Plots

601 Group 3 Project 2 Plots

12/12/2021

```
data <- read.csv("/Users/seola/Desktop/modifieddata.csv", header=TRUE)


data$state <- as.character(data$state)
data$state[data$state == "Alabama"] <-"Southeast"
data$state[data$state == "Alaska"] <-"West"
data$state[data$state == "Arizona"] <-"Southwest"
data$state[data$state == "Arkansas"] <-"Southeast"
data$state[data$state == "California"] <-"West"
data$state[data$state == "Alabama"] <-"Southeast"
data$state[data$state == "Alaska"] <-"West"
data$state[data$state == "Arizona"] <-"Southwest"
data$state[data$state == "Arkansas"] <-"Southeast"
data$state[data$state == "California"] <-"West"
data$state[data$state == "Colorado"] <-"West"
data$state[data$state == "Connecticut"] <-"Northeast"
data$state[data$state == "District of Columbia"] <-"Southeast"
data$state[data$state == "Delaware"] <-"Midwest"
data$state[data$state == "Florida"] <-"Southeast"
data$state[data$state == "Georgia"] <-"Southeast"
data$state[data$state == "Hawaii"] <-"West"
data$state[data$state == "Idaho"] <-"West"
data$state[data$state == "Illinois"] <-"Midwest"
data$state[data$state == "Indiana"] <-"Midwest"
data$state[data$state == "Iowa"] <-"Midwest"
data$state[data$state == "Kansas"] <-"Midwest"
data$state[data$state == "Kentucky"] <-"Southeast"
data$state[data$state == "Louisiana"] <-"Southeast"
data$state[data$state == "Maine"] <-"Northeast"
data$state[data$state == "Maryland"] <-"Northeast"
data$state[data$state == "Massachusetts"] <-"Northeast"
data$state[data$state == "Michigan"] <-"Midwest"
data$state[data$state == "Minnesota"] <-"Midwest"
data$state[data$state == "Mississippi"] <-"Midwest"
data$state[data$state == "Missouri"] <-"Midwest"
data$state[data$state == "Montana"] <-"West"
data$state[data$state == "Nebraska"] <-"Midwest"
data$state[data$state == "Nevada"] <-"West"
data$state[data$state == "New Hampshire"] <-"Northeast"
data$state[data$state == "New Jersey"] <-"Northeast"
data$state[data$state == "New Mexico"] <-"Southwest"
data$state[data$state == "New York"] <-"Northeast"
```

```
data$state[data$state == "North Carolina"] <-"Southeast"
data$state[data$state == "North Dakota"] <-"Midwest"
data$state[data$state == "Ohio"] <-"Midwest"
data$state[data$state == "Oklahoma"] <-"Southwest"
data$state[data$state == "Oregon"] <-"West"
data$state[data$state == "Pennsylvania"] <-"Northeast"
data$state[data$state == "Rhode Island"] <-"Northeast"
data$state[data$state == "South Carolina"] <-"Southeast"
data$state[data$state == "South Dakota"] <-"Midwest"
data$state[data$state == "Tennessee"] <-"Southeast"
data$state[data$state == "Texas"] <-"Southwest"
data$state[data$state == "Utah"] <-"West"
data$state[data$state == "Vermont"] <-"Northeast"
data$state[data$state == "Virginia"] <-"Southeast"
data$state[data$state == "Washington"] <-"West"
data$state[data$state == "West Virginia"] <-"Southeast"
data$state[data$state == "Wisconsin"] <-"Midwest"
data$state[data$state == "Wyoming"] <-"West"

data$GDP.USD. <- as.numeric(gsub(",","", data$GDP.USD.))
data$Personal.Income <- as.numeric(gsub(",","", data$Personal.Income))
```

Summary

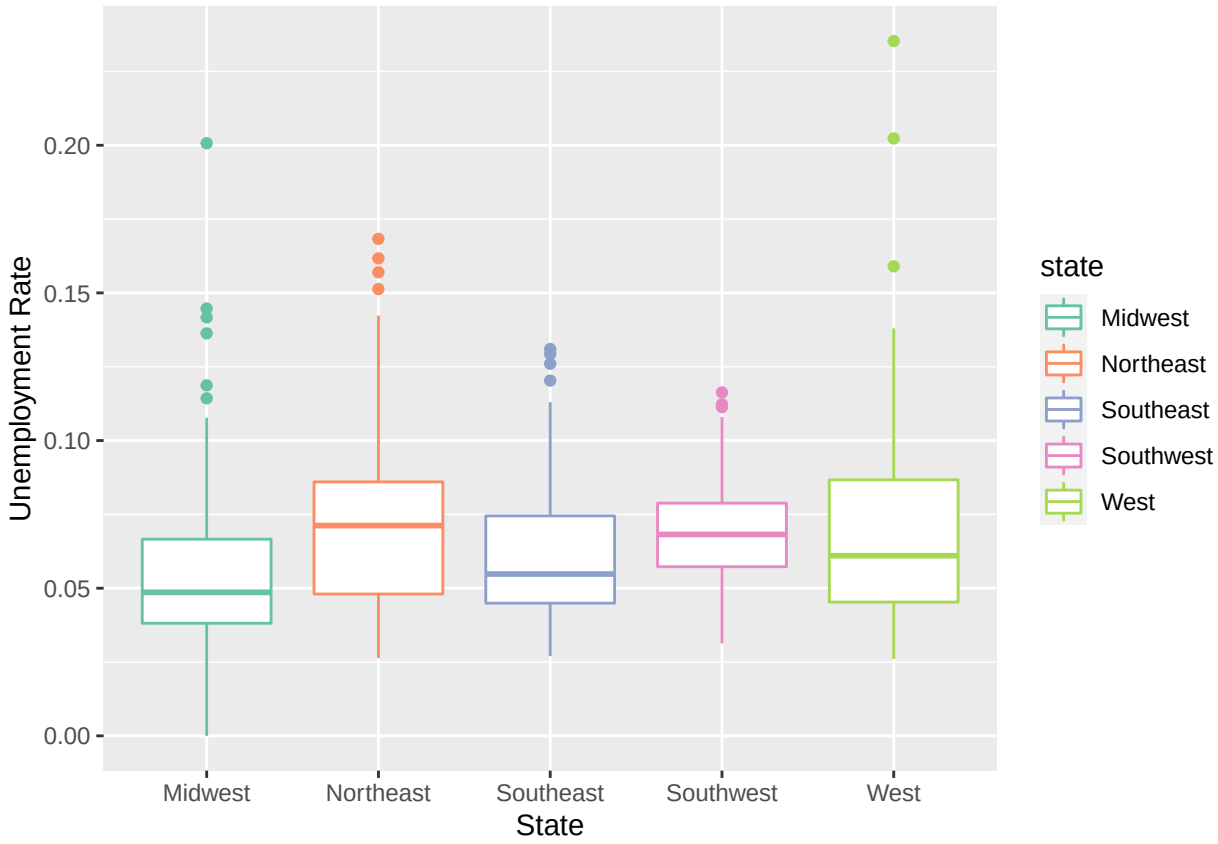
```
summary(data)
```

```
##     state              quarter          unemployment.rate   new.cases
##  Length:357         Length:357         Min.   :0.00000   Min.   :      0
##  Class :character   Class :character   1st Qu.:0.04130   1st Qu.:  11036
##  Mode  :character   Mode  :character   Median :0.05800   Median :  47832
##                                        Mean   :0.06504   Mean   : 110122
##                                        3rd Qu.:0.07900   3rd Qu.: 122955
##                                        Max.   :0.23530   Max.   :1420927
##      death       lockdown.days   vaccination.rate    GDP.USD.
##  Min.   :    0   Min.   : 0.00   Min.   :0.0000   Min.   :  30804
##  1st Qu.:  151   1st Qu.: 0.00   1st Qu.:0.0000   1st Qu.:  93891
##  Median :  651   Median : 0.00   Median :0.0000   Median : 243555
##  Mean   : 1770   Mean   : 5.86   Mean   :0.1638   Mean   : 438332
##  3rd Qu.: 2033   3rd Qu.: 0.00   3rd Qu.:0.3839   3rd Qu.: 559479
##  Max.   :30054   Max.   :90.00   Max.   :0.6791   Max.   :6187637
##  Personal.Income
##  Min.   :  35158
##  1st Qu.:  91750
##  Median : 241175
##  Mean   : 397164
##  3rd Qu.: 530063
##  Max.   :3052951
```

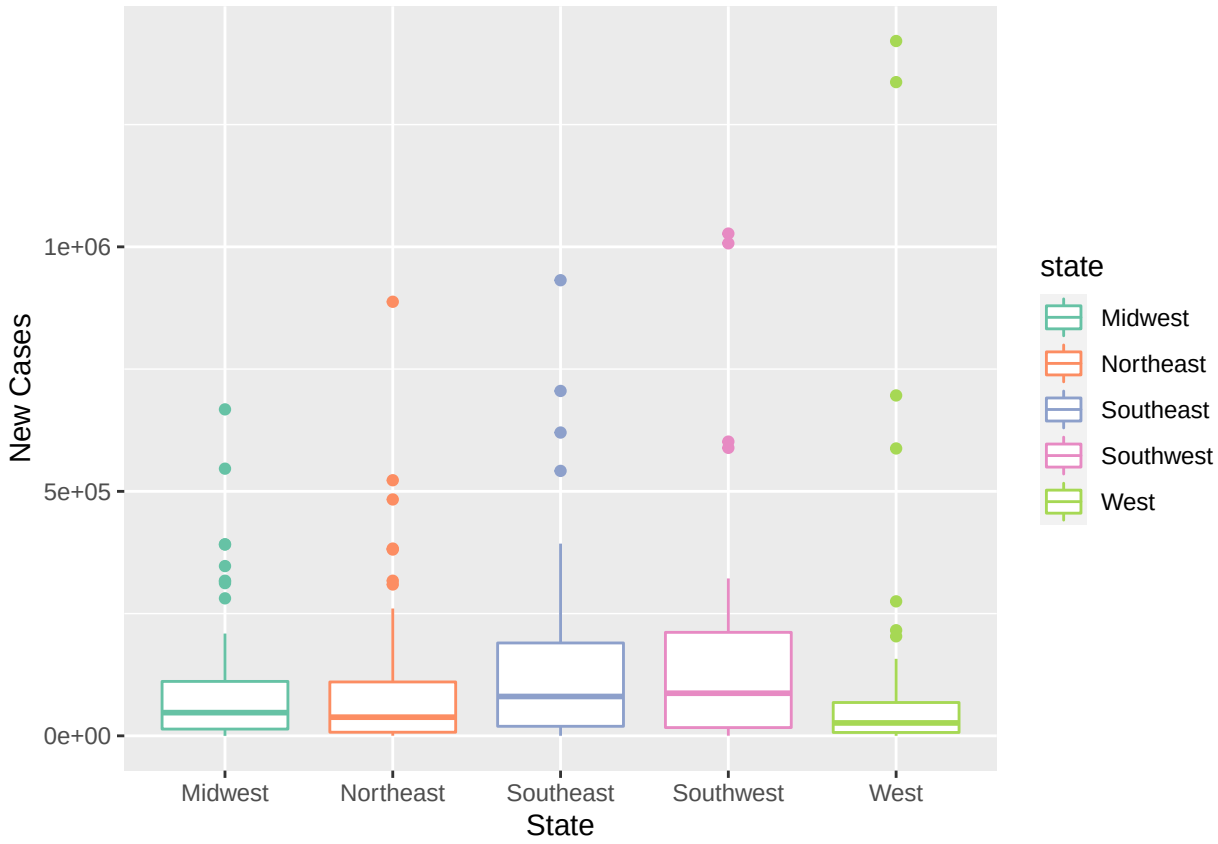
Boxplots categorized by states

```
p1 = ggplot(data, aes(x=state, y=unemployment.rate, col=state)) +
  geom_boxplot() +
  labs(x = "State", y = "Unemployment Rate") +
  scale_color_brewer(palette = "Set2")
```

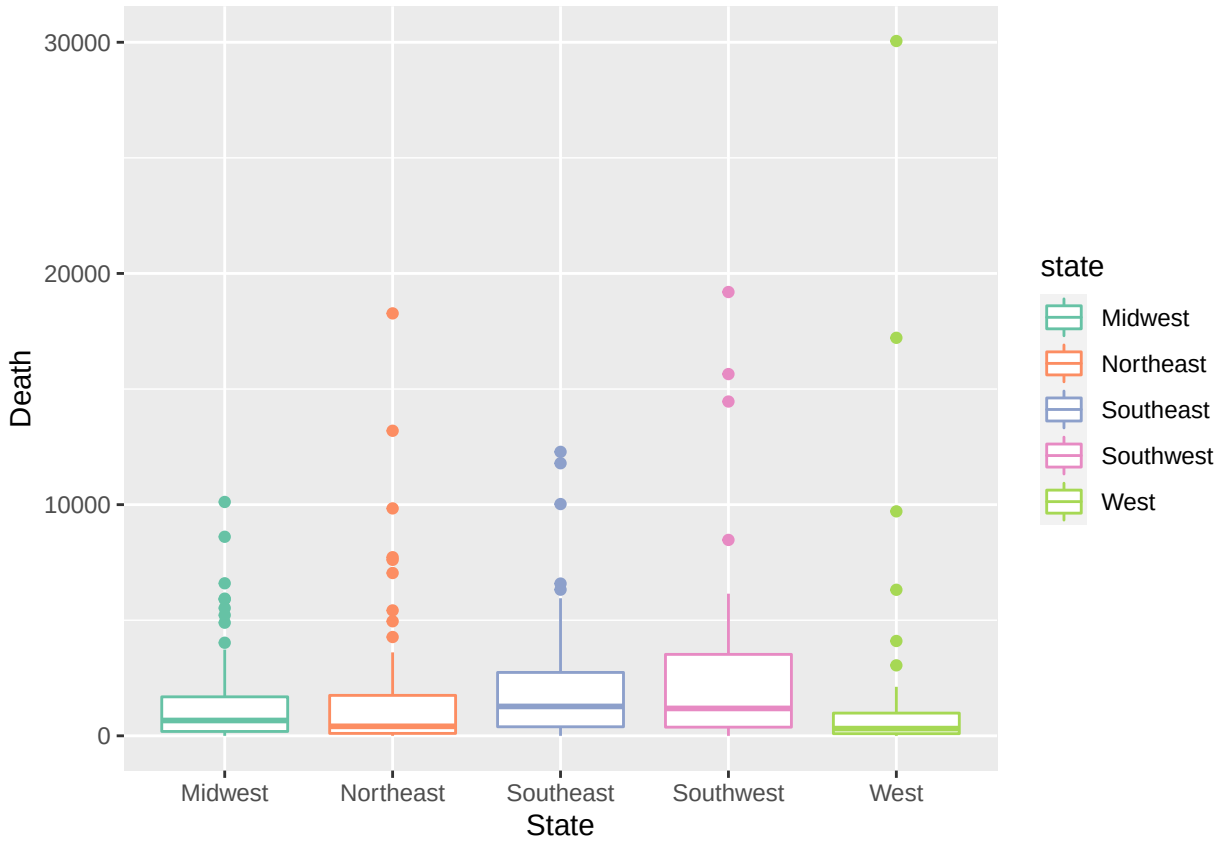
```r
p2 = ggplot(data, aes(x=state, y=new.cases, col=state)) +
  geom_boxplot() +
  labs(x = "State", y = "New Cases") +
  scale_color_brewer(palette = "Set2")

p3 = ggplot(data, aes(x=state, y=death, col=state)) +
  geom_boxplot() +
  labs(x = "State", y = "Death") +
  scale_color_brewer(palette = "Set2")

p4 = ggplot(data, aes(x=state, y=lockdown.days, col=state)) +
  geom_boxplot() +
  labs(x = "State", y = "Lockdown Days") +
  scale_color_brewer(palette = "Set2")

p5 = ggplot(data, aes(x=state, y=vaccination.rate, col=state)) +
  geom_boxplot() +
  labs(x = "State", y = "Vaccination Rate") +
  scale_color_brewer(palette = "Set2")

p6 = ggplot(data, aes(x=state, y=GDP.USD., col=state)) +
  geom_boxplot() +
  labs(x = "State", y = "GDP") +
  scale_color_brewer(palette = "Set2")

p7 = ggplot(data, aes(x=state, y=Personal.Income, col=state)) +
  geom_boxplot() +
  labs(x = "State", y = "Personal Income") +
  scale_color_brewer(palette = "Set2")

p1
```


Unemployment Rate
0.20
0.15
0.10
0.05
0.00
Midwest
Northeast
Southeast
Southwest
West
State
state
Midwest
Northeast
Southeast
Southwest
West

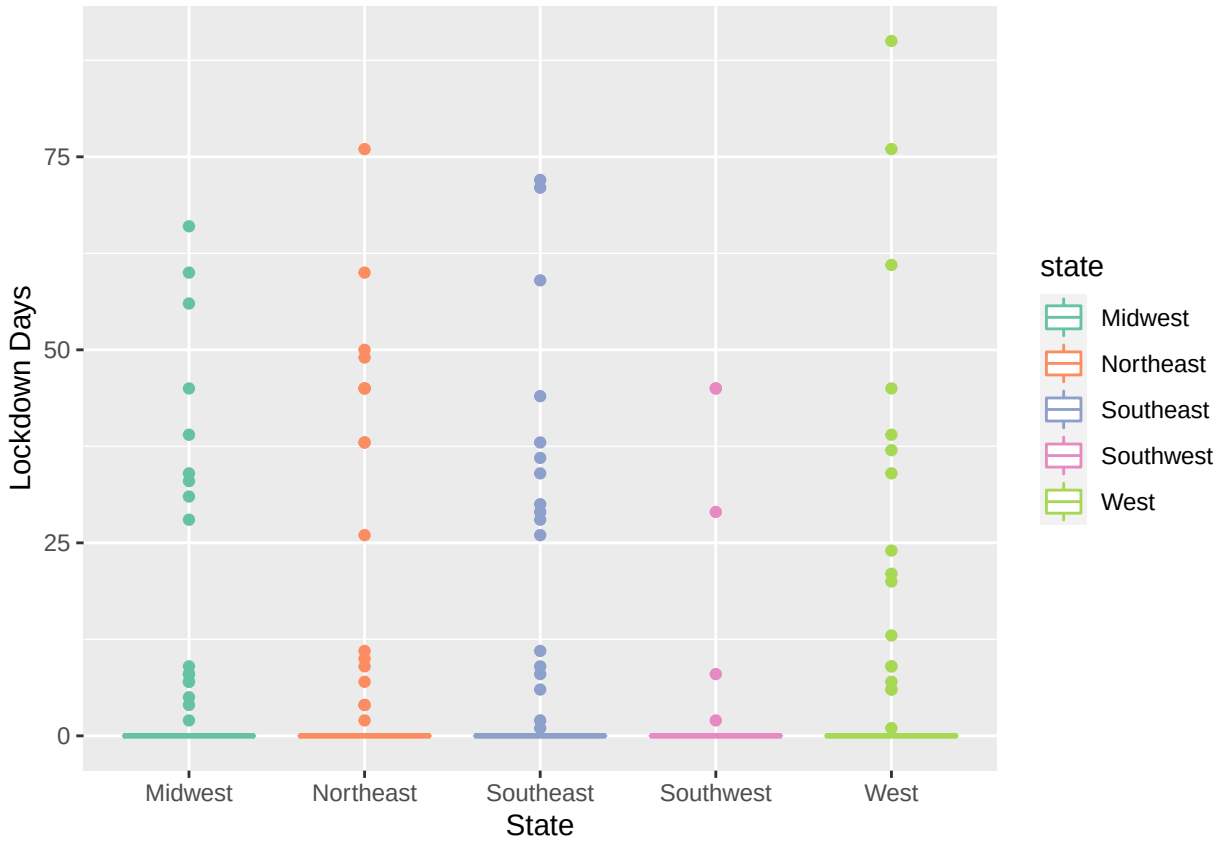
```
p2
```


New Cases
1e+06
5e+05
0e+00
Midwest
Northeast
Southeast
Southwest
West
State
state
Midwest
Northeast
Southeast
Southwest
West

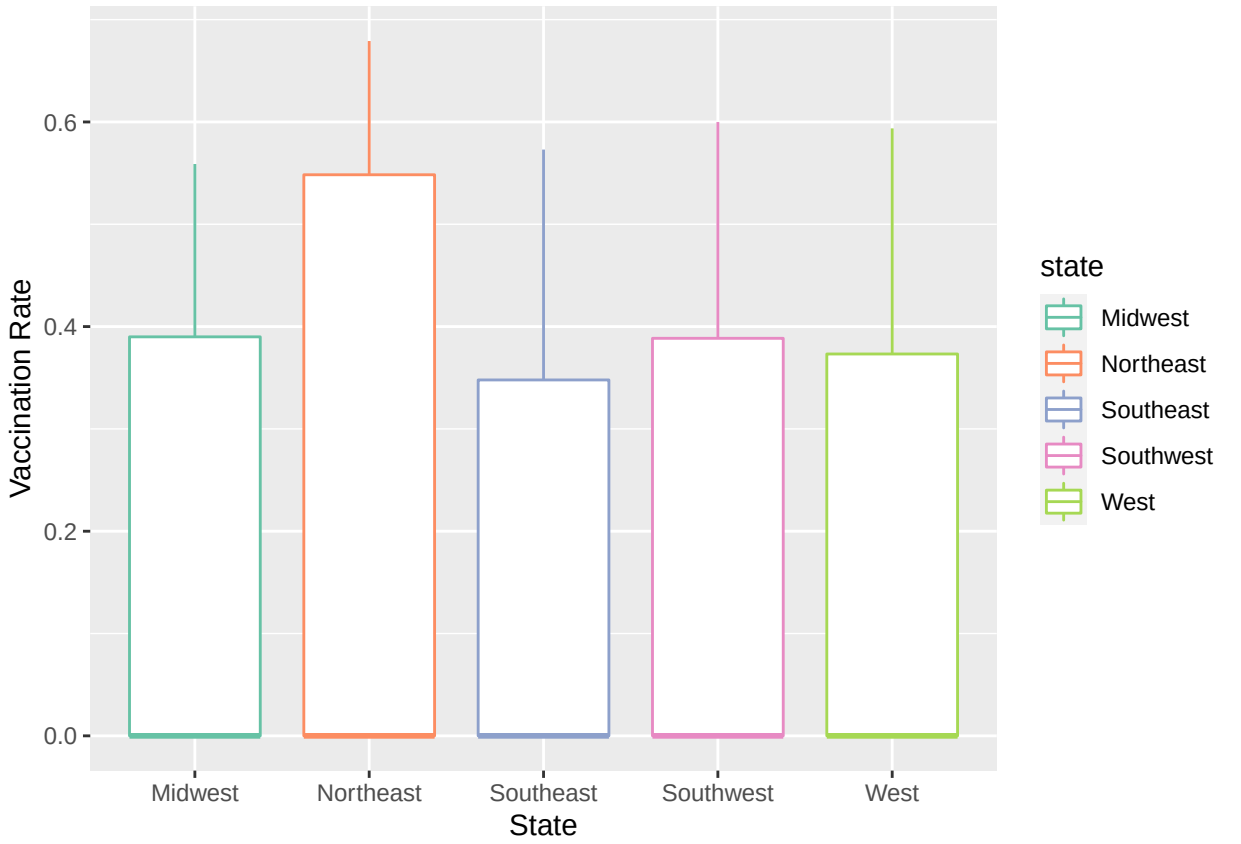
```
p3
```


30000
20000
10000
0
Death
Midwest
Northeast
Southeast
Southwest
West
State
state
Midwest
Northeast
Southeast
Southwest
West

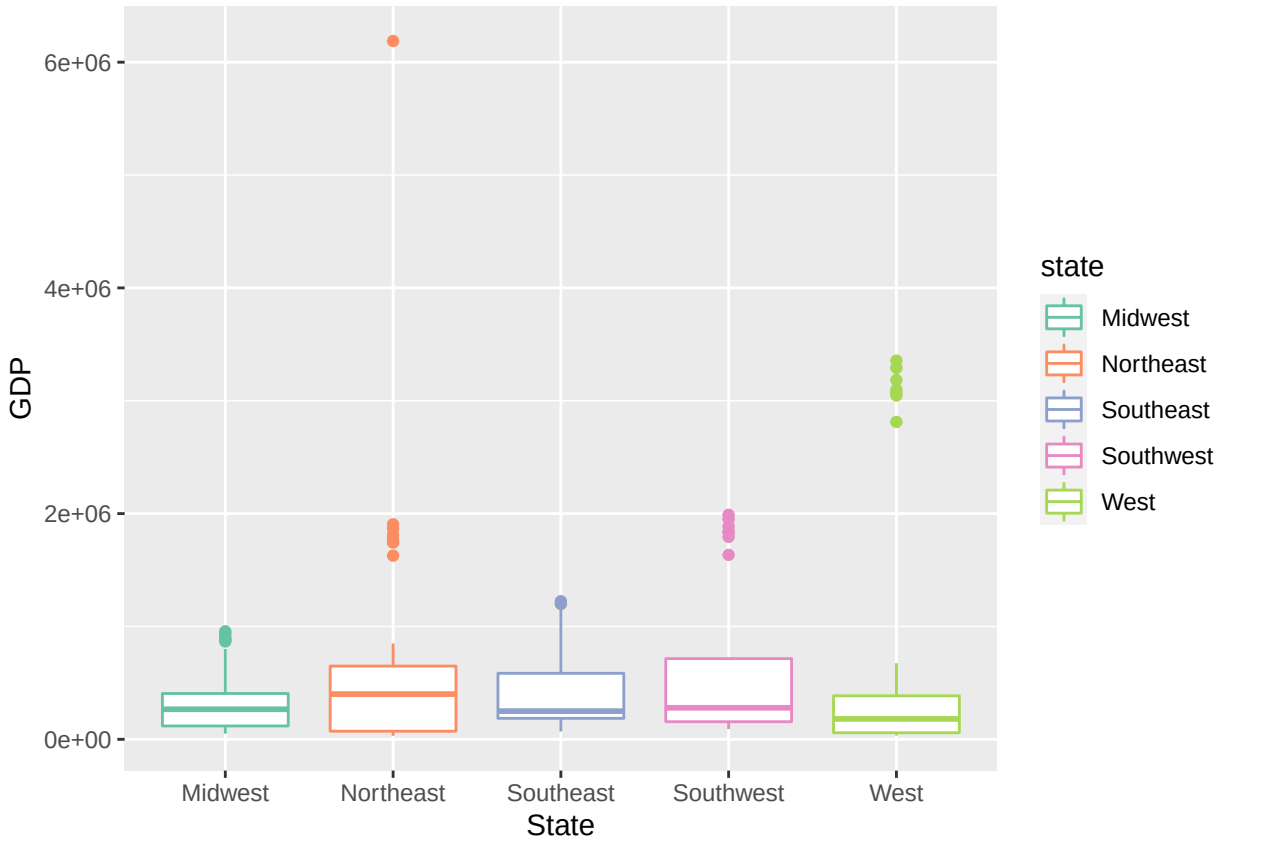
```
p4
```


Lockdown Days
75
50
25
0
Midwest
Northeast
Southeast
Southwest
West
State
state
Midwest
Northeast
Southeast
Southwest
West

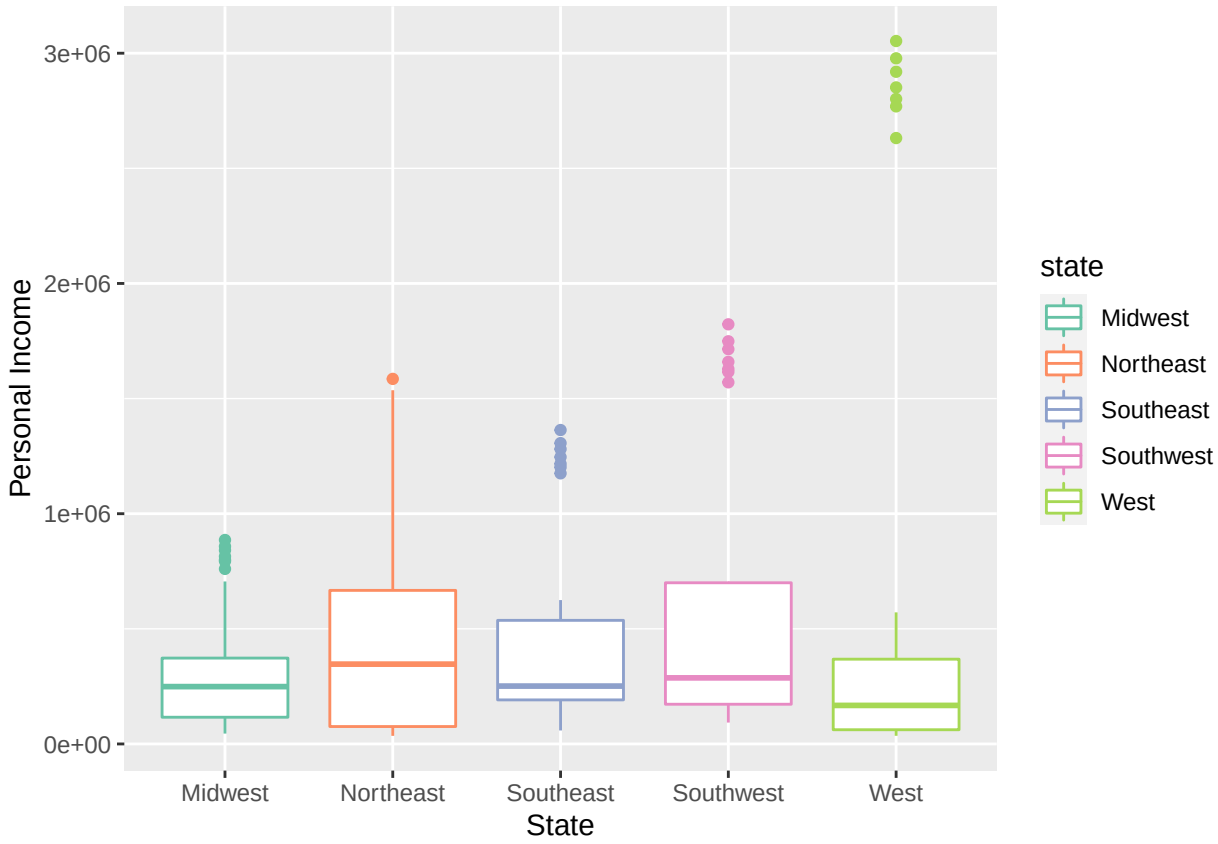
```
p5
```


Vaccination Rate
0.6
0.4
0.2
0.0
Midwest
Northeast
Southeast
Southwest
West
State
state
Midwest
Northeast
Southeast
Southwest
West

```
p6
```


6e+06
4e+06
2e+06
0e+00
GDP
Midwest
Northeast
Southeast
Southwest
West
State
state
Midwest
Northeast
Southeast
Southwest
West

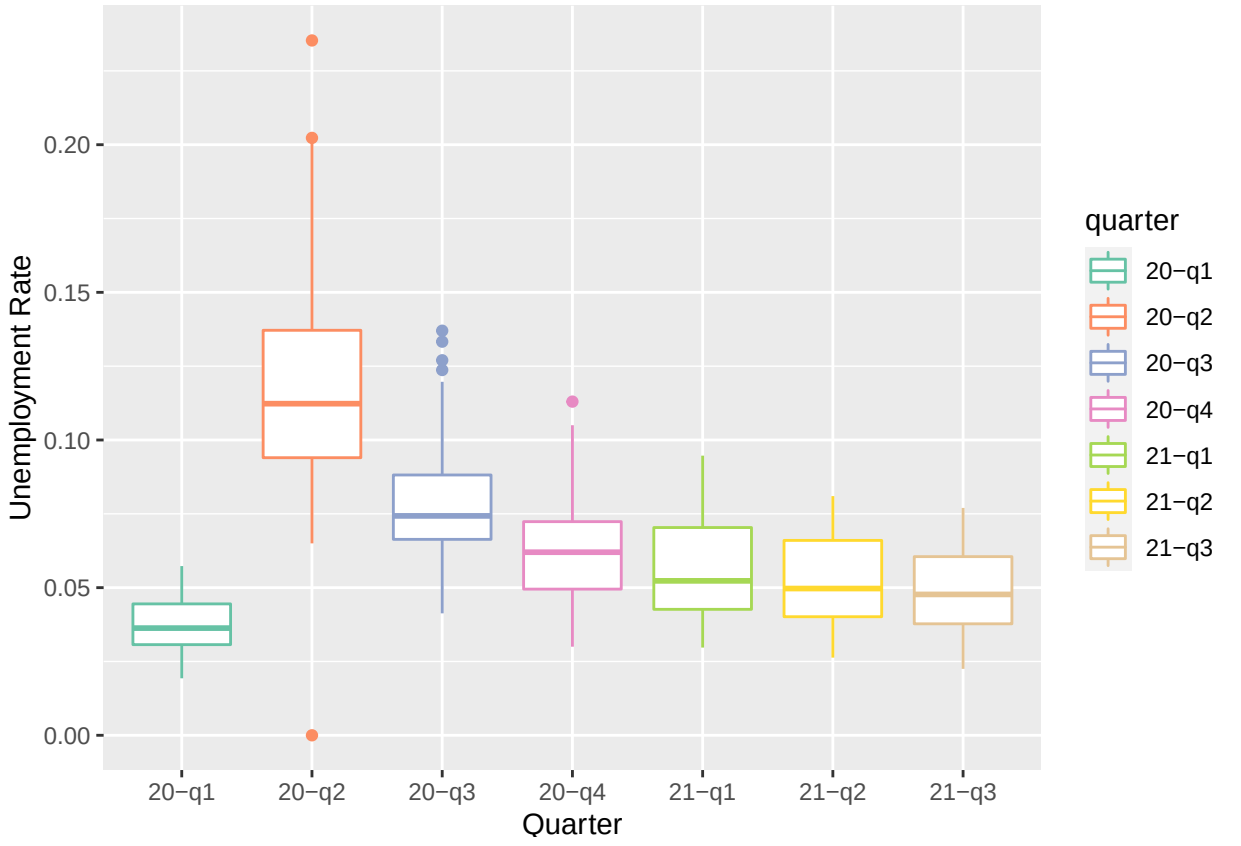
```
p7
```



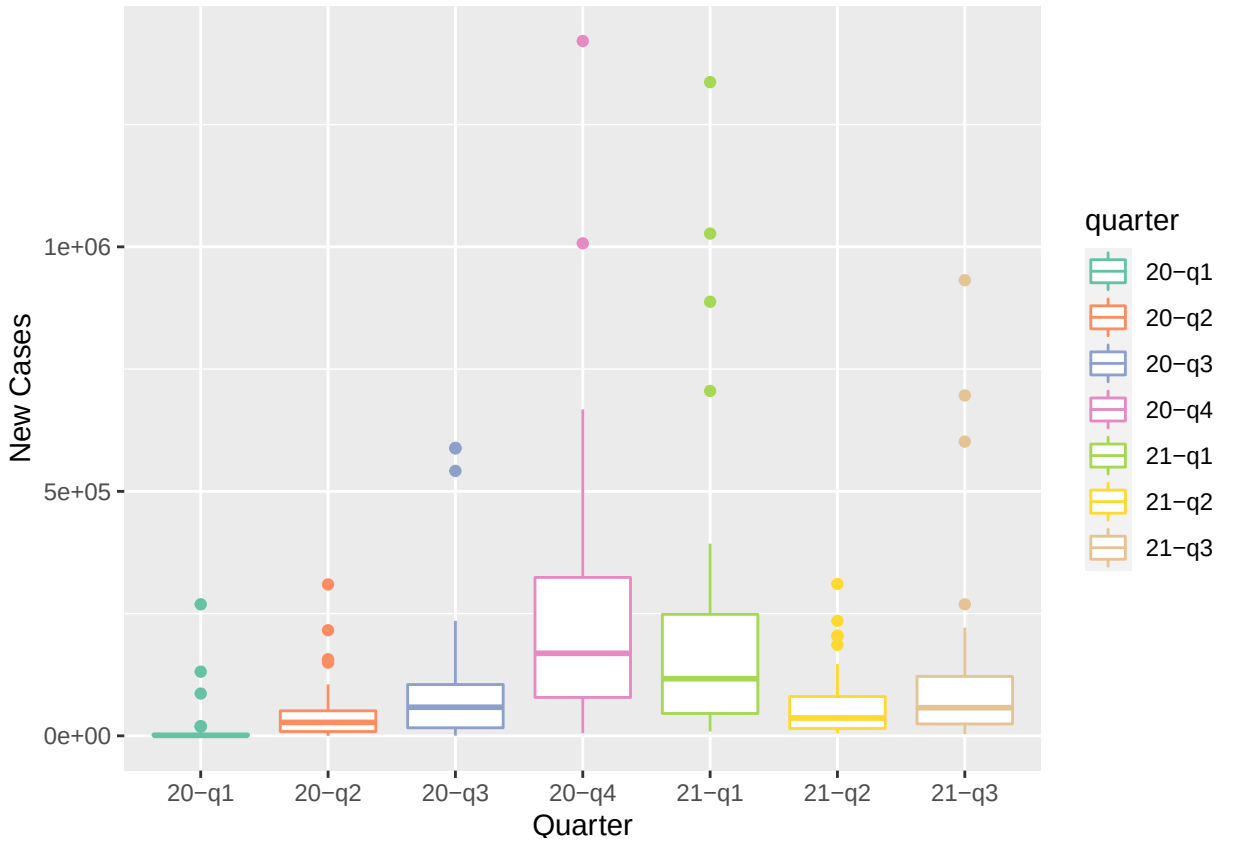
Boxplots categorized by quarter

```
p11 = ggplot(data, aes(x=quarter, y=unemployment.rate, col=quarter)) +
  geom_boxplot() +
  labs(x = "Quarter", y = "Unemployment Rate") +
  scale_color_brewer(palette = "Set2")

p22 = ggplot(data, aes(x=quarter, y=new.cases, col=quarter)) +
  geom_boxplot() +
  labs(x = "Quarter", y = "New Cases") +
  scale_color_brewer(palette = "Set2")

p33 = ggplot(data, aes(x=quarter, y=death, col=quarter)) +
  geom_boxplot() +
  labs(x = "Quarter", y = "Death") +
  scale_color_brewer(palette = "Set2")

p44 = ggplot(data, aes(x=quarter, y=lockdown.days, col=quarter)) +
  geom_boxplot() +
  labs(x = "Quarter", y = "Lockdown Days") +
  scale_color_brewer(palette = "Set2")

p55 = ggplot(data, aes(x=quarter, y=vaccination.rate, col=quarter)) +
  geom_boxplot() +
  labs(x = "Quarter", y = "Vaccination Rate") +
  scale_color_brewer(palette = "Set2")

p66 = ggplot(data, aes(x=quarter, y=GDP.USD., col=quarter)) +
```

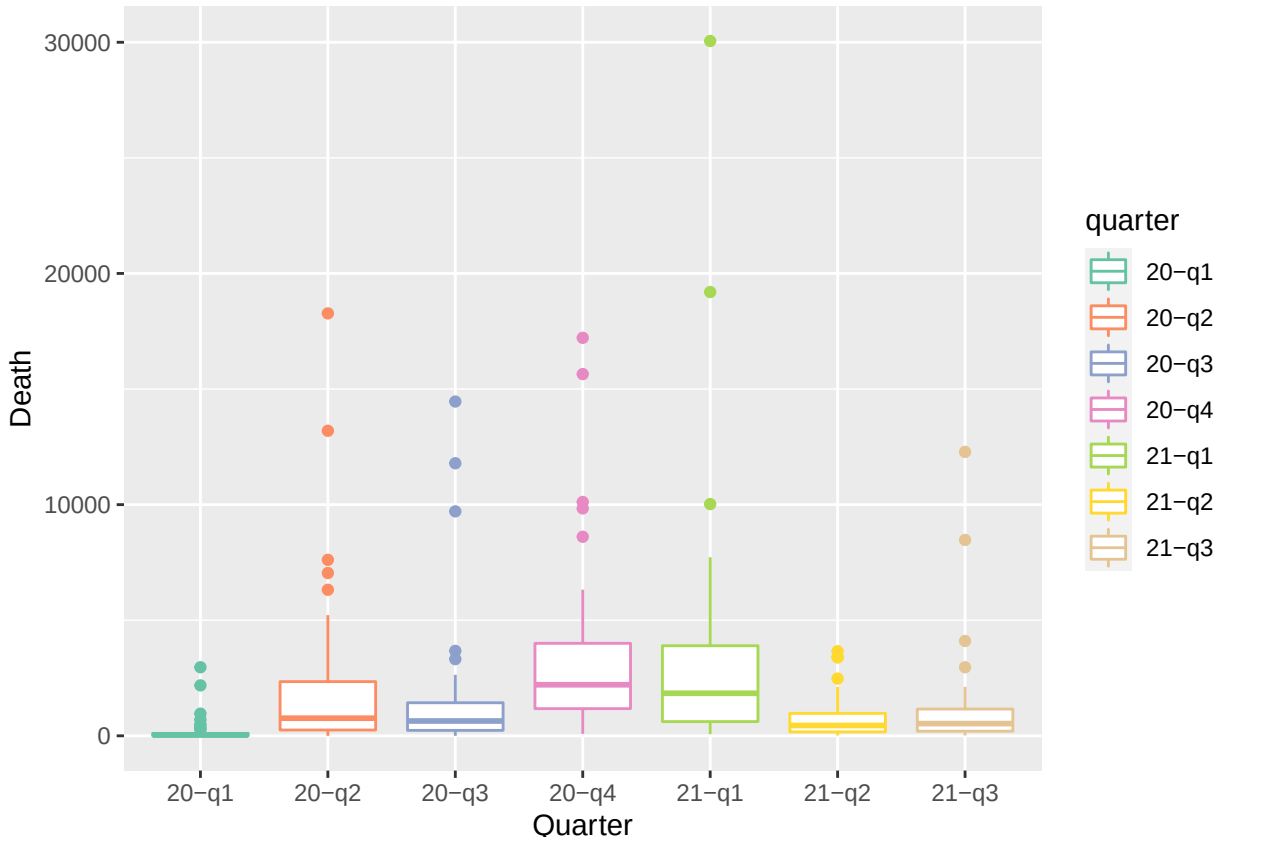
```
  geom_boxplot() +
  labs(x = "Quarter", y = "GDP") +
  scale_color_brewer(palette = "Set2")

p77 = ggplot(data, aes(x=quarter, y=Personal.Income, col=quarter)) +
  geom_boxplot() +
  labs(x = "Quarter", y = "Personal Income") +
  scale_color_brewer(palette = "Set2")

p11
```



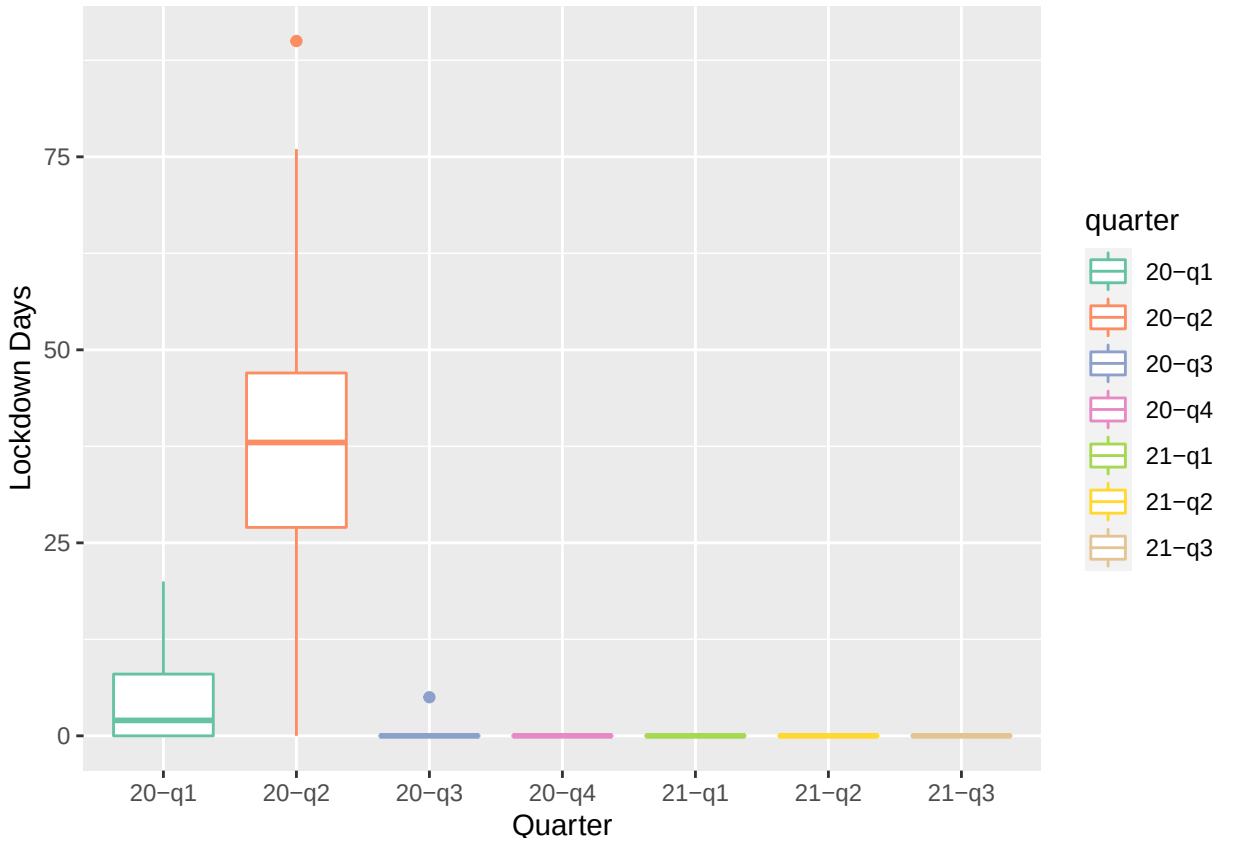

```
p22
```


New Cases
1e+06
5e+05
0e+00
20−q1
20−q2
20−q3
20−q4
21−q1
21−q2
21−q3
Quarter
quarter
20−q1
20−q2
20−q3
20−q4
21−q1
21−q2
21−q3

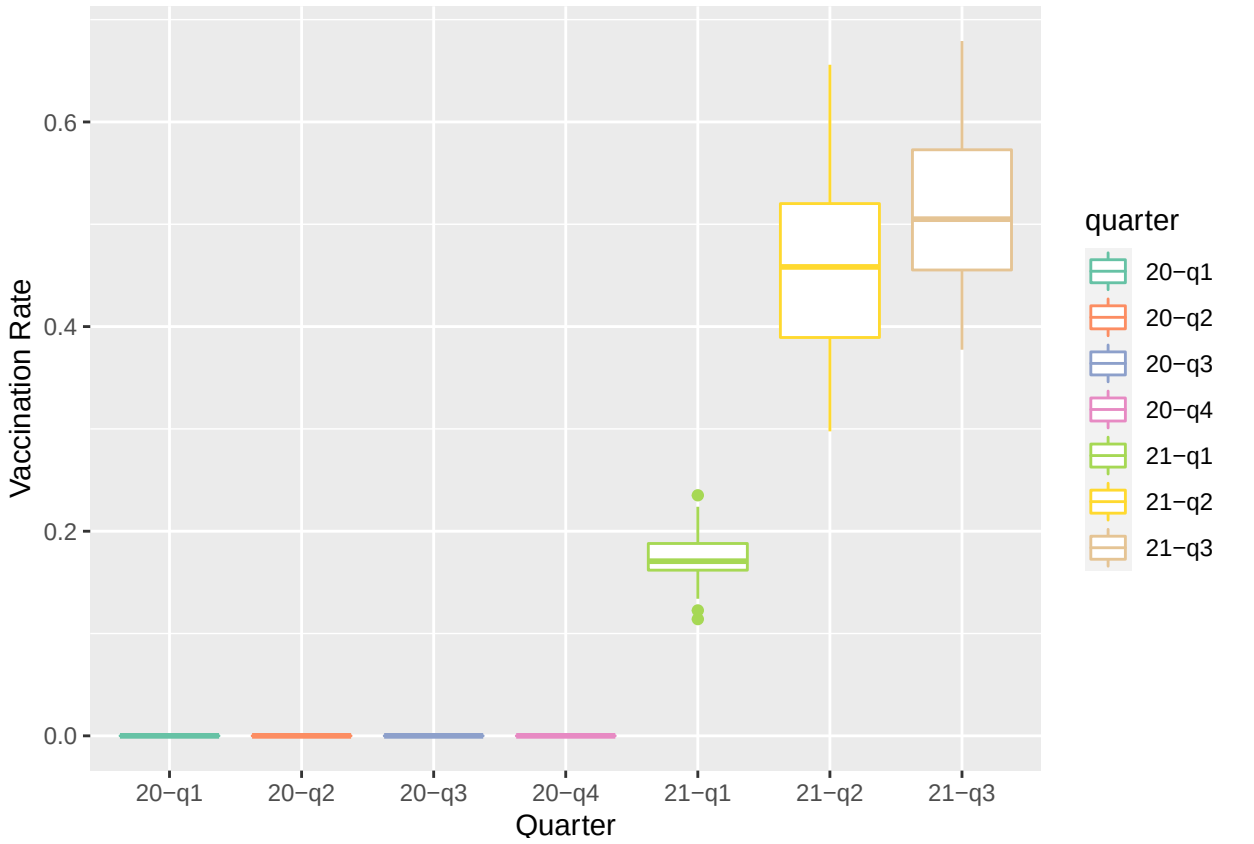
```
p33
```


30000
20000
10000
0
Death
20−q1
20−q2
20−q3
20−q4
21−q1
21−q2
21−q3
Quarter
quarter
20−q1
20−q2
20−q3
20−q4
21−q1
21−q2
21−q3

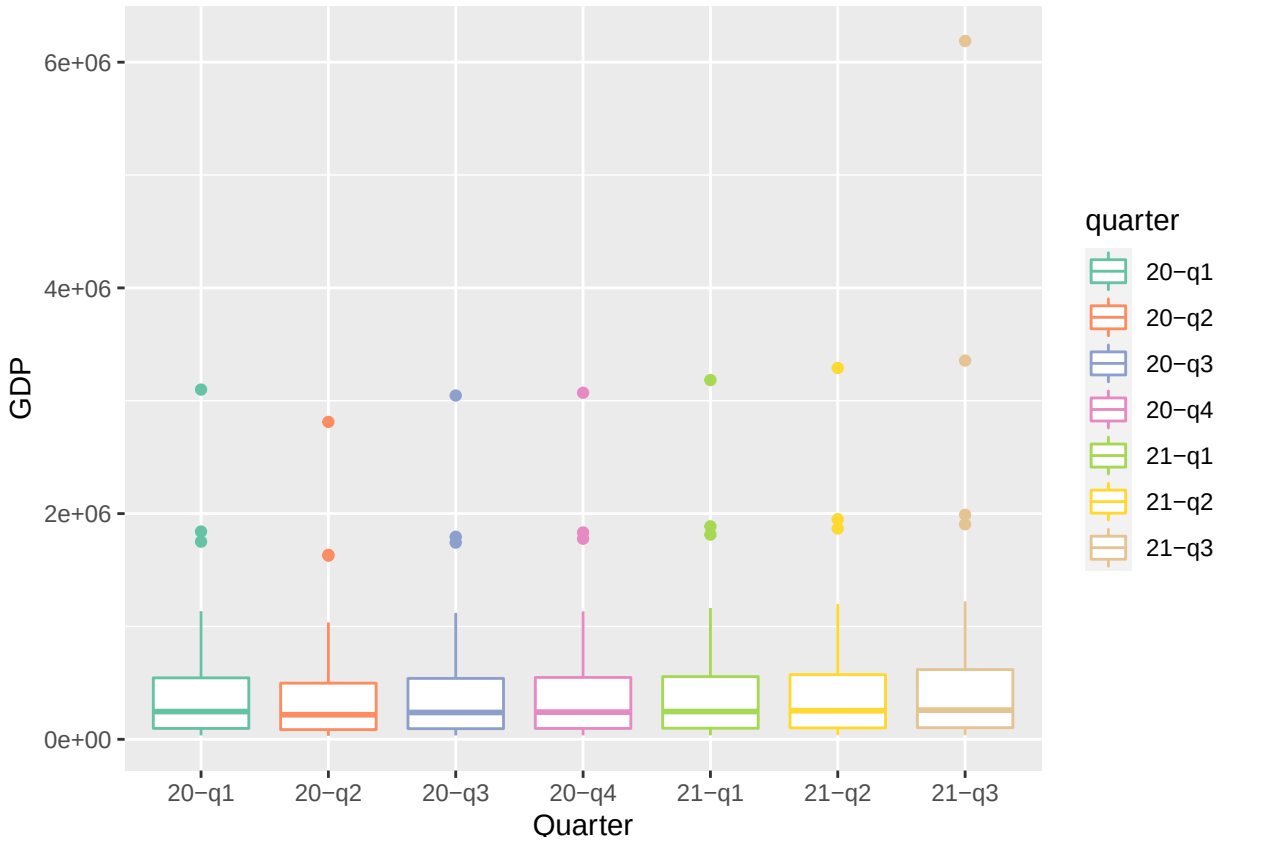
```
p44
```


Lockdown Days
75
50
25
0
20−q1
20−q2
20−q3
20−q4
21−q1
21−q2
21−q3
Quarter
quarter
20−q1
20−q2
20−q3
20−q4
21−q1
21−q2
21−q3

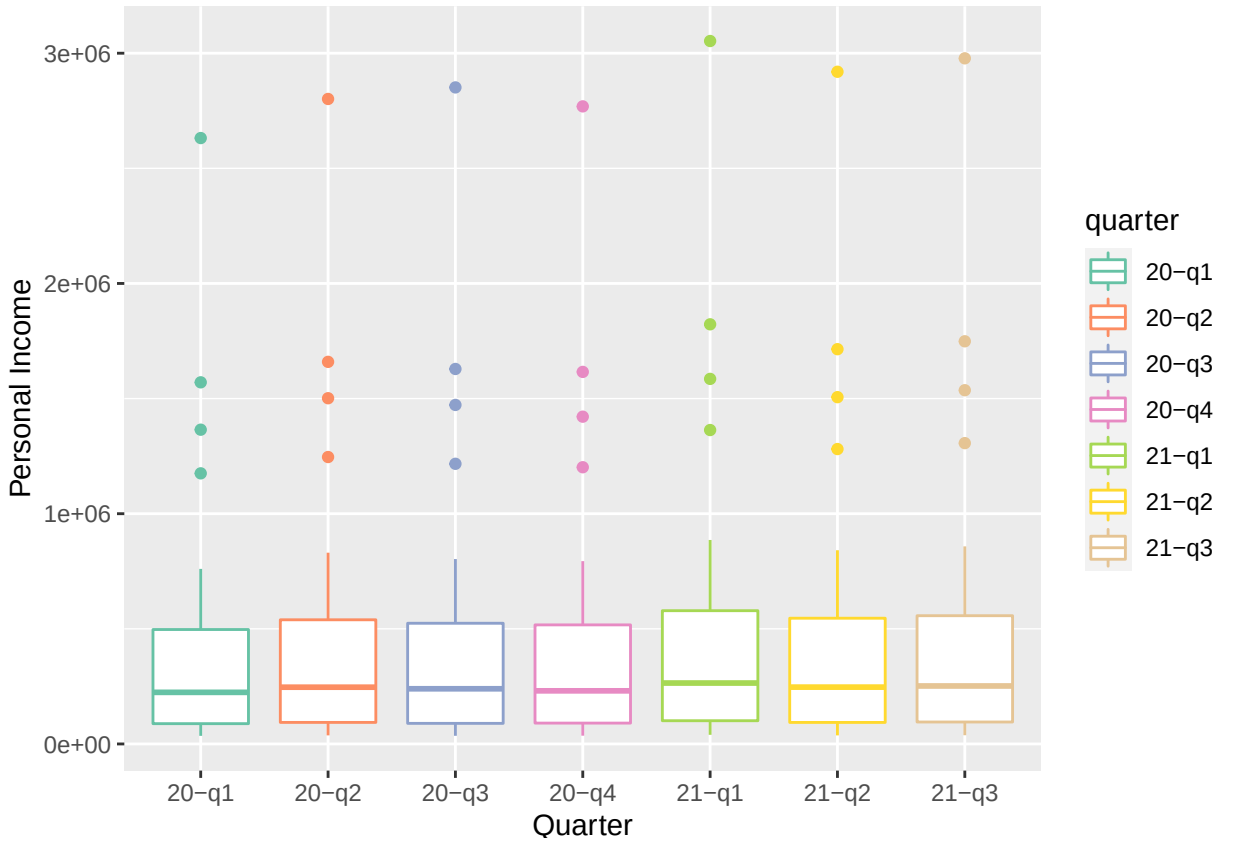
```
p55
```


Vaccination Rate
0.6
0.4
0.2
0.0
20−q1
20−q2
20−q3
20−q4
21−q1
21−q2
21−q3
Quarter
quarter
20−q1
20−q2
20−q3
20−q4
21−q1
21−q2
21−q3

```
p66
```


6e+06
4e+06
2e+06
0e+00
GDP
20−q1
20−q2
20−q3
20−q4
21−q1
21−q2
21−q3
Quarter
quarter
20−q1
20−q2
20−q3
20−q4
21−q1
21−q2
21−q3

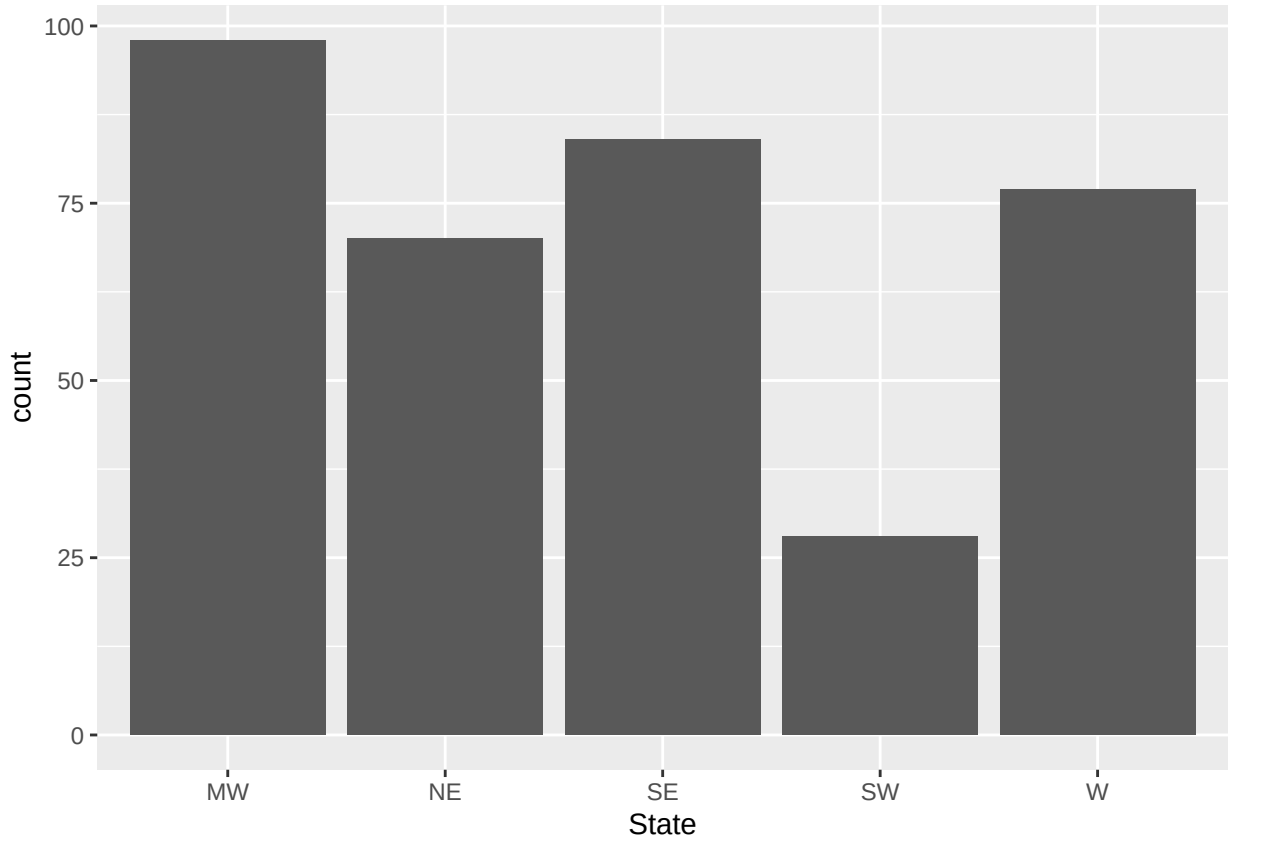
p77



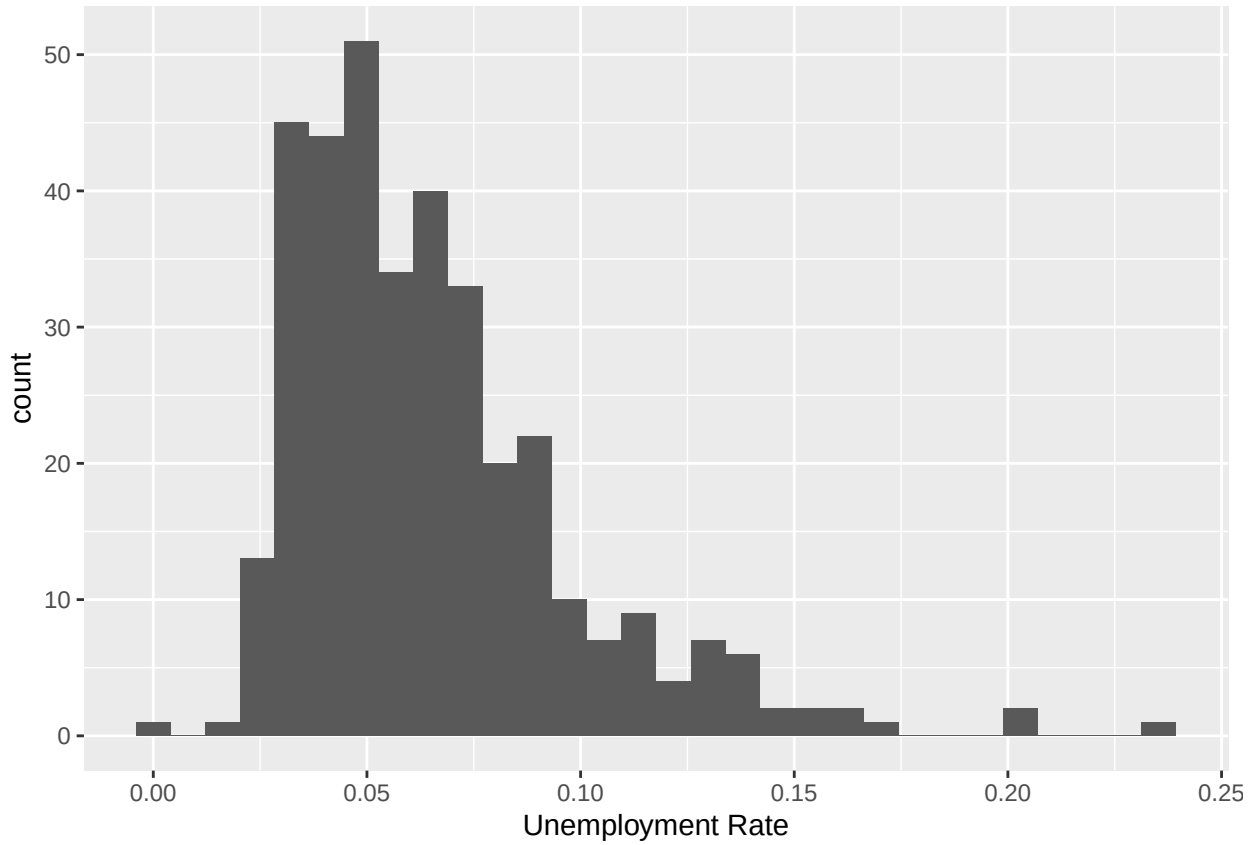
Histogram

```
data$state[data$state == "Southeast"] <-"SE"
data$state[data$state == "West"] <-"W"
data$state[data$state == "Southwest"] <-"SW"
data$state[data$state == "Northeast"] <-"NE"
data$state[data$state == "Midwest"] <-"MW"


p111 = ggplot(data, aes(x=state)) +
  geom_bar() +
  labs(x="State")

p222 = ggplot(data, aes(x=unemployment.rate)) +
  stat_bin(bins = 30) +
  geom_histogram() +
  labs(x="Unemployment Rate")

p333 = ggplot(data, aes(x=new.cases)) +
  stat_bin(bins = 30) +
  geom_histogram() +
  labs(x="New Cases")

p444 = ggplot(data, aes(x=death)) +
  stat_bin(bins = 30) +
  geom_histogram() +
  labs(x="Death")
```

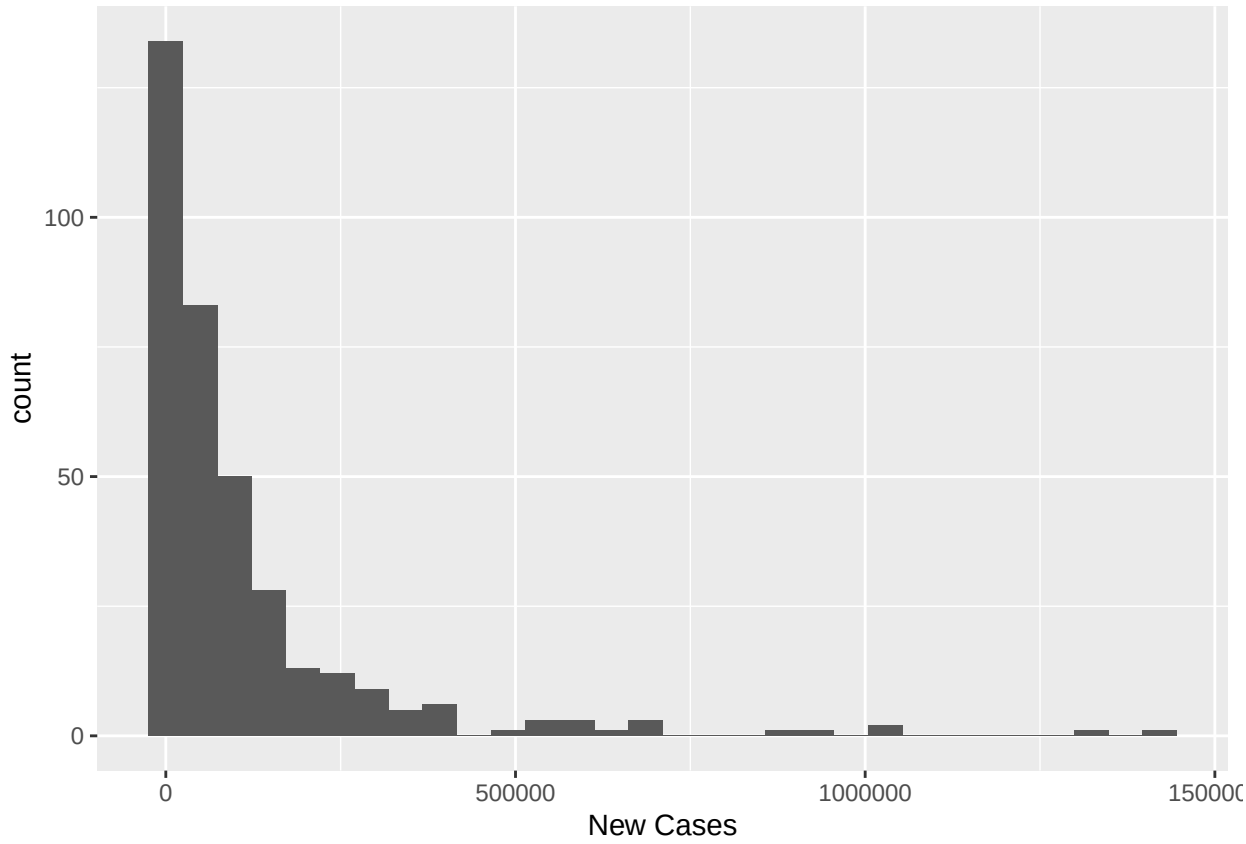
```
p555 = ggplot(data, aes(x=lockdown.days)) +
  stat_bin(bins = 30) +
  geom_histogram() +
  labs(x="Lockdown Days")

p666 = ggplot(data, aes(x=vaccination.rate)) +
  stat_bin(bins = 30) +
  geom_histogram() +
  labs(x="Vaccination Rate")

p777 = ggplot(data, aes(x=GDP.USD.)) +
  stat_bin(bins = 30) +
  geom_histogram() +
  labs(x="GDP")

p888 = ggplot(data, aes(x=Personal.Income)) +
  stat_bin(bins = 30) +
  geom_histogram() +
  labs(x="Personal Income")

p111
```



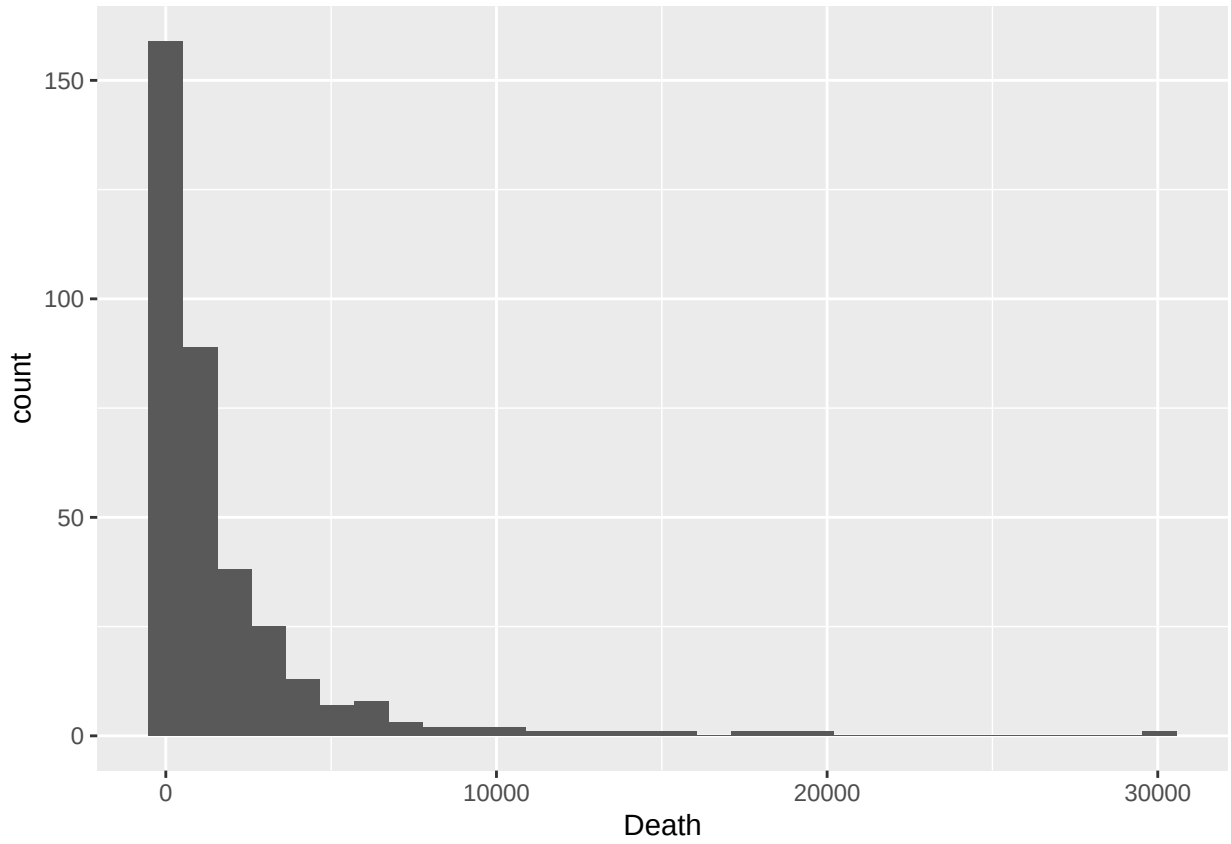

```
p222
```


50
40
30
20
10
0
count
0.00
0.05
0.10
0.15
0.20
0.25
Unemployment Rate

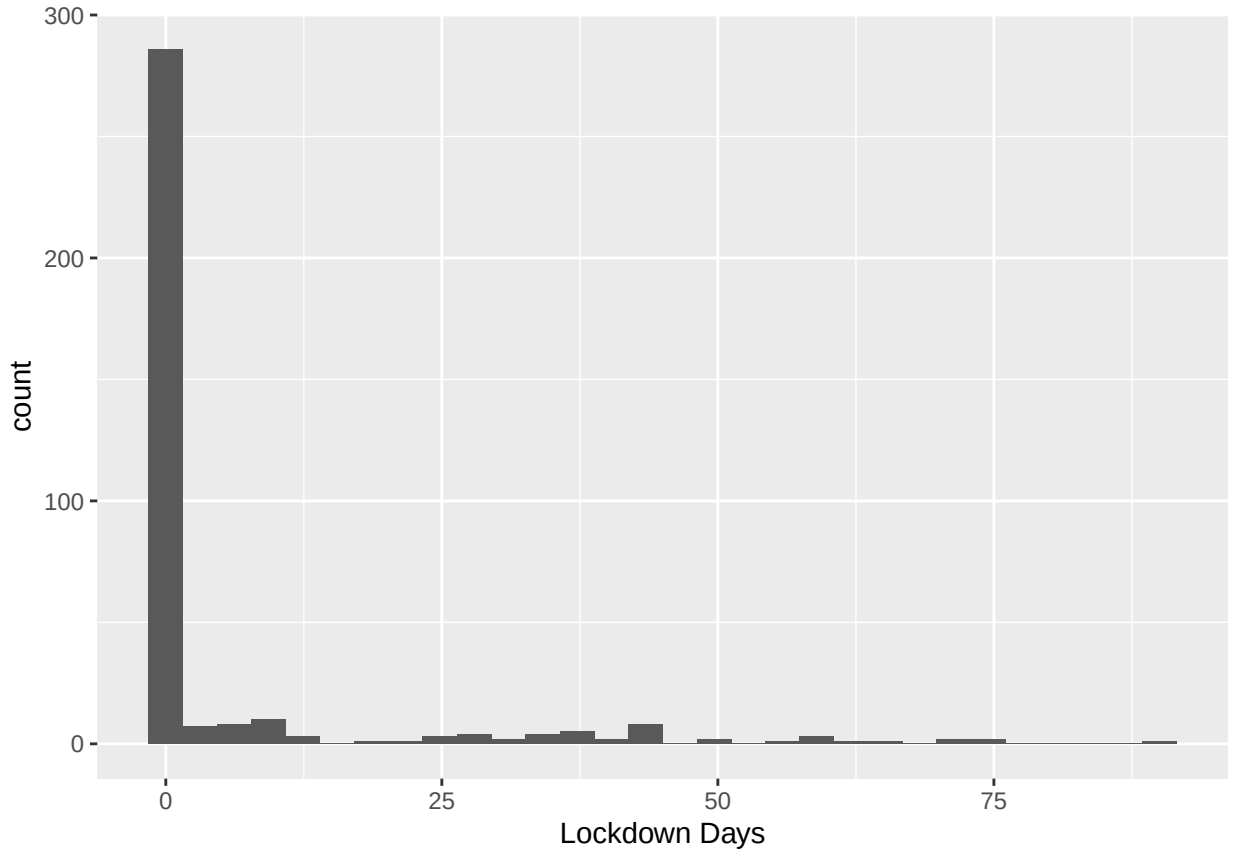
```
p333
```



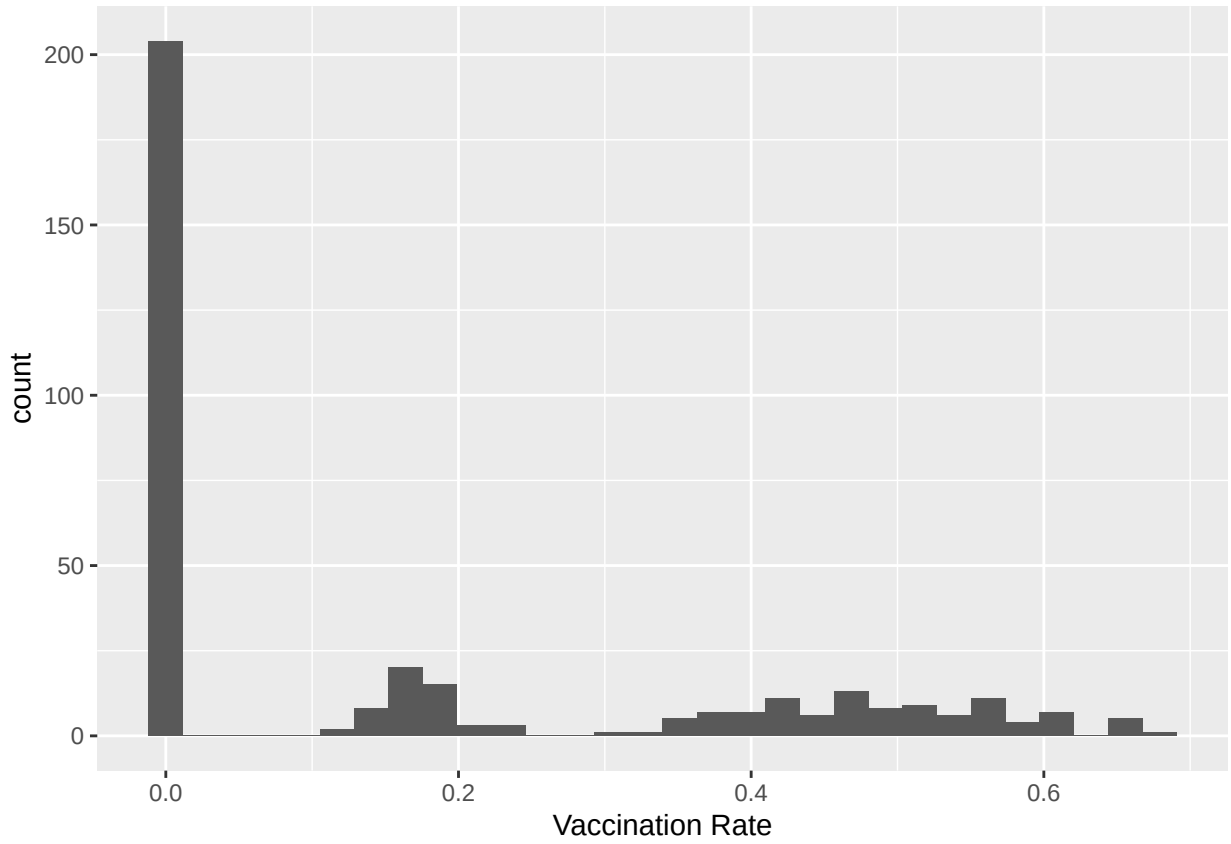

```
p444
```



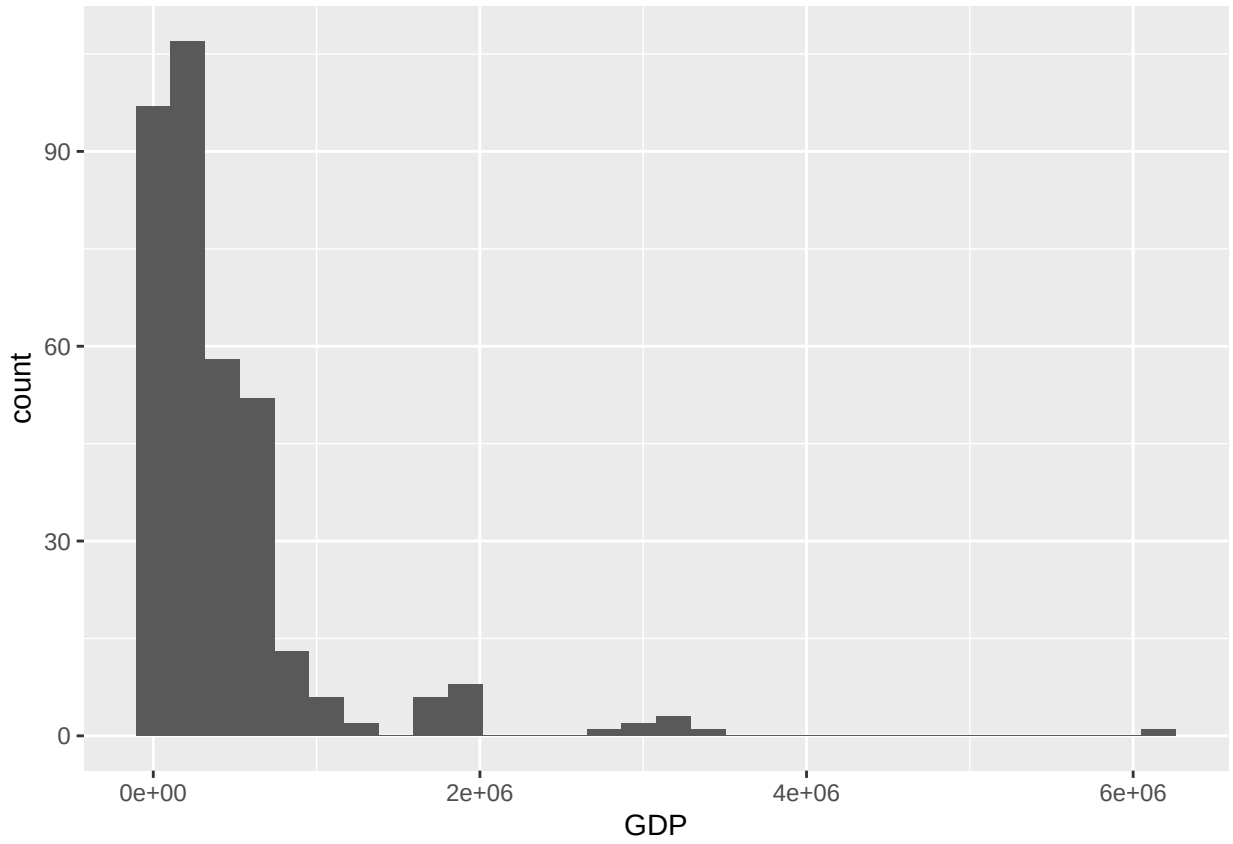

```
p555
```



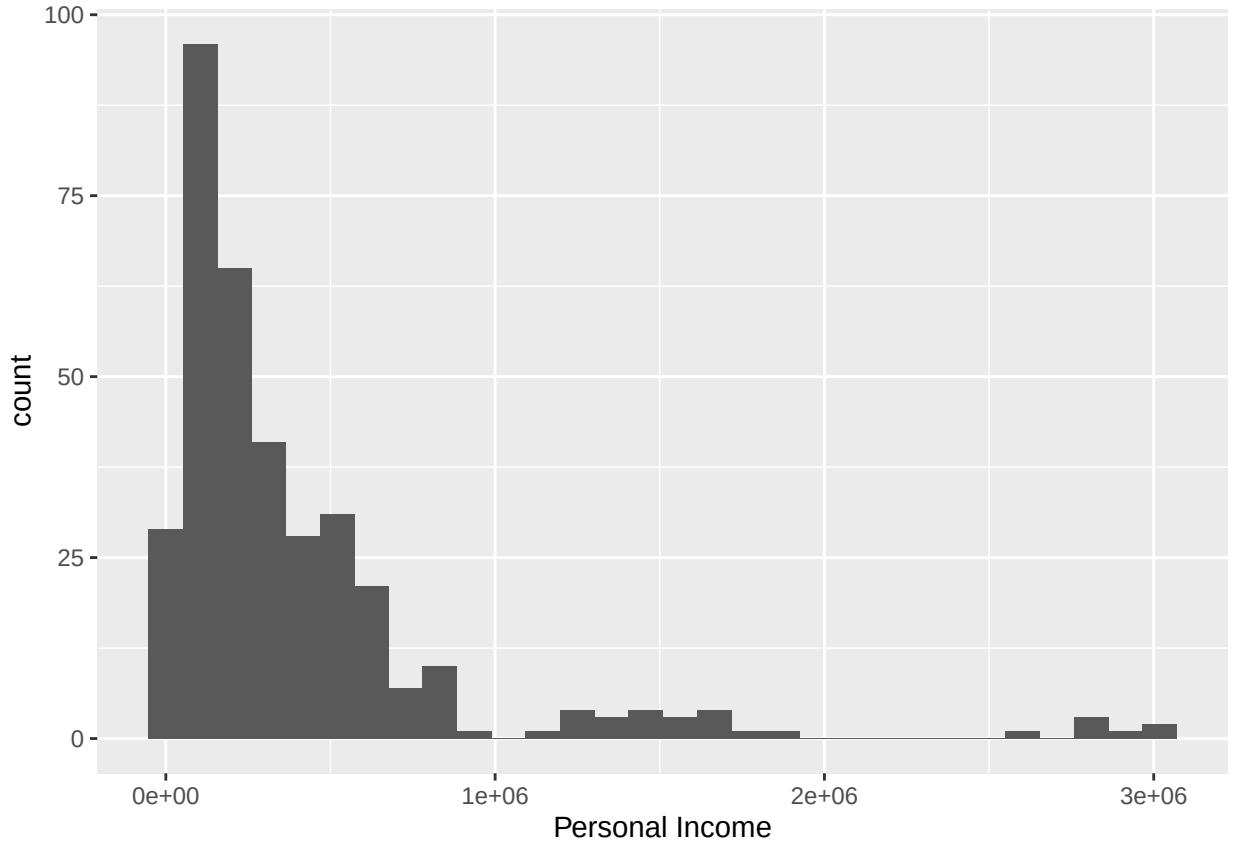

```
p666
```



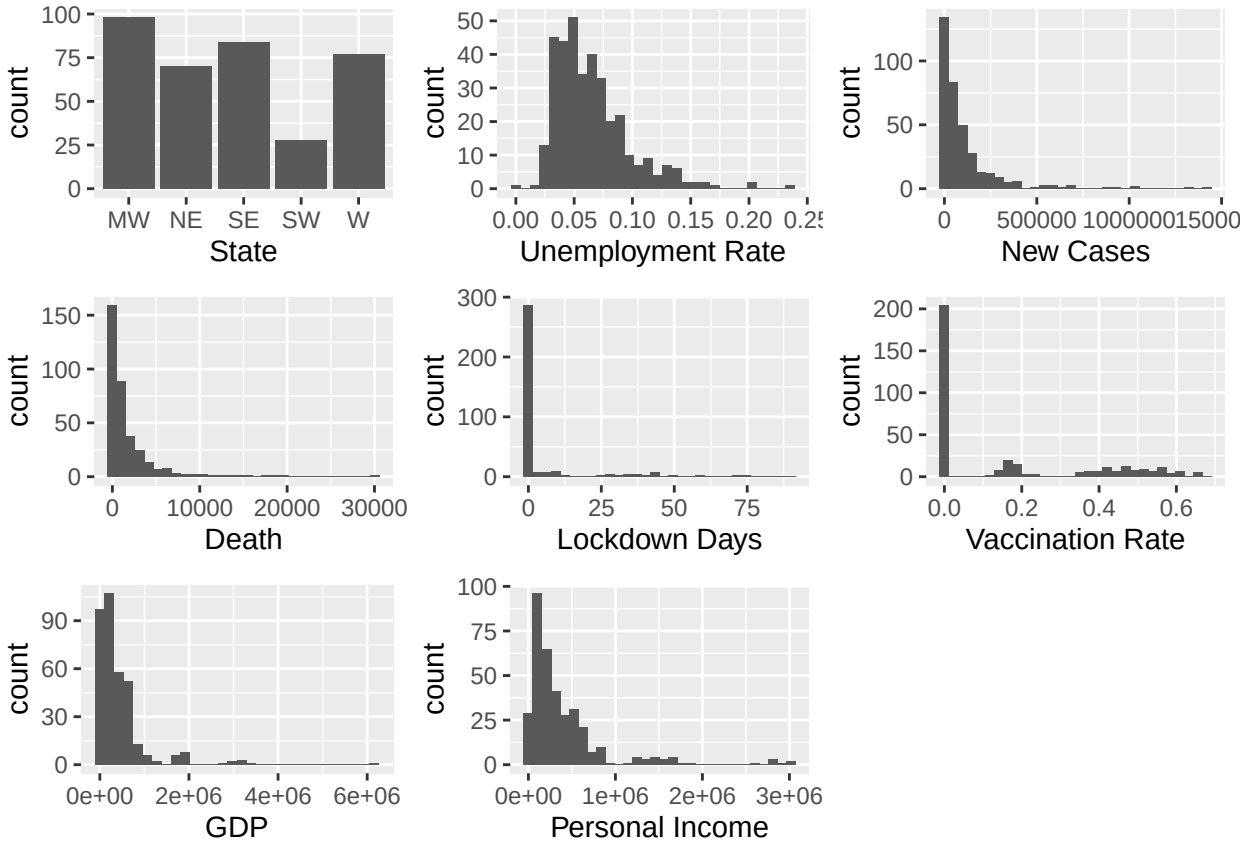

```
p777
```




```
p888
```

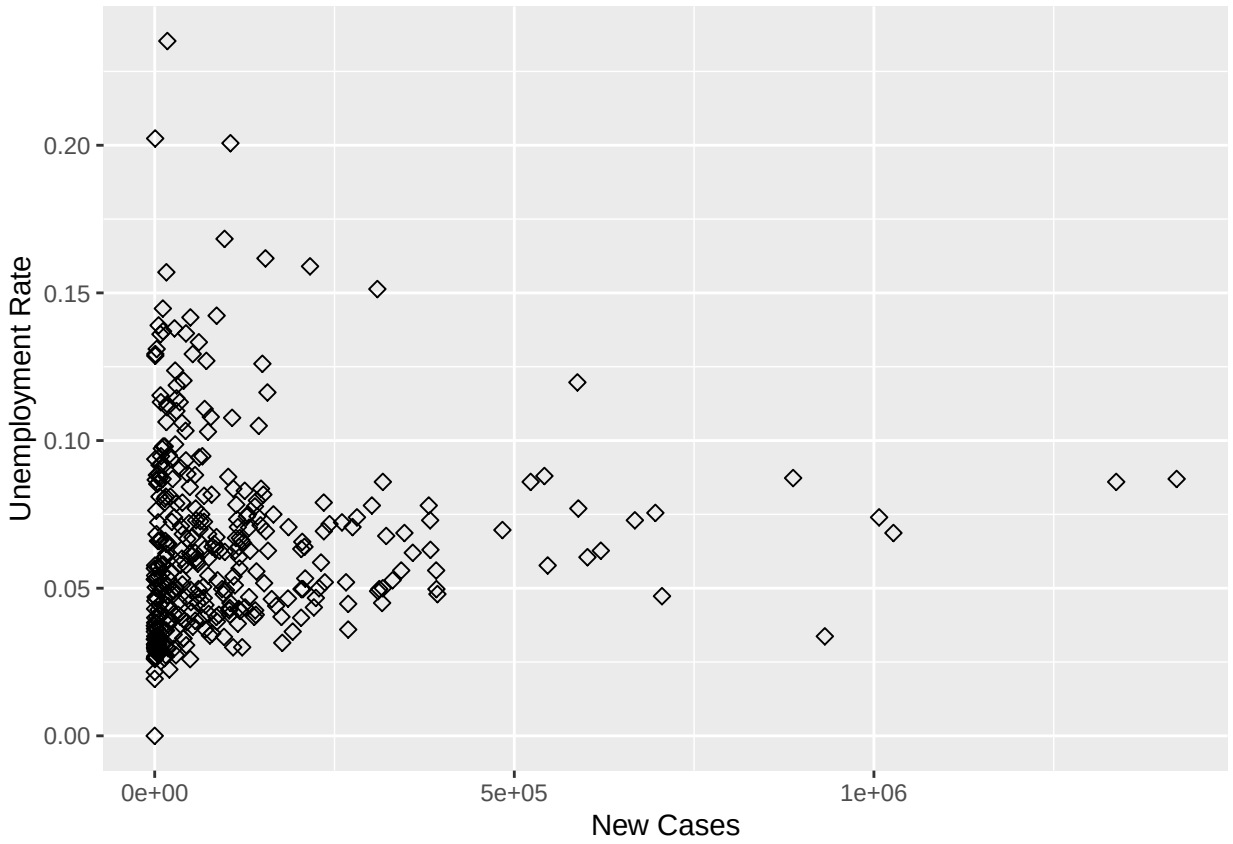



```
ggarrange(p111,p222,p333,p444,p555,p666,p777,p888, nrow = 3, ncol = 3)
```

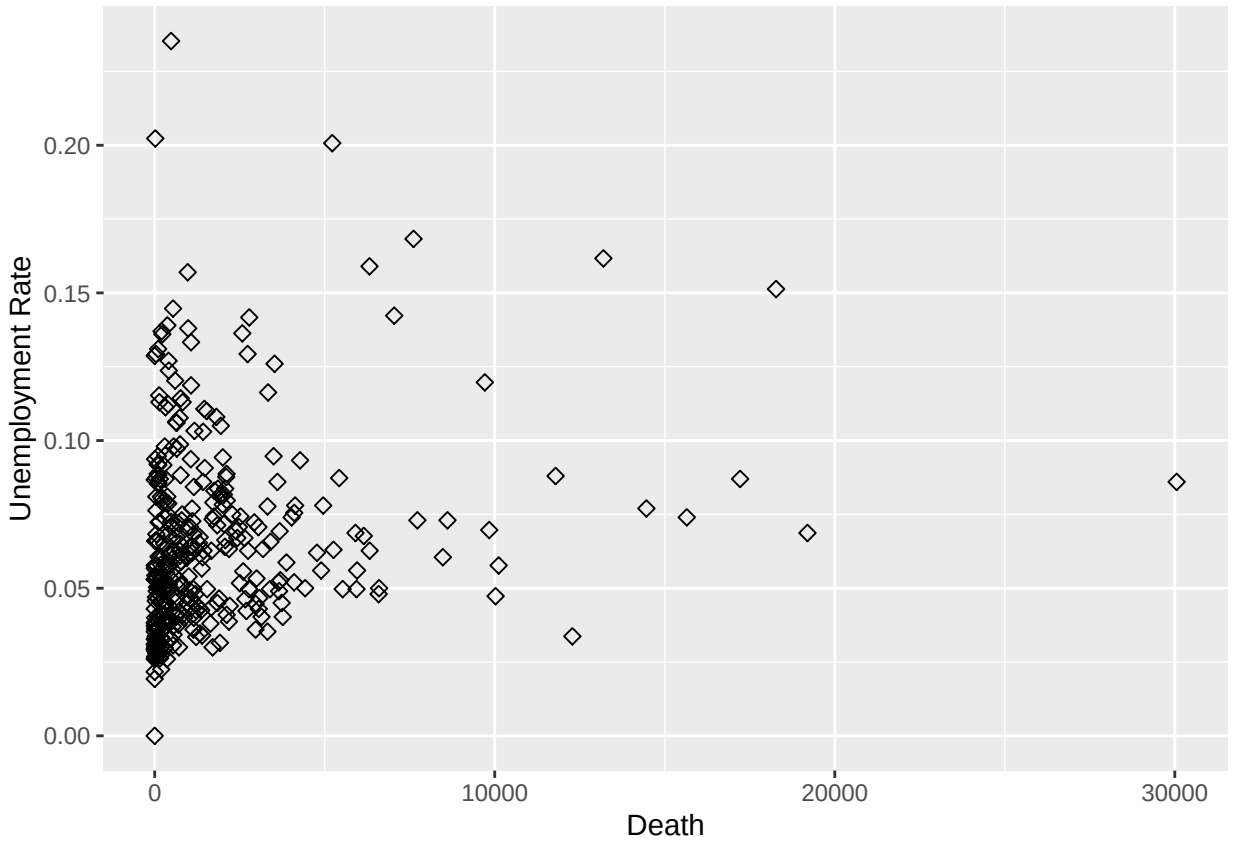


Scatter Plots

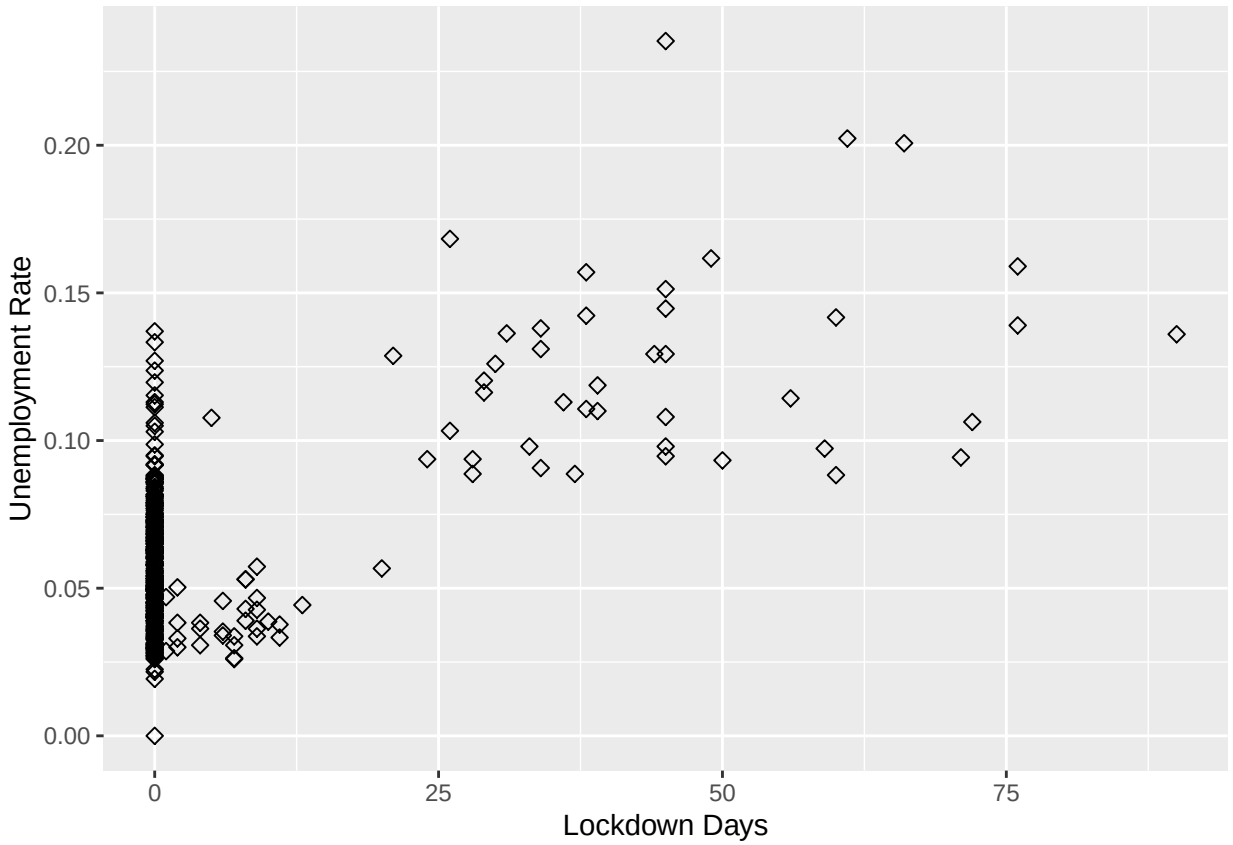
```
p1111 = ggplot(data, aes(x=new.cases, y=unemployment.rate,)) +
  geom_point(size=2, shape=23) +
  labs(x = "New Cases", y = "Unemployment Rate") +
  scale_color_brewer(palette = "Set2")

p2222 = ggplot(data, aes(x=death, y=unemployment.rate)) +
  geom_point(size=2, shape=23) +
  labs(x = "Death", y = "Unemployment Rate") +
  scale_color_brewer(palette = "Set2")

p3333 = ggplot(data, aes(x=lockdown.days, y=unemployment.rate)) +
  geom_point(size=2, shape=23) +
  labs(x = "Lockdown Days", y = "Unemployment Rate") +
  scale_color_brewer(palette = "Set2")

p4444 = ggplot(data, aes(x=vaccination.rate, y=unemployment.rate)) +
  geom_point(size=2, shape=23) +
  labs(x = "Vaccination Rate", y = "Unemployment Rate") +
  scale_color_brewer(palette = "Set2")

p1111
```


0.20
0.15
0.10
0.05
0.00
Unemployment Rate
0e+00
5e+05
1e+06
New Cases

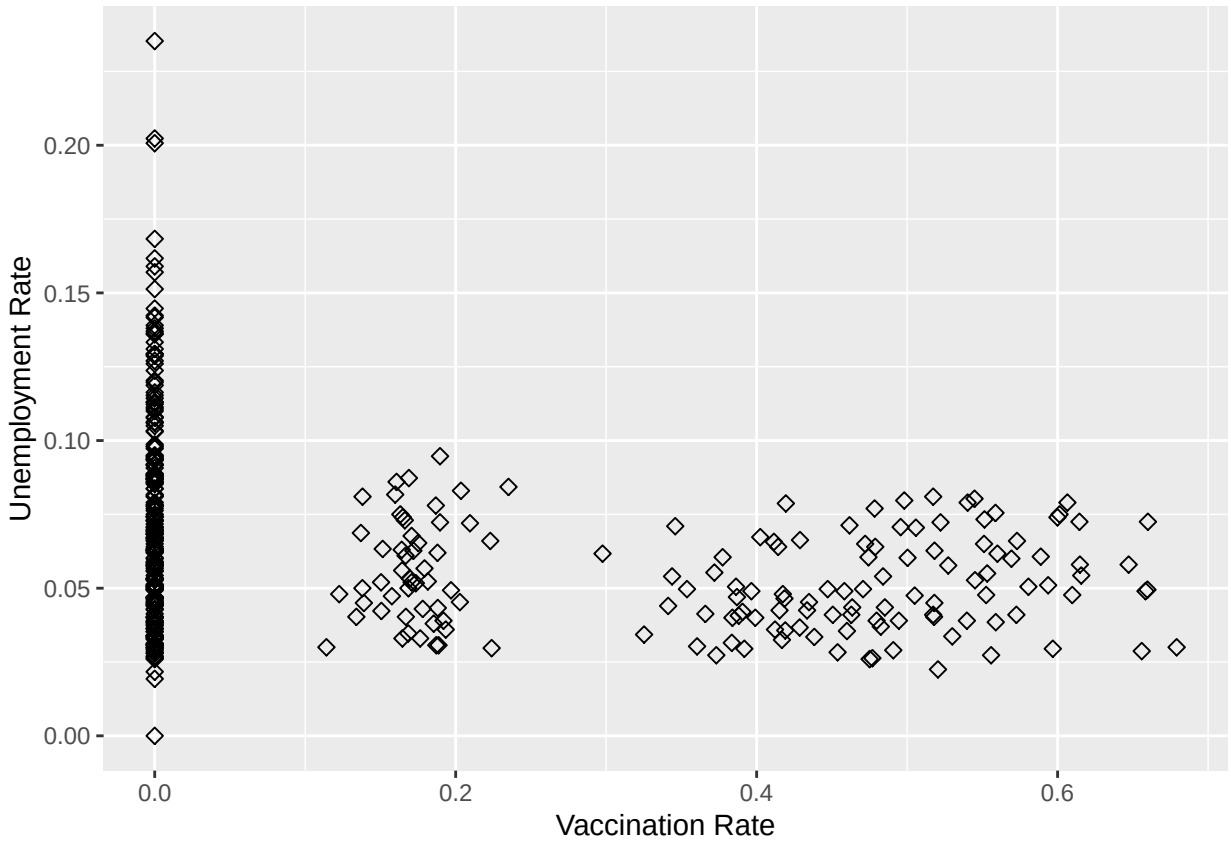
```
p2222
```


Unemployment Rate
0.20
0.15
0.10
0.05
0.00
0
10000
20000
30000
Death

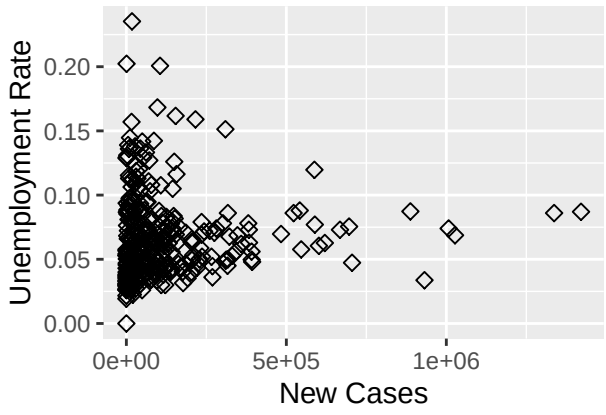
```
p3333
```

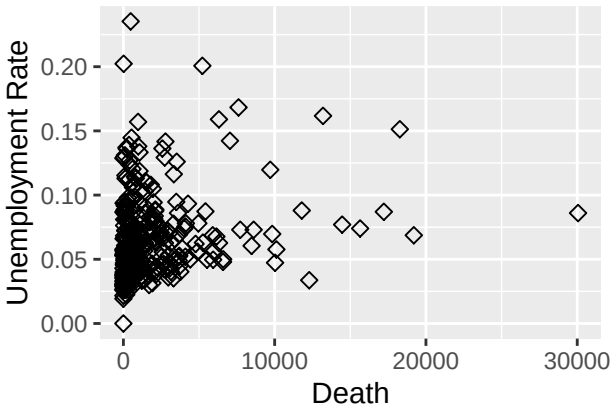

Unemployment Rate
0.20
0.15
0.10
0.05
0.00
0
25
50
75
Lockdown Days

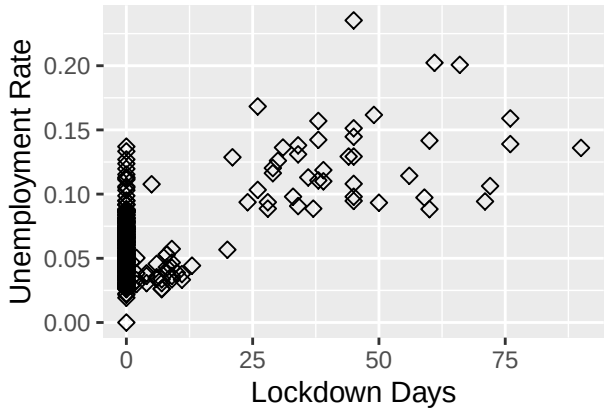
```
p4444
```

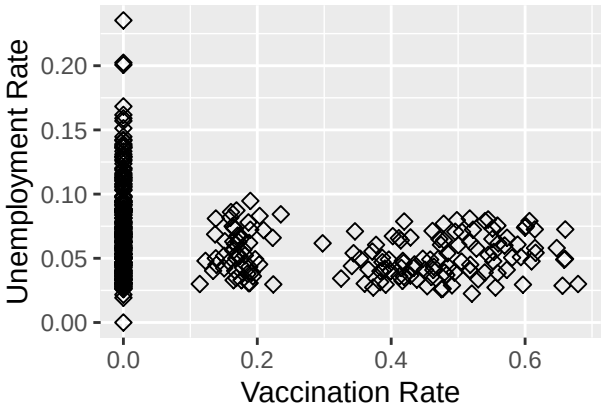



```
ggarrange(p1111,p2222,p3333,p4444, nrow = 2, ncol = 2)
```


Unemployment Rate
0.20
0.15
0.10
0.05
0.00
0e+00
5e+05
1e+06
New Cases


Unemployment Rate
0.20
0.15
0.10
0.05
0.00
0
10000
20000
30000
Death


Unemployment Rate
0.20
0.15
0.10
0.05
0.00
0
25
50
75
Lockdown Days


Unemployment Rate
0.20
0.15
0.10
0.05
0.00
0.0
0.2
0.4
0.6
Vaccination Rate