

Descriptive Methods in Data Mining

2022-2023

Project Description

Our Imaginary Enterprise:

Wonderful Wines of the World

Wonderful Wines of the World (WWW) is a ten-year-old enterprise that seeks out small, unique wineries worldwide and brings its wines to its customers. Its mission is to delight its customers with well-made, unique, and exciting wines that would never travel far beyond their points of origin.

WWW sells wines through catalogs (electronic and physical), websites, mobile app, and ten small stores in major cities around the USA. Customers can purchase at the stores, by telephone (after looking at the catalog), or through the website/mobile app.

Through aggressive promotion in wine and food magazines, WWW now has 350,000 customers in its database. Most customers are highly involved in wine, entertain frequently, and have sufficient money to indulge their passion for wine. WWW sometimes offers wine accessories as well – wine racks, cork extractors, etc.

WWW is trying to make use of the database it started about five years ago. So far, it has simply mass-marketed everything. All customers get the catalog, and there are no loyalty programs or attempts to identify target markets for cross selling opportunities.

Now, WWW wants to "get smart" about its database, start differentiating customers and develop more focused programs.

WWW has provided one random sample of 10 000 of its customers from its active database. These customers have purchased something from WWW in the past 18 months (after 18 months with no purchase, a person is eliminated from the active database). With these data you should perform 2 different segmentations, using the clustering algorithms discussed in class. The objective of the first segmentation is to identify engagement/value segments, in other words, we want to know which are the different groups of customers that we have and how

valuable they are to the company. This should lead you to select a subset of the variables available in the database. The second segmentation should be focused on buying behavior, by this we mean that we would like to understand the different groups of customers, based on the different types of wines they tend to buy more frequently (e.g. dry red wine, sweet red wines, etc.). Finally, you should be able to give a consolidated overview of the two segmentations, so that we can develop marketing plans for each segment.

Variables in the customer table

Variable	Description
CUSTID	customer ID number
DAYSWUS	number of days as a customer
AGE	customer's age
EDUCATION	years of education
INCOME	household income
KIDHOME	1=has child under 13 yo living at home
TEENHOME	1=has teen (13 to 19 yo) living at home
FREQ	number of purchases in past 18 months
RECENCY	number of days since last purchase
MONETARY	total sales to this person in past 18 months
LTV	customer lifetime value
PERDEAL	% purchases bought on discount (units)
DRYRED	% of wines that were dry red wines
SWEETRED	% of wines that were sweet red wines
DRYWH	% of wines that were dry white wines
SWEETWH	% of wines that were sweet white wines
DESSERT	% of wines that were dessert wines (port, sherry, etc.)
EXOTIC	% of wines that were exotic wines
WEBPURCH	% of purchases made on website/app
WEBVISIT	average number of visits to website/app per month
ACCESS	number of accessories bought in past 18 months

Note: DRYRED + SWEETRED + DRYWH + SWEETWH + DESSERT = 100%