

Empirical study
Deal Finder Service
Project

Manoj Paidimarri	180101054
Gali Jaya Prakash Reddy	180101025
Bedada Ajay Kumar	180101014
Doddavula Likhith Kumar Reddy	180101098

TABLE OF CONTENTS

1. Introduction to Empirical Study :	3
2. Identification of Research Question :	3
2.1 Validity of the research question :	3
2.2 Research Hypothesis :	4
3. Determination Of Variables:	4
3.1 Identification and Assignment of Tasks :	5
4. Design Of Experiment :	6
5. Analysis Of Empirical Data :	10
5.1 Tests of Normality :	10
5.2 Kruskal-Wallis test :	15
6. Conclusion:	17

Empirical Study

1. Introduction to Empirical Study :

Empirical study is broadly divided into four stages

- 1) Identification of research question
- 2) Determination of variables
- 3) Design of experiment
- 4) Analysis of empirical data

In identification of the research question stage, we are supposed to observe the relationship between the user's deal judgment behavior and the deals.

To observe this relationship, we will set up an experiment, we present a deal to the user and ask him/her to judge the deal. We change the deals and ask the user to judge it again. We repeat the process a few times and complete our experiment.

2. Identification of Research Question :

RQ: How likely are you to buy a product (on a scale of 1-10) depends on the product's price, the Rating given for the product, and number of Images shown for the product in a deal?

2.1 Validity of the research question :

Here we framed a more specific question, i.e., a testable question. RQ has internal validity. However, we may not get the true answer due to a lack of external validity. So to balance this trade-off, we framed multiple testable questions using different rating scales.

RQ1: How likely are you to buy a product(on a scale of 1-10) depends on the product's price, the Rating given for the product, and number of Images shown for the product in a deal?

RQ2: How likely are you to buy a product(on a scale of 1-3) depends on the product's price, the Rating given for the product, and number of Images shown for the product in a deal?

RQ3: How likely are you to buy a product(on a scale of 1-5) depends on the product's price, the Rating given for the product, and number of Images shown for the product in a deal?

Here we can capture three different judgment behaviors in terms of the three rating scales in the question and we are likely to arrive at a generalized answer for an untestable question from the specific answers to multiple testable research questions, this is a better approach than having only untestable questions and user feedback.

2.2 Research Hypothesis :

In this step, we start with two hypotheses : null hypothesis and alternative hypothesis

The two hypotheses for RQ is

H_0 : The likeliness to buy a product does not depend on the product's price, the Rating given for the product, and the Images shown for the product in a deal.

H_1 : The likeliness to buy a product depends on the product's price, the Rating given for the product, and the Images shown for the product in a deal.

In null hypothesis, we are essentially stating that the test condition is not going to affect the judgment. In the alternative hypothesis, it is just the opposite. We are stating that the test condition does affect outcome.

3.Determination Of Variables:

Here the likeliness to buy a product is a variable that takes values from the rating scale i.e it can take values between 1 to 10 (integer only). First variable

(Product price) can take any integer. For the second variable(rating) can take values between 1 to 5. Third value (number of images shown) can take any integer. Now we know variables and their values,by using this we generated the test conditions by varying values systematically. We recorded the data manually.

So,the scales we used Interval scale for product price, nominal scale for rating and number of images shown.

For the experiment we are taking the deals through systematic variations of variables. For the number of images we planned to have values 1,3 to see the effect of less number of images shown on likeliness to buy and similarly for more number of images shown in second case. For product price we take three values 800,1000,1200. Three values 2,3.5,4.8 are taken for the rating. Therefore altogether $3 \times 3 \times 2 = 18$ deals for our empirical study. Here the likeliness to buy a product is a dependent variable i.e depends on the factors of the deal.

We performed a controlled experiment.For our experiment,we used only the participants of age between 20 and 30.So conclusions we drew will be applicable to the users of age between 20 and 30 only.

3.1 Identification and Assignment of Tasks :

In the study on the likelihood to buy a product, the task for participants is to rate the deal. We have 15 participants.We divide them into 3 groups (5 in each).We will ask each group to rate only 6 deals.Each participant in a group rates all 6 deals. Since we are distributing the tasks to participants, we call the study as "between - subject".

The confounding variable practise effect is taken care of by a counterbalancing sequence in which tasks are given to the participants. To do this we have used latin square method but the recordings are shown in the following way.

4. Design Of Experiment :

The deals are described as follows,

Deal Number	Price	Rating	No of Images
1	800	2	1
2	800	2	3
3	800	3.5	1
4	800	3.5	3
5	800	4.8	1
6	800	4.8	3
7	1000	2	1
8	1000	2	3
9	1000	3.5	1
10	1000	3.5	3
11	1000	4.8	1
12	1000	4.8	3
13	1200	2	1
14	1200	2	3
15	1200	3.5	1
16	1200	3.5	3
17	1200	4.8	1
18	1200	4.8	3

Test conditions for the experiment :

Factor	Levels		
Price	800	1000	1200
Rating	2	3.5	4.8
No of Images	1	3	

4.1 RQ1 (Rating 1-10)

Group 1:

	D14	D3	D9	D16	D7	D15
P1	5	7	7	8	5	6
P2	6	6	6	8	4	5
P3	3	5	6	8	3	6
P4	2	6	6	7	4	6
P5	4	6	6	8	4	6

Group 2:

	D4	D8	D1	D12	D11	D6
P6	10	4	2	10	7	8
P7	9	3	1	9	7	8
P8	8	6	2	10	8	10
P9	9	5	3	10	8	9
P10	9	5	3	10	8	10

Group 3:

	D5	D13	D2	D18	D17	D10
P11	7	4	6	9	8	7

P12	6	4	5	10	8	8
P13	7	3	4	8	7	8
P14	8	2	4	9	8	8
P15	7	2	4	10	8	9

4.2 RQ2 (Rating 1-3)

Group 1 :

	D14	D3	D9	D16	D7	D15
P1	2	2	2	3	1	2
P2	2	2	2	3	1	2
P3	1	2	2	3	1	2
P4	1	2	2	2	1	2
P5	1	2	2	3	1	2

Group 2 :

	D4	D8	D1	D12	D11	D6
P6	3	1	1	3	2	3
P7	3	1	1	3	2	3
P8	3	2	2	3	3	3
P9	3	2	3	3	3	3
P10	3	2	3	3	3	3

Group 3 :

	D5	D13	D2	D18	D17	D10
P11	2	2	2	3	3	2
P12	2	2	2	3	3	3

P13	2	1	2	3	2	3
P14	3	1	2	3	3	3
P15	2	1	2	3	3	3

4.3 RQ3(Rating 1-5)

Group 1:

	D14	D3	D9	D16	D7	D15
P1	3	4	4	4	3	3
P2	3	3	4	4	2	3
P3	2	3	3	4	2	3
P4	1	3	3	4	2	3
P5	2	3	3	4	2	3

Group 2:

	D4	D8	D1	D12	D11	D6
P6	5	2	1	5	4	4
P7	5	1	1	5	4	4
P8	4	3	1	5	4	5
P9	5	3	2	5	4	5
P10	5	3	2	5	4	5

Group 3:

	D5	D13	D2	D18	D17	D10
P11	4	2	3	5	4	4
P12	3	2	3	5	4	4

P13	4	2	2	4	4	4
P14	4	1	2	5	4	4
P15	4	1	2	5	4	5

5. Analysis Of Empirical Data :

5.1 Tests of Normality :

5.1.1 RQ1 (Rating 1-10)

Group 1:

Tests of Normality

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
d14	.136	5	.200 [*]	.987	5	.967
d3	.300	5	.161	.883	5	.325
d9	.473	5	<.001	.552	5	<.001
d15	.473	5	<.001	.552	5	<.001
d7	.300	5	.161	.883	5	.325
d16	.473	5	<.001	.552	5	<.001

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Group 2:

Tests of Normality

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
d4	.300	5	.161	.883	5	.325
d8	.237	5	.200 [*]	.961	5	.814
d1	.231	5	.200 [*]	.881	5	.314
d12	.473	5	<.001	.552	5	<.001
d11	.367	5	.026	.684	5	.006
d6	.241	5	.200 [*]	.821	5	.119

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Group 3:

Tests of Normality

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
d5	.300	5	.161	.883	5	.325
d13	.241	5	.200 [*]	.821	5	.119
d2	.349	5	.046	.771	5	.046
d18	.231	5	.200 [*]	.881	5	.314
d17	.473	5	<.001	.552	5	<.001
d10	.300	5	.161	.883	5	.325

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

5.1.2 RQ2 (Rating 1-3) :

Group 1:

Tests of Normality						
	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
d14	.367	5	.026	.684	5	.006
d3	.	5	.	.	5	.
d9	.	5	.	.	5	.
d16	.473	5	<.001	.552	5	<.001
d7	.	5	.	.	5	.
d15	.	5	.	.	5	.

a. Lilliefors Significance Correction

Group 2:

Tests of Normality						
	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
d4	.	5	.	.	5	.
d8	.367	5	.026	.684	5	.006
d1	.241	5	.200 [*]	.821	5	.119
d12	.	5	.	.	5	.
d11	.367	5	.026	.684	5	.006
d6	.	5	.	.	5	.

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Group 3:

Tests of Normality

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
d5	.473	5	<.001	.552	5	<.001
d13	.367	5	.026	.684	5	.006
d2	.	5	.	.	5	.
d18	.	5	.	.	5	.
d17	.473	5	<.001	.552	5	<.001
d10	.473	5	<.001	.552	5	<.001

a. Lilliefors Significance Correction

5.1.3 RQ-3(Rating 1-5)

Group-1:

Tests of Normality

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
d14	.231	5	.200 [*]	.881	5	.314
d3	.473	5	<.001	.552	5	<.001
d9	.367	5	.026	.684	5	.006
d16	.	5	.	.	5	.
d7	.473	5	<.001	.552	5	<.001
d15	.	5	.	.	5	.

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Group-2:

Tests of Normality

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
d4	.473	5	<.001	.552	5	<.001
d8	.349	5	.046	.771	5	.046
d1	.367	5	.026	.684	5	.006
d12	.	5	.	.	5	.
d11	.	5	.	.	5	.
d6	.367	5	.026	.684	5	.006

a. Lilliefors Significance Correction

Group-3:

Tests of Normality

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
d5	.473	5	<.001	.552	5	<.001
d13	.367	5	.026	.684	5	.006
d2	.367	5	.026	.684	5	.006
d18	.473	5	<.001	.552	5	<.001
d17	.	5	.	.	5	.
d10	.473	5	<.001	.552	5	<.001

a. Lilliefors Significance Correction

By the above tests for normality we observed that the data is not following normal distribution. So we have performed a non parametric test. Since we have a between-subject design with three factors we have applied the "KRUSKAL- WALLIS" test.

5.2 Kruskal-Wallis test :

5.2.1 RQ1 (Rating 1-10)

Group 1 :

Test Statistics ^{a,b}						
	d5	d13	d2	d18	d17	d10
Kruskal-Wallis H	4.000	4.000	4.000	4.000	4.000	4.000
df	4	4	4	4	4	4
Asymp. Sig.	.406	.406	.406	.406	.406	.406

a. Kruskal Wallis Test

b. Grouping Variable: persons

Group 2 :

Test Statistics ^{a,b}						
	d14	d3	d9	d16	d7	d15
Kruskal-Wallis H	4.000	.000	.000	4.000	.000	.000
df	4	4	4	4	4	4
Asymp. Sig.	.406	1.000	1.000	.406	1.000	1.000

a. Kruskal Wallis Test

b. Grouping Variable: persons

Group 3 :

Test Statistics ^{a,b}						
	d5	d13	d2	d18	d17	d10
Kruskal-Wallis H	4.000	4.000	.000	.000	4.000	4.000
df	4	4	4	4	4	4
Asymp. Sig.	.406	.406	1.000	1.000	.406	.406

a. Kruskal Wallis Test

b. Grouping Variable: persons

5.2.2 RQ2 (Rating 1-3)

Group 1,2,3 :

Test Statistics^{a,b}

	d14	d3	d9	d16	d7	d15
Kruskal-Wallis H	4.000	4.000	4.000	.000	4.000	.000
df	4	4	4	4	4	4
Asymp. Sig.	.406	.406	.406	1.000	.406	1.000

a. Kruskal Wallis Test

b. Grouping Variable: persons

Test Statistics^{a,b}

	d4	d8	d1	d12	d11	d6
Kruskal-Wallis H	4.000	4.000	4.000	.000	.000	4.000
df	4	4	4	4	4	4
Asymp. Sig.	.406	.406	.406	1.000	1.000	.406

a. Kruskal Wallis Test

b. Grouping Variable: persons

Test Statistics^{a,b}

	d5	d13	d2	d18	d17	d10
Kruskal-Wallis H	4.000	4.000	4.000	4.000	.000	4.000
df	4	4	4	4	4	4
Asymp. Sig.	.406	.406	.406	.406	1.000	.406

a. Kruskal Wallis Test

b. Grouping Variable: persons

5.2.3 RQ3 (Rating 1-5)

Group 1,2,3:

Test Statistics ^{a,b}						
	d4	d8	d1	d12	d11	d6
Kruskal-Wallis H	4.000	4.000	4.000	.000	.000	4.000
df	4	4	4	4	4	4
Asymp. Sig.	.406	.406	.406	1.000	1.000	.406

a. Kruskal Wallis Test

b. Grouping Variable: persons

Test Statistics ^{a,b}						
	d14	d3	d9	d16	d7	d15
Kruskal-Wallis H	4.000	4.000	4.000	.000	4.000	.000
df	4	4	4	4	4	4
Asymp. Sig.	.406	.406	.406	1.000	.406	1.000

a. Kruskal Wallis Test

b. Grouping Variable: persons

Test Statistics ^{a,b}						
	d5	d13	d2	d18	d17	d10
Kruskal-Wallis H	4.000	4.000	4.000	4.000	.000	4.000
df	4	4	4	4	4	4
Asymp. Sig.	.406	.406	.406	.406	1.000	.406

a. Kruskal Wallis Test

b. Grouping Variable: persons

6.Conclusion:

By the above analysis we refute the null hypothesis and conclude that likelihood to buy an item depends on the factors price, rating, number of images.