# Extra\_credit\_Part\_1

# Kat Lee

## 5/17/2021

## Introduction

In this extra credit assignment we'll practice some tidyverse functions with a dataset about Covid-19 vaccinations across different counties in California. For more information about this data see https://data.ca. gov/dataset/covid-19-vaccine-progress-dashboard-data-by-zip-code.

First let's load tidyverse and read in dataset.

library(tidyverse)
vaccines = read\_csv("coveragebyzip\_od.csv")

Question 1: Fill in the following code to filter the dataframe for only zip codes for which vaccination information has not been redacted.

```
vaccines_filtered = vaccines %>%
filter([fill in here])
```

Question 2: Select the columns we're interested in (all columns but "VEM source" and "Redacted").

vaccines\_selected = vaccines\_filtered %>%
select([fill in here])

Question 3: Group by county and summarize average percent of the population fully vaccinated

```
vaccine_summary = vaccines_selected %>%
group_by([fill in here]) %>%
summarise([fill in here])
```

Question 4: We're going to make violin plots of the distributions of the percent of the population fully vaccinated or partially vaccinated across all zip codes in San Diego.

We must first prepare out dataframe for plotting by first filtering your vaccines\_selected dataframe for only zip codes in San Diego and then pivoting it longer so that Percent of Population Fully Vaccinated and Percent of Population Partially Vaccinated are stored in one column and their values in another column. Fill in the following code to do so:

```
vaccines_pivoted = vaccines_selected %>%
filter([fill in here]) %>%
pivot_longer(cols = [fill in here], names_to = "Metric", values_to = "Percentage")
```

Question 5: Now plot the distributions of the percent of the population fully vaccinated or partially vaccinated across all zip codes in San Diego as violin plots.

```
vaccines_pivoted %>%
ggplot(aes(x = [fill in here], y = [fill in here], fill = Metric)) +
geom_violin()
```

# Submission:

Please knit your Rmarkdown to pdf and submit on gradescope.

## **Rubric:**

Each question is worth 1 point for a total of 5 possible extra credit points, graded on the accuracy / functionality of your code.