

Lab 4

Math 241, Week 4

```
# Put all necessary libraries here
library(tidyverse)
library(ggplot2)
library(tidyuesdayR)
library(readr)
library(datapasta)
library(dplyr)
#tinytex::install_tinytex()

#using the libraries
covid_survey <- read_csv("C:/Users/agjjo/OneDrive/Documents/GitHub/math241/labs/lab04/data/covid-survey")
#View(covid_survey)
dim(covid_survey)
```

```
## [1] 1122  14
```

Problem 1: COVID survey - interpretation

first thing that jumps out to me is that these nursing students do not seem to trust the vaccine Look at age * safe: the mid 20's students have the highest trust (older/younger rated lower with lower variation) the only clear indicator for if someone is queer was the gender question but, its seems like queer(in this case, nonbinary) students were more likely to trust the vaccine North american indigenous students had high variability all responses. This makes me think that that grouping on it's own is a poor predictor and if you wanted better info about that population(those populations) you would have to bring in more context ###

Problem 2: COVID survey - reconstruct

```
# Creating a function to delete rows with a certain minimum number of NA's
delete.na <- function(Df, n = 0) {
  Df[rowSums(is.na(Df)) <= n, ]
}

# Assuming 'covid_survey' is your data frame
covid_cleaned <- delete.na(covid_survey, 12)

# Relabeling survey responses, this doesn't work but U get the idea
#covid_cleaned <- covid_cleaned %>%
  #mutate(exp_race = case_when(
    #exp_race == 1 ~ "American Indian / Alaskan native",
    #exp_race == 2 ~ "Asian",
    #exp_race == 3 ~ "Black / African American",
    #exp_race == 4 ~ "Native Hawaiian / Other Pacific Islander",
    #exp_race == 5 ~ "White",
    #TRUE ~ as.character(exp_race) # Handling other cases
```

```
#))  
  
# Showing dimensions of the cleaned data  
dim(covid_cleaned)
```

```
## [1] 1112 14
```