TOM3020 Fall 2020 Class Project

Data Analytics With Best Buy

Lab 1 Practice Hypothesis Testing

1) Shivani Patel (sbp@cpp.edu)
2) Mackenzie Shields (mgshields@cpp.edu)
3) Khalil Meade, (kmmeade@cpp.edu)
4) Michelle Scober, mscober@cpp.edu
5) Gabriel Chipukites (gtchipukites@cpp.edu)
6) Mary Gomez (marygomez@cpp.edu)
7) Omie (ojghalambor@cpp.edu)
8) Finn Eckhardt (feckhardt@cpp.edu)
9) Derek Hirose (dhirose@cpp.edu)

Example: On Black Friday, Best Buy wanted to analyze their data on laptop sales. The data below provides information given during that day. We have a 95% Confidence Interval.

Mean	Z- Test value	Value from Z-Table	p-Value	Double p-value
4	1.770978158	0.9616	0.0384	0.0768

Total Sales (Thousands $)	Transaction #
4.0	1
3.5	2
3.5	3
4.5	4
4.5	5
4.0	6
4.0	7
5.5	8
5.0	9
4.5	10
5.0	11
4.0	12
5.0	13
5.5	14
5.0	15
5.0	16
3.5	17
4.0	18
4.0	19
4.0	20
3.5	21
3.5	22
4.0	23
3.0	24

Black Friday Sales

Mean	4.25
Standard Error	0.141164926
Median	4
Mode	4
Standard Deviation	0.691564075
Sample Variance	0.47826087
Kurtosis	-0.835314479
Skewness	0.268883853
Range	2.5
Minimum	3
Maximum	5.5
Sum	102
Count	24
Confidence Level(95.0%	0.292021898

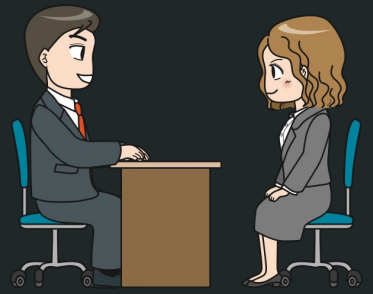
Null Hypothesis: Mean of laptop sales <4.25

Alternative: Mean of laptop sales = 4.25

Conclusion: Best Buy, does not reject the null hypothesis based on the P - value.

Lab 2 Practice Chi Square Lab

1. Xiomara Rodriguez (xerodriguez@cpp.edu)
2. Belinda Gallardo
3. Arturo Sandoval (Arturos1@cpp.edu)
4. Breeana Berumen (Bhberumen@cpp.edu)
5. Alex Thomas (atthomas@cpp.edu)
6. Ethan Nguyen (eknguyen@cpp.edu)
7. Emmanuel Rodriguez (emmanuelr1@cpp.edu)
8. Celenia Perez celeniaperez@cpp.edu
9. Yesenia chavez ychavez@cpp.edu



Example: Best Buy wants to determine if there was a significant difference in the number of new hires each week during the 10 weeks prior to Black Friday. The confidence interval is 95%.

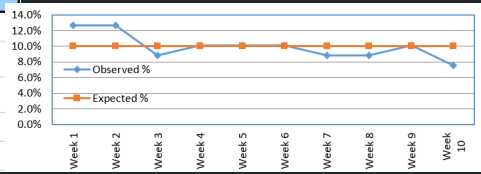
Null Hypothesis: There is not a significant difference in the number of new hires each week

Alternative Hypothesis: There is a significant difference in the number of new hires each week

One-Way Chi-Square Test (Uniform Distribution/Multinomial Distribution)

Uniform	Weekly Hires (Thousands)										
	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 8	Week 9	Week 10	Total
Observed	5.0	5.0	3.5	4.0	4.0	4.0	3.5	3.5	4.0	3.0	39.5
Observed %	12.66%	12.66%	8.86%	10.13%	10.13%	10.13%	8.86%	8.86%	10.13%	7.59%	100.00%
Expected	3.95	3.95	3.95	3.95	3.95	3.95	3.95	3.95	3.95	3.95	39.5
Expected %	10.00%	10.00%	10.00%	10.00%	10.00%	10.00%	10.00%	10.00%	10.00%	10.00%	100.00%
χ^2 - Value	0.279113924	0.279113924	0.051265823	0.000632911	0.000632911	0.000632911	0.051265823	0.051265823	0.000632911	0.228481013	0.943037975
											Total

Overall χ^2 - Value	0.94303797		16.9190	← χ^2 - Critical Value	α =	0.05
P-Value	0.99956		9	Degrees of Freedom	P-Value ≤ α → Significant	
Significant?	No				Expected Value < 5 → Violation	



Conclusion: Do not reject the null hypothesis based on p-value

Lab 3 Practice Simple Regression Lab

Example: Best Buy ran TikTok ads for their computers for 9 consecutive months during March through November.

Television Ads	Computers Sold
5	20
4	15
6	20
2	18
6	25
5	23
4	19
2	11
3	21

Null hypothesis: There was no significant change in computer sales after the ad run.

Alternative hypothesis: There was significant change in computer sales after the ad run.

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.681161638
R Square	0.463981176
Adjusted R Square	0.387407059
Standard Error	3.261179866
Observations	9

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	64.44183007	64.44183007	6.059242871	0.043366497
Residual	7	74.44705882	10.63529412		
Total	8	138.8888889			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	11.51764706	3.270757507	3.521400481	0.009709356	3.783534536	19.25175958	3.783534536	19.25175958
TikTok Ads (Hundreds)	1.847058824	0.750363234	2.461552939	0.043366497	0.072731723	3.621385924	0.072731723	3.621385924

1) $H_o: \beta_1 = 0$
2) $H_1: \beta_1 \neq 0$
3) P-value = 0.043366497
4) P-Value 0.043366497 < 0.05
5) Reject H_o? There is a significant Linear Relationship between On-Line Ads and Computers Sold.
6) # Computers Sold = 11.51764706 + 1.847058824(On-Line Ads)

1) $H_o: \beta_1 = 0$
2) $H_1: \beta_1 \neq 0$
3) Critical Values: ± 2.365 from T-Table df=7 (2 tail)
4) T-Stat = 2.461552939 > 2.365
5) Reject H_o? There is a significant Linear Relationship between On-Line Ads and Computers Sold.
6) # Computers Sold = 11.51764706 + 1.847058824*(On-Line Ads)

1.Chad Norman (cdnorman@cpp.edu)
2.Mandy Phan (mandyphan@cpp.edu)
3. Isaiah McClenton (icmcclenton@cpp.edu)
4. Kevin Su (ksu@cpp.edu)
5. Nina Pangilinan (nlpangilinan@cpp.edu)
6. Leslie Osayamwen (loosayamwen@cpp.edu)
7. William Sim (wsim@cpp.edu)
8. Madeline Metis
9. (memetis@cpp.edu)

Lab 4 Multiple Regression Lab

1. Jensen Gomez (Jvgomez@cpp.edu)
2. Madeline Metiz (memetis@cpp.edu)
3. George Johnson (grjohnson@cpp.edu)
4. Marc Schreiber (maschreiber@cpp.edu)
5. Camrin Ware chware@cpp.edu

Example: During Black Friday, Best Buy launched multiple online Ads campaigns to compete with Amazon's dominance during the COVID 19 Pandemic. Amazon is beating most traditional brick and mortar stores. This Multiple Linear Regression model predicts the number of computers sold at Best Buy **(Y)** based on # of Amazon Ads **(X1)** and and # of Best Buy Ads **(X2)**.

Best Buy attempts to compete with online retailer Amazon by outclassing them with a higher # of Ads than Amazon.

Model at 86% effectiveness shows that a higher number of Best Buy Ads generates more Computer Sales.

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.930990256
R Square	0.866742857
Adjusted R Squa	0.82232381
Standard Error	1.756318902
Observations	9

Amazon Ads	Best Buy Ads	Best Buy Computers Sold
5	6	20
4	15	15
6	6	20
2	7	18
6	4	25
5	5	23
4	8	19
2	12	11
5	6	21

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	2	120.3809524	60.19048	19.512864	0.002366309
Residual	6	18.50793651	3.084656		
Total	8	138.8888889			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	19.61904762	3.39497304	5.778852	0.0011735	11.31184785	27.92625	11.3118479	27.9262474
X Variable 1	-0.73015873	0.20992026	-3.47827	0.0131703	-1.243815102	-0.2165	-1.2438151	-0.21650236
X Variable 2	1.174603175	0.49971017	2.350569	0.0570141	-0.048143564	2.39735	-0.04814356	2.39734991

1) H_o: $\beta_1 = \beta_2 = 0$
2) H_1: At Least one of the β's $\neq 0$
3) P-value = 0.01
4) P-Value < 0.05
5) Reject H_o: There is a significant Linear Relationship between (Amazon Ads & Best Buy Ads) and Computers Sold.
6) # Computers Sold = 19.62 + -0.73*(Amazon Ads) – 1.17*(Best Buy Ads)

1) H_o: $\beta_1 = \beta_2 = 0$
2) H_1: At Least one of the β's $\neq 0$
3) F-Critical Value = 19.51
4) F-Test Stat = 19.51 > 5.14
5) Reject H_o: There is a significant Linear Relationship between (Amazon Ads & Best Buy Ads) and Computers Sold.
6) # Computers Sold = 19.62 + -0.73*(Amazon Ads) – 1.17*(BestBuy Ads)

Lab 5 Practice ANOVA Lab

-Manuel Arias (maarias@cpp.edu)
-Jakob M (jtmacedo@cpp.edu)
-Jasmine Rivas (jarivas@cpp.edu)
-Sharar Kamal (skamal@cpp.edu)
-Elizabeth Polanco (epolanco@cpp.edu)
-Noe Ramirez (noeramirez@cpp.edu)

Example: Best Buy sold their TVs on Black Friday at a discount. The following is a dataset for the numbers of TVs sold based on TV ads, mobiles ads, and magazine ads. With an alpha of 0.05.

TV Ads	Mobile Ads	Magazine Ads
20	42	2
14	33	5
9	15	3
12	9	7
15	24	4

We reject the null hypothesis because our P-value is less than 0.05. (0.006 < 0.05)

Along with our F-critical value being less than our F-table value.

Anova: Single Factor

SUMMARY

Groups	*Count*	*Sum*	*Average*	*Variance*
TV Ads	5	70	14	16.5
Mobile Ads	5	123	24.6	177.3
Magazine Ads	5	21	4.2	3.7

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	1040.93333	2	520.466667	7.90582278	0.00645249	3.88529383
Within Groups	790	12	65.8333333			
Total	1830.93333	14				

Contributors:

1. Adrian Carrion
2. Alex T. Thomas
3. Arturo Sandoval
4. Belinda Gallardo
5. Bilal Elgabalawy
6. Breeana Berumen
7. Celenia Perez
8. Gabriel Chipukites (Chip)
9. Dgee_11
10. Emmanuel Rodriguez
11. Ethan Nguyen
12. Finn Eckhardt
13. Isaiah McClenton
14. Jenson Gomez
15. Mackenzie Shields
16. Kevin Su
17. Khalil Meade
18. Leslie Osayamwen
19. Mandy Phan
20. Marc Schreiber
21. Mary Gomez
22. Nina Pangilinan
23. Noe Ramirez
24. Omid Ghalambor
25. Shivani Patel
26. Talor Hern
27. William Sim
28. Xiomara Rodriguez
29. Chad Norman
30 Yesenia Chavez
31. Noe Ramirez
32. Manuel Arias
33. Michelle Scober
34.
35.
36.