Digital Marketing Analyst

Liankai Ren

1. Identify and Correct Errors

In this step, used the **Find and Replace** function to correct any typographical errors in the **Region** and **Salesperson** columns. This helped standardize the data and avoid discrepancies in further analysis.

The dataset now correctly reflects:

- **4 regions**: East, West, South, North.
- **4 salespeople**: Ali Khan, Jane Smith, Maria Lee, John Doe.

After addressing the typographical errors, the next challenge was to fix the missing values in the transaction dates. Given that the transactions were not ordered by time, chose a more logical approach based on the assumption that the transaction date should be related to the salesperson was necessary, the product being sold, and the region. The reasoning was that, within the same period, a salesperson would likely focus on selling the same type of goods in the same region, with a higher success rate and probability of completing transactions.

To fill in the missing dates:

- Create subsets based on salesperson, product, and region.
- Calculate the average date within each subset, from the earliest transaction date to the latest.

For example, the first missing date was for Maria Lee, who was selling Clothing in the North region. After creating the subset based on the salesperson, product, and region, the earliest transaction as 2024-01-09 and the latest transaction as 2024-02-02. The average date would be 2024-01-21 to fill in the missing value.

During the data cleaning process, there was an error in the date column where TXN1030 was recorded as occurring on 2024-02-31, which is an invalid date since February only has 29 days in 2024. I corrected this error by changing the date to 2024-02-29, the last valid day in February.

2: Analysis

Sales by Product and Region

Based on the data from the **pivot table**, an analysis has been conducted of the sales by both **region** and **product**. The key findings are:

- **East** is the **top-selling region**, with sales almost twice as high as the **West** region, which ranks second in total sales.
- **Furniture** is the **top-selling product**, with significantly higher sales compared to other products.
- **Clothing** and **Electronics** sales are very close to each other, may indicating similar demand and performance for these two product categories.

Top 3 Salespeople Based on Total Sales

Based on the analysis of total sales, the **top 3 salespeople** are:

1. **Ali Khan**: Ali Khan had the **best performance**, accounting for **41% of the total sales count**. The total sales were more than **twice** that of **Jane Smith**, who placed second.
2. **Jane Smith**: Despite contributing only **19% of the total sales count**, Jane Smith had the **highest average transaction value**. This higher average transaction value helped to

secure the **second place in total sales**, outperforming others with a smaller number of sales but higher-value transactions.
3. **Maria Lee**: Maria Lee completed the top 3 ranking, showing consistent sales performance, though with lower sales compared to Ali Khan and Jane Smith.

Discount Analysis

Based on the discount rules provided, the analysis of the discount data revealed the following:

- There was a total of **128 transactions** that received a discount.
- Out of these, **90 transactions** received a discount greater than **$300**, meaning over **90% of the transactions** were eligible for the **15% discount**.
- The overall **total discount rate** for all transactions is **14%**.

Given that the majority of discount are receiving the **15% discount** (due to the $2,000 threshold), the total discount rate is higher than expected. A possible adjustment could be to **raise the $2,000 threshold** to lower the number of transactions qualifying for the 15% discount. This adjustment could help bring the **total discount rate closer to 10%**, which might be a more balanced target.

Sales Performance Evaluation

All regions successfully **met their target sales**. However, there were notable differences in performance across the regions:

- **East**: The **East region** showed the highest total sales, demonstrating its dominance in overall sales performance.
- **West**: The **West region** had the most impressive **growth above the target**, indicating strong performance and a significant improvement over its target sales.
- **North**: While the **North region** met its target, its growth above the target was the **least**, suggesting that it could benefit from further strategies or adjustments to improve its performance and achieve higher sales in the future.

Conclusion

In this report, I successfully completed various Excel tasks involving data cleaning, validation, calculation, analysis and decision-making. Key steps included correcting errors in the dataset, filling missing dates using a logical approach, and addressing invalid date entries. I also analyzed sales performance, identifying the top regions and products, and provided insights on discount analysis and sales performance across regions.

Overall, the data showed strong performance in the East and West regions. Adjustments to the discount rules could help to raise the total sales. The analysis provided valuable insights into how to optimize sales strategies for future periods.