week6_assignment

Lunhan Zhang

2022-09-21

```
data <- read.csv("COVID-19 Activity.csv")
```

```
library(dplyr)
```

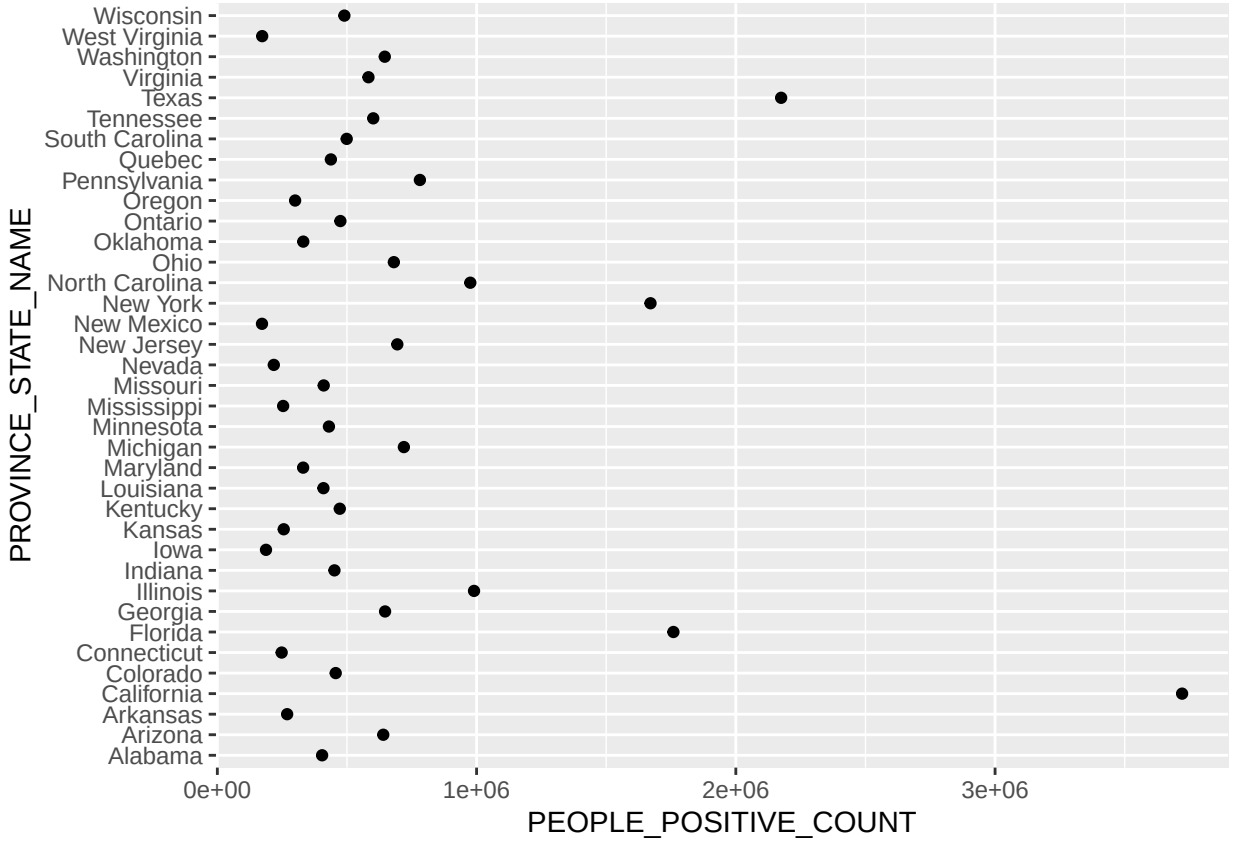
```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

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

```
smallerdata <- data %>%
select(PEOPLE_POSITIVE_CASES_COUNT, CONTINENT_NAME, PROVINCE_STATE_NAME, REPORT_DATE, PEOPLE_DEATH_NEW_C
filter(CONTINENT_NAME=='America') %>%
rename("COUNTRY_NAME"="CONTINENT_NAME")%>%
filter(REPORT_DATE>= ('2022-01-01'))%>%
  group_by(PROVINCE_STATE_NAME)  %>%
                    summarise(PEOPLE_POSITIVE_COUNT = sum(PEOPLE_POSITIVE_NEW_CASES_COUNT),
                              PEOPLE_DEATH_COUNT = sum(PEOPLE_DEATH_NEW_COUNT),
                              .groups = 'drop') %>%
  slice(-c(1)) %>%
  filter( PEOPLE_POSITIVE_COUNT>100000 & PEOPLE_DEATH_COUNT> 1000)
```

```
library(ggplot2)
ggplot(data = smallerdata) +
  geom_point(mapping = aes(x = PEOPLE_POSITIVE_COUNT, y = PROVINCE_STATE_NAME))
```

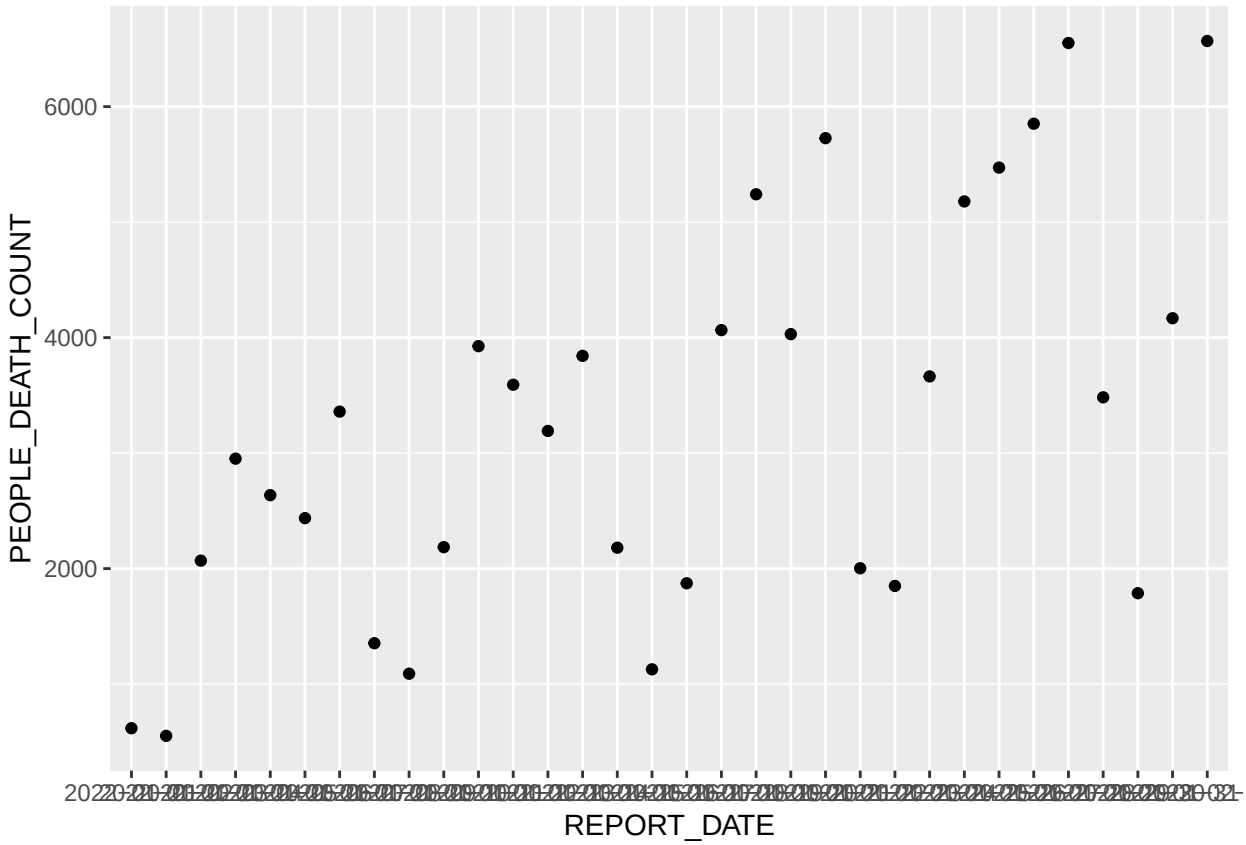


This plot is showing the number of people who are tested positive from some of the states.

```
library(dplyr)
datedata <- data %>%
select(PEOPLE_POSITIVE_CASES_COUNT, CONTINENT_NAME, PROVINCE_STATE_NAME, REPORT_DATE, PEOPLE_DEATH_NEW_C
filter(CONTINENT_NAME=='America') %>%
rename("COUNTRY_NAME"="CONTINENT_NAME")%>%
filter(REPORT_DATE>= ('2022-01-01')&REPORT_DATE<= ('2022-02-01'))%>%
  group_by(REPORT_DATE)  %>%
                    summarise(PEOPLE_POSITIVE_COUNT = sum(PEOPLE_POSITIVE_NEW_CASES_COUNT),
                              PEOPLE_DEATH_COUNT = sum(PEOPLE_DEATH_NEW_COUNT))
library(ggplot2)

ggplot(data = datedata) +
  geom_point(mapping = aes(x = REPORT_DATE, y = PEOPLE_DEATH_COUNT))+
  geom_smooth(mapping = aes(x = REPORT_DATE, y = PEOPLE_DEATH_COUNT))
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```



This plot is for the trend in 2022 Feb, I have couple questions in this plot. 1. How can I reset the x-scale as a proper name? 2. This graph only shows the point but not the line. Could you help me check this code?