Project Tittle: Retail Store Purchase Patterns

Project Description: To ensure patterns and relationship between products purchased together by customer. Optimizing inventory management, marketing strategies, cross-selling and upselling opportunities. Thereby increasing overall sales and customer satisfaction.

Data preprocessing:

- Import necessary libraries and load dataset on Jupyter notebook
- Checking Data columns and null values. Delete some columns and fill null values with perfect match of those data.
- Then dealing with some duplicates data.

Analytical part:

Season vs city by total items:

City / Season	Atlanta	Boston	Chicago	Dallas	Houston	Los Angeles	Miami	New York	San Francisco	Seattle
Fall	94	134	128	152	105	171	159	159	198	147
Spring	148	121	147	151	160	115	132	95	157	176
Summer	98	110	141	89	115	161	130	109	148	117
Winter	178	115	208	157	145	166	152	54	205	156

Insights:

- Winter has the highest sales volume, with more total items sold across almost all cities where sales increase significantly during this winter, while Summer has the lowest.
- Comparatively less items sold in New York

City wise items and costs:

City	Total_Cost	Total_Items	Transaction_ID
Atlanta	4577.86	518	90
Boston	4775.43	480	86
Chicago	5554.17	624	110
Dallas	5160.77	549	103
Houston	5025.67	525	94
Los Angeles	5551.83	613	108
Miami	6008.34	573	104
New York	3882.62	417	76
San Francisco	6986.52	708	125
Seattle	5195.44	596	104

Insights:

- San Francisco also reports high sales, with $6986.52 in total costs, respectively.
- Miami city has also total costs with $6008.34 in sales and 596 items sold.
- Atlanta has the lowest sales in terms of total cost, with $4577.86.
- These cities indicates where marketing efforts or store expansion might be beneficial.

Payment methods with no of users
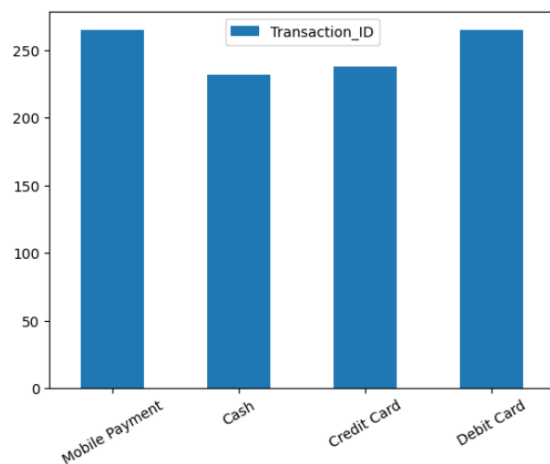
Insights:

- The bar chart shows the number of users for each payment method, with a high mobile payment usage rate indicating a strong consumer demand for digital payment methods.
- Though it is still frequently used, cash is the least chosen option, suggesting that some clients still use traditional ways to pay.

No of stores in different cities:
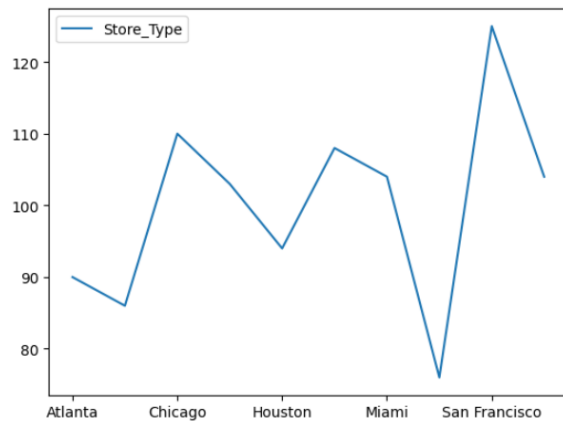
Insights:

- San Francisco may have the most extensive or vast market presence because it has the most stores.
- With the least amount of stores, Houston and Atlanta may offer fewer selections or a less diverse store structure.
- Strong market diversity is also seen in other cities, where the number of stores is significant.

Stores with discount and without discount in different season:
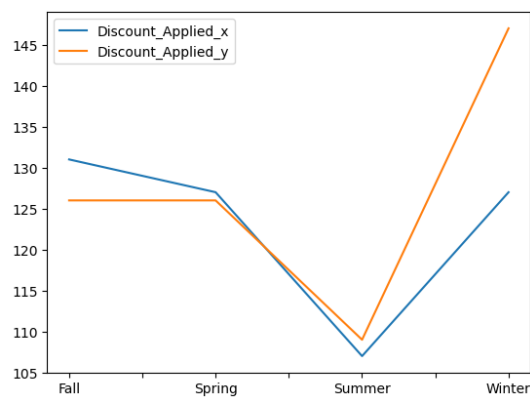
Insights:

- During Winter, there is a significant increase in transactions with discounts applied, indicating that discounts might be used as a strategy to drive sales in this peak season, and Summer has the least.
- Winter had the biggest sales, which is consistent with the higher discount application all over this season. This indicates that discount strategies are useful for maintaining strong sales at this time.

Monthly transactions and sold items:
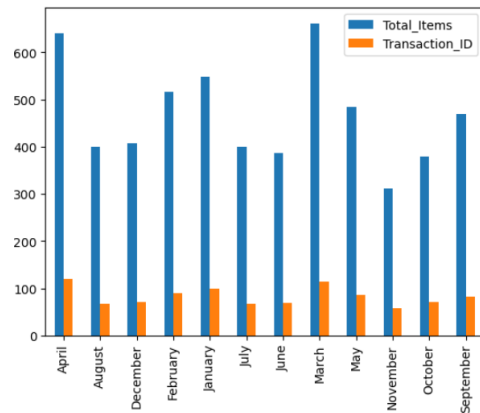
Insights:

- March and April show a peak in both transactions and total items sold, suggesting possible seasonal campaigns or events that boost sales in these months.
- December also shows a peak in both transactions and total items sold, likely due to holiday shopping. This month appears, reinforcing the importance of Winter for sales.

Top 10 Heavy buyers:

	Customer_Name	Total_Items	Total_Cost
0	James Smith	26	149.29
1	Melissa Freeman	15	107.81
2	Michele Davis	14	102.88
3	Michael Johnson	14	141.37
4	Jessica Crawford	13	103.56
5	Derrick Chen	10	75.92
6	Victor Stein	10	64.22
7	Kimberly Gomez	10	96.09
8	Andrea Weeks	10	45.61
9	Christopher Small	10	71.17

Insights:

- James Smith is a high-value customer, indicated by their highest purchase volume (26 items) and total cost (149.29).
- Also, it appears Michele Davis and Melissa Freeman make large purchases on frequently.
- The top customers often have between 10 and 26 items overall, and their overall costs range between 45 and 149.

Promotion with their transactions:
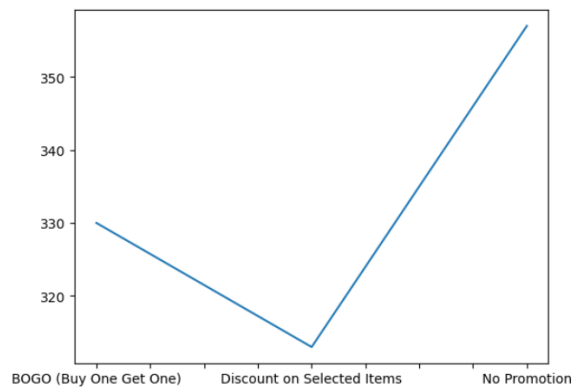
Insights:

- No Promotion has the highest number of transactions.
- BOGO (Buy One Get One) has the lowest transaction count among the promotion types.
- The "No Promotion" category has the highest transaction count. This may indicate that customers are purchasing even without discount, suggesting a strong customer base.

Customer segment with their transactions:

Customer_Category	
Senior Citizen	6847.91
Teenager	6803.86
Homemaker	6688.73
Young Adult	6619.92
Student	6541.71
Middle-Aged	6511.28
Professional	6428.38
Retiree	6276.86

Insights:

- Spending by senior citizens may be an indicator of a loyal and valuable customer within this age range, maybe supported by loyalty programs or specific offerings that target them.
- Given their high spending levels, teenagers and stay-at-home moms may have certain demands or preferences that the shop successfully meets.
- It's possible that retirees spend least could be a sign of their low level of engagement or purchasing power.

Store types with total costs:

```
Store_Type
Convenience Store     9658.78
Department Store      9386.65
Supermarket           8694.25
Specialty Store       8575.86
Pharmacy              8505.58
Warehouse Club        7897.53
```

Insights:

- The main fact that convenience stores generate the most spending indicates that, while possibly higher pricing, people may prefer the accessibility and quick assistance these establishments provide.
- The strong performance of supermarkets and department stores suggests that these shop types may be able to meet a wide range of consumer needs by providing a range of product categories and competitive pricing.
- Lower spending for pharmacies and warehouse clubs can be a result of a smaller customer or fewer visits. At these store formats, promoting specific products or incentives can help improve sales volume.

Conclusion:

Winter sees the highest sales, driven by effective discount strategies, while mobile payments are the most popular method. Convenience stores lead in spending due to accessibility, and Senior Citizens, Teenagers, and Homemakers are the top-spending segments. San Francisco and Miami are strong markets, but cities like Atlanta may benefit from more marketing efforts. Seasonal peaks in March, April, and December suggest targeting promotions during these months to maximize sales.

All resources are given in this link: https://github.com/sakib007q/customer_purchase_pattern.git