Cluster Analysis of Wine Market Segmentation – A Consumer Based Study in the Mid-Atlantic USA

Ramu Govindasamy¹, Surendran Arumugam², Jingkun Zhuang¹, Kathleen M. Kelley³ and Isaac Vellangany¹

¹Department of Agricultural, Food and Resource Economics, Rutgers - The State University of New Jersey
²Imayam Institute of Agriculture & Technology, Thuraiyur, Tiruchirappalli, Tamil Nadu, India
³Department of Plant Science, The Pennsylvania State University, USA

*Corresponding author: govind@sebs.rutgers.edu

ABSTRACT

The U.S. wine market has been steadily increasing over the past 15 years. The number of wineries has increased from 2688 in 1999 to 8862 in 2016. About 7% of all those wineries are located in the Mid-Atlantic region, which includes New Jersey, New York, and Pennsylvania. However, competition has been rising as the market grows. Many foreign wine companies from Europe, South America, and Oceania, are either selling or planning to sell their products to the fast-growing U.S. wine market. It is important to promote local wine industry in the U.S. In this connection, this study has attempted to predict the segment of Mid-Atlantic wine market - based on purchasing behavior, attitudes, and social demographic attributes. A Cluster Analysis used to segment the Mid-Atlantic wine market into four clusters namely Class 1 Detractors, Class 2 Enthusiasts, Class 3 Neutral and Class 4 Advocators. Class 1. Detractors are the cluster that is the most unlikely to buy local wine. Neatly, 67.4% of Detractors indicated that they had never obtained local wine before. 2. Enthusiasts and Class 4 Advocators are the target market of Mid-Atlantic local wineries and hence we should pay more attention to these two market segments. 74.5% of Class 2 indicated that they had bought wine from the Mid-Atlantic wine region. About 60% of Class 4 Advocators stated that they had purchased local wine before. The characteristics of Class 4 are very similar to Class 2. The chance of Class 3 Neutral to buy local wine is 50/50. They drink and buy wine more frequently than consumers in other clusters. Typically, we do not recommend Mid-Atlantic wineries to target this market segment, unless they want to expand their market beyond Class 2 and Class 4. The assessment of perception of consumers will help the producers, wholesalers, and retailers to target ultimate consumers and specific market segments.

Keywords: Wine, Purchase Behavior, Consumer Behavior, Logistic Regression, Cluster Analysis, Market Segmentation, Marketing Strategy, Decision Making, Mid-Atlantic, NY, NJ, PA.

In 1975, many considered the U.S. wine business still as a backwater — quiet but soon to make some big waves. The consumption estimated to 31.8 mhl (Millions of Hectolitres/1 Hectolitre = 100 liters) in 2016, confirmed its position as the top global consumer since 2011, and saw growing domestic demand around 31.0 mhl compared with 2015 consumption (+2.5% of change) (OIV-International Organization of Vine and Wine, 2017). The prediction of sales growth ranges from 10 to 14 per cent for the premium wine segment in 2017, up from nine to 13 per cent in 2016 (McMillan, 2017). The research conducted related to wine consumption in several countries and the consumer purchasing behavior (Bruwer et al. 2014, Goldsmith et al. 1998; Goldsmith and Hofacker, 1991; Orth and Bourrain, 2005a).

In the U.S., wine consumer behavior exhibits different aspects of an individual that influence the consumer’s final choice in the decision-making process. Understanding consumer behavior in the context of the US market is essential, as the USA has
accounted for the most significant volume of wine sales since 2010 (Wine Institute, 2014). A consumer’s perceptions, attitudes, and characteristics influence purchasing decision (Brager, 2014; Janeen et al. 2015). Similarly, wine consumer’s level of knowledge and experience in purchasing wine can also affect their choice (Mitchell & Greatorex, 1989; O’Cass and McEwen, 2004; Miller and Chadee, 2008; Janeen et al. 2015). In addition, demographic characteristics of consumers play a significant role in the wine consuming decision (Dodd, 1995; Dodd, Laverie, Wilcox, & Duhan, 2005; Johnson & Bastian, 2007; Janeen et al. 2015). Likewise, People of different marital status have differences in their alcohol consumption. The alcohol consumption either increases or decreases as people’s marital status varies (Power, Rodgers, & Hope, 1999).

In particular, region was more important for high involvement consumers and price was more important for low involvement consumers. Many studies used simulated choices to measure the importance of price, region, brand, and awards (Oliver, 1999; Lockshin et al. 2006; Miller and Chadee, 2008; Ha and Jang, 2013). Previous experience, personal recommendations, and the taste of the wine were the significant factors, also preferences based on age, involvement level, and the geographical part influenced wine consumers purchasing decision (Casini et al. 2009; Janeen et al. 2015). Regarding wine purchasing, online wine purchases were of higher priced wines, and the size of the shopping basket was more substantial in the online environment, probably because of the shipping charges for 6 or 12 bottles. A review of the postal code of the online purchasers compared to a separate sample of in-store purchasers showed the online purchasers tended to live in city centers, where parking and transporting wine would be difficult. The authors speculated that online purchases were aimed at expensive and hard to find wines, whereas in-store purchases were convenient (Stening and Lockshin, 2001). In addition, there are differences among online wine buyers, where some are very comfortable buying online, and others are not (Bressolles and Durrieu (2010). This contradicts the result of the study which was published by Bruwer and Lesschaeve (2012) about the socio-demographic profile, attitudes, and perceptions of tourists at the winery (Alebaki and Iakovidou, 2011; Kolyesnikova and Dodd, 2008; Gill et al. 2007). There is evidence that tourism benefits winery substantially, but attracting tourists is similar to drawing any other type of buyer: higher involvement and heavier buyers are more likely to visit and buy wine.

In general, ranges of several factors affect consumers’ purchasing behavior, which leads to difference in ways in which consumers’ approach wines. Socio-demographic differences, personal characteristics, and sensory preferences are essential to distinguish new from longer-term wine buyers. Therefore, it is crucial to understand consumers’ behavior for wine and it adds some commentary about the way forward and a discussion to provide the most usable results for the development of marketing in the Mid-Atlantic wine sector. Against this background, this study has attempted to predict the segment of Mid-Atlantic wine market based on the purchasing behavior, attitudes, and social demographic attributes. The findings of this study may be useful to local wineries to understand better the wine consumers purchasing behavior, beliefs, and social demographic characteristics to target specific wine market target in the mid-Atlantic region of the United States.

Methodology

The central research question is about how to target this wine market segment in the Mid-Atlantic region. We can refine this question into several small objectives. To identify the demographics and behaviors that describe Mid-Atlantic wine buyers. To identify wine consumers’ preferences on different wine attributes. To segment wine consumers into several groups and study the characteristics of each group. This study used the data that Penn State University collected via internet survey in 2009. This survey helped us to quantify consumer wine purchases and preferred varieties, identify the demographics and behaviors that describe Mid-Atlantic wine buyers. This study especially investigated consumer segmentation by employing Cluster Analysis (Punj, 1983; Donald et al. 1997). More discussions were made on how to maintain business with current buyers, as well as how to target other less likely buyers by giving an understanding of their preferences.

The marketing cost is one of the concerns of local wineries. Local wineries cannot afford the cost if
the marketing strategy is dependent upon targeting an entire mass market. The importance of market segmentation is that it allows a business to target consumers with specific needs and wants (Donald et al. 1997; Hofstede et al. 1999; Antreas, 2000; Bruwer, J. & Li E., 2007; Torres & Kunc, 2016). In the end, this benefits the company because they can use their corporate resources more efficiently and make better strategic marketing decisions. In this section, we employed Cluster Analysis to class wine consumers into several groups. Different groups will have different demographics and preferences. Many cluster analysis methods are available out there. We used the \texttt{hclust} function in R to achieve the hierarchical clustering (Ward, 1963). We used the Ward linkage when applying the hierarchical clustering. The hierarchical clustering method defines the cluster distance between two clusters to be the maximum distance between their individual components. At every stage of the clustering process, the two nearest groups were merged into a new cluster. The process is repeated until the whole data set is agglomerated into one single cluster (Kaufman and Rousseeuw, 1990; Legendre and Legendre, 2012).

Two-Way Contingency and Chi-Square Independence Test of Wine Consumer Clusters

To study the differences between the four market segments, we performed Cluster Analysis, a two-way contingency table and Chi-square independence tests. The Chi-square independence test is to test whether two variables are associated or not. In the case of the state variable, our hypotheses are:

- $H_0$: State and Wine Consumer Clusters are not associated.
- $H_1$: State and Wine Consumer Clusters are associated.

The idea behind the chi-square independence test is to compare the observed frequencies with the frequencies we would expect if the null hypothesis of non-association is correct. Equation (1) is the test statistic used for this comparison. Represents the expected frequencies whereas refers to observed frequencies. Equation (2) was used to estimate $E$.

$$
\chi^2 = \sum (O - E)^2 / E \quad \ldots \text{(1)}
$$

$$
E = \frac{\text{row} \times \text{column}}{n} \quad \ldots \text{(2)}
$$

The two-way contingency table shows us the distribution of the data in each group, which allows us to compare the difference of the levels in the categorical variables in each group. Based on the two-way contingency table, derived the Chi-square tests to test if each of the variables is associated with the response variable.

RESULTS AND DISCUSSION

Fig. 1 is the dendrogram of our cluster analysis. From the dendrogram, it is not very clear about how many clusters can be chosen. It can be cut at 2, or 3 or 4 clusters.

![Dendrogram of Cluster Analysis](image1)

**Fig. 1: Dendrogram of Cluster Analysis**

To decide the optimal number of clusters, we plotted an elbow plot as shown in Fig. 2. The elbow is very clear; it appears at the fourth cluster.

![Elbow plot of optimal number of clusters](image2)

**Fig. 2: Elbow plot of optimal number of clusters**

According to the elbow plot, we chose to keep four clusters. We employed a simple ANOVA to test whether there were significant differences between any two classes. ANOVA results as below shows that there is significant difference at least between two classes.

$$
\begin{array}{|c|c|c|c|c|}
\hline
\text{Df} & \text{Sum Sq.} & \text{Mean Sq.} & \text{F value} & \text{Pr(>F)} \\
\hline
\text{Class} & 3 & 5.71 & 1.9049 & 7.739 & 4.01E-05 \quad *** \\
\text{Residuals} & 1242 & 305.7 & 0.2461 & & \\
\hline
\end{array}
$$
Table 1 is a cross table of the BUY variable and wine consumer clusters. 67.4% of Class1 do not buy local wine, which means Class-1 is very unlikely to buy local wine, so we named this class as Detractors. 74.5% of Class-2 buys local wine. This is a very high percentage. We call Class-2 Enthusiasts. The third class is considered as Neutral since about 50% of Class-3 buys local wine. The last class has 59.3% who purchases local wine. It is not as high as Class-2 but is still more likely to buy local wine compared to Class-1 and Class-3.

In Table 2, we can see that the Chi-Squared test results for state, age, income, education, and marital status are significantly related to the wine consumer clusters.

Table 3 contain variables that were tested to be statistically significant. Demographics such as state, age, gender, education, income, occupation and marital status are all associated with the market

Table 1: Cross table of BUY variable and clusters

<table>
<thead>
<tr>
<th>BUY</th>
<th>BUY Buy Class1 Detractors (n=574)</th>
<th>BUY Buy Class2 Enthusiasts (n=306)</th>
<th>BUY Buy Class3 Neutral (n=189)</th>
<th>BUY Buy Class4 Advocators (n=177)</th>
<th>BUY Buy Total (N=1246)</th>
</tr>
</thead>
<tbody>
<tr>
<td>YES</td>
<td>32.6%</td>
<td>74.5%</td>
<td>49.2%</td>
<td>59.3%</td>
<td>49.2%</td>
</tr>
<tr>
<td>NO</td>
<td>67.4%</td>
<td>25.5%</td>
<td>50.8%</td>
<td>40.7%</td>
<td>50.8%</td>
</tr>
</tbody>
</table>

Table 2: Contingency Table and Independence Test of Wine Consumer Clusters

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Class1 (n=574)</th>
<th>Class2 (n=306)</th>
<th>Class3 (n=189)</th>
<th>Class4 (n=177)</th>
<th>Total (n=1246)</th>
<th>Chi-Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>State</td>
<td>NJ</td>
<td>21.6%</td>
<td>20.6%</td>
<td>13.2%</td>
<td>16.9%</td>
<td>19.4%</td>
</tr>
<tr>
<td></td>
<td>NY</td>
<td>44.1%</td>
<td>42.8%</td>
<td>65.1%</td>
<td>50.8%</td>
<td>47.9%</td>
</tr>
<tr>
<td></td>
<td>PA</td>
<td>34.3%</td>
<td>36.6%</td>
<td>21.7%</td>
<td>32.2%</td>
<td>32.7%</td>
</tr>
<tr>
<td>Age</td>
<td>21-24</td>
<td>20.2%</td>
<td>13.1%</td>
<td>22.2%</td>
<td>17.5%</td>
<td>18.4%</td>
</tr>
<tr>
<td></td>
<td>25-34</td>
<td>23.9%</td>
<td>25.8%</td>
<td>41.3%</td>
<td>26.0%</td>
<td>27.3%</td>
</tr>
<tr>
<td></td>
<td>35-44</td>
<td>27.9%</td>
<td>29.1%</td>
<td>24.3%</td>
<td>28.2%</td>
<td>27.7%</td>
</tr>
<tr>
<td></td>
<td>45-64</td>
<td>28.0%</td>
<td>32.0%</td>
<td>12.2%</td>
<td>28.2%</td>
<td>26.6%</td>
</tr>
<tr>
<td>Q1a Wine Drink. Freq.</td>
<td>Daily</td>
<td>3.5%</td>
<td>2.6%</td>
<td>28.0%</td>
<td>5.1%</td>
<td>7.2%</td>
</tr>
<tr>
<td></td>
<td>A few times a week</td>
<td>20.6%</td>
<td>25.5%</td>
<td>39.2%</td>
<td>28.2%</td>
<td>25.7%</td>
</tr>
<tr>
<td></td>
<td>About once a week</td>
<td>15.2%</td>
<td>19.6%</td>
<td>18.5%</td>
<td>29.4%</td>
<td>18.8%</td>
</tr>
<tr>
<td></td>
<td>2 to 3 times a month</td>
<td>23.7%</td>
<td>29.4%</td>
<td>11.6%</td>
<td>22.0%</td>
<td>23.0%</td>
</tr>
<tr>
<td></td>
<td>About once a month</td>
<td>14.6%</td>
<td>9.2%</td>
<td>1.1%</td>
<td>9.6%</td>
<td>10.5%</td>
</tr>
<tr>
<td></td>
<td>A few times a year</td>
<td>22.5%</td>
<td>13.7%</td>
<td>1.6%</td>
<td>5.6%</td>
<td>14.8%</td>
</tr>
<tr>
<td>Q2 Wine Buying Freq.</td>
<td>Daily</td>
<td>0.3%</td>
<td>0.0%</td>
<td>12.2%</td>
<td>1.7%</td>
<td>2.2%</td>
</tr>
<tr>
<td></td>
<td>A few times a week</td>
<td>3.3%</td>
<td>1.6%</td>
<td>21.7%</td>
<td>3.4%</td>
<td>5.7%</td>
</tr>
<tr>
<td></td>
<td>About once a week</td>
<td>9.2%</td>
<td>11.4%</td>
<td>30.2%</td>
<td>16.4%</td>
<td>14.0%</td>
</tr>
<tr>
<td></td>
<td>2 to 3 times a month</td>
<td>17.2%</td>
<td>25.8%</td>
<td>22.2%</td>
<td>27.7%</td>
<td>21.6%</td>
</tr>
</tbody>
</table>
Cluster Analysis of Wine Market Segmentation – A Consumer Based Study in the Mid-Atlantic USA

Table 3: Profile of Wine Consumer Clusters

<table>
<thead>
<tr>
<th></th>
<th>Class1</th>
<th>Class2</th>
<th>Class3</th>
<th>Class4</th>
<th>Total</th>
<th>Chi-Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(n=574)</td>
<td>(n=306)</td>
<td>(n=189)</td>
<td>(n=177)</td>
<td>(n=1246)</td>
<td></td>
</tr>
<tr>
<td>State</td>
<td>—</td>
<td>More PA</td>
<td>63%</td>
<td>53%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>less NY</td>
<td>is from NY</td>
<td>is less PA than average</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>—</td>
<td>Oldest</td>
<td>Youngest</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Drinking Freq.</td>
<td>Least frequent</td>
<td>Slightly less frequent than average</td>
<td>Most frequent, heavy wine drinkers</td>
<td>Moderate wine drinkers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Buying Freq.</td>
<td>Least frequent</td>
<td>Slightly less frequent than average</td>
<td>Most frequent.</td>
<td>Moderate</td>
<td></td>
<td></td>
</tr>
<tr>
<td>First Drink was wine</td>
<td>25%YES, lower than average (33%)</td>
<td>26% YES, lower than average</td>
<td>65% YES,</td>
<td>37% YES</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>75% female</td>
<td>65% female</td>
<td>37% female</td>
<td>58% female</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Education</td>
<td>lower</td>
<td>higher</td>
<td>higher</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Family Income</td>
<td>Lowest</td>
<td>middle</td>
<td>middle</td>
<td>Highest</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Occupation</td>
<td>More retired and unemployed</td>
<td>More retired and unemployed</td>
<td>Less retired &amp; unemployed</td>
<td>Less retired and unemployed</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Marital Status</td>
<td>0.8%</td>
<td>1.3%</td>
<td>0.6%</td>
<td>2.3%</td>
<td>1.1%</td>
<td>***</td>
</tr>
</tbody>
</table>

Chi-Squared values:
- F=88.828, p < 2.2e-16
- F=50.593
- F=23.27
- F=31.806
- F = 19.583
- F = 0.02067
- F = 9.628e-06
- p = 2.244e-15
- p = 0.006842
- p = 0.0007108
- p = 9.628e-06
- p = 0.006842
- p = 0.02067
- ***
segments that we derived from Cluster Analysis. There is a correlation between buying and drinking frequency with our market clusters.

Most of Class1 the Detractors are infrequent wine drinkers. They tend to have lower education level and less income. Most of Class2 the Enthusiasts are in their 40’s and 50’s. They are moderate wine drinkers. Usually, they drink wine once a week or two to three times a month. They tend to have a higher education level and mid-level income. Many of them are married or in a partnership at least. Class3 Neutral are frequent wine drinkers. Most of them are in their 20’s and 30’s. They also have a higher education level and mid-level income like Class2. Many of them are male. Class2 Advocators are moderate wine drinkers. They tend to have a higher education level and the highest income compared to other classes. Many of them are single.

Class 2 the Enthusiasts and Class4, the Advocators, are the target market of local wineries. If local wineries want to target larger markets, Class3 the Neutral can also be a likely potential market. However, Class1 the Detractors is the market segment that local wineries should avoid.

CONCLUSION

We have employed cluster techniques to address local wine market segments in the Mid-Atlantic region. This paper classified local wine market into four clusters namely Class1 Detractors, Class2 Enthusiasts, Class3 Neutral and Class4 Advocators.

Segment of Unlikely Buyers - Class1 Detractors

Class1 Detractors is the cluster that is the most unlikely to buy local wine. Mid-Atlantic wineries should avoid this market segment when designing marketing strategies. 67.4% of Class1 Detractors indicated that they had never bought local wine before. Consumers in this cluster are infrequent wine drinkers. Compared to other clusters, less of Class1 Detractors have a Bachelor’s degree. Besides, they tend to have a lower income level.

Segments of Likely Buyers- Class2 Enthusiasts and Class4 Advocators

Class2 Enthusiasts and Class4 Advocators are the target market of Mid-Atlantic local wineries. More attention should be paid to these two market segments. Class2 Enthusiasts are the most likely to purchase local wine compared to other three clusters. 74.5% of Class2 indicated that they had bought wine from the Mid-Atlantic wine region. Most of the Class2 Enthusiasts are in their 40s or 50s. They are moderate wine drinkers (once a week, or 2 to 3 times a week). Many of them have bachelor’s degree or a higher education level. They tend to have a mid-income level as well. About 60% of Class4 Advocators stated that they had bought local wine before. The characteristics of Class4 are very similar to Class2. Many of them are moderate wine drinkers with a Bachelor’s degree. However, unlike many of Class2, they are married or in a partnership, while most of Class4 are single.

Segment of Neutral Buyers - Class3 Neutral

The chance of Class3 Neutral to buy local wine is 50/50. Most consumers in Class3 are males in their 20s or 30s. They drink and buy wine more frequently than consumers in other clusters buy. Many of them hold a Bachelor’s degree and have a mid-income level. Typically, we do not recommend Mid-Atlantic wineries to target this market segment, unless they want to expand their market beyond Class2 and Class4. The results provide implications for wine market to target ultimate consumers and specific market segments, including offering wine brands with a wider array of varietals, various price tiers, creative packaging and sustainable messages in their presentation. However, further research will be needed to examine the relationships between on-farm wine production and location, land and operator characteristics, the mix of products and marketing outlets, and relative costs and returns associated with wine industries.

REFERENCES


Cluster Analysis of Wine Market Segmentation – A Consumer Based Study in the Mid-Atlantic USA


