Data Analysis: Artificial Intelligence in Online Retail in China

Donna Casa

Dinesh Rawat

Aditya Prashant Gurjar

Mohammad Sabir Shah

Rohit Mahendran

Radhika Jayakrishnan

Cape Breton University

Table of Contents

Objective of the Study:

The report aims to establish relationships between two variables, which are customer loyalty and the level of usage of a certain artificial intelligence application (chatbots) as case # 1. For case # 2, the relationship of another two variables, level of income and incidence of online shopping, are being established. IBM SPSS was used as an analysis tool for the data gathered by marketing.

Data Collection:

In the marketing survey questionnaires, 252 respondents were asked which of the seven artificial intelligence applications have they used while online shopping, which are chatbots, virtual assistance device, personalized product recommendations, personalized offers with augmented reality views, virtual research solution, smart display, and aisle kiosk. Customer loyalty was measured by a 5-point Likert scale, Strongly Disagree, Disagree, Neutral, Agree, and Strongly Agree. Income range was measured as Less than CNY 35,000, CNY 35,000 – CNY 75,000, CNY 75,000- CNY 100,00, and More than CNY 100,000. Lastly, Online shopping incidence was measured as 1 purchase a month, Up to 5 purchases each month, 5-10 purchases a month, and More than 10 purchases a month.

Testing Hypotheses for Case # 1 using Chi-Square:

Null Hypothesis (H0): Level of usage of application 1 is not related to customer loyalty
Alternative Hypothesis (H1): Level of usage of application 1 is related to customer loyalty
Dependent Variable: Customer loyalty

Independent Variable: Level of usage of application 1 (Chatbots)

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Usage_Chatbot * I am more loyal as a customer when I shop on line with artificial intelligence tools	249	98.8%	3	1.2%	252	100.0%

Usage_Chatbot * I am more loyal as a customer when I shop on line with artificial intelligence tools Crosstabulation

Count

		I am more loyal as a customer when I shop on line with artificial intelligence tools					
		Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree	Total
Usage_Chatbot	not using	12	27	48	64	21	172
	using	9	5	15	36	12	77
Total		21	32	63	100	33	249

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	8.062[a]	4	.089
Likelihood Ratio	8.484	4	.075
Linear-by-Linear Association	1.086	1	.297
N of Valid Cases	249		

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 6.49.

Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	.180	.089
	Cramer's V	.180	.089
N of Valid Cases		249	

Analysis:

Chi-square tests are performed to find the correlation coefficient or to check if 2 variables are dependent on each other. From the Chi-Square test table, we observe that 0 cells (0.0%) have an expected count of less than 5. If the percentage was over 20, it would imply that a chi-square table assumption has been violated. But according to the Chi-Square test table, the minimum expected value is 6.49 and the assumption has not been violated.

The Asymptotic significance or the p-value observed is 0.089 or 8.9%. Assuming that the significance level is 5%, it is observed that the p-value is higher than the significance level. Hence, there is no correlation between customer loyalty and the usage of AI tools (Chatbots).

Testing Hypotheses for Case # 2 using Correlation and Ordinal Regression:

Null Hypothesis (H0): Level of income is not related to online shopping incidence
Alternative Hypothesis (H1): Level of income is related to online shopping incidence
Dependent Variable: Shopping incidence
Independent Variable: Level of income

Correlations

		What is your annual income range?	How often do you make online purchase?
What is your annual income range?	Pearson Correlation	1	.289**
	Sig. (2-tailed)		<.001
	N	252	252
How often do you make online purchase?	Pearson Correlation	.289**	1
	Sig. (2-tailed)	<.001	
	N	252	252

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

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	What is your annual income range?[b]	.	Enter

a. Dependent Variable: How often do you make online purchase?

b. All requested variables entered.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.289[a]	.083	.080	.85840

a. Predictors: (Constant), What is your annual income range?

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	16.752	1	16.752	22.734	<.001[b]
	Residual	184.212	250	.737		
	Total	200.964	251			

a. Dependent Variable: How often do you make online purchase?

b. Predictors: (Constant), What is your annual income range?

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	1.593	.200		7.965	<.001
	What is your annual income range?	.293	.062	.289	4.768	<.001

a. Dependent Variable: How often do you make online purchase?

Analysis:

In the correlation table, total income will be perfectly correlated with itself and the frequency of online purchases will perfectly correlate with itself as well, which is why their Pearson values are 1. The value 0.289 shows the strength of the correlation between the variables. It's a moderate strength positive correlation. The significance is <.001, therefore, the relation between the variables is statistically significant.

In the Anova test, since the significance value is less than 0.01, we can say that the variables are highly dependent on each other. The frequency of online purchases is highly dependent on the person's income.

Conclusion:

For case # 1, we failed to reject the Null Hypothesis (H0): Level of usage of application 1 is not related to customer loyalty using the Chi-square test at a 5% significance level. Hence, the Alternative Hypothesis (H1): Level of usage of application 1 is related to customer loyalty is not supported.

For case #2, we reject the Null Hypothesis (H0): Level of income is not related to online shopping incidence using correlation and ordinal regression tests at a 5% significance level. Hence, we accept the Alternative Hypothesis (H1): Level of income is related to online shopping incidence. There is a significant relationship between customer income and shopping incidence.