7. Scenario: You are working as a data analyst for an e-commerce company. You have been given

a dataset containing information about customer orders, stored in a Pandas DataFrame named

order\_data. The DataFrame has columns for customer ID, order date, product name, and order

quantity. Your task is to analyze the data and answer specific questions about the orders.

Question: Using Pandas DataFrame operations, how would you find the following information

from the order\_data DataFrame:

1. The total number of orders made by each customer.

2. The average order quantity for each product.

3. The earliest and latest order dates in the dataset.

Code:

import pandas as pd

df = pd.read\_csv(r"C:\Users\jampa\Downloads\customer\_orders.csv")

df['Order Date'] = pd.to\_datetime(df['Order Date'])

total\_orders\_per\_customer = df.groupby('Customer ID').size()

average\_quantity\_per\_product = df.groupby('Product Name')['Order Quantity'].mean()

earliest\_order\_date = df['Order Date'].min()

latest\_order\_date = df['Order Date'].max()

print("1. Total number of orders per customer:")

print(total\_orders\_per\_customer)

print("\n2. Average order quantity per product:")

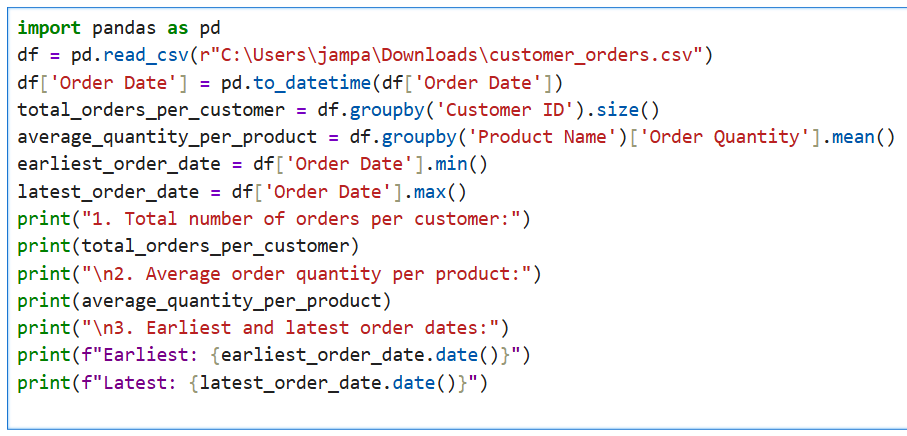
print(average\_quantity\_per\_product)

print("\n3. Earliest and latest order dates:")

print(f"Earliest: {earliest\_order\_date.date()}")

print(f"Latest: {latest\_order\_date.date()}")

output:



A screenshot of a computer

AI-generated content may be incorrect.

Dataset:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Customer ID | Order Date | Product Name | Order Quantity | |
| 101 | 1/10/2024 | Laptop | 1 |  |
| 102 | 1/12/2024 | Mouse | 2 |  |
| 103 | 2/1/2024 | Keyboard | 1 |  |
| 104 | 2/10/2024 | Laptop | 1 |  |
| 105 | 2/15/2024 | Mouse | 3 |  |
| 105 | 3/1/2024 | Monitor | 2 |  |
| 106 | 3/5/2024 | Laptop | 1 |  |
| 2 | 2 | 2 | 2 |  |
|  |  |  |  |  |
|  |  |  |  |  |