**Dataset Visualization for Repeat Buyers Prediction for E-Commerce**

Using data visualization allows us to understand the pattern and trends in the data, such as which merchants have the most repeat and non-repeat buyers, the distribution of age range and gender among users, the trend of user actions over time, and more. By presenting the data in a visual format, we can easily identify patterns, relationships, and trends that might not be apparent from looking at raw data alone.

User Behavior Data Visualization:

**Fig. 1** shows that users have clicked on items over 6000 times but have only purchased, marked favorites and added to cart less than 1000 times. Add to Cart is almost negligible because people don't want to just add to the cart without buying.

Chart, bar chart, waterfall chart

Description automatically generated

**Fig. 2** shows the majority of the users only click on the items without purchasing them due to various reasons such as checking out what's available before buying or they simply don't like the item/s.

<see image on the next page>

Chart, pie chart

Description automatically generated

**Fig. 3** shows that category 737 is the most popular and has the most actions among other categories. It is recommended that items in categories that have count of actions between 500 - 1000 be promoted for better sales.

A picture containing histogram

Description automatically generated

**Fig. 4** shows that there is an increase in user actions, especially clicks and purchases on the last week of September and during the promotional Double 11 Day.The surge in user activity indicates that it is best to optimize marketing and sales strategies during these periods. Visualizing the trend of user actions overtime helps in making data-driven decisions to improve user engagement.

A picture containing chart

Description automatically generated

User Profile Data Visualization

**Fig. 5** shows the distribution of Age Range and Gender to help better understand the customer's base. Based on this chart, Female within the age range of 25-29 are the target audience. Merchants can focus on products and promotions that are more likely to appeal to female buyers.

<see image on the next page>

Chart, bar chart

Description automatically generated

**Fig. 6** shows that most activities are made by users between 25 to 29 years old. Selling items appropriate to this age range would increase the chance of converting non-repeat buyers to repeat buyers.

A picture containing diagram

Description automatically generated

**Fig. 7** shows that most of the buyers are female. It is recommended to plan for product enhancement and promotions for female buyers and launching new items for men to increase their interest in buying.

**Diagram

Description automatically generated**

**Fig. 8** shows the frequency of action type by age range and gender. This can provide insights into the shopping behavior of different age and gender groups.

For example, male buyers in age between 40-49 has a higher frequency of adding items to their cart but a lower frequency of making purchases. This may indicate that these buyers are more price-sensitive than female buyers in the same age range.

On the other hand, female buyers between age range 40-49 have lower frequency on adding items to cart but higher frequency on purchasing the items. This might indicate that these buyers are more impulsive buyers or have a higher disposal income than male buyers in the same age range.

<see image on the next page>

Chart

Description automatically generated

Training Data Visualization

**Fig. 9** shows the distribution of those who are making purchases versus those who are not. This provides a clear and concise representation of repeat vs. non-repeat buyers. This helps in understanding the proportion of customers repeat buyers versus non-repeat buyers.

This chart shows that there is an overwhelming number of Non-repeat buyers. It is recommended to focus more on improving the overall customer experience to encourage repeat purchases.

This chart provides valuable information that can be helpful for businesses to make decisions about their marketing and customer retention strategies.

<see image on the next page>

Chart, bar chart

Description automatically generated

**Fig. 10** shows the behavior of customers on the platform. Merchant ID 3828 is identified as the merchant with the most repeat buyers. Identifying the top merchants with the most repeat buyers should focus on providing better support and resources, which could lead to increased customer satisfaction and loyalty.

Chart, bar chart

Description automatically generated

**Fig. 11** shows which merchants have the most non-repeat buyers. Merchant ID 4044 is identified as the merchant with the most non-repeat buyers. The merchants on this list should identify potential areas for improvement in customer retention strategies. This can be used for comparing a merchant's performance to its competitors in terms of customer retention.

As noticed on Fig. 10, 9 merchants (4044, 3828, 4173, 1102, 4976, 1892, 3734, 598, 361) also have the most repeat buyers. This could suggest that these merchants are popular and have a lot of customers but are struggling to keep them as there are overwhelming numbers of non-repeat buyers.

Chart, bar chart, histogram

Description automatically generated

**Fig. 12** shows which users have the most repeat purchases. By identifying these users, businesses can analyze the behavior of these repeat purchasers and come up with a new strategy to attract new customers and improve customer retention.

<see image on the next page>

Chart, bar chart, histogram

Description automatically generated

**Fig. 13** shows which users have the most non-repeat purchases. It is important to identify these users and analyze their behavior. Analyzing the behavior of these users can help identify potential issues with the products and services and come up with new marketing strategies to encourage repeat purchases.

Chart, bar chart

Description automatically generated

* END