

Project Report: Retail Customer Retention Analytics – WALMART

1. Project Title

Retail Customer Retention Analytics – WALMART

2. Project Overview

Walmart, one of the largest multinational retail corporations, operates across hypermarkets, discount department stores, and e-commerce platforms worldwide. With increasing competition from Amazon, Target, and other retailers, retaining customers has become a critical challenge. Walmart collects vast data on customer purchases, loyalty programs, and online interactions, but their current reporting lacks analytical depth to: Boost brand engagement

You are hired as a Power BI Analyst to design a Customer Retention Dashboard for Walmart that consolidates fragmented data and delivers real-time, actionable insights. Understand why customers are churning?

- Identify loyal vs. at-risk customers
 - Measure the impact of loyalty tiers, promotions, and influencer-driven campaigns
 - Guide region- and channel-specific retention strategies
-

3. Project Objective

Develop an interactive Customer Retention Analytics Dashboard in Power BI using Walmart data that will:

- Consolidate customer demographics, purchase history, store/e-commerce performance, and membership data.
 - Enable dynamic segmentation of high-value, repeat, and churned customers.
 - Provide actionable insights to improve retention, loyalty engagement, and regional strategies.
-

4. Dataset Description

a.Customer_Demographics.csv

Columns: Customer_ID, Age, Gender, Region, Income_Level, Membership_Since, Preferred_Channel (Store/Online) Engagement Summary Dataset

b .Customer_Transactions.csv

Columns: Transaction_ID, Customer_ID, Store_ID, Product_Category (Groceries, Electronics, Apparel, Home & Living), Transaction_Date, Amount, Promotion_Applied (Yes/No)

c. Store_Locations.csv

Columns: Store_ID, Store_Type (Supercenter, Neighborhood Market, Sam's Club, Online), Region, Opening_Year

d. . Loyalty_Program.csv

Columns: Customer_ID, Loyalty_Tier (Basic, Plus, Premium, Elite), Points_Earned, Points_Redeemed

e.Churn_Labelled_Customers.csv

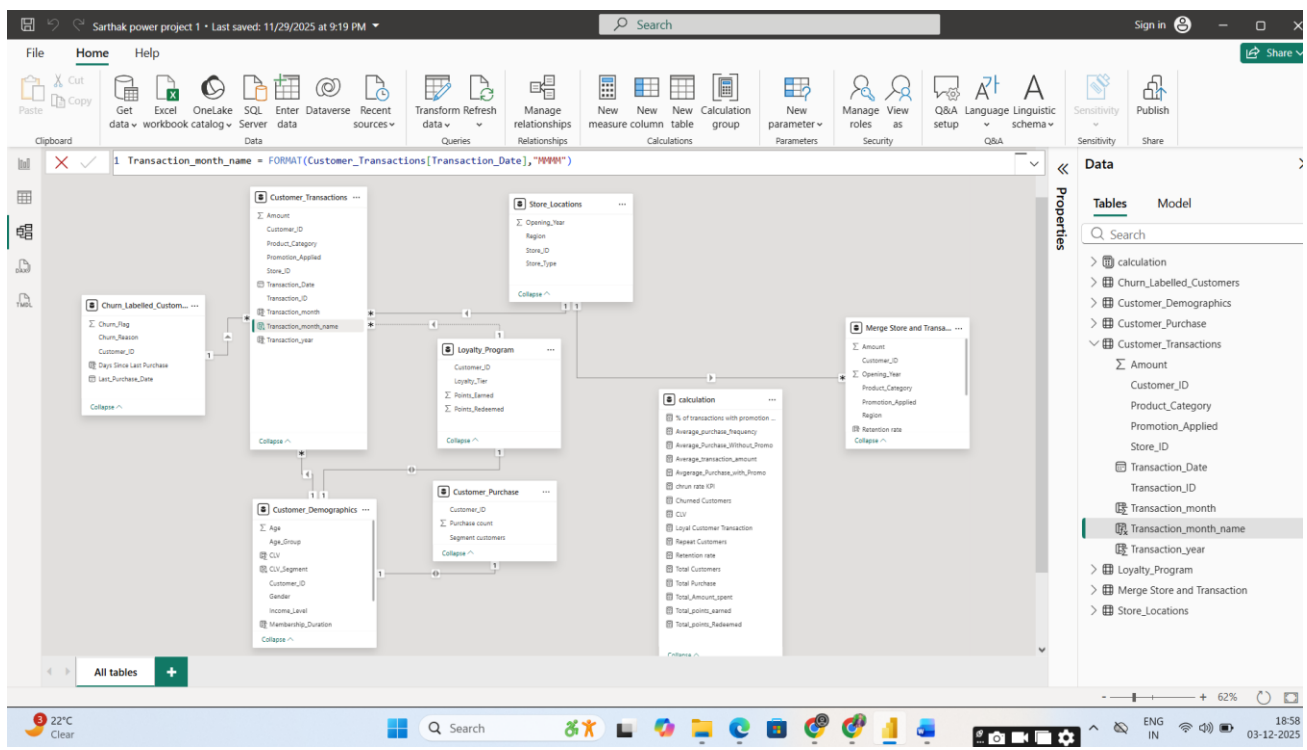
Columns: Customer_ID, Last_Purchase_Date, Churn_Flag (0 = Active, 1 = Churned), Churn_Reason (Inactivity, Competitor, Low Engagement)

5. Task-wise Implementation**Task 1: Data Modeling & Cleaning**

- Load and transform datasets in Power Query
- Handle duplicates, missing values, and ensure correct data types
- Create calculated columns:
 - $\text{Membership_Duration} = \text{Today} - \text{Membership_Since}$
 - Extract Transaction_Year, Transaction_Month
- Create a basic Data Model view
 - One-to-Many: Customer_Demographics → Transactions, Loyalty_Program, Churn_Labelled_Customers
 - Many-to-One: Transactions → Store_Locations

Power Query editor load and handle duplicate operations and so on .



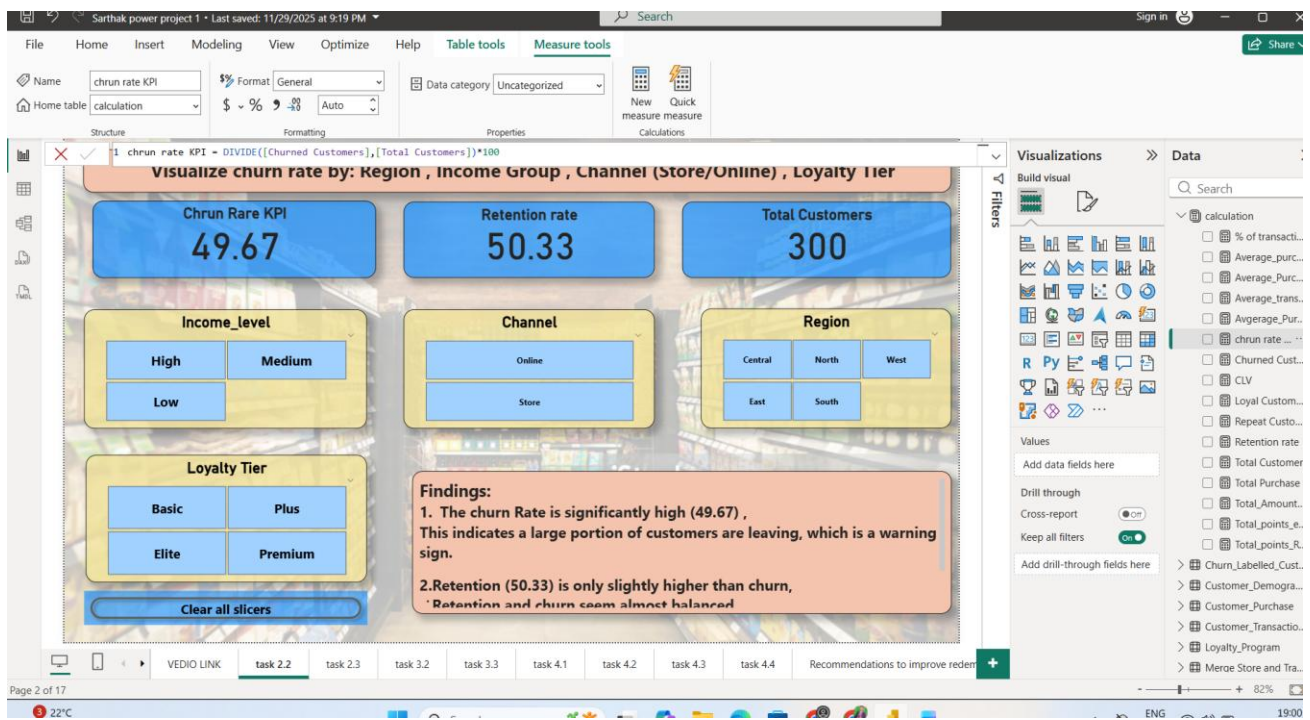


Task 2: Churn & Retention Metrics

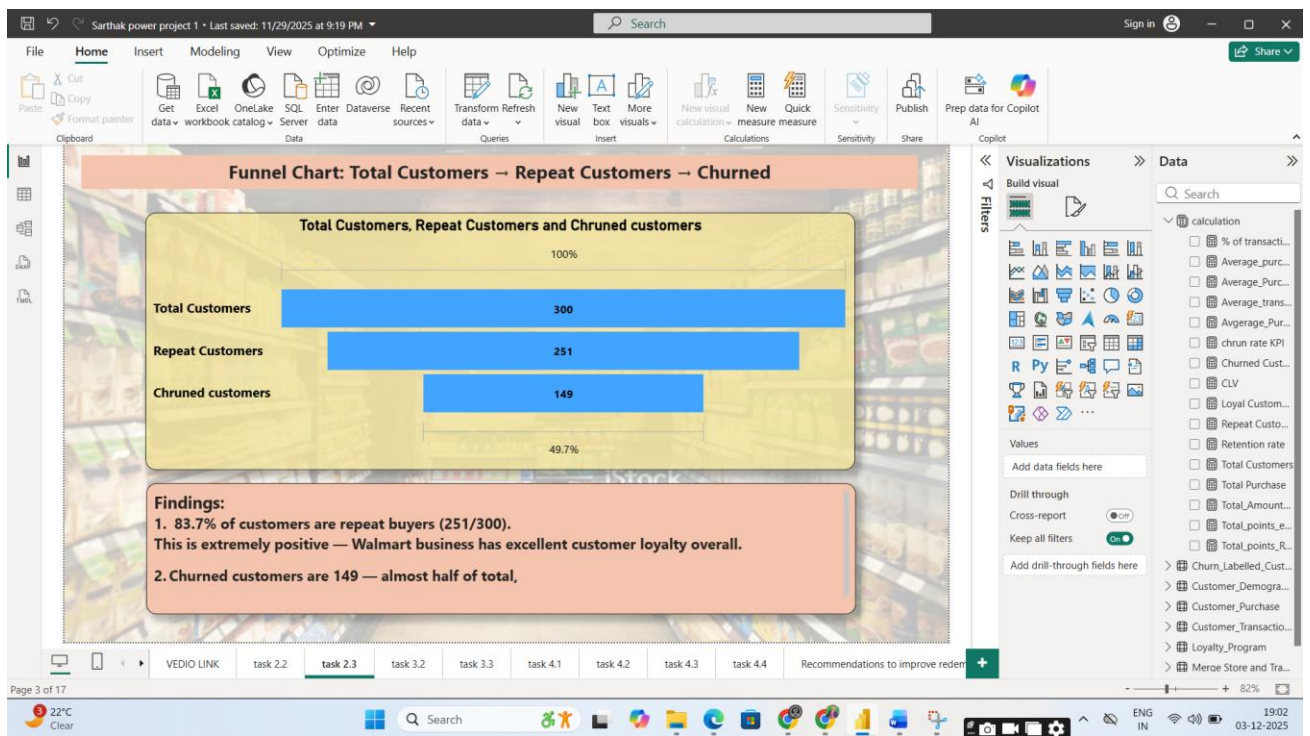
- Create Churn Rate KPI = $(\text{Churned Customers} / \text{Total Customers}) * 100$
- Visualize churn rate by:
 - Region
 - Income Group
 - Channel (Store/Online)
 - Loyalty Tier
- Funnel Chart: Total Customers → Repeat Customers → Churned

Soln: 1 task is done using creating Churn rate KPI measure, 2 task is done visualizing churn rate for that, I use cards and slicers, 3 task is done visualizing Total Customers Repeat Customers Churned by using funnel chart and measures for each.

.Churn rate measure and visualization by churn rate



Funnel chart



Findings:

- The churn Rate is significantly high (49.67),
This indicates a large portion of customers are leaving, which is a warning sign.
- Retention (50.33) is only slightly higher than churn,
 - Retention and churn seem almost balanced.
 - The business is retaining 50% of customers

3. 83.7% of customers are repeat buyers (251/300). This is extremely positive — Walmart business has excellent customer loyalty overall.

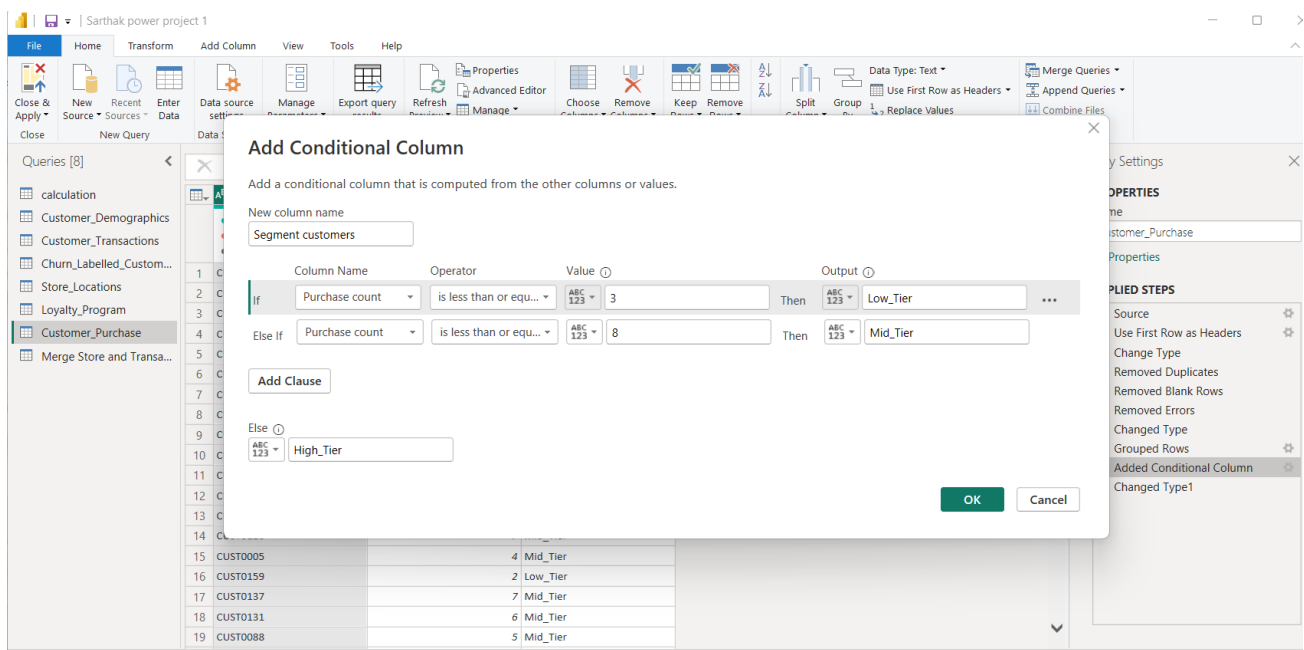
4. Churned customers are 149 — almost half of total,

Task 3: Repeat Purchase Analysis

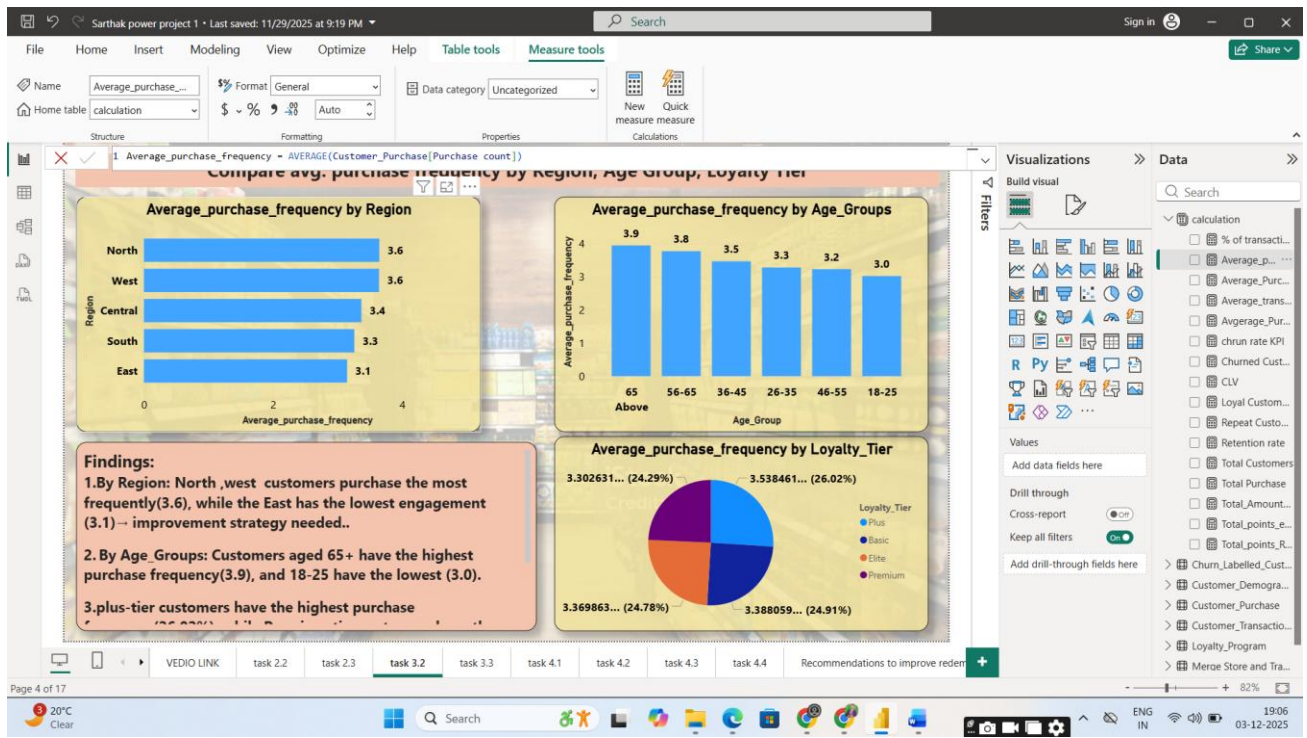
- Segment customers:
 - Low-Tier: 0–3 purchases
 - Mid-Tier: 4–8 purchases
 - High-Tier: 9+ purchases
- Compare avg. purchase frequency by Region, Age Group, Loyalty Tier
- Identify most purchased product categories by loyal customers

Soln: 1 task is done using Power query editor need to create duplicate transaction and group by customer id and add conditional column , 2 task is done visualiazing avg. purchase frequency and for that first create measure and visuliazue using different chart. For 3 task for loyal customers create mesasure first and then visualize using bar chart.

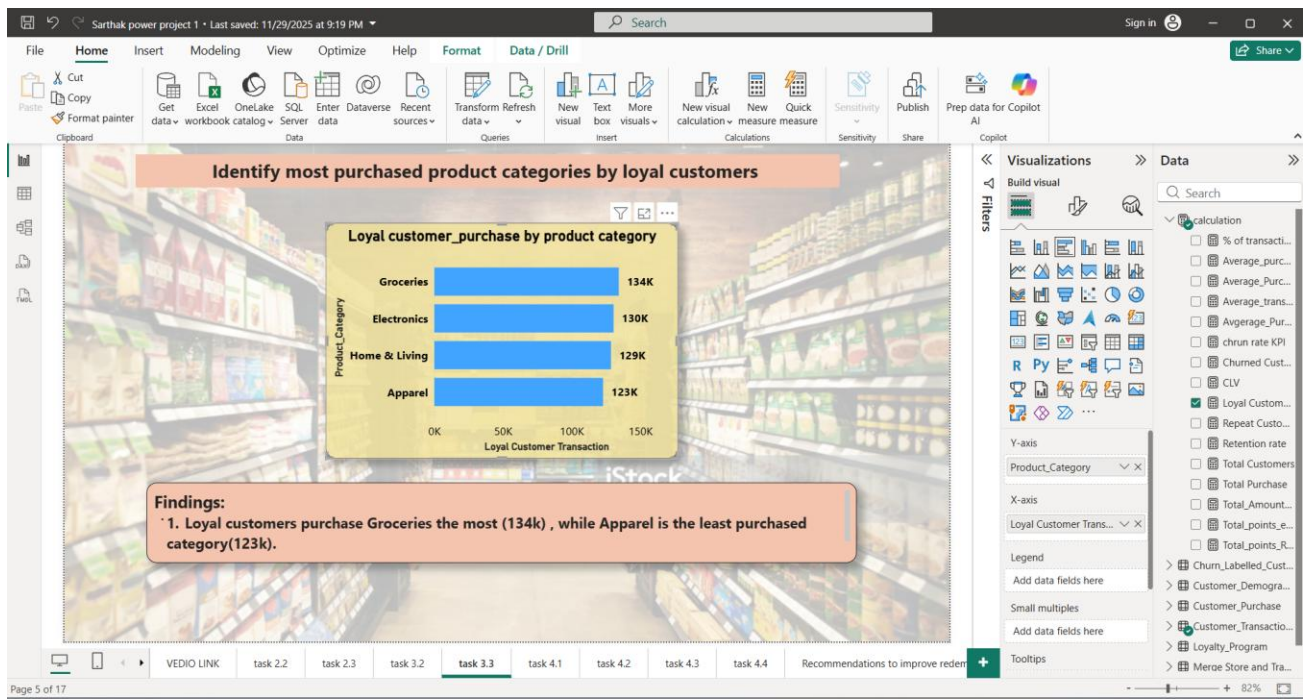
Segment customer in power query editor using conditional column and group by.



- Compare avg. purchase frequency by Region, Age Group, Loyalty Tier



.Identify most purchased product categories by loyal customers



Findings:

1. By Region:

North ,west customers purchase the most frequently(3.6), while the East has the lowest engagement (3.1)→ improvement strategy needed..

2. By Age_Groups: Customers aged 65+ have the highest purchase frequency(3.9), and 18-25 have the lowest (3.0).

3.plus-tier customers have the highest purchase frequency(26.02%), while Premium-tier customers have the lowest (24.29%) despite their higher status. Overall, purchase frequency is very similar across all tiers, indicating minimal impact of loyalty tier on buying behavior.

4.

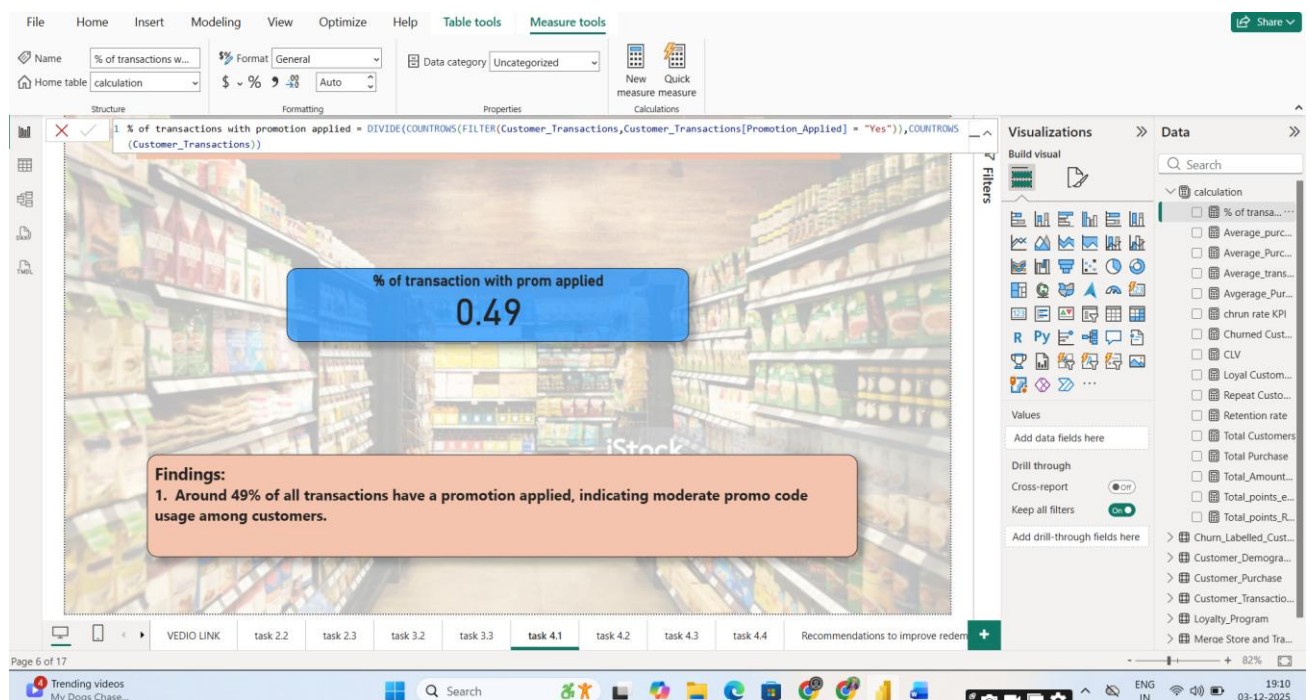
Loyal customers purchase Groceries the most (134k) , while Apparel is the least purchased category(123k).

Task 4: Promotion & Loyalty Impact

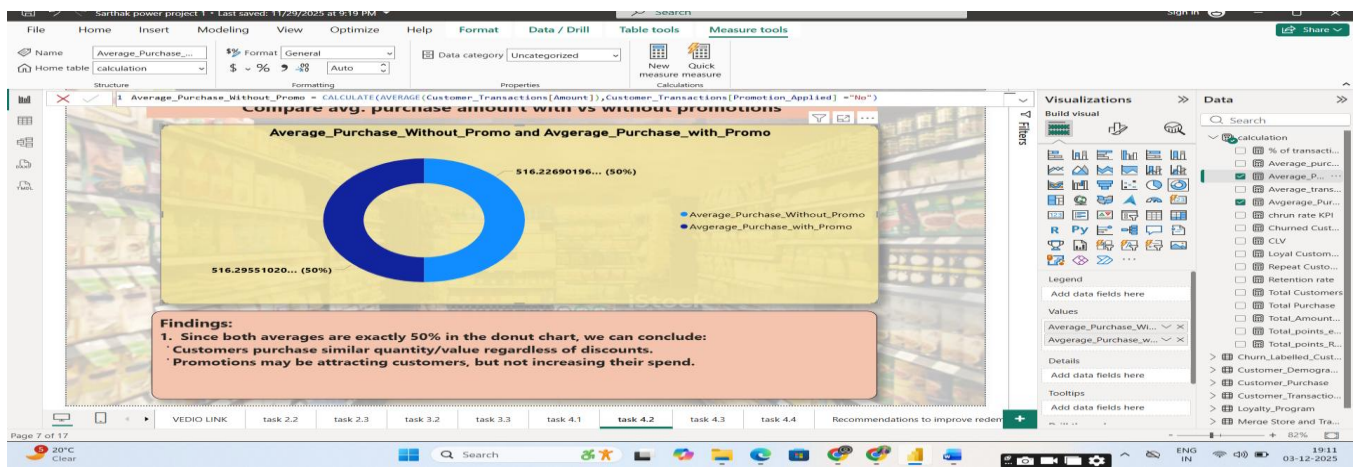
- % of transactions with promotion applied
- Compare avg. purchase amount with vs without promotions
- Churn rate across loyalty tiers
- Points Earned vs Redeemed by Tier (clustered column chart)
- Recommendations to improve redemption & retention

Soln: 1task is done using Creating measure and used card to visualize, and group by customer id ,2 task is done using Creating measure and visuliazize using Donut chart. For 3 task Visulize using bar chart , For 4 created using measur and Visulize using clusterd column chart ,

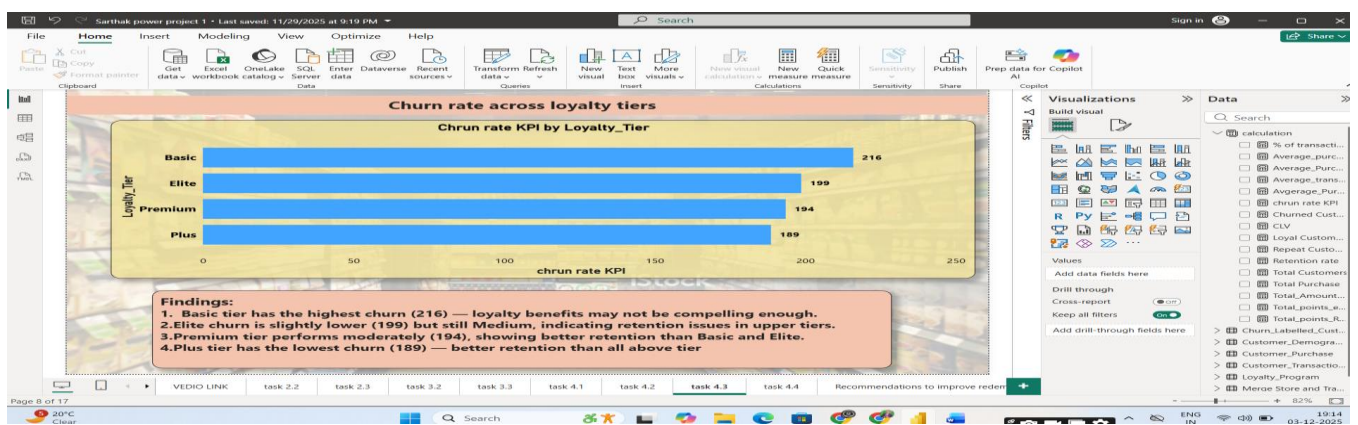
• % of transactions with promotion applied



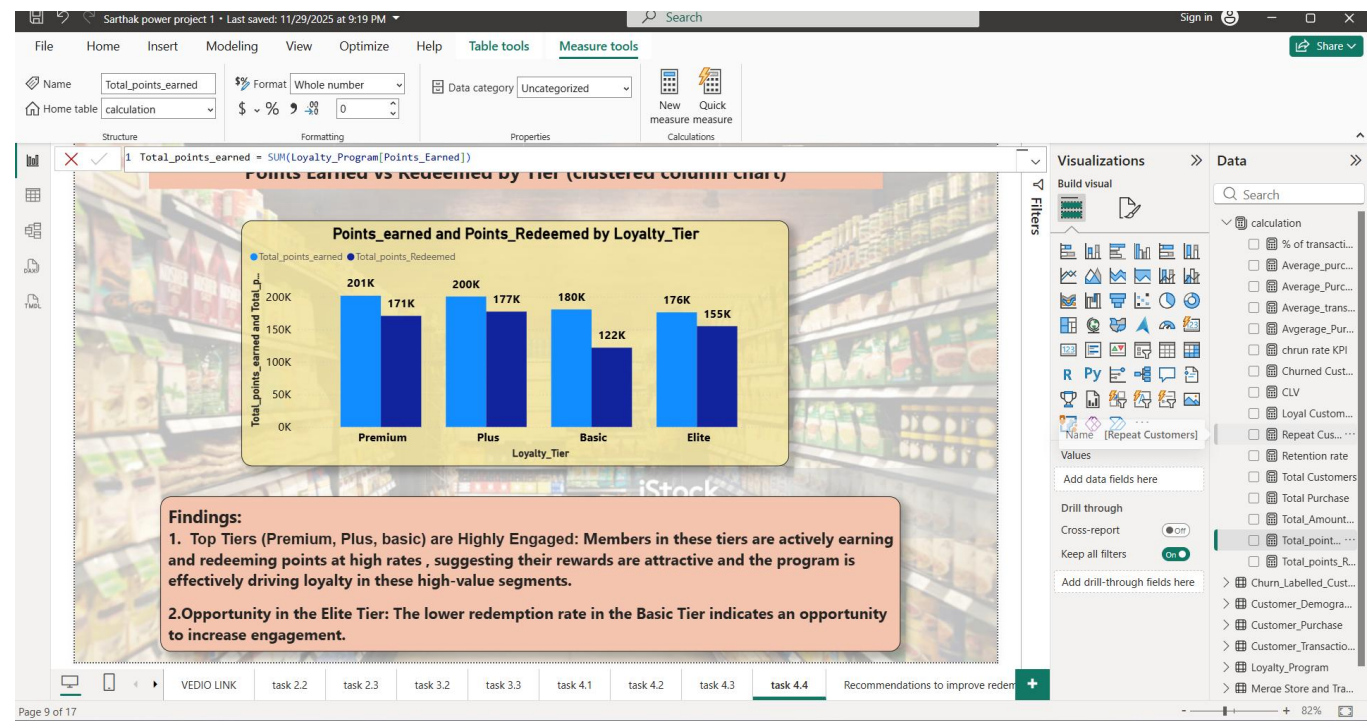
• Compare avg. purchase amount with vs without promotions



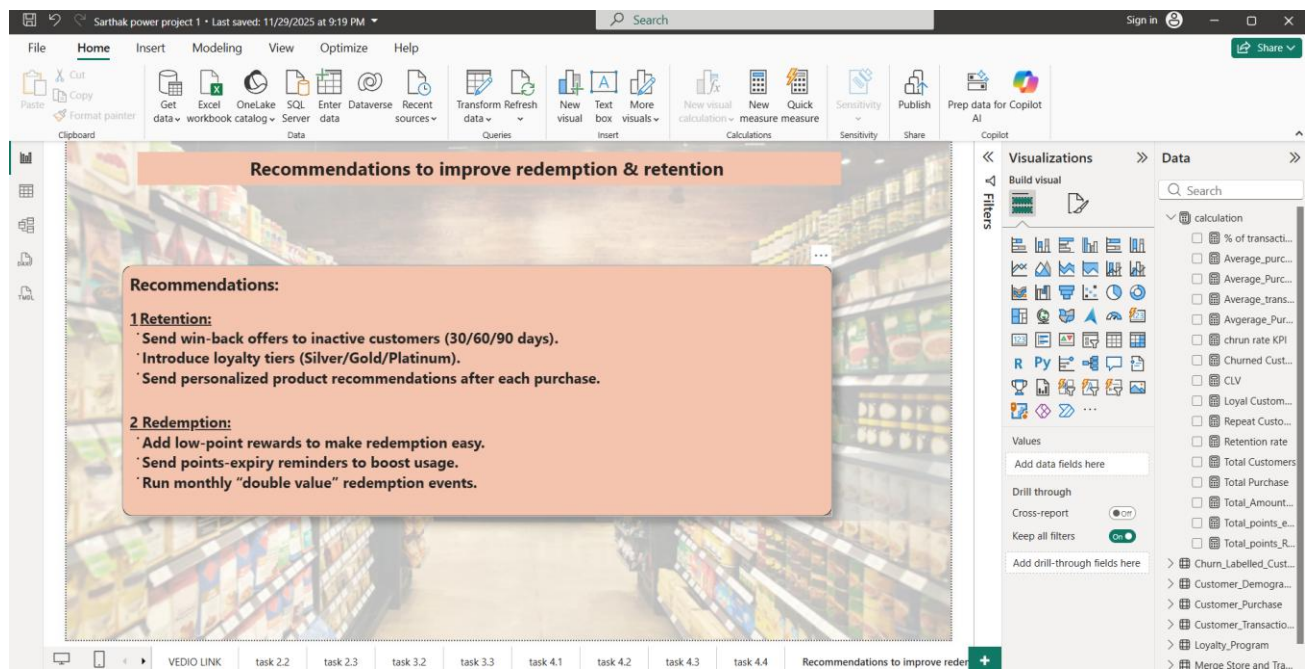
• Churn rate across loyalty tiers



• Points Earned vs Redeemed by Tier (clustered column chart)



• Recommendations to improve redemption & retention



Findings:

- Around 49% of all transactions have a promotion applied, indicating moderate promo code usage among customers.
- Since both averages are exactly 50% in the donut chart, we can conclude:

 - Customers purchase similar quantity/value regardless of discounts.
 - Promotions may be attracting customers, but not increasing their spend.
- Basic tier has the highest churn (216) — loyalty benefits may not be compelling enough.
 - Elite churn is slightly lower (199) but still Medium, indicating retention issues in upper tiers.
 - Premium tier performs moderately (194), showing better retention than Basic and Elite.
 - Plus tier has the lowest churn (189) — better retention than all above tier
- Top Tiers (Premium, Plus, basic) are Highly Engaged: Members in these tiers are actively earning and redeeming points at high rates, suggesting their rewards are attractive and the program is effectively driving loyalty in these high-value segments. Opportunity in the Elite Tier: The lower redemption rate in the Basic Tier indicates an opportunity to increase engagement.

Recommendations:

1 Retention:

- Send win-back offers to inactive customers (30/60/90 days).
- Introduce loyalty tiers (Silver/Gold/Platinum).
- Send personalized product recommendations after each purchase.

2 Redemption:

- Add low-point rewards to make redemption easy.
- Send points-expiry reminders to boost usage.
- Run monthly “double value” redemption events.

Task 5: Store & Channel Performance vs Retention

- Merge store data with transactions
- Visualize:
 - Avg. transaction amount by Store Type
 - Churn rate by store type
 - Correlation between store opening year & retention

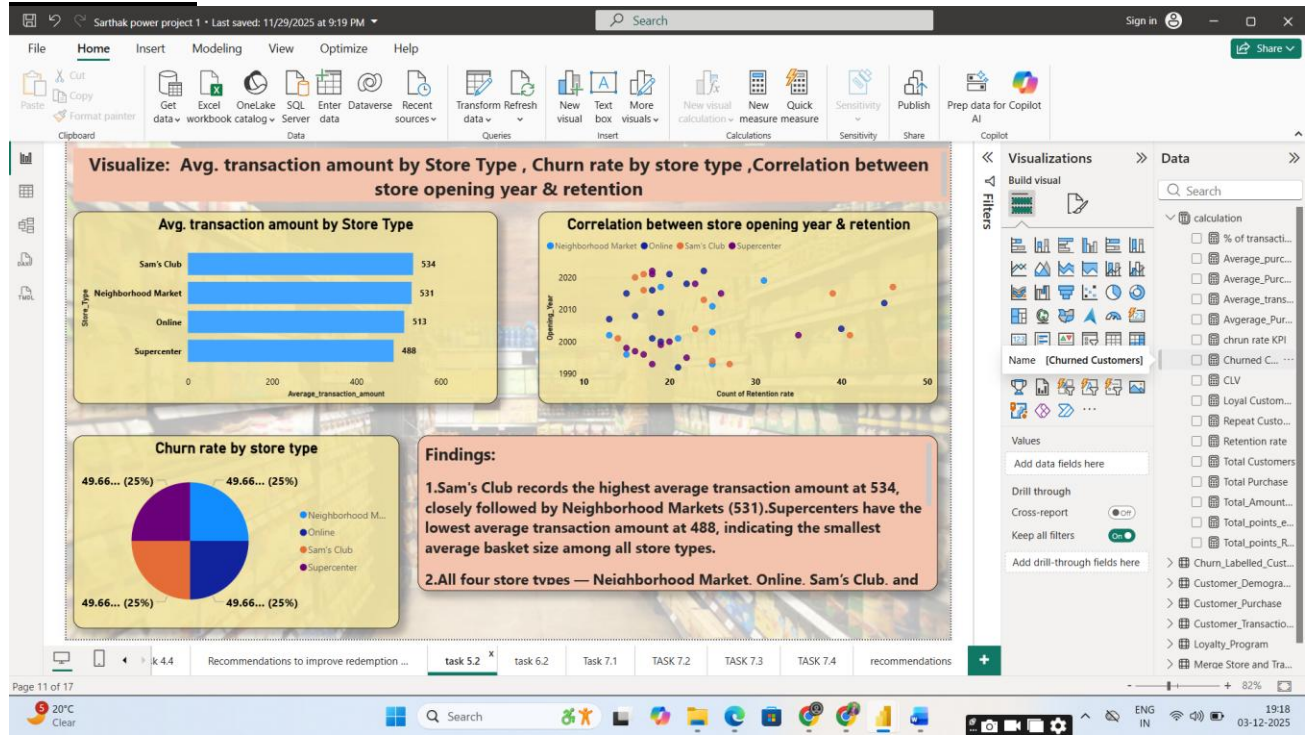
Soln: Merge store data with transactions using power query editor. for 2 tasks create Measure and visualize using different charts.

• Merge store data with transactions

The screenshot displays the Microsoft Power Query Editor interface. The main area shows a table with 10 columns: Transaction_ID, Customer_ID, Store_ID, Product_Category, Transaction_Date, and Amount. The table contains 19 rows of data. The interface includes a ribbon with tabs like File, Home, Transform, Add Column, View, Tools, and Help. The 'Transform' tab is active, showing options like Merge Queries, Merge Queries as New, and Merge Queries as Existing. The 'Merge Queries' button is highlighted. The 'Properties' pane on the right shows the 'Name' property set to 'Merge Store and Transaction'. The 'Applied Steps' pane shows the steps: Source, Expanded Store_Locations, Renamed Columns, and Removed Columns. The status bar at the bottom indicates '10 COLUMNS, 999+ ROWS' and 'Column profiling based on top 1000 rows'.

	Transaction_ID	Customer_ID	Store_ID	Product_Category	Transaction_Date	Amount
1	TXN00001	CUST0294	STORE011	Groceries	14-04-2025	
2	TXN00048	CUST0299	STORE011	Electronics	24-08-2024	
3	TXN00013	CUST0256	STORE001	Electronics	20-11-2023	
4	TXN00002	CUST0129	STORE035	Electronics	29-03-2024	
5	TXN00003	CUST0155	STORE024	Electronics	19-10-2023	
6	TXN00046	CUST0084	STORE024	Home & Living	03-07-2024	
7	TXN00004	CUST0189	STORE033	Electronics	27-12-2023	
8	TXN00041	CUST0156	STORE033	Apparel	17-03-2024	
9	TXN00006	CUST0110	STORE004	Electronics	01-01-2024	
10	TXN00005	CUST0209	STORE041	Groceries	29-03-2024	
11	TXN00022	CUST0202	STORE041	Electronics	30-07-2025	
12	TXN00007	CUST0229	STORE008	Home & Living	02-10-2024	
13	TXN00019	CUST0088	STORE007	Home & Living	18-06-2024	
14	TXN00040	CUST0174	STORE007	Electronics	16-09-2025	
15	TXN00008	CUST0221	STORE009	Home & Living	02-02-2024	
16	TXN00029	CUST0196	STORE009	Home & Living	26-04-2024	
17	TXN00009	CUST0062	STORE048	Electronics	04-03-2025	
18	TXN00023	CUST0217	STORE048	Electronics	22-05-2024	
19						

Visualization



Findings:

1. Sam's Club records the highest average transaction amount at 534, closely followed by Neighborhood Markets (531). Supercenters have the lowest average transaction amount at 488, indicating the smallest average basket size among all store types.

2. All four store types — Neighborhood Market, Online, Sam's Club, and Supercenter — show the same churn value: 49.66% (25% share each).

This means no store type is performing better or worse in terms of churn rate.

3..Positive Trend: As the variable on the Y-axis (Store Opening Year) increases (moving upwards toward more recent years), the maximum observed values of the X-axis variable (Retention Rate) also generally increase (moving rightwards toward higher retention).

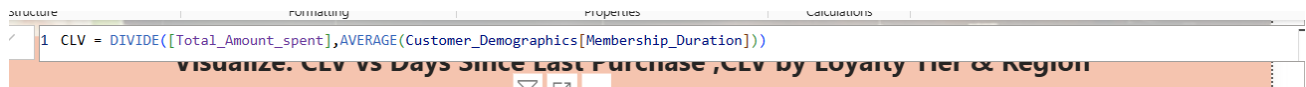
Task 6: Customer Lifetime Value (CLV) Analysis

- $CLV = \text{Total Amount Spent} / \text{Membership Duration (Years)}$
- Segment customers into Low, High CLV

- Above Average CLV as High
- Below Average CLV value- Low
- Visualize:
 - CLV vs Days Since Last Purchase
 - CLV by Loyalty Tier & Region

Soln: Create measure for clv and clv segment, and visualize clv using different charts.

. CLV = Total Amount Spent / Membership Duration (Years)



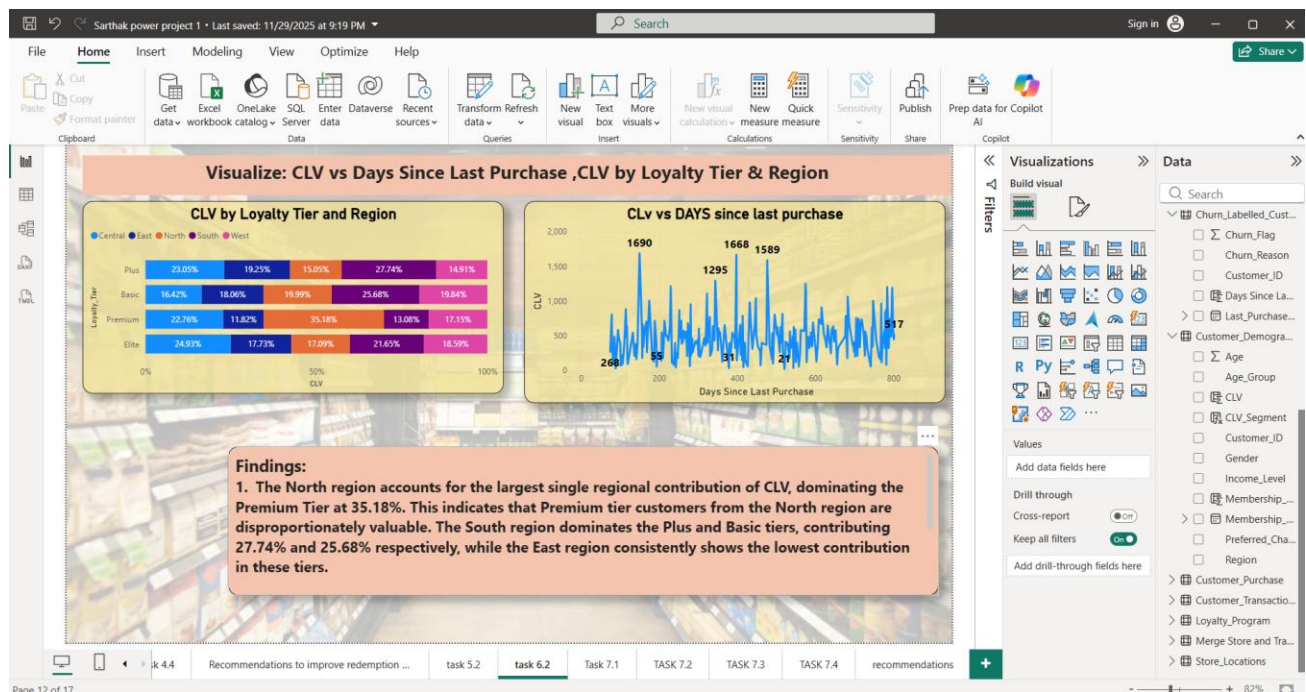
- **. Segment customers into Low, High CLV**

```
CLV_Segment =
VAR AvgCLV =
    CALCULATE(
        AVERAGE(Customer_Demographics[CLV]),
        ALL(Customer_Demographics)
    )
RETURN
    IF(Customer_Demographics[CLV] >= AvgCLV, "High CLV", "Low CLV")
```



- **Visualize:**

- **CLV vs Days Since Last Purchase**
- **CLV by Loyalty Tier & Region**



Findings:

1.

The North region accounts for the largest single regional contribution of CLV, dominating the Premium Tier at 35.18%. This indicates that Premium tier customers from the North region are disproportionately valuable. The South region dominates the Plus and Basic tiers, contributing 27.74% and 25.68% respectively, while the East region consistently shows the lowest contribution in these tiers.

2. High CLV values appear even at large gaps (such as 1295, 1589, and 1690 days), showing that some high-

value customers still return after long breaks, while low CLV values also appear at very small gaps (like 21, 31, and 55 days), meaning recent purchases don't always represent high-value customers. Overall, this indicates that CLV is not strongly linked to how recently a customer purchased, as customers of all value levels show different and inconsistent buying behaviors.

Task 7: Final Dashboard & Executive Summary

- Multi-page Power BI Report:

- Page 1: KPIs (Churn, CLV, Repeat Rate)
- Page 2: Loyalty & Promotion Impact
- Page 3: Store/Channel Insights
- Page 4: Segmentation (Churned, Repeat, High-Value)

- Slicers: Region, Channel, Income, Loyalty Tier

- Top 3 recommendations for Walmart :

- Which customers to prioritize for retention?

1. High-Spend Customers who stopped purchasing in last 60–90 days

→ They have high lifetime value; losing them hurts the most.

2. Customers with high points balance but low recent activity

→ They are close to churning but can be pulled back through rewards.

3. Frequent shoppers shifting to fewer visits

→ Decline in visit frequency is an early churn signal.

- Which channels are underperforming?

1. Push store shoppers to online by giving app-only discounts.

2. Improve store experience with in-store exclusive deals and faster checkout.

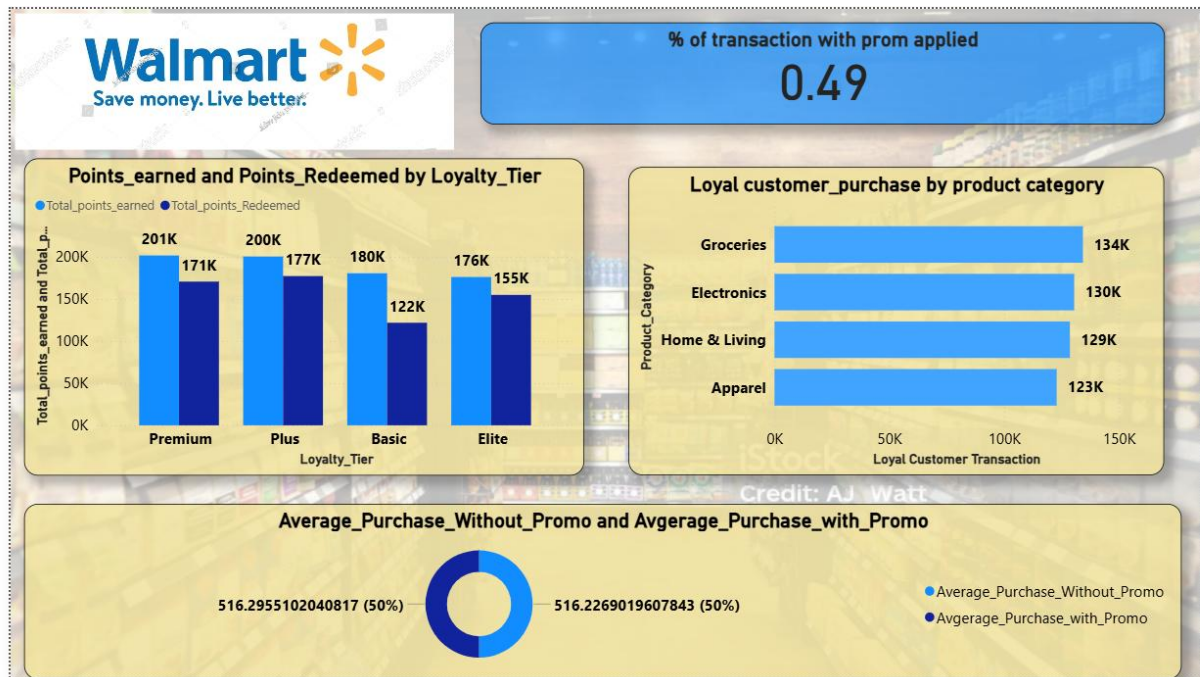
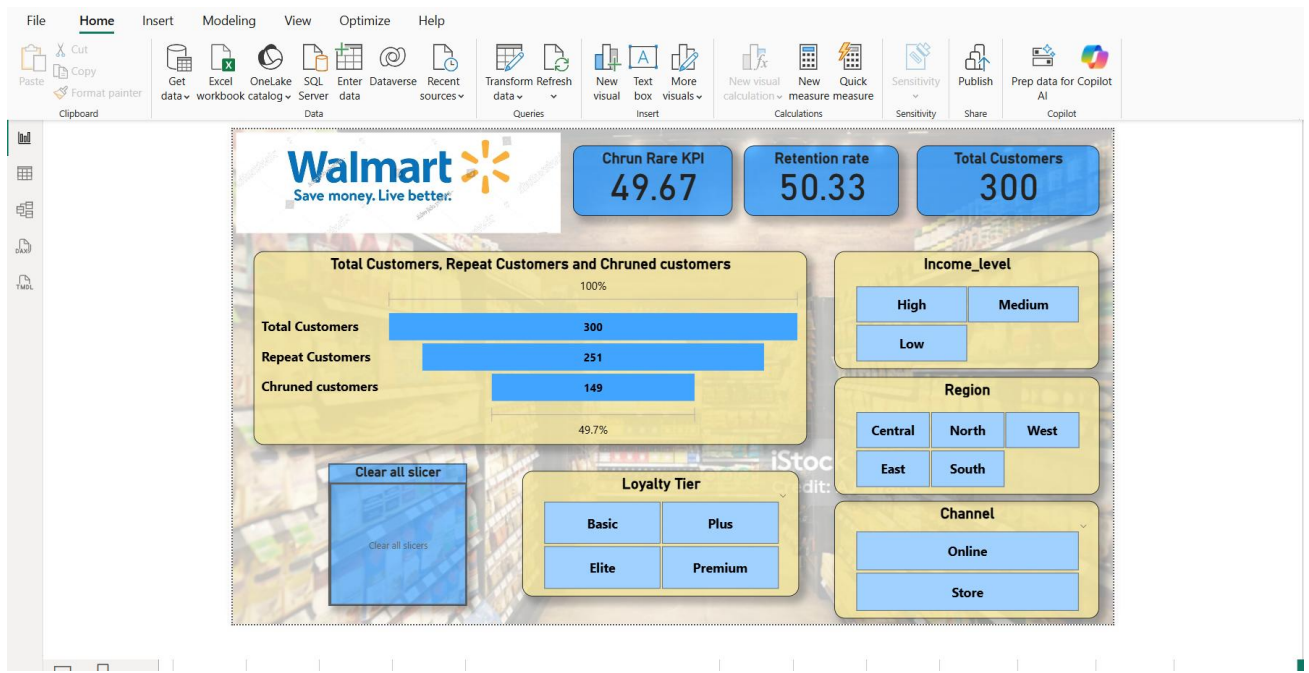
3. Promote loyalty program at billing counters so store customers earn & redeem more.

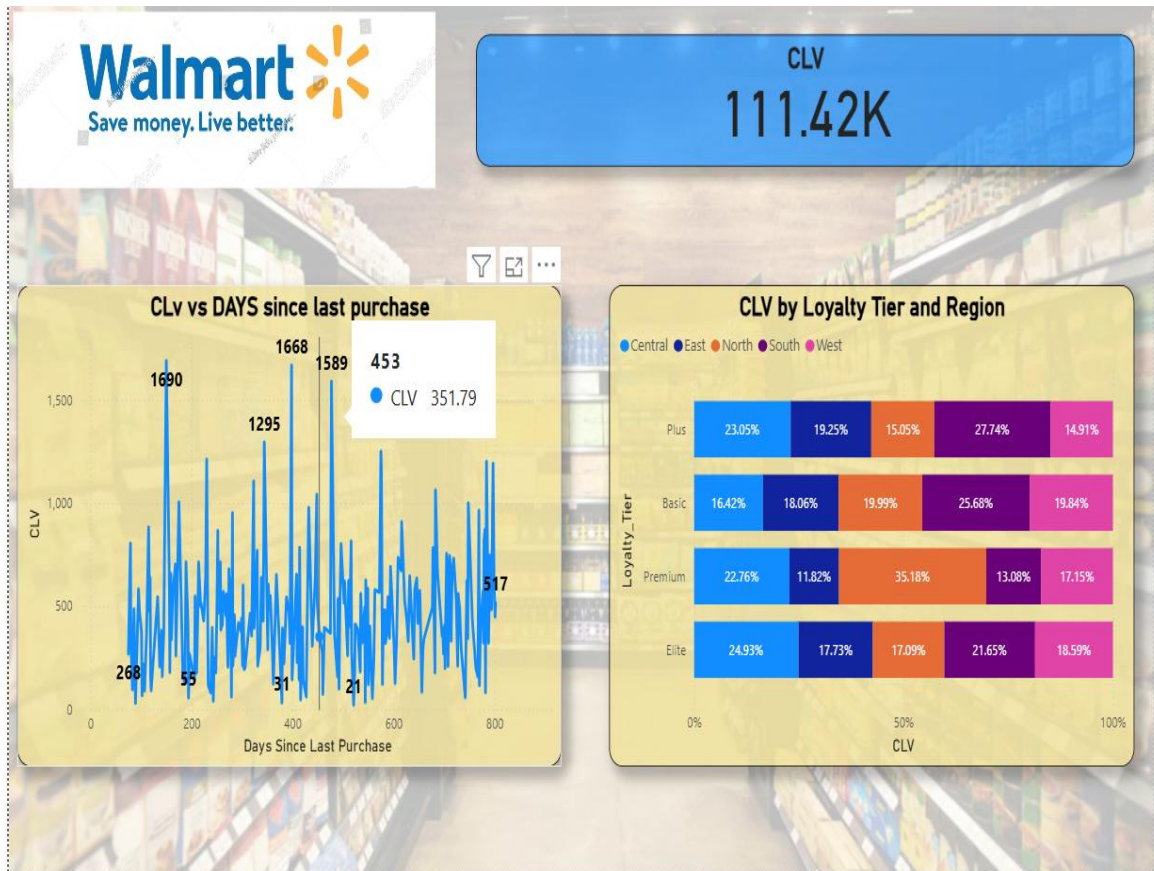
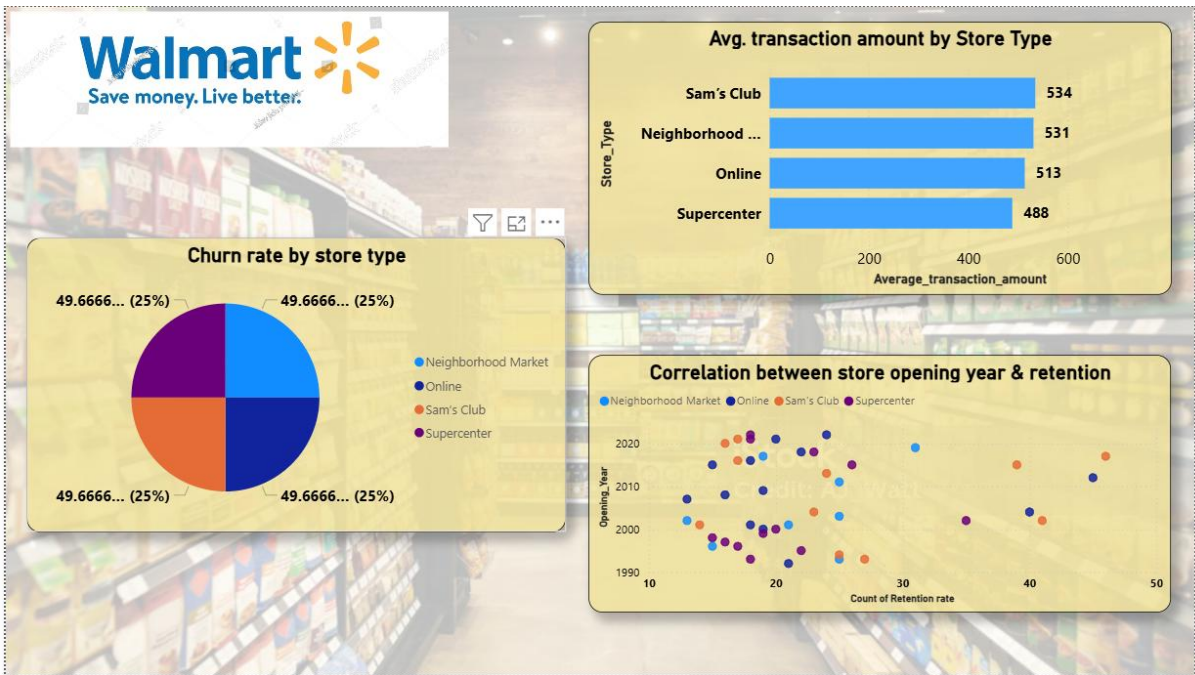
- How to strengthen loyalty program engagement

1. Create clear upgrade benefits to push customers from Basic → Plus → Premium → Elite.

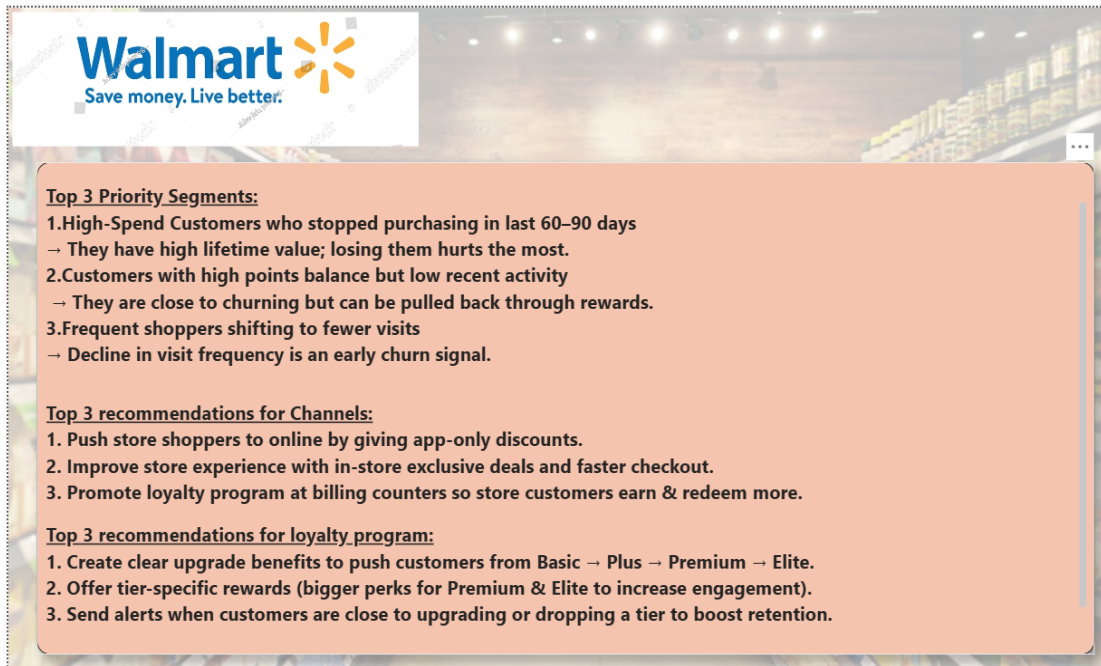
2. Offer tier-specific rewards (bigger perks for Premium & Elite to increase engagement).
3. Send alerts when customers are close to upgrading or dropping a tier to boost retention.

- **Multi-page Power BI Report:**





.Recommendation



8. Conclusion

This project provides Interactive Customer Retention Analytics Dashboard in Power BI using Walmart data that will:

- Consolidate customer demographics, purchase history, store/e-commerce performance, and membership data.
- Enable dynamic segmentation of high-value, repeat, and churned customers.
- Provide actionable insights to improve retention, loyalty engagement, and regional strategies.

9. Tools & Techniques Used

- **Microsoft Power bi (Power Query, Measure ,calculated column, Charts)**
 - **Data Cleaning & Transformation**
 - **Visualization (Line Charts, Bar Charts, Scatter Plots)**
 - **Loyalty & Promotion Impact Analysis**
 - **Store/Channel Insights Analysis**
 - **Customer Lifetime Value (CLV) Analysis**
-

10. References

- Retail Customer Retention Analytics – WALMART
 - Marketing Analytics Frameworks (Churn rate kpi ,retention rate)
 - Studies Trends and technologies
-