First Year Exam

Diana Ofelia Lopez Cantu A59019192

First, libraries need to be called:

```
library(lubridate)
```

```
Loading required package: timechange
```

```

Attaching package: 'lubridate'
```

```
The following objects are masked from 'package:base':

    date, intersect, setdiff, union
```

```
library(ggplot2)
library(ggpubr)
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
```

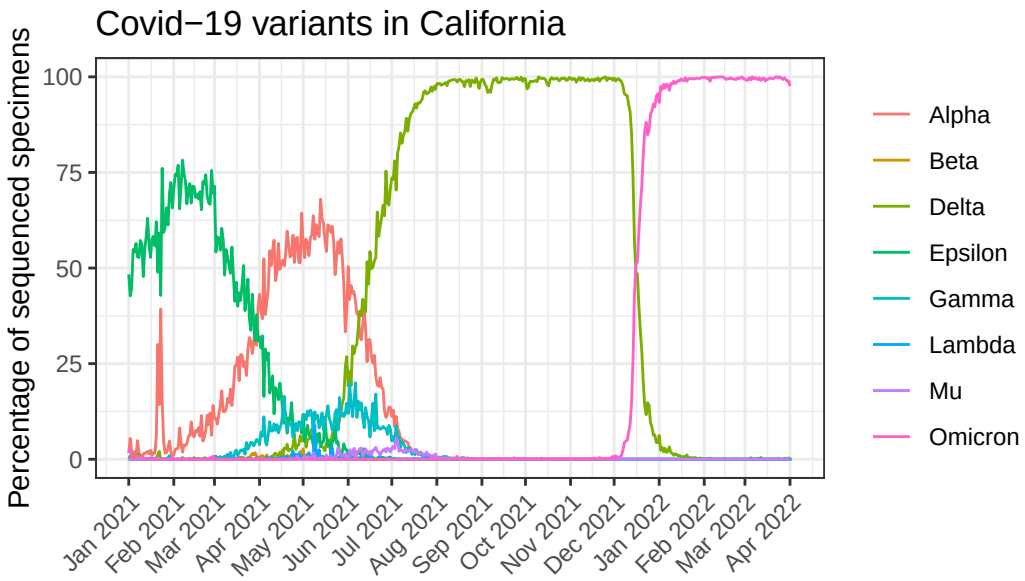
```r
library(scales)
```

```r
covid <- read.csv("/Users/dianalopezcantu/Desktop/covid19_variants.csv")
#covid
```

```r
covid$date <- ymd(covid$date)
df <- data.frame(covid$date, covid$variant_name, covid$percentage)
df2 <- dplyr::filter(df,
                     covid$variant_name
%in% c("Alpha", "Beta", "Delta", "Epsilon", "Gamma","Lambda", "Omicron","Mu"))
df3 <- subset(df2, covid$date < "2022-01-01")
#df3
```

```r
#Create a plot

ggplot(df3) + aes(x = covid.date, y = covid.percentage) +
  theme_bw() +
  geom_line(aes(colour =covid.variant_name)) +
  xlab("") + ylab("Percentage of sequenced specimens") +
  labs(title ="Covid-19 variants in California",
       caption = "Data Source: <https://www.cdph.ca.gov/>", colour = NULL) +
  scale_x_date(date_labels = "%b %Y", date_breaks = "1 month") +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```



Data Source: <https://www.cdph.ca.gov/>