**Studying Student Evaluations**

At this point in the semester you have enough exposure to R to work on a laboratory project with only minimal exposure to commands and techniques. The goal of today’s laboratory is to have you study Lafayette College student evaluations, and since we cannot gain access to the database of all the evaluations, we will have to be content with examining a sample.

**Warm-Up:** In this model for what you will do, I will describe how I explored two questions:

1. Is there a correlation between expected grades (Question 9 on the student evaluation form) and general student satisfaction (Question 8)?
2. Do students believe they learn more (Question 4) when the professor is male (the instructor’s sex is not on the form … but at Lafayette we know almost all of the professors)?

The summary information from student evaluations can be found at:

<https://fac-eval.lafayette.edu/>

which can only be accessed using your Lafayette credentials. Since I do not have access to the full database of student evaluations, I had to work with a sample, and decided to look at a random sample of ten classes all from Spring 2016. In that term there are 536 courses whose evaluations are available. I used the following function in R to get a random selection of 10 numbers from 1 through 536,

>sample.int(536,10,replace=FALSE)

which gave me a set of ten randomly chosen numbers without replacement. I then looked at those evaluations to create the following table.

|  |  |  |  |
| --- | --- | --- | --- |
| **Sex** | **Learn** | **Satis** | **Grade** |
| F | 4.3 | 4.2 | 3.7 |
| M | 3.8 | 2.7 | 3.2 |
| M | 4.2 | 4.0 | 3.3 |
| F | 4.9 | 4.5 | 3.4 |
| F | 4.8 | 3.3 | 3.3 |
| M | 4.7 | 4.1 | 3.6 |
| M | 4.9 | 4.8 | 3.8 |
| M | 4.5 | 4.2 | 3.6 |
| M | 3.5 | 3.0 | 3.8 |
| M | 4.9 | 3.0 | 3.7 |

To examine the first question—expected grades and student satisfaction—I created a scatterplot of these two quantitative variables and found a positive correlation of r=0.46. (We learned these techniques in Lab 3.)



There may be a correlation between expected grades and student satisfaction. This is what many national studies have indicated, and there are many possible explanations for the correlation. For example, if students are doing well in a class it is reasonable to believe they are generally satisfied with the course. Having only ten data points, though, doesn’t give me much confidence that there actually is a positive correlation, or that the correlation’s strength is indeed as high as r=0.46.

To examine if there is gender bias in the amount students believe they learn from instructors, I looked at the two most commonly used notions of “center”, disaggregated by sex. (We learned how to do such things in Lab 2.)

|  |  |  |  |
| --- | --- | --- | --- |
|  | n | mean | median |
| Female | 3 | 4.67 | 4.8 |
| Male | 7 | 4.36 | 4.5 |

These statistics indicate that Lafayette students might have some gender bias against male faculty members. In much larger, national studies, it is usually the case that the bias goes the other direction. That, plus the fact that the sample for men is very small and for women it is tiny, gives me very little confidence in this result.

In both cases my sample of ten observations is too small to be convinced of the outcome. We have not yet studied the notion of confidence for either of these settings, but our brief exposure to hypothesis testing and confidence intervals should be sufficient for everyone to believe that ten observations is too few. On the other hand, creating a larger sample by hand is rather time consuming. That’s why you’ll work collaboratively during the laboratory session to collect a large enough sample that you can have some confidence in your findings.

PS: If my sample size was larger, then using the grading rubric on the next page, this discussion would earn 7/10 points.

**Laboratory Assignment!**

For this laboratory you need to analyze a sample of Lafayette College student evaluations. You will be working with the people at your pod to collect the sample. Because of this, everyone at your pod will need to agree on the hypotheses you want to test. Everyone needs to examine if there is a correlation between expected grades and student satisfaction. The second question you address is completely up to you, except that it must be based on a categorical variable. As examples, you could consider:

* Does instructor’s sex impact student satisfaction (#8)?
* Does the field (maybe STEM vs non-STEM?) impact “amount learned” (#4)?
* Is student satisfaction higher for First Year Seminars than other 100 level courses?
* Do students believe instructors are more effective (#7) in low enrollment courses (maybe under 20?) than larger classes?

Once your pod has agreed on the second question, you then need to decide on a sampling strategy. You might want to create a Google sheet to collect your data, as that will allow everyone in the pod to contribute data and access the results. I encourage you to spend the bulk of your laboratory time thinking about your strategy and collecting your sample.

Next Friday (March 8th) you will need to submit an essay that describes your study, your analysis of the results, what conclusions you draw, and suggestions for how you might strengthen the evidence for your conclusions. I expect it should require two pages; you can’t hand in anything longer than three pages. While you will work with your pod-mates on the data collection, everyone needs to write their own essay.

**Grading Rubric**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **1 point** | **2 points** | **3 points** |
| **Well written?** | No major grammatical or structural issues. | Nicely presented, well organized, and clearly written. | Outstanding presentation of your ideas. |
| **Sampling** | Your sample has at least 50 observations. | You used a valid sampling strategy. | **Bonus:** Your pod came up with the largest sample in the class.[[1]](#footnote-1) |
| **Statistics** | Your statistical summary is correct. | Clear and succinct presentation and discussion of stats. | Insightful design of your study. |
| **List of pod-mates** | You should include this! |  |  |
| **Next steps** | You should include two ways you could improve your study and strengthen your results |  |  |

1. If there is a tie between two or more groups for “largest sample” then no bonus points will be given. [↑](#footnote-ref-1)