International Journal of Business and Economics
Vol. 7, No. 2, 2022, pp. 173-190

http://ijbe.ielas.org/index.php/ijbe/index
ISSN (online) 2545-4137

Measuring the Impacts of Consumers' Demographics on Spending Propensity and Expenses on Firms Profitability using Market Basket Optimization Model

Zahidul Karim

Associate Professor, Department of Management Studies
Jahangirnagar University
Email: zahidulkarim@dusty.tamiu.edu

Sumaya Fatema Binte Shahid

Graduated in HRM, Department of Management Studies, Jahangirnagar University
E-mail: sumayaniloy@gmail.com

Shahpar Shams

Assistant Professor, Department of Marketing
Jahangirnagar University, Bangladesh
E-mail: shahparshams@juniv.edu

Received: 2022-10-26 Accepted: 2022-11-04 Published online: 2022-11-19

Abstract

Market Basket Analysis (MBA) and its importance in selecting the right basket of goods have obtained considerable interest among the managers and executives of many retail stores. The present study has explored the literature gap in measuring the determinants for managing the optimum market basket of goods for consumers. The study found that some customer demographic and firm specific expense variables provide important prediction power to measure the customer spending propensity and firms' profitability respectively. We have used two different market basket optimization models to measure and analyze the results. The results show significant influence of age and advertising expenses on consumers' propensity to expense and firms' profitability respectively. The study has used two regression models to analyze the market basket using WarpPLS and R Software. Finally, the study suggests for future study to measure the impacts of consumer behavioral and psychological aspects and industry types on the spending propensity and firms' profitability through market basket analysis.

Keywords: Market Basket, Consumer Demographics, Spending Propensity, Firms' Profitability.

---

1. Introduction

Market basket analysis (MBA) has become one of the important competitive measures for retailers and marketing firms to optimize the sales and revenues along

International Journal of Business and Economics
Vol. 7, No. 2, 2022, pp. 173-190
http://ijbe.ielas.org/index.php/ijbe/index
ISSN (online) 2545-4137

with satisfying customer needs. Many firms fail to measure optimum market baskets for the consumers that made them unprofitable and maintain the sustainable competitive advantage over the rivals. Consumers are concerned about the time and cost when they buy basket of goods from retail stores. Pandemic of Covid19 and effect of lockdown influences consumers to use e-commerce platform to complete shopping for large basket of goods. We also found that up-sell feature also referred to as '*frequently bought together*' in many e-commerce websites (e.g., Amazon, eBay and Daraz) and predicts all possible items that consumers might buy together with the items that they have added to shopping cart. Customers are given more choice to add these items with this feature.

MBA has brought increasing attention to other field of studies (e.g., economics, management, and organizational behavior). Russell & Petersen (2000) mentioned that choice of market basket helps consumer to choose appropriate items from a set of available product groups in each shopping attempt. The important attributes of choosing market basket depends on related and interdependent demand of items for ultimate basket. Mamiya et al. (2017) mentioned that unhealthy diet is one of principle causes of increasing mortality and disability worldwide. Effective measurement and analysis of dietary patterns derived from monitoring and investigation of nutrition at societal level.

Griva et al. (2018) mentioned that market basket analytics is a useful technique for the retailer business in obtaining information about consumer buying behavior and preferences. In this study the author proposed a business analytics approach that captures customer visit information from the sales data of market basket. These attributed purchased categories for the basket of customers from their visit to retail stores and explore shopping intention of consumers behind the visit (e.g., a 'breakfast' visit to purchase egg, milk, bread, butter, etc).

Cavique, L. (2007) defined market basket as a bundle of products that customer purchase together in each visit of a retail store. MBA is a powerful method to utilize cross-selling strategies. This is important for retailers to specify large baskets of goods because it deals with many items. Mild & Reutterer (2003) noted that there is a growing interest among retail managers to learn about customer behavior for cross purchase. Aguinis et al. (2013) exemplifies that MBA provides important insights for management (e.g., HRM, organizational behavior, entrepreneurship, and strategic management). Mamiya et al. (2017) mentioned that now marketing firms and department can easily pool electronic data from a sample of retail stores through scanner technologies.

Boztuğ & Reutterer (2008) highlighted that that market baskets analysis has been derived from consumers shopping behavior that induce consumers to buy items from multiple categories and frequently purchase the desired items interdependently from various categories. Explanatory models to explain the dynamics of consumer behavior explicitly allow us to categorize purchase dependencies. Some models

International Journal of Business and Economics
Vol. 7, No. 2, 2022, pp. 173-190
http://ijbe.ielas.org/index.php/ijbe/index
ISSN (online) 2545-4137

facilitate estimating effects of marketing-mix variables on purchase incidences for individual and across-category based on the predefined set of product categories. Kim (2012) mentioned that "extended use of market basket analysis, network-leveled analysis is expected to be more effectively and efficiently used in personalized services, such as cross selling, up selling, and personalized product display utilizing the deep relation between products."

Aguinis et al. (2013) mentioned that the objective of MBA is to explore different bundle of products, items, and categories that consumer frequently purchase to fill their basket from their trip to retail stores. The study illustrated that MBA provides important insights apart from marketing perspective and extended to management domains, such as HRM for managing employee benefits, organizational behavior to explain dysfunctional behavior, entrepreneurship to measure entrepreneurial identities, and strategic management for corporate social responsibility.

2. Research Objectives

The objective of Market Basket Analysis (MBA) is to measure the relationships between bundle of products, items, or categories. There is substantial literature gap to optimize market basket based on the unit sales and price. Retail stores and their managers always struggle to optimize the market basket for greater profitability. It is always difficult to predict the customer behavior due to the heterogeneity in the customer choice, taste, and preferences. However, the sales pattern based on unit and price can provide understanding about the market basket optimization. This study will depict the pattern of market basket in measuring customer choice.

3. Literature Review

Russell & Petersen (2000) mentioned that Market basket is a process of decision-making by which consumer choose their desired items from a set of product groups and visit the retail store specifically to obtain them. The core philosophy of market basket analysis is to demonstrate the interdependence of consumer demand across different group of products selected for final basket. Mild & Reutterer (2003) mentioned that managers want to learn about cross-purchase behavior of customers in the long term. Author mentioned that collaborative filtering algorithms are often used to predicting customer preferences, choice, and ranking customers for online shopping.

Brin et al. (1997) discussed about difficulties of analyzing market-basket data and articulates its importance. They have presented algorithm to observe large number of items that uses limited passes over the data compared to traditional algorithms by using fewer item sets than different sampling techniques. They examine the methods of recording items to improve efficiency of the algorithm. They also present an innovative way to create "implication rules," that are normalized by antecedent and measure consequent of actual implications. This demonstrates how to produce sensible results than other methods.

Mamiya, Moodie & Buckeridge (2017) articulated that having unhealthy diet is one of the important reasons for early death and disability. A proper assessment and development of appropriate dietary patterns needs monitoring of nutrition intake at the societal level. However, observing nutrition intake basically depends on dietary surveys that often fail to assess food selection at desired level. Thus, marketing firms obtain and preserve transaction data with digital scanner from geographically dispersed retail stores that represents the sales pattern of retailers.

Kamakura (2012) mentioned that market basket analysis (MBA) is a forceful and widely used techniques in modem trading transactions that posed several shortcomings associated with issues of collecting data for purchase sequence from group purchase. Nowadays, retailers can generate sequence of purchase data from virtual records. Advanced technology made it easy to track the sequences instead of extrapolating them from group purchases. This study evaluates and compares conventional market basket analysis methods with a sequential method and offers a basis for sequence of purchase analysis. Aguinis et al. (2013) mentioned that MBA is a data-extracting process initiated by marketing and now using in different fields such as pharmaco-epidemiology, immunology, geophysics, and nuclear science. The author explains how MBA can bridge between lamented micro–macro and science–practice.

Cavique, L. (2007) mentioned that market basket refers to collection of goods that consumers buy together in each visit to retail store. The market basket analysis is a vigorous method to apply cross-selling marketing strategies. Especially for retailers it is important to attract customers and explore large baskets for them as they deal with thousands of product categories. Although application of algorithmic method can identify large set of products, they are not suitable in computational time. This method uses compressed data and is collect them by converting market basket problem into a high weighted group problem.

Hwang (2020) mentioned that the market basket analysis uses binary matrix to tackle collective filtering (CF) and target marketing. Effective selection of variable improves forecasting power and identifies significant users or items in collective filtering and target marketing. Thus, the study suggested two approaches variable selection: one is Pearson correlation and random forests approach by evaluating performance in different empirical settings.

Aiello et al. (2020) presented Tesco Grocery 1.0 dataset with 420 million food items purchased by 1.6 million loyalty card owners who buy groceries at 411 Tesco stores in Greater London throughout the year 2015. The study compares volume of food purchase to census population data and match nutrient and energy intake to

International Journal of Business and Economics
Vol. 7, No. 2, 2022, pp. 173-190
http://ijbe.ielas.org/index.php/ijbe/index
ISSN (online) 2545-4137

official data of food-related illnesses to measure validity and evaluate the extent to which the dataset is ecologically reliable. The data is used to link food purchases from several geographical indicators and suggests for future studies on health outcomes, cultural aspects, and economic factors.

To introduce and promote new product to customers, many firms send product catalogues to potential and existing customers. This is a way to fulfill their expectations instead of giving a large member of available products. The accumulated information can provide direction for advertisement and promotion. Rather than spending for adverting a particular product category, retailers would advertise product groups as a bundle that corresponds to customer shopping objectives (e.g., breakfast products advertisements). The proposed model for behavioral segmentation and categorizing customer visits are useful for business decisions. This provide indirect information but is has significant impact on customer satisfaction (Griva et al., 2018).

Multi category data driven model is a useful tool to design target marketing programs that integrate customer heterogeneity. Marketing analysts and retail marketing managers can be benefited directly from proposed methodology, such as: a data-driven strategy for selecting product categories included in models to predict cross-category effects. The data compression task warrants ensuring that the selected categories will effectively provide the structures of consumers' dynamic decision-making processes and information about segment-specific cross-category purchase and related marketing-mix effects become available. Retail marketing managers can use this information to optimize value in developing marketing strategies for target customers within their loyalty programs (Boztuğ & Reutterer, 2008).

Rahman (2020) mentioned that business to business firms spend more money on advertising to sales promotion and increase brand equity. A large portion of firm's spending occurred for advertising to attract end-user and pull them towards products categories that can be a complement to push promotional efforts. However, advertising efficiency vary across firms and implies that higher level of efficiency lead to higher profitability of firms.

Mild & Reutterer (2003) found that retail managers have found greater interest in managing and learning cross-category purchase behavior of customers. Most recently, exploring cross-category purchase patterns among retail market basket analysis has obtained more interest because of growing potential of promotion within the framework used in online environments. Walters & Jamil (2003) stated that different variety of shopping trips influence search and purchase behavior of consumers and has carried long interest among practitioners and researchers in marketing. The results show that consumer primarily visit retail store to purchase products that offer price specials with flyers and advertised price specials than consumers of other types of shopping trips.

International Journal of Business and Economics
Vol. 7, No. 2, 2022, pp. 173-190
http://ijbe.ielas.org/index.php/ijbe/index
ISSN (online) 2545-4137

4. Research Methods

This is an empirical study based on the secondary data analysis. The study employed several statistical tools to analyze and predict the customer buying behavior in the market basket optimization. Here, we have utilized two different research softwires (R Software, version 4.1.2 and WarpPLS, version 7.0) to conduct factor analysis, variance analysis, sample selection, correlation, and regression analysis. WarpPLS is useful analytical tool to measure the path coefficients (Kock, 2021). The study will also conduct hypothesis tests, descriptive statistics, correlation, and regression analysis. P value, F value and T-statistics will be measured to test the hypothesis. The study will measure the following two hypotheses:

H1: There is positive association between advertising expenses and firm's profitability
H2: There is positive association between promotional spending and firm's profitability
H3: There is positive association administrative expenses and firm's profitability
H4: Consumer age is positively associated with their spending score
H5: Consume gender is positively associated with their spending score
H6: Consumer annual income is positively associated with their spending score

Data Collection: The study is based on the secondary data collection. The study includes extensive literature from renowned journal articles. The US supermarket datasets from Kaggle.com are gathered and information from other websites was used to test the hypothesis. Four different data sets were used. There data consists of four categories: 50 supermarket branch, Ads CTR optimization, market basket optimization, and supermarket customer members. All empirical tests were conducted using R, SPSS and WarpPLS.

Description about the Market Basket Data:

The study used the dataset of XYZ supermarkets that have been operating since 2008 and business flourished until 2016. These supermarkets have large database, but they were less concerned to use them in achieving superior business performance. Their annual revenues sharply decreased by 10% and the declining trends shows to sustain in every year. The study has been analyzing these available datasets to estimate potential revenue growth of 50 supermarkets in the United States. There were four dataset that are summarized below:

1. 50 *Supermarket*Branches.csv: The data includes information on about different advertising, promotional expenses and profits of 50 supermarket branches in the USA. The data includes spending on advertisement, promotion and administration and profits.

2. Ads*CTR*Optimisation.csv: These data include the numbers of Click-Through Rates (CTR) from 10000 users of 10 different advertisements. This basically indicates the member customers of click on a specific ad. when they watch it first and second time.

3. Market*Basket*Optimisation.csv: This dataset includes 7500 sales data in a week.

4. Supermarket_CustomerMembers.csv: This dataset is used for customer segmentation.

These datasets on U.S. Supermarket are open and free for use. The data are available to use and legally permissible for everyone who want to use these data for any project analytics.

Structure of the Research

The study has developed six hypotheses based on the literature from previous studies. The study has reported descriptive statistics for all the four sets of data mentioned above. It has developed two models to forecast market basket. The first model explores whether the profitability of firms is influenced by advertising, promotional and administrative expenses. Thus, three hypotheses have been developed to test and validate the relationships. Different statistical measures such as: correlation, VIF, F statistics, t-statistics, regression, ANOVA was analyzed for testing the model. The second model explores the relationships of consumer demographics with the spending propensity. Hence, the influence of customer age, gender and annual income were measured on their spending propensity (measured by spending score). Here another three hypotheses were developed based on the prior literature and test and validate the relationships using the same statistical tools used for the first model. Finally, the study analyses the statistical results and presented the findings. The study concludes with presenting the results of hypothesis and proposing the future scope for analyzing market basket and profit optimization.

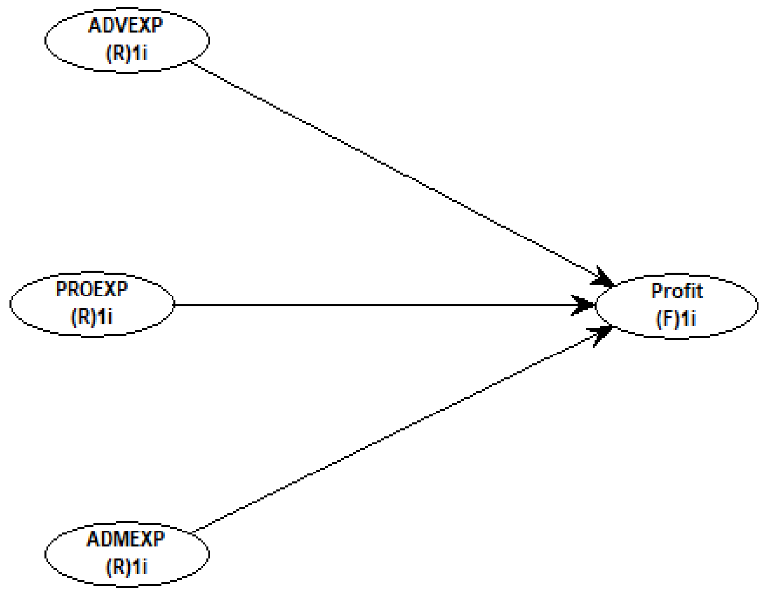
Proposed Models

Two models have been used to measure the market basket optimization. The first model measures the relationship of advertising expenses, promotional expenses, and administrative expenses with firm’s profit. The second model examines the relationships of consumer’s demographics (e.g., age, gender, and annual income) on their spending propensity (spending score): Both the models are presented with visual representation: using these two model the study has attempted to give direction to the retail store managers to decide how effectively manage their portfolio of expenses and determines the customers motives for spending on necessary goods. This help store managers and marketing executives to manage basket of goods and understand their relationships variables on daily turnover of stocks of products.

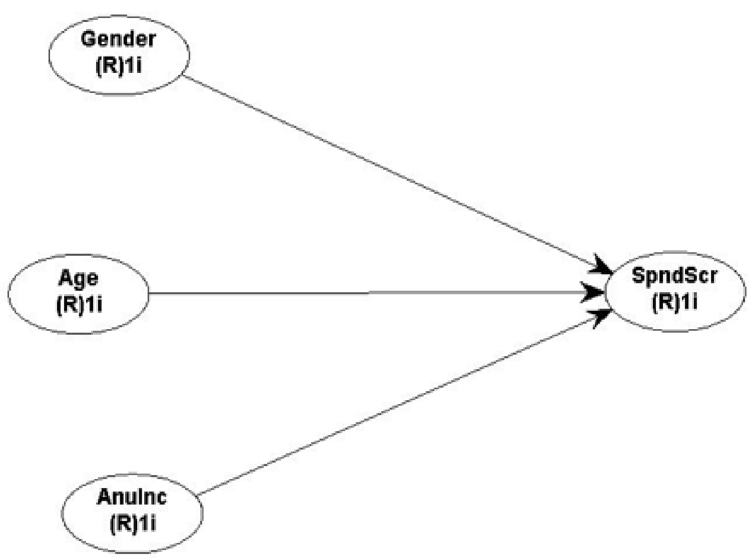
http://ijbe.ielas.org/index.php/ijbe/index
ISSN (online) 2545-4137

The study developed hypotheses based on the literature and proposed Model 1 and Model 2 to optimize the market basket:

Model 1: The relationships of advertising, promotional and administrative expenses on profit



Model 2: The relationships of gender, age, and annual income on consumer' spending score



International Journal of Business and Economics
Vol. 7, No. 2, 2022, pp. 173-190
http://ijbe.ielas.org/index.php/ijbe/index
ISSN (online) 2545-4137

Results and Discussion

The table 1 shows the descriptive statistics of 50 supermarket branches. All these supermarkets spend money on three categories to increase profit. The expenses incurred for advertising, promotion, and administrative purposes. For the descriptive table1, it is observed from the mean of all the three categories that 5o supermarket branches collectively spend more on administrative purposes with a mean value of 211025.09 and 121344.63 and 73721.61 for promotional and advertising purposes respectively. This implies that firms are more concerned on efficient management of administrative activities to increase profitability. The mean value of profit for 50 supermarket branches is 112012.63.

Table 1: Descriptive statistics for 50 Supermarket Branches

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation	SE
Advertising Expenses	50	0.00	165349.20	73721.61	45902.25	6491.55
Promotional Expenses	50	51283.14	182645.56	121344.63	28017.80	3962.31
Administrative Expenses	50	0.00	471784.10	211025.09	122290.31	17294.4
Profit	50	14681.40	192261.83	112012.63	40306.18	5700.15

4.1.2 Descriptive Statistics of 50 supermarket branches, R software, version:

The following table explains the frequency of click by customer place on particular ad. There are different set of advertisement for different product. The advertisement categories are presented from ad. 1 to ad. 10 and N represents two types of frequency that customer click on the advertisement. The first row (0) shows the number of click that customer place when they see the advertisement first time for each category and the second rows (1) represents the number of clicks that customer place on the same set of advertisement when they see the advertisement second time. The table shows that the number of clicks on advertisement for first time is more than the customers see the posted ad in the second time.

Table 2: Ads CTR Optimization

Ad. 1	**N**	**Ad.2**	**N**	**Ad.3**	**N**	**Ad.4**	**N**	**Ad.5**	**N**
0	8297	0	8705	0	9272	0	8804	0	7305
1	1703	1	1295	1	728	1	1196	1	2695
Ad.6	**N**	**Ad.7**	**N**	**Ad.8**	**N**	**Ad.9**	**N**	**Ad.10**	**N**
0	9874	0	8888	0	7909	0	9048	0	9511
1	126	1	1112	1	2091	1	952	1	489

International Journal of Business and Economics
Vol. 7, No. 2, 2022, pp. 173-190
http://ijbe.ielas.org/index.php/ijbe/index
ISSN (online) 2545-4137

Descriptive Statistics of Ads CTR Optimization using R software, version: 4.1.2

The following table shows that the company has retrieved the turnover information of the items sold in different cluster. There is total 20 categories within which various items are sold together for market basket. The categories are represented by V1 to V20. Each categories presents that when customer buy V1, he/she buy the bundle of goods together. Similarly, when consumer buy V2, V3 and so on, he/she also buy the bundle of goods within the specific category.

Table 3: Market Basket Optimization for Group of Products

```
                V1                  V2                  V3                   V4                  V5                V6
 mineral water     : 577                 :1754               :3112               :4156                :4972              :5637
 burgers           : 576   mineral water: 484  mineral water: 375  mineral water: 201  green tea    : 153  french fries: 107
 turkey            : 458   spaghetti    : 411  spaghetti    : 279  eggs         : 181  eggs         : 134  eggs        : 102
 chocolate         : 391   eggs         : 302  eggs         : 225  french fries : 174  french fries: 130  green tea   : 100
 frozen vegetables: 373   ground beef  : 291  milk         : 213  spaghetti    : 167  chocolate    : 115  chocolate   :  71
 spaghetti         : 354   french fries : 243  french fries : 180  milk         : 149  milk         : 114  pancakes    :  69
 (Other)           :4772   (Other)      :4016  (Other)      :3117  (Other)      :2473  (Other)      :1883  (Other)     :1415
             V7                   V8                   V9                   V10                  V11                  V12
            :6132                :6520                :6847                :7106                :7245                :7347
 green tea     :  96   green tea     :  67   green tea      :  57   green tea     :  31   low fat yogurt:  22   green tea      :  15
 french fries  :  81   pancakes      :  44   low fat yogurt :  38   french fries  :  19   green tea     :  20   french fries   :  10
 pancakes      :  69   low fat yogurt:  43   frozen smoothie:  35   low fat yogurt:  17   fresh bread   :  14   frozen smoothie:  10
 eggs          :  59   french fries  :  40   french fries   :  34   tomato juice  :  17   french fries  :  12   low fat yogurt :   9
 low fat yogurt:  55   chocolate     :  38   fresh bread    :  28   pancakes      :  14   light mayo    :   9   fresh bread    :   7
 (Other)       :1009   (Other)       : 749   (Other)        : 462   (Other)       : 297   (Other)       : 179   (Other)        : 103
            V13                  V14                  V15                  V16                  V17
            :7414                :7454                :7476                :7493                :7497
 green tea     :   8   green tea      :   4   magazines     :   3   antioxydant juice:   1   antioxydant juice:   1
 fresh bread   :   6   french fries   :   3   fresh bread   :   2   cake             :   1   french fries     :   1
 low fat yogurt:   6   frozen smoothie:   3   green tea     :   2   chocolate        :   1   frozen smoothie  :   2
 escalope      :   4   cottage cheese :   2   low fat yogurt:   2   frozen smoothie  :   1
 french fries  :   4   eggplant       :   2   pancakes      :   2   magazines        :   1
 (Other)       :  59   (Other)        :  33   (Other)       :  14   (Other)          :   3
            V18                V19              V20
           :7497              :7498            :7500
 frozen smoothie:   1   cereals   :   1   olive oil:   1
 protein bar    :   2   mayonnaise:   1
 spinach        :   1   spinach   :   1
```

Table generated using Market Basket Optimization, R software, version: 4.1.2

The following table shows the descriptive statistics for supermarket members where those are categorized by gender, age annual income and spending score. The mean age of customer is 38.85, annual income is 60.56 and spending score is 50.2. The standard error for annual income and spending score are very high that is 1.8571 and 1.8259 respectively. Dummy variable used for gender classification (1 denoted female and 2 denoted male). The standard error for age is 0.9833.

International Journal of Business and Economics
Vol. 7, No. 2, 2022, pp. 173-190
http://ijbe.ielas.org/index.php/ijbe/index
ISSN (online) 2545-4137

Table 4: Supermarket Customer Members

	n	mean	sd	median	skew	kurtosis	se
Customer ID	200	100.5	57.87918	100.5	0	-1.21801	4.092676
Gender*	200	1.44	0.497633	1	0.239936	-1.95212	0.035188
Age	200	38.85	13.96901	36	0.47831	-0.70785	0.987758
Annual Income ('000)	200	60.56	26.26472	61.5	0.317031	-0.15456	1.857196
Spending Score	200	50.2	25.82352	50	-0.04651	-0.85754	1.825999

Table generated using R software, version: 4.1.2

The simple multiple regression outcomes for 50 supermarket branches are presented in the following table. The model is precise and statistically significant. The model shows that the R square vale is 0.951 and adjusted R square value is 0.948 this means the predictor variables measures 95.1% variability of criteria (dependent variable) variable that is profit of the firm. The model is statistically significant at .001% level. The ANOVA test results show the F value is 295.978 that is very good and statistically significant. The following regression model describe that the variables of gender, age and annual income describe the spending ability (measured in terms of score) by 95.1%. This means that the variability of spending power can be explained 95.1% by the three explanatory variables. Hence, the model is precise and statistically significant.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.975[a]	0.951	0.948	9232.33484	0.951	295.978	3	46	0.000

The following table represents the ANOVA where the F value is 295.978 and p value is <.001, this implies that the model is has power to describe the variables and is statistically significant. The higher F value and lower p value represents more explanatory power of the model.

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	75683964196.193	3	25227988065.398	295.978	.000[b]
	Residual	3920856300.957	46	85236006.543		
	Total	79604820497.149	49			

a. Dependent Variable: Profit

b. Predictors: (Constant), Administrative Expenses, Promotional Expenses, Advertising Expenses

The following regression coefficient shows positive relationships between advertising expenses and profitability of firm and this is statistically significant at .001%

International Journal of Business and Economics
Vol. 7, No. 2, 2022, pp. 173-190
http://ijbe.ielas.org/index.php/ijbe/index
ISSN (online) 2545-4137

significance level. The $1 increase of advertising expenses will lead to an increase in profit by $0.918. On the other hand, there is negative relationship between promotional expenses and profit, an increase in promotional expenses by $1 will lead to decrease in profit by $0.019% and is statistically insignificant. Thought the last variable advertising expenses has positive correlation with profit, but it is statistically insignificant. Hence, the regression coefficient estimates that advertising expenses has only the significant impacts on firm's profitability. The table also shows that there is no multicollinearity among the variables since VIF is less than 3 (Standard VIF is 5).

Regression Coefficient

Coefficients[a]

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B		Collinearity Statistics
	B	Std. Error	Beta			Lower Bound	Upper Bound	VIF
(Constant)	50122.193	6572.35		7.626	0.000	36892.73	63351.65	
Advertising Expense	0.806	0.045	0.918	17.846	0.000	0.715	0.897	2.469
Promotional Expense	-0.027	0.051	-0.019	-0.526	0.602	-0.130	0.076	1.175
Administrative Expense	0.027	0.016	0.083	1.655	0.105	-0.006	0.060	2.327

a. Dependent Variable: Profit

The regression table using R also shows that advertising expenses have only statistically significant impacts the profitability of the firm while promotion and administrative expenses have insignificant impacts on firms' profitability. There is no multicollinearity among the variables. In the above table, the VIF score for all the three variables is less than 5.0 (standard value). The VIF for advertising expense is 2.469, promotional expense is 1.175 and administrative expense is 2.327. In addition, the following table shows the Eigenvalue is less than 1 for all the three variables, namely advertising expenses, promotional expenses, and administrative expenses.

Dimension	Eigenvalue	Condition Index	Variance Proportions			
			(Constant)	Advertisement Spend	Promotion Spend	Administrative Spend
1	3.661	1.000	.00	.01	.00	.01
2	.242	3.891	.04	.12	.05	.11
3	.077	6.875	.02	.67	.02	.62
4	.020	13.536	.94	.21	.93	.26

International Journal of Business and Economics
Vol. 7, No. 2, 2022, pp. 173-190
http://ijbe.ielas.org/index.php/ijbe/index
ISSN (online) 2545-4137

Correlation Matrix

The above correlation matrix shows that there is high positive correlation between profit and advertising expenses and is statistically significant at 0.01 level. The correlation between profit and promotional expenses is positive but low. In addition, the table shows that relation between profits and administrative expenses is highly positive and statistically significant at 0.01 level. This implies that increase of advertising expenses and administrative expenses positively associated with profit. If firms increase these expenses, their profits also increase. There is positive and high correlation between advertising expenses and administrative expenses while the relationship is negative and low correlations between promotional expenses and administrative expenses. The relationship of promotional expenses with profit and advertising expenses are positive but low. The following table shows the correlations among profit, advertising expenses, promotional expenses, and administrative expenses.

Correlations

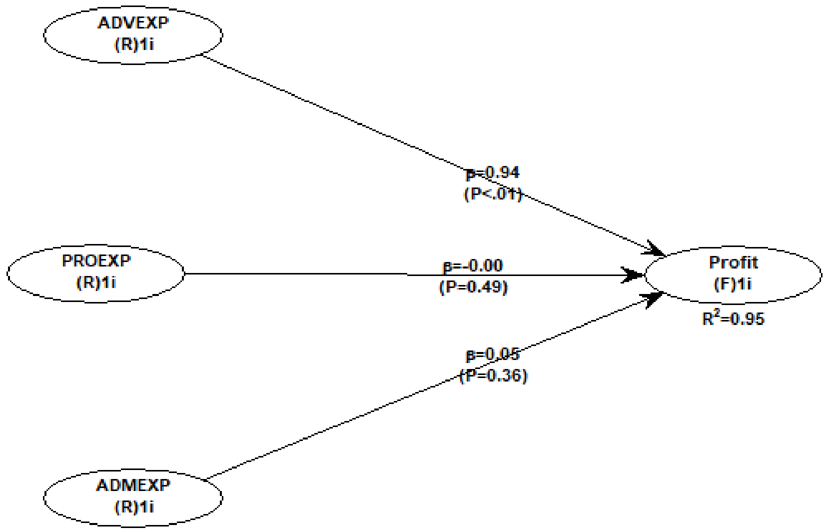
		Advertising Expenses	Promotional Expenses	Administrative Expenses	Profit
Advertising Expenses	Pearson Correlation	1			
Promotional Expenses	Pearson Correlation	0.242	1		
Administrative Expenses	Pearson Correlation	.724**	-0.032	1	
Profit	Pearson Correlation	.973**	0.201	.748**	1

**. Correlation is significant at the 0.01 level (2-tailed). N=50

Results of Model 1

The WarpPLS 7.0 has been used to measure the path coefficient, p value and R square. The following model shows the path coefficient and p value between predictor and criteria variables. The beta coefficient of advertising expenses is 0.94 ($P<.01$) and is statistically significant since P value is <.05. The beta coefficient for promotional expenses is 0.00 that is statistically insignificant. The beta coefficient for administrative expenses is 0.05 and statistically insignificant. Hence, the advertising expenses has strong predictive power of firm's profitability that implies increase in advertising expenses will lead to higher profit for the firm. However, the predictive power of the model is very high since the R-square value is 0.95 that means all the three variables combinedly predict 95% variability of profit.

International Journal of Business and Economics
Vol. 7, No. 2, 2022, pp. 173-190
http://ijbe.ielas.org/index.php/ijbe/index
ISSN (online) 2545-4137



The simple multiples regression results are presented in the following table for supermarket customer members. This model shows that the independent variables age, gender and annual income together can explain 10.9% variability of profit. Though the model has less explanatory power to predict the variability of profit, there must be other independent variables that need to be included to increase the explanatory power.

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics		
					R Square Change	F Change	Sig. F Change
1	.330[a]	0.109	0.095	24.567	0.109	7.960	0.000

a. Predictors: (Constant), Annual Income ($), Age, Gender

b. Dependent Variable: SPNDSCOR

The ANOVA table shows that the F statistics is 7.960 that is not very high, but the regression is statistically significant at .001% significance level.

ANOVA[a]

Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	14412.481	3	4804.160	7.960	.000[b]
Residual	118291.519	196	603.528		
Total	132704.000	199			

a. Dependent Variable: SPNDSCOR

b. Predictors: (Constant), Annual Income k$, Age, Genre

The regression coefficient in the following tables shows that gender has insignificant relations with spending score. The gender used dummy variable (1 for female and 2 for male). There is also negative relationship between age and spending score. If the age increases the spending score decrease and this is highly significant (P<0.001) in explaining the changes of spending score. However, the annual income variable has positive relationships with spending score that is desirable, but the coefficient is low (.008) and statistically insignificant.

Coefficients[a]

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B	
	B	Std. Error	Beta			Lower Bound	Upper Bound
(Constant)	75.943	7.974		9.524	0.000	60.217	91.669
Gender	-2.013	3.512	-0.039	-0.573	0.567	-8.939	4.913
Age	-0.600	0.125	-0.325	-4.806	0.000	-0.847	-0.354
Annual Income ($)	0.008	0.066	0.008	0.119	0.905	-0.123	0.139

a. Dependent Variable: SPNDSCOR

The regression table using R also shows that the age of consumers has only statistically significant impacts on the spending propensity while the gender and income has insignificant impacts on the spending propensity.

Analysis of variance for Gender

The variance analysis of the gender shows that the male and female have no significant difference on explaining the variability of spending. The p value is greater than .05 that implies that the difference of gender is not statistically significant on the spending differential.

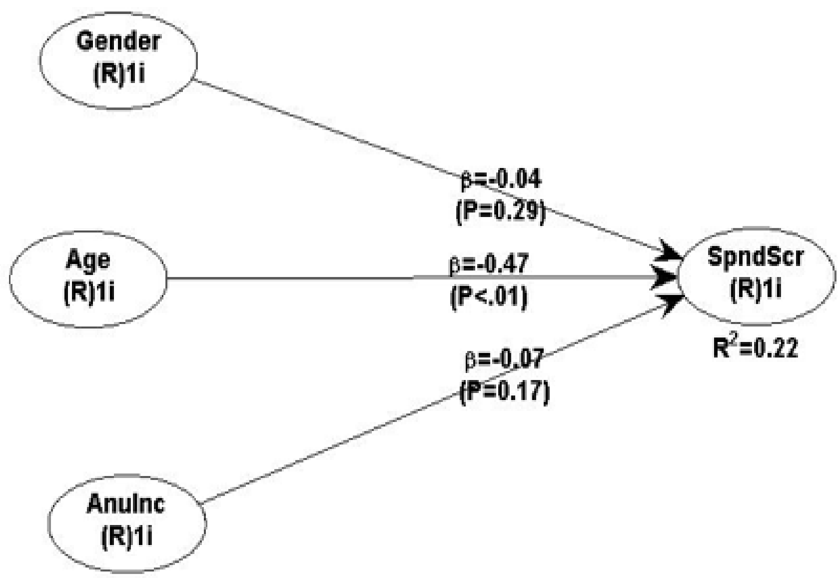
	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Gender	1	448.0917	448.0917	0.670837	0.413744659
Residuals	198	132255.9	667.9591	NA	NA

Dummy	***Gender***	***Average Spending Score***
0	***Male***	***51.5***
1	***Female***	***48.5***

The group means for male and female from the ANOVA are 51.5 and 48.5 respectively. This indicates that there is no significant difference between the average spending propensity (measured by spending score) between male and female. Thus, gender does not significantly affect the spending power of male and female.

Results of Model 2

When WarpPLS 7.0 has been used to measure the path coefficient, p value and R square, the model shows beta coefficient of gender is -0.04 and is statistically insignificant since P value is >.05. The beta coefficient for age is -0.47 that is statistically significant at 1% (p<0.01) significance level. The beta coefficient for annual income is -0.07 and statistically insignificant since p value is >.05. Hence, the age has strong predictive power on consumers' spending propensity. This implies that if the age of consumer increase by 1 year, their spending score will decrease by .47 percent. However, the predictive power of the model is 0.22 that means all the three variables combinedly predict 22% variability of consumer spending propensity.



Results of Hypotheses Test: Form the above table; we see the outcomes of our hypotheses test. The H1 is supported because the null hypothesis is rejected at .001% significance level and P value is <.05. Thus, there is positive and significant association between advertising expenses and firm's profit. In contrast, H2 and H3 are not supported since null hypothesis cannot be rejected. The p value is >.05. Thus, there are insignificant relations of promotional expenses and administrative expenses with profit. We see, H4 is supported that means is age of the consumers positively affects their spending score. H5 and H6 are not supported that means the gender difference and annual income level of consumers does not affect the spending score although is not congruence with previous literature.

Hypothesis	Description	P-Value	Results/Outcomes
H1	There is positive association between advertising expenses and firm's profitability	0.000<0.5	***Supported***
H2	There is positive association between promotional spending and firm's profitability	0.602>.05	***Not supported***
H3	There is positive association administrative expenses and firm's profitability	0.105>0.5	***Not supported***
H4	Consumer age is positively associate with their spending score	0.000<.05	***Supported***
H5	Consume gender is positively associated with their spending score	0.567>.05	***Not supported***
H6	Consumer annual income is positively associated with their spending score	0.905>.05	***Not supported***

5. Conclusion and Future Research

This study measures the impacts of firms' different expenses on their profitability and the demographic variables of consumers to measure their spending propensity by using two market basket optimization models. In the first model, this study found strong and positive association of advertising expenses with profitability. In the second model, the study found positive and strong relationships consumer age and their spending propensity. This information will assist retailers to manage consumers' market basket and firms' profitability. However, the study could not consider other factors (e.g., consumer income, lifestyle, and psychological factors) to measure the market basket optimization due to the lack of data availability. The study could not consider other firm level variables (e.g., research and development expenses, new product, and process development) on firms' profitability. The study confined analysis of data that are only available for some product categories, consumer demographics, and firm level variables such as advertising, promotion and administrative expenses, consumers spending score and firm's profit to predict the market basket optimization. The study suggests for measuring the impacts of consumer behaviors and types of industry on spending propensity and firm's profitability respectively.

References

Aiello, L. M., Quercia, D., Schifanella, R., & Del Prete, L. (2020). Tesco Grocery 1.0, a large-scale dataset of grocery purchases in London. *Scientific data*, *7*(1), 1-11.

Aguinis, H., Forcum, L. E., & Joo, H. (2013). Using market basket analysis in management research. *Journal of Management*, *39*(7), 1799-1824.

Boztuğ, Y., & Reutterer, T. (2008). A combined approach for segment-specific market basket analysis. *European Journal of Operational Research*, *187*(1), 294-312.

Brin, S., Motwani, R., Ullman, J. D., & Tsur, S. (1997, June). Dynamic itemset counting and implication rules for market basket data. In *Proceedings of the 1997 ACM SIGMOD international conference on Management of data* (pp. 255-264).

Cavique, L. (2007). A scalable algorithm for the market basket analysis. *Journal of Retailing and Consumer Services*, *14*(6), 400-407.

Griva, A., Bardaki, C., Pramatari, K., & Papakiriakopoulos, D. (2018). Retail business analytics: Customer visit segmentation using market basket data. *Expert Systems with Applications*, *100*, 1-16.

Hwang, W. Y. (2020). Variable selection for collaborative filtering with market basket data. *International Transactions in Operational Research*, *27*(6), 3167-3177.

Kamakura, W. A. (2012). Sequential market basket analysis. *Marketing Letters*, *23*(3), 505-516.

Kaggle (2022). U.S. Supermarket Data. A United States supermarket dataset for marketing analysis purposes. Retrieve from https://www.kaggle.com/datasets/sindraanthony9985/marketing-data-for-a-supermarket-in-united-states

Kim, H. K., Kim, J. K., & Chen, Q. Y. (2012). A product network analysis for extending the market basket analysis. *Expert Systems with Applications*, *39*(8), 7403-7410.

Kock, N. (2021). WarpPLS User Manual: Version 7.0. *Laredo, TX: ScriptWarp Systems*.

Mamiya, H., Moodie, E. E., & Buckeridge, D. L. (2017). A novel application of point-of-sales grocery transaction data to enhance community nutrition monitoring. In *AMIA Annual Symposium Proceedings* (Vol. 2017, p. 1253). American Medical Informatics Association.

Mild, A., & Reutterer, T. (2003). An improved collaborative filtering approach for predicting cross-category purchases based on binary market basket data. *Journal of Retailing and consumer Services*, *10*(3), 123-133.

Rahman, M., Rodríguez-Serrano, M. Á., & Lambkin, M. (2020). Advertising efficiency and profitability: evidence from the pharmaceutical industry. *Industrial Marketing Management*, *89*, 619-629.

Russell, G. J., & Petersen, A. (2000). Analysis of cross category dependence in market basket selection. *Journal of Retailing*, *76*(3), 367-392.

Walters, R. G., & Jamil, M. (2003). Exploring the relationships between shopping trip type, purchases of products on promotion, and shopping basket profit. *Journal of Business Research*, *56*(1), 17-29.