

Case Study Prompt

Understanding customer value is one of the biggest challenges for marketing teams around the world. As we accumulate more and more data, it is important to harness the true potential of the data and discover actionable insights to make smart business decisions.

In this exercise, we have transaction data of a retail store (attached). Please categorize the customers into different segments using RFM analysis and suggest data-driven action items for each segment. The tool for analysis and visualization can be of your choice.

Assuming the above data points are related to an eCommerce setup. How would the action items/recommendations change for each segment? Are there any additional digital metrics you would want to track to understand online customer behavior better? Please share those metrics.

Notes:

- I discuss the case in this document.
- Find the GitHub repository containing the Jupyter notebook [here](#), where I carried out the initial EDA and tried to impute missing customer IDs.
- I carried out customer segmentation in this [Excel file](#).

Summary of data

- Total number of unique transactions - 25900
- Total sales made between 01-Dec-2010 and 09-Dec-2011 - \$9,747,747
- Total number of customers - 4372 (excluding missing customer id data). Out of which (refer to the appendix for definitions):
 - Core customers - 1216
 - Loyal customers (frequent buyers) - 1375
 - Whale (high avg. order value) customers - 881
 - Promising (regular but low order value) customers - 835
 - New customers - 130
 - Lapsed (lost) customers - 131
 - Other customers - 543

Key findings and recommendations

- **Core customers** account for 27.8% of total customers (excluding customers with missing cust_id) but **generate more than 60% of total revenue**. These are customers that we simply cannot afford to lose.
 - Add value through loyalty programs. These are loyal customers who have shown a higher willingness to pay, so discounts may not be the best strategy for these customers. Instead, we should focus on value addition through added benefits, privileged customer care service, and more exclusive services and benefits.
 - Increase sales by introducing new products and product recommendations. These customers trust us and will be willing to try new products and/or purchase more on our platform.
- Loyal customers are the most frequent buyers. Loyalty programs such as discount coupons and free shipping are effective for repeat buyers. We can also reach out to these customers for online reviews and word-of-mouth.
- Whale or high-ticket customers generate maximum revenue per order. These customers have demonstrated a high willingness to pay. Consider premium offers, subscription tiers, luxury products, or value add and cross/up-sells to increase average order value (AOV). Don't waste margin on discounts.
- Promising customers are loyal customers who frequently buy from us but place low order values. We can try increasing monetization through product recommendations based on past purchases and incentives tied to spending thresholds (pegged to our store AOV).
- Most new customers never graduate to loyal. Having clear strategies in place for first-time buyers such as triggered welcome emails, discounts on cart/wishlist items, and push notifications will help retain new customers.
- Lapsed customers are customers who were once regular buyers but haven't purchased from us lately. We can experiment with price deals, new product launches, or other retention strategies to get back these customers.

Q. How would the action items/recommendations change for each segment?

As we can see, the actionable change for different customer segments. This is because customers at different stages have different priorities and motivations. For example, a new customer may have a different favorite store and may have purchased with us only because the product was out of stock at the other store. New customers need to be reminded and incentivized with discounts before some of them turn into regular customers.

On the other hand, a loyal customer who purchases frequently from us will be willing to purchase more from us, in which case introducing new products or effective product recommendations can help increase sales.

At the same time, our core customers will be watching out for priority care services and added benefits.

Q. Are there any additional digital metrics you would want to track to understand online customer behavior better?

1. Dwell time

This is how long our customers spend on our website/app. Theoretically, the longer the dwell time, the more likely a customer is to make a purchase (although we can still get customers who pop onto our website, search for a specific item and purchase within minutes).

Similarly, we can record time stamps every time a customer visits our website/app and check for whether or not lapsed customers (who haven't purchased in recent times) have also stopped visiting our website/app. We can strategize differently for lapsed customers who were recently seen on our platform versus those who were not seen on our platform in recent times.

2. Website/app traffic patterns

We can track metrics such as bounce rates (for top landing pages), average engagement time, and other metrics on Google Analytics to better understand how our customers interact with our website/app.

Heatmaps can give us deeper insights into how visitors interact with a given page and help us solve UI/UX-related issues.

3. Customer demography

Demographic data such as age and gender (in addition to country/ region) can help us improve product recommendations.

4. Interests

For large e-commerce platforms with a huge range of products e.g., Amazon, categorizing customers based on interests (product categories the customer purchases frequently) can be very helpful.

Appendix 1 - Definitions

Customer Segment	Description
Core	Customers marked with '1' in this column are our core customers. They purchase frequently, spend more on every order, and have placed an order recently. These customers have a score of 4-5 in each of the RFM metrics e.g., recency, frequency, and monetary. These few customers account for the maximum revenue.
Loyal	These are loyal customers. They purchase very frequently i.e., they've a score of 4-5 in the frequency column.
Whale	These are big-ticked customers with maximum average order value (avg_order_value = 4 or 5).
Promising	Promising customers are customers who purchase frequently (3-4 in frequency) but place low-value orders (≤ 3 in avg_order_value).
New	These are first-time buyers with a maximum of one or two orders and a recency score of 5.
Lapsed	These were regular customers in the past but haven't purchased from us in a while.

Appendix 2 - Cool EDA Findings

- Nearly 25% of rows have missing customer IDs.
- InvoiceNo starting with 'C' are likely return orders.
- There are two instances of adjustments for bad debts which are indicated with a -ve unit price.
- The company seems to like to give away a lot of free items on certain orders.
- Some of the free items that are faulty are also refunded for and thus a -ve order quantity for free items.

Check out my Jupyter Notebook for more cool EDA findings.