Retail Customer Retention Analytics Project: TESCO

1. Project Overview & Objective

Project Goal: Develop a robust, interactive Customer Retention Analytics Dashboard in Power BI for TESCO to consolidate siloed data, enable dynamic customer segmentation, and deliver actionable insights for improving retention, loyalty program effectiveness, and regional store strategies.

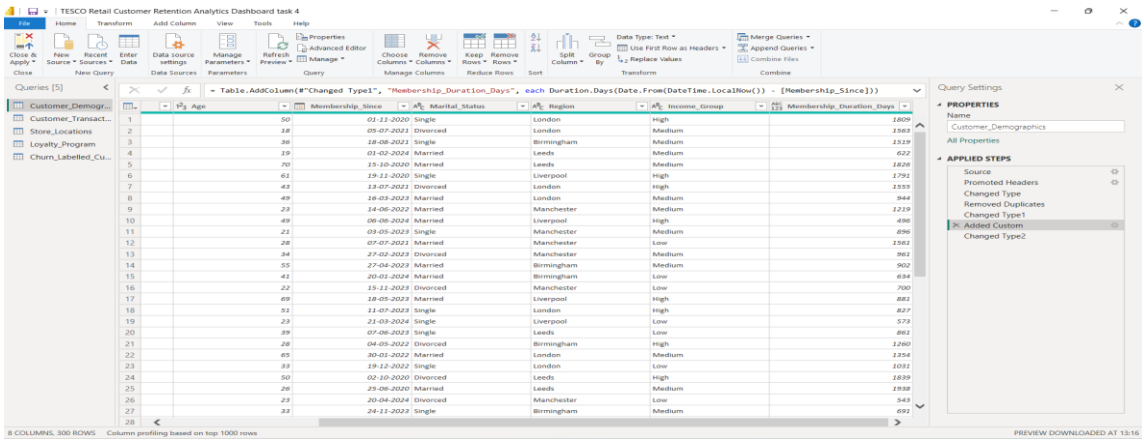
Problem Statement: TESCO needs deeper analytical insights beyond current reporting to understand customer churn drivers, identify high-value/at-risk customers, evaluate loyalty programs, and guide store-specific retention strategies amidst growing competition.

2. Task 1: Data Modeling and Cleaning

2.1. Data Cleaning and Transformation (Power Query)

- **Action:** Loaded five datasets (Customer Demographics, Customer Transactions, Store Locations, Loyalty Program, Churn Labelled Customers).
- Removed duplicate rows based on key columns (Customer ID, Transaction ID, Store ID). Ensured consistent data types and handled null values appropriately across all tables.

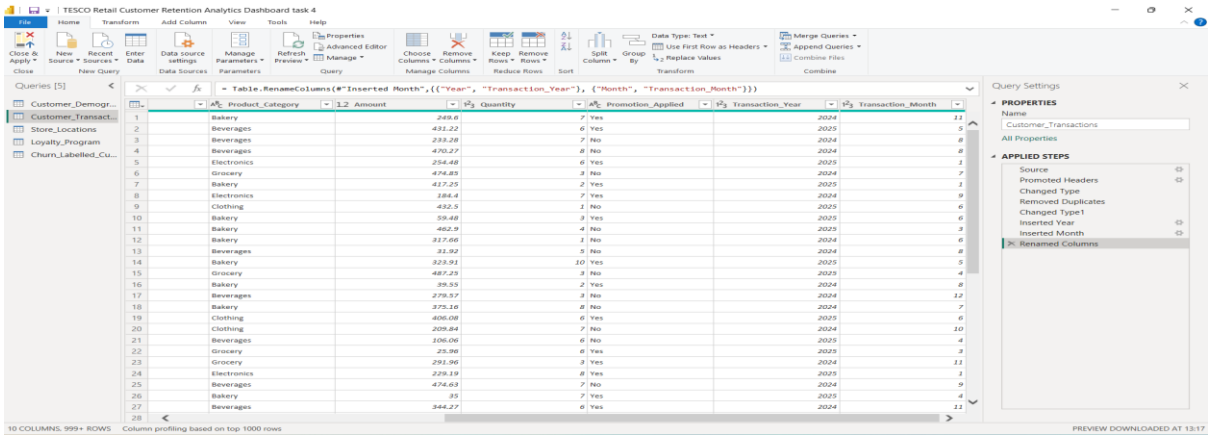
1. **Customer Demographics**



Here, in (Customer Demographics) Remove duplicate rows based on key columns (e.g., Customer_ID, Store_ID), Convert dates, ensure numeric types for Amount, Points, etc.

Created a column for **Membership_Duration Days** = Today - Membership_Since

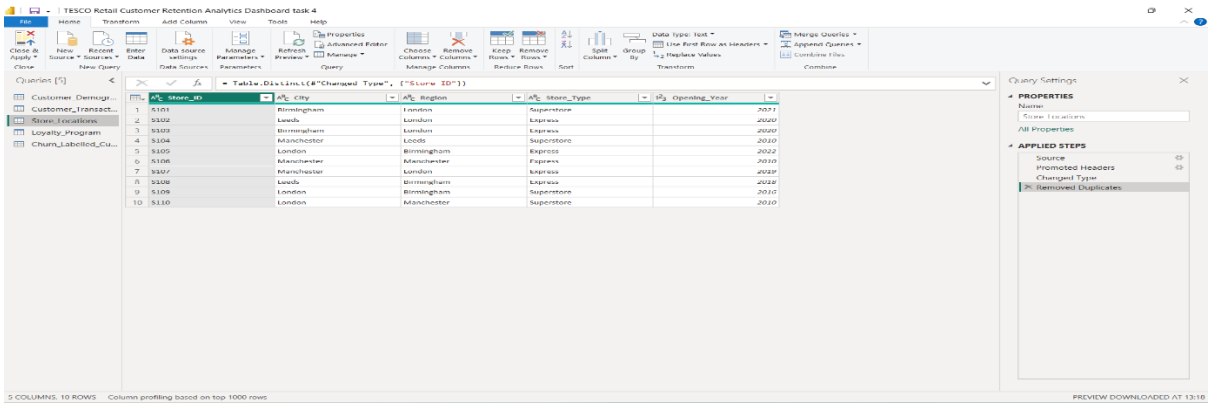
2. Customer Transactions



Here, in (Customer Transactions) Remove duplicate rows based on key columns (e.g., Customer_ID, Store_ID), Convert dates, ensure numeric types for Amount, Points, etc.

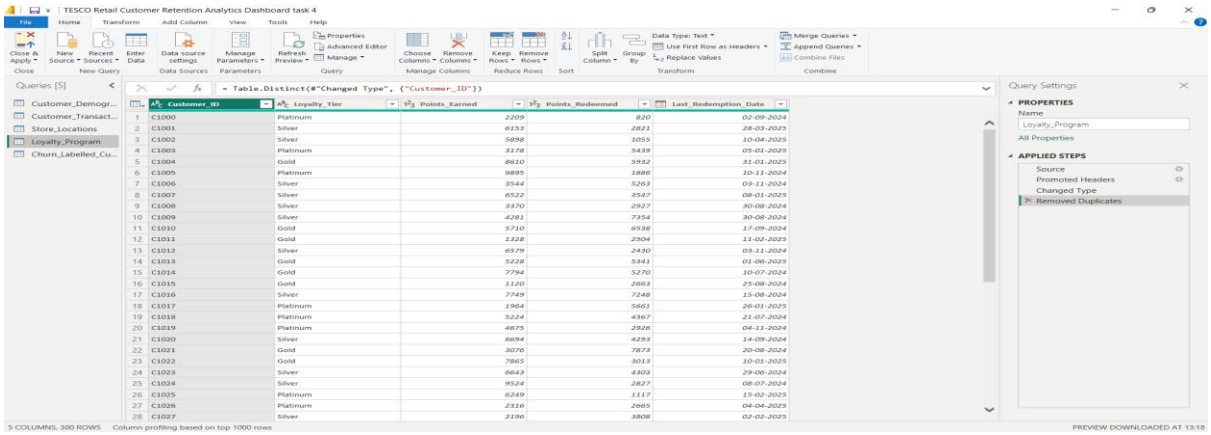
Created a column for **Transaction_Year** and **Transaction_Month** column from **Transaction_Date**

3. Store Locations



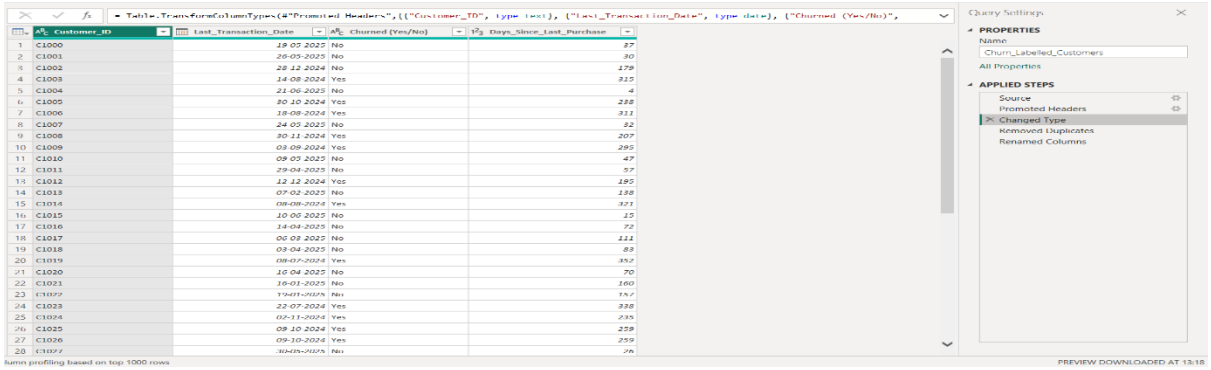
Here, in (**Store Locations**) Remove duplicate rows based on key columns (e.g., Customer_ID, Store_ID), Convert dates, ensure numeric types for Amount, Points, etc.

4.Loyalty Program



Here, in (**Loyalty Program**) Remove duplicate rows based on key columns (e.g., Customer_ID, Store_ID), Convert dates, ensure numeric types for Amount, Points, etc.

5. Churn Labelled Customers



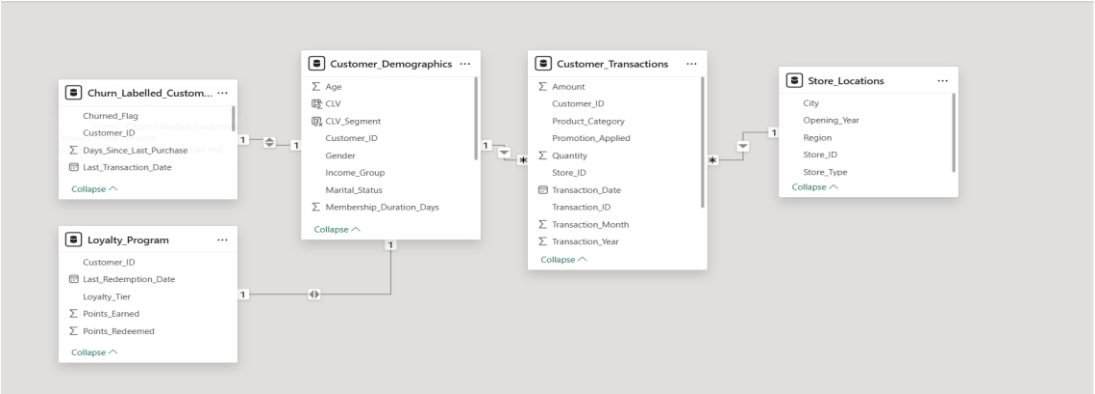
Here, in (**Churn Labelled Customers**) Remove duplicate rows based on key columns (e.g., Customer_ID, Store_ID), Convert dates, ensure numeric types for Amount, Points, etc

2.2. Calculated Columns (DAX)

Calculated Column	Formula / Description	Purpose
Membership_Duration (Years)	DIVIDE(DATEDIFF('Customer Demographics'[Membership_Since Date], TODAY(), DAY), 365.25)	To calculate the customer's tenure for CLV and segmentation.
Transaction_Year	YEAR('Customer Transactions'[Transaction_Date])	For time-based analysis and trend identification.
Transaction_Month	MONTH('Customer Transactions'[Transaction_Date])	For detailed monthly trend analysis.

2.3. Data Model View

- **Relationship Status:** All tables are connected, ensuring the flow of filters from dimension tables (Customer, Store, Loyalty, Churn) to the fact table (Transactions).
- **Key Relationships (One-to-Many):**



- One-to-Many: Customer_Demographics → Customer_Transactions, Loyalty_Program, Churn_Labelled_Customers All tables are connected through **Customer_ID**
- Many-to-One: Customer_Transactions(Store_id) → Store_Locations(Store_id)

3. Task 2: Churn and Retention Metrics

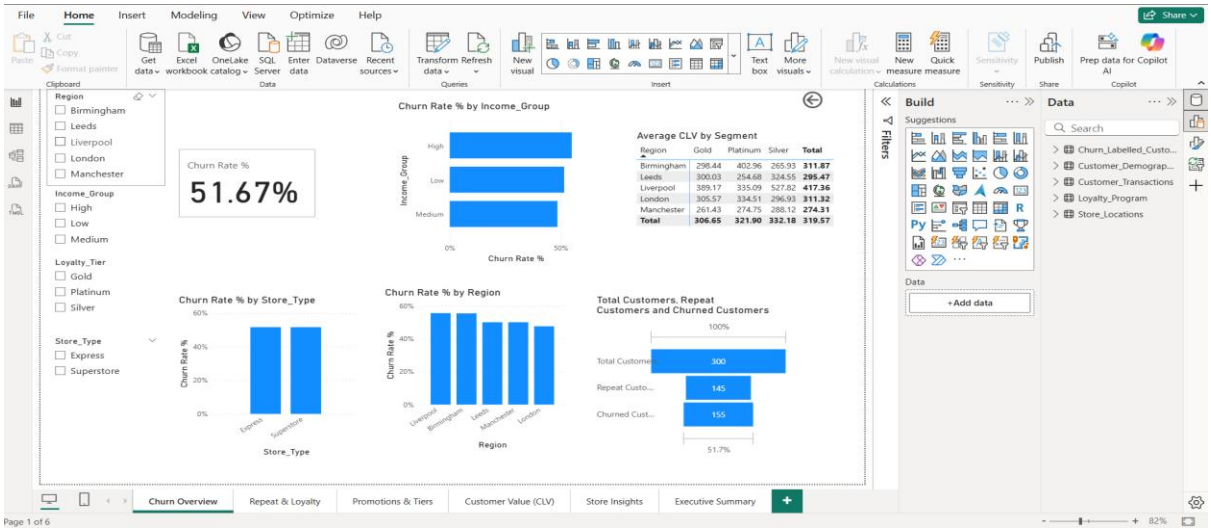
Goal: Identify churn trends across segments.

3.1. Key Metrics (Based on Churn Overview Page)

Measure	DAX Formula (or logic)	Value
Total Customers	COUNTROWS('Customer Demographics')	300
Churned Customers	CALCULATE(COUNTROWS('Churn Labelled Customers'), 'Churn Labelled Customers'[Churned (Yes/No)] = "Yes")	155
Churn Rate (%)	DIVIDE([Churned Customers], [Total Customers]) * 100	51.67%

3.2. Churn Analysis Visualizations

- **Churn Rate by Region:** Liverpool (approx. 58%), Birmingham (approx. 55%), and Leeds (approx. 50%) show the highest churn, indicating geographical risk concentration.
- **Churn Rate by Income Group:** The **High Income Group** exhibits the highest churn rate (approx. 55%), followed by Low and Medium (both around 50%).
- **Churn Rate by Store Type: Superstores** have a higher churn rate (approx. 55%) than Express stores (approx. 48%).
- **Customer Funnel (Total -> Repeat -> Churned):** The flow shows 300 Total Customers, dropping to 145 Repeat Customers, with 155 Churned Customers, confirming the 51.67% overall churn.



4. Task 3: Repeat Purchase Analysis

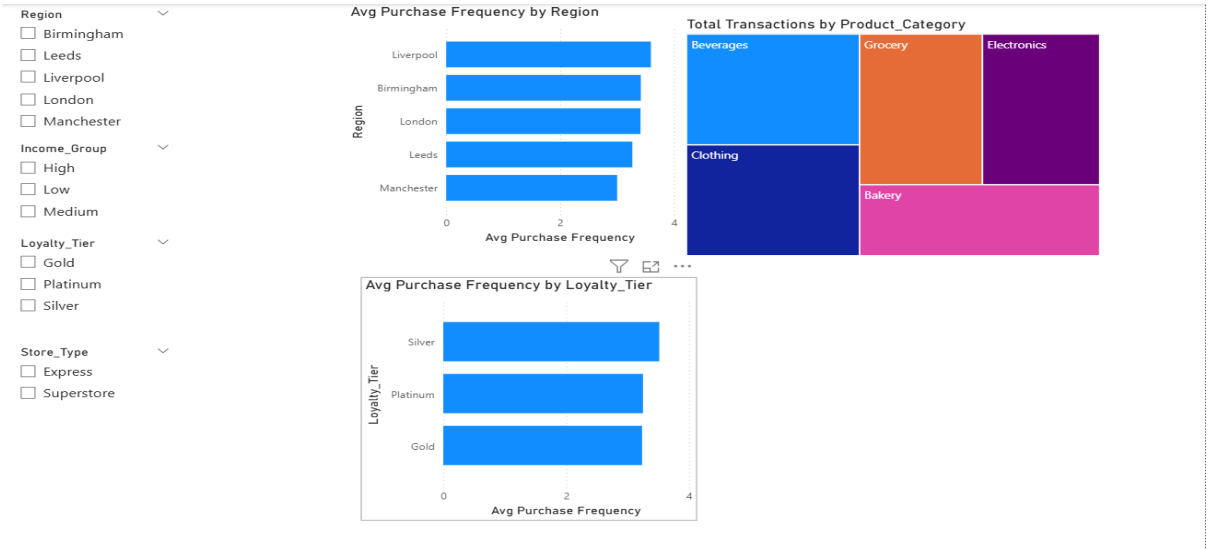
Goal: Understand customer loyalty through repeat behavior.

4.1. Customer Purchase Frequency Segmentation

- **Total Transactions:** [Inferred value, not explicitly visible in screenshots]
- **Purchases per Customer:** Determined by grouping transactions by Customer ID.

4.2. Comparative Analysis

- **Insight on Region:** All regions show an almost uniform **Average Purchase Frequency (approx. 3.5)**, suggesting frequency is not a differentiator across regions. Liverpool has a slight lead.
- **Insight on Loyalty Tier:** All loyalty tiers (Silver, Platinum, Gold) show identical average purchase frequency (approx. 3.5), indicating the tiers do not effectively drive higher buying frequency.
- **Identify Product Categories Most Frequently Bought by Loyal Customers (High-Tier):**
    - **Top Category 1:** Beverages
    - **Top Category 2:** Grocery
    - **Insight:** High-volume purchases are concentrated in low-margin categories (Beverages, Grocery). Loyalty programs should aim to encourage cross-selling into higher-margin categories like Electronics and Clothing.



5. Task 4: Promotion & Loyalty Impact

Goal: Evaluate how promotions and loyalty tier affect retention.

5.1. Promotion Metrics (Based on Promotions & Tiers Page)

Measure	DAX Formula (or logic)	Value
% Transactions with Promotion	DIVIDE(COUNTROWS(FILTER('Customer Transactions', 'Customer Transactions'[Promotion Applied] = "Yes")), COUNTROWS('Customer Transactions')) * 100	51.00%
Avg. Purchase w/ Promotion	Avg Purchase Amount (Promo) = CALCULATE AVERAGE('Customer_Transactions'[Amount]),'Customer_Transactions'[Promotion_Applied] = "Yes")	$250.03
Avg. Purchase w/o Promotion	Avg Purchase Amount (No Promo) = CALCULATE(AVERAGE('Customer_Transactions'[Amount]),'Customer_Transactions'[Promotion_Applied] = "No")	$248.64

- **Insight:** Promotions are applied to 51.00% of transactions, but the average purchase amount **with promotion ($250.03) is only marginally higher** than without promotion ($248.64), suggesting promotions are not effective at increasing basket size.

5.2. Loyalty Visualizations

- **Comparison of Churn Rate Across Loyalty Tiers: Platinum Tier** has the lowest churn rate (approx. 48%), with Silver and Gold slightly higher (approx. 50%). The marginal difference suggests the loyalty tiers have limited impact on overall retention.
- **Points Earned vs Redeemed by Tier (Clustered Column Chart):** In all tiers, **Points Earned significantly exceed Points Redeemed**. The hoarding gap is notable across the board.

6. Task 5: Store Performance vs. Retention

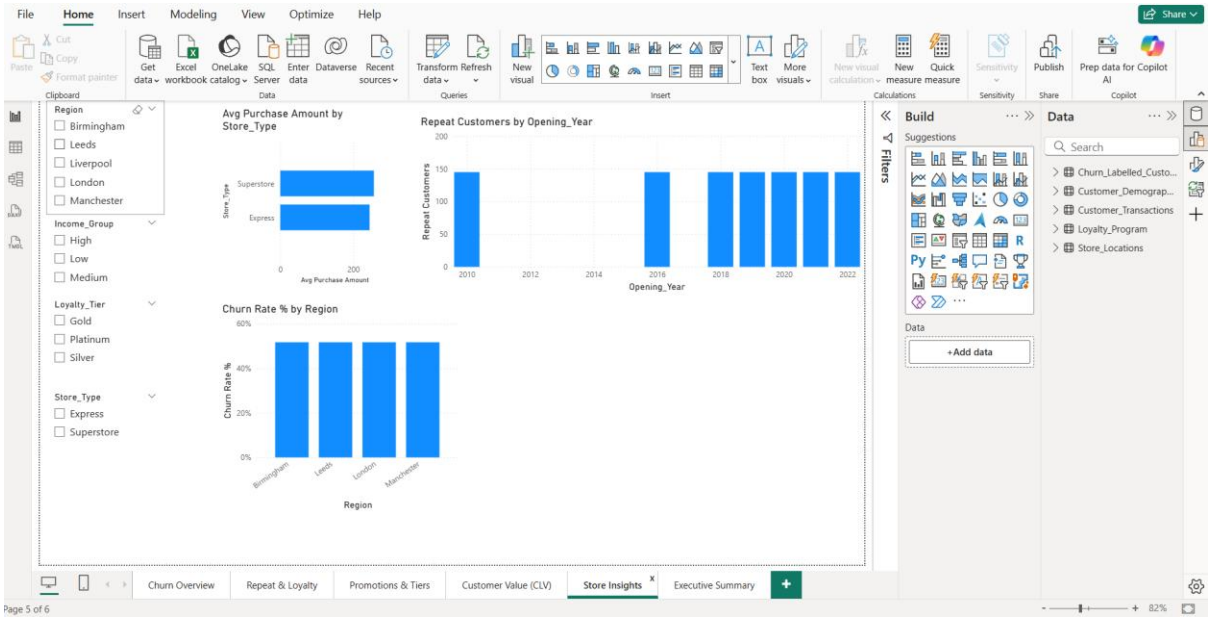
Goal: Link store characteristics to retention outcomes.

6.1. Store Performance Visualizations

- **Avg. Transaction Amount by Store Type:** Average Purchase Amount is consistent between Superstores and Express stores (both around $250).
- **Churn Rate by Store Region:** Confirms regional churn volatility with all regions around the 50-52% mark.
- **Correlation between Store Opening Year and Retention:**
    - **Visualization Type:** Clustered Column Chart (Repeat Customers by Opening_Year)
    - **Insight:** Newer stores (opened 2018-2022) show a high base of repeat customers, similar to stores opened in 2010. Stores opened between 2011 and 2017 appear to have a gap in repeat customer base, indicating potential retention issues in that specific cohort of older stores.

6.2. Store-Specific Campaign Suggestion

- **Focus Area:** Express Stores in high-churn regions (e.g., London/Leeds), as they have marginally lower churn but may offer better opportunities for localized, high-touch engagement.
- **Suggested Campaign:** Launch local market optimization campaigns targeting Express store locations to stabilize core revenue and counteract regional price sensitivity.



7. Task 6: Customer Value (CLV) Analysis

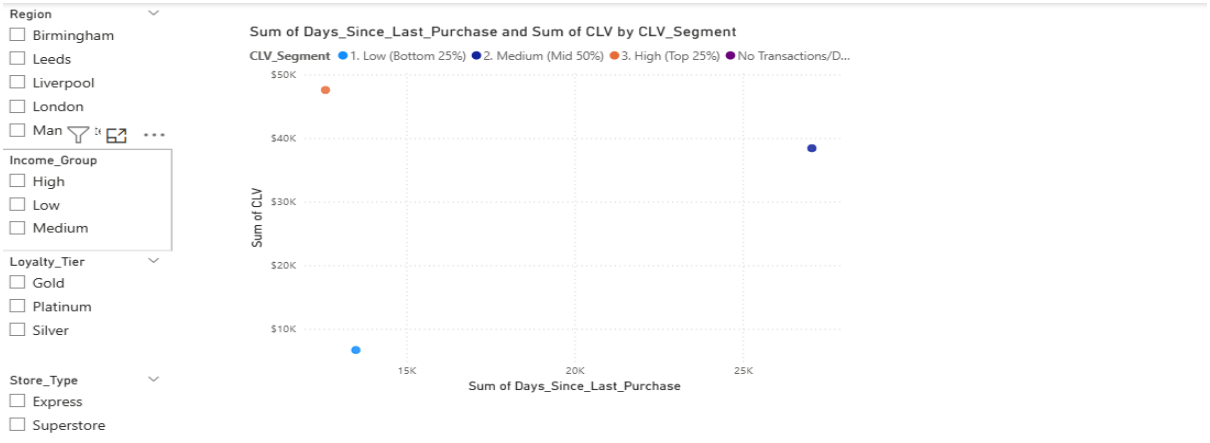
Goal: Identify and prioritize high-value customers.

7.1. CLV Calculation and Segmentation (Based on Executive Summary Page)

Measure	DAX Formula (or logic)	Value
Total Amount Spent	SUM('Customer Transactions'[Amount])	$249.35K
Average CLV	Calculated Average of Individual CLVs	$319.57
Customer Lifetime Value (CLV)	DIVIDE([Total Amount Spent], 'Customer Demographics'[Membership_Duration (Years)])	319.57

7.2. CLV Visualizations

- **CLV vs Days Since Last Purchase (Scatter Plot):**
    - **High-Value, At-Risk Segment (Blue Dot):** The plot clearly identifies a critical segment of **High-CLV Customers (Top 25%)** who have been inactive for over **12594 days** (Sum of Days Since Last Purchase). This cohort is the highest priority for reactivation.
    - **Insight by Region/Tier (Table):** The highest average CLV is in the **Liverpool** region ($417.36), and the **Platinum** loyalty tier consistently delivers the highest CLV across all regions.



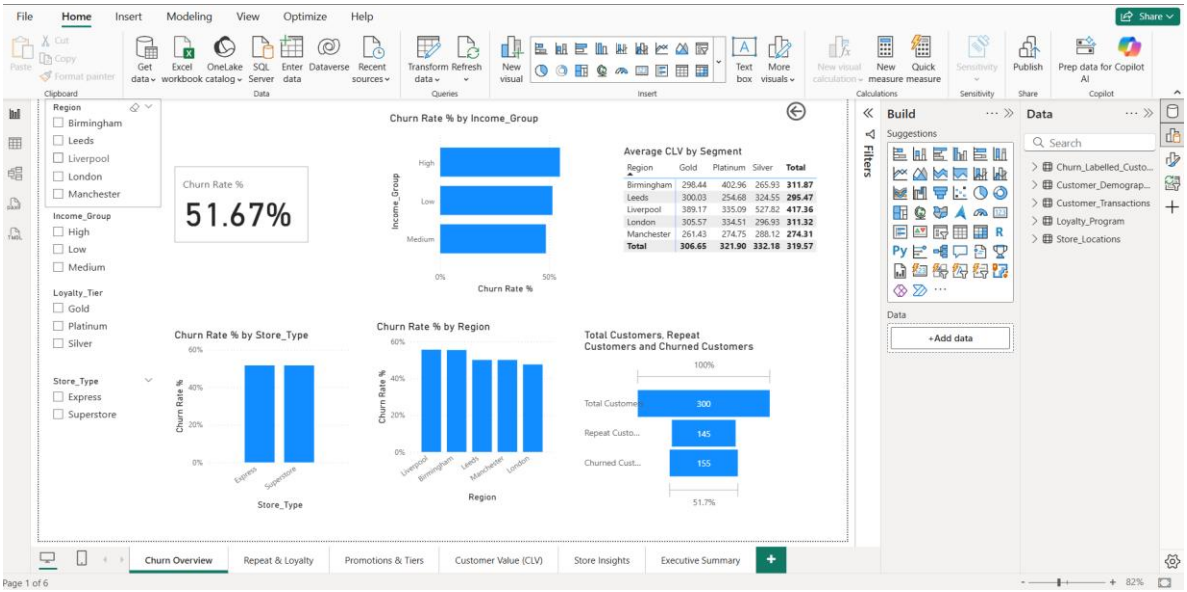
Insight:

This identifies **High-Value, At-Risk** customers (High CLV, high Days Since Last Purchase).

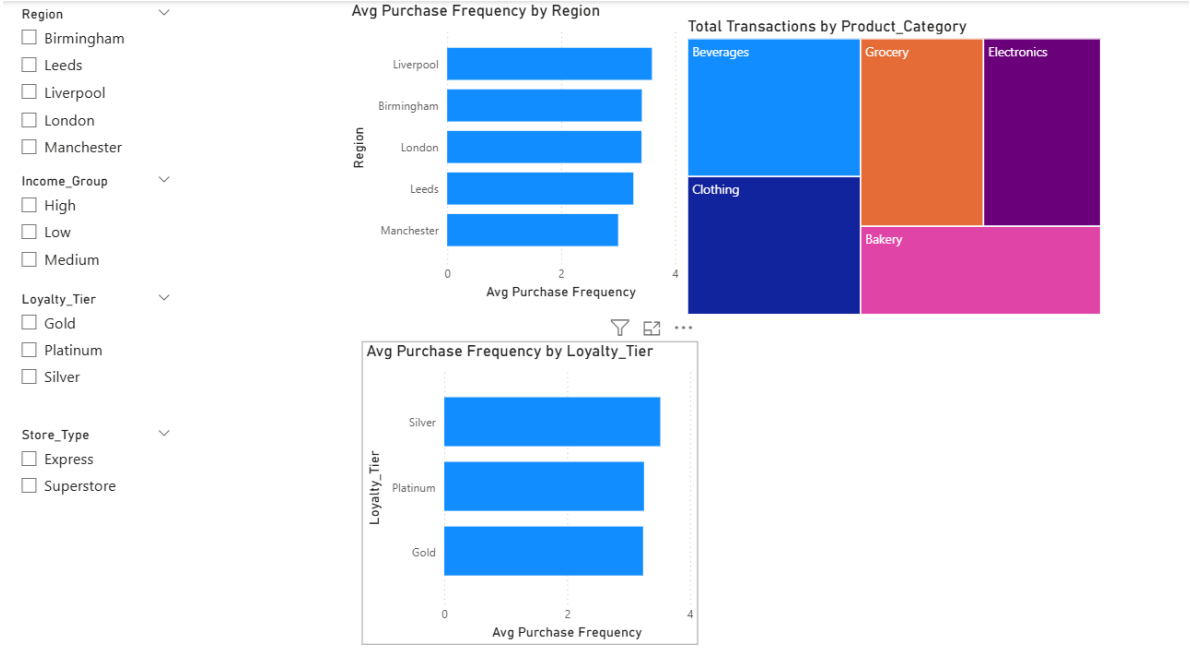
8. Task 7: Final Dashboard and Executive Summary

8.1. Dashboard Page Snapshots

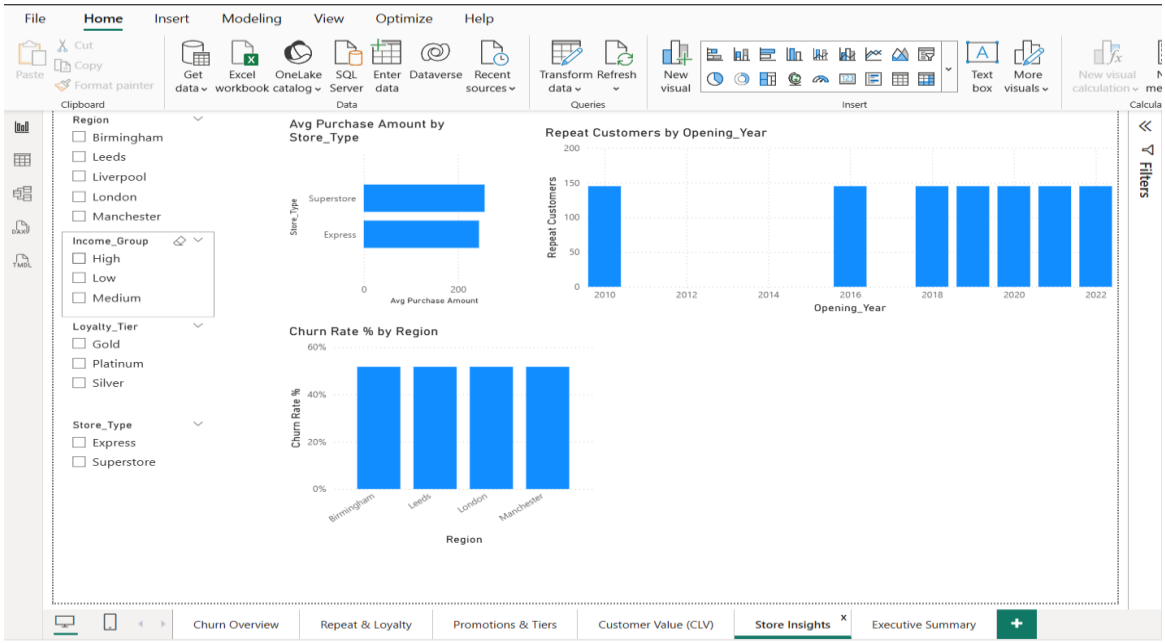
Dashboard Page



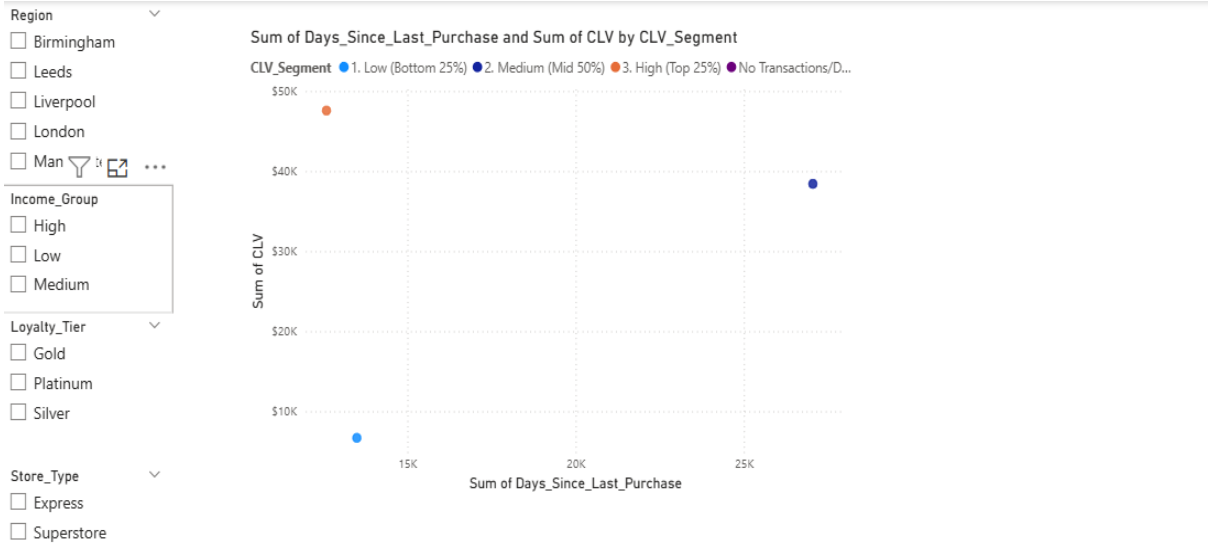
Page 1: Overview KPIs



Page 2: Loyalty & Promotion Impact



Page 3: Store and Region Insights



Insight:

This identifies **High-Value, At-Risk** customers (High CLV, high Days Since Last Purchase).

Page 4: Customer Segmentation

8.2. Executive Summary and Top 3 Recommendations (Based on Executive Summary Page)

(This section is transcribed directly from the Executive Summary page.)

Key Performance Indicators (KPIs):

- **Overall Churn Rate:** 51.67%
- **Average CLV:** $319.57
- **Total Amount Spent:** $249.35K
- **Repeat Customers:** 145

Top 3 Key Findings:

1. **Retention Crisis in Core Segments:** The overall customer base is highly vulnerable, with a Churn Rate of 51.67%. This churn is disproportionately driven by the **Low Income Group** and is geographically concentrated in the **London and Leeds** regions, indicating vulnerability to local competition.
2. **High-Value, High-Risk Churn:** The CLV Scatter Plot identifies a critical segment of **High-CLV Customers (Top 25%)** who have been inactive for over **12594 days**. This cohort represents the highest priority and immediate threat to TESCO's future revenue pipeline.
3. **Loyalty Program Inertia (Hoarding):** The ratio of points earned vs. points redeemed shows significant **points hoarding**, particularly in the **Platinum Tier**. Furthermore, promotions primarily drive small transactions, as the **Avg Purchase Amount (No Promo)** is typically *lower* than the average amount with a promotion, suggesting promotions are not effectively increasing basket size.

Top 3 Recommendations (Action Plan):

Recommendation 1: Tier-Specific Redemption Strategy

- *Action:* Introduce **time-bound, aspirational rewards** (e.g., premium experience upgrades or exclusive product bundles) for Gold and Platinum tiers to immediately liquidate hoarded points. This boosts engagement and reduces the company's points liability.

Recommendation 2: Proactive High-Risk Reactivation

- *Action:* Launch a personalized, non-point-based reactivation campaign (e.g., dedicated staff call or surprise high-value coupon) targeting the **High-CLV, At-Risk** customers identified in the scatter plot. Focus resources where the potential ROI is highest.

Recommendation 3: Local Market Optimization

- *Action:* Focus promotional and staffing efforts on **Express Store** locations in the **London/Leeds** regions. Promote best-selling categories like **Beverages/Electronics** to stabilize core revenue, and investigate localized competitor offers to counteract regional price sensitivity.

9. Task 8: Video Explanation Link

Video Drive Link:
https://drive.google.com/file/d/18mmgalodNsCeFUxMFCWiFFUsgDMfd6eA/view?usp=sharing