**1.Optimizing Stock for New Store Opening in New York Based on Subcategory Profitability**

Your task is to provide a list of four subcategories from each category, ranked based on the total quantity sold in New York.

**Example approach\_Sixth Question**

1: Filter the DataFrame df to include only sales made in New York by using the condition df['state'] == 'New York'.

sales\_ny = df.loc[df['state'] == 'New York']

2: Calculate the total quantity sold for each subcategory within each category in New York. Group the DataFrame by 'category' and 'sub\_category' columns and sum the 'quantity' column.

subcategory\_profit = sales\_ny.groupby(['category','sub\_category'])['quantity'].sum().reset\_index()

3: Sort the subcategory\_profit DataFrame by 'category' and 'quantity' columns in descending order to get the subcategories with the highest quantities sold within each category. Then, group the DataFrame by 'category' and select the top 4 subcategories using .head(4).

subcategories\_ranked = subcategory\_profit.sort\_values(['category', 'quantity'], ascending=False).groupby('category').head(4)

4: Print the result, displaying the top 4 subcategories for each category, ranked by total profit in New York. Select the columns 'category', 'sub\_category', and 'quantity'.

print("The top 4 subcategories for each category, ranked by total quantity sold in New York are:")

print(subcategories\_ranked[['category', 'sub\_category', 'quantity']])

You output should look like this:

A close-up of a computer screen

AI-generated content may be incorrect.