Chaudhary_Homework_6

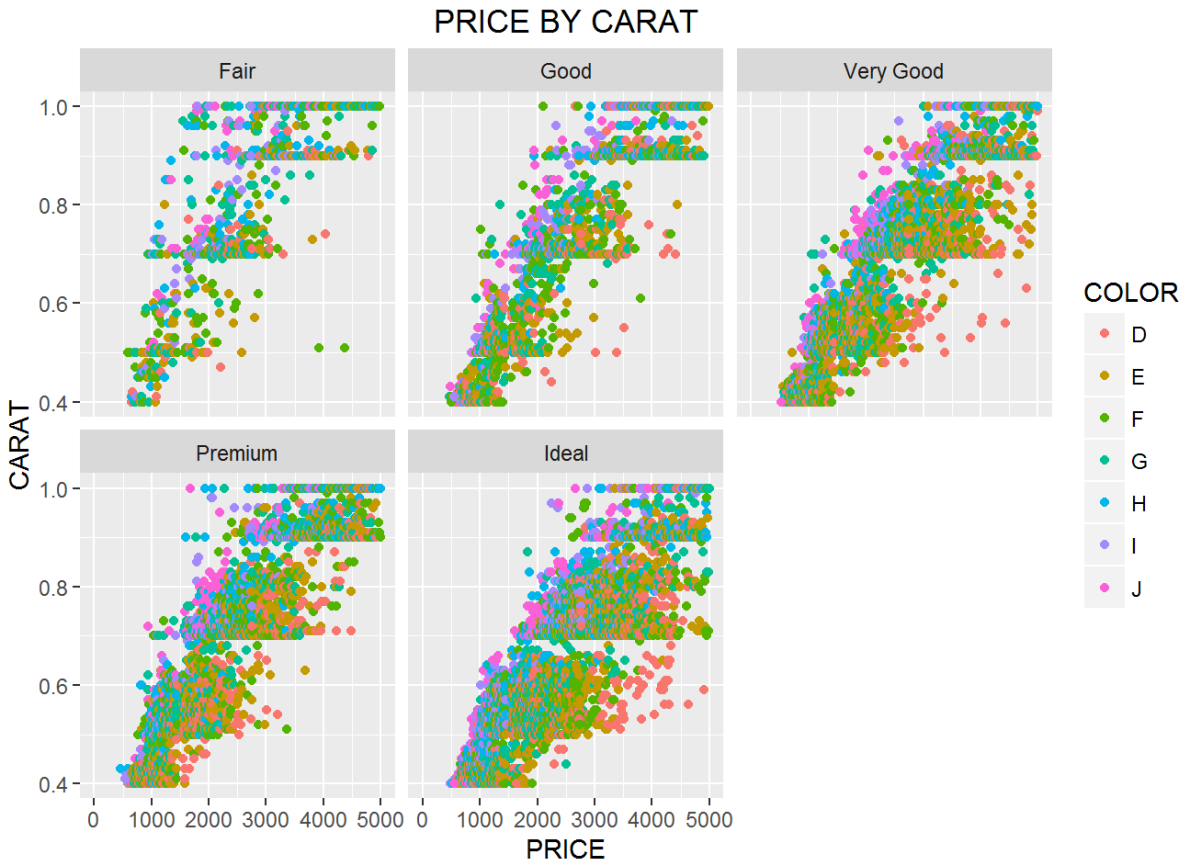
Jyoti Chaudhary

November 3, 2016

PROBLEM 1

1. The diamonds dataset is often used to illustrate the ggplot2 package and can be loaded with data(diamonds). Please make a first plot of caret by price, colored by the variable color and using the variable cut to facet_wrap. Please touch up the plot by capitalizing the axes labels and the legend title and by giving an overall title.

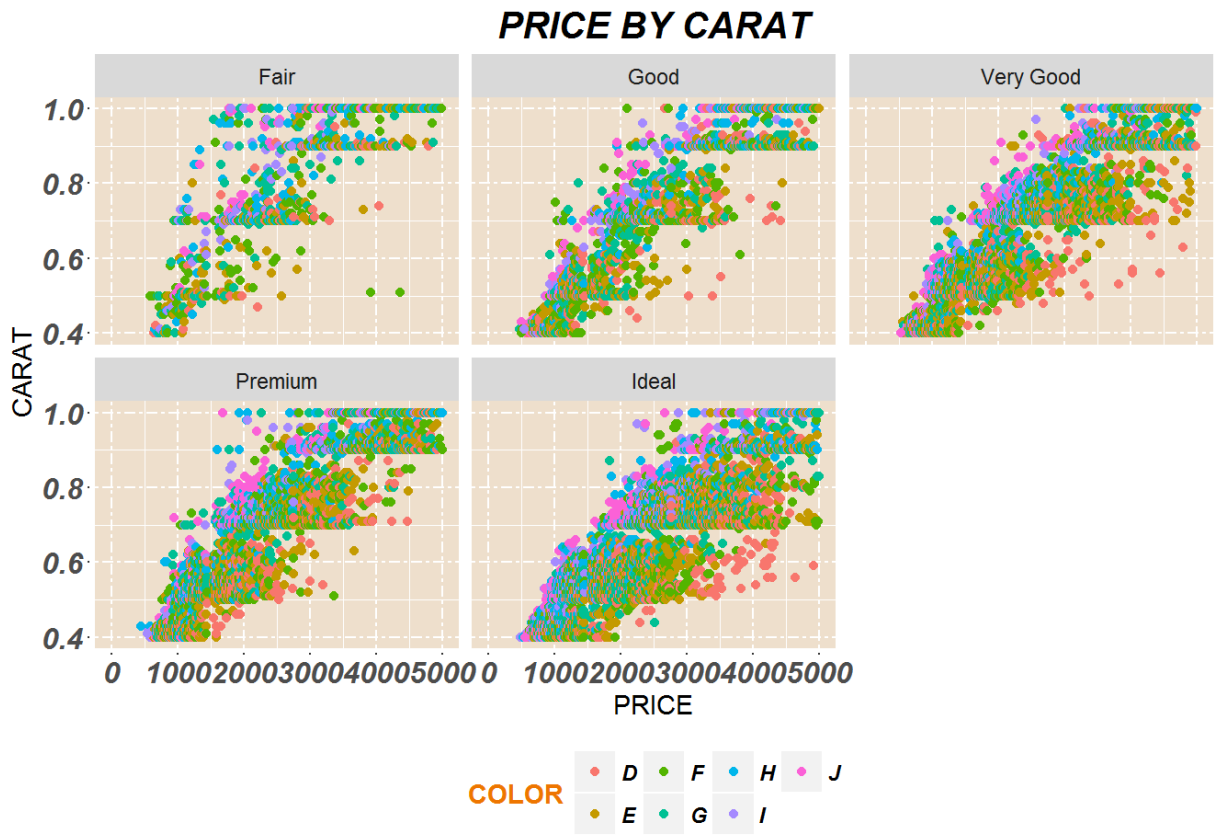
```
## Loading DIAMONDS dataset into R
data(diamonds)

# Basic plot with capitalized axis and label titles
diamond_plot <- ggplot(diamonds, aes(price, carat, color=color)) + geom_point() + facet_wrap(~cut) + scal
e_colour_discrete(name = "COLOR") + labs(title="PRICE BY CARAT", y = "CARAT", x = "PRICE")  + ylim(c(0.4,
1)) + xlim(c(0,5000))

diamond_plot
```



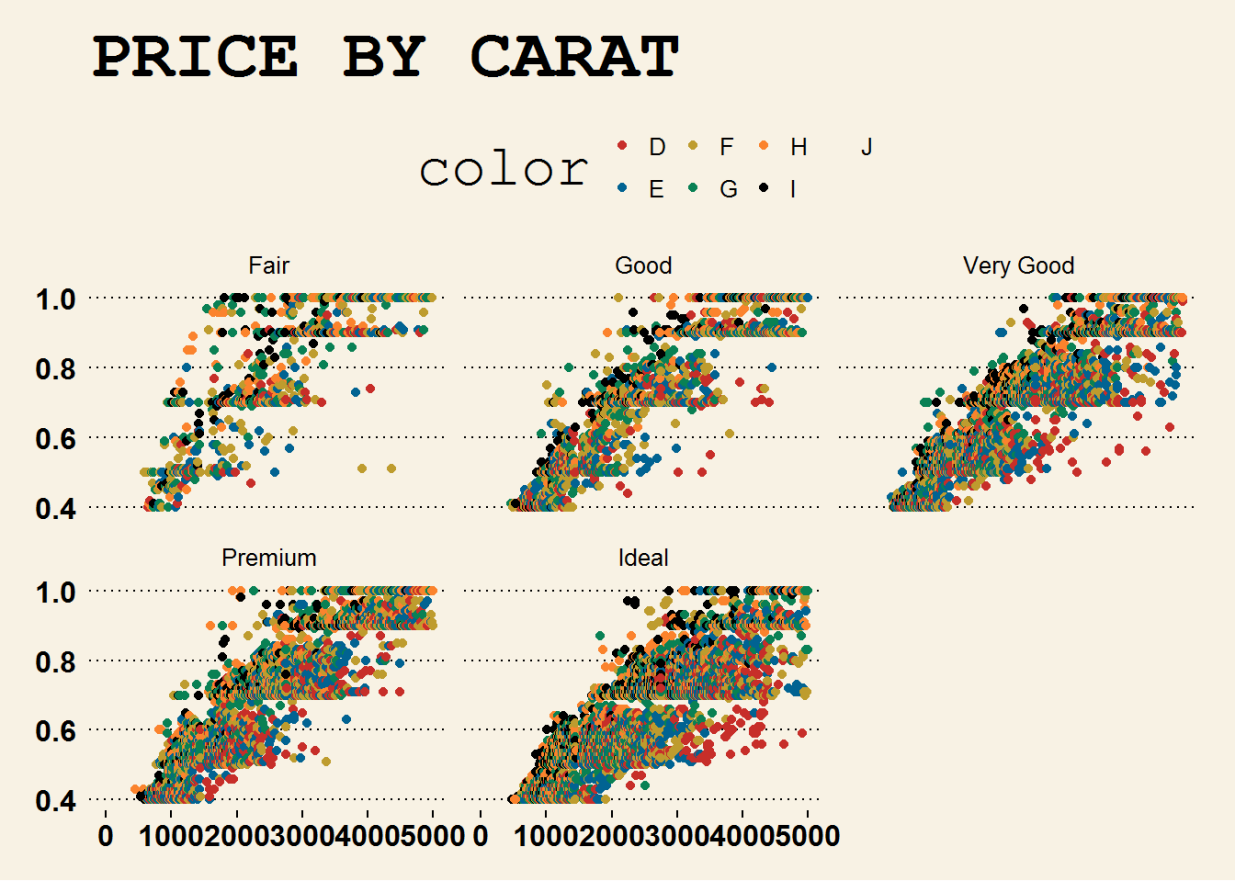
PROBLEM 1

After that, please make a total of 6 versions of that same plot using different combinations of themes and color palettes. Your choice as to the combinations.

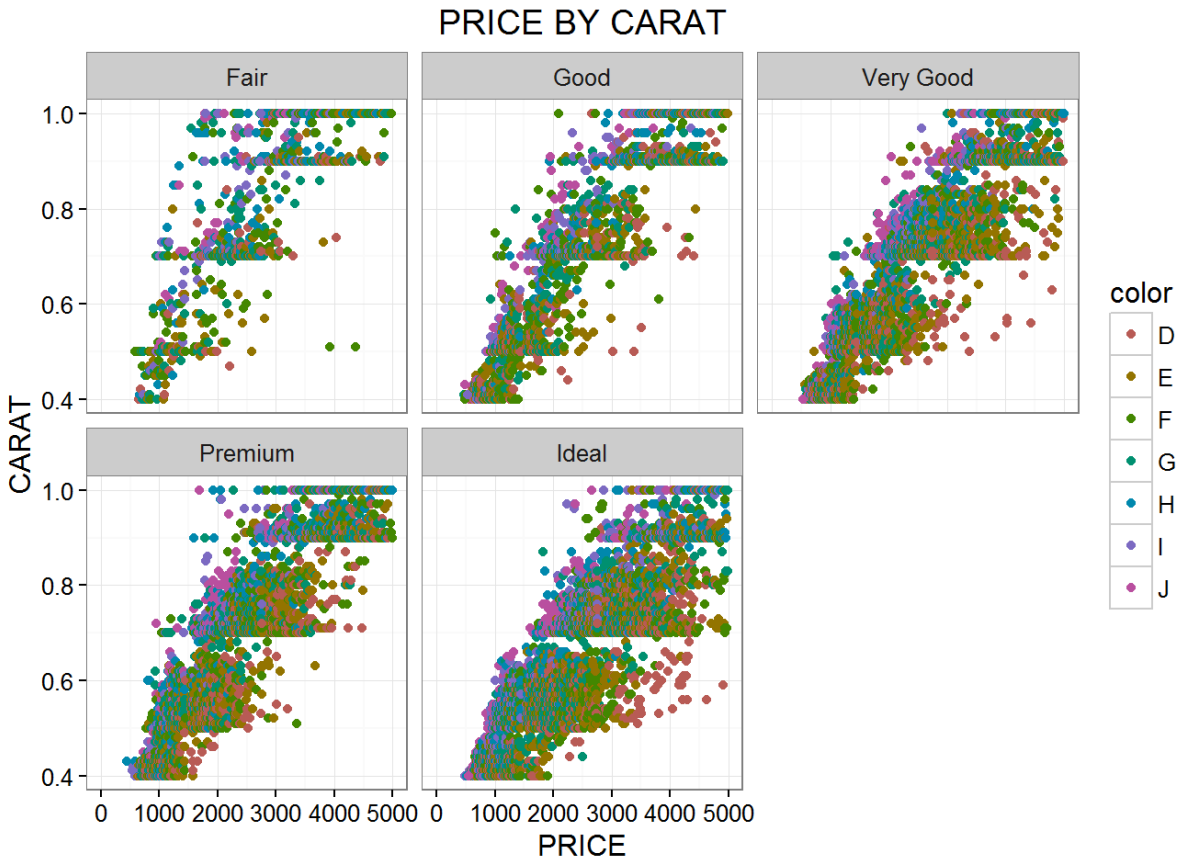
```
# PLOT 1 --> Panel background set to "antiquewhite2". Axis text converted to Bold & Italic. Legend positi
on change to bottom and direction to Horizontal.

diamond_plot + theme(panel.background = element_rect(fill = "antiquewhite2"), panel.grid.major = element_
line(linetype = "twodash"), axis.ticks = element_line(linetype = "dotted"),
    axis.text = element_text(size = 12, face = "bold.italic"),
    plot.title = element_text(size = 15,
        face = "bold.italic"), legend.text = element_text(face = "bold.italic"),
    legend.title = element_text(face = "bold",
        colour = "darkorange2"), legend.position = "bottom",
    legend.direction = "horizontal")
```



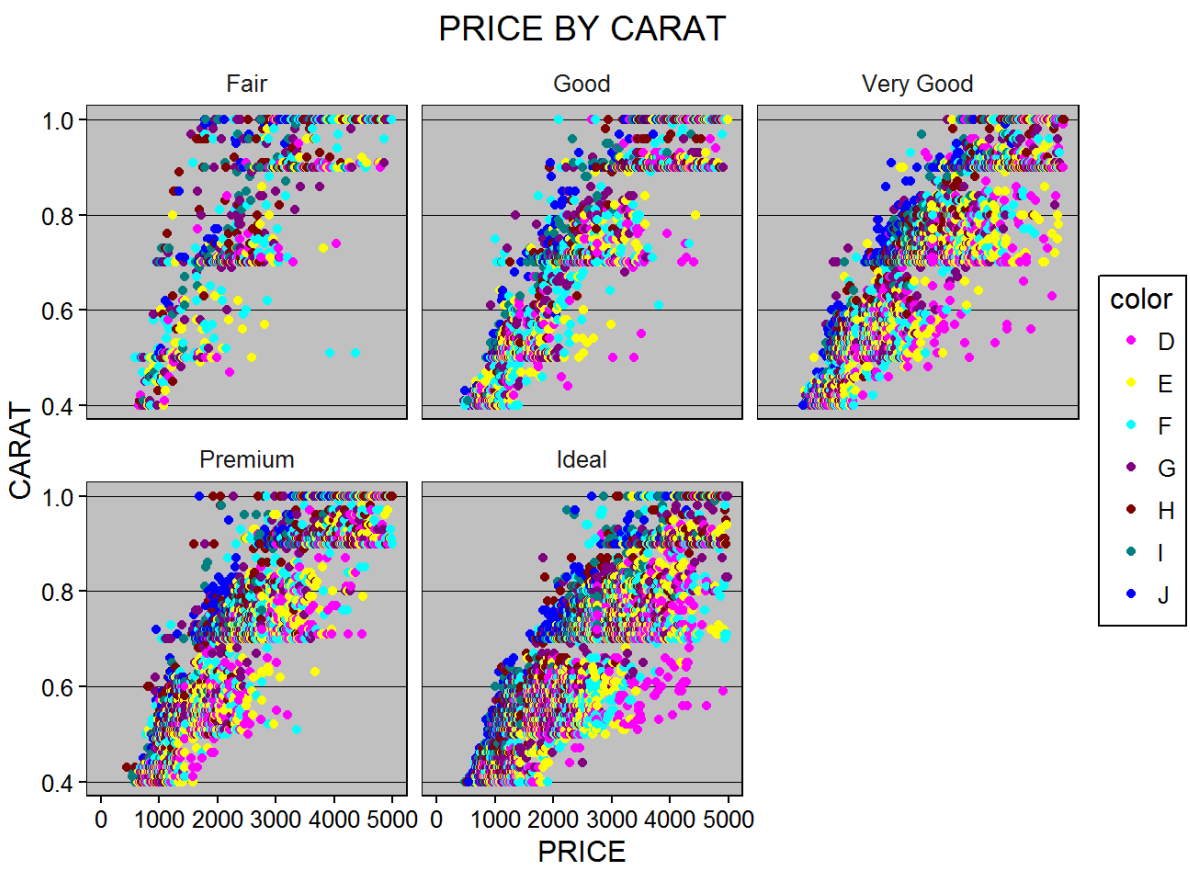

```
# PLOT 2 --> Wall Street Journal theme applied
diamond_plot + theme_wsj() + scale_color_wsj()
```



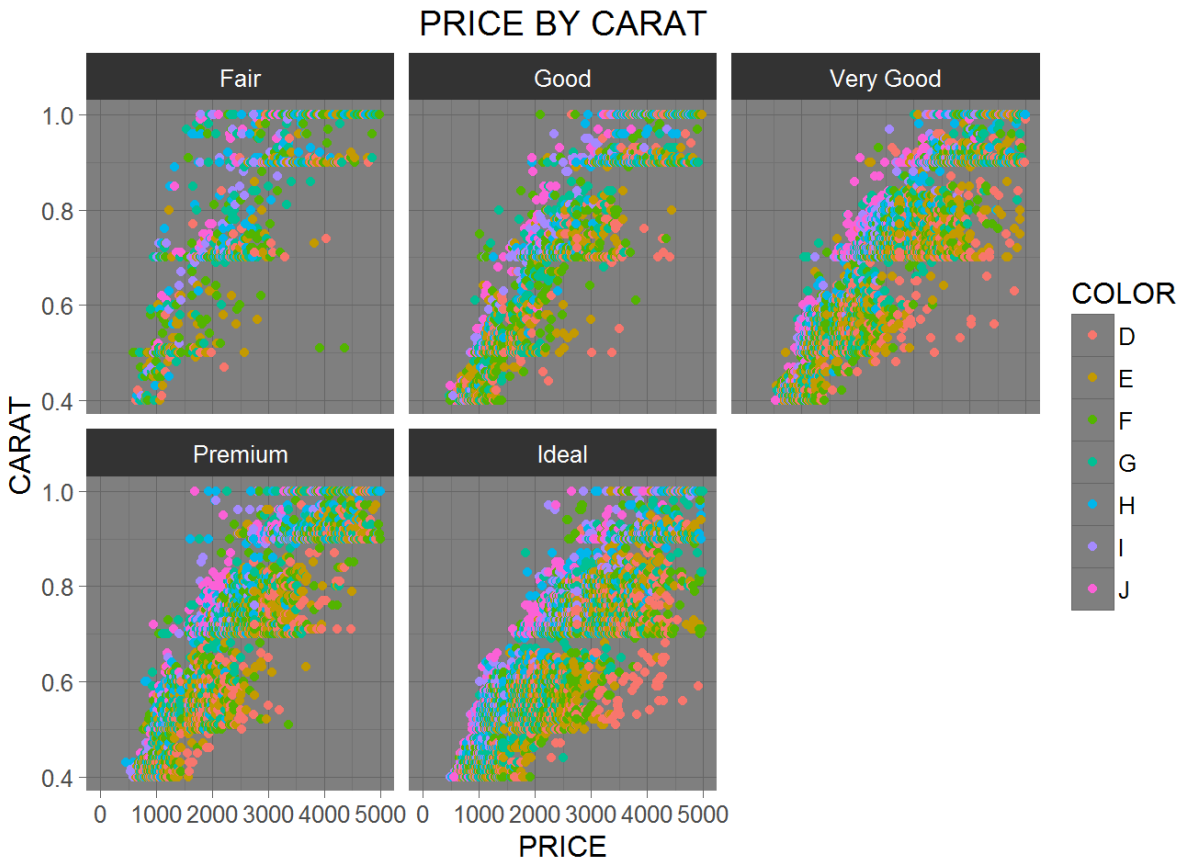

```
# PLOT 3 --> theme blank and white. chrome = 70 & luminance = 50
diamond_plot + scale_color_hue(c=70, l=50) + theme_bw()
```



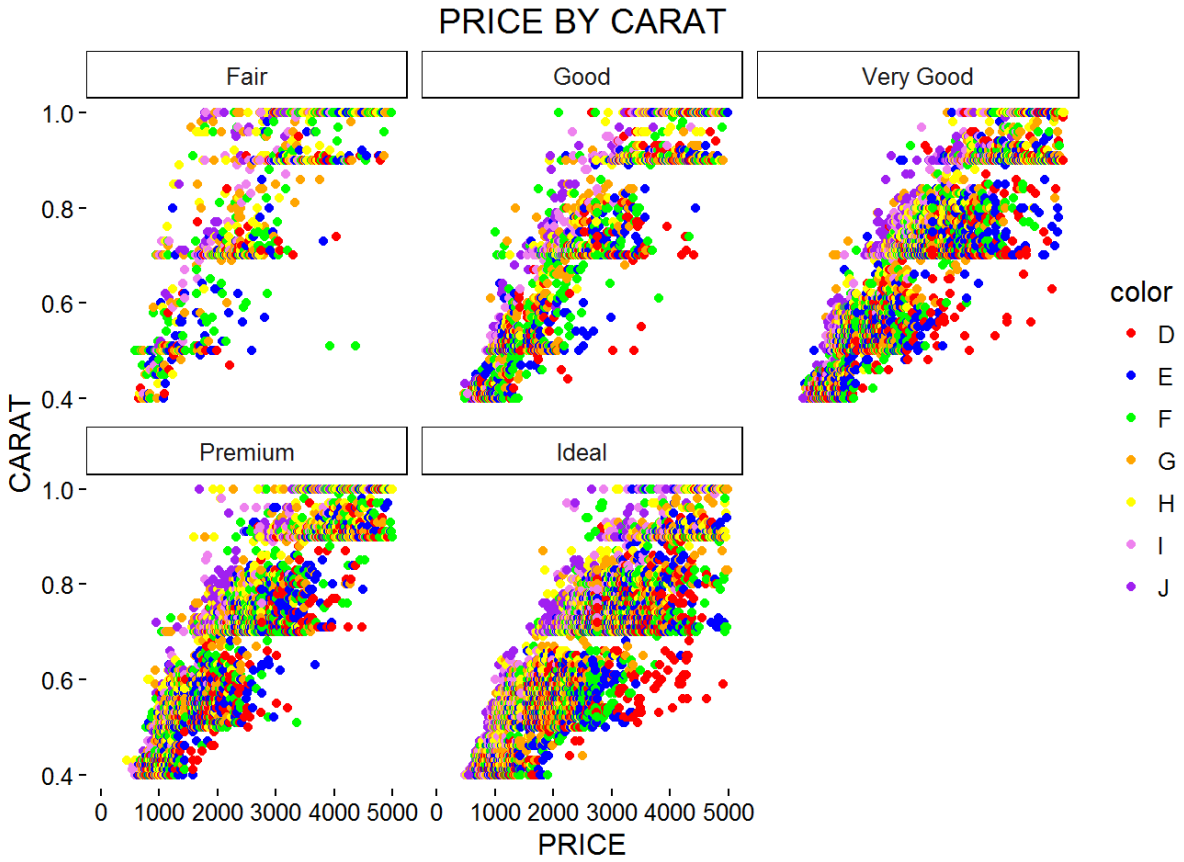
```
# PLOT 4 --> theme excel.
diamond_plot + theme_excel() + scale_color_excel()
```




```
# PLOT 5 --> theme tufte.
diamond_plot + theme_dark() + scale_fill_viridis(discrete = TRUE)
```




```
# PLOT 6 --> theme CLASSIC.
diamond_plot + theme_classic() + scale_color_manual(values=c("red", "blue", "green", "orange", "Yellow",
"violet", "purple"))
```



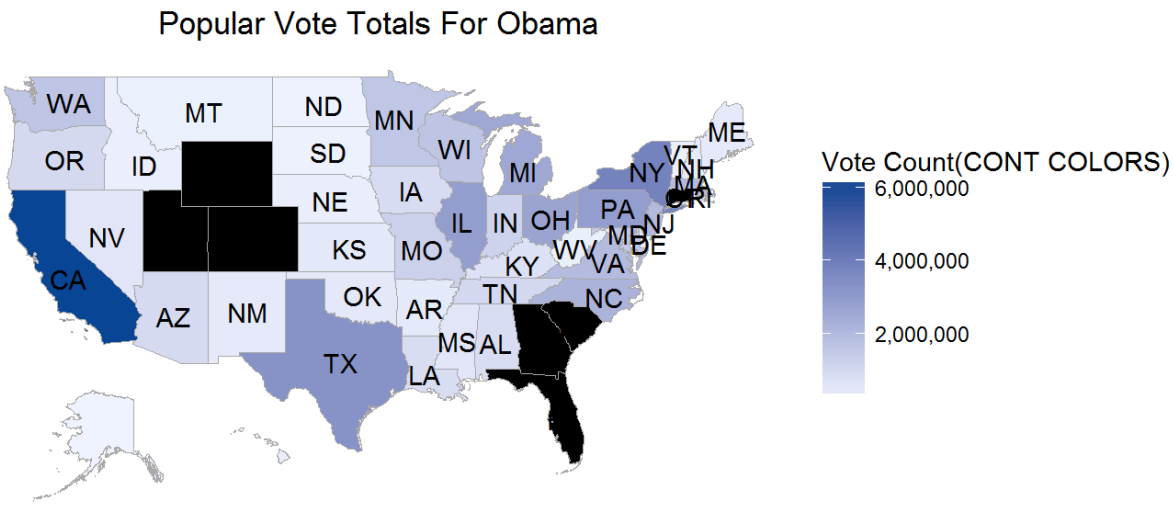
PROBLEM 2

2. The website https://www.theguardian.com/news/datablog/2012/nov/07/us-2012-election-county-results-download (https://www.theguardian.com/news/datablog/2012/nov/07/us-2012-election-county-results-download) has county-level election results for the 2012 US Presidential election. Download the Excel file and then read that into R (in submitting your homework, you may assume that the file is already downloaded to the working directory). Please make the following maps: a The popular vote totals for Obama

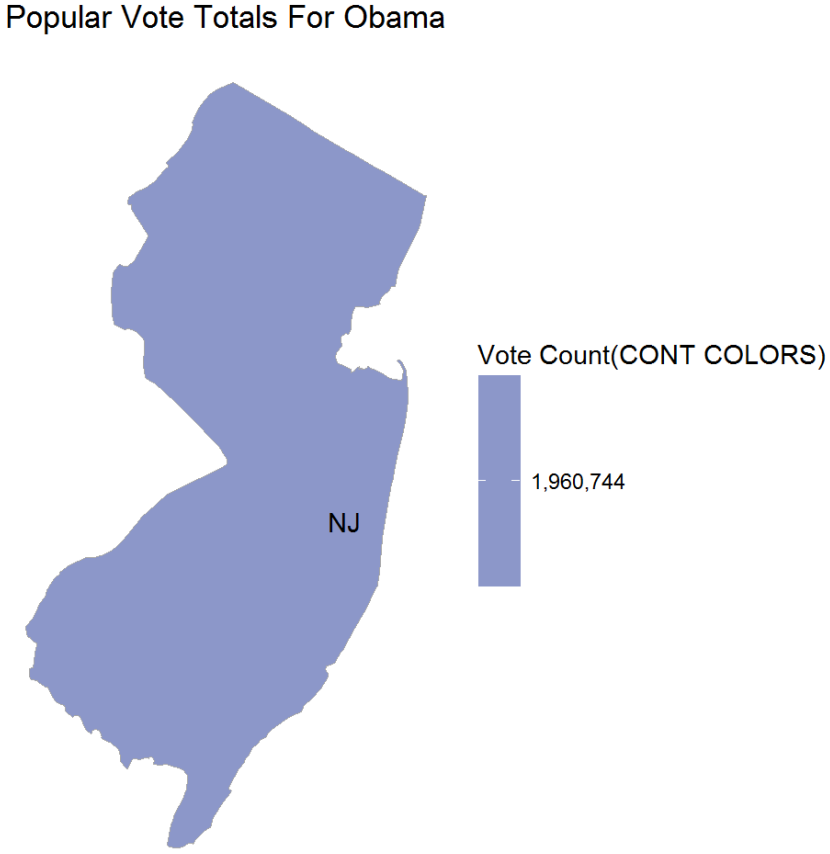
```
## Load excel into R
election <- read_excel("US_elect_county.xls", sheet = "OBAMA V ROMNEY ONLY", na="", col_names = TRUE, col
_types = NULL)
```

```
## DEFINEDNAME: 21 00 00 01 0b 00 00 00 02 00 00 00 00 00 00 0d 3b 00 00 00 00 1f 12 00 00 d5 00
## DEFINEDNAME: 21 00 00 01 0b 00 00 00 02 00 00 00 00 00 00 0d 3b 00 00 00 00 1f 12 00 00 d5 00
## DEFINEDNAME: 21 00 00 01 0b 00 00 00 02 00 00 00 00 00 00 0d 3b 00 00 00 00 1f 12 00 00 d5 00
## DEFINEDNAME: 21 00 00 01 0b 00 00 00 02 00 00 00 00 00 00 0d 3b 00 00 00 00 1f 12 00 00 d5 00
## DEFINEDNAME: 21 00 00 01 0b 00 00 00 02 00 00 00 00 00 00 0d 3b 00 00 00 00 1f 12 00 00 d5 00
## Unknown type: 517
## Unknown type: 517
```

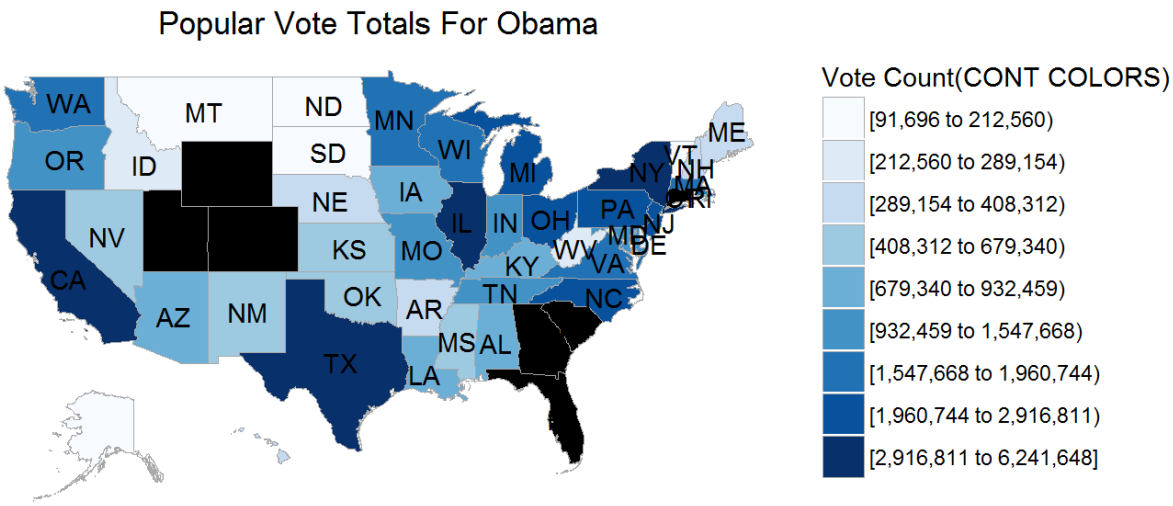
```
names(election) <- make.names(names(election))

colnames(election) <- c("State.Postal", "State.Name", "FIPS", "Obama.vote", "X", "Romney.vote", "Y")

election_state <- filter(election, election$'FIPS' == "0")

election_state_Obama <- select(election_state, State.Name, Obama.vote)

election_state_Obama$State.Name <- str_to_lower(election_state_Obama$State.Name)

colnames(election_state_Obama) <- c("region", "value")

# Remove duplicate rows
election_state_Obama1 <- election_state_Obama[!duplicated(election_state_Obama[,c('region')]),]

# MAP WITH CONTINOUS COLORS
state_choropleth(election_state_Obama1,
                 title  = "Popular Vote Totals For Obama",
                 legend = "Vote Count(CONT COLORS)", num_colors = 1)
```



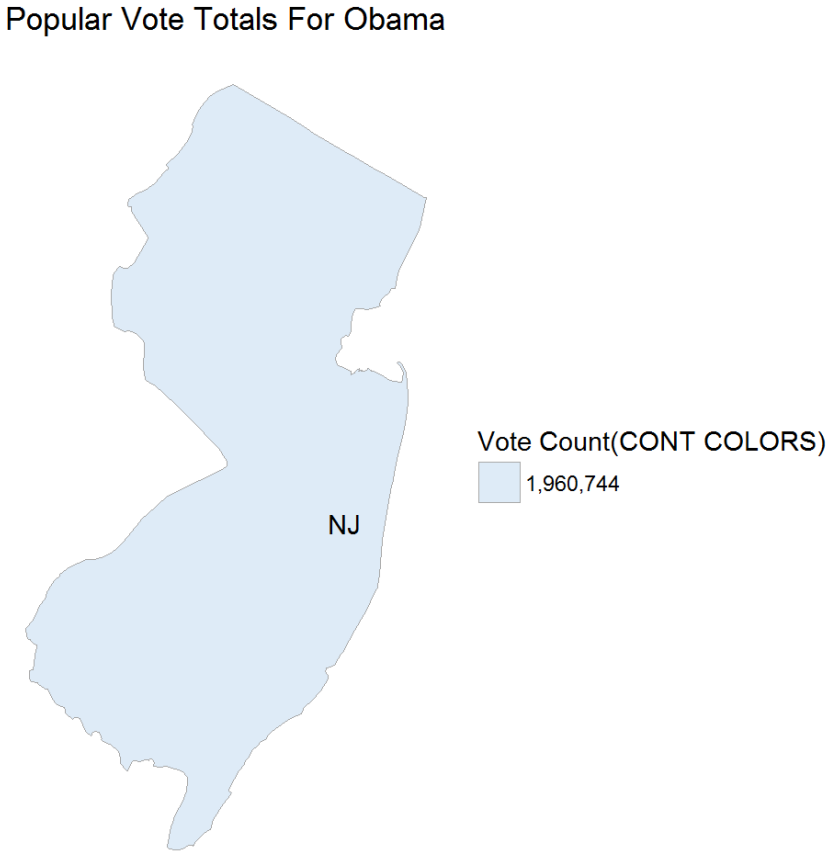

```
# NJ MAP WITH CONTINOUS COLORS
state_choropleth(election_state_Obama1,
                 title  = "Popular Vote Totals For Obama",
                 legend = "Vote Count(CONT COLORS)", num_colors = 1, zoom = "new jersey" )
```




```
# MAP WITH BINNED COLORS
state_choropleth(election_state_Obama1,
                 title  = "Popular Vote Totals For Obama",
                 legend = "Vote Count(CONT COLORS)", num_colors = 9)
```

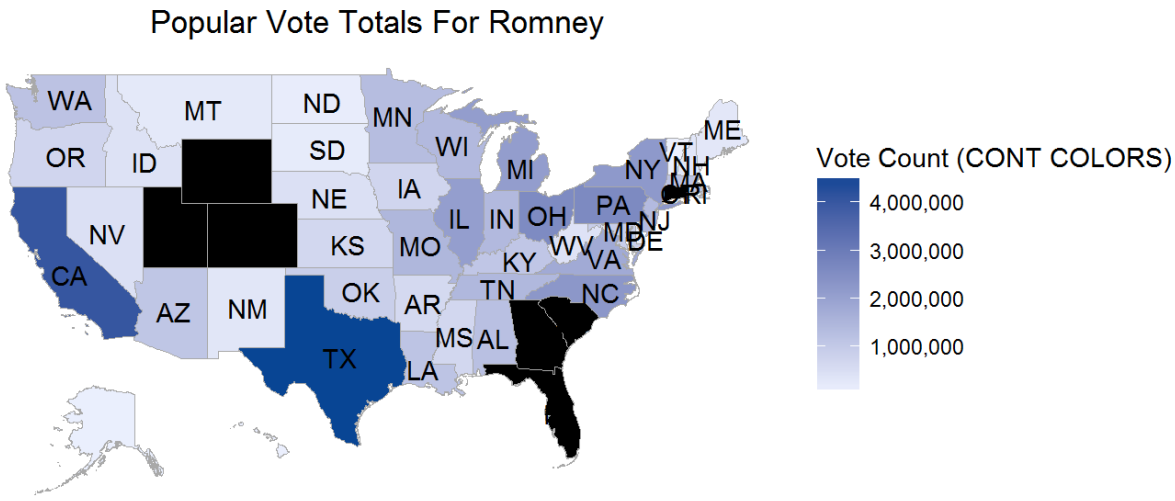



```
# NJ MAP WITH BINNED COLORS
state_choropleth(election_state_Obama1,
                 title  = "Popular Vote Totals For Obama",
                 legend = "Vote Count(CONT COLORS)", num_colors = 9, zoom = "new jersey")
```

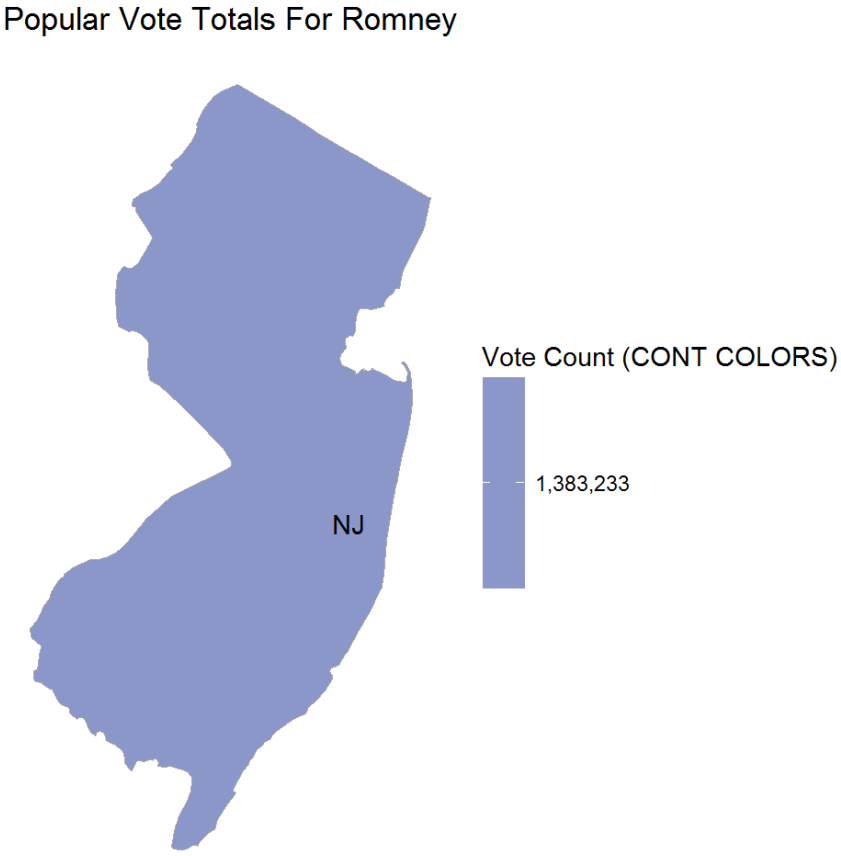


b The popular vote totals for Romney

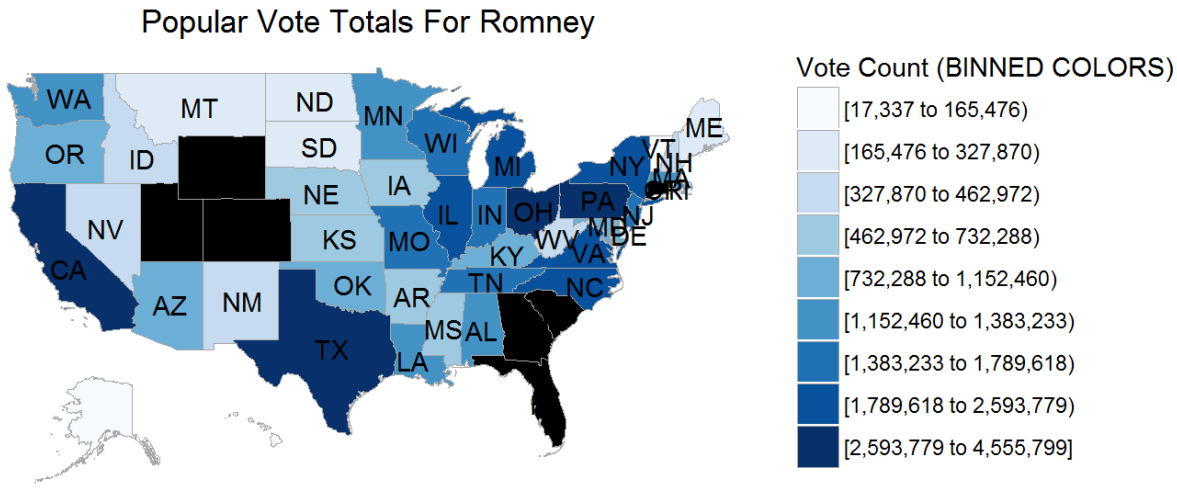
```
election_state_Romney <- select(election_state, State.Name, Romney.vote)

election_state_Romney$State.Name <- str_to_lower(election_state_Romney$State.Name)

colnames(election_state_Romney) <- c("region", "value")

# Remove duplicate rows
election_state_Romney1 <- election_state_Romney[!duplicated(election_state_Romney[,c('region')]),]

# MAP WITH CONTINOUS COLORS
state_choropleth(election_state_Romney1,
                 title  = "Popular Vote Totals For Romney",
                 legend = "Vote Count (CONT COLORS)", num_colors = 1)
```



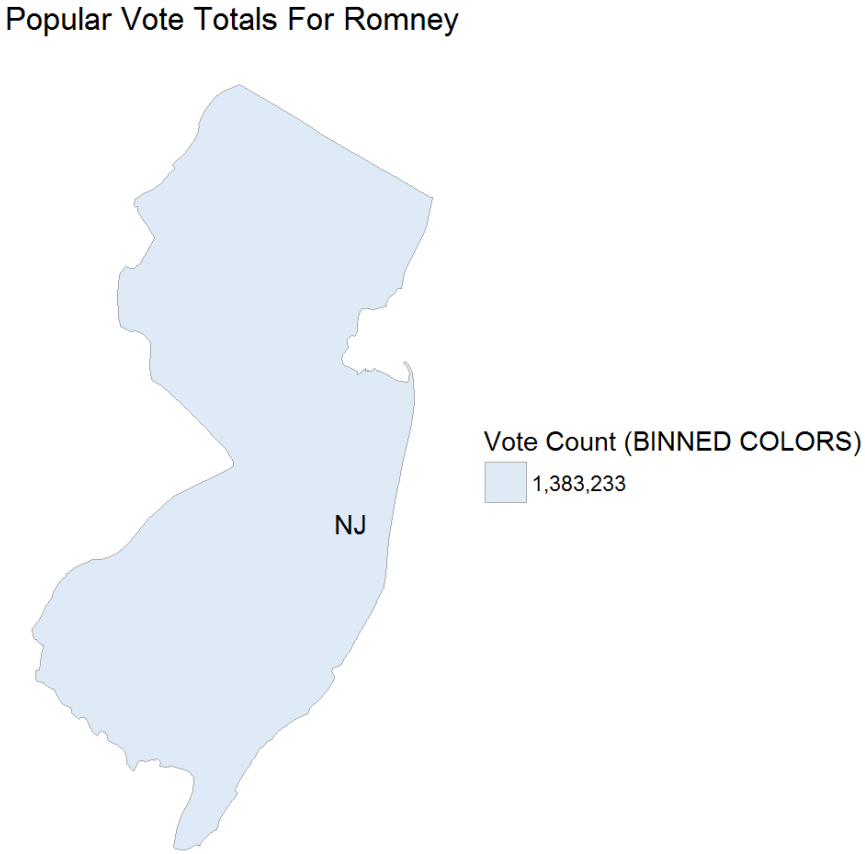

```
# NJ MAP WITH CONTINOUS COLORS
state_choropleth(election_state_Romney1,
                 title  = "Popular Vote Totals For Romney",
                 legend = "Vote Count (CONT COLORS)", num_colors = 1, zoom = "new jersey")
```




```
# MAP WITH BINNED COLORS
state_choropleth(election_state_Romney1,
                 title  = "Popular Vote Totals For Romney",
                 legend = "Vote Count (BINNED COLORS)", num_colors = 9)
```

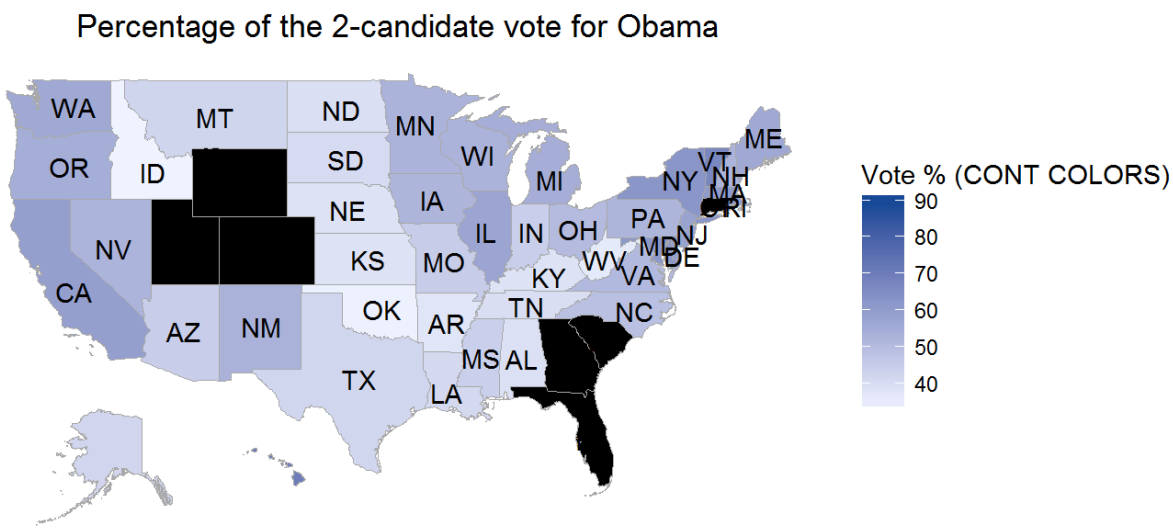



```
# NJ MAP WITH BINNED COLORS
state_choropleth(election_state_Romney1,
                 title  = "Popular Vote Totals For Romney",
                 legend = "Vote Count (BINNED COLORS)", num_colors = 9, zoom = "new jersey")
```

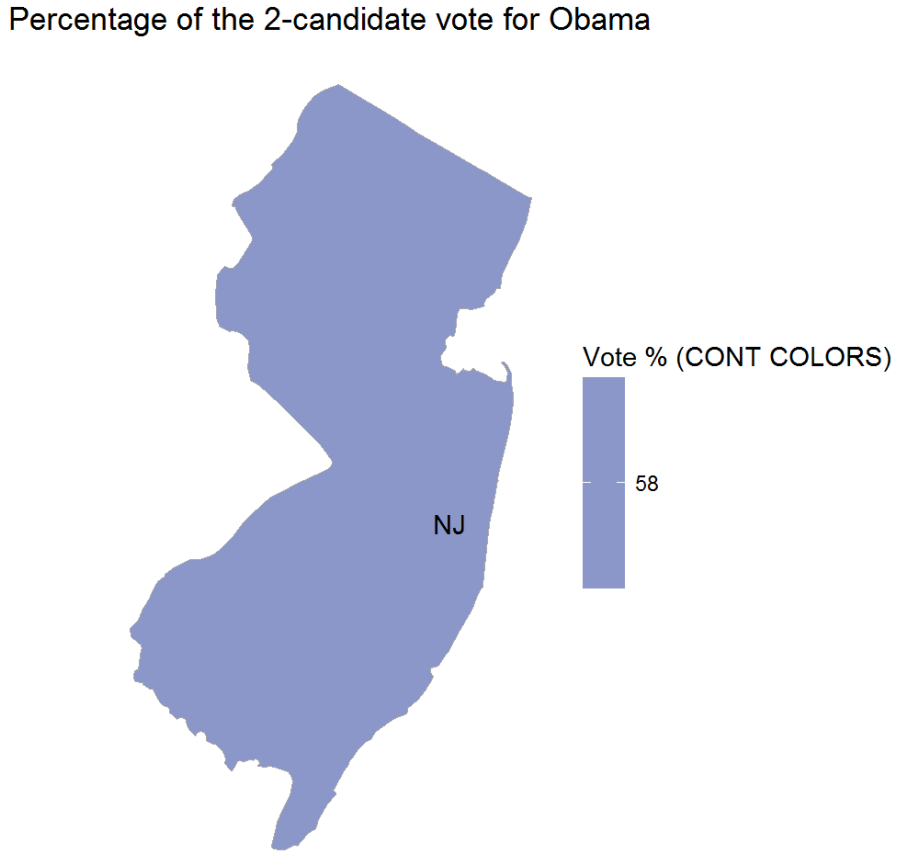


c The percentage of the 2-candidate vote for Obama

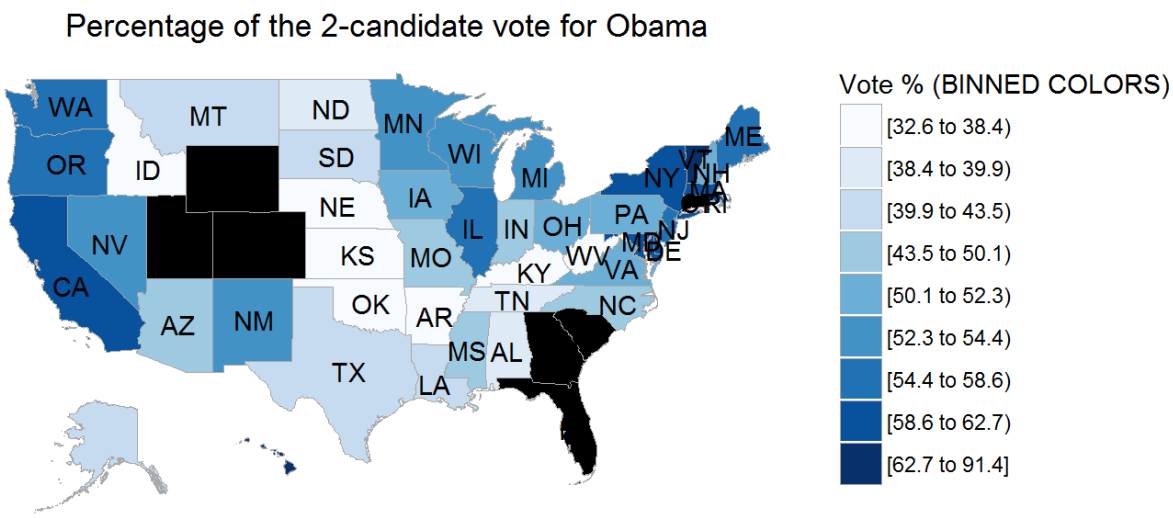
```
election_state_Obamapercent <- select(election_state, State.Name, X)

election_state_Obamapercent$State.Name <- str_to_lower(election_state_Obamapercent$State.Name)

colnames(election_state_Obamapercent) <- c("region", "value")

# Remove duplicate rows
election_state_Obamapercent1 <- election_state_Obamapercent[!duplicated(election_state_Obamapercent[,c('r
egion')]),]

# MAP WITH CONTINOUS COLORS
state_choropleth(election_state_Obamapercent1,
                 title  = "Percentage of the 2-candidate vote for Obama",
                 legend = "Vote % (CONT COLORS)", num_colors = 1)
```




```
# NJ MAP WITH CONTINOUS COLORS
state_choropleth(election_state_Obamapercent1,
                 title  = "Percentage of the 2-candidate vote for Obama",
                 legend = "Vote % (CONT COLORS)", num_colors = 1, zoom = "new jersey")
```

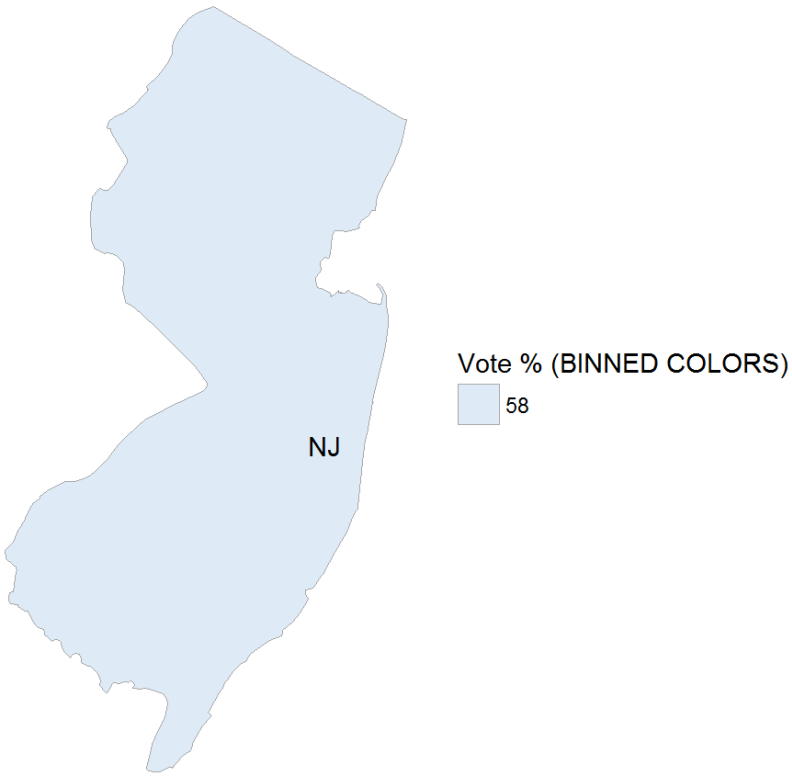



```
# MAP WITH BINNED COLORS
state_choropleth(election_state_Obamapercent1,
                 title  = "Percentage of the 2-candidate vote for Obama",
                 legend = "Vote % (BINNED COLORS)", num_colors = 9)
```




```
# NJ MAP WITH BINNED COLORS
state_choropleth(election_state_Obamapercent1,
                 title  = "Percentage of the 2-candidate vote for Obama",
                 legend = "Vote % (BINNED COLORS)", num_colors = 9, zoom = "new jersey")
```

Percentage of the 2-candidate vote for Obama



Problem 3

3. The weather history on October 17 for the closest (small) airport to Piscataway can be found at https://www.wunderground.com/history/airport/KSMQ/2016/10/17/DailyHistory.html (https://www.wunderground.com/history/airport/KSMQ/2016/10/17/DailyHistory.html)?req_city=Piscataway&req_state=NJ&req_statename=New+Jersey&reqdb.zip=08854&reqdb.magic=1&reqdb.wmo=99999 Download (and clean) the table at the bottom of the page, using html_node("#obsTable"), not html_node("table"). Please plot the time of day against temperature and dew point for the day (I suggest using a line and points for each, and separate colors for the temp/dew point).

```
url_weather <- "https://www.wunderground.com/history/airport/KSMQ/2016/10/17/DailyHistory.html?req_city=Piscataway&req_state=NJ&req_statename=New+Jersey&reqdb.zip=08854&reqdb.magic=1&reqdb.wmo=99999"

hourly_weather <-
  url_weather %>%
  read_html() %>%
     html_nodes("#obsTable") %>%
       html_table(fill = TRUE)

hourly_weather <- data.frame(hourly_weather)

hourly_weather1 <- select(hourly_weather, Time..EDT., Temp., Dew.Point)


colnames(hourly_weather1) <- c("Time", "Temp", "Dew")

hourly_weather1$Time <- str_sub(hourly_weather1$Time, end = 5)
hourly_weather1$Temp <- as.double(str_sub(hourly_weather1$Temp, end = 4))
hourly_weather1$Dew <- as.double(str_sub(hourly_weather1$Dew, end = 4))

require(data.table)

test_data_long <- melt(hourly_weather1, id="Time")  # convert to long format

ggplot(data=test_data_long,
       aes(x=Time, y=value, colour=variable)) +
  geom_line() + geom_point()
```