Analysis Task

Dataset Analysis Questions

1. **Combine the Tables:**
    - Merge the provided tables into a single dataset for analysis.
2. **Sales Analysis:**
    - **Q1**: Calculate the "Sales" column values. How are they derived based on the dataset attributes?
3. **City Performance:**
    - **Q2**: Identify the city that generates the maximum sales. Provide a summary of how you arrived at your conclusion.
4. **Discount and Profitability Analysis:**
    - **Q3**: Analyze the impact of discount rates on the profitability of products across various sub-categories and regions. What patterns or trends can you identify?
5. **Quantity Statistical Measures:**
    - **Q4**: Compute the following statistical measures for the "Quantity" column:
        - Mean
        - Median
        - Standard Deviation

        Hint: Organize the "Quantity" data in ascending order and use statistical tools or programming methods for your calculations. Focus on the spread of the data to interpret the results.
6. **Discount Variability Analysis:**
    - **Q5**: Calculate the coefficient of variation (CV) for the "Discount" column and compare the relative variability between product types ("Apparel" and "Accessories").

        Hint:
        - First, compute the mean and variability (spread) of the "Discount" column for each product type.
        - Then, interpret which product type has a more consistent discount pattern, focusing on the relative variability.
7. **Chi-Square Test for Independence:**
    - **Q6**: Conduct a chi-square test of independence between the "Ship Mode" and "Region" columns.

        Hypotheses:
        - Null Hypothesis (H_0): Ship Mode and Region are independent.
        - Alternative Hypothesis (H_1): Ship Mode and Region are not independent.

        Note: Use a significance level (α) of 0.05 for this test.

        Hint:
        - Create a contingency table by counting the frequency of each "Ship Mode" within each "Region."

- Calculate the expected frequencies for each cell based on the row and column totals.