ADVANCED QUANTIATIVE RESEARCH REPORT

Linh K. T. Vo (1661618)

HAN University of Applied Sciences

Lecturer: Tijmen Weber

Date: 20/01/2023

Abstract

To improve the business performance of an online shop for next periods, customer behavior of making online purchases on the website is taken into account. They were asked to fill in a short survey at the end of their purchases. Some statistical analyses are carried out to come to final conclusions about the behavior of the customers. The last part of the report is to form judgements of an article by Al-Zyoud (2018) on three domains: sample, survey and analyses.

Contents

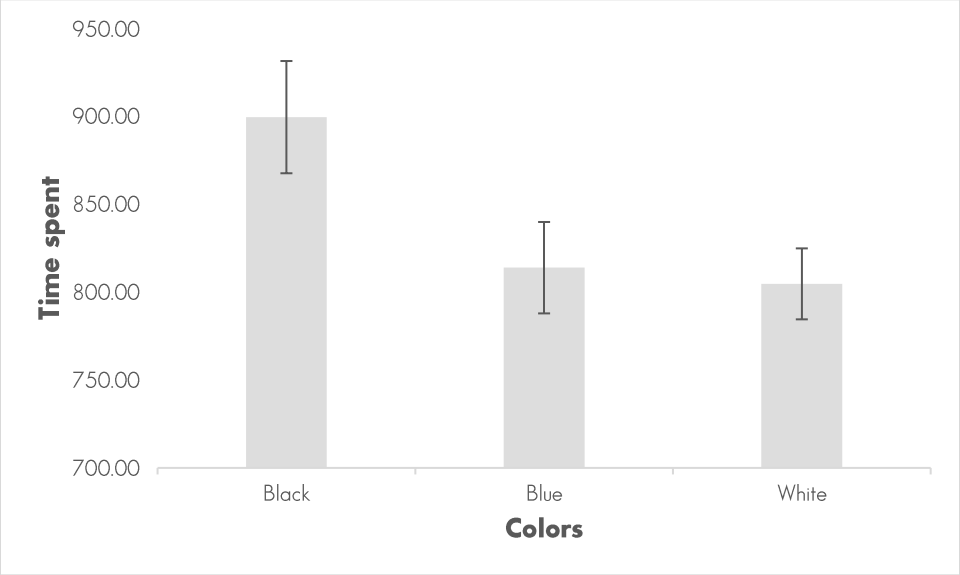
Average time spent on website between colors

Bar chart

As can be seen from Figure 1, black color has the highest average time spent (899.65) compared to other colors. The difference between blue and white is not significant (7.09). The error bars are calculated at 95 percent confidence intervals. The margin of error of three colors is 31.94, 26.00 and 20.20 respectively. Black significantly influences the time customers spent on website.

Figure 1

Average time spent on website between colors



ANOVA and TUKEY test

As can be inferred from Table 1, there is statistically significant difference between the three colors $F(2, 1077) = 15.414$, $p = 2.5E\text{-}0.7$ because p-value is lower than 0.05. The F value here is quite high, leading to lower the p-value. Tukey test is conducted for post hoc analyses. Based on Table 2, only two group pairs (Black-Blue and Black-White) are significantly different from each other.

Table 1

ANOVA test between colors

	df	F	P value
Between Groups	2	15.414	2.5E-07
Within Groups	1077		
Total	1079		

Table 2

TUKEY test between colors

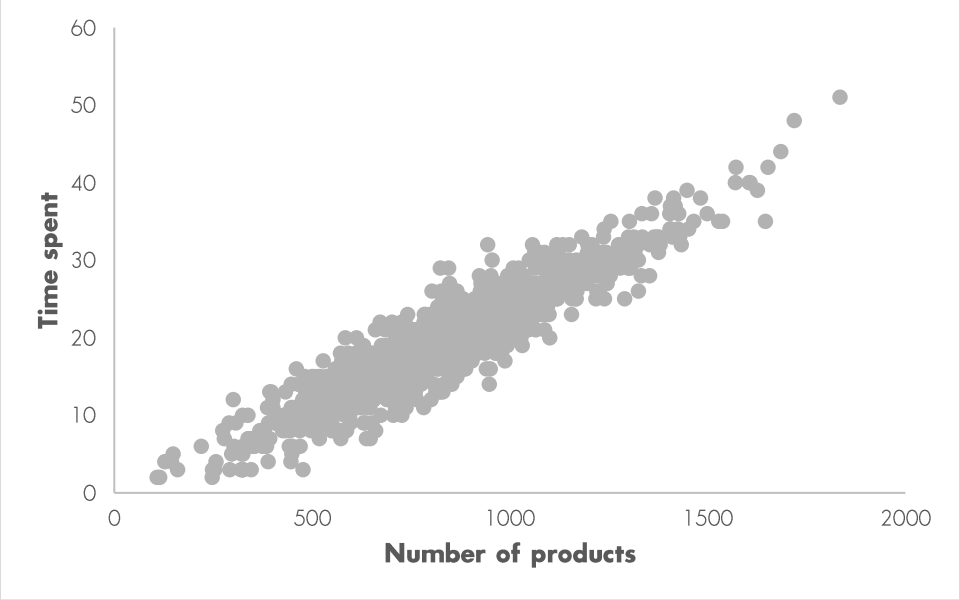
Group 1	Group 2	p-value
Black	Blue	1.2E-0.5
Black	White	2.3E-0.6
Blue	White	0.92607

Average time spent between number of products and average number of pictures

Scatterplot

Figure 2

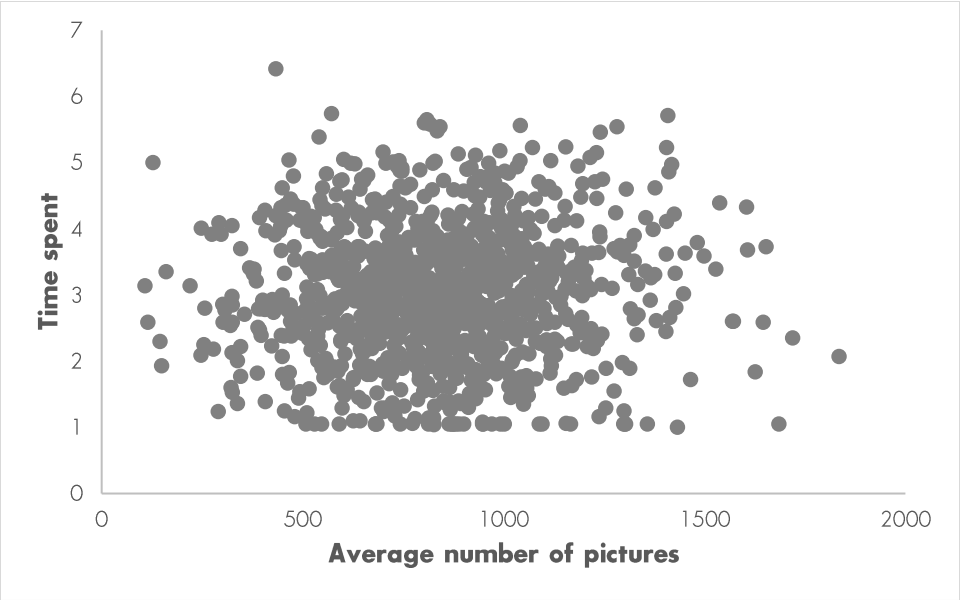
Average time spent on website versus number of products



As can be seen from Figure 2, the data points cluster tightly. Therefore, there is strong positive relationship between time spent and number of products. As the number of products increases, the time spent on the website correspondingly tends to increase.

Figure 3

Average time spent on website versus average number of pictures



As can be observed from Figure 3, the data points disperse more than they do in Figure 2. There is no relationship between the average number of pictures and time spent on the website.

Correlation matrix

Table 3

Correlation matrix between three variables

	Sample size	p-value	Correlation coefficient
Time spent vs Number of products	1080	0.000	0.922
Time spent vs average number of pictures	1080	0.331	0.030
Number of products vs average number of pictures	1080	0.403	0.025

Based on the data from Table 3, there is only one positive correlation between the two variables: $r(1078) = 0.922$, $p = 0$. The more the number of products are browsed, the more time

customers spend on the website. There is no statistically significant correlation between two variables: $r(1078) = 0.030$, $p = 0.331$ and $r(1078) = 0.025$, $p = 0.403$.

Purchase amount spent on the website

Regression model

From the Table 4, most of the variables are significant except average number of pictures and ease of purchase with rating 2. Some of the effects have negative influence and the rest has positive influence on the purchase amount spent on the website. Customers spend more money on the website if they browse more products on PC and the products purchased have higher ratings. Indeed, review rating has the largest impact on the purchase spent. The shipping time has negative influence on the amount of purchase spent on the website. Besides, the average number of pictures has no significant effect on the purchase amount spent. From the categorical variables, search engine and other have opposite influence compared to friends and family and social media advertisement. There are large differences between the finding website ways with social media advertisement. This tool must be the most efficient one to increase customers' reach. Beside that, ease of purchase with rating 4 and 5 are significant and positively affect the purchase amount spent. The others are not statistically significant. The R-squared is also high.

Table 4

The effect of other variables on purchase amount spent on the website

	Model	
	Coefficient	Standard error
Intercept	142.86***	18.92
Number of products browsed	1.72***	0.11
Average number of pictures	0.58	0.78
Shipping time	-11.43***	0.80
Review rating	10.99***	1.56
Age	-0.94***	0.07
PC	74.63***	1.75
Social media advertisement (ref)		

	Model	
	Coefficient	Standard error
Search engine	-65.25***	1.77
Other	-90.07***	10.90
Friends or Family	64.14***	2.97
Ease of purchase 3 (ref)		
Ease of purchase 2	11.86	17.26
Ease of purchase 4	42.27***	2.15
Ease of purchase 5	82.48***	2.79
Model Parameters		
R-Squared	0.8726	
Adjusted R-Squared	0.8710	
N	969	

*Note. *** p<0.001 ** p<0.01 *p<0.05. All continuous variables are standardized.*

Time spent on the website versus Number of products browsed

It is unwise to combine two variables: number of products browsed and time spent on the website in the same model. As can be seen from Table 5, the VIF of two variables are higher than 5, which means that there is multicollinearity. In addition to that, from Table 3, there is high correlation coefficient between these two variables. These two variables correlate very highly with each other. Therefore, the model doesn't work.

Table 5

The effect of other variables including time spent on the purchase amount

	Model		
	Coefficient	Standard error	Vif
Intercept	131.61***	8.12	
Number of products browsed	1.86***	0.28	6.52
Time spent on website	-0.00	0.00	6.55
Average number of pictures	0.58	0.78	1.01
Shipping time	-11.43***	0.80	1.01
Review rating	10.99***	1.56	1.01
Age	-0.94***	0.07	1.00
PC	74.68***	1.76	1.01
Social media advertisement (ref)			
Search engine	-65.23***	1.77	1.23
Other	-89.93***	10.90	1.02
Friends or Family	64.08***	2.98	1.23
Ease of purchase 3 (ref)			

	Model		
	Coefficient	Standard error	Vif
Ease of purchase 2	11.61	17.27	1.02
Ease of purchase 4	42.28***	2.15	1.66
Ease of purchase 5	82.51***	2.79	1.66
Model Parameters			
R-Squared	0.8726		
Adjusted R-Squared	0.8709		
N	969		

*Note. *** p<0.001 ** p<0.01 *p<0.05. All continuous variables are standardized.*

Standardized variables

Based on the data from Table 6, the review rating has a positively stronger effect on the purchase amount compared to pictures and shipping time. A one-standard-deviation increase in review rating leads to an average increase in purchase mount of 4.36. However, the shipping time has the largest influence on the amount of purchase spent. With a one-standard-deviation increase in shipping time, it can make average of purchase amount decrease by 11.

Table 6

The effect of standardized variables on the purchase amount

	Model	
	Coefficient	Standard error
Intercept	208.01***	2.12
Pictures	1.31***	2.13
Shipping time	-11.54	2.13
Review rating	4.36	2.12
Model Parameters		
R-Squared	0.0347	
Adjusted R-Squared	0.0317	
N	969	

*Note. *** p<0.001 ** p<0.01 *p<0.05. All continuous variables are standardized.*

Predictions about the amount customers purchase

This model can be used to accurate predictions about how much a customer will spend because it has good R-squared (R-squared = 0.8726). However, according to Parsimony, it is best to pick the simplest model that fits the purpose. To make a better prediction, this model can have

higher adjusted R-squared by leaving out two variables that are not relevant: pictures and ease of purchase rating 2. As can be seen from Table 7, the adjusted R-Squared is increased up to 0.8711.

Table 7

The effect of other variables on the amount of purchase

	Model 2		Model 1	
	Coefficient	Standard error	Coefficient	Standard error
Intercept	132.84***	7.67	142.86***	18.92
Number of products browsed	1.72***	0.11	1.72***	0.11
Average number of pictures			0.58	0.78
Shipping time	-11.48***	0.80	-11.43***	0.80
Review rating	11.08***	1.55	10.99***	1.56
Age	-0.94***	0.07	-0.94***	0.07
PC	74.57***	1.75	74.63***	1.75
Social media advertisement (ref)				
Search engine	-65.31***	1.77	-65.25***	1.77
Other	-90.45***	10.89	-90.07***	10.90
Friends or Family	64.24***	2.97	64.14***	2.97
Ease of purchase 3 (ref)				
Ease of purchase 2			11.86	17.26
Ease of purchase 4	42.12***	2.14	42.27***	2.15
Ease of purchase 5	82.30***	2.77	82.48***	2.79
Model Parameters				
R-Squared	0.8725		0.8726	
Adjusted R-Squared	0.8711		0.8710	
N	969		969	

*Note. *** p<0.001 ** p<0.01 *p<0.05. All continuous variables are standardized.*

With the following information of a customer in Table 8, the prediction of the amount purchased by a customer calculated as 212.68.

Table 8

The information of a customer

Products	Device	Pictures	Shipping time	Review rating	Find website	Ease of purchase	Age
17	PC	3.5	2	4.3	Search engine	4	27

Missing values on the survey questions

By creating dummy variables and using MICE, there are some changes in the outcome of new model. In overall, the new R-Squared decreases from 0.8726 to 0.8580 as well as adjusted R-Squared. The number of observations increases from 969 to 1080. In the new model, average number of pictures and ease of purchase rating 2 are still not statistically significant. New dummy variables for missing values have influence on the purchase amount spent on the website. These two variables are just missing values at random. There is not much difference between models. Multiple imputation has not produced any strange results. The conclusions still stay the same so this can be a good solution.

Table 9

The effect of other variables on the amount of purchase

	Model 3		Model 1	
	Coefficient	Standard error	Coefficient	Standard error
Intercept	130.84***	8.12	142.86***	18.92
Number of products browsed	1.80***	0.11	1.72***	0.11
Average number of pictures	0.74	0.78	0.58	0.78
Shipping time	-11.19***	0.79	-11.43***	0.80
Review rating	10.97***	1.59	10.99***	1.56
Age	-0.96***	0.07	-0.94***	0.07
PC	74.64***	1.74	74.63***	1.75
Social media advertisement (ref)				
Search engine	-66.27***	1.80	-65.25***	1.77
Other	-80.17***	10.50	-90.07***	10.90
Friends or Family	63.10***	3.01	64.14***	2.97
NA	-38.60***	4.51		
Ease of purchase 3 (ref)				
Ease of purchase 2	11.75	18.15	11.86	17.26
Ease of purchase 4	40.65***	2.18	42.27***	2.15
Ease of purchase 5	81.09***	2.85	82.48***	2.79
Ease of purchase NA	43.26***	4.63		
Model Parameters				
R-Squared	0.8580		0.8726	
Adjusted R-Squared	0.8562		0.8710	
N	1080		969	

*Note. *** p<0.001 ** p<0.01 *p<0.05. All continuous variables are standardized.*

Article by Al-Zyoud (2018)

Sample

The sample size is quite small only 150 participants and limited within the Jordanian iron and metal industry (Al-Zyoud, 2018, p. 108). The population size is more than 2 billion people who use social networking (Al-Zyoud, 2018, p. 102). It has selection bias while the participants are only in iron and metal industry.

Survey

The survey has an operationalization with three constructs: functional branding, intentional branding and social media marketing. However, only functional branding has two clear dimensions: usability and customer interactions. There is no mention about the dimensions of the other two constructs.

Drop-off and pick-up survey method has some advantages but there are other more efficient survey methods that can be used: online survey, mail survey…

Analyses

Because the sample is bias, the result for the demographic question of age is not valid. The range of age is above 25 while social networking is widely used from 18 to 29 in USA (Social media usage statistics by age, n.d.).

It is good that the test is tested with Cronbachs' alpha to calculate the correlation coefficient between all items of a construct.

Functional branding and intentional branding both have significant influence on social media marketing. Functional branding and intentional branding can have multicollinearity. This should be taken into account.

For the H3c hypothesis, it is impossible to use mean of each dimension to jump to conclusion that social media marketing strategies are more strongly inclined towards ability and intention. These dimensions coming from three constructs with different indicators/items. Regression model should be conducted to analyze the relationships and check the spuriousness. Because correlation is not equal causation. These can be statistically significant relationship between variables but it could be coincidence or there is a third variable that does matter.

References

Al-Zyoud, M. F. (2018). Social media marketing, functional branding strategy and intentional branding. *Problems and Perspectives in Management, 16*(3), 102-116. doi:10.21511/ppm.16(3).2018.09