Retail Customer Retention Analytics – IKEA

Task 1: Data Modeling and Cleaning

Used Power Query Editor to load and transform the datasets.

Removed duplicate rows based on key columns in general most of the data was clean

Formatted columns: Convert dates, ensure numeric types for Amount, Points, etc. Handle missing or null values appropriately (e.g., filter or replace)

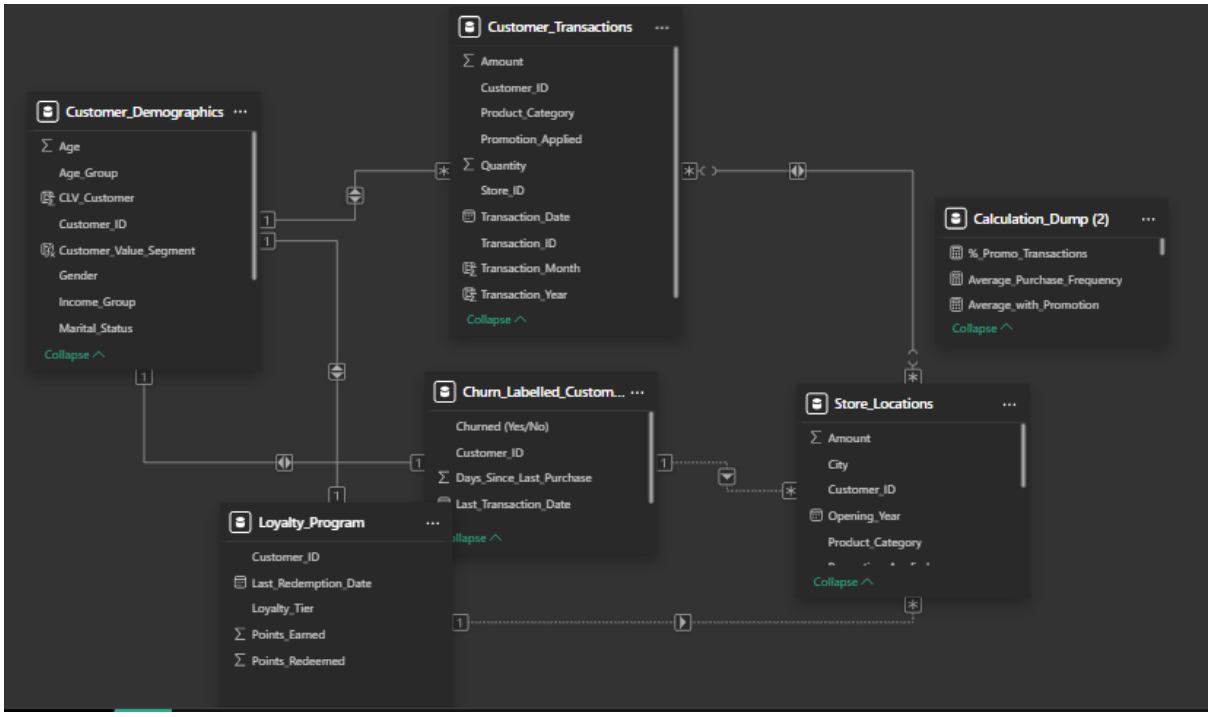
Created calculated columns:

Membership_Duration = TODAY() - Customer_Demographics[Membership_Since].[Date]

Transaction_Month = MONTH(Customer_Transactions[Transaction_Date])

Transaction_Year = YEAR(Customer_Transactions[Transaction_Date])

Basic Model View



Created Basic Model View using common key between the tables.

One-to-Many and One-to-One in the above data model

For Different Measure I have created a new table named Calculation dump. All measures where simple DAX function are used have been created in this model

Task 2: Churn and Retention Metrics

Create a Churn Rate card: (Churned Customers / Total Customers) * 100

Created a measure Named **Churn Rate** and the formula used is :

```
Churn_Rate = SUMX(Churn_Labelled_Customers,IF(Churn_Labelled_Customers[Churned (Yes/No)]="Yes",1,0))/COUNTA(Churn_Labelled_Customers[Customer_ID])
```

Visualized churn rate by:

Region : Used Clustered Column Chart

X – Axis : Region Y – Axis: Churn_Rate

Income Group: Used Clustered Column Chart

X – Axis : Income_Group Y – Axis: Churn_Rate

Store Type (via store join): Used Donut Chart

Legend: Store_Type Values: Churn_Rate

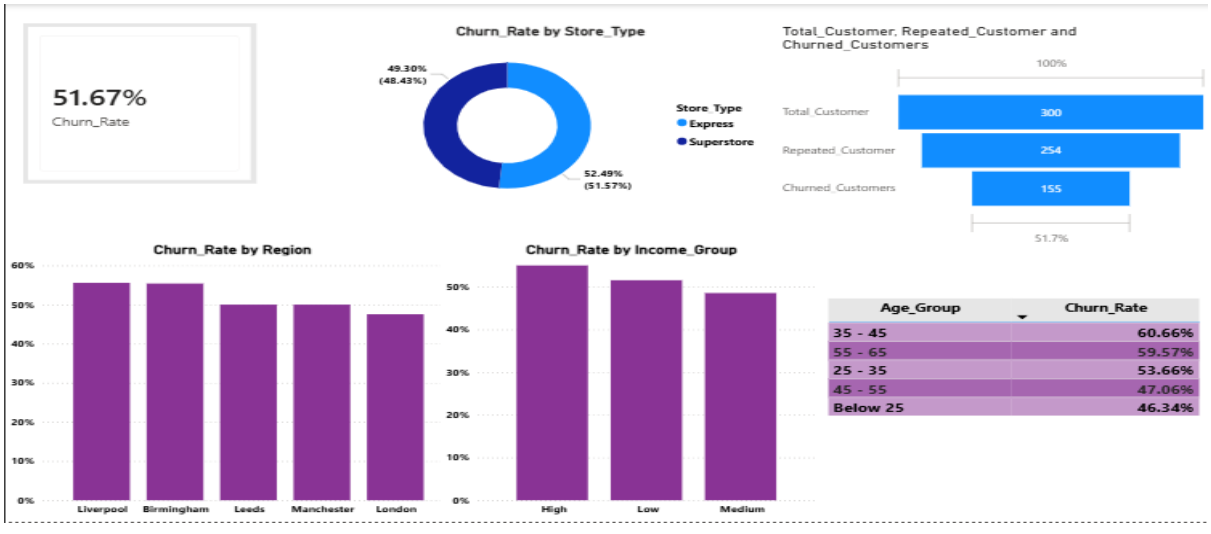
Identify top 5 segments with highest churn % : Used Table

Columns: Age_Group , Churn_Rate and Applied filter for Top 5 in filter pane

Created funnel chart: Total Customers → Repeat → Churned

Measures Created :

```
1. Total_Customer = DISTINCTCOUNT(Customer_Demographics[Customer_ID])
2. Repeated_Customer =
   CALCULATE(DISTINCTCOUNT(Customer_Transactions[Customer_ID]),FILTER(VALUES(Customer_Transactions[Customer_ID]),CALCULATE(COUNT(Customer_Transactions[Customer_ID]))>1))
3. Churned_Customers =
   CALCULATE(DISTINCTCOUNT(Churn_Labelled_Customers[Customer_ID]),Churn_Labelled_Customers[Churned (Yes/No)] = "Yes")
```



Task 3: Repeat Purchase Analysis

Created a measure:

Low-Tier Customers: (2-4 purchases)

```
Low_Tier_Customer = CALCULATE( DISTINCTCOUNT(Customer_Transactions[Customer_ID]), FILTER(
VALUES(Customer_Transactions[Customer_ID]), [Total_Purchase] >= 2 && [Total_Purchase] <= 4
))
```

Mid-Tier Customers: (5-10 purchases)

```
Mid_Tier_Customer = CALCULATE( DISTINCTCOUNT(Customer_Transactions[Customer_ID]), FILTER(
VALUES(Customer_Transactions[Customer_ID]), [Total_Purchase] >= 5 && [Total_Purchase] <= 10
))
```

High-Tier Customers: (11+ purchases)

```
High_Tier_Customer = CALCULATE( DISTINCTCOUNT('Customer_Transactions'[Customer_ID]),
FILTER( VALUES('Customer_Transactions'[Customer_ID]), [Total_Purchase] >= 11 ))
```

Compared avg. purchase frequency by:

Region: Used Clustered Column Chart

X – Axis : Region Y – Axis: Average_Purchase_Frequency

Age Group: Used Clustered Column Chart

X – Axis : Age_Group Y – Axis: Average_Purchase_Frequency

Loyalty Tier: Used Clustered Bar Chart

X-Axis : Average_Purchase_Frequency Y – Axis: Loyalty_Tier

Identify product categories most frequently bought by loyal customers :

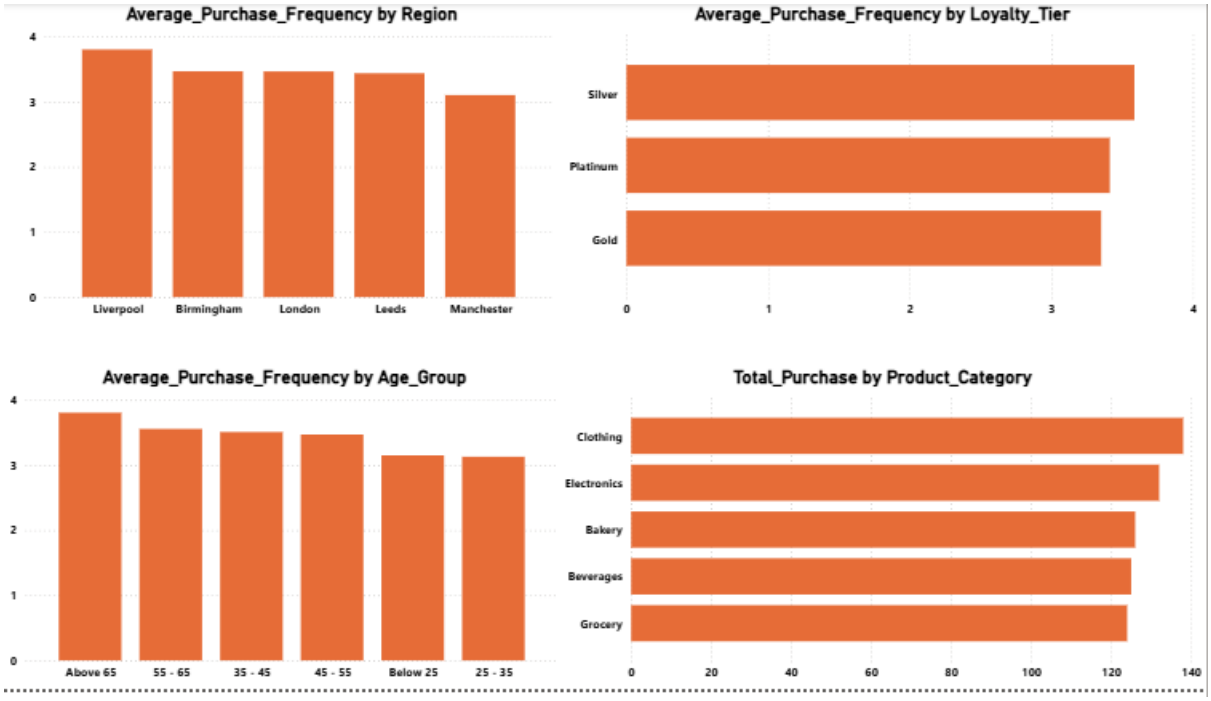
For that I have created a measure

```
Total_Purchase = DISTINCTCOUNT(Customer_Transactions[Transaction_ID])
```

to count Total Purchase

Used Clustered Bar Chart

X – Axis : Total_Purchase Y – Axis : Product_Category



Task 4: Promotion & Loyalty Impact

Created Measure:

% of transactions where promotion was applied -

```
%_Promo_Transactions =
DIVIDE(CALCULATE(COUNT(Customer_Transactions[Transaction_ID]),Customer_Transactions[Promotion_Applied] = "Yes"),COUNT(Customer_Transactions[Transaction_ID]))
```

Avg. purchase amount with vs without promotion –

```
Average_with_Promotion =
CALCULATE(AVERAGE(Customer_Transactions[Amount]),Customer_Transactions[Promotion_Applied] = "Yes")
```

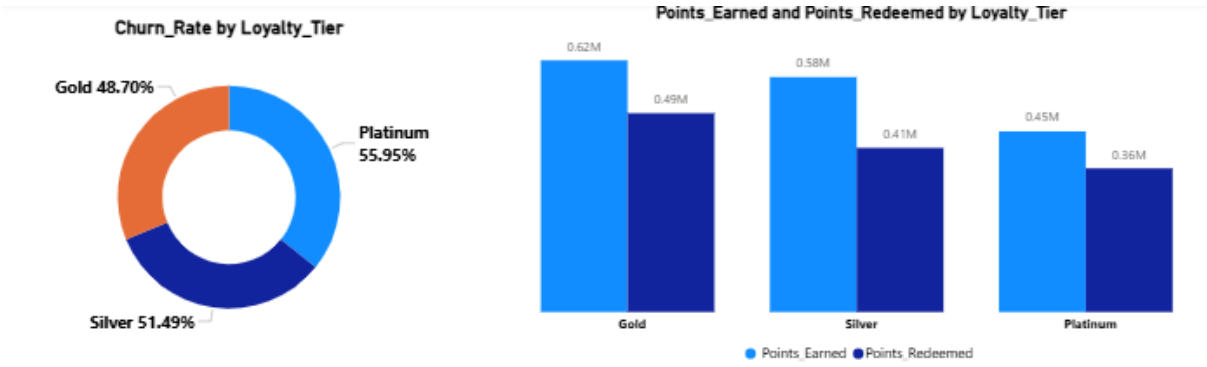
```
Avg_Without_Promo =
CALCULATE(AVERAGE(Customer_Transactions[Amount]),Customer_Transactions[Promotion_Applied] = "No")
```

Compare churn rate across loyalty tiers – Used Donut Chart

Legend : Loyalty_tier Values : Churn_rate

Show Points Earned vs Redeemed by Tier using clustered column chart -

X – Axis : Loyalty_Tier Y – Axis : Points_Earned , Points_Redeemed



Despite being the highest loyalty tier Platinum shows highest churn rate, indicating they are getting less reward, Experiential reward and tier protection strategies should be applied to improve retention.
As for Silver Level Customer they have low redemption rate so there should be bonus points on first redemption and simple rewards should be given

Task 5: Store Performance vs Retention

Merge Store_ID from transactions with Store_Locations : Merged Store_Id from Transaction with Store_Location using Power Query Editor by Click on Transform Data tab in Power BI then Using the Merge functionality in Power Query Editor.

Visualize:

Avg. transaction amount by store type : Used Donut Chart

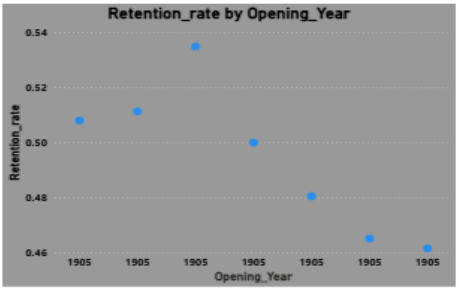
Legend : Store_Type Values : Average of Amount

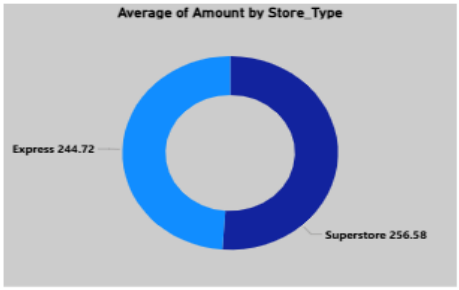
Churn rate by store region : Used Donut Chart

Legend : Region Values : Churn_Rate

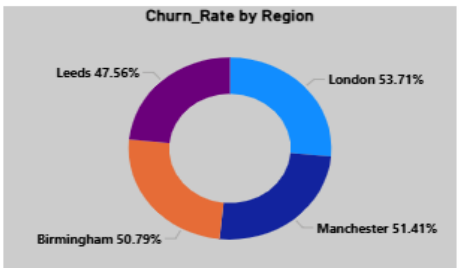
Correlation between store opening year and retention : Used Scatter Chart

X – Axis : Opening_Year Y – Axis : Retention_Rate





London and Manchester are having highest churn rate . But the store specific campaign should be run in every region because there is not much gap between churn rate of every region and should focus more in these region



Task 6: Customer Value (CLV) Analysis

Calculate CLV = Total Amount Spent / Membership Duration (in years)

Calculated Two CLV's , Total CLV and CLV by Customer

```
CLV =
DIVIDE(CALCULATE(SUM(Customer_Transactions[Amount]),ALLEXCEPT(Customer_Transactions,Custome
r_Transactions[Customer_ID])),AVERAGE(Customer_Demographics[Membership_Duration])/365)
```

```
CLV_Customer =
VAR TotalSpend =
    CALCULATE(
        SUM('Customer_Transactions'[Amount]),
        FILTER(
            'Customer_Transactions',
            'Customer_Transactions'[Customer_ID]
                = 'Customer_Demographics'[Customer_ID]
        )
    )

VAR MembershipYears =
    'Customer_Demographics'[Membership_Duration] / 365

RETURN
IF(
    ISBLANK(TotalSpend)
        || ISBLANK(MembershipYears)
        || MembershipYears = 0,
    0,
    DIVIDE(TotalSpend, MembershipYears)
)
```

Segmented customers into:

Low (Bottom 25%) , Medium (Mid 50%) and High (Top 25%)

Created a Column **Customer_Value_Segmentation** using formula :

```
Customer_Value_Segment =
VAR CustomerRank =
    RANKX(
        ALL('Customer_Demographics'),
        'Customer_Demographics'[CLV_Customer],
        ,
        ASC,
        DENSE
    )
```

```
VAR TotalCustomers =
    COUNTROWS(
        ALL('Customer_Demographics')
    )

VAR RankPercent =
    DIVIDE(CustomerRank, TotalCustomers)

RETURN
IF(
    RankPercent <= 0.25,
    "Low (Bottom 25%)",
    IF(
        RankPercent <= 0.75,
        "Medium (Mid 50%)",
        "High (Top 25%)"
    )
)
```

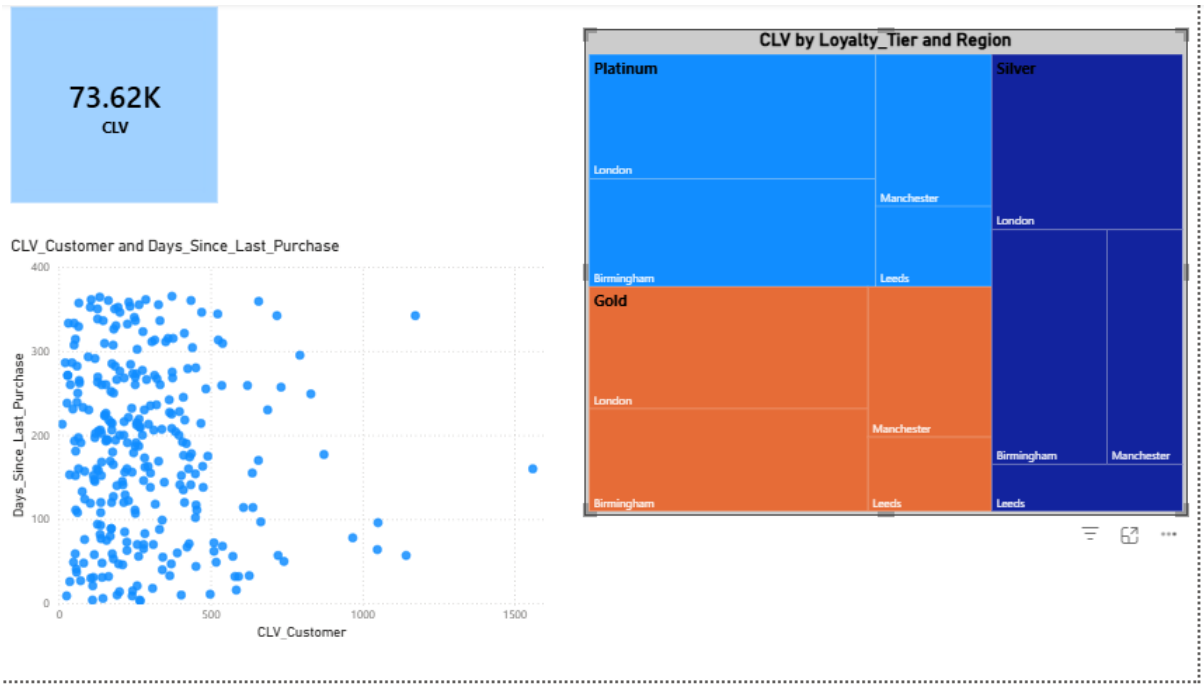
Visualize:

CLV vs Days Since Last Purchase : Used Scatter Plot

X – Axis : CLV_Customer Y – Axis : Days_Since_Last_Purchase

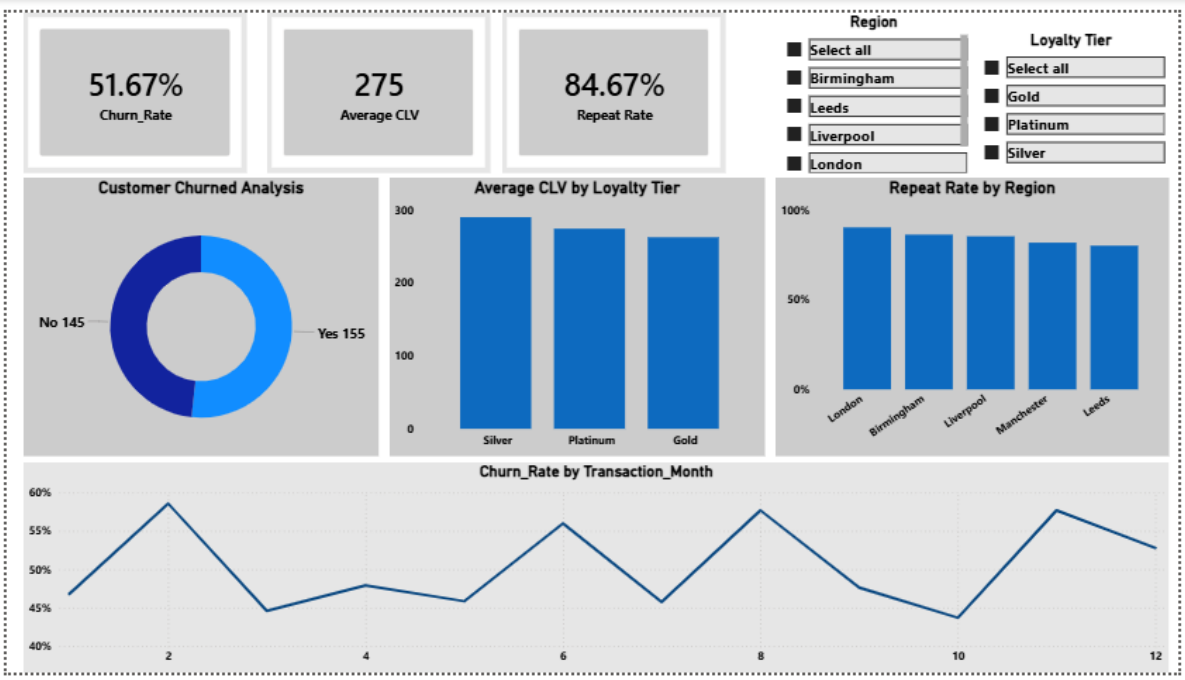
CLV by Loyalty Tier and Region : Used Treemap

Category - Loyalty Tier Details – Region Values – CLV

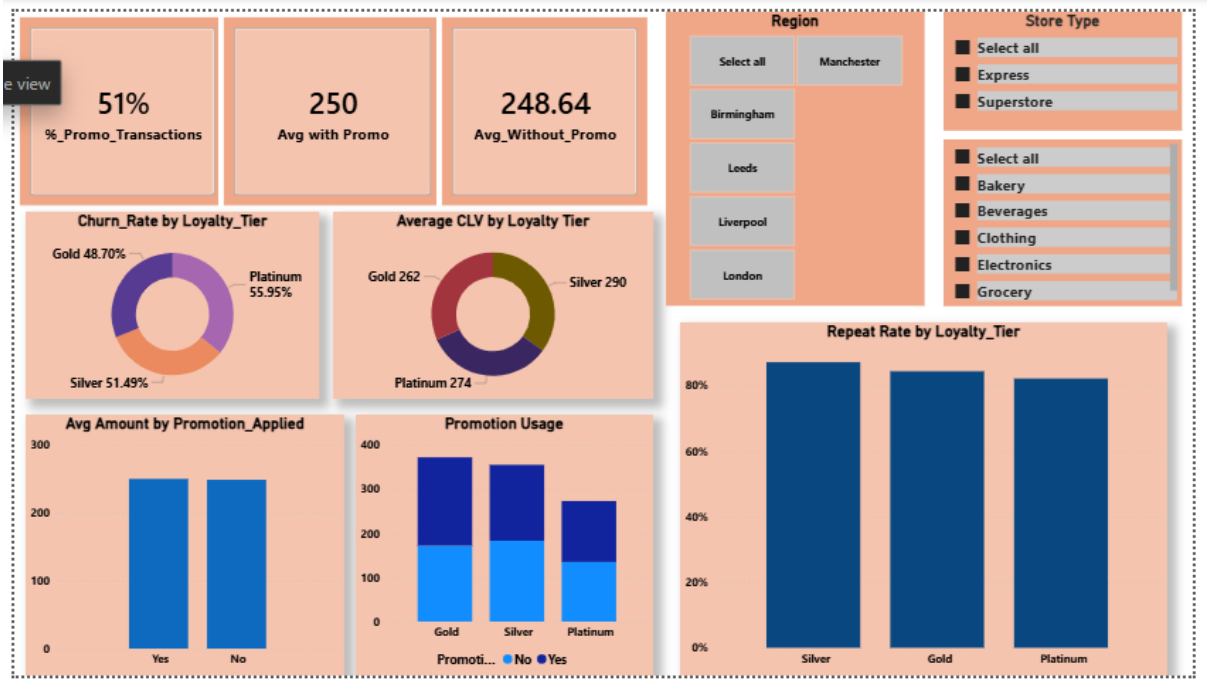


Task 7: Final Dashboard and Executive summary

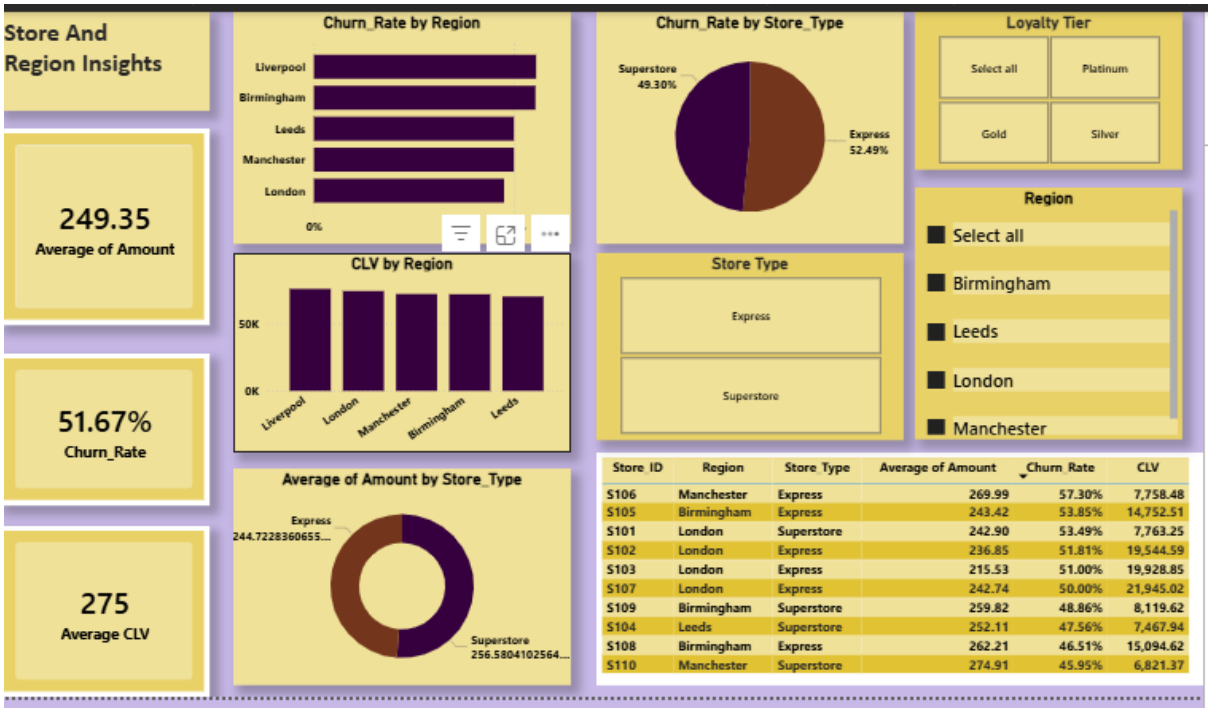
Page 1: Overview KPIs (Churn, CLV, Repeat Rate)



Page 2: Loyalty & Promotion Impact

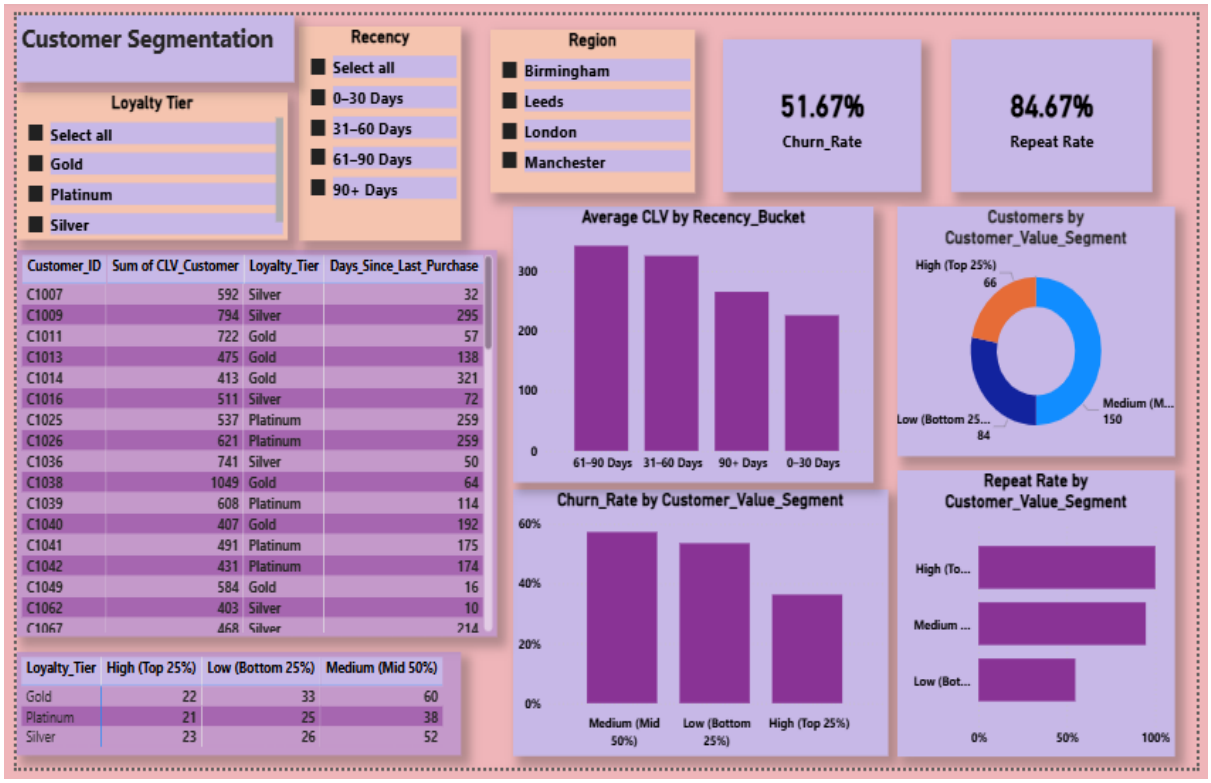


Page 3: Store and Region Insights



Store_ID	Region	Store_Type	Average of Amount	Churn_Rate	CLV
S106	Manchester	Express	269.99	57.30%	7,758.48
S105	Birmingham	Express	243.42	53.85%	14,752.51
S101	London	Superstore	242.90	53.49%	7,763.25
S102	London	Express	236.85	51.81%	19,544.59
S103	London	Express	215.53	51.00%	19,928.85
S107	London	Express	242.74	50.00%	21,945.02
S109	Birmingham	Superstore	259.82	48.86%	8,119.62
S104	Leeds	Superstore	252.11	47.56%	7,467.94
S108	Birmingham	Express	262.21	46.51%	15,094.62
S110	Manchester	Superstore	274.91	45.95%	6,821.37

Page 4: Customer Segmentation (Churned, Repeat, High-Value)



Customer_ID	Sum of CLV_Customer	Loyalty_Tier	Days_Since_Last_Purchase
C1007	592	Silver	32
C1009	794	Silver	295
C1011	722	Gold	57
C1013	475	Gold	138
C1014	413	Gold	321
C1016	511	Silver	72
C1025	537	Platinum	259
C1026	621	Platinum	259
C1036	741	Silver	50
C1038	1049	Gold	64
C1039	608	Platinum	114
C1040	407	Gold	192
C1041	491	Platinum	175
C1042	431	Platinum	174
C1049	584	Gold	16
C1062	403	Silver	10
C1067	468	Silver	214

Loyalty_Tier	High (Top 25%)	Low (Bottom 25%)	Medium (Mid 50%)
Gold	22	33	60
Platinum	21	25	38
Silver	23	26	52

Create a card or textbox summarizing top 3 recommendations:

- **What should IKEA do to retain more customers?**
- **Where should they focus next?**

1. Strengthen Retention for Medium & Low-Value Customers (High Churn Risk)

Medium and low customer value segments show the highest churn, especially among Expre ss and Super Store shoppers aged 25–55.

Action:

Introduce personalized discounts, bundle offers, and targeted loyalty rewards for these seg ments.

Push Silver loyalty tier upgrades with spend-based incentives to reduce churn.

2.Optimize Express Stores in High-Churn Regions

Express stores in Liverpool, Leeds, Manchester, London, and Birmingham consistently show higher churn across multiple age groups.

Action:

Improve in-store experience, checkout speed, and product availability in Express formats.

Run region-specific promotions focused on high-churn age groups (Below 25, 25–35, 45–55, 55–65)

3.Leverage High-Purchase & High-CLV Segments for Growth

Customers aged 65+, Silver loyalty tier, and Clothing category buyers show higher purchase f requency and CLV, especially within 31–90 day recency buckets.

Action:

Focus re-engagement campaigns on 31–90 day inactive customers.

Crosssell Electronics and Grocery to frequent Clothing buyers to increase basket size and rep eat rate.

Video Presentation:

https://drive.google.com/file/d/1HlZwwwj_3hR8LDvL5QSqeiKsih-ZzLxY/view?usp=drive_link