2014 3rd International Conference on Business, Management and Governance
IPEDR vol.82 (2014) © (2014) IACSIT Press, Singapore
DOI: 10.7763/IPEDR.2014.V82.9

Starbucks Revenue: Analyzing Consumer Purchases

Jennifer Appel, Simon Shi, and Fatemeh Emdad[+]
Babson College

Abstract. The purpose of this research is to identify relationships between different consumer purchasing behaviors and their implications on total revenue. A primary data sample of 106 Starbucks customers was collected at the Starbucks located at 101 Linden Square in Wellesley, Massachusetts in 2012. The variables analyzed included: weather, items purchased, sale amount, group size, and customer characteristics and behaviors. This study conducted multiple hypothesis tests, principal component analysis, and a logistic regression to further investigate these relationships. Ultimately, the study found that customer characteristics and behaviors, specifically whether the customer stayed or left after purchasing, had the largest influence over the sale amount.

Keywords: Hypothesis test, Principal Component Analysis, Logistic Regression, revenue

1. Introduction

In order to be financially successful, it is imperative for a company to focus on two important factors: customers and revenue. For Starbucks specifically, their loyal customers represent the most important factor, and this study was conducted so they could better understand their customers' behaviors. The variables included in the study give better insight into the customers' buying patterns and the factors that affect revenue the most. Through the findings, Starbucks should be able to enhance their customers' in-store experiences and increase their operating efficiency. Ultimately, from the results of the study, Starbucks can determine what affects the average sale price per customer the most.

In order to further understand how the variables impact Starbucks revenue, four hypothesis tests were completed to check the original hypotheses. The first hypothesis involves discovering the mean sale amount for each observed transaction. A hypothesis test was created to test whether the mean sale amount was greater than $3.50 per person. The next hypothesis was that the mean sale amount for customers who stayed after purchasing would be greater than for those that left. A hypothesis test was run to determine the mean sale amounts for customers who stayed versus left after purchase. Additionally, a hypothesis was formed around the proportion of each gender in the population. It was observed that many more women came to the Starbucks during the observed time periods, and so it is important to investigate the proportion of each gender in the entire population. Finally, a hypothesis was created about the average age of the customer. The hypothesis was to determine whether the average age of a customer was equal or not equal to 40. With all of the information from the hypothesis tests, Starbucks can more effectively manage their stores and customer relationships.

2. Method

The primary data collection process began by randomly selecting a local coffee shop to survey, which gave each local coffee shop an equal opportunity to be selected. The data was collected in February 2012 at the Starbucks located at 101 Linden Street, Wellesley, MA. In total, 106 Starbucks customers were observed. The variables recorded were weather, the customer's age and gender, items purchased, group size, sale

[+] Corresponding author. Tel.: + 9063709279;
E-mail address: femdad@babson.edu

amount, and stay. The weather refers to whether it was rainy or sunny at the time of data collection. The surveyors wanted to determine whether weather had an effect on the customer's purchases. Gender is referring to whether the customer involved in the transaction was male "0" or female "1". The items purchased refers to the drinks, such as regular coffee and specialty items, including holiday drinks, tea, decaf items, and mixed drinks, food, and paraphernalia. Group size refers to the size of the group in the transaction. For example, with a group size of three, the total sale amount should be higher, assuming they were purchasing for each person in the group. Finally, stay refers to whether or not the customer stayed in the store after purchasing their drink. The surveyors wanted to investigate the mean sale amount for customers who stayed versus left. All of these components that were collected can hopefully help Starbucks improve their understanding of the factors that contribute the most to each transaction amount and ultimately to revenue.

Once all of the data was collected, the researchers completed multiple hypothesis tests, Principal Component Analysis (PCA), and Logistic Regression. The four hypothesis tests involved the mean sale amount, the mean sale amount for customers who stayed versus left after purchasing, the proportion of each gender in the population, and the average age of surveyed customers. The researchers were interested to see how the different variables affected the total transaction sale amount per customer and what implications this could have on a broader level for Starbucks.

Before completing these tests, the data was surveyed and standardized. Any Z score values below -3 or above 3 indicate an outlier and the corresponding outlier were removed. Some of these "outliers" were a result of the customer purchasing paraphernalia, which made the sale price significantly higher. Therefore, it was decided that these observations should still be included because paraphernalia represents an important part of Starbucks business and their overall revenue.

Principal Component Analysis (PCA) is a powerful classic technique for statistical data analysis, feature extraction and data compression. Pearson was the first to introduce PCA technique [1]. Tufts and Kumaresan developed this technique in [2]. Then Vaccaro Tufts and Boudreaux-Bartels reviewed it in [3]. PCA works on the correlated data and transforms the original correlated data into a number of uncorrelated data called principal components [4]. PCA was completed to find a smaller number of principal components that would explain most of the variability in the original variables. By completing this analysis, there is a significant gain in efficiency in terms of describing the variance-covariance structure of the variables [5]. This test will produce a smaller number of factors that will assist in simplifying the data, with a minimum loss of information, so it is easier to work with. In this way, PCA was a useful analysis to apply before the logistic regression [6].

After PCA, the researchers used logistic regression in hopes of being able to classify new data [7, 8]. Logistic Regression is used for categorical variables, and for this test, the categorical variable stay was used as the dependent variable. Through this test, Starbucks will gain better insight into the buying patterns and behaviors of their customers. Ultimately, they will be able to classify whether new customers will stay or leave the store after purchasing based on the other independent variables.

3. Results

3.1. Hypothesis Tests

Beginning with the hypothesis tests, the researchers used Minitab to try and test the original hypotheses. The first hypothesis test conducted related to the mean sale amount being greater than $3.50 for the 103 observations. H0: $\mu \leq \$3.50$ vs. H1: $\mu > \$3.50$. Based on the sample, the mean sale amount for Starbucks customers is approximately $4.03. Therefore with a p-value of 0.015, which is less than the significance level of 0.05, there is sufficient evidence to reject the null hypothesis and it can be concluded that the mean sale amount for Starbucks customers is greater than $3.50.

The next hypothesis test conducted used the variables sale and stay to try and determine if the mean sale amount was higher for the customers who stayed or left after purchasing. H0: $\mu_{sale.didnotstay} \leq \mu_{sale.stay}$ vs. H1: $\mu_{sale.didnotstay} > \mu_{sale.stay}$. The original hypothesis was that customers who stayed after purchasing would spend more on average than those who left. The mean sale amount for customers who left,

designated with a 0, is $4.78, while the mean sale amount for customers who stayed, designated with a 1, is $3.42. From this information, it is apparent that the mean sale amount for customers that did not stay is higher than for those that stayed. To confirm this, the p-value is 0.002, which again is below the alpha level of 0.05. Therefore, there is sufficient evidence to reject the null hypothesis and conclude that the mean sale amount for those who did not stay is greater than those who stayed.

The next hypothesis test was run as a proportion test. The researchers wanted to discover the proportion of the overall population that was male to female. The original hypothesis was that the proportion of males to females in the population was not equal. H0: $\pi M = \pi F$ vs. H1: $\pi M \neq \pi F$. The percentage of males (designated as sample 1) in the population is about 34%, while the percentage of females (designated as sample 2) in the population is about 66%. When looking at the p-value of 0.000, which is less than the significance level of 0.05, there is sufficient evidence to reject the null hypothesis that there were equal proportions of each gender in the population. Therefore, it can be concluded that there is a higher proportion of females in the population than males, and this is logical based on the original proportion of males to females that were observed in the sample.

The last hypothesis test was conducted to understand the average age of a Starbucks customer. It was hypothesized that the average age would be approximately 40 years old. H0: $\mu = 40$ vs. H1: $\mu \neq 40$. Based on the output, it can be concluded that the mean age for Starbucks customers is approximately 36 years old. Therefore, with a p-value of 0.008, which is less than the significance level of 0.05, there is sufficient evidence to reject the null hypothesis and it can be concluded that the mean age for Starbucks customers is not equal to 40, but is approximately 36. Additionally, as can be seen in the t-distribution graph, the t-value of -2.72 falls under the shaded rejection region because it is smaller than the t critical value of -1.983.

3.2. Principal Component Analysis

PCA was completed to find a smaller number of principal components that would explain most of the variability in the original data. For Starbucks, PCA will be particularly important in identifying the underlying factors that affect the total sale amount for their customers. All of the variables were included in the following analysis.

When looking at the correlation matrix, it is important that there are some higher correlation values between variables in order to continue with the analysis. If there is no correlation between the original variables, then it would not be logical to run PCA. In this case, the correlations above .30 are used. The correlations are between the following variables: Sale and Paraphernalia, Regular Coffee and Specialty, Food and Specialty, and Food and Drink Size. The high correlation between Sales and Paraphernalia can be attributed to the fact that sales increase drastically with the purchase of paraphernalia. Additionally, for the variables Regular Coffee and Specialty, they are negatively correlated due to the fact that they are not normally purchased together—if one is purchased, the other one usually is not. Furthermore, the negative correlation between Food and Specialty drinks and Food and Drink Size can be attributed to the total amount a customer is willing to spend at Starbucks. If they are purchasing food, their total will increase significantly and they may have less money available to spend on a larger or more expensive specialty drink if they have a budgetary constraint.

For the next section with the KMO and Bartlett's test, there is some conflicting information. Typically, the acceptable value for KMO is .50 in order to continue with the analysis. In this case, the KMO value is .408, however the researchers feel that PCA will contribute to their analysis and will continue with the test. The significance level for Bartlett's Test of Sphericity is .000, which is below the significance level of 0.05. This says that there is some correlation between the grouping variables and so it is satisfactory to continue running PCA.

The Communalities show how much of the variance in each variable has been accounted for by the extracted factors. The extraction percentage on each variable is at least 50%, with most of the values around .70 and .80. Therefore, the analysis has accounted for a significant amount of the variance in each of the variables.

Table 1: Total Variance Explained

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	2.244	20.401	20.401	2.244	20.401	20.401	1.957	17.789	17.789
2	2.028	18.440	38.841	2.028	18.440	38.841	1.829	16.628	34.417
3	1.585	14.406	53.247	1.585	14.406	53.247	1.695	15.405	49.822
4	1.123	10.205	63.452	1.123	10.205	63.452	1.309	11.901	61.723
5	1.040	9.453	72.905	1.040	9.453	72.905	1.230	11.183	72.905
6	.890	8.091	80.997						
7	.850	7.732	88.728						
8	.512	4.658	93.387						
9	.496	4.508	97.895						
10	.132	1.202	99.097						
11	.099	.903	100.000						

Extraction Method: Principal Component Analysis.

As can be seen from Table 1, SPSS has extracted five principal components. The first five principal components account for 20.401%, 18.440%, 14.406%, 10.205%, and 9.453% of variation in the data respectively and they cumulatively account for 72.905% of the variability in the data.

In the non-rotated component matrix with values below .4 suppressed, it is unclear what the extracted factors are. Therefore, further steps must be taken in order to find component matrices that are clear. Using Varimax rotation, Table 2, the new rotated component matrix shows a much clearer picture of the extracted factors. Through rotation, the numbers of factors on which the variables under investigation have high loadings are reduced, therefore making interpretation easier. The components can be interpreted as follows:

- Component 1 – Types of drinks a customer orders
- Component 2 – The effect purchasing paraphernalia has on sale price
- Component 3 – What factors influence staying
- Component 4 – The effect gender has on group size
- Component 5 – How other characteristics, such as weather, age, and group size, influence the transaction

Table 2: Rotated Component Matrix

Rotated Component Matrix[a]

	Component				
	1	2	3	4	5
Weather					.760
Gender				-.822	
Age					.537
Group_Size				.627	-.488
Stay			-.504		
Sale		.935			
Regular_Coffee	.941				
Specialty	-.913				
Food			-.667		
Paraphernalia		.902			
Size			.844		

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 10 iterations.

3.3. Logistic Regression

As previously mentioned, logistic regression is used to classify data into different groups. For Starbucks, the test will be used to classify the transactions into whether or not the customer stays after purchasing. From this information, Starbucks can learn about customer behaviors in order to improve the service in their stores.

In this case, there are 92 degrees of freedom because there are 103 data points, which are then reduced by eleven variables. With the chi-squared distribution, the rejection region is any value above 115.4. Therefore, at an alpha of .05, this is not in the rejection region. Looking at the different R-Squared tests, Cox & Snell's test says that 25.6% of the variation of the dependent variable is explained by the logistic regression and Nagelkerke's test says that 34.3% of the variation in the dependent variable is explained by the predictors. The latter is more reliable and this number should be reported.

Table 3: Hosmer and Lemeshow Test

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	11.794	8	.161

For the Hosmer and Lemeshow Test, the following hypothesis test was used: Ho: Observed = Predicted vs. H1: Observed ≠ Predicted, Table 3. The significance of the data is .161, which is greater than .05, and therefore there is not sufficient evidence to reject the null hypothesis, meaning the model is a good fit.

There are many variables that have very high significance levels which include weather, gender, age, group size, specialty, food, paraphernalia, and size of drink. After re-running the analysis with some insignificant variables removed, the classification rate for occupants that did not stay increased to 67.4% while the rate for those who did stay increased to 77.2%. The overall classification rate has increased by 2.9% to 72.8% and two less observations have been misclassified. Removing any further variables will decrease the overall classification percentage.

After completing the logistic regression on the original data, additional data was collected on a short trip to Starbucks. Then, using the original data as a training data set, these new data points were classified. The classification rate of the new data based on the training data is slightly higher than the overall classification rate. All of the occupants that did stay were classified correctly. Out of eleven cases, there were only two misclassifications. Therefore, this logistic regression may be a good predictor for future cases, as it is consistent and has a good classification rate.

The coefficients for the variables are used to create a regression equation, and the probability that a person will stay at Starbucks can be calculated using the following equation.

$$Probability = \frac{e^{1.488-.637Gender-.612Sale+2.448Specialty+1.361Food-.012Age-.405Size+2.7RegularCoffee+4.266Paraphernalia}}{1+e^{1.488-.637Gender-.612Sale+2.448Specialty+1.361Food-.012Age-.405Size+2.7RegularCoffee+4.266Paraphernalia}}$$

Additionally, looking at the model, there are a few variables that stand out. Regular coffee, specialty drinks, and food have high multipliers which means customers are more likely to stay if they purchase any of those items. Gender, age, sale, and drink size are all less significant as they have multipliers of less than 1. Then, paraphernalia is 96.506, which means that when a customer purchases paraphernalia, the odd ratio will be 96 times as large.

Ultimately, by running this logistic regression, Starbucks can hopefully classify their customers buying behaviors in the future. This information can be particularly valuable as the mean sale amount for customers who did not stay tends to be higher than for those that do stay. Therefore, Starbucks can use this information in determining the seating area in their new stores as well as optimizing the configuration of their current stores.

4. Discussion & Conclusion

Overall, this study was designed to identify relationships between consumer purchasing behaviors and their implications on total revenue for Starbucks. Beginning with the primary sample from a local Starbucks, multiple hypothesis tests, principal component analysis, and logistic regression were run to further examine these relationships. Based on the findings, the researchers have discovered which factors contribute the most to each transaction amount and the relevance of a customer staying or leaving after purchasing. Although this information is extremely valuable, there are always ways to improve the tests in hopes of achieving a better output for analysis.

For the future, the researchers have some recommendations about the data collection process and different methods to run as well. Although the data was good to work with, it may be beneficial to use fewer categorical variables. For example, if it was possible to obtain the customer's income per year, it would be interesting to examine the effect of income on total amount spent. Additionally, it was questionable whether or not to include the PCA due to the KMO test and the scree plot. It is the hope that in the future with new variables in the data these will improve so the analysis will be even better and more comprehensive.

Ultimately, through this study, it was confirmed that customer characteristics and behaviors, specifically whether the customer stayed or left after purchasing, had the largest influence over the sale amount. Furthermore, with the logistic regression, Starbucks will now be able to classify their customers buying behaviors in the future. It is the hope that this study will give Starbucks better insight of their customers and ultimately be able to use this information to give them a competitive advantage.

5. References

[1] A. Adams and J. F. Fournier. Sobolev Spaces, volume 140 of Pure and Applied Mathematics. Academic Press, Boston, 2003.

[2] M. Aizerman, E. Braverman, and L. Rozonoer. Theoretical foundations of the potential function method in pattern recognition learning. Automation and Remote Control, 25:821-837, 1964.

[3] M. Akkhtar and J. Asenstorfer. Iterative detection for MC-CDMA system with base station antenna array for fading channels. Proceedings of IEEE Blobecom '98, pages 241-246, 1998.

[4] Berenson, Mark L., David M. Levine, and Timothy C. Krehbiel Basic Business Statistics: Concepts and Applications. Prentice Hall, 2012.

[5] Hosmer, D. & Lemeshow, S. Applied Logistic Regression. New York: John Wiley & Sons, Inc. 2000.

[6] Johnson, Richard A., and Dean W. Wichern. Applied Multivariate Statistical Analysis. Pearson Prentice Hall, 2007.

[7] Lay, David C. Linear Algebra and Its Applications. Addison-Wesley, 2012. Print. Long, J. Scott (1997).

[8] Regression Models for Categorical and Limited Dependent Variables. Thousand Oaks, CA: Sage Publications.