Customer Purchase Behavior Analysis - Team Yantraksh

Team Members: Chethan Vasthaw Tippani, Mithuna Somireddy

This is our project where we tried to understand how Amazon customers behave, who are the loyal ones, who might leave, and how much money they might bring in future. We used machine learning and visual dashboards to get insights.

1. Process We Followed

- First, we cleaned the data (removed duplicates, filled missing values).

- Then we created some new columns like CLV, loyalty score, etc.

- We applied ML models like clustering, linear regression, and logistic regression.

- Finally, we visualized everything using Power BI.

2. Data Cleaning & Feature Engineering

- Removed missing values by filling with mean/median for numbers.

- Removed duplicate customer entries.

- Changed date columns to proper datetime format.

- Created new columns like:

  - CLV: Customer lifetime value (future worth)

  - Loyalty Score: Based on how often and how much they shop

  - Discount Applied, Return Status, Customer Segment, Preferred Channel

3. ML Models

We used 3 models mainly:

- KMeans (to group customers)

- Linear Regression (to predict CLV)

- Logistic Regression (to predict churn)

4. Code Explanation

Step 1: Reading and Cleaning Data

We read the Excel data and cleaned it. Also converted the date columns and filled null values.

Step 2: Adding Features

We added columns like CLV (future money we expect), loyalty score, discount used or not, etc.

Step 3: Clustering (KMeans)

We grouped customers using KMeans so we can know who are high spenders, regular buyers etc.

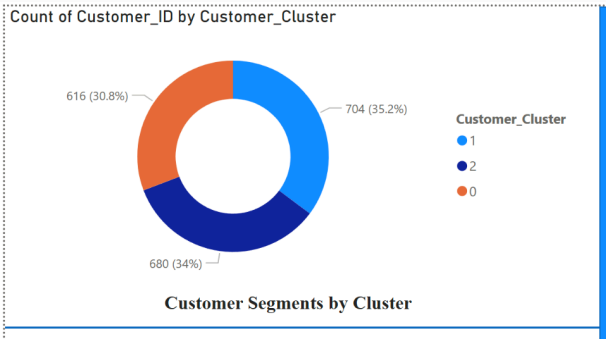
Step 4: Predicting CLV (Linear Regression)

We trained a linear regression model to predict how much a customer might spend in future.

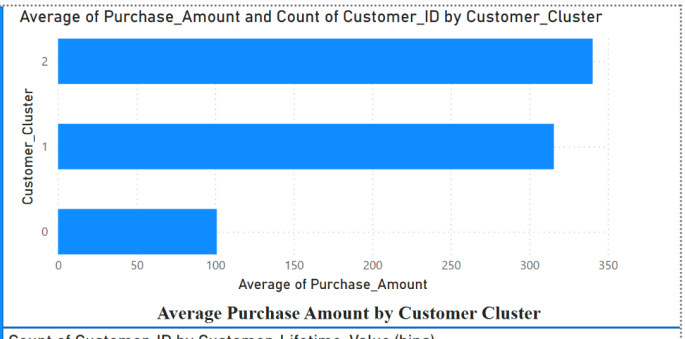
Step 5: Churn Prediction (Logistic Regression)

We predicted if a customer is likely to leave Amazon based on their past behavior.

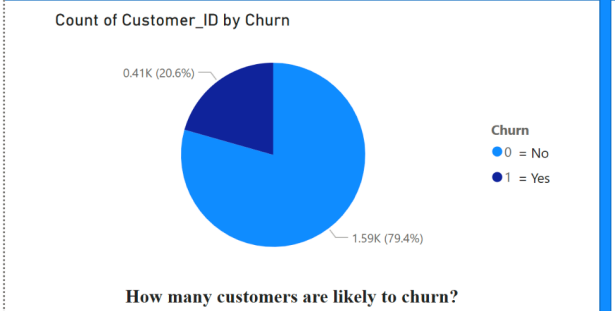
5. Power BI Visuals (Attached Below)



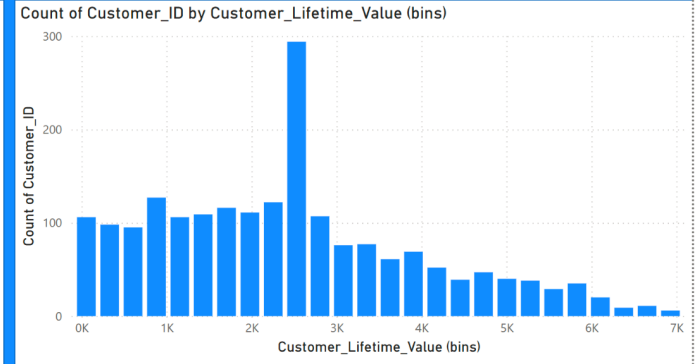
Customer Segments by Cluster



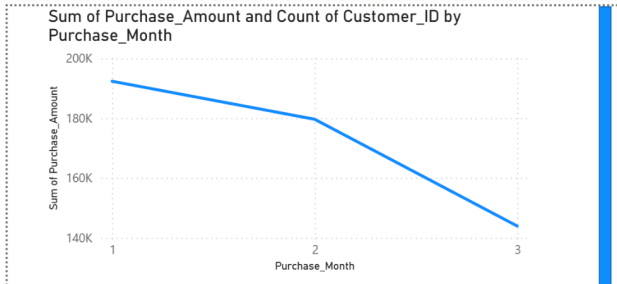
Average Purchase Amount by Customer Cluster



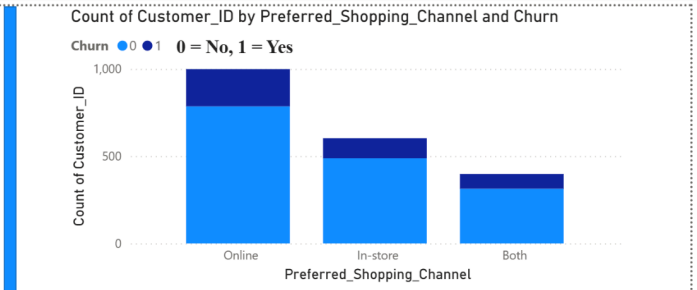
How many customers are likely to churn?



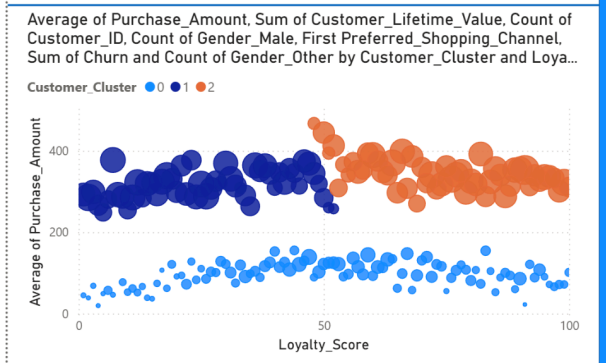
Customer Lifetime Value Distribution



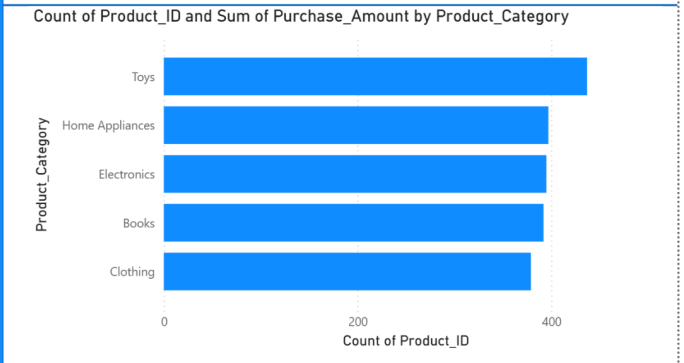
Monthly Purchase Trends



Churn Rate by Shopping Channel



Loyalty Score vs Purchase Amount



Top Product Categories

6. Conclusion

So overall, this project helped us understand how customers behave and how Amazon can improve marketing, offers, and retention using data.