Questions ——> DATA MANIPULATION

Question-1, we were asked to Classify the customers into their discount bracket and find the counts of distinct customers falling in each discount category:

```
aryansharma@Aryans-MacBook-Air G-7(ProcDNA) % python3 test.py
Discount Bracket
10%    493
20%     54
30%      3
Name: count, dtype: int64
```

Explanation: The code classifies customers into three discount categories based on their number of transactions made by same customers and total dollar purchases. The output shows that there are 3 distinct customers falling into the '30% discount' category, 54 distinct customers falling into the '20% discount' category, 493 distinct customers falling into the '10% discount' category.

Question-2, we were asked to Create a table containing Customer ID, number of transactions made, category of discount coupon, and their total dollar purchase in last 2 years:

```
aryansharma@Aryans-MacBook-Air G-7(ProcDNA) % python3 test1.py
    Customer ID  Number of Transactions  Total Dollar Purchase Discount Category
0      AA-10375                       2                     71      10% discount
1      AA-10480                       5                    624      20% discount
2      AA-10645                       2                   1010      10% discount
3      AB-10015                       2                    136      10% discount
4      AB-10060                       4                    275      10% discount
..          ...                     ...                    ...               ...
545    WB-21850                       6                   1355      20% discount
546    XP-21865                       5                    539      20% discount
547    YC-21895                       3                    265      10% discount
548    ZC-21910                       7                   1198      20% discount
549    ZD-21925                       5                    747      20% discount

[550 rows x 4 columns]
aryansharma@Aryans-MacBook-Air G-7(ProcDNA) %
```

Explanation: The code creates a table with Customer ID, the number of transactions made, the category of the discount coupon, and the total dollar purchase for each customer in the last 2 years.

Question-3, we were asked to Find the top 10 customers based on their purchase amounts in the last 6 months:

```
Top 10 Customers based on Purchase Amount in the Last 6 Months:
Series([], Name: Dollar Sales, dtype: int64)
```

Explanation: The code tries to find the top 10 customers based on their purchase amounts in the last 6 months. However, the output shows an empty Series, indicating that there are no customers with purchases in the last 6 months in the given dataset.

Question-4, we were asked to Find the top 2 salespersons of StyleMore along with their bonuses:

```
Top 2 Salespersons of StyleMore with Bonuses:
Sales Person Name
Mike Davidson    165356
Jane Austin       91596
Name: Dollar Sales, dtype: int64
```

Explanation: The code identifies the top 2 salespersons of StyleMore along with their bonuses. The output shows that 'Mike Davidson' has the highest sales amount with a bonus of 165356, and 'Jane Austin' comes next with a bonus of 91596.

```
Top Selling Products in Each Category in 2022:
Category  Product Name
Jeans     Chinos               31000
Shirts    Full Sleeve Shirt    18331
Shoes     Formal Shoes         73350
Name: Dollar Sales, dtype: int64
aryansharma@Aryans-MacBook-Air G-7(ProcDNA) %
```

Question-5, we were asked to Rank the top selling product in each category over 2022 on the basis of their dollar sales:

Explanation:
The top selling products in each category for the year 2022, ranked on the basis of their dollar sales:

- Category: Jeans
  - Product Name: Chinos
  - Dollar Sales: $31,000
- Category: Shirts
  - Product Name: Full Sleeve Shirt
  - Dollar Sales: $18,331
- Category: Shoes
  - Product Name: Formal Shoes
  - Dollar Sales: $73,350