**PRACTICE PROBLEM SET**

**FACTOR ANALYSIS AND DATA REDUCTION METHODS**

**AEM 6700**

**NAME: ANDRES CASTANO ZULUAGA**

**NETID: AC986**

1. In the winery tasting room experience problem, the factor analysis is used as a data reduction method to detect dimensions or factors that in some sense describe in more broader way the experience of the customers (visitors) in the winery tasting room. This exercise is important, especially from a managerial point of view because the NYS winery operators want to understand what are the factors driving the consumer satisfaction of the visitors, and given that the survey instrument contains 25 questions regarding their tasting room experience, it seems very difficult (and maybe inefficient) to the operators to evaluate or implement strategies based on 25 different variables. Instead, it would be helpful if they can identify some more general evaluative dimensions. The factor analysis is designed to find this more general evaluative dimensions.
2. The general idea of the rotation of the factors is the redistribution of the variance among the number of factors that were selected. This redistribution is important because in the process of select the factors, we were keeping them due to the amount of variance that they can explain. Typically, the first factor account for a big part of the variance, and then, we select another factor due to its ability to explain the remaining variance, and so on. This procedure causes that the first factor is quite related with a big part of the variables, and that some variables can be heavily correlated with different factors (cross-loaded variables), which make very difficult the process to identify general factors.

The rotation help us to mitigate the problem of cross loaded variables, and at the same time keep the factors orthogonal to each other. After this, we use the new factor loadings (after the rotation) to create the factors. The factor loadings tell us the correlation of each variable with each factor. Variables with factors loadings +- 50 are considered necessary for practical significance. We create each factor with the variables that show higher factor loadings. If after the rotation some cross-loaded variable persist, we might remove it of the analysis and respecify the factor model, or we could ignore it and interpret the solution as it is. Which strategy implement depends on the objective of the analysis and the theoretical implications.

1. The purpose of factor analysis is to identify interrelated set of variables. To apply factor analysis, we need to demonstrate that the variables are sufficiently correlated to produce representative factors. One of the measures to quantify the degree of intercorrelations among the variables and therefore the appropriateness of the factor analysis is the Kaiser-Meyer-Olin (KMO) criterion. The Kaiser-Meyer-Olin test return values between 0 and 1. Greater KMO values show that correlations between pairs of variables can be explained by other variables, which means the factor analysis can be applied. Usually, a KMO equal or greater than 0.5 on overall, and for each variable, is used as rule of thumb. In our case, the overall KMO was 0.8616, and all the variables individually shown a KMO greater that 0.7. This result suggest that the factor analysis is appropriate for the winery testing room problem.
2. The only factor that shows a statistically significant effect on the overall consumer satisfaction is the “tasting protocol”. This factor is related with perception that the customers have about the flexibility in the choice of wines tasted, the space available for testing, the waiting time for the tasting to start, and the waiting time between samples. Given that the factor was calculated as the mean of the observed answers in this factor the interpretation must be done in general terms. In our case, the slope partial coefficient for the “tasting protocol” factor is 0.24, which means that an increase in one point in the tasting protocol measure, increase the overall satisfaction in 0.24 points. I will not interpret the remaining coefficients because are not statistical significant.
3. Figure 1 to 4 portray the Average Marginal Effect (AME) of different levels of consumer satisfaction on different measures of sales performance.

|  |  |
| --- | --- |
| **Figure 1:** Impact of overall consumer satisfaction on the probability to purchase. | **Figure 2:** Impact of overall consumer satisfaction on the Average number of bottles purchased. |
| purchase-impact.png | numberofbottles-impact.png |
| **Figure 3:** Impact of overall consumer satisfaction on the Average amount of purchase ($). | **Figure 4:** Impact of overall consumer satisfaction on Probability of repurchase. |
| moneyspent-impact.png | re-purchase-impact.png |

Note: The marginal effect presented is the Average Marginal Effect (AME).

1. **Table 1: Impact of Converting a Visitor from “Satisfied” (score=4) to “Highly Satisfied” (score=5) on Various Measure of Sales Performance:**

|  |  |  |
| --- | --- | --- |
| **Performance Measure** | **Sample Average** | **Impact** |
| Probability of purchase | 70% | 8.87 pp. |
| Average number of bottles purchased | 4.5 | 0.91 bottles |
| Average amount of purchase ($) | $62 | 8.23 ($) |
| Probability of repurchase | 88% | -0.997 pp. |

**Note:**   The impact reported is the difference between the Average Marginal Effects evaluated for customers with a level of satisfaction 4 and others with a level of satisfaction of 5. Stata coded attached.

Table 1 summarizes the results for the impact of our measure of costumer overall satisfaction on various measures of sales performance (probability of purchase, probability of repurchase, average number of bottles purchased, and average amount of purchase ($)). The results show that an increase in the consumer satisfaction from 4 to 5 increases the probability to purchase in 8.87 pp. and decreases the probability to repurchased in 0.99 pp.

On the other hand, the results suggest an increase in the consumer satisfaction from 4 to 5 increases the average number of bottles purchased in 0.91 bottles and the average amount of purchase in in 8.23 ($).