In [1]:

```python
import pandas as pd
import numpy as np
np.random.seed(42)

num_transactions = 100
transaction_ids = np.arange(1, num_transactions + 1)
product_ids = np.random.randint(1, 20, num_transactions)
customer_ids = np.random.randint(1, 50, num_transactions)

dates = pd.to_datetime('2023-01-01') + pd.to_timedelta(np.random.randint(

units_sold = np.random.randint(1, 10, num_transactions)
unit_prices = np.round(np.random.uniform(5.0, 100.0, num_transactions), 2


data = {
    'TransactionID': transaction_ids,
    'ProductID': product_ids,
    'CustomerID': customer_ids,
    'Date': dates,
    'UnitsSold': units_sold,
    'UnitPrice': unit_prices
}

df = pd.DataFrame(data)


df['TotalRevenue'] = df['UnitsSold'] * df['UnitPrice']

print(df.head())

```

```
   TransactionID  ProductID  CustomerID       Date  UnitsSold  UnitPrice  \
0              1          7          22 2023-09-04          1      74.35
1              2         15          27 2023-11-20          1       9.53
2              3         11          35 2023-07-22          5      58.77
3              4          8           1 2023-07-03          3      20.07
4              5          7          35 2023-05-03          4      16.42

   TotalRevenue
0         74.35
1          9.53
2        293.85
3         60.21
4         65.68
```

In [2]:

```python
# Total revenue generated by all transactions
total_revenue = df['TotalRevenue'].sum()
print(f"Total Revenue: ${total_revenue:.2f}")

```

```
Total Revenue: $26575.09
```

In [3]:

```
# Average amount spent per transaction
average_spent = df['TotalRevenue'].mean()
print(f"Average Amount Spent per Transaction: ${average_spent:.2f}")

```

```
Average Amount Spent per Transaction: $265.75
```

In [4]:

```
# Top 5 products sold by quantity
top_products = df.groupby('ProductID')['UnitsSold'].sum().nlargest(5)
print("Top 5 Products Sold by Quantity:")
print(top_products)

```

```
Top 5 Products Sold by Quantity:
ProductID
8     47
12    46
7     39
15    38
1     31
Name: UnitsSold, dtype: int32
```

In [5]:

```
# Total amount spent by each customer
customer_spending = df.groupby('CustomerID')['TotalRevenue'].sum()
print("Total Amount Spent by Each Customer:")
print(customer_spending)

```

```
Total Amount Spent by Each Customer:
CustomerID
1     2019.35
2     1148.99
3      611.32
4      419.88
5       64.98
6      876.69
7      580.14
8      238.84
9      348.48
11     298.88
12     133.12
13     999.04
14     401.54
15    1264.52
17     346.45
19     126.24
20     166.44
22     571.99
23     819.83
24     301.87
25     364.84
26    1137.62
27     546.81
28    1141.80
29     118.94
30     260.28
32    1164.48
33     705.67
35     359.53
36      17.21
37    1187.31
39    1551.10
41    1039.56
42    1238.28
43     138.66
44     991.34
45     443.54
47      13.72
49    2415.81
Name: TotalRevenue, dtype: float64
```

In [6]:

```
# Total sales per month
monthly_sales = df.groupby(df['Date'].dt.to_period('M'))['TotalRevenue'].
print("Total Sales per Month:")
print(monthly_sales)

```

```
Total Sales per Month:
Date
2023-01     651.05
2023-02    1602.80
2023-03     138.79
2023-04    4327.44
2023-05    4718.54
2023-06    1761.06
2023-07    2126.11
2023-08    1234.94
2023-09    2552.21
2023-10    2674.39
2023-11    1112.88
2023-12    3674.88
Freq: M, Name: TotalRevenue, dtype: float64
```

In [7]:

```python
import pandas as pd
import numpy as np

# Generate sample data
np.random.seed(42)

num_transactions = 100
transaction_ids = np.arange(1, num_transactions + 1)
product_ids = np.random.randint(1, 20, num_transactions)
customer_ids = np.random.randint(1, 50, num_transactions)
dates = pd.to_datetime('2023-01-01') + pd.to_timedelta(np.random.randint(
units_sold = np.random.randint(1, 10, num_transactions)
unit_prices = np.round(np.random.uniform(5.0, 100.0, num_transactions), 2

data = {
    'TransactionID': transaction_ids,
    'ProductID': product_ids,
    'CustomerID': customer_ids,
    'Date': dates,
    'UnitsSold': units_sold,
    'UnitPrice': unit_prices
}

df = pd.DataFrame(data)
df['TotalRevenue'] = df['UnitsSold'] * df['UnitPrice']

# Calculations
total_revenue = df['TotalRevenue'].sum()
average_spent = df['TotalRevenue'].mean()
top_products = df.groupby('ProductID')['UnitsSold'].sum().nlargest(5)
customer_spending = df.groupby('CustomerID')['TotalRevenue'].sum()
monthly_sales = df.groupby(df['Date'].dt.to_period('M'))['TotalRevenue'].

# Results
print(f"Total Revenue: ${total_revenue:.2f}")
print(f"Average Amount Spent per Transaction: ${average_spent:.2f}")
print("Top 5 Products Sold by Quantity:")
print(top_products)
print("Total Amount Spent by Each Customer:")
print(customer_spending)
print("Total Sales per Month:")
print(monthly_sales)

```

```
Total Revenue: $26575.09
Average Amount Spent per Transaction: $265.75
Top 5 Products Sold by Quantity:
ProductID
8     47
12    46
7     39
15    38
1     31
Name: UnitsSold, dtype: int32
Total Amount Spent by Each Customer:
CustomerID
1     2019.35
2     1148.99
3      611.32
4      419.88
5       64.98
6      876.69
7      580.14
8      238.84
9      348.48
11     298.88
12     133.12
13     999.04
14     401.54
15    1264.52
17     346.45
19     126.24
20     166.44
22     571.99
23     819.83
24     301.87
25     364.84
26    1137.62
27     546.81
28    1141.80
29     118.94
30     260.28
32    1164.48
33     705.67
35     359.53
36      17.21
37    1187.31
39    1551.10
41    1039.56
42    1238.28
43     138.66
44     991.34
45     443.54
47      13.72
49    2415.81
Name: TotalRevenue, dtype: float64
Total Sales per Month:
Date
2023-01     651.05
2023-02    1602.80
2023-03     138.79
```

```
2023-04    4327.44
2023-05    4718.54
2023-06    1761.06
2023-07    2126.11
2023-08    1234.94
2023-09    2552.21
2023-10    2674.39
2023-11    1112.88
2023-12    3674.88
Freq: M, Name: TotalRevenue, dtype: float64
```

In [ ]:

```
1
```