Covid-19 variants in California

Deepak (PID: A59010602)

17/07/2022

Importing libraries

```
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 4.1.3
```

```
library(dplyr)
```

```
## Warning: package 'dplyr' was built under R version 4.1.3
```

```
##
## 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(scales)
```

```
## Warning: package 'scales' was built under R version 4.1.3
```

Loading dataset

```
dataset = read.csv('covid19_variants.csv')
dataset$date = as.Date(dataset$date)
head(dataset,20)
```

```
##          date       area area_type variant_name specimens percentage
## 1  2021-01-01 California     State      Epsilon        28      47.46
## 2  2021-01-01 California     State        Other        29      49.15
## 3  2021-01-01 California     State        Alpha         1       1.69
## 4  2021-01-01 California     State         Beta         0       0.00
## 5  2021-01-01 California     State        Total        59     100.00
## 6  2021-01-01 California     State        Gamma         0       0.00
```

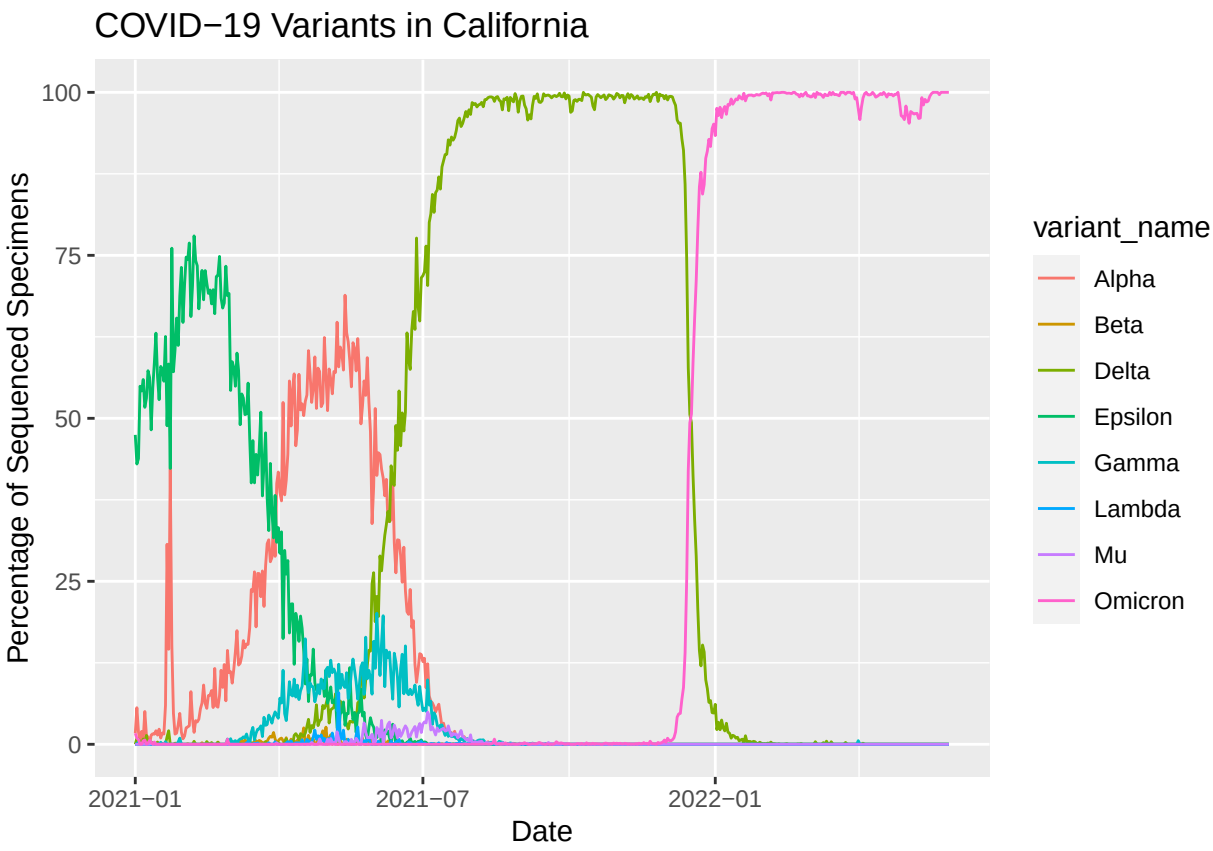
```
## 7  2021-01-01 California     State           Mu         0       0.00
## 8  2021-01-01 California     State        Delta         0       0.00
## 9  2021-01-01 California     State      Omicron         1       1.69
## 10 2021-01-01 California     State       Lambda         0       0.00
## 11 2021-01-02 California     State        Delta         1       0.93
## 12 2021-01-02 California     State      Omicron         1       0.93
## 13 2021-01-02 California     State      Epsilon        46      42.99
## 14 2021-01-02 California     State         Beta         0       0.00
## 15 2021-01-02 California     State           Mu         0       0.00
## 16 2021-01-02 California     State        Gamma         0       0.00
## 17 2021-01-02 California     State        Other        53      49.53
## 18 2021-01-02 California     State       Lambda         0       0.00
## 19 2021-01-02 California     State        Total       107     100.00
## 20 2021-01-02 California     State        Alpha         6       5.61
##    specimens_7d_avg percentage_7d_avg
## 1                NA                NA
## 2                NA                NA
## 3                NA                NA
## 4                NA                NA
## 5                NA                NA
## 6                NA                NA
## 7                NA                NA
## 8                NA                NA
## 9                NA                NA
## 10               NA                NA
## 11               NA                NA
## 12               NA                NA
## 13               NA                NA
## 14               NA                NA
## 15               NA                NA
## 16               NA                NA
## 17               NA                NA
## 18               NA                NA
## 19               NA                NA
## 20               NA                NA
```

Selecting for variants of interest

```
dataset  = dataset%>%
  filter(dataset$variant_name != 'Other' & dataset$variant_name != 'Total')
```

Plotting the data

```
ggplot(dataset, aes(x=date, y=percentage, color=variant_name)) +
  geom_line() + ggtitle("COVID-19 Variants in California") +
  xlab("Date") + ylab("Percentage of Sequenced Specimens") +
  scale_x_date(labels = date_format("%Y-%m"))
```


COVID−19 Variants in California
Percentage of Sequenced Specimens
100
75
50
25
0
2021−01
2021−07
2022−01
Date
variant_name
Alpha
Beta
Delta
Epsilon
Gamma
Lambda
Mu
Omicron