

# Lab 3: Your responses

Your name: {Name here}

In the spaces below, insert your code and/or explanation as notes. Look for the angle brackets-{like these}-for where a response is needed.

When you are done, click on the arrow next to “knit” and choose “Knit as PDF” to save.

Load the data and the tidyverse package in the code chunk below

```
library(tidyverse)

## Loading tidyverse: ggplot2
## Loading tidyverse: tibble
## Loading tidyverse: tidyr
## Loading tidyverse: readr
## Loading tidyverse: purrr
## Loading tidyverse: dplyr

## Conflicts with tidy packages -----

## filter(): dplyr, stats
## lag():    dplyr, stats

cps_data<-read_csv("https://github.com/jshannon75/geog4300/raw/master/Labs/Lab%203-Confidence%20interval")

## Parsed with column specification:
## cols(
##   .default = col_integer(),
##   STATE = col_character(),
##   Region = col_character(),
##   CPSID = col_double(),
##   EARNWEEK = col_double()
## )

## See spec(...) for full column specifications.
```

**Question 1 (4 points)** *Using the information on this variable in the codebook (p. 12), create a subset of records without the NIU or missing response records. Then use group\_by and summarise to calculate the number of individuals grouped in each status as shown above. Use the resulting data frame to calculate an estimate of the national food insecurity rate.*

*#Calculate the national food insecurity rate following the steps above.*

**Question 2 (3 points)** *Using the formula for confidence intervals for proportions shown in class, calculate a confidence interval for the rate you identified in question 1. Interpret what that confidence interval tells you.*

*#Calculate the margin of error*

{Interpret what this confidence interval tells you.}

**Question 3 (5 points)** *Adapting the above function, create an estimated food insecurity rate for each state from these data. To do so, you'll need to create counts for each response (food secure, low food security, very low food insecurity), sum the latter two, and divide by the total responses within each state.*

*#Calculate state food insecurity rates*

**Question 4 (3 points)** *Create a command that estimates the margin of error for each state based on the national rate you calculated in question 1 and the total responses for each state.*

*#Calculate the margin of error for each state*

**Question 5 (3 points)** *Compare the margin of error you calculated for Georgia to the national margin of error. How do they differ? Mathematically, why are they different?*

{Your response here}

**Question 6 (5 points)** *Create a column in your state food insecurity estimates that converts each state's food insecurity rate to a z score based on the whole population. What are the z scores for Wisconsin, Washington, and Mississippi? What do those z scores tell you?*

*#Convert to a z score*

{What do the three z scores tell you?}

**Question 7 (4 points)** *Pick a probabilistic sampling strategy (or combination of strategies) discussed in our text or in lecture that would be appropriate for this research question: random, systemic, stratified, and cluster. Describe how this strategy could be used to create a sample for use in this proposed study. Describe one strength and weakness of this approach.*

{Your response here}

**Question 8 (3 points)** *Health officials would like to do a related survey of household food insecurity with enough responses to allow for margins of error under 2% (with 95% confidence). Assume that the rate is similar to the one you identified for Georgia in question 5. Use R to compute how big a sample they would need.*

*#Compute the sample size here.*