

An Analysis on Experimental Data

by Finley Clarke

FILE	11222_FINLEY_CLARKE_AN_ANALYSIS_ON_EXPERIMENTAL_DATA_192216_1076257949.DOCX (286.19K)		
TIME SUBMITTED	29-MAY-2020 04:05PM (UTC+0100)	WORD COUNT	1436
SUBMISSION ID	129559582	CHARACTER COUNT	6881



An Analysis on Experimental Data

FINLEY CLARKE
HND in Applied Computing – Data
Visualisation

TABLE OF CONTENTS/ABSTRACT

AN INTRODUCTION	2
THE ANALYSIS	3
IN CONCLUSION	7

An analysis on data collected over a 40-day period on a retail company's website that records data of sales on two versions of their website in order to determine the more popular version of the website.

Day	Old website	New website	Day	Old website	New website
1	15	15	21	76	95
2	18	17	22	72	91
3	26	24	23	68	89
4	39	37	24	63	85
5	61	59	25	60	82
6	84	83	26	54	80
7	96	99	27	51	76
8	104	112	28	48	74
9	110	126	29	45	72
10	113	130	30	41	68
11	111	128	31	39	62
12	108	125	32	37	61
13	103	119	33	35	57
14	99	116	34	31	55
15	98	113	35	76	103
16	94	110	36	29	52
17	89	106	37	26	48
18	86	104	38	22	44
19	81	101	39	19	43
20	80	99	40	16	41

Label

AN INTRODUCTION

The following report is a detailed analysis on a collection of data put together by a retail company who has recently been experimenting with their E-Commerce site's appearance and the impact it has on their customer-base and their spending habits on the website. Utilising a process known as "*A/B testing*", the company arranges for its web server to switch between the new and old design of the website for each visitor on a random basis. The company will record the results from these selected customer's actions (sales essentially) over a 40-day testing period. The company, in a bid to encourage and boost sales, is also offering a 15% discount on all orders made during this 40-day testing period and so they hope this should help in the data collecting process. After this testing period is wrapped up, the company wishes to determine whether they should keep the new design of the site or stick with the original site's design. In order to determine this, the data (shown above below the table of contents) needs to be properly analysed and studied to determine the patterns made by customers and their transactions. In this analysis, various methods will be used ranging from the T-Test, hypotheses, suitable testing, and the study of outliers in the data presented.



THE ANALYSIS

Label

An Overview of the Data provided –

The table of data for this analysis can be consulted in the Table of Contents/Abstract section. The results for the first 7 days (Days 1-7) of testing begin relatively low (starting sales beginning at 15) but had a gradual climb up to 96-99 sales by the end of that week. In addition to this, the results were quite closely matched and did not deviate from each other too much. The results begin to become much more interesting from day 8 onwards. The results for the old and new website gradually drift away from each other with the new website's share of the sales edging out ahead of the old website. Despite this, the sales for both websites gradually increase into the hundreds between days 7-13 before the old website's sales cap off. The sales on the new website stay in the hundreds until day 19. The sales of both sites begin to decline around the day 11-12 mark. From day 12 onwards, the sales would gradually decline although the new website would retain its lead ahead of the old website. The only other notable anomaly in the data shown is on day 35 where the new website's sales spike at 103. A substantial difference from the previous day which was at 55 sales.

2

3

The Hypothesis & Null Hypothesis –

It is hypothesised before the testing period that the new website yields better results compared to the old website. This can be told by the fact that the sales between the two sites over the 40-day period begins to deviate from each other with the new website's sales peaking higher than the old website's sales. It only takes one brief glance at the data in order to identify this, the data is quite clear.

The Null Hypothesis theorises that before the testing period that the new and old website hold no difference between each other and that they are identical in every way. There is no evidence or data provided anywhere in the testing period that would state otherwise.

Good

4

What is a T-Test –

The T-Test (sometimes referred to as the "Student's T-Test") is what is used to compare two individual means and can determine whether they are detached from each other or connected. In addition to this, the T-Test can also tell one how significant these differences between the two means are. It allows one to know whether these differences are occurrences by chance or deliberate. The T-Test is one of the most popular methods to employ in the study of statistics. The T-Test has had its roots traced back to William Sealy Gosset, an English statistician born in Canterbury, Kent in 1876, who is credited with developing the "t-statistic" and published said development under the pseudonym of "Student". Despite this, an earlier variant called "The T-Distribution" also appears in an earlier paper published in 1895 by Karl Pearson, an English mathematician born in 1857.

Good

Good Citation

Testing the Hypothesis & Null Hypothesis using T-Tests –

T-Test: Two-Sample Assuming Unequal Variances

	Old	New
Mean	64.30769231	81.69230769
Variance	952.0607287	935.7975709
Observations	39	39
Hypothesized Mean Difference	0	
df	76	
t Stat	-2.49869181	
P(T<=t) one-tail	0.007312224	
t Critical one-tail	1.292790268	
P(T<=t) two-tail	0.014624447	
t Critical two-tail	1.665151353	

Alpha @ 0.10

The first T-Test performed was at an Alpha level of 0.10. The results are not suitable, however. This is because this T-Test is not disproving the null-hypothesis which means that the old and new website cannot be proved to substantially different. This means that further T-Tests must be performed in order to prove this hypothesis.



T-Test: Two-Sample Assuming Unequal Variances -

	Old	New
Mean	63.075	80.025
Variance	988.4301282	1022.999359
Observations	40	40
Hypothesized Mean Difference	0	
df	78	
t Stat	-2.390271823	
P(T<=t) one-tail	0.009624153	
t Critical one-tail	1.664624645	
P(T<=t) two-tail	0.019248306	
t Critical two-tail	1.990847069	

Alpha @ 0.05

The second T-Test was more conclusive than the first and was performed this time at an Alpha level of 0.05. But these tests still do not fully prove that the old and new websites are truly distinct from one another. This means that yet another T-Test has to be performed in order to finally prove that the two websites are actually distinct and different from each other

T-Test: Two-Sample Assuming Unequal Variances

	Old	New
Mean	63.075	80.025
Variance	988.4301282	1022.999359
Observations	40	40
Hypothesized Mean Difference	0	
df	78	
t Stat	-2.390271823	
P(T<=t) one-tail	0.009624153	
t Critical one-tail	2.375110958	
P(T<=t) two-tail	0.019248306	
t Critical two-tail	2.640340015	

Alpha @ 0.01

The third T-Test was the final test performed and was adjusted to an Alpha level of 0.01. Again, like the previous two T-Tests performed, the old and new websites were not able to be distinguished from each other due to the “t-stat” and “t-critical two-tail” results (see above). Since the alpha cannot be decreased any further, it means that the hypothesis cannot be proven over the null hypothesis and thus is a failed data evaluation.

Outliers (Anomalies in the Data) –

The only notable anomaly in the data that can be brought up is the spike in results on day 35. The New Website’s sales bumped up from Day 34’s 55 sales up to a substantial 103 sales. There are various reasons or theories that could explain this spike. Perhaps a substantial group of individuals gained an interest in a rising trend that the site housed which caused them to rush to the site which by off chance, the “A/B testing” gave them the new version of the website which caused the spike. There could have easily been a miscalculation of the server’s behalf too (for whatever reason this could be can be anything from a server going down to a basic glitch in the system). Or it could really be as simple as more people simply visiting the website just by off-chance that particular day.



Just in-case this particular outlier may have caused a blip in the T-Test that prevented it from producing valid results, the T-Test is going to be performed once more with the outlier data removed to see what results it yields and whether there are any striking differences between the T-Tests with the outlier data included and a T-Test with the outlier data removed.

T-Test: Two-Sample Assuming Unequal Variances -

	Old	New
Mean	62.74358974	79.43589744
Variance	1009.932524	1035.673414
Observations	39	39
Hypothesized Mean Difference	0	
df	76	
t Stat	-2.304823627	
P(T<=t) one-tail	0.011953967	
t Critical one-tail	1.292790268	
P(T<=t) two-tail	0.023907933	
t Critical two-tail	1.665151353	

Analysing the data with the outlier data removed reveals that it made practically no difference to the results and does not miraculously prove the difference between the old and new website concretely. In order to stop these outlier data samples from popping up in future testing, keep an eye on whatever is recording the data during the testing process to ensure that no mistakes or mishaps are made which could potentially compromise the data and allow anomalies like this pop up and ruin an analysis like this one with false data.



IN CONCLUSION

In conclusion, the analysis failed in its bid to reject the proposed null hypothesis and was not able to prove the hypothesis was true. This was due to the data not being conclusive enough to draw any sort of reasonable conclusion after all of the T-tests were performed and analysed. Potential reasons for this are also unable to be determined too as even with the outliers data removed from the testing and that amended data being put through a T-test, it did not yield any tangible results either.

In order to perform a proper analysis that may yield better results next time, perhaps a more careful pre-preparation before the testing process should be applied. That means making sure that the server is ready to perform and is properly equipped to deal with the tasks required of it and doesn't get overwhelmed which can lead to crashes, glitches and other anomalies that can jeopardise the end data results which may have led to the data results we see in this analysis.

An Analysis on Experimental Data

GRADEMARK REPORT

FINAL GRADE

43/0

GENERAL COMMENTS

Instructor

Student FEEDBACK FORM

This section details the *extent* to which the assessment criteria are demonstrated by you, which in turn determines your mark. The marks available for each category of skill are shown. Lecturers will use the space provided to comment on the achievement of the task(s), including those areas in which you have performed well and areas that would benefit from development/improvement.

Generic Assessment Criteria	Marks available	Marks awarded
1. Engagement with Literature Skills		0
	0	0
2. Knowledge and Understanding Skills		0
	0	0
3. Cognitive and Intellectual Skills		0

Descriptions of the dataset. You have given a good written description of the data, but you should have included the data descriptive analysis from Excel. This should be included to draw some initial conclusions, mean, median, standard deviation. This should be included to show how you develop the hypothesis. You should also use some visualisation of the data to support your text.	10	3
Development of appropriate hypotheses for the analysis. You have included a good statement about the null hypothesis. However, you should also have included a statement about the alternative hypothesis. To show that you are able to present, evaluate and interpret qualitative and quantitative data, in order to develop lines of argument and make sound judgements in statistical analysis.	10	4

Selection (and justification) of appropriate techniques to perform the analysis. You have used appropriate testing methods to perform the analysis. You should mention that the tests become – more ‘stringent’, by altering the level of significance. (Alpha) You also do not explain why you chose a two tailed test over a one tailed test? Or why you chose the two sample with unequal variances!	15	6
Consideration of outliers and other factors. You have considered the outliers and how they may have occurred. To improve here you need to expand on how this could skew the overall results. For example – could removing the outliers give a significant difference between the means?	10	4
4. Practical Application Skills		0

Application of selected analysis techniques. You have performed all the selected tests (T-Test two tailed) accurately, and given consideration the different alpha levels. However, you have not explained what the alpha levels actually signify. You have also not indicated how you calculated the P value? The probability of any significance between the means.	20	11
Interpretation of the results of the analyses. You have not interpreted all your results correctly. You have not interpreted the lower critical value of the two tails. You have taken into account the various alpha levels and then compared the tests with the outlier data removed. Although you only performed one test with the outlier removed and did not indicate the alpha level used. (it was 0.1) Your test should be – IF $t_{stat} < -t_{critical}$ two-tail OR $t_{stat} > +t_{critical}$ two-tail – do not reject null hypothesis. That is the observed difference between the means is not convincing enough to reject.	15	6

Presentation of the results in both textual and, where appropriate, diagrammatical form. You have included some good description of the figures. To improve you should have included some other graphs to show the means etc – Line graphs and/or box and whisker for example.	10	4
5. Transferable Skills for Life and Professional Practice		0
Presentation of the paper in an appropriate professional format. Correct use of technical English. Suitable labelling and referencing strategies. You have demonstrated that you can communicate the results of your study/work adequately although not always accurately., You show some structured and coherent arguments. Overall the report was well structured. To improve, You should give more consideration to the labelling of visualisations and referencing them – Figure 1 for example	10	5
		43

PAGE 1

PAGE 2

QM

Label

Please label your figures and use this label as a reference in your text. e.g. Figure 1

PAGE 3



Comment 1

A good introduction

PAGE 4

QM

Label

Please label your figures and use this label as a reference in your text. e.g. Figure 1



Comment 2

You should use a visualisation here to support your text. A comparative line chart would illustrate these data nicely.



Comment 3

YOU should provide more of the descriptive statistics - One of the tools in the data analysis set. You should include a discussion of the mean, median, standard deviation of the two sets of data.

QM

Good

Good



Comment 4

You should also state that the alternative hypothesis is that there is a difference between the 2 sites.

QM

Good

Good

QM

Good Citation

Your reference supports your evaluation.

PAGE 5



Comment 5

You are not 'proving' anything.

the analysis is to decide whether to reject the null hypothesis or not.

PAGE 6



Comment 6

the spike also happened on the old website data!

PAGE 7



Comment 7

It doesn't matter - datasets are usually 'cleaned' before analysis takes place.

PAGE 8
