_date]

product\_sales = recent\_sales.groupby('Product')['Quantity sold'].sum()

top\_5\_products = product\_sales.sort\_values(ascending=False).head(5)

print("Top 5 products sold in the past month:")

print(top\_5\_products)

Output:



