Assignment3_XinHuang

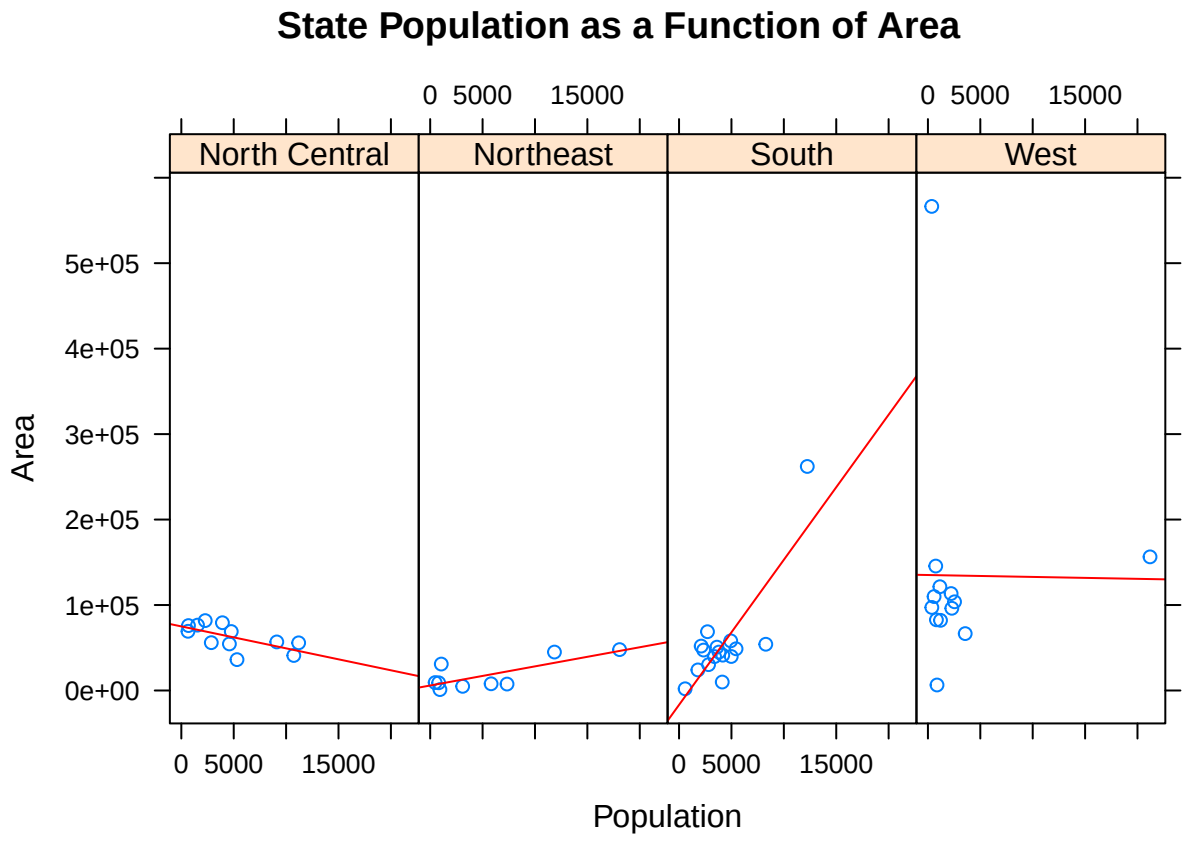
Xin Huang

Answer for Question 1

1a

Run the code to generate the figure:

```
#generate the data
myState <-as.data.frame(cbind(state.x77, region = state.region))
myState <- cbind(myState, regionName = levels(state.region)[state.region])
myState$StateName <- rownames(myState)
colnames(myState)[6] <- "HSgrad"

xyplot(Area ~ Population | as.factor(regionName),
       main = "State Population as a Function of Area",
       myState,
       layout = c(4, 1),
       panel = function(x, y, ...) {
           panel.xyplot(x, y, ...)
           panel.lmline(x, y, col = 2)
           })
```



1b

Run the code to generate the figure:

```
westState <- myState %>%
    subset(regionName == 'West') %>%
    mutate(density = Population / Area)
westState[order(westState$density), 11: 12]
```

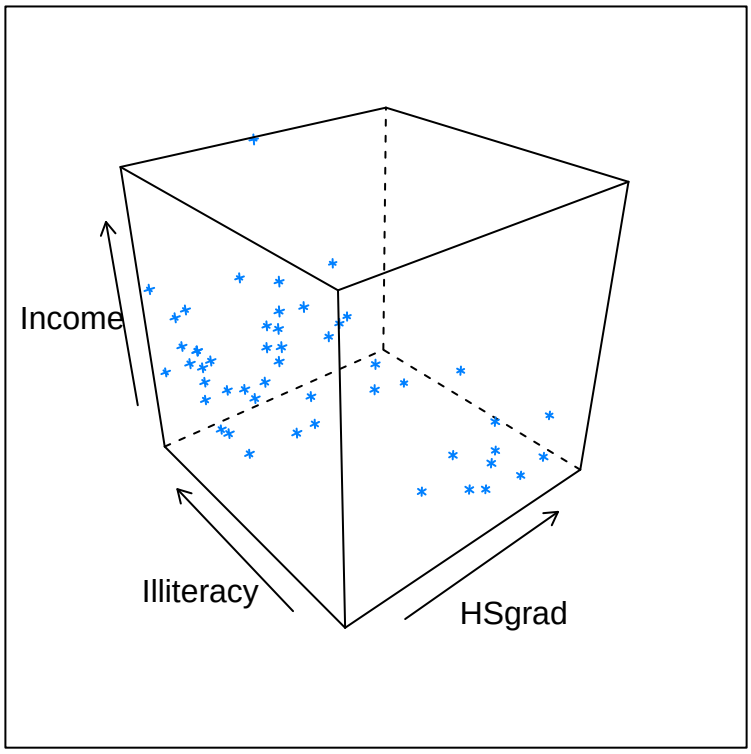
```
##     StateName      density
## 1      Alaska 0.0006443845
## 13    Wyoming 0.0038681934
## 7     Montana 0.0051240839
## 8      Nevada 0.0053690542
## 9  New Mexico 0.0094224624
## 6       Idaho 0.0098334482
## 11       Utah 0.0146535763
## 2     Arizona 0.0195032491
## 10     Oregon 0.0237461532
## 4    Colorado 0.0244877898
## 12 Washington 0.0534625207
## 5      Hawaii 0.1350972763
## 3  California 0.1355708904
```

That state is: California

1c

Run the code to generate the figure:

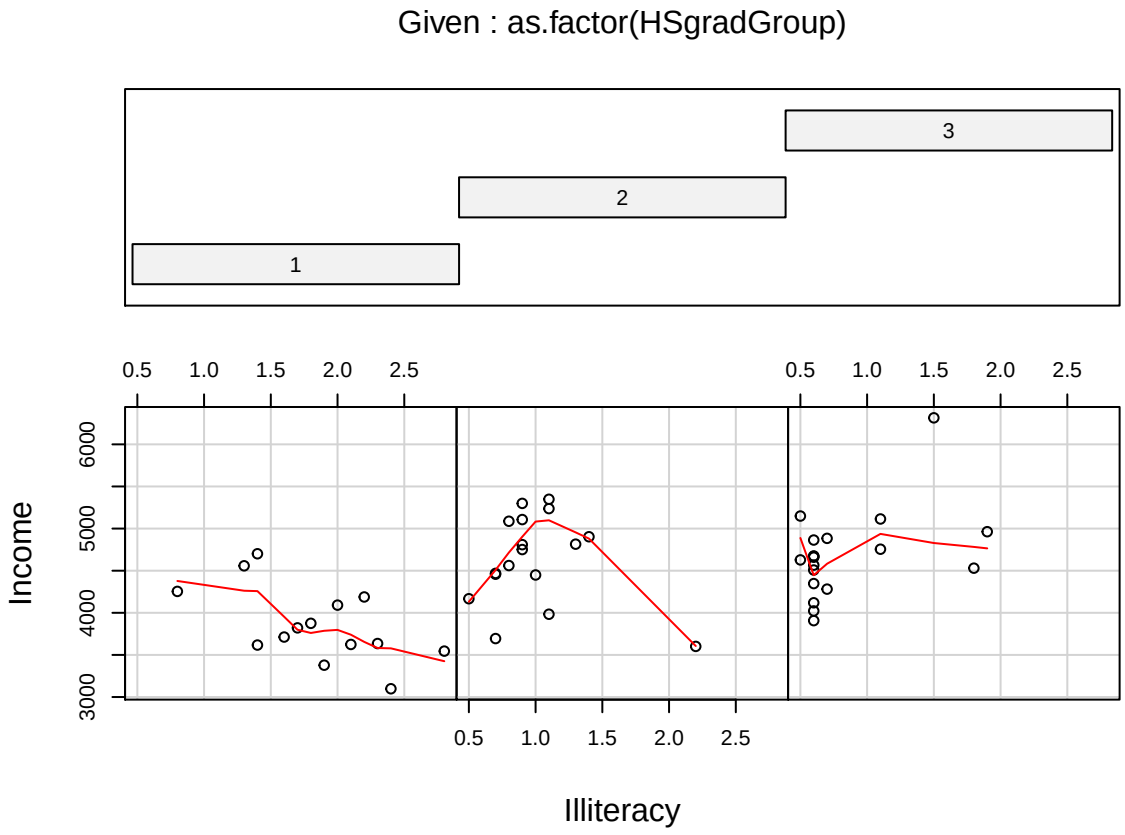
```
cloud(myState$Income ~ myState$Illiteracy * myState$HSgrad,
      xlab = "HSgrad",
      ylab = "Illiteracy",
      zlab = "Income",
      main = "3D plot of States",
      distance = .4)
```



1d

Run the code to generate the figure:

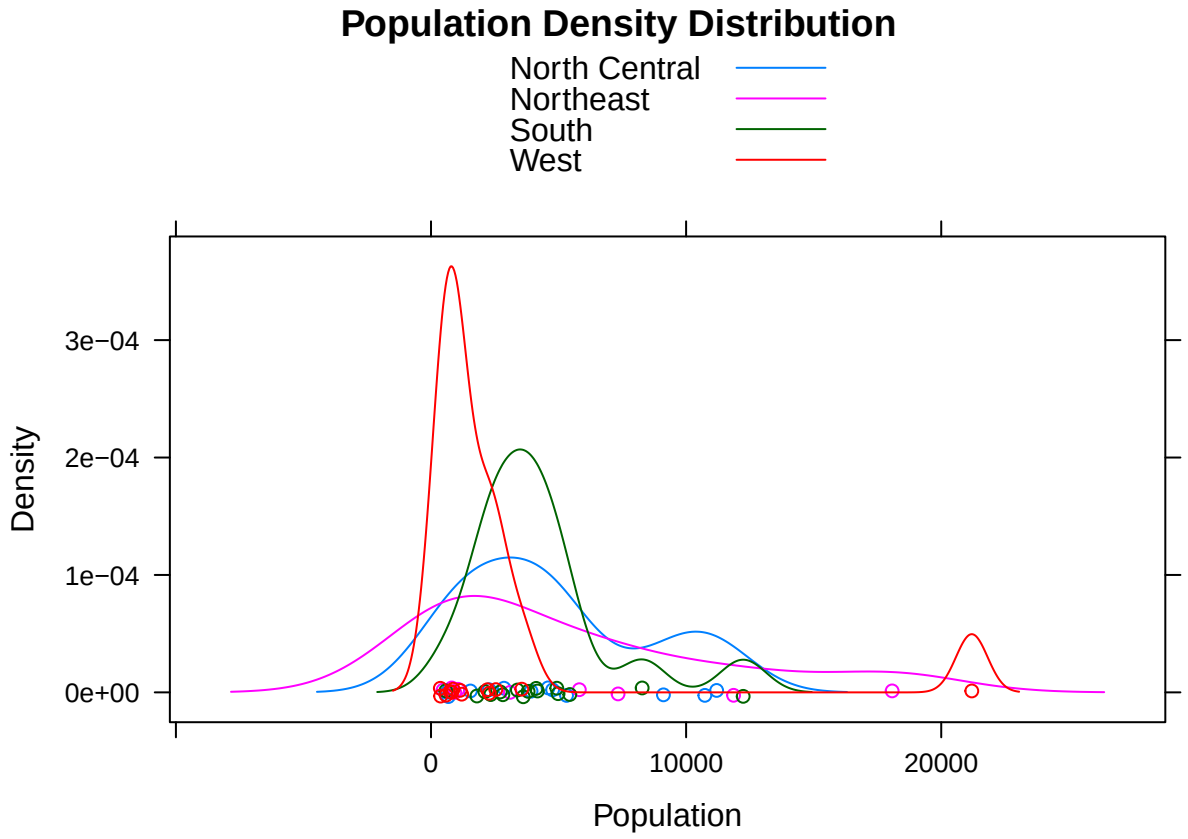
```
# sort HS Grad
group1 <- myState %>%
    subset(HSgrad < 50) %>%
    mutate(HSgradGroup = 1)

group2 <- myState %>%
    subset(HSgrad >= 50 & HSgrad <= 57) %>%
    mutate(HSgradGroup = 2)

group3 <- myState %>%
    subset(HSgrad > 57) %>%
    mutate(HSgradGroup = 3)

newState <- rbind(group1, group2, group3)
coplot(Income ~ Illiteracy | as.factor(HSgradGroup),
       data = newState,
       panel = panel.smooth, rows = 1)
```



1e

Use `densityplot()` to plot the figure:

```
densityplot(~ Population, data = myState, group = regionName,
            auto.key = TRUE, lwd=1,
            main="Population Density Distribution")
```

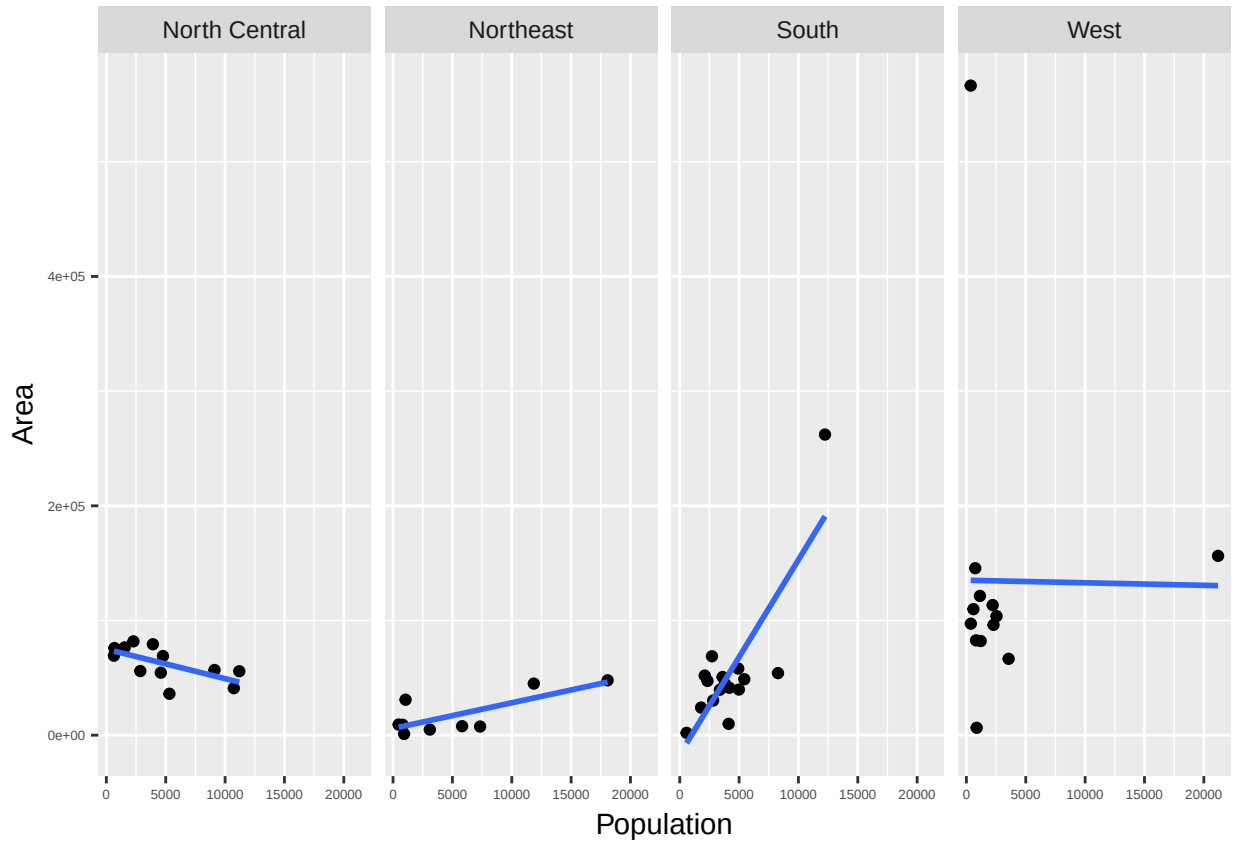


Answer for Question 2

2a

Run the code to generate the figure:

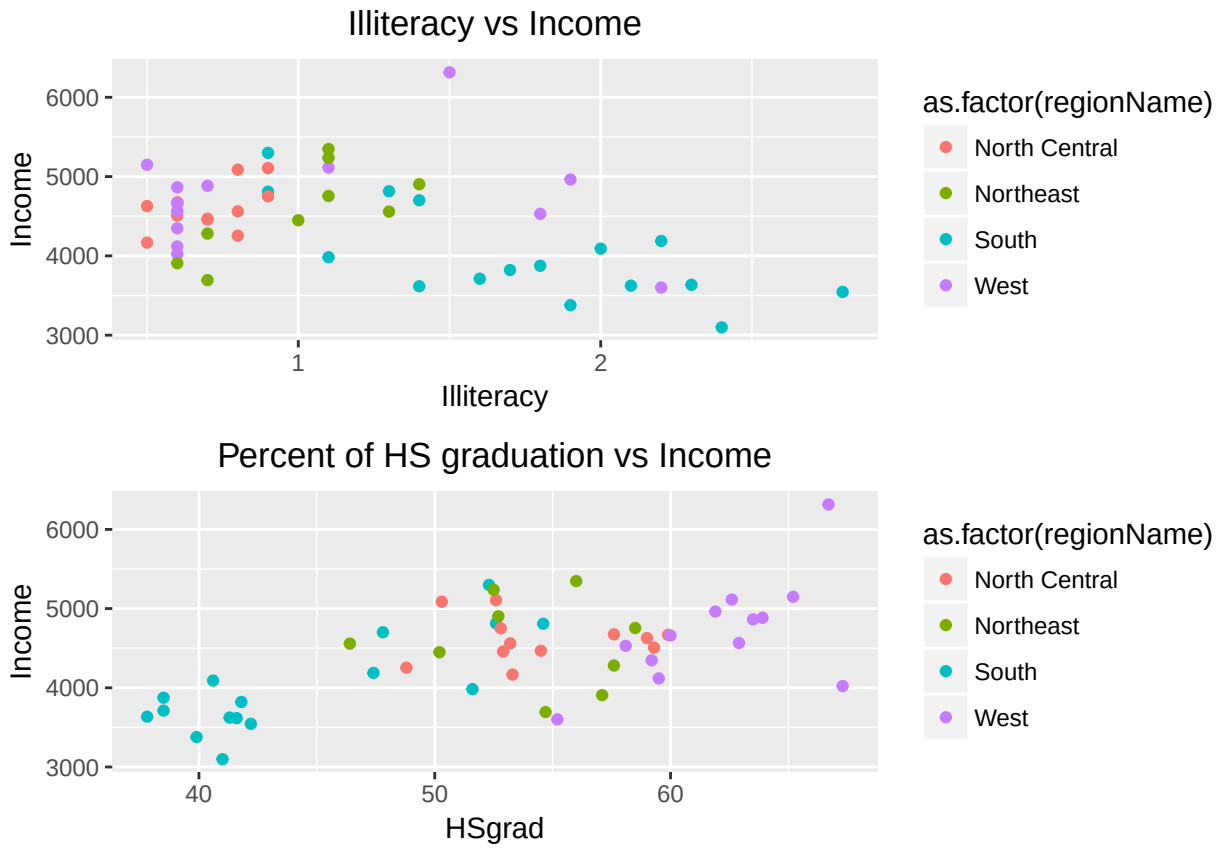
```
myState <-as.data.frame(cbind(state.x77, region = state.region))
myState <- cbind(myState, regionName = levels(state.region)[state.region])
myState$StateName <- rownames(myState)
colnames(myState)[6] <- "HSgrad"

qplot(Population, Area, data = myState) +
    geom_smooth(method = "lm",  se = FALSE) +
    facet_grid(~regionName) +
    theme(axis.text=element_text(size=5))
```



2b

Run the code to generate the figure:

```
p1 <- qplot(Illiteracy, Income, data = myState,
            col = as.factor(regionName),
            main = "Illiteracy vs Income") +
            theme(plot.title = element_text(hjust = 0.5))
p2 <- qplot(HSgrad, Income, data = myState,
            col = as.factor(regionName),
            main = "Percent of HS graduation vs Income") +
            theme(plot.title = element_text(hjust = 0.5))
grid.arrange(p1, p2, ncol = 1, nrow = 2)
```

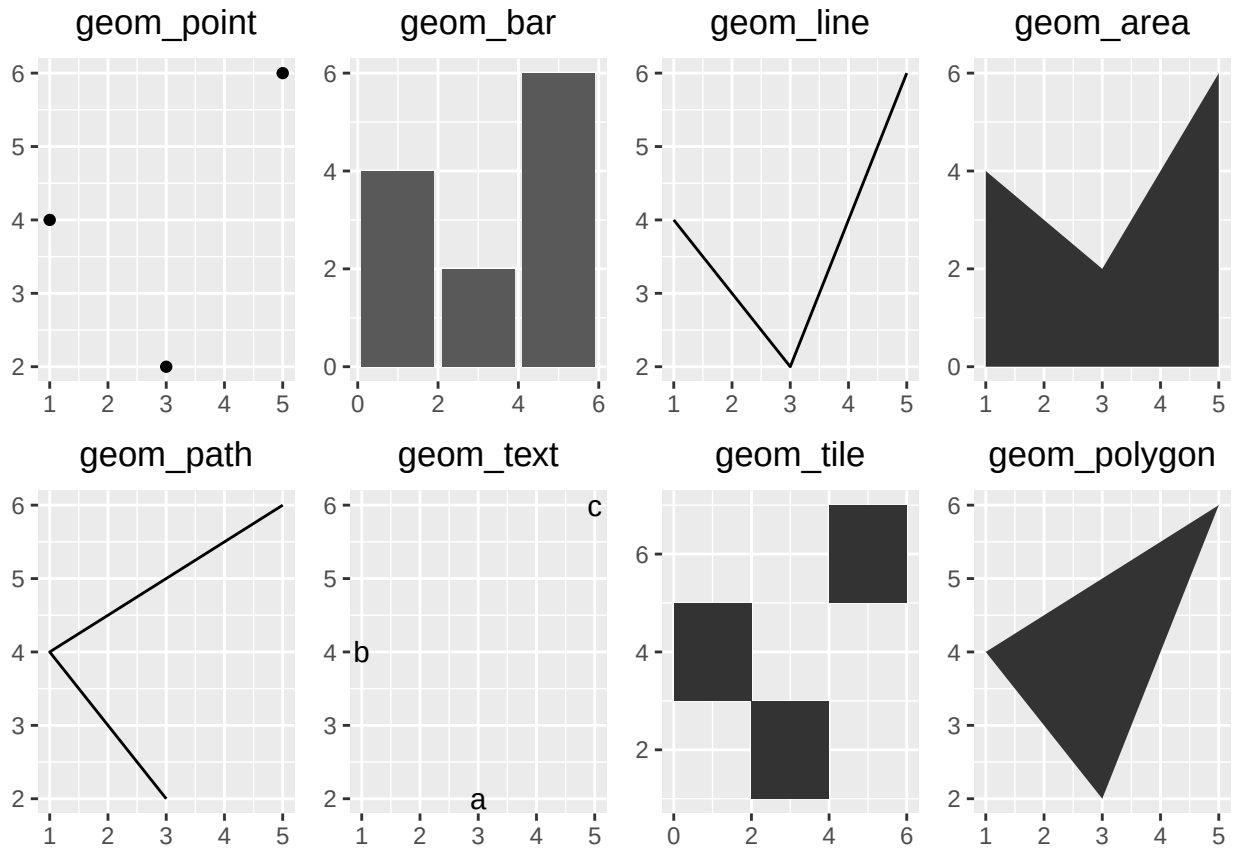


Answer for Question 3

Run the code to generate the figure:

```
df<- data.frame(
    x = c(3,1,5),
    y = c(2,4,6),
    label = c("a","b","c"))
myPlot <- ggplot(df, aes(x, y, label = label)) +
    xlab(NULL) +
    ylab(NULL)

p1 <- myPlot + geom_point() + ggtitle("geom_point") +
    theme(plot.title = element_text(hjust = 0.5))
p2 <- myPlot + geom_bar(stat="identity") + ggtitle("geom_bar") +
    theme(plot.title = element_text(hjust = 0.5))
p3 <- myPlot + geom_line() + ggtitle("geom_line") +
    theme(plot.title = element_text(hjust = 0.5))
p4 <- myPlot + geom_area() + ggtitle("geom_area") +
    theme(plot.title = element_text(hjust = 0.5))
p5 <- myPlot + geom_path() + ggtitle("geom_path") +
    theme(plot.title = element_text(hjust = 0.5))
p6 <- myPlot + geom_text() + ggtitle("geom_text") +
    theme(plot.title = element_text(hjust = 0.5))
p7 <- myPlot + geom_tile() + ggtitle("geom_tile") +
```

```
    theme(plot.title = element_text(hjust = 0.5))
p8 <- myPlot + geom_polygon() + ggtitle("geom_polygon") +
    theme(plot.title = element_text(hjust = 0.5))
grid.arrange(p1, p2, p3, p4,
             p5, p6, p7, p8,
             ncol = 4, nrow = 2)
```

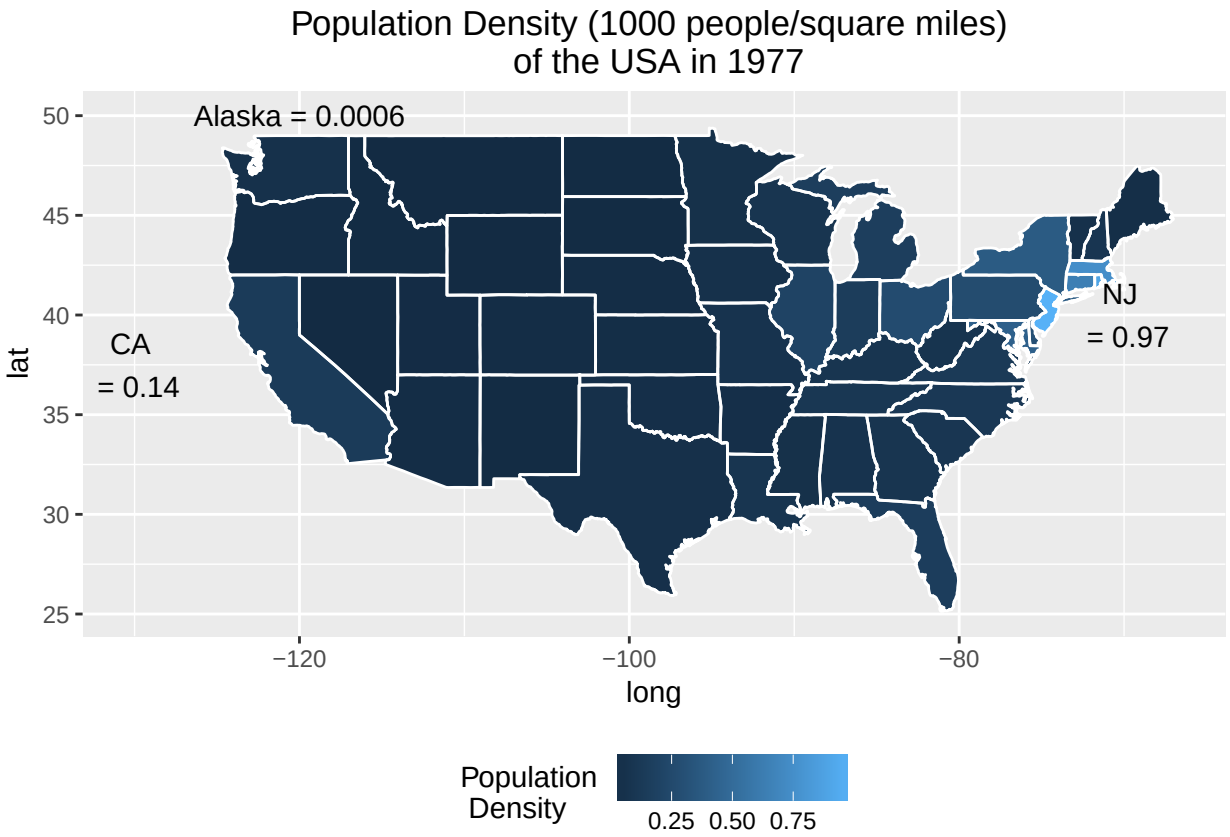


Answer for Question 4

Run the code to generate the figure:

```
myState <-as.data.frame(cbind(state.x77, region = state.region))
myState <- cbind(myState, regionname = levels(state.region)[state.region])

colnames(myState) <- tolower(colnames(myState))
myState$statename <- tolower(rownames(myState))
myState$`Population \n Density` <- myState$population / myState$area
colnames(myState)[6] <- "hsgrad"
colnames(myState)[4] <- "lifeexp"

stateMap <- map_data("state")
colnames(stateMap)[5] <- "statename"
myStateMap <- merge(stateMap, myState)
orState <- myStateMap[order(myStateMap$order), ]
```

```
ggplot(orState, aes(long, lat, group = group, fill = `Population \n Density`)) +
    geom_polygon() +
    borders("state", colour= "white", size = 0.5) +
    annotate("text",label = "NJ \n = 0.97", x = -70, y = 40) +
    annotate("text",label = "CA \n = 0.14", x = -130, y = 37.5) +
    annotate("text",label = "Alaska = 0.0006", x = -120, y = 50) +
    ggtitle("Population Density (1000 people/square miles) \n of the USA in 1977") +
    theme(plot.title = element_text(hjust = 0.5),
          legend.position="bottom")
```

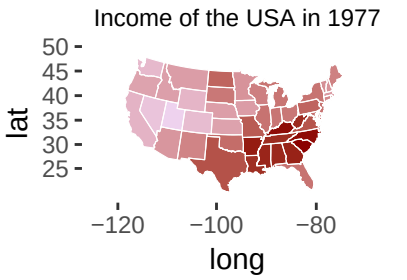


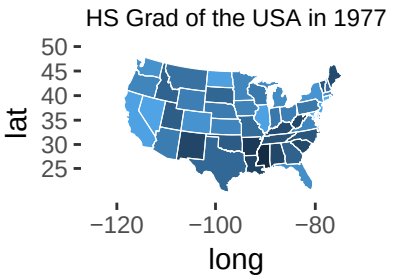
Answer for Question 5

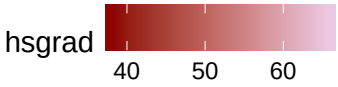
Run the code to generate the figure:

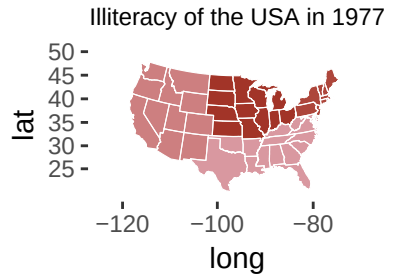
```
#plot map grouped by state and filled with hsgrad
p1 <- ggplot(orState, aes(long, lat, group = group, fill = hsgrad)) +
    geom_polygon() +
    borders("state", colour= "white", size = 0.2) +
    coord_map("albers", lat0 = 39, lat1 = 45) +
    ggtitle("Income of the USA in 1977") +
    theme(plot.title = element_text(hjust = 0.5, size = 9),
          legend.position="bottom",
          panel.background = element_blank()) +
    scale_fill_continuous(low='darkred', high='thistle2', guide='colorbar')

#plot map grouped by state and filled with income
```

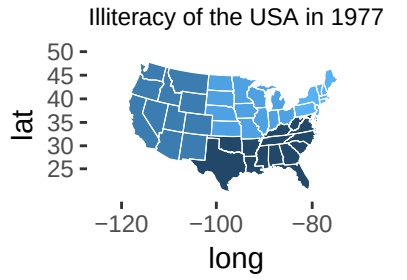
```r
p2 <- ggplot(orState, aes(long, lat, group = group, fill = income)) +
    geom_polygon() +
    borders("state", colour= "white", size = 0.2) +
    coord_map("albers", lat0 = 39, lat1 = 45) +
    ggtitle("HS Grad of the USA in 1977") +
    theme(plot.title = element_text(hjust = 0.5, size = 9),
          legend.position="bottom",
          panel.background = element_blank())

#plot map grouped by region and filled with illiteracy
p3 <- ggplot(orState, aes(long, lat, map_id = region, fill = illiteracy)) +
    geom_map(map = orState, data=orState) +
    borders("state", colour= "white", size = 0.2) +
    coord_map("albers", lat0 = 39, lat1 = 45) +
    ggtitle("Illiteracy of the USA in 1977") +
    theme(plot.title = element_text(hjust = 0.5, size = 9),
          legend.position="bottom",
          panel.background = element_blank()) +
    scale_fill_continuous(low='darkred', high='thistle2', guide='colorbar')

#plot map grouped by region and filled with income
p4 <- ggplot(orState, aes(long, lat, map_id = region, fill = income)) +
    geom_map(map = orState, data=orState) +
    borders("state", colour= "white", size = 0.2) +
    coord_map("albers", lat0 = 39, lat1 = 45) +
    ggtitle("Illiteracy of the USA in 1977") +
    theme(plot.title = element_text(hjust = 0.5, size = 9),
          legend.position="bottom",
          panel.background = element_blank())

grid.arrange(p1, p2, p3, p4, ncol = 2, nrow = 2)
```

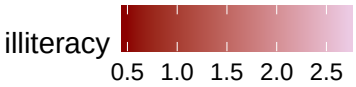












By comparing those four figures, you can see that:
1. The higher rate of hsGrad of one state, the higer income it has 2. The lower illiteracy of one big region, the higher income it has