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:

#7

data=pd.read\_excel(r"C:\Users\hares\Documents\order\_data.xlsx")

data['Order Dat e'] = pd.to\_datetime(data['Order Date'])

orders\_per\_customer = data.groupby('Customer ID').size().reset\_index(name='Total Orders')

avg\_quantity\_per\_product = data.groupby('Product Name')['Order Quantity'].mean().reset\_index()

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

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

print("\n1. Total Orders by Each Customer:\n", orders\_per\_customer)

print("\n2. Average Order Quantity per Product:\n", avg\_quantity\_per\_product)

print(f"\n 3. Order Date Range: {earliest\_date.date()} to {latest\_date.date()}")

output :

1. Total Orders by Each Customer:

Customer ID Total Orders

0 C001 3

1 C002 3

2 C003 2

2. Average Order Quantity per Product:

Product Name Order Quantity

0 Keyboard 1.0

1 Laptop 1.0

2 Monitor 1.0

3 Mouse 2.0

3. Order Date Range: 2024-01-05 to 2024-01-22

Dataset :

|  |  |  |  |
| --- | --- | --- | --- |
| **Customer ID** | **Order Date** | **Product Name** | **Order Quantity** |
| C001 | 05-01-24 | Laptop | 1 |
| C002 | 07-01-24 | Mouse | 2 |
| C001 | 10-01-24 | Keyboard | 1 |
| C003 | 12-01-24 | Laptop | 1 |
| C002 | 15-01-24 | Monitor | 1 |
| C003 | 18-01-24 | Mouse | 3 |
| C001 | 20-01-24 | Monitor | 1 |
| C002 | 22-01-24 | Mouse | 1 |