

EXPLORATORY DATA ANALYSIS



DATA DICTIONARY

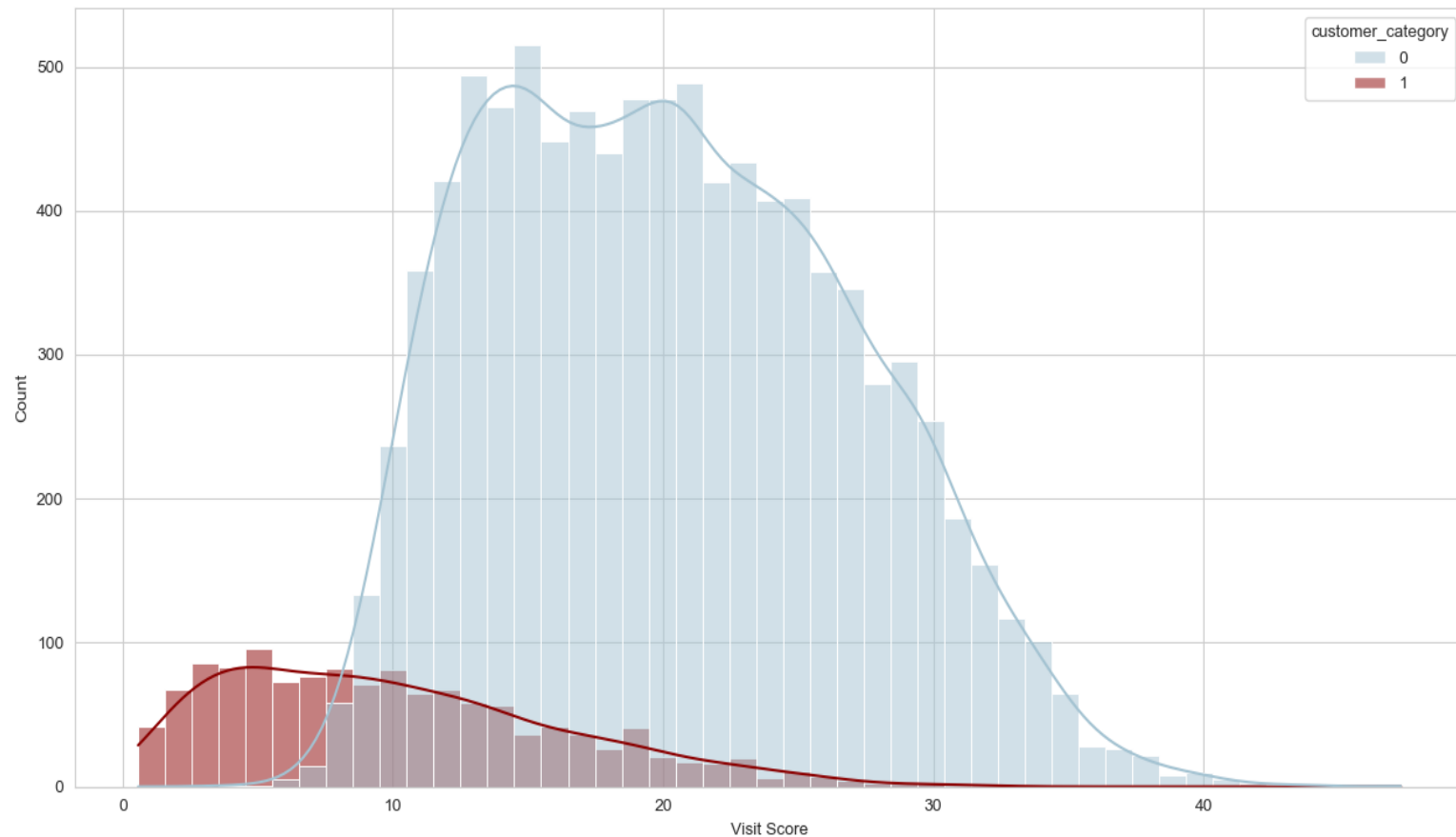
S.No.	Column Name	Description
1	Visit Score	A score based on how regularly the customer visits the website
2	Product Search Score	Quality or price range of product that the customer searches for.
3	Ctr Score	How many of the searched links does the customer click
4	Stay Score	A score based on the time spent on an avg. by the customer
5	Frequency Score	A score based on how many times in a day the customer visit the website
6	Product Variation Score	A score based on how many varieties of products does a customer search for
7	Order Score	A score based on the no. of orders that has been successfully delivered and not returned
8	Affinity Score	An internal overall score calculated which signifies the affinity of the customer towards the website
9	X_1	Anonymized feature based on loyalty of the customer
10	Active Segment	The categorization of the customers based on their activity
11	Category	The cluster/group to which the customer should belong to



VISIT SCORE ANALYSIS

How visit score is distributed among the two type of customers ?

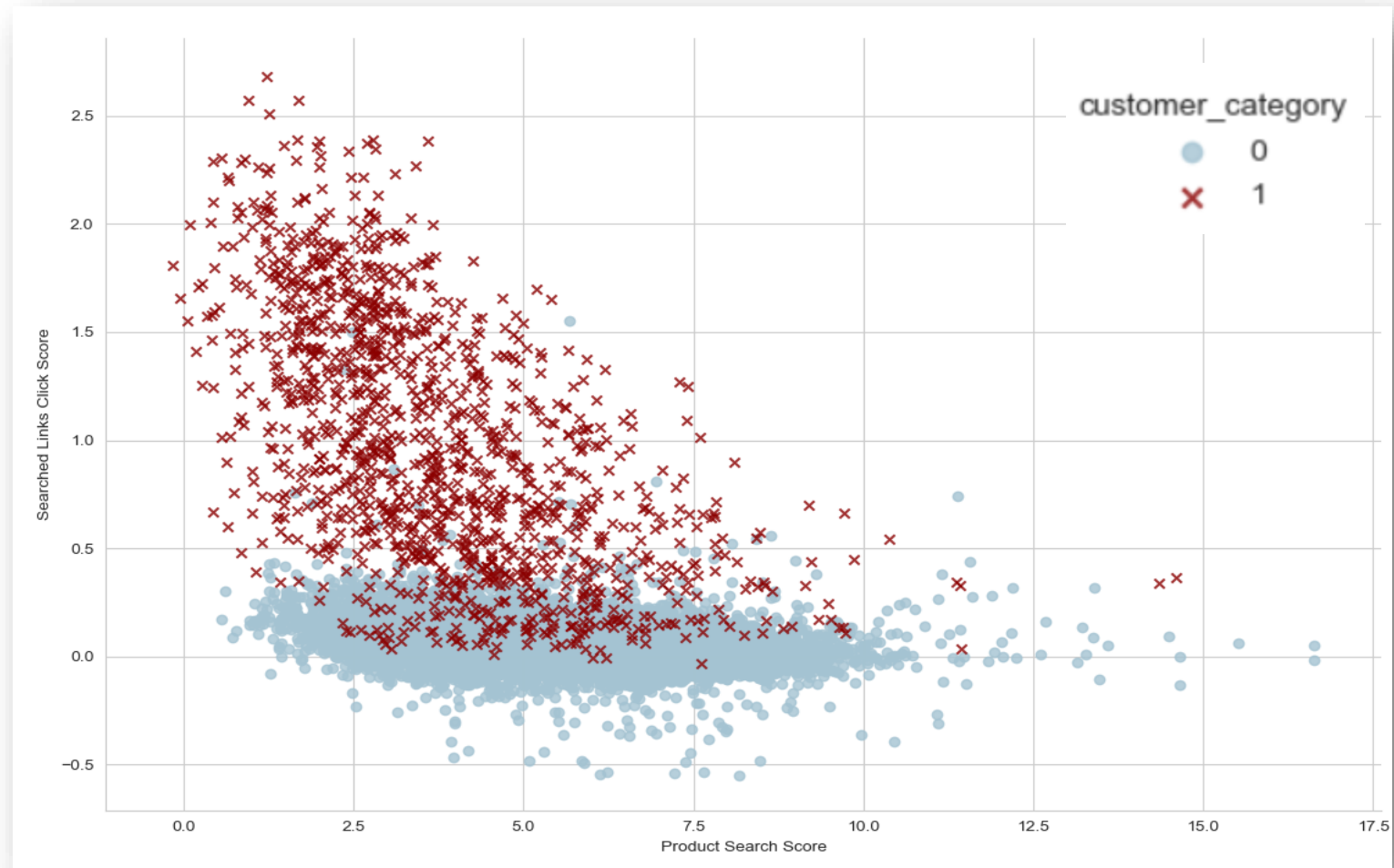
- Category '0' customers have a normal distribution of visit score.
- Most of the Category '1' customers have visit score between 0 and 12. Thus, it seems that category '1' customers are not very regular.



SEARCH ANALYSIS

Analysis of searching process among different users based on quality/price of products they search for ?

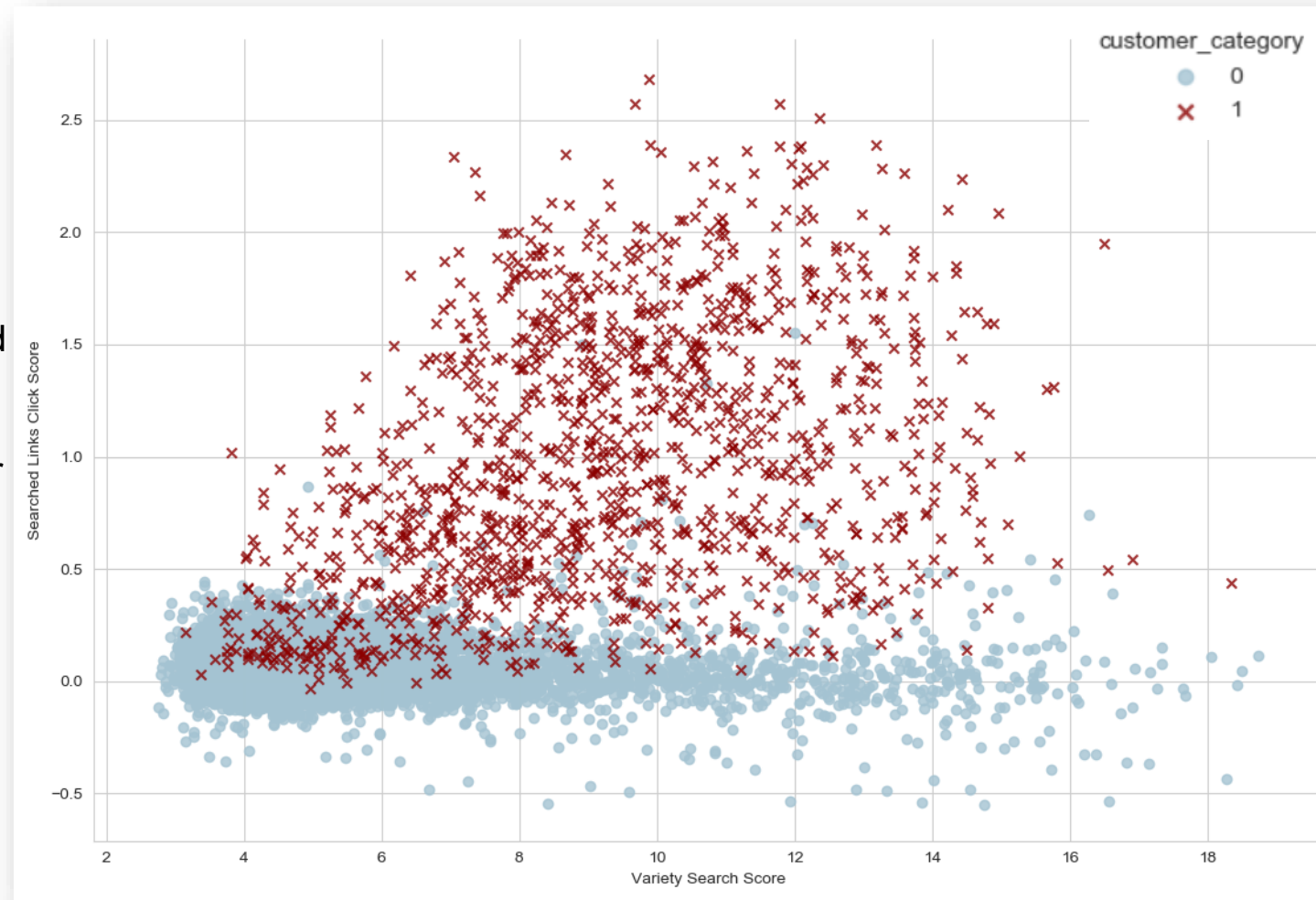
- Customers that tend to click more on the searched links fall under category '1'. But after an increase in score of product search, they tend to click less. We can infer that category '1' customers are more attracted towards cheaper products.
- Clicking rate of category '0' customers remain constant. There's not much rise in the same. There's something with the website which is not prompting these users to click more after searching for better and expensive products



SEARCH ANALYSIS

Analysis of searching process among different users based on variety of products they search for ?

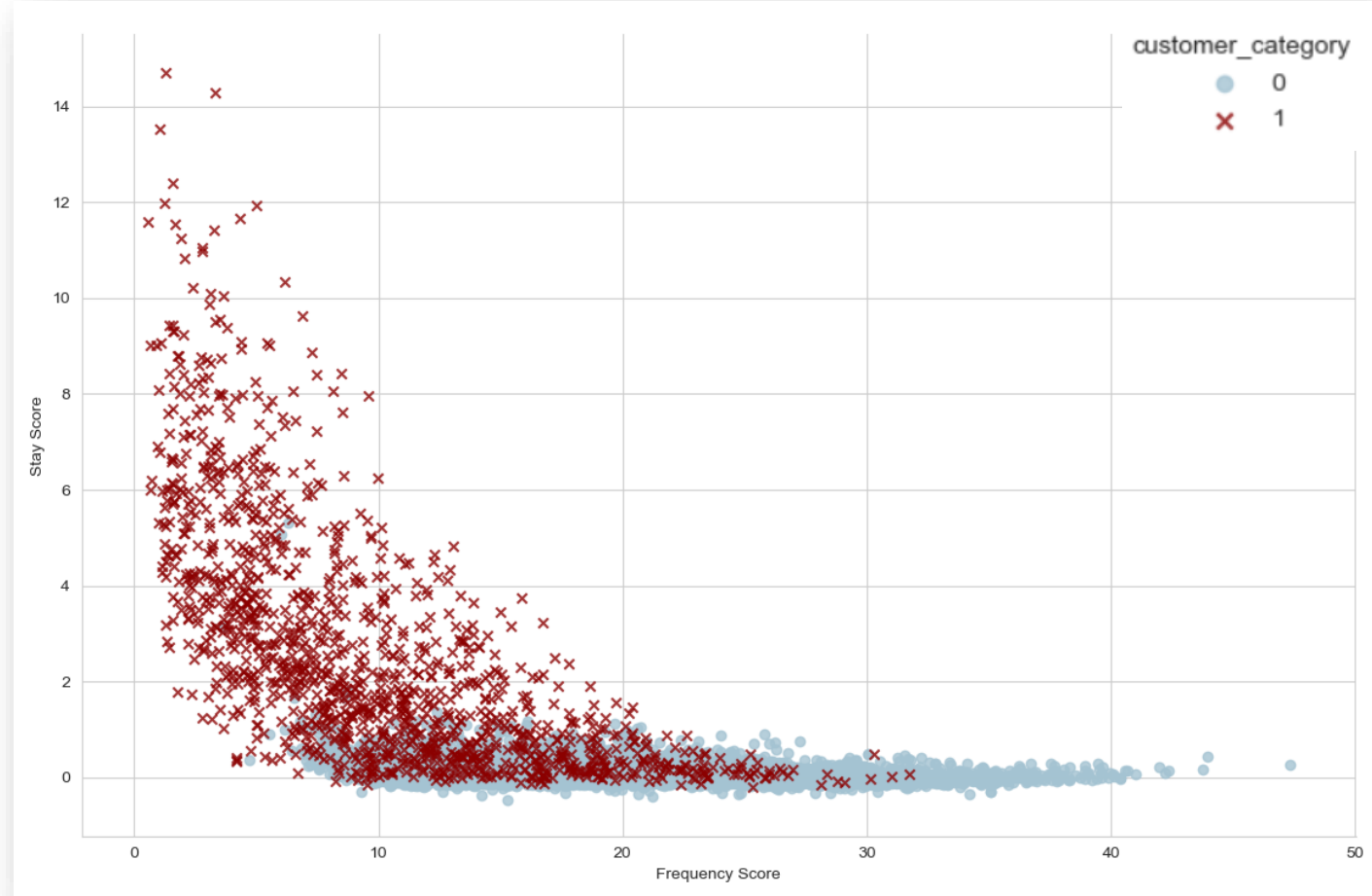
- Category '1' customers have a higher clicking rate. And with increase in varieties, clicking rate also increases. We can infer that these customers are more prone to click on searched products when there's variety in their searches.
- For category '0' customers, overall clicking rate is less as compared to the other category of customers.
- Also, on an average their clicking rate remains constant with changes in variety of searches.



STAY ANALYSIS

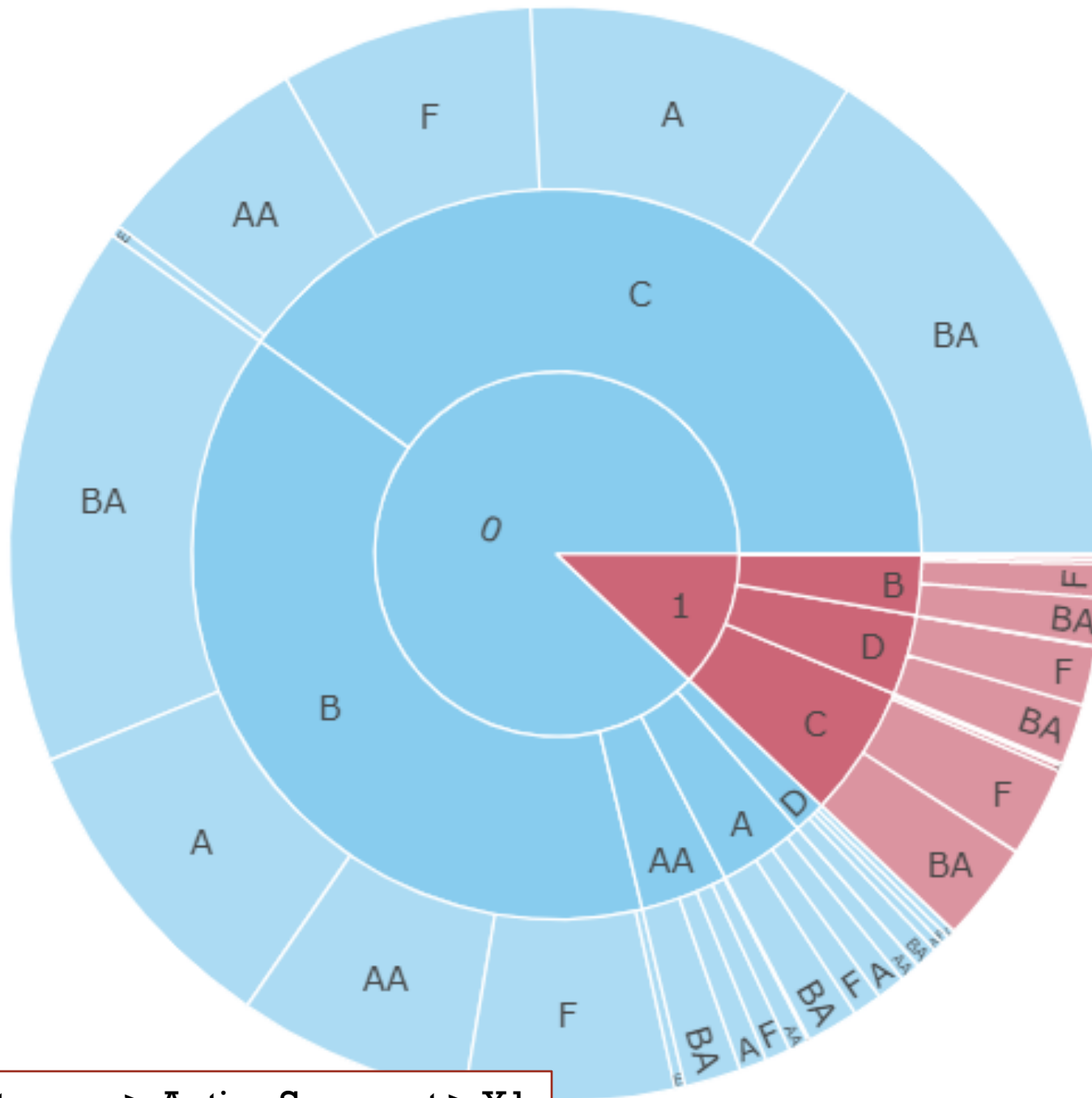
Analysis of how long a customer stay on the website based on their frequency of visits in a day?

- For category '1' customers, as their frequency increases, they then to stay more on our website.
- This means that category '1' customers slowly gets fond of our website. But most of these customers have very less frequency.
- For category '0' customers, there's no rise in stay score even after an increase in their frequency



LOYALTY & ACTIVITY ANALYSIS

- How is the distribution among customers based on their loyalty and activity?
- Most of the category '0' customers fall under activity segment of 'B' and 'C'. In these customers, loyalty of 'BA' level is most prevalent.
- Almost one half of category '1' customers are of 'C' active segment and another half is divided between 'B' and 'D'.



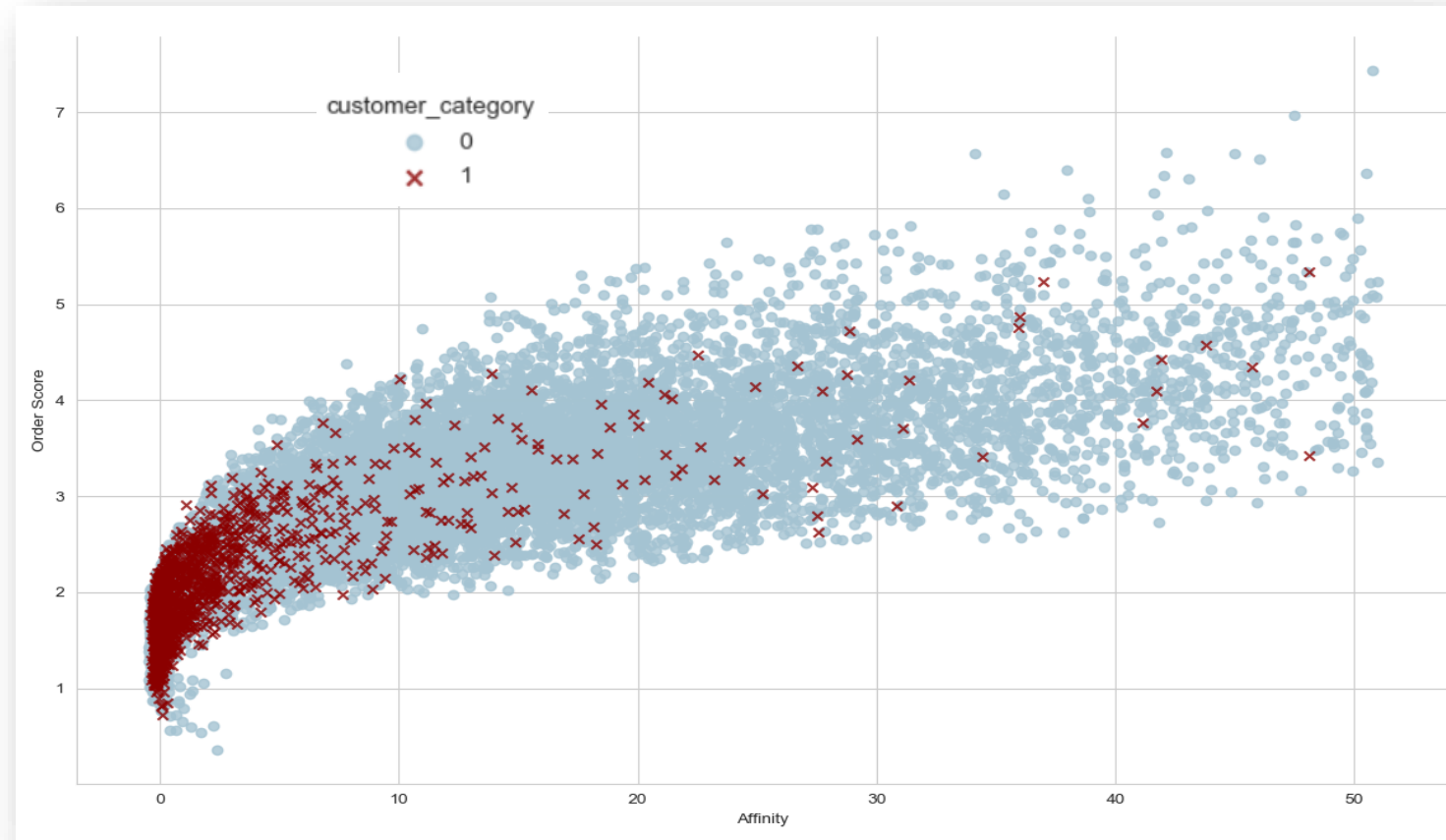
Category > Active Segment > X1



AFFINITY ANALYSIS

Does affinity towards the website affects orders?

- With increase in affinity, number of non-returned orders also increases.
- This would align with the hypothesis that as you like a website more, you are prone to make more number of orders.
- Also, category 1 users has lower affinity score and tend to make less orders.



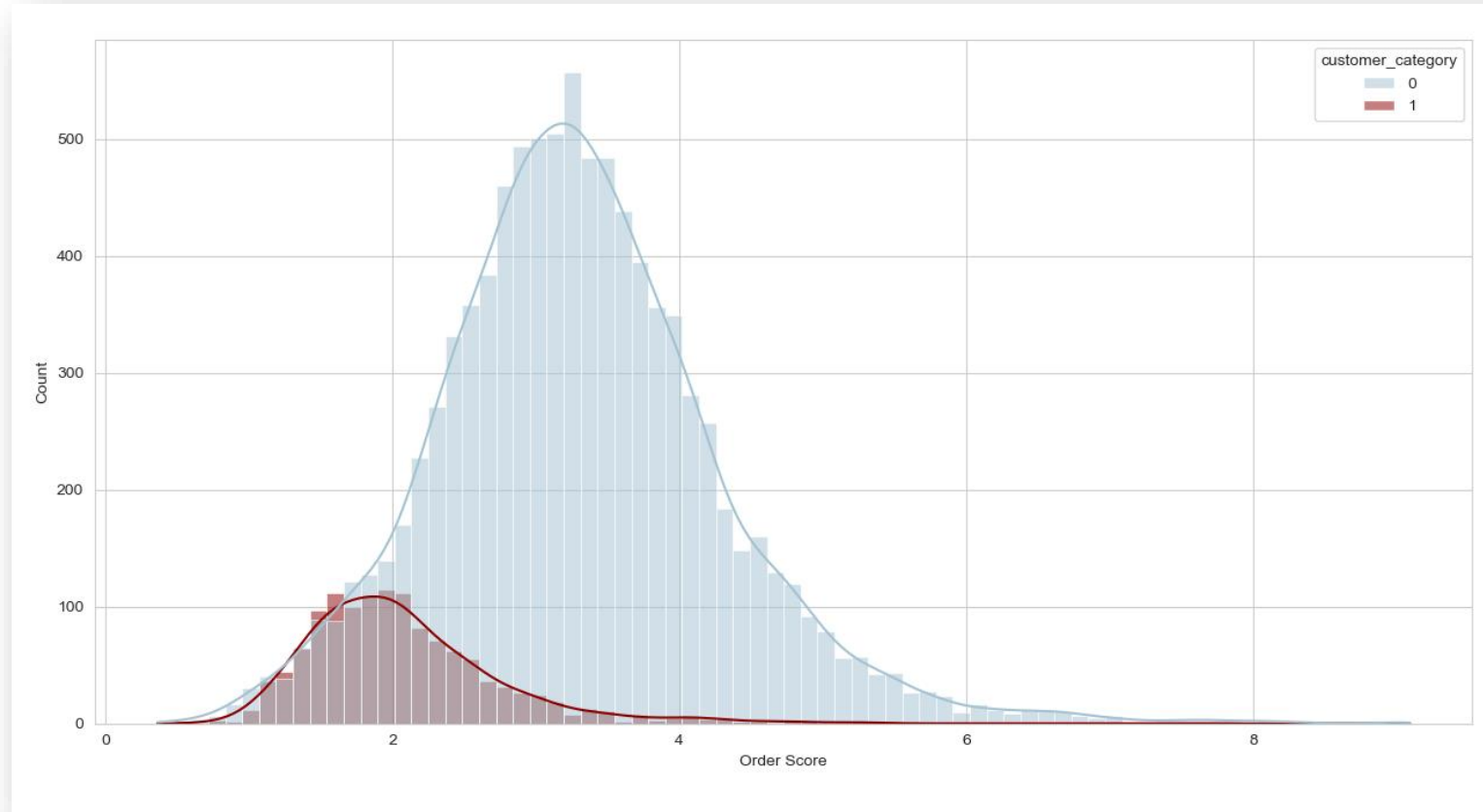
****Removed datapoints with outliers in Affinity Score**



ORDER SCORE ANALYSIS

How order score is distributed among the two type of customers ?

- Category '0' customers have a normal distribution of order score.
- Most of the Category '1' customers have visit score between 1 and 3. These customers do not have much non-returned orders as compared to the other category.





CONCLUSION

Based on above analysis, following are the habits of category 'l' customers that one can infer:-

- They are very irregular.
- They are not very loyal and have less affinity towards the website.
- Even after staying more on the website and having a higher clicking rate, they tend to make less orders.