## Bioinformatics First Year Exam

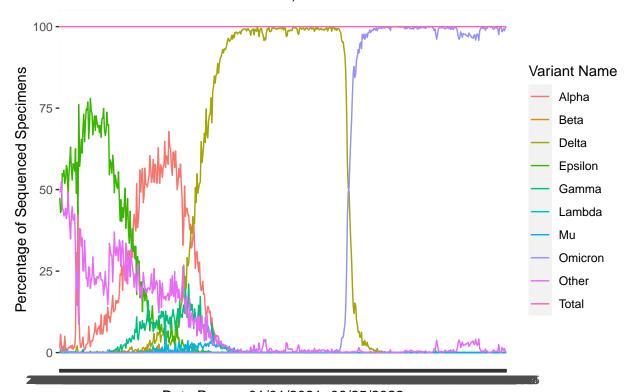
## Crisandra Jade Diaz, A59002951

## 7/17/2022

```
##Load packages into library
library(dplyr)
## Attaching package: 'dplyr'
## The following objects are masked from 'package:stats':
##
##
       filter, lag
## The following objects are masked from 'package:base':
##
##
       intersect, setdiff, setequal, union
library(lubridate)
##
## Attaching package: 'lubridate'
## The following objects are masked from 'package:base':
##
##
       date, intersect, setdiff, union
library(ggplot2)
library(tibble)
Data taken from https://data.chhs.ca.gov/dataset/covid-19-variant-data
##Reading CSV file and naming it as "covid.data"
covid.data <- read.csv("covid19_variants.csv")</pre>
\#\#Loading data from .csv file into ggplot to create a line graph
```

```
##Taking data from covid.data and used the following the format for a line graph to plot the data with
ggplot(data = covid.data, aes(x=date, y=percentage, group=variant_name, color=variant_name) )+
    ##insert line graph
    geom_line() +
    ## Label x-axis
    xlab("Date Range: 01/01/2021-06/25/2022") +
    ##Label y-axis
    ylab("Percentage of Sequenced Specimens") +
    ##Label Legend Title
    labs(colour= "Variant Name") +
    ##Label graph title
    ggtitle("Covid 19 Variants in California, USA")
```

## Covid 19 Variants in California, USA



Date Range: 01/01/2021-06/25/2022