COVID-19 Variant Data

PID:A59010419

7/15/2022

```
variants <- read.csv( 'covid19_variants.csv' )
head(variants)
```

```
##          date       area area_type variant_name specimens percentage
## 1 2021-01-01 California     State        Gamma         0       0.00
## 2 2021-01-01 California     State         Beta         0       0.00
## 3 2021-01-01 California     State       Lambda         0       0.00
## 4 2021-01-01 California     State        Alpha         1       1.69
## 5 2021-01-01 California     State      Epsilon        28      47.46
## 6 2021-01-01 California     State        Other        29      49.15
##   specimens_7d_avg percentage_7d_avg
## 1               NA                NA
## 2               NA                NA
## 3               NA                NA
## 4               NA                NA
## 5               NA                NA
## 6               NA                NA
```

date

```
library(lubridate)
```

```
##
## Attaching package: 'lubridate'
```

```
## The following objects are masked from 'package:base':
##
##     date, intersect, setdiff, union
```

```
variants$date <- ymd(variants$date)
```

drop

```
library(dplyr)
```

```
## Warning: package 'dplyr' was built under R version 4.1.2
```

```
##
## 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
```
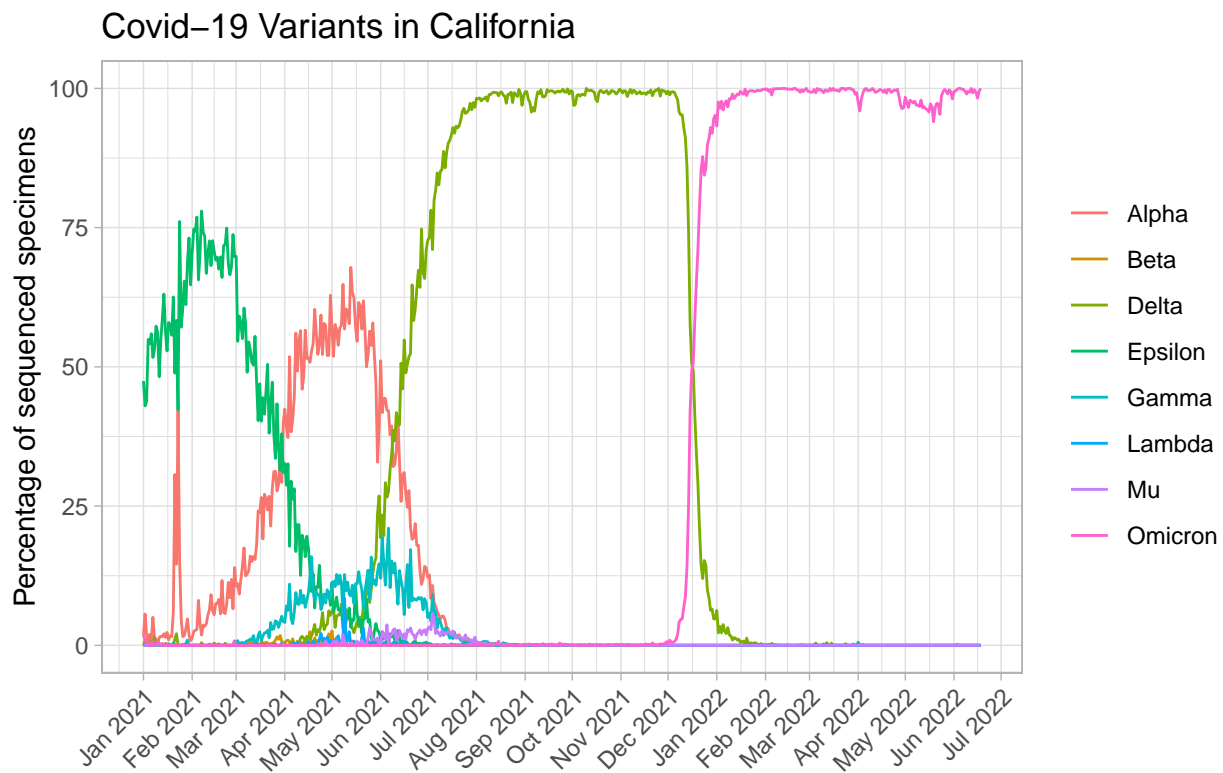
```r
 data<- filter(variants, variant_name != "Total")
 data<- filter(data, variant_name != "Other")
```

plot

```r
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 4.1.2
```

```r
ggplot(data) +
  aes(date,
      percentage,color=variant_name) +
  geom_line()+ scale_x_date(date_breaks = "1 month", date_labels = "%b %Y")+labs(x="", y="Percentage of
      caption = "Data Souce: <https://www.cdph.ca.gov/>")  + theme_light() + theme(legend.title=element
```


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