Business Problem:

FLO store is putting up a new brand of women's shoes for sale. The new brand of the store is above the general customer preferences. For this reason, it is requested to contact customers in the profile who may be interested in the promotion of the brand and the sale of products. Shoppers from loyal customers (champions, loyal\_customers) and women's categories are customers who need to be contacted in particular. The ID numbers of these customers will be forwarded to the sales and marketing department with a CSV file.

metin, ekran görüntüsü, renklilik, yazı tipi içeren bir resim

Açıklama otomatik olarak oluşturuldu

For each transaction in the dataset, the following records are kept:

* Date of the transaction,
* Unique customer number of the person making the purchase,
* Amount paid in the purchase,
* Information about the category (Women, Men, Children...) from which the purchase was made,
* Other specific details.

To extract the target customers to introduce the new brand from the dataset:

* Firstly, the dataset should be prepared for analysis, cleaning empty or incorrect values,
* RFM metrics should be calculated,
* RF scores should be calculated, and segments should be defined,
* According to the requests of the marketing and sales departments, the ID numbers of relevant customers should be extracted from the dataset.

Step 1: Data Preparation:

* Total purchase amount and total purchase count were calculated for each customer. These calculations were added as new variables to the dataset,
* Missing values and returns were removed from the dataset,
* The data type of the purchase date was initially recorded as "object". It was changed to "datetime".

Step 2: Calculation of RFM Metrics:

* The analysis date is set to 2 days after the latest purchase date in the dataset. Using each customer's latest purchase date, the recency value was calculated,
* The number of purchases made by each customer was recorded as frequency,
* The total amount of purchases made by each customer was recorded as monetary value,
* Finally, these three variables were stored in a new dataframe named RFM along with customer ID numbers. This created a dataframe where we can access RFM values for each customer with their ID.

Step 3: Calculation of RF Score and Definition of Segments:

* Each of the RFM metrics was converted into scores ranking from 1 to 5 and recorded accordingly,
* The recency and frequency values were combined to create an RF variable. The RF variable consists of values like 11, 23, 43, 55, 45, etc. The first digit represents recency, and the second digit represents frequency.
* Segments were created based on RF values. These segments correspond to the segments in the RFM table. Then, segment naming was done for each customer based on their RF score.

After this three-step process, the dataset is ready. Now, from among the customers, any customer type in any segment (for example, "loyal customers" segment with "adult" customers) can be easily selected. Actions can be taken accordingly.

Step 4:

For the promotion of the new brand, the target segments will be "loyal customers" and "champions," specifically "Female" customers in terms of gender. By filtering the segments created through RFM analysis and the existing gender information in the dataset, appropriate customer IDs are extracted and sent to the sales and marketing department.