E-COMMERCE CHURN RATE

E-COMMERCE

Abhishek Rana(2401010021)

Aayush Aryan(2401010118)

Kush Puri(2401010246)

MD Sajjan(2401010272)

Kaustubh Hiwanj(2401010217)

Kushagra Masheshwari(2401010248)

PROBLEM

Customer churn significantly affects revenue stability and long-term growth. The organization observed fluctuations in customer retention across various countries and cities but lacked structured insights to identify:

- High-churn geographies

- Patterns influencing customer drop-off

- The relationship between sales performance and churn behavior

A data-driven analysis was required to understand churn distribution geographically and support targeted retention strategies.

APPROACH

We analyzed the customer dataset to evaluate churn patterns country-wise and city-wise. The study included:

- Segmentation of customers by geography

- Comparison of churn rates across countries and cities

- Evaluation of sales performance and order behavior

- Assessment of cart abandonment trends

We calculated key performance indicators (KPIs) to measure overall business health and customer behavior.

KEY INSIGHTS

- The overall churn rate of **28.70%** indicates that nearly 3 out of 10 customers discontinue engagement.

- A high **cart abandonment rate (57.09%)** suggests potential friction in the purchasing journey.

- Certain countries and cities exhibit significantly higher churn rates, indicating regional performance gaps.

- Despite strong total sales (14.4M), retention inefficiencies may be limiting long-term revenue growth.

KEY RECOMMENDATIONS

1. Implement targeted retention campaigns in high-churn countries and cities.

2. Optimize checkout experience to reduce cart abandonment.

3. Introduce loyalty or reward programs to improve repeat purchases.

4. Conduct region-specific customer feedback analysis to identify localized issues.

5. Monitor churn and abandonment KPIs monthly to track improvement trends.

PROBLEM STATEMENT

Formal Problem Definition

Customer churn represents a critical challenge affecting revenue sustainability and customer lifetime value. The organization lacks detailed visibility into churn behavior across different geographical regions, particularly at the country and city levels.

Without identifying where churn is concentrated and understanding associated purchasing behaviors, the company is unable to design targeted retention strategies.

Therefore, the primary problem is:

To analyze customer churn patterns geographically and identify actionable insights to reduce churn and improve customer retention.

DATASET LINK

https://www.kaggle.com/datasets/dhairyajeetsingh/ecommerce-customer-behavior-dataset

STEP1 : REDUCING DATAr

1. Reduced Dataset from 50,001 rows to 10,001 rows using pandas script.

Python Script Used

```
import pandas as pd
```

```
# Load dataset
df =
pd.read_csv("/content/ecommerce_customer_churn_datas
et.csv")

# Randomly sample 10,000 rows
df_sample = df.sample(n=10000, random_state=42)

# Save new dataset
df_sample.to_csv("ecommerce_sample_10000.csv",
index=False)
```

STEP2 : CLEANING DATA

1. Age Column:
   a. Removed Outlier : There existed values like **150** in the **age column** , so we got rid of it by dropping it from dataset.
   b. After taking care of Outlier:

MIN	5
MAX	75

   c. We imputed blank cells with **median** of the entire column.
   d. Reason : To minimize skewness in age column.
2. Gender Column:

a. It's a **<u>categorical column</u>** , so we can fill blank cells using **<u>mode</u>** of the gender column.
b. But there were no missing values.

	Count
Male	4,699
Female	5,302

3. Session Duration

a.

	Count
Missing	687
Mean	27.6
Median	26.8
MIN , MAX	1 , 69

b. Imputed missing values with the <u>**median**</u> of entire column

4. Pages per session Column

a.

	Count
Missing	605
Mean	8.7
Median	8.5

b. Imputed the blank cells using <u>**median**</u>

5. Wishlist Items Column

a.

	Count
Missing	786

b. We just replaced missing values with **0** (i.e. assuming there we no wishlist items)

6. Total_Purchase Column

a.

	Count
-ve Values	15
Mean	13
Decimal	2,272

b. Total purchase can't be negative so **replaced -ve values** with **0**.

c. We converted decimal values into integer using **floor** values.

7. Days_since_last_purchase Column

a.

	Count
Missing	573
Mean	29
MIN , MAX	0 , 287

b. Replace missing values with **mean** of the entire column.

8. Discount_Usage_Rate Column

a.

	Count
Missing	711
Mean	41.9

b. Imputed missing values with **0** (i.e. won't be getting any discount)

9. Returns_Rate Column

a.

	Count
Missing	905
Mean	6.62

b. Imputed missing values with **0** .

10. Email_Open_Rate , Customer_Service Calls , Product_Reviews Written Columns

a. Email_Open_Rate , Customer_Service Calls , Product_Reviews Written had 538 , 31 , 699 missing values respectively.

b. Replace all these blank cells with the **median** of respective column.

11. Social_Media_Engagement_Score Column

a. Had 1,174 missing values.

b. Created a separate column named **Social_Score_missing** , to get insight of missing values for further comparison with

churned rate and get an idea how Social_Media_Engagement affects churn rate.

12. Payment method Diversity Column

a.

	Count
Missing	490
Mean	2
MIN , MAX	1 , 5

b. All the tuples that had missing payment method , they had average order value , so replaced empty cells with **mean**.

13. Credit_Balance Column

a.

	Count
Missing	1,105
Mean	1,965
MIN , MAX	0 , 7,197

b. For empty cells we made an assumption that their credit balance was **0** , because if we would have used mean , that would have inferred to false balance of 1,965 making data biased.

STEP3 : CREATING PIVOT TABLES & GAPHS

1. Executive Overview

This dashboard provides a comprehensive analysis of customer behavior, purchasing trends, engagement performance, and churn dynamics across eight countries:

Australia, Canada, France, Germany, India, Japan, UK, and USA.

The objective is to evaluate business performance, identify behavioral patterns, and uncover opportunities for improving retention and conversion rates.

2. Key Performance Indicators (KPI Summary)

E-Commerce Customer Insights & Churn Analysis Dashboard

Total Customers	Total Order	Total Sales	Avg Order Value	Avg Cart Abandonm…	Churn Rate
9999	130279	14448598.04	123.77	57.09	28.70%

Total Customers: 9,999
 Total Orders: 130,279
 Total Sales: 14,448,598.04
 Average Order Value: 123.77
 Average Cart Abandonment Rate: 57.09%
 Churn Rate: 28.70%

The company demonstrates strong revenue and transaction volume. However, cart abandonment and churn remain key areas for optimization.

3. Dashboard Slicers (Interactive Filtering)

The dashboard includes slicers that allow dynamic filtering by:

- **Country**

- **Churn Status (0 = Retained, 1 = Churned)**

- **Gender**

These slicers enable focused analysis of customer behavior across specific segments, helping identify:

- Country-level performance differences

- Behavioral patterns of churned vs retained customers

- Gender-based purchasing trends

This enhances the analytical flexibility and usability of the dashboard.

4. Average Order Value by Country

Average Order Value by Country

Avg order value across regions
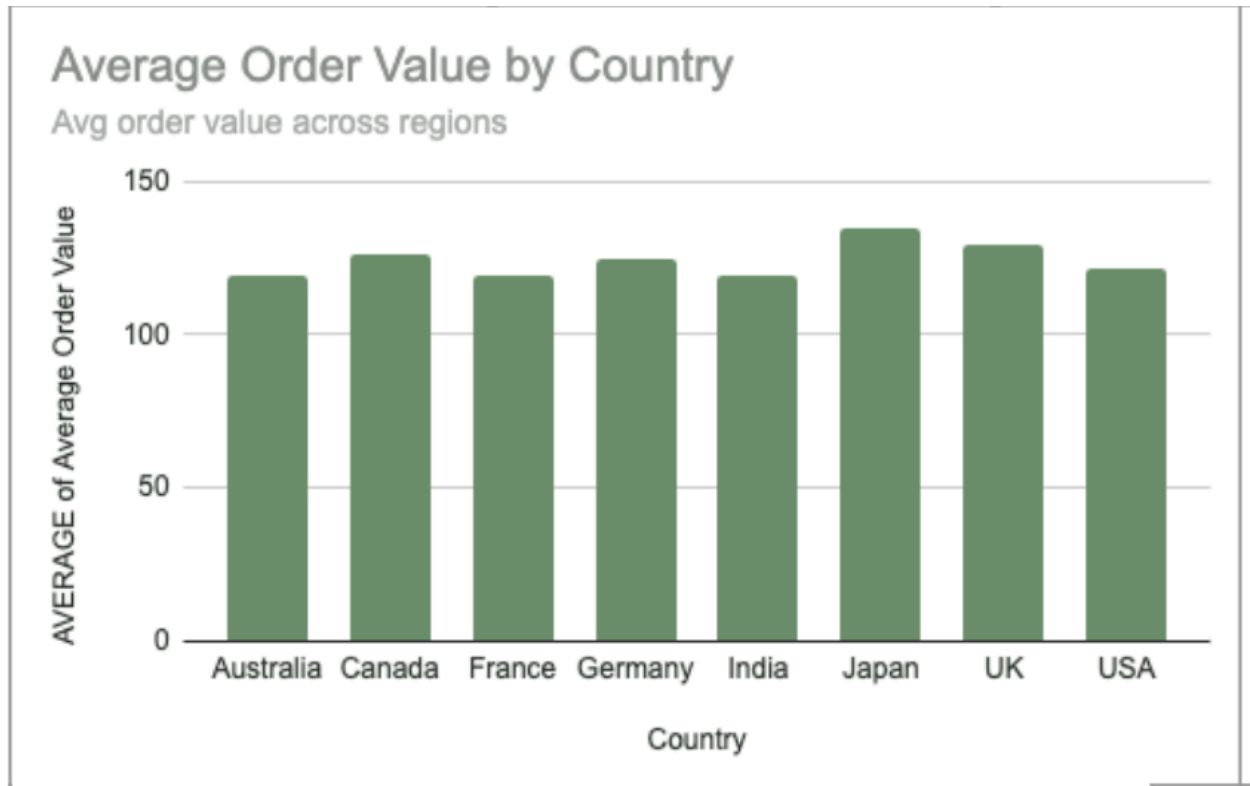
The Average Order Value ranges approximately between 119 and 135 across countries.

- Japan records the highest AOV.

- Germany and UK also show strong performance.

- Other countries remain stable with minimal variation.

This indicates consistent pricing strategy and balanced spending behavior globally.
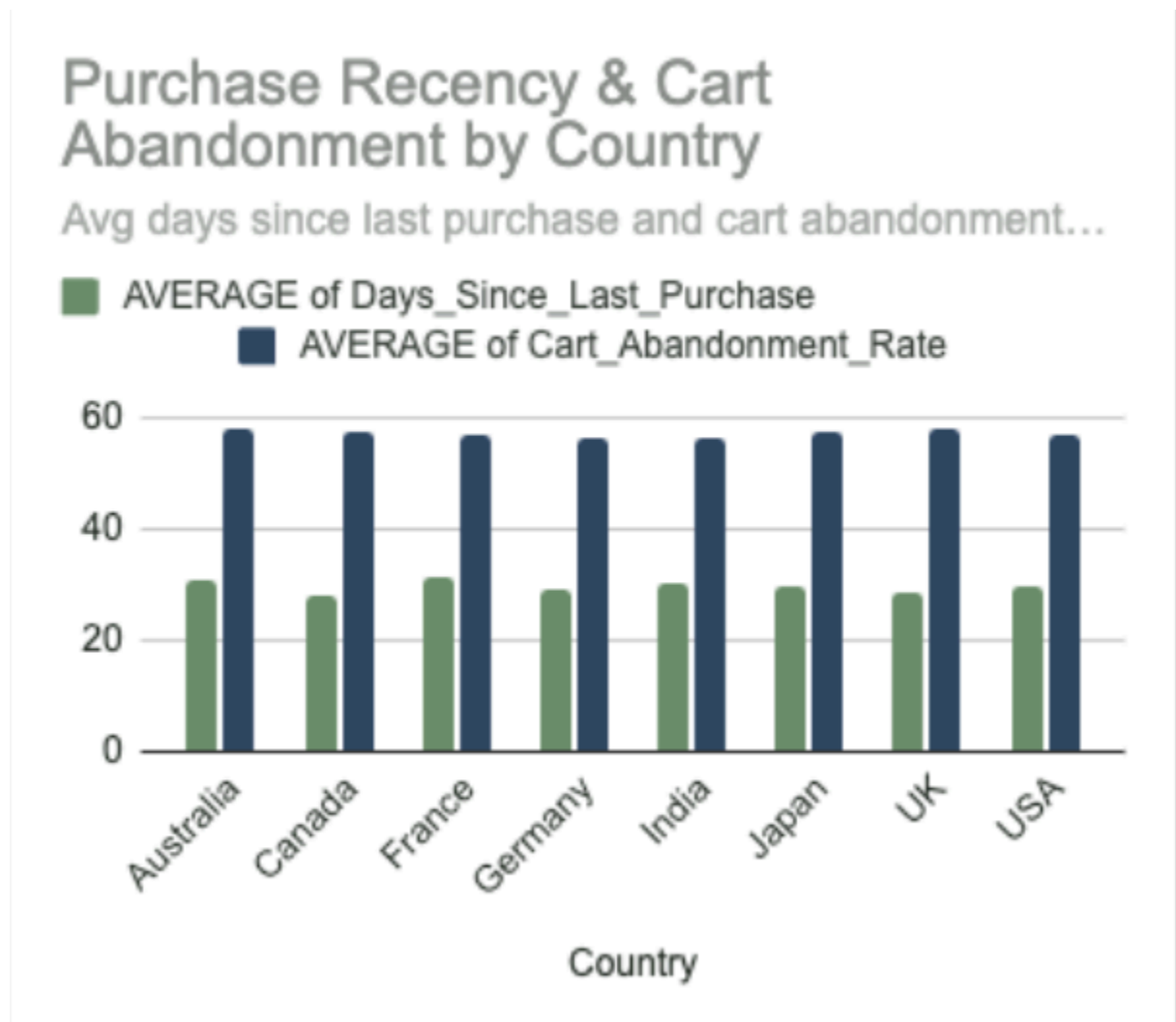
5. Discount Usage by Country
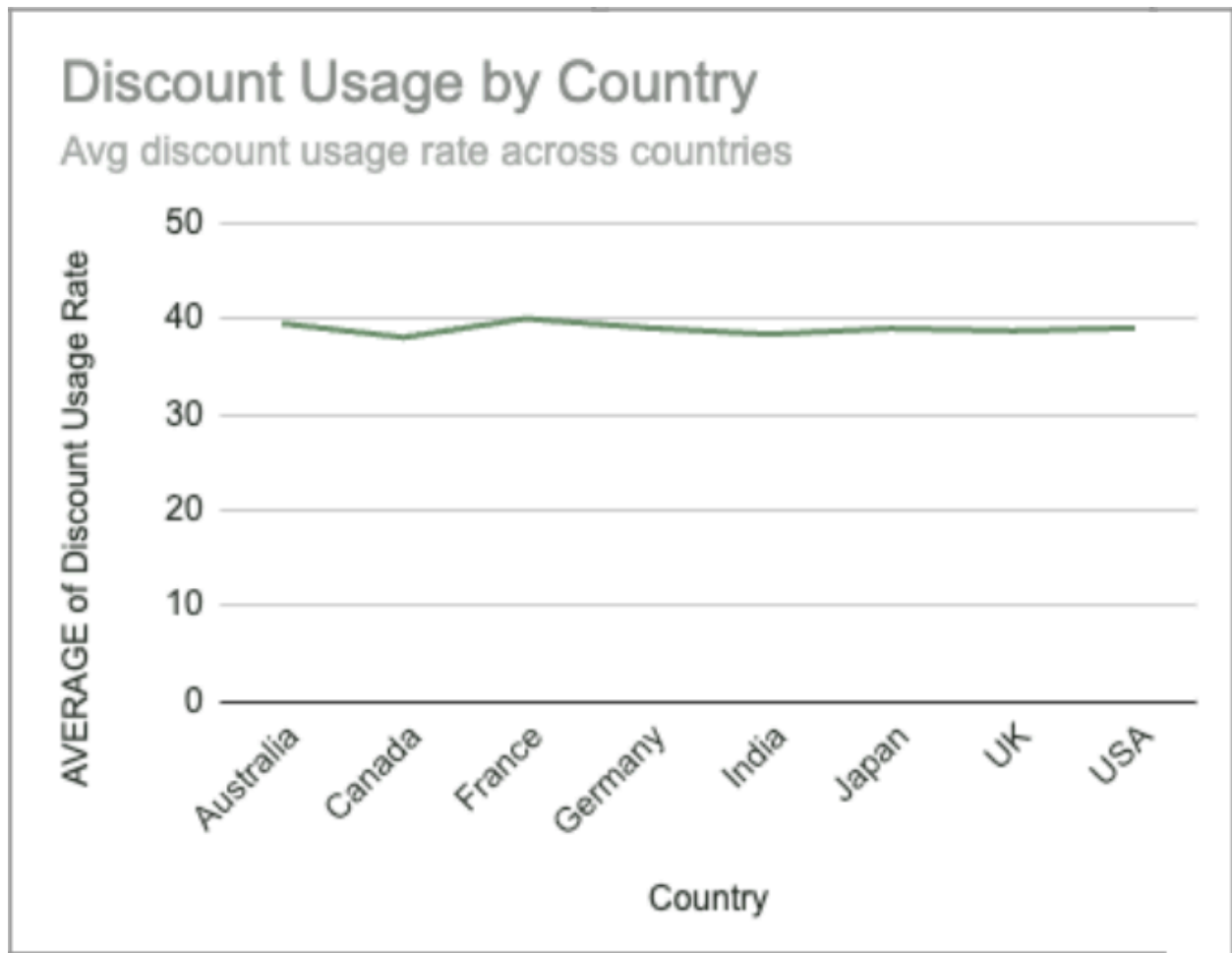
Average discount usage is approximately 39% across all regions.

- Very small variation between countries.

- Customers moderately respond to promotions.

This suggests stable promotional effectiveness without excessive discount dependency.

6. Purchase Recency & Cart Abandonment by Country

This graph compares:

- Average Days Since Last Purchase (~29 days)

- Average Cart Abandonment Rate (~56–58%)

Customers exhibit a monthly purchasing cycle across all countries. However, the cart abandonment rate remains consistently high, highlighting a major opportunity to improve checkout optimization.
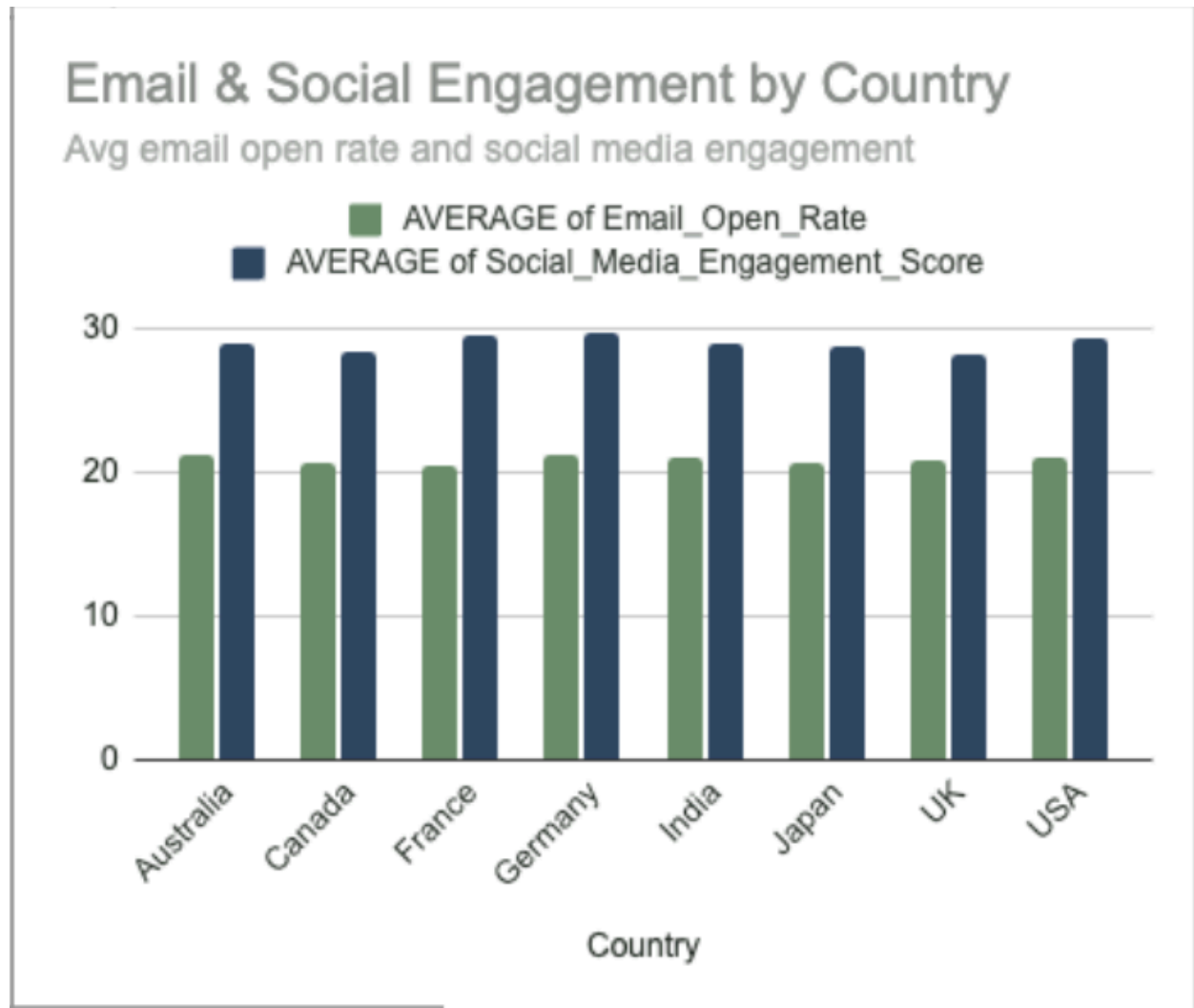
7. Session Duration by Country
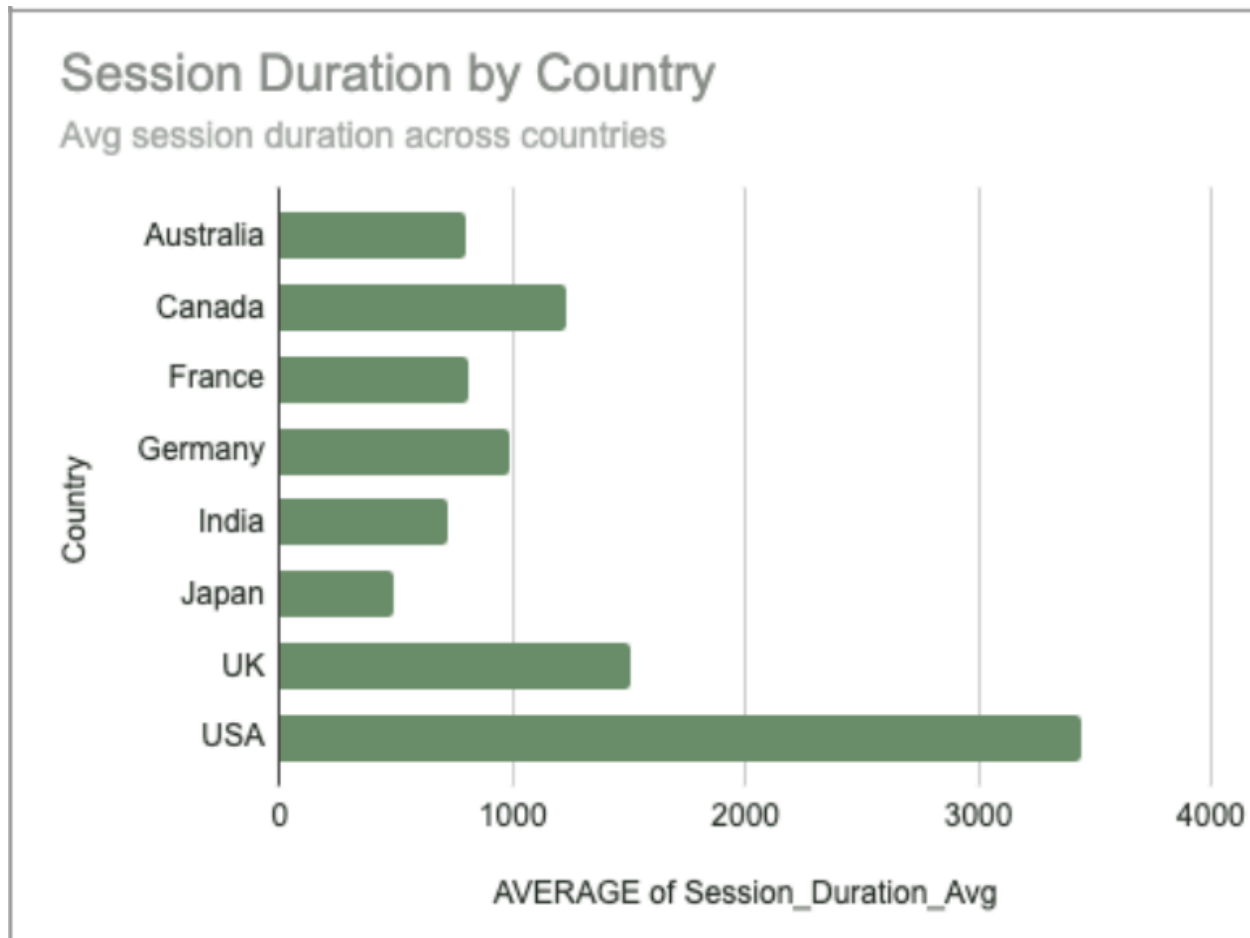
The United States shows the highest average session duration.

- UK and Canada follow.

- Japan and India show relatively lower session duration.

Higher session duration reflects stronger engagement and deeper browsing behavior.

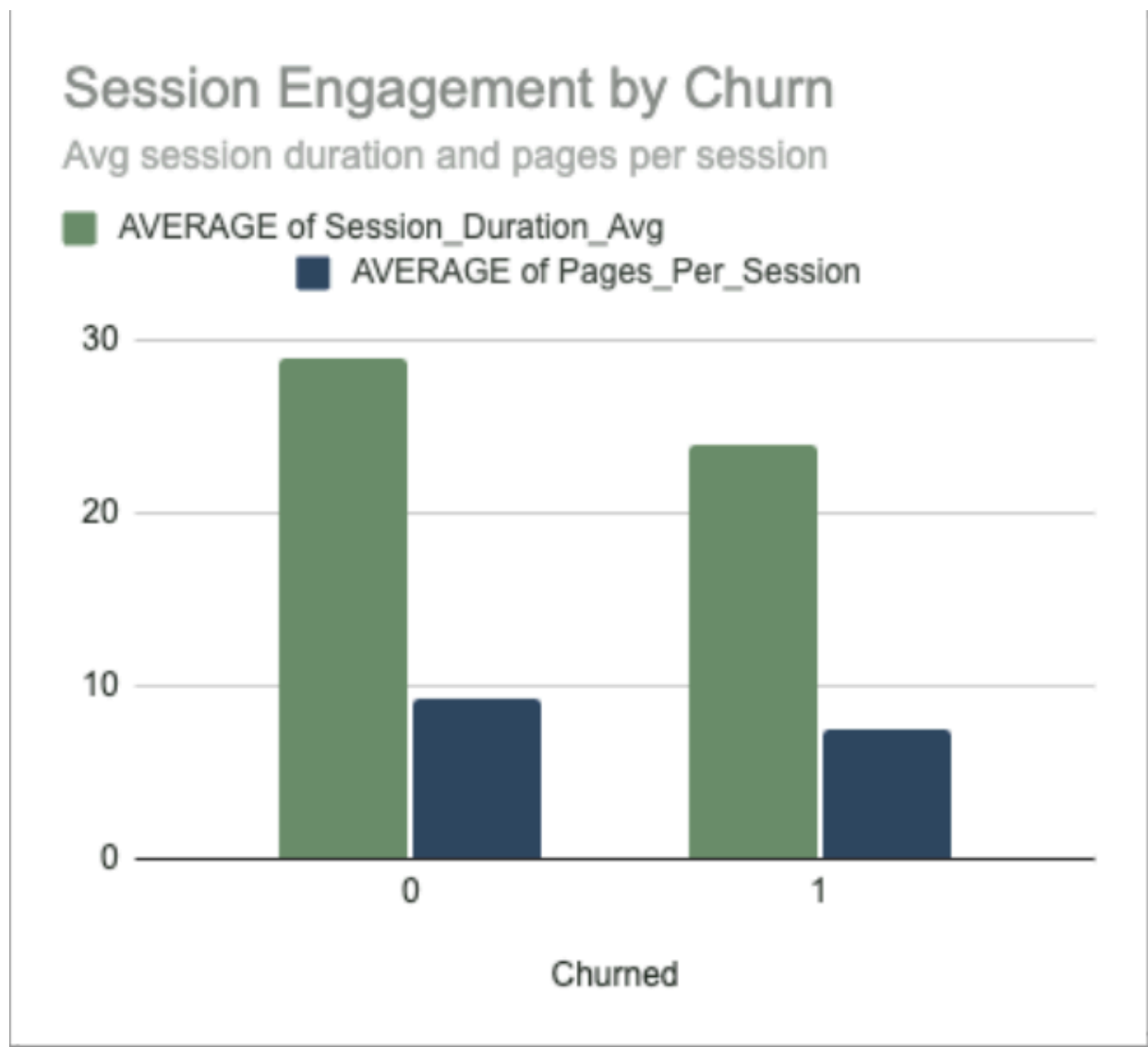## 8. Email & Social Engagement by Country

Email Open Rate ≈ 20–21%
 Social Media Engagement Score ≈ 28–30

Engagement levels are steady across all countries with no extreme variation.
 Social media engagement slightly exceeds email performance.
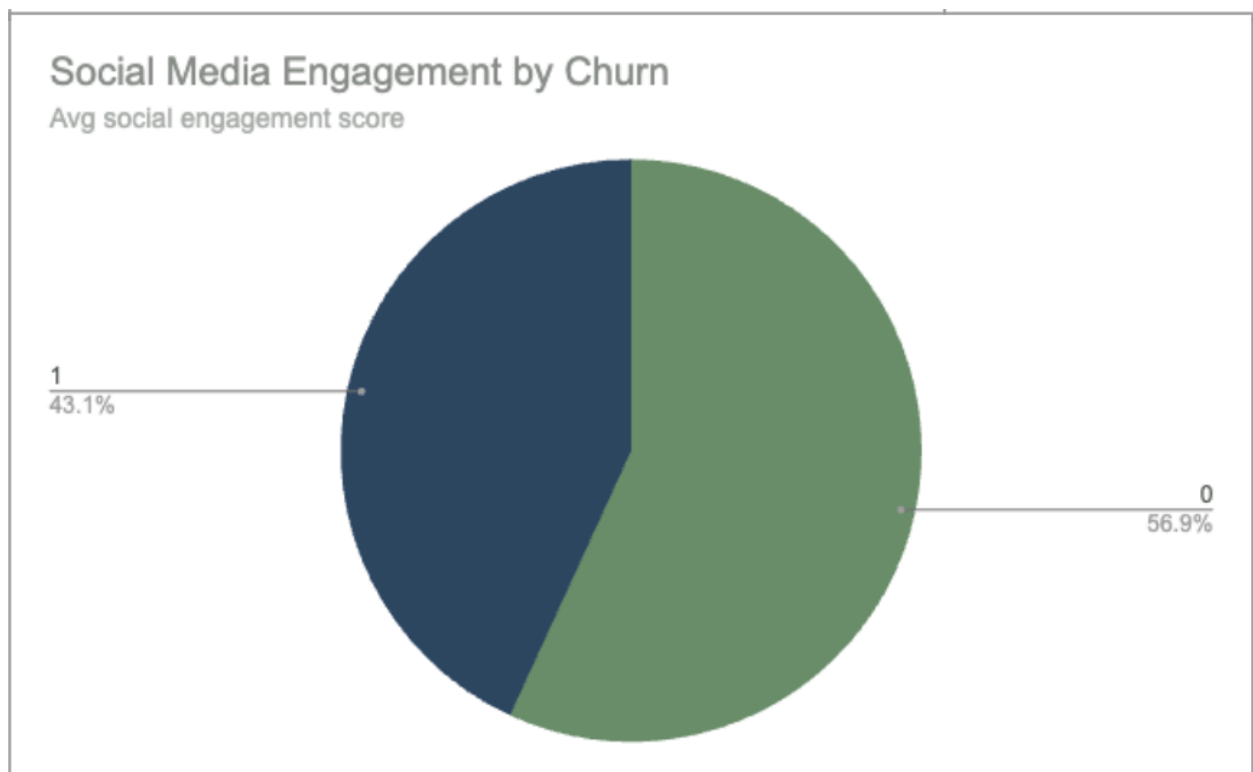
9. Session Engagement by Churn

This visualization compares retained (0) vs churned (1) customers.

- Retained customers spend more time per session.

- Retained customers view more pages per session.

- Churned customers show reduced engagement.

Lower engagement levels are directly associated with churn risk.
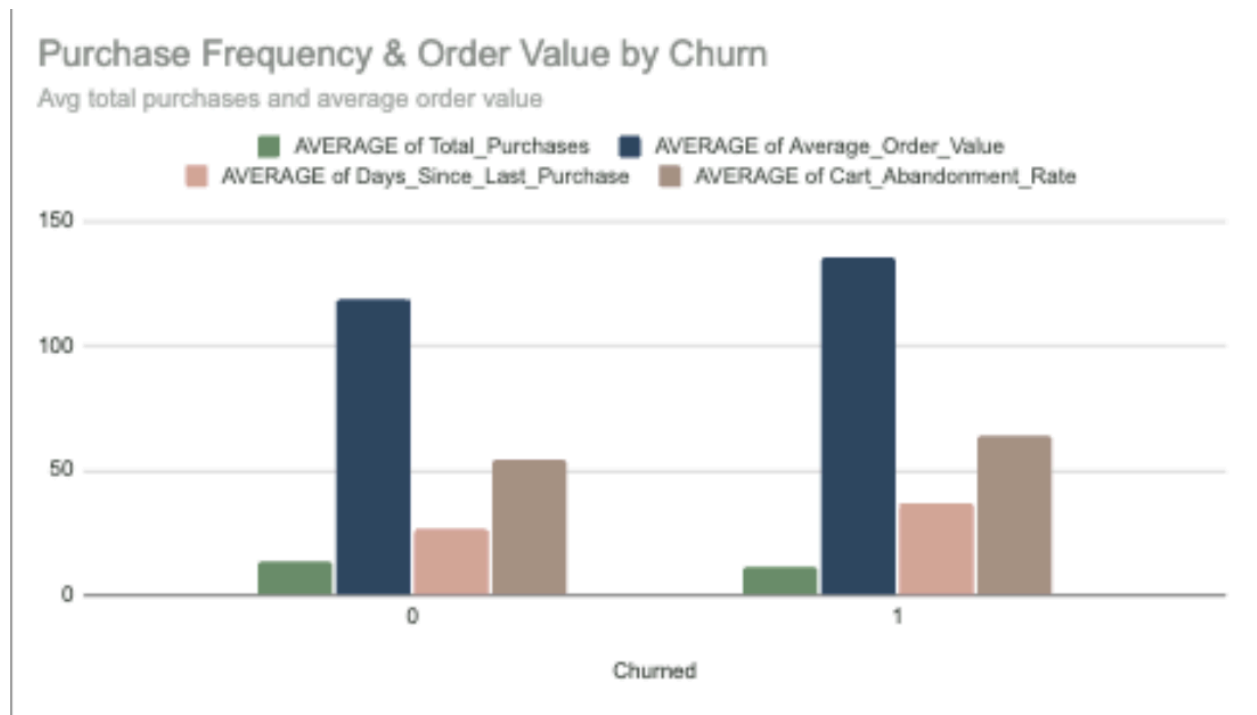
10. Social Media Engagement by Churn

Retained customers demonstrate higher social engagement scores compared to churned users.

This reinforces the importance of digital interaction in improving customer retention.

11. Purchase Frequency & Order Value by Churn

Purchase Frequency & Order Value by Churn
Avg total purchases and average order value

Legend:
- AVERAGE of Total_Purchases
- AVERAGE of Average_Order_Value
- AVERAGE of Days_Since_Last_Purchase
- AVERAGE of Cart_Abandonment_Rate

X-axis: Churned (0, 1)

This graph compares:

- Total Purchases

- Average Order Value

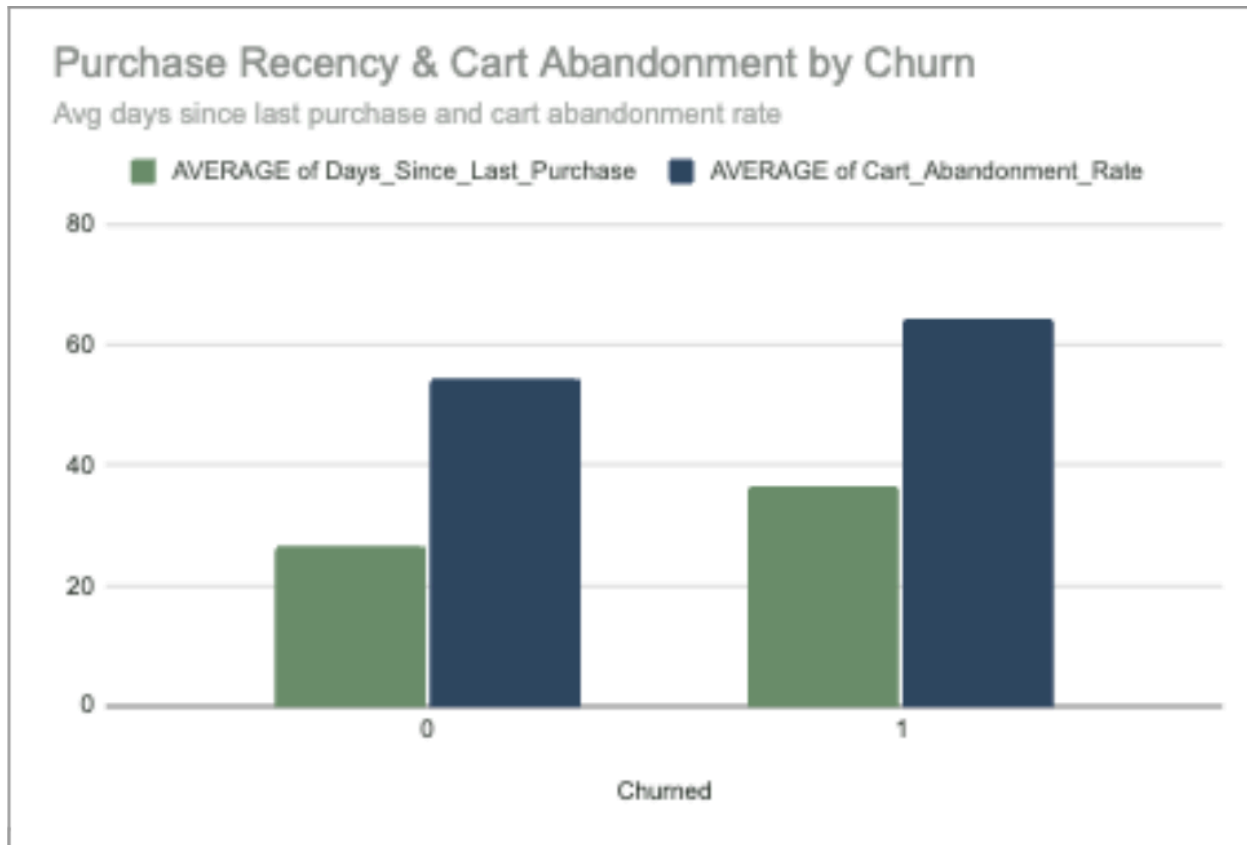- Days Since Last Purchase

- Cart Abandonment Rate

Churned customers show:

- Lower purchase frequency

- Higher days since last purchase

- Higher cart abandonment

This indicates declining purchase activity before churn occurs.

12. Purchase Recency & Cart Abandonment by Churn

Purchase Recency & Cart Abandonment by Churn

Churned customers:

- Have significantly higher days since last purchase

- Exhibit higher cart abandonment rates

This suggests disengagement and friction during checkout contribute to churn.

13. Missing Social Engagement by Churn

Missing Social Engagement by Churn

% of customers with missing social data

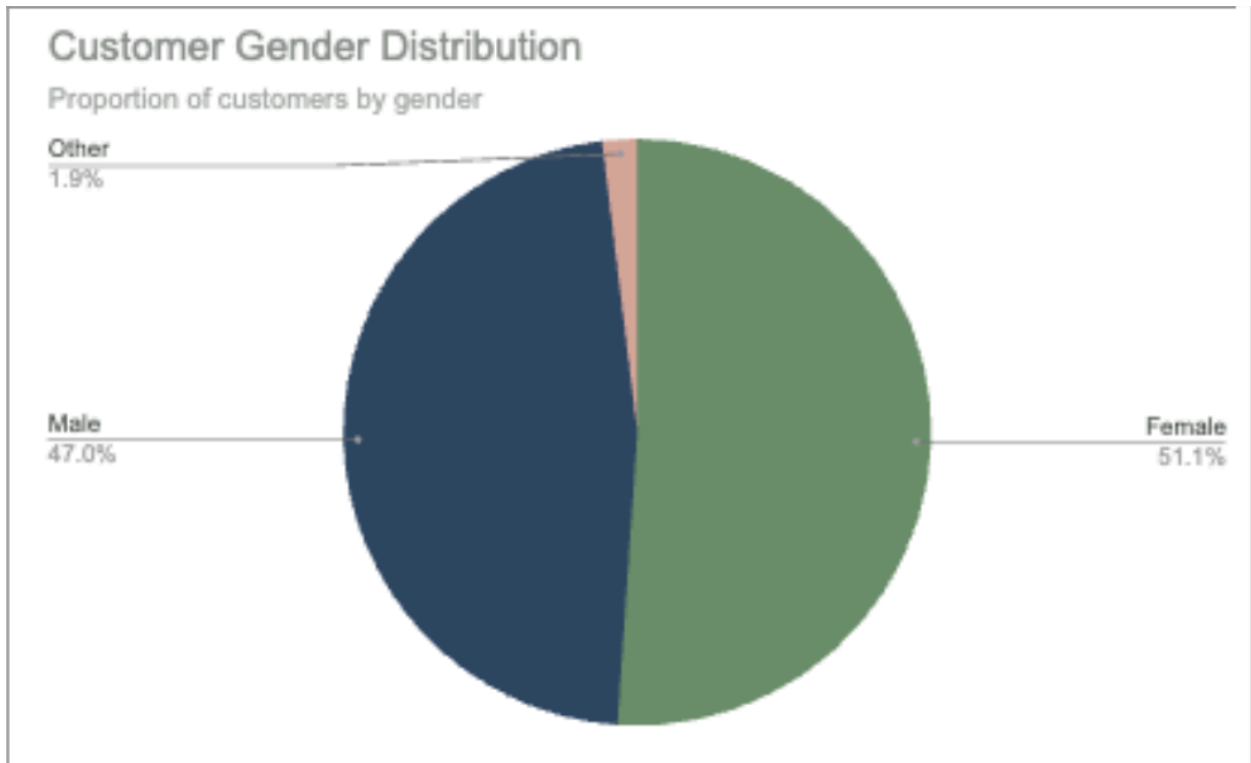
Churned
28.7%
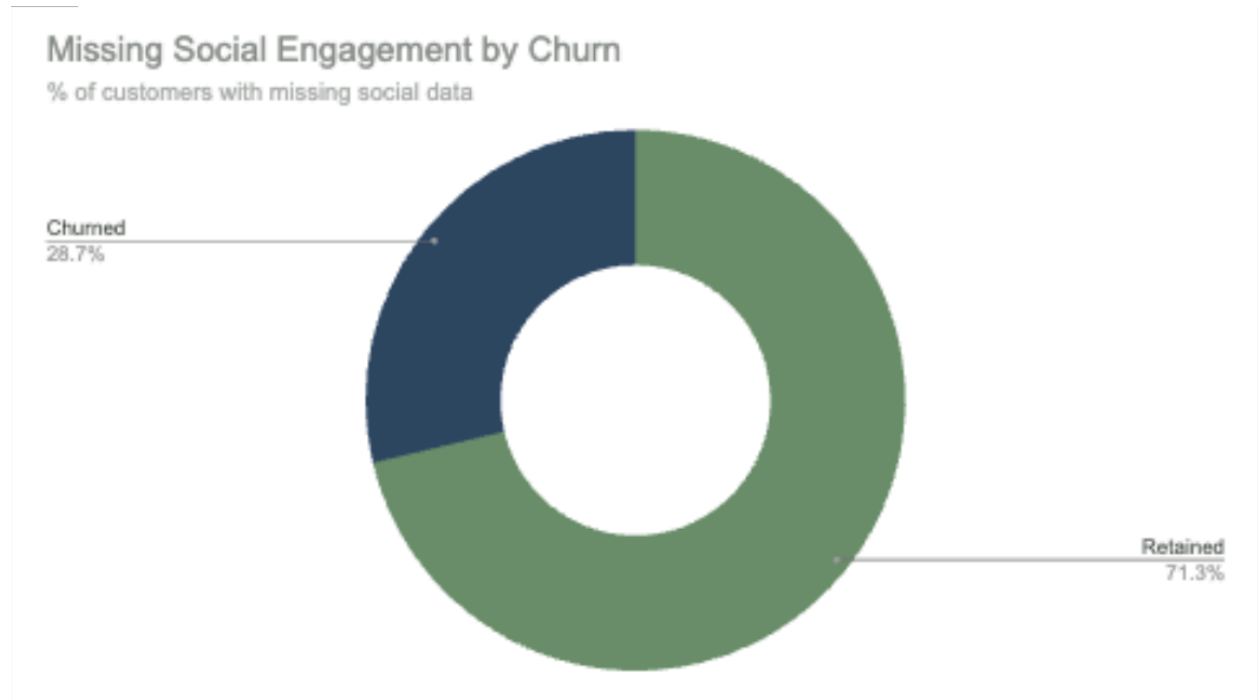
Retained
71.3%

Approximately:

Retained: 71.3%
 Churned: 28.7%

Customers with weaker or missing social engagement signals show higher churn probability, indicating digital inactivity as a risk indicator.
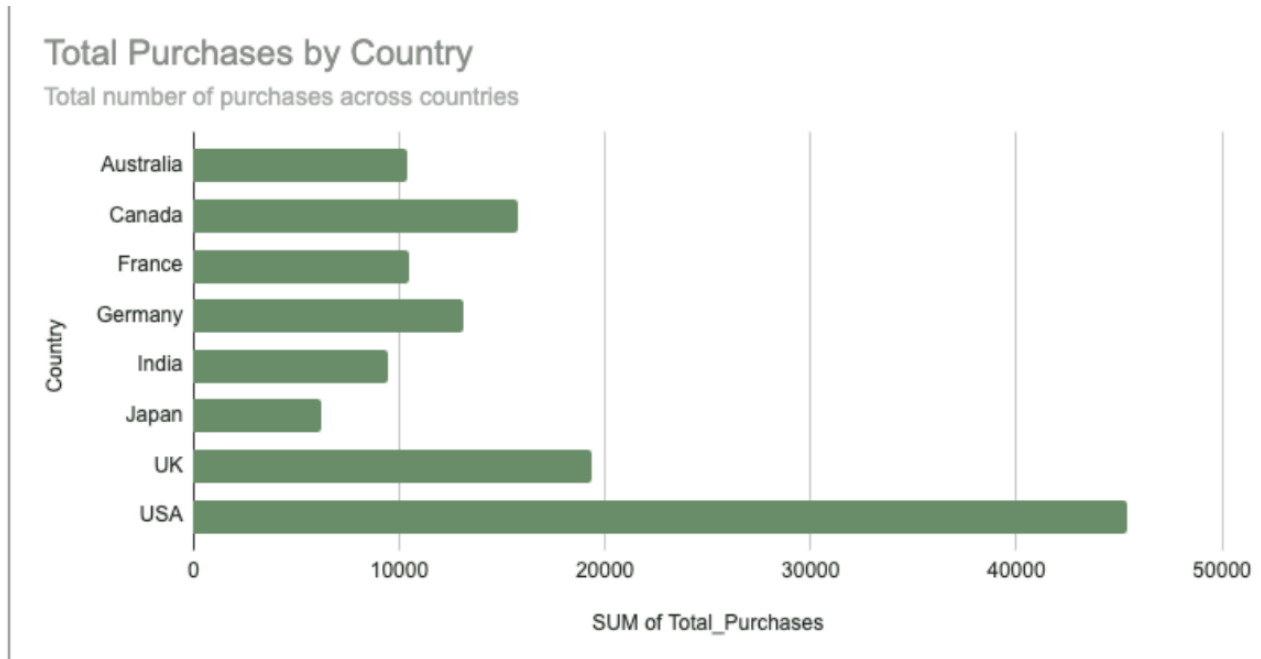
14. Customer Gender Distribution

Customer Gender Distribution

Proportion of customers by gender

Female: 51.1%
 Male: 47.0%
 Other: 1.9%

The customer base is well balanced with a slight female majority.

15. Total Purchases by Country

Total Purchases by Country

Total number of purchases across countries

USA leads significantly in total purchases.

UK and Canada follow.

Other countries show moderate but stable transaction volume.

This highlights the USA as the strongest revenue-driving market.

16. Overall Business Insights

The business demonstrates:

- Strong total sales and order volume

- Stable average order values across markets

- Consistent monthly purchasing behavior

- Balanced demographic distribution

- Uniform digital engagement performance

However:

- Cart abandonment remains high (57%)

- Churn rate is significant (28.7%)

- Lower engagement correlates strongly with churn

17. Strategic Implications

To improve performance, the business should focus on:

- Optimizing checkout process to reduce abandonment

- Increasing engagement campaigns to reduce churn

- Implementing personalized retention strategies

- Leveraging high-performing markets (USA, UK)

- Strengthening digital interaction touchpoints