```python
# retail_sales_analysis.py

import pandas as pd

# ----------------------------
# 1. Simulated Retail Dataset
# ----------------------------
data = pd.DataFrame({
    'Order_ID': [1001, 1002, 1003, 1004, 1005, 1006, 1007, 1008],
    'Order_Date': pd.to_datetime(['2024-01-15', '2024-03-23', '2024-07-11', '2024-10-05', '2024-11-21',
                                  '2024-12-15', '2024-02-12', '2024-08-30']),
    'Product_Category': ['Office Supplies', 'Consumer Electronics', 'Office Supplies', 'Consumer Electronics',
                         'Furniture', 'Furniture', 'Office Supplies', 'Consumer Electronics'],
    'Sub_Category': ['Binders', 'Phones', 'Paper', 'Accessories', 'Chairs', 'Tables', 'Storage', 'Accessories'],
    'Region': ['West', 'East', 'East', 'West', 'South', 'South', 'North', 'West'],
    'Sales_Representative': ['Ravi Sharma', 'Anita Singh', 'Ravi Sharma', 'Anita Singh', 'John Doe', 'John Doe',
                             'Kumar Patel', 'Ravi Sharma'],
    'Units_Sold': [15, 10, 40, 5, 8, 7, 20, 12],
    'Unit_Price': [25, 200, 3, 30, 150, 120, 7, 25],
    'Cost_Price': [15, 150, 2, 20, 100, 80, 5, 18],
    'Discount_%': [0.0, 10.0, 0.0, 5.0, 15.0, 0.0, 0.0, 5.0]
})

# ----------------------------
# 2. Feature Engineering
# ----------------------------
data['Total_Sales'] = data['Units_Sold'] * data['Unit_Price']
```

```
data['Total_Cost'] = data['Units_Sold'] * data['Cost_Price']
data['Profit'] = data['Total_Sales'] - data['Total_Cost']
data['Profit_Margin_%'] = round((data['Profit'] / data['Total_Sales']) * 100, 1)
data['Quarter'] = data['Order_Date'].dt.to_period('Q')
data['Year'] = data['Order_Date'].dt.year

# ----------------------------
# 3. KPI Overview
# ----------------------------
print("\n--- Overall KPIs ---")
print(f"Total Sales: ₹{data['Total_Sales'].sum()}")
print(f"Total Profit: ₹{data['Profit'].sum()}")
print(f"Profit Margin: {round((data['Profit'].sum() / data['Total_Sales'].sum()) * 100, 1)}%")
print(f"Average Discount: {data['Discount_%'].mean():.1f}%")
print(f"Units Sold: {data['Units_Sold'].sum()}")

# ----------------------------
# 4. Grouped Insights
# ----------------------------

# By Product Category
category_summary = data.groupby('Product_Category').agg({
    'Units_Sold': 'sum',
    'Total_Sales': 'sum',
    'Total_Cost': 'sum',
    'Profit': 'sum'
})
category_summary['Profit_Margin_%'] = (category_summary['Profit'] / category_summary['Total_Sales'] * 100).round(1)

print("\n--- Sales by Product Category ---")
print(category_summary)

```

```
60 # By Region
61 region_summary = data.groupby('Region').agg({
62     'Total_Sales': 'sum',
63     'Profit': 'sum'
64 })
65 region_summary['Profit_Margin_%'] = (region_summary[
   'Profit'] / region_summary['Total_Sales'] * 100).
   round(1)
66
67 print("\n--- Profit by Region ---")
68 print(region_summary)
69
70 # By Sales Representative
71 rep_summary = data.groupby('Sales_Representative').
   agg({
72     'Total_Sales': 'sum',
73     'Profit': 'sum'
74 })
75 rep_summary['Profit_Margin_%'] = (rep_summary['
   Profit'] / rep_summary['Total_Sales'] * 100).round(1
   )
76
77 print("\n--- Sales Rep Performance ---")
78 print(rep_summary)
79
80 # ----------------------------
81 # 5. Seasonality (Quarter)
82 # ----------------------------
83 seasonal = data.groupby('Quarter').agg({
84     'Total_Sales': 'sum',
85     'Profit': 'sum'
86 })
87 seasonal['Profit_Margin_%'] = (seasonal['Profit'] /
   seasonal['Total_Sales'] * 100).round(1)
88
89 print("\n--- Seasonality (Quarterly Sales & Profit
   ) ---")
90 print(seasonal)
91
```