

Customer understanding and segmentation

Kohei Arai

1

3.: Analytical methods for gaining a deeper understanding of customers and grouping them

- Customer analysis: Analyze customer attributes and behavioral patterns to understand customer profiles
- RFM analysis: Classify customers based on the most recent purchase date (Recency), purchase frequency (Frequency), and purchase amount (Monetary)
- Decile analysis: Divide customers into 10 groups based on purchase amount and other indicators to understand the characteristics of top customers
- k-means clustering: Unsupervised learning method that automatically classifies customers into groups with similar characteristics
- Customer segmentation: Classify customers based on attributes such as age, gender, and purchasing behavior
- Cohort analysis: Analyze changes in behavior over time for customer groups acquired during a specific period

2

Customer Segmentation

- There are various approaches to customer analysis.
- An example of using KMeans clustering to cluster customer data and identify customer segments.
- Using two features, the number of purchases and the average purchase amount, we can use KMeans clustering: visualization to see how customer segments are formed and consider strategies and actions for different segments based on that.
- The methods, features, and number of clusters used may differ depending on the actual data and purpose → Select the appropriate approach according to the data and goals.

3

RFM analysis

- RFM analysis (Recency, Frequency, Monetary) is one of the methods that are useful for customer segmentation and marketing strategy planning.
- RFM consists of the following three elements:
 - Recency: The time since the customer's last purchase
 - Frequency: The frequency of purchases by the customer
 - Monetary: The total amount of money that the customer has purchased in the past
- A simple example of RFM analysis using Python generates random data and calculates the RFM score

4

RFM analysis

- Calculate Recency (number of days since last purchase), Frequency (purchase frequency), and Monetary (purchase amount) for each customer and calculate a score for each.
- Finally, add up these scores to calculate the RFM score.
- Using the RFM score, you can segment customers and apply different marketing strategies to each segment.

5

The different flavors of customer segmentation

A basic approach to segmentation relies on quantitative insights such as sales and conversion rates, limiting the perspective of customers to only representations of what they're like. A more advanced approach uses qualitative insights based on behavioral and transactional data, enabling organizations to have a clear idea of who customers are.

	QUALITATIVE INSIGHTS		QUANTITATIVE INSIGHTS	
TYPE	Persona-based	Behavior-based	Predictive	
WHAT IS IT?	An approach that blends data-driven segments with qualitatively developed representations of customers	An approach that divides customers into groups based on behavioral, transactional and other data sources	An approach that parses customers based on their propensities to take specific actions	
WHAT IS IT FOR?	Describing	Grouping	Predicting	
HOW DO YOU DO IT?	Conduct interviews with customers to determine archetypes	Apply a clustering algorithm (i.e., k-means)	Apply classification or logistic regression	
WHEN SHOULD YOU USE IT?	To customize for user types even if you have limited data and analytics resources	To discover new, unknown segments in large amounts of data	To identify which customers are likely to click, buy, churn, etc.	
	BASIC		ANALYTICS SOPHISTICATION → ADVANCED	

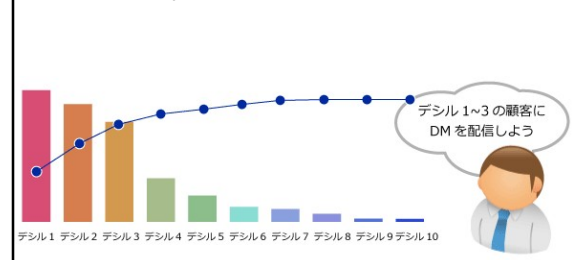
6

Decile analysis

- Decile analysis is a method of dividing data into 10 groups and analyzing the characteristics of each group.
- In this example of performing decile analysis using Python, random data is generated, divided into deciles, and the mean and median are calculated.
- Randomly assign data to deciles from 0 to 9 and calculate the mean and median for each decil.
- When applying decil analysis, it is important to select appropriate indicators and calculation methods depending on the data set and the purpose of the analysis.

7

Decil Analysis

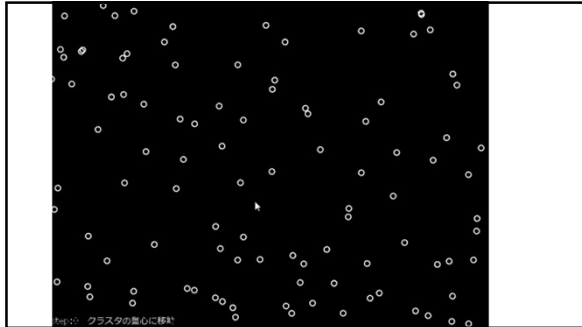


8

k-means clustering

- Python code using the k-means clustering algorithm as an example of clustering.
- This example uses the scikit-learn library.
- Generate two-dimensional random data and apply k-means clustering
- The `n_clusters` parameter specifies the number of clusters
- Visualize the data using the predicted clusters to which the data points belong
- This code is just an example, and when applying it to real data, you need to consider the appropriate number of clusters and data preprocessing
- There are various other clustering methods, so you need to choose the most suitable method depending on the nature of the problem

9



10

Cohort analysis

- Cohort analysis is a method of grouping customers who registered or purchased during a specific period into the same group (cohort) and tracking and comparing their subsequent behavior and transactions.
- In this example of basic cohort analysis in Python, we divide customers into cohorts based on their registration date and track their subsequent behavior (e.g. usage and purchases).

11

Cohort analysis

- Create a cohort for each month of registration and calculate the monthly difference from the first purchase: Then, calculate the average monthly purchase amount for each cohort and visualize it → You can visually understand the impact of the passage of time since a customer registered and the purchasing trends for each cohort.
- Use cohort analysis to make decisions about marketing, product improvements, etc.

12

The average monthly purchase amount for each cohort

- Create a cohort for each month of registration and calculate the monthly difference from the first purchase:
- Then, calculate the average monthly purchase amount for each cohort and visualize it → You can visually understand the impact of the passage of time since a customer registered and the purchasing trends for each cohort. Use cohort analysis to make decisions about marketing, product improvements, etc.

13