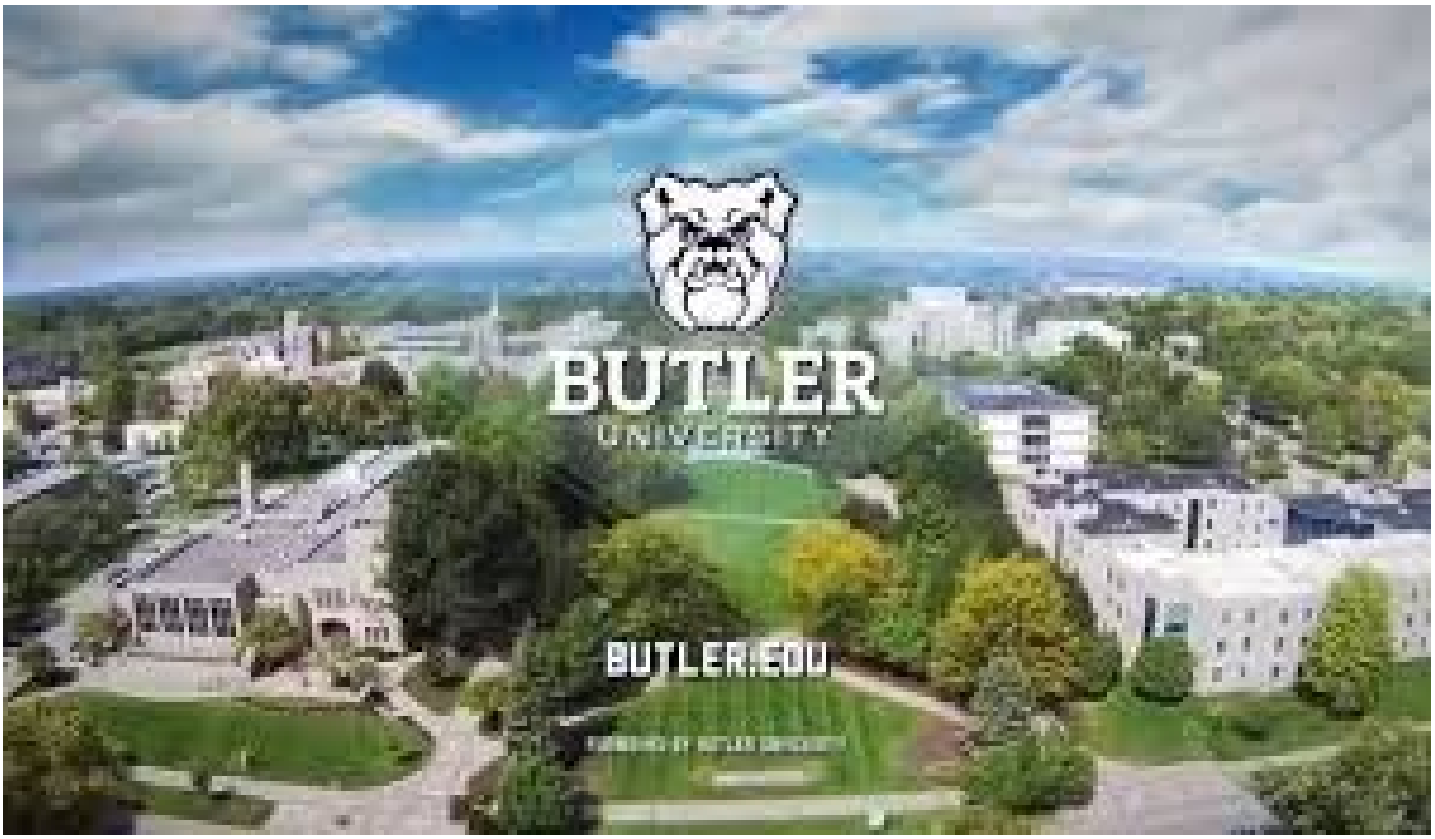


Statistical Theory: Group Project 2
Opinion of Diversity Efforts Over Time

Lauren Tibbets, Abigale Wynn, Rosa Florence

Table of Contents

Chapter 1: Introduction

Question: At the 5% significance level, do the data provide evidence to say that there is an association between year in school at Butler University and opinion on how Butler is doing with their efforts towards “inclusion and diversity”?

Background: On March 28, 2017, the Butler Collegian featured a Letter to the Editor titled, “*Black Student Union addresses lack of black staff members”.* The article outlined key complaints towards diversity efforts on campus by a club centered around racial diversity. This begs the question if this is an isolated case or if other groups on campus share similar views of the university’s efforts towards diversity and inclusion. Butler works very hard to promote a positive and inclusive community, but perhaps over time, its practices are not keeping up with turbulent political factors. We decided to test whether or not there is an association between a student’s year in school and opinion on Butler’s inclusion efforts because we wanted to highlight the difference in opinions from students who may have seen varying inclusion efforts throughout multiple years at Butler versus students who have not been at Butler as long. We expected to see a difference in opinions between classes. This difference means that the “qualitative variables”, year in school at Butler and opinion on how Butler is doing with their efforts towards “inclusion and diversity”, would depend on one another.

We collected data via anonymous survey sent out to various listservs on campus chosen specifically to promote diversity among class while minimizing any skewing opinions from outside variables. The survey was sent to members of greek life, honors program, math department, computer science department, athletics, and each class Facebook page. The survey had four questions; the first two questions were required, and the last two questions were presented as optional and for further research.

Question one: What is your class/year in school?

a. First year
b. Sophomore
c. Junior
d. Senior
e. Graduate

Question two: How strongly do you believe that Butler is doing a good job with their efforts towards “inclusion and diversity”?

a. Strongly agree
b. Agree
c. Neutral
d. Disagree
e. Strongly disagree

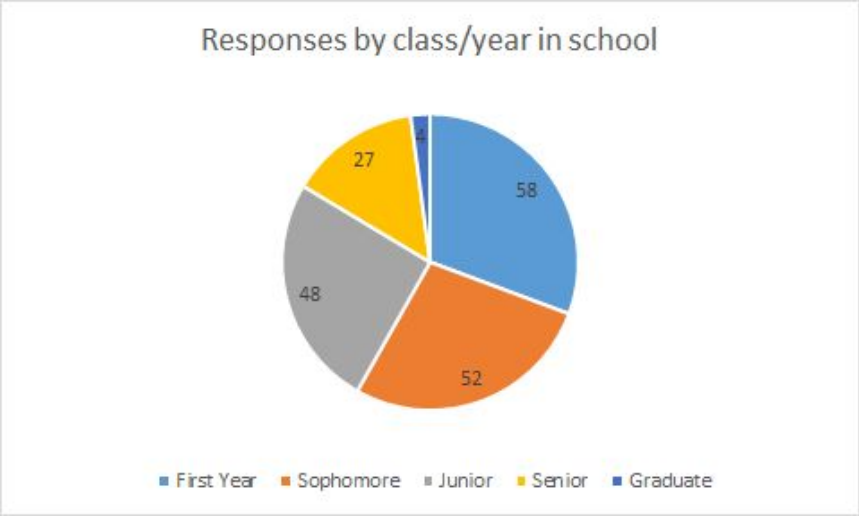
Question three: What is your ethnicity/race? *
Question four: What gender/sex do you identify with?*

> **Because the third and fourth questions were optional, we provided blanks to type an answer for the responses rather than a multiple choice option.*

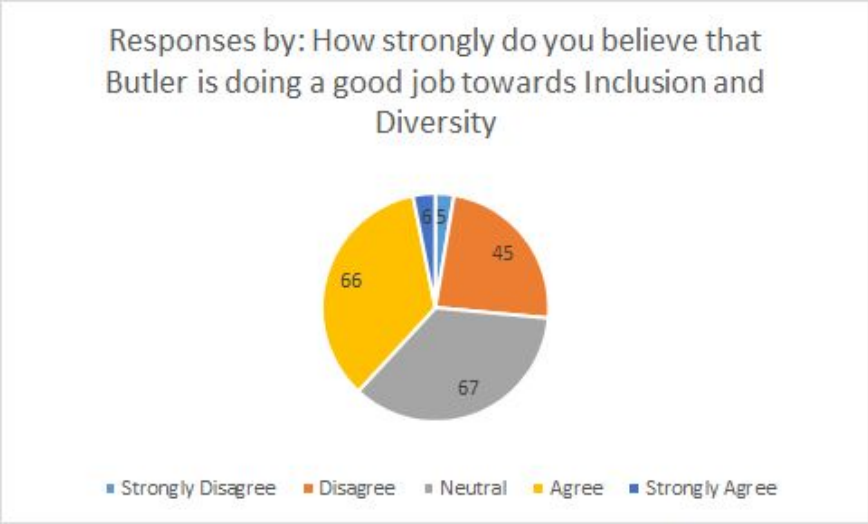
The Chi-square test of independence is the appropriate test for this question because it is used to determine if there is a significant relationship between two variable or if they are independent. In our case, we have two categorical/qualitative variables, meaning that neither “year in school” nor “opinion on Butler’s efforts towards ‘inclusion and diversity’” can be described by an average. Rather, these variables are described by bins and numbers of people who fit into those bins. We are trying to determine if these variables are related.

Chapter 2: Data Descriptions

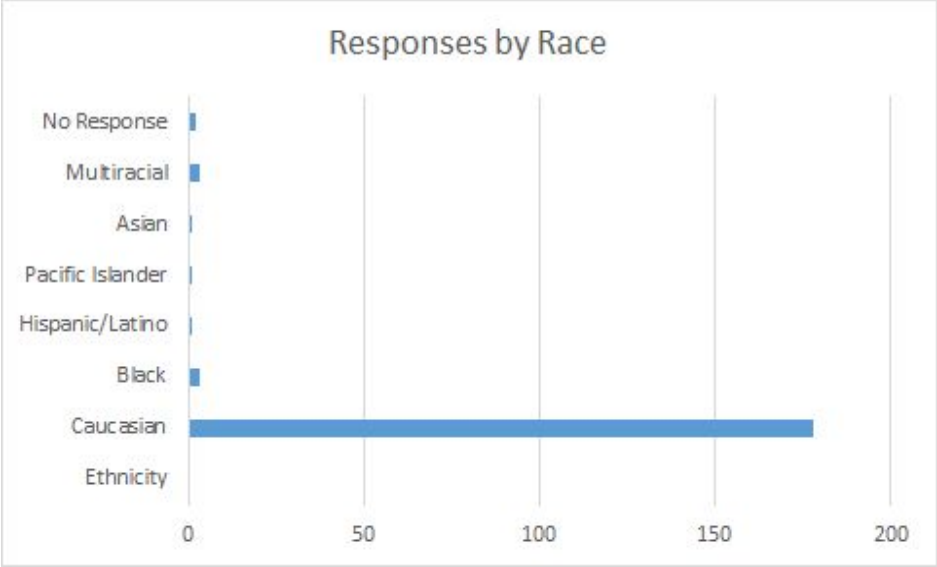
The charts below display the number of responses for each variable tested, including the optional questions. The bar graph and pie chart are appropriate for this data set because we are describing categorical variables.

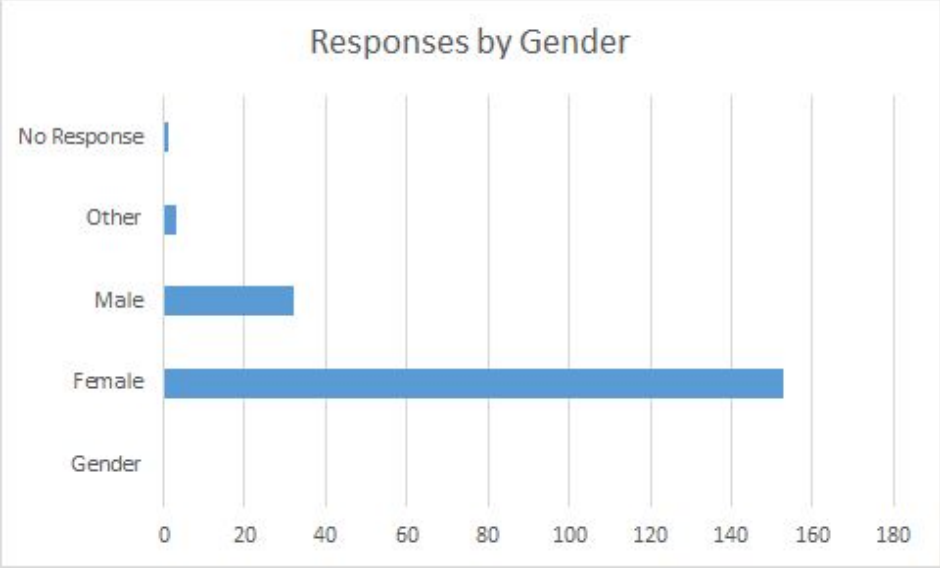


The first pie chart shows the amount of responses for each categorical variable of class/year in school. This was one of the required questions on the survey. For the five “class” variables, the following amount of responses was collected: 58 first year, 52 sophomore, 48 junior, 27 senior, and 4 responses. The data show more lowerclassmen responses than upperclassmen responses; this discrepancy should be kept in mind when performing and interpreting the Chi-square test of independence. Seniors responded the least, so we do not have as many opinions coming from those who have observed the university’s efforts over a longer time span.



The second pie chart shows the amount of responses for each opinion on how well Butler is doing at achieving its goals towards diversity and inclusion. The following amount of responses were collected: 5 strongly disagree, 45 disagree, 67 neutral, 66 agree, and 6 agree. This data is interesting because it is very heavy in the middle responses of agree, neutral, and disagree, and very few people responded with either strongly agree or strongly disagree. This might signal that the people who responded to our survey may not be heavily influenced by the university's efforts towards increased diversity, or the university's diversity efforts have not been extremely noticeable to the student population.





Chapter 3: Hypothesis Test

We will be using the Chi-square test of independence because it is used to determine if there is a significant relationship between two qualitative variables or if they are independent. The data involves two qualitative variables, meaning both opinion and year in school are not described by averages. Rather, the variables are measured by fitting into certain “bins”.

The following table shows the observed values from the survey.

Observed						
Belief/Class Year	First Year	Sophomore	Junior	Senior	Graduate	**Totals**
Strongly Disagree	1	2	2	0	0	5
Disagree	14	15	10	6	0	45
Neutral	17	19	19	9	3	67
Agree	23	14	17	12	0	66
Strongly Agree	3	2	0	0	1	6
Totals	58	52	48	27	4	189

Based on the above observed values the expected values were calculated using the expected value formula: [(Column Total x Row Total)/Sample Size], and are displayed below:

Expected						
Belief/Class Year	First Year	Sophomore	Junior	Senior	Graduate	**Totals**
Strongly Disagree	1.534391534	1.375661376	1.26984127	0.714285714	0.105820106	5
Disagree	13.80952381	12.38095238	11.42857143	6.428571429	0.952380952	45
Neutral	20.56084656	18.43386243	17.01587302	9.571428571	1.417989418	67
Agree	20.25396825	18.15873016	16.76190476	9.428571429	1.396825397	66
Strongly Agree	1.841269841	1.650793651	1.523809524	0.857142857	0.126984127	6
Totals	58	52	48	27	4	189

Hypotheses:

H_0: a Butler student’s class is independent from the student’s opinion on Butler’s inclusion efforts

H_a: a Butler student’s class is not independent from the student’s opinion on Butler’s inclusion efforts

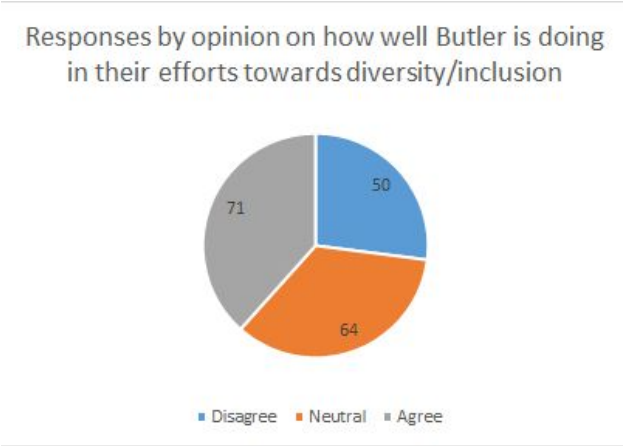
Assumptions:

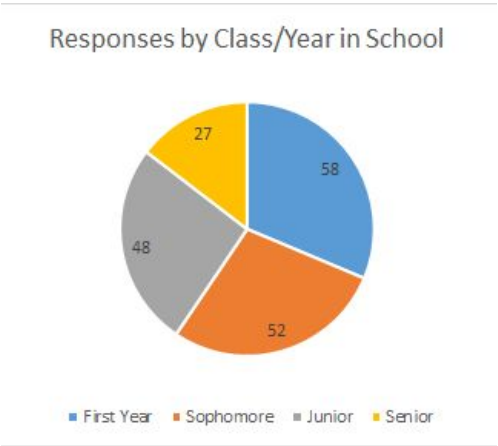
→ Due to the nature of the survey, the data represents a simple random sample

→ All expected frequencies must be greater than one [*not met*]

→ At most 20% of the expected frequencies can be less than five [*not met*]

Multiple assumptions for the Chi-square test of independence are not met: not all expected values are greater than one and more than 20% of the expected values are less than five. Because the sample data as a whole does not fit all assumptions, we concluded it would be best to finish the testing using a data selection that **does** fit those assumptions. For the rest of the test, only class years from first year to senior will be considered. Any options under “strongly disagree” will be combined with “disagree” and those under “strongly agree” will be combined with “agree”. The following pie charts and tables represent the data that will be considered as our observed values from now on.





Expected					
Belief/Class Year	First Year	Sophomore	Junior	Senior	**Totals**
Disagree	15.67567568	14.05405405	12.97297297	7.297297297	50
Neutral	20.06486486	17.98918919	16.60540541	9.340540541	64
Agree	22.25945946	19.95675676	18.42162162	10.36216216	71
Totals	58	52	48	27	185

Observed					
Belief/Class Year	First Year	Sophomore	Junior	Senior	**Totals**
Disagree	15	17	12	6	50
Neutral	17	19	19	9	64
Agree	26	16	17	12	71
Totals	58	52	48	27	185

((Oi-Ei)^2)/Ei)					
Belief/Class Year	First Year	Sophomore	Junior	Senior	**Totals**
Disagree	0.029123952	0.617515593	0.072972973	0.230630631	0.950243148
Neutral	0.468151503	0.056797362	0.34531426	0.012415541	0.882678665
Agree	0.628570678	0.784492401	0.109708476	0.258875777	1.781647333
Totals	1.125846133	1.458805356	0.527995709	0.501921948	3.614569146

The new data selection meets all of the assumptions, so we will continue with the Chi-square test of independence. The Chi-square test statistic is found by adding [(Oi-Ei)^2/Ei] for each point. The test statistic adds to 3.6146 after adding all of the value in the table above. The degrees of freedom (which can be found by the following equation: [number or rows-1][number of columns-1]) = 6 in this test.

Both the test statistic and degrees of freedom will be used to find the p-value on a calculator. The chi-square value is .7287 according to a calculator. Because the p-value is larger than the alpha value of .05, we fail to reject the null hypothesis and conclude that a Butler student's year in school and their opinion on the effectiveness of Butler's diversity and inclusion efforts are independent.

Chapter 4: Conclusion

At the beginning of the project, we believed a student's opinion of Butler's efforts towards increased inclusion and diversity would be dependent upon their year in school. We expected there would be a difference in opinions throughout the different classes showing that university efforts to improve diversity and inclusion are perceived by students to have increased or decreased in different years. This dependence would also have concluded whether or not Butler's efforts have been noticed by students and aided in creating a more inclusive campus.

However, at the 5% significance level, the data provide evidence to say that a student's opinion of Butler's inclusion and diversity efforts is independent of the student's year in school. This independence may show that Butler's increase in efforts towards creating a more inclusive campus has not been reflected in students' perception of diversity and inclusion on campus.

Although the null hypothesis was true and had a very high probability of occurring due to chance (>72%), there are underlying factors that may have skewed this result. In regards to what inspired the question, minority groups such as the Black Student Union were the groups that felt victimized. However, due to the random nature of our surveying, our data did not include many of these minority groups. In fact, as the chart below shows, of 175 responses included in our test, only 7 total responses identified with any racial group other than Caucasian. And the same can be said for gender in that the responses were also mainly from females. Although Butler has more female students than males, the distribution responses to our survey was much more skewed than the 60-40 female-male population distribution on campus. Because our data did not truly represent all opinions on campus, we cannot make an accurate statement about overall student opinion of Butler's efforts towards diversity and inclusion based on the dependence between class year and opinion of Butler's inclusion efforts.

Ethnicity		**Gender**	
Caucasian	166	Female	143
Black	1	Male	29
Hispanic/Latino	1	Other	2
Pacific Islander	1	No Response	1
Asian	1	**Total**	**175**
Multiracial	3		
No Response	2		
Total	**175**		