

Hypotheses Testing for customer satisfaction towards online shopping

Presented to:

Dr. Nermin Elgohary

Presented by:

Omar Elsayed Darwesh

Mustafa Nasser Ali

Rana Ali Ibrahim

Overview

Conducted a survey to figure out factors that may correlate with and may affect the customer satisfaction for online shopping experience.

Ten factors/variables are collected by ten questions through an online survey namely Age, Product Name, Platform Name, Overall Satisfaction Rate, Delivery Time Rate, Platform Design Rate, Packaging Quality Rate, Discount Percent, Waiting Days Rate and Price.

Dropped three factors namely Age, Product Name and Platform Name since they will not help figuring out effect as the number of observations is very small.

Therefore, all factors used in the following statistical tests are:

The phenomena/dependent variable is Overall Satisfaction Rate.

The six independent variables that may correlate with the phenomena are Delivery Time Rate, Platform Design Rate, Packaging Quality Rate, Discount Percent, Waiting Days Rate and Price.

Files attached to the report:

- Customer Satisfaction towards online shopping Feedback (Survey Responses)-Dataset (.xlsx file)
- Customer Satisfaction towards online shopping-SPSS-Data-Document (.spv file)
- Customer Satisfaction towards online shopping-SPSS-Output-Document (.sav file)

Normality Testing

Note that in this very normality testing, the kurtosis and skewness shall be within the range between +3 and -3 to be considered normal.

Tested for normality and found that

- The dependent variable Overall Satisfaction Rate is normally distributed. Its skewness is -1.251 and its kurtosis is 2.054
- The independent variable Delivery Time Rate is normally distributed. Its skewness is -0.037 and its kurtosis is 0.828
- The independent variable Platform Design Rate is normally distributed. Its skewness is -0.052 and its kurtosis is -1.465
- The independent variable Packaging Quality Rate is normally distributed. Its skewness is -1.146 and its kurtosis is 2.214
- The independent variable Discount Percent is normally distributed. Its skewness is 0.944 and its kurtosis is 0.119
- The independent variable Waiting Days Rate is normally distributed. Its skewness is 0.305 and its kurtosis is -.549
- The independent variable Price is NOT normally distributed. Its skewness is 3.772 and its kurtosis is 16.279

Therefore, Pearson Correlation test will be used for all normally distributed variables and Spearman correlation test for non-normally distributed variables.

Hypothesis Testing and Correlation test

Constructed a hypothesis between the phenomena and each variable of the independent variables.

Tested if there's a significant relationship between the phenomena and each variable of the independent variables using Pearson correlation test for normally distributed variables and Spearman correlation test for not normally distributed variables.

NOTE that in this very correlation test the significance level is 0.05.

1- The hypothesis for Delivery Time Rate and Overall Satisfaction Rate:

- Null Hypothesis: there is no relationship between Delivery Time Rate and Overall Satisfaction Rate.
- Alternative Hypothesis: there is a significant relationship between Delivery Time Rate and Overall Satisfaction Rate.

P-value = 0.003

Pearson Correlation coefficient (r) = 0.481

- Since the P-value is smaller than the significance level, the Null hypothesis is rejected. So, there is a significant relationship between Delivery Time Rate and Overall Satisfaction Rate. It's a direct moderate relationship.

2- The hypothesis for Platform Design Rate and Overall Satisfaction Rate:

- Null Hypothesis: there is no relationship between Platform Design Rate and Overall Satisfaction Rate.
- Alternative Hypothesis: there is a significant relationship between Platform Design Rate and Overall Satisfaction Rate.

P-value = 0.003

Pearson Correlation coefficient (r) = 0.486

- Since the P-value is smaller than the significance level, the Null hypothesis is rejected. So, there is a significant relationship between Platform Design Rate and Overall Satisfaction Rate. It's a direct moderate relationship.

3- The hypothesis for Packaging Quality Rate and Overall Satisfaction Rate:

- Null Hypothesis: there is no relationship between Packaging Quality Rate and Overall Satisfaction Rate.
- Alternative Hypothesis: there is a significant relationship between Packaging Quality Rate and Overall Satisfaction Rate.

P-value = 0.0

Pearson Correlation coefficient (r) = 0.740

- Since the P-value is smaller than the significance level, the Null hypothesis is rejected. So, there is a significant relationship between Packaging Quality Rate and Overall Satisfaction Rate. It's a direct strong relationship.

4- The hypothesis for Discount Percent and Overall Satisfaction Rate:

- Null Hypothesis: there is no relationship between Discount Percent and Overall Satisfaction Rate.
- Alternative Hypothesis: there is a significant relationship between Discount Percent and Overall Satisfaction Rate.

P-value = 0.744

Pearson Correlation coefficient (r) = 0.062

- Since the P-value is greater than the significance level, we fail to reject the Null hypothesis. So, there is NO significant relationship between Discount Percent and Overall Satisfaction Rate.

5- The hypothesis for Waiting Days Rate and Overall Satisfaction Rate:

- Null Hypothesis: there is no relationship between Waiting Days Rate and Overall Satisfaction Rate.
- Alternative Hypothesis: there is a significant relationship between Waiting Days Rate and Overall Satisfaction Rate.

P-value = 0.032

Pearson Correlation coefficient (r) = 0.359

- Since the P-value is smaller than the significance level, the Null hypothesis is rejected. So, there is a significant relationship between Waiting Days Rate and Overall Satisfaction Rate. It's a direct moderate relationship.

6- The hypothesis for Price and Overall Satisfaction Rate:

- Null Hypothesis: there is no relationship between Price and Overall Satisfaction Rate.
- Alternative Hypothesis: there is a significant relationship between Price and Overall Satisfaction Rate.

P-value = 0.278

Spearman Correlation coefficient (r) = 0.186

- Since the P-value is greater than the significance level, we fail to reject the Null hypothesis. So, there is NO significant relationship between Price and Overall Satisfaction Rate.

Hypothesis Testing and Regression test

Constructed a hypothesis between the phenomena and each variable of the independent variables.

Tested if there's a significant impact of each variable of the independent variables on the phenomena using simple and multiple Regression test.

Included only four of the six independent variables that resulted in a significant relationship with the phenomena by the correlation test namely Delivery Time Rate, Platform Design Rate, Packaging Quality Rate and Waiting Days Rate.

NOTE that in this very regression test the significance level is 0.05.

▪ Simple Linear Regression test

1- The hypothesis for Delivery Time Rate and Overall Satisfaction Rate:

- Null Hypothesis: there is no significant impact of Delivery Time Rate on Overall Satisfaction Rate.
- Alternative Hypothesis: there is a significant impact of Delivery Time Rate on Overall Satisfaction Rate.

P-value = 0.003

R Square = 0.231

- Since the P-value is smaller than the significance level, the Null hypothesis is rejected. So, there is a significant impact of Delivery Time Rate on Overall Satisfaction Rate.
- Since the R Square = 0.231, this means that 23.1% of the variation in the dependent variable (Overall Satisfaction Rate) is explained by the variation in the independent variable (Delivery Time Rate).

2- The hypothesis for Platform Design Rate and Overall Satisfaction Rate:

- Null Hypothesis: there is no significant impact of Platform Design Rate on Overall Satisfaction Rate.
- Alternative Hypothesis: there is a significant impact of Platform Design Rate on Overall Satisfaction Rate.

P-value = 0.003

R Square = 0.237

- Since the P-value is smaller than the significance level, the Null hypothesis is rejected. So, there is a significant impact of Platform Design Rate on Overall Satisfaction Rate.
- Since the R Square = 0.237, this means that 23.7% of the variation in the dependent variable (Overall Satisfaction Rate) is explained by the variation in the independent variable (Platform Design Rate).

3- The hypothesis for Packaging Quality Rate and Overall Satisfaction Rate:

- Null Hypothesis: there is no significant impact of Packaging Quality Rate on Overall Satisfaction Rate.
- Alternative Hypothesis: there is a significant impact of Packaging Quality Rate on Overall Satisfaction Rate.

P-value = 0.000

R Square = 0.548

- Since the P-value is smaller than the significance level, the Null hypothesis is rejected. So, there is a significant impact of Packaging Quality Rate on Overall Satisfaction Rate.
- Since the R Square = 0.548, this means that 54.8% of the variation in the dependent variable (Overall Satisfaction Rate) is explained by the variation in the independent variable (Packaging Quality Rate).

4- The hypothesis for Waiting Days Rate and Overall Satisfaction Rate:

- Null Hypothesis: there is no significant impact of Waiting Days Rate on Overall Satisfaction Rate.
- Alternative Hypothesis: there is a significant impact of Waiting Days Rate on Overall Satisfaction Rate.

P-value = 0.032

R Square = 0.129

- Since the P-value is smaller than the significance level, the Null hypothesis is rejected. So, there is a significant impact of Waiting Days Rate on Overall Satisfaction Rate.
- Since the R Square = 0.129, this means that 12.9% of the variation in the dependent variable (Overall Satisfaction Rate) is explained by the variation in the independent variable (Waiting Days Rate).

▪ Multiple Linear Regression test

- The hypothesis for Delivery Time Rate, Platform Design Rate, Packaging Quality Rate and Waiting Days Rate and Overall Satisfaction Rate:

- Null Hypothesis: there is no significant impact of Delivery Time Rate, Platform Design Rate, Packaging Quality Rate and Waiting Days Rate combined on Overall Satisfaction Rate.
- Alternative Hypothesis: there is a significant impact of Delivery Time Rate, Platform Design Rate, Packaging Quality Rate and Waiting Days Rate combined on Overall Satisfaction Rate.

Delivery Time Rate P-value = 0.309

Platform Design Rate P-value = 0.605

Packaging Quality Rate P-value = 0.000

Waiting Days Rate P-value = 0.637

Adjusted R Square = 0.529

- Since the P-value of Packaging Quality Rate is smaller than the significance level and all other p-values is greater than the significance level, the Null hypothesis is rejected. So, there is a significant impact of Packaging Quality Rate on Overall Satisfaction Rate.
- Despite the fact that all other independent variables have significant impact on the dependent variable when tested individually through simple linear regression, these very independent variables turned out to be not significant at all and don't have significant impact on the customer satisfaction rate in the presence of Packaging Quality Rate.
- Since the R Square = 0.529, this means that 52.9% of the variation in the dependent variable (Overall Satisfaction Rate) is explained by the variation in the independent variables combined (Delivery Time Rate, Platform Design Rate, Packaging Quality Rate and Waiting Days Rate).