# Lab 2 Instructions

**BSTA 513/613**

Nicky Wakim

## Directions

You can download the `.qmd` file for this lab here.

The above link will take you to your **editing** file. **Please do not remove anything from this editing file!!** You will only add your code and work to this file.

### Purpose

The purpose of this lab is to explore our data further, set up the unadjusted odds ratio, and create code to later help us present our final model.

### Grading

**This lab is graded out of 12 points.** The TAs will go through and grade your lab. They will make sure each section is complete and will follow the rubric below. I have instructed them that completion and clear effort is all that is needed to receive 100%. Nicky will go through the labs to give you feedback.

### Rubric

| | 4 points | 3 points | 2 points | 1 point | 0 points |
|---|---|---|---|---|---|
| Formatting | Lab submitted on Sakai with `.html` file. Answers are written in complete sentences with no major grammatical nor spelling errors. With little editing, the answer can be incorporated into the project report. | Lab submitted on Sakai with `.html` file. Answers are written in complete sentences with grammatical or spelling errors. With editing, the answer can be incorporated into the project report. | Lab submitted on Sakai with `.html` file. Answers are written in complete sentences with major grammatical or spelling errors. With major editing, the answer can be incorporated into the project report. | Lab submitted on Sakai with `.html` file. Answers are bulletted or do not use complete sentences. | Lab *not* submitted on Sakai with `.html` file. |
| Code/Work | All tasks are directly followed or answered. This includes all the needed code, in code chunks, with the requested output. | All tasks are directly followed or answered. This includes all the needed code, in code chunks, with the requested output. In a few tasks, the code syntax or output is not quite right. | Most tasks are directly followed or answered. This includes all the needed code, in code chunks, with the requested output. | Some tasks are directly followed or answered.This includes all the needed code, in code chunks, with the requested output. In a few tasks, the code syntax or output is not quite right. | More than a quarter of the tasks are not completed properly. |

| | 4 points | 3 points | 2 points | 1 point | 0 points |
|---|---|---|---|---|---|
| Reasoning* | Answers demonstrate understanding of research context and investigation of the data. Answers are thoughtful and can be easily integrated into the final report. | Answers demonstrate understanding of research context and investigation of the data. Answers are thoughtful, but lack the clarity needed to easily integrate into the final report. | Answers demonstrate some understanding of research context and investigation of the data. Answers are fairly thoughtful, but lack connection to the research. | Answers demonstrate some understanding of research context and investigation of the data. Answers seem rushed and with minimal thought. | Answers lack understanding of research context and investigation of the data. Answers seem rushed and without thought. |

*Applies to questions with reasoning

## Lab activities

> **Note**
>
> I have left it up to you to load the needed packages for this lab.

### Restate research question

> **Task**
>
> Please restate your research question below using the provided format (1 sentence). You can change the wording if you'd like, but please make sure it is still clear. It's repetitive, but it helps me contextualize my feedback as I look through your lab.

In this study, we will investigate the association between food insecurity and _____.

### Make sure variables are coded correctly

Use `class()` to determine the class of each of the 11 variables you selected from Lab 1 (including the outcome). A tidyverse equivalent to the `apply()` function that we learned last quarter is `map()`. Please take a look at the description of the `map()` function.

Make sure the class that `R` recognizes is the class that you expect the variable to be. Categorical variables should be factors amd numeric variables should be numeric. **It is very important that your outcome, food insecurity, is a factor with the reference level set to "No."** For example, if I am using age, but the class is character, I will need to convert age to a numeric variable. If I have a categorical covariate that is recognized as a character, I should convert it to a factor with a specific reference level.

```
df_name %>% map(class)
```

**!** Task

- Use `class()` to determine the class of each of the 11 variables you selected from Lab 1 (including the outcome).
- Change the variable type to the appropriate type.

### Consider potential confounders and effect modifiers

For each of the 10 predictor variables, fill out the below table. This does not need to be extensive reasoning. If you would like to present this information in another way, you may.

| Variable name | Confounder, Effect modifier, or nothing? | Reasoning (1-2 sentences) |
|---|---|---|

### Bivariate exploratory data analysis

Use `ggpairs()` (introduced in BSTA 512 Lesson 13) to quickly look at the relationship between variables.

> **!** Task
>
> Using `ggplot` or tables, visualize your variables. Get a sense of each variable's distribution. Do you notice anything out of the ordinary?

### Multivariate exploratory data analysis

### Fit simple logistic regression

### Fit a mulivariable logistic regression