A/B Testing Data Analysis for Retail Ecommerce Webpage

Chi Square hypothesis testing for discrete KPIs

To conduct Chi Square test, I have used Excel. The level of significance is 0.05.

The method of conducting Chi Square hypothesis test is explained in the Appendix B.

Hypothesis testing for repeat sessions

Null hypothesis: There is no association between existence of Haptic feature and whether a session is new or repeated.

Alternate hypothesis: There is an association between existence of Haptic feature and whether a session is new or repeated.

The result is as follows:

Chi-square value	4.02
df	1
p-value	0.04487457

Since p-value is less than 0.05, we **reject** the null hypothesis and accept the alternate hypothesis.

This means that, there is a statistically significant association between existence of Haptic feature and whether a user is new or not. Further viewing the results, we see that:

	With Haptic	**Without Haptic**
New visits	68.92%	69.97%
Repeat visits	**31.08%**	30.03%

Webpage with Haptic feature saw 31.08% repeat visits, whereas that without Haptic saw 30.03% repeat visits. **Percentage of repeat visits was more by 1.05% in the case of website version with Haptic feature**.

Hypothesis testing for bounced sessions

Null hypothesis: There is no association between existence of Haptic feature and whether a session bounced or not.

Alternate hypothesis: There is an association between existence of Haptic feature and whether a a session bounced or not.

The result is as follows:

Chi-square value	14.59
df	1

p-value	0.000133753

Since p-value is less than 0.05, we **reject** the null hypothesis and accept the alternate hypothesis.

This means that, there is a statistically significant association between existence of Haptic feature and whether a session is bounced or not. Further viewing the details, we see that:

	With Haptic	**Without Haptic**
Bounced sessions	**32.28%**	33.72%
Not bounced sessions	67.72%	66.28%

Webpage with Haptic feature saw 32.28% bounced sessions, whereas that without Haptic saw 33.72% bounced sessions. **Percentage of bounced sessions reduced for version with haptic feature, which is a positive for the ecommerce.**

Hypothesis testing for converted sessions

Null hypothesis: There is no association between existence of Haptic feature and whether a session is converted or not.

Alternate hypothesis: There is an association between existence of Haptic feature and whether a session is converted or not.

The result is as follows:

Chi-square value	16.76
df	1
p-value	0.0000425

Since p-value is less than 0.05, we **reject** the null hypothesis and accept the alternate hypothesis.

This means that, there is a statistically significant association between existence of Haptic feature and whether a session is converted or not. Further viewing the details, we see that:

	With Haptic	**Without Haptic**
Converted sessions	**5.21%**	4.51%
Not converted sessions	94.79%	95.49%

Webpage with Haptic feature saw 5.21% converted sessions, whereas that without Haptic saw 4.51% converted sessions. **Percentage of converted sessions increased for version with haptic feature, which is a big positive for the ecommerce.**

Two-sample T-Test hypothesis testing for continuous KPIs

I have used the two-sample T-Test for hypothesis testing to compare the two samples of when Haptic feature is used, and when Haptic feature is not used. The result is computed using Excel.

KPI	Null Hypothesis	Alternate Hypothesis	Two-tail p-value	Conclusion
Average session duration	Means of session duration are equal.	Means of session duration are unequal.	0.011718254	p-value is less than 0.05; So we reject null hypothesis.
Average revenue per session	Means of revenue per session are equal.	Means of revenue per session are unequal.	0.974303394	p-value is more than 0.05; So we cannot reject the null hypothesis.
Average order value	Means of order value are equal.	Means of order value are unequal.	0.022022017	p-value is less than 0.05; So we reject null hypothesis.
Average revenue per user	Means of revenue per user are equal.	Means of revenue per user are unequal.	0.974308182	p-value is more than 0.05; So we cannot reject the null hypothesis.

Summary of findings

KPI	Results
Number of new users	Statistically significant difference; webpage with Haptic feature has a slightly **worse** rate of bringing in new users than webpage without Haptic feature.
Bounce rate	Statistically significant difference; webpage with Haptic feature has a slightly **better** rate of reducing bounced session than webpage without Haptic feature.
Conversion	Statistically significant difference; webpage with Haptic feature has a **better** conversion rate than webpage without Haptic feature.
Average session duration	Means of session duration are significantly unequal; average session duration is **higher** for webpage with Haptic feature.

Average revenue per session	There is no significant statistical difference between the means of revenue per session for the two.
Average order value	Means of order value are significantly unequal; average order value is **lower** for webpage without Haptic feature.
Average revenue per user	There is no significant statistical difference between the means of revenue per user for the two.

Decision

In order to decide whether we should go with Haptic feature or not, let's see if the results fulfilled our primary and secondary KPIs.

Primary KPI	Conversion Rate	**Increased with Haptic feature**
Secondary KPI	Average Revenue per User	**No statistically significant difference between the features**

Since the Haptic feature successfully increased conversion rate and had no negative effect on Average Revenue per User, the ecommerce **should implement the Haptic feature**.

This is further supported by the facts that:

- Haptic feature showed reduction in Bounce Rate compared to no Haptic feature, thus increasing the probability of better SERP ranking and greater traffic.
- Haptic feature showed higher Average Session Duration, which improves the ecommerce's chances for revenue generation.'

Recommendations

We have decided to keep the Haptic feature because it's the best for higher conversion, but statistical analyses showed that there are some drawbacks that need to be addressed:

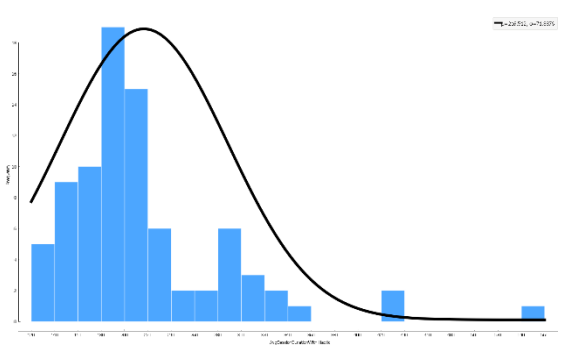
- **To gain more new users**: It is seen that webpage with Haptic feature that a slightly worse number of new users compared to that without the feature. To increase new user traffic, the ecommerce can –
    - Improve SEO and use the right keywords, especially those that will attract customers who care to know about the touch and feel of the product.
    - Introducing the Haptic feature for similar high-involvement products such as – leather goods, could translate into greater traffic and more new users.

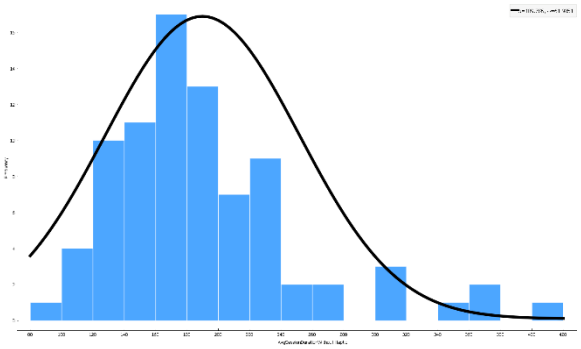
- **To achieve a greater average order value**: It is seen that average order value is lower for webpage with Haptic feature compared to that without the feature. To increase order value, the ecommerce can –
    - Create a minimum order amount for ‘free shipping’
    - Bundle Men’s Suit product with complementary ones such as – tie, socks, belts etc.
    - Upsell or recommend complementary products
    - Set up a customer loyalty program with cashback or discount incentives
- **Optimizing profits**: To optimize profit, the website should spend less than the mean value of average revenue per session with Haptic feature, which is BDT 605.57, **to acquire a new visit**. Moreover, the website should spend less than the mean value of average revenue per user with Haptic feature, which is BDT 1211.11, **to acquire a new customer**.

Appendix

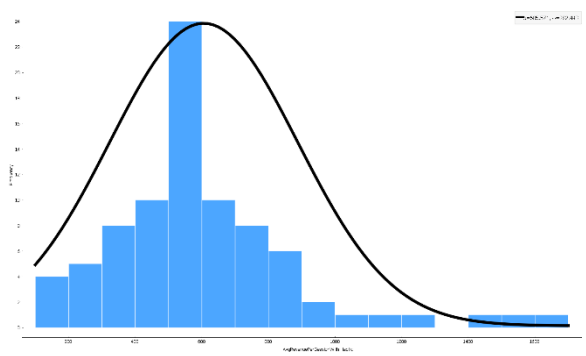
A: Normality to continuous metrics

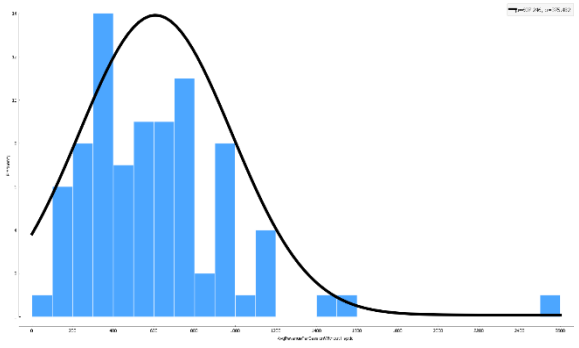
- Average Session Duration With Haptics & Average Session Duration Without Haptics



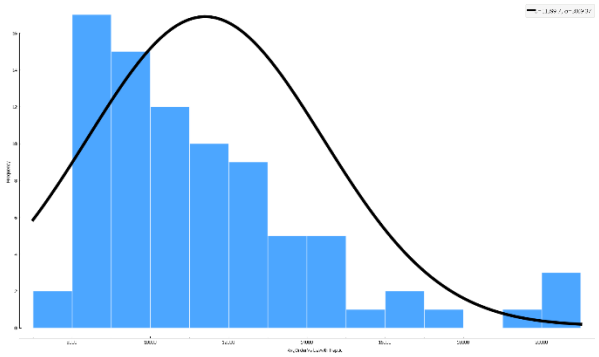


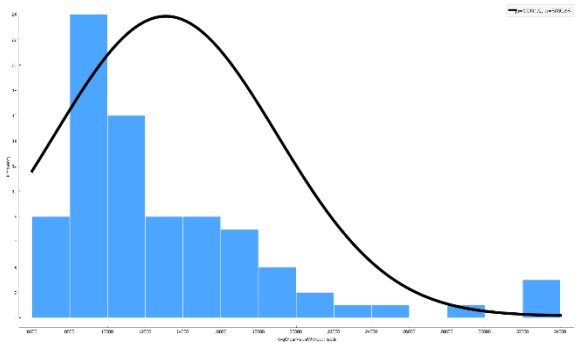
- Average Revenue/Session With Haptics & Average Revenue/Session Without Haptics



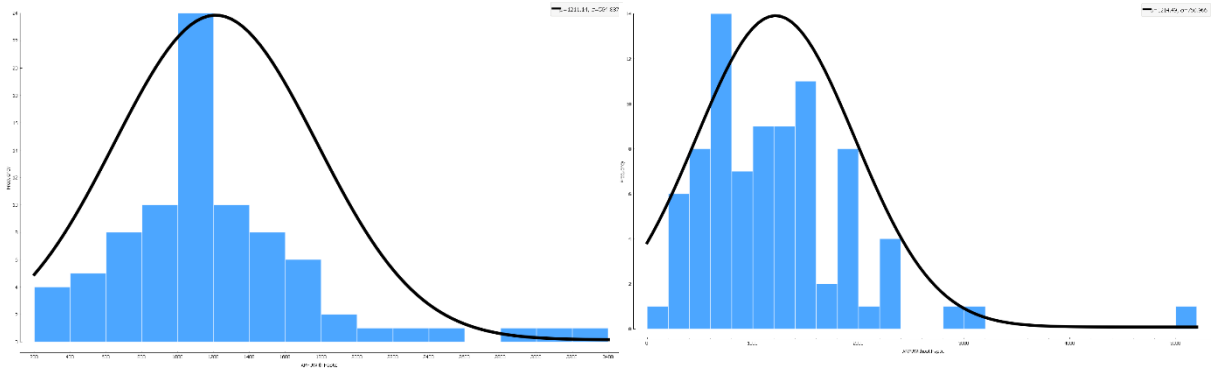


- Average Order Value With Haptics & Average Order Value Without Haptics





- Average Revenue Per User With Haptics & Average Revenue Per User Without Haptics



B: Process of Chi Square Hypothesis Testing

First, the null and alternate hypotheses are formed.

For the Chi Square test, the format followed is:

	With Haptic	**Without Haptic**
Outcome A	Sum of all Outcome As under this Segment	Sum of all Outcome As under this Segment
Outcome B	Sum of all Outcome Bs under this Segment	Sum of all Outcome Bs under this Segment

Then, I have calculated the Expected value for each combination using the formula:

$$\boldsymbol{Expected\ value = \frac{(row\ total\ *\ column\ total)}{Overall\ total}}$$

In the next step, I have calculated the difference between observed and expected values using the following formula = $\boldsymbol{\frac{(Observed\ value - Expected\ value)^2}{Expected\ value}}$

Finally, the degree of freedom is taken as 1 and the Chi Square value is calculated as the sum of all the calculated differences between observed and expected values. The P-value is determined using the Excel function CHISQ.DIST.RT.

C: Chi Square in Excel

Number of new users

	With Haptic	Without Haptic	
New users	11089	10509	**21598**
Not new users	5000	4510	**9510**
	16089	**15019**	31108

Expected values (E)

	With Haptic	Without Haptic
New users	11170.45	10427.55
Not new users	4918.55	4591.45

(O-E)^2/E

	With Haptic	Without Haptic
New users	0.59	0.64
Not new users	1.35	1.44

Chi-square value	4.02
df	1
p-value	0.044874573

Bounce rate

	With Haptic	Without Haptic	
Bounced sessions	10386	10128	**20514**
Not bounced session	21792	19910	**41702**
	32178	**30038**	62216

Expected values (E)

	With Haptic	Without Haptic
Bounced sessions	10609.80	9904.20
Not bounced session	21568.20	20133.80

(O-E)^2/E

	With Haptic	Without Haptic
Bounced sessions	4.72	5.06
Not bounced session	2.32	2.49

Chi-square value	14.59
df	1
p-value	0.00013375

Conversion rate

	With Haptic	Without Haptic	
Converted sessions	1678	1354	**3032**
Not converted sessions	30500	28684	**59184**
	32178	**30038**	62216

Expected values (E)

	With Haptic	Without Haptic
Converted sessions	1568.14	1463.86
Not converted sessions	30609.86	28574.14

(O-E)^2/E

	With Haptic	Without Haptic
Converted sessions	7.70	8.24
Not converted sessions	0.39	0.42

Chi-square value	16.76
df	1
p-value	0.0000425

D: T-Tests in Excel

Average session duration

	With Haptic	*Without Haptic*
Mean	216.5118081	189.7861451
Variance	5230.849956	3889.006895
Observations	83	83
Hypothesized Mean I	0	
df	161	
t Stat	2.549608643	
P(T<=t) one-tail	0.005859127	
t Critical one-tail	1.654373057	
P(T<=t) two-tail	**0.011718254**	
t Critical two-tail	1.974808092	

Average revenue per session

	With Haptic	*Without Haptic*
Mean	605.5714468	607.2455408
Variance	80746.71961	142706.2157
Observations	83	83
Hypothesized Mean I	0	
df	152	
t Stat	-0.032264567	
P(T<=t) one-tail	0.487151697	
t Critical one-tail	1.654940175	
P(T<=t) two-tail	**0.974303394**	
t Critical two-tail	1.975693928	

Average order value

	With Haptic	*Without Haptic*
Mean	11399.69699	13091.89133
Variance	9166776.438	35003791.92
Observations	83	83
Hypothesized Mean D	0	
df	122	
t Stat	-2.319651754	
P(T<=t) one-tail	0.011011009	
t Critical one-tail	1.657439499	
P(T<=t) two-tail	**0.022022017**	
t Critical two-tail	1.979599878	

Average revenue per user

	With Haptic	*Without Haptic*
Mean	1211.143494	1214.491062
Variance	322987.6489	570826.2724
Observations	83	83
Hypothesized Mean I	0	
df	152	
t Stat	-0.032258553	
P(T<=t) one-tail	0.487154091	
t Critical one-tail	1.654940175	
P(T<=t) two-tail	**0.974308182**	
t Critical two-tail	1.975693928	

Please find the Excel files in the link below:

https://drive.google.com/drive/folders/13PynL9M3jfu03IcQtt7hx-MnCmlTMaRp?usp=sharing