Assignment 7.2

Problem Statement

1. Write a program to create barplots for all the categorical columns in mtcars.

2. Create a scatterplot matrix by gear types in mtcars dataset.

3. Write a program to create a plot density by class variable

```
library(psych)
describe(mtcars)
```

```
##       vars  n   mean     sd median trimmed    mad   min    max  range  skew
## mpg      1 32  20.09   6.03  19.20   19.70   5.41 10.40  33.90  23.50  0.61
## cyl      2 32   6.19   1.79   6.00    6.23   2.97  4.00   8.00   4.00 -0.17
## disp     3 32 230.72 123.94 196.30  222.52 140.48 71.10 472.00 400.90  0.38
## hp       4 32 146.69  68.56 123.00  141.19  77.10 52.00 335.00 283.00  0.73
## drat     5 32   3.60   0.53   3.70    3.58   0.70  2.76   4.93   2.17  0.27
## wt       6 32   3.22   0.98   3.33    3.15   0.77  1.51   5.42   3.91  0.42
## qsec     7 32  17.85   1.79  17.71   17.83   1.42 14.50  22.90   8.40  0.37
## vs       8 32   0.44   0.50   0.00    0.42   0.00  0.00   1.00   1.00  0.24
## am       9 32   0.41   0.50   0.00    0.38   0.00  0.00   1.00   1.00  0.36
## gear    10 32   3.69   0.74   4.00    3.62   1.48  3.00   5.00   2.00  0.53
## carb    11 32   2.81   1.62   2.00    2.65   1.48  1.00   8.00   7.00  1.05
##       kurtosis    se
## mpg      -0.37  1.07
## cyl      -1.76  0.32
## disp     -1.21 21.91
## hp       -0.14 12.12
## drat     -0.71  0.09
## wt       -0.02  0.17
## qsec      0.34  0.32
## vs       -2.00  0.09
## am       -1.92  0.09
## gear     -1.07  0.13
## carb      1.26  0.29
```

```
library(ggplot2)
```

```
##
## Attaching package: 'ggplot2'
```

```
## The following objects are masked from 'package:psych':
##
##     %+%, alpha

library(car)

## Loading required package: carData

##
## Attaching package: 'car'

## The following object is masked from 'package:psych':
##
##     logit

library(corrgram)
library(reshape)
library(dplyr)

##
## Attaching package: 'dplyr'

## The following object is masked from 'package:reshape':
##
##     rename

## The following object is masked from 'package:car':
##
##     recode

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

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

library(gridExtra)

##
## Attaching package: 'gridExtra'

## The following object is masked from 'package:dplyr':
##
##     combine

data=mtcars
name=mtcars
mtcars$am <- as.factor(mtcars$am)
levels(mtcars$am) <- c("Automatic", "Manual")
head(mtcars)
```

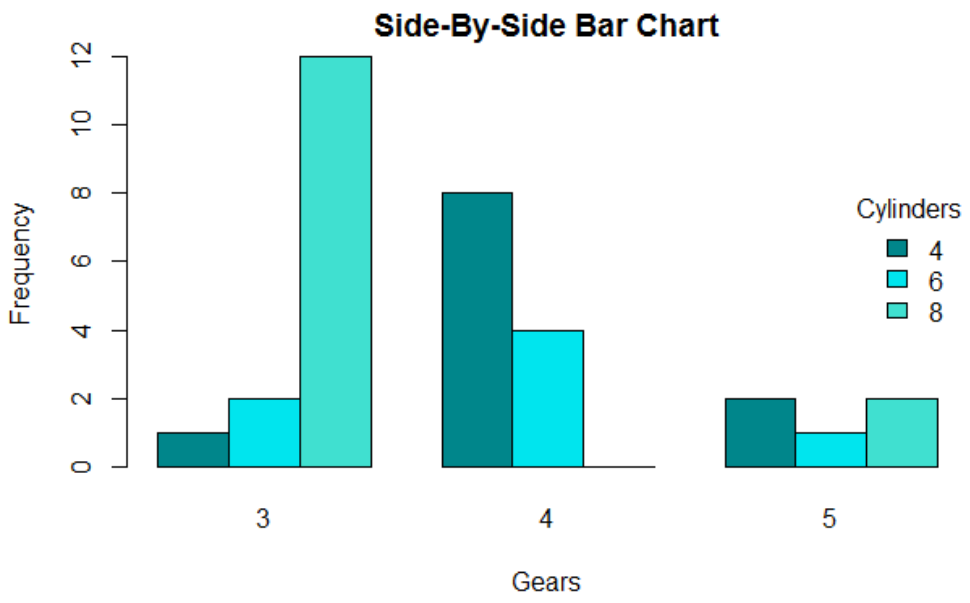
```
##                    mpg cyl disp  hp drat    wt  qsec vs        am gear
## Mazda RX4         21.0   6  160 110 3.90 2.620 16.46  0    Manual    4
## Mazda RX4 Wag     21.0   6  160 110 3.90 2.875 17.02  0    Manual    4
## Datsun 710        22.8   4  108  93 3.85 2.320 18.61  1    Manual    4
## Hornet 4 Drive    21.4   6  258 110 3.08 3.215 19.44  1 Automatic    3
## Hornet Sportabout 18.7   8  360 175 3.15 3.440 17.02  0 Automatic    3
## Valiant           18.1   6  225 105 2.76 3.460 20.22  1 Automatic    3
##                   carb
## Mazda RX4            4
## Mazda RX4 Wag        4
## Datsun 710           1
## Hornet 4 Drive       1
## Hornet Sportabout    2
## Valiant              1
```

```
summary(mtcars)
```

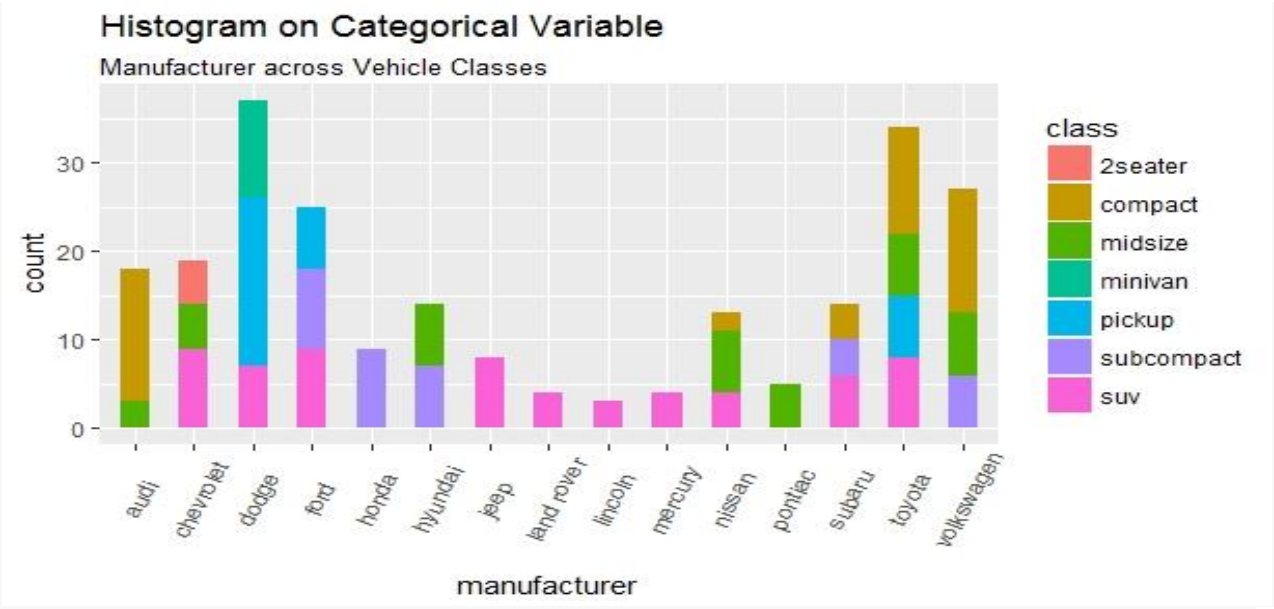
```
##       mpg             cyl             disp             hp
##  Min.   :10.40   Min.   :4.000   Min.   : 71.1   Min.   : 52.0
##  1st Qu.:15.43   1st Qu.:4.000   1st Qu.:120.8   1st Qu.: 96.5
##  Median :19.20   Median :6.000   Median :196.3   Median :123.0
##  Mean   :20.09   Mean   :6.188   Mean   :230.7   Mean   :146.7
##  3rd Qu.:22.80   3rd Qu.:8.000   3rd Qu.:326.0   3rd Qu.:180.0
##  Max.   :33.90   Max.   :8.000   Max.   :472.0   Max.   :335.0
##       drat             wt             qsec             vs
##  Min.   :2.760   Min.   :1.513   Min.   :14.50   Min.   :0.0000
##  1st Qu.:3.080   1st Qu.:2.581   1st Qu.:16.89   1st Qu.:0.0000
##  Median :3.695   Median :3.325   Median :17.71   Median :0.0000
##  Mean   :3.597   Mean   :3.217   Mean   :17.85   Mean   :0.4375
##  3rd Qu.:3.920   3rd Qu.:3.610   3rd Qu.:18.90   3rd Qu.:1.0000
##  Max.   :4.930   Max.   :5.424   Max.   :22.90   Max.   :1.0000
##          am          gear           carb
##  Automatic:19   Min.   :3.000   Min.   :1.000
##  Manual   :13   1st Qu.:3.000   1st Qu.:2.000
##                 Median :4.000   Median :2.000
##                 Mean   :3.688   Mean   :2.812
##                 3rd Qu.:4.000   3rd Qu.:4.000
##                 Max.   :5.000   Max.   :8.000
```

1. Write a program to create barplots for all the categorical columns in mtcars.

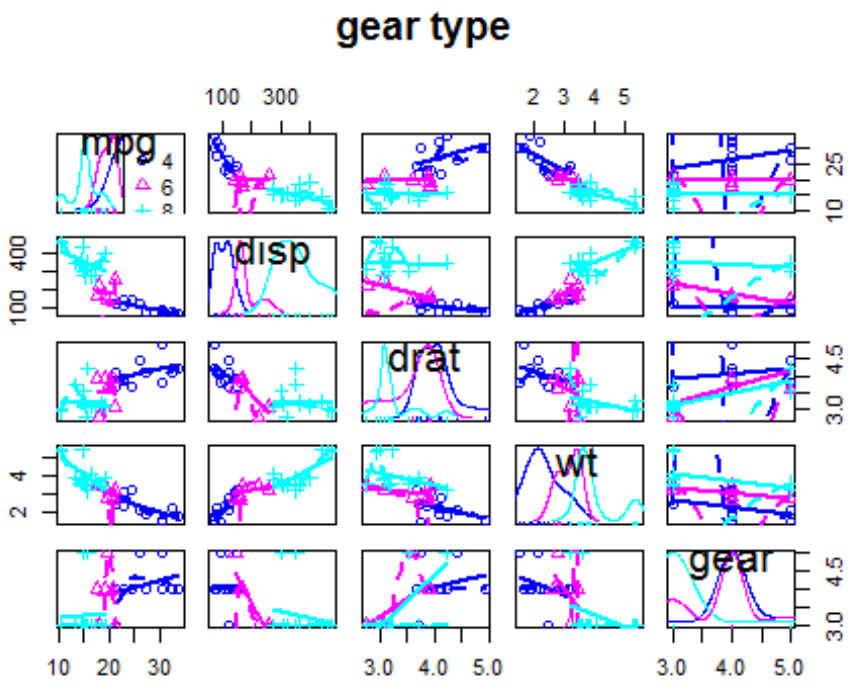
```
 table1 <- table(mtcars$cyl, mtcars$gear, dnn=c("Cylinders", "Gears")) #
Creates a contingency table
addmargins(table1) #Displays the table (Not necessary
barplot(table1, ylab="Frequency", xlab="Gears", main="Side-By-Side Bar
Chart", col=c("turquoise4", "turquoise2", "turquoise" ), beside=TRUE,
width=.3)
legend("right", title="Cylinders", legend= sort(unique(mtcars$cyl)), fill
=c("turquoise4", "turquoise2", "turquoise" ), box.lty=0)
legend("right", title="Cylinders", legend= sort(unique(mtcars$cyl)), fill
=c("turquoise4", "turquoise2", "turquoise" ), box.lty=0)
```



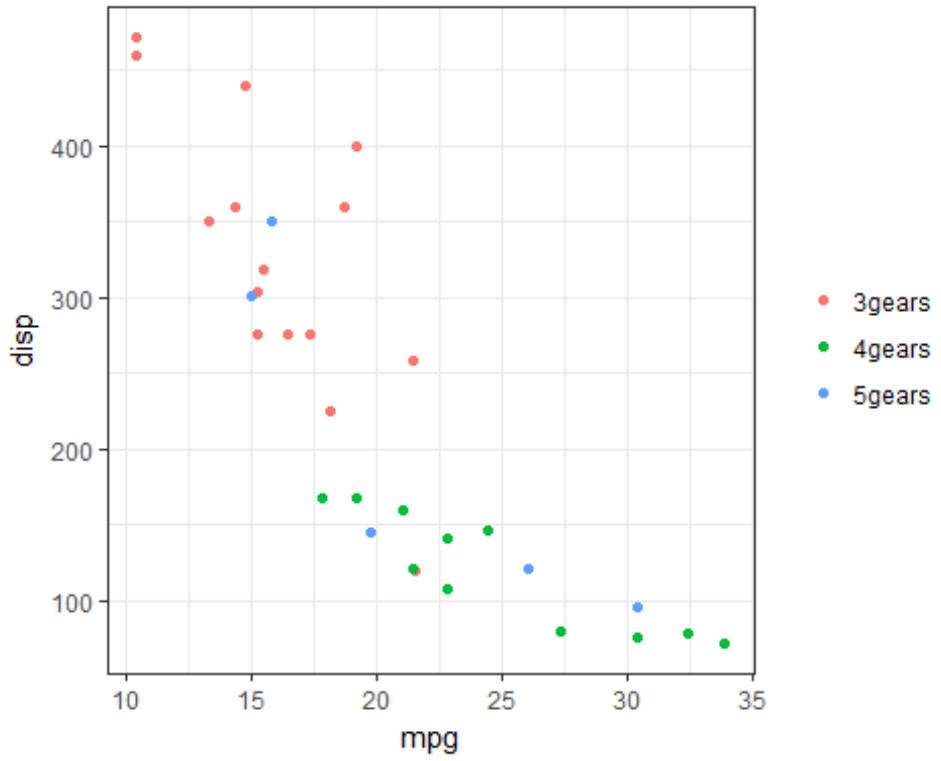

```
> # Histogram on a Categorical variable
> g <- ggplot(mpg, aes(manufacturer))
> g + geom_bar(aes(fill=class), width = 0.5) +
+   theme(axis.text.x = element_text(angle=65, vjust=0.6)) +
+   labs(title="Histogram on Categorical Variable",
+        subtitle="Manufacturer across Vehicle Classes")
```



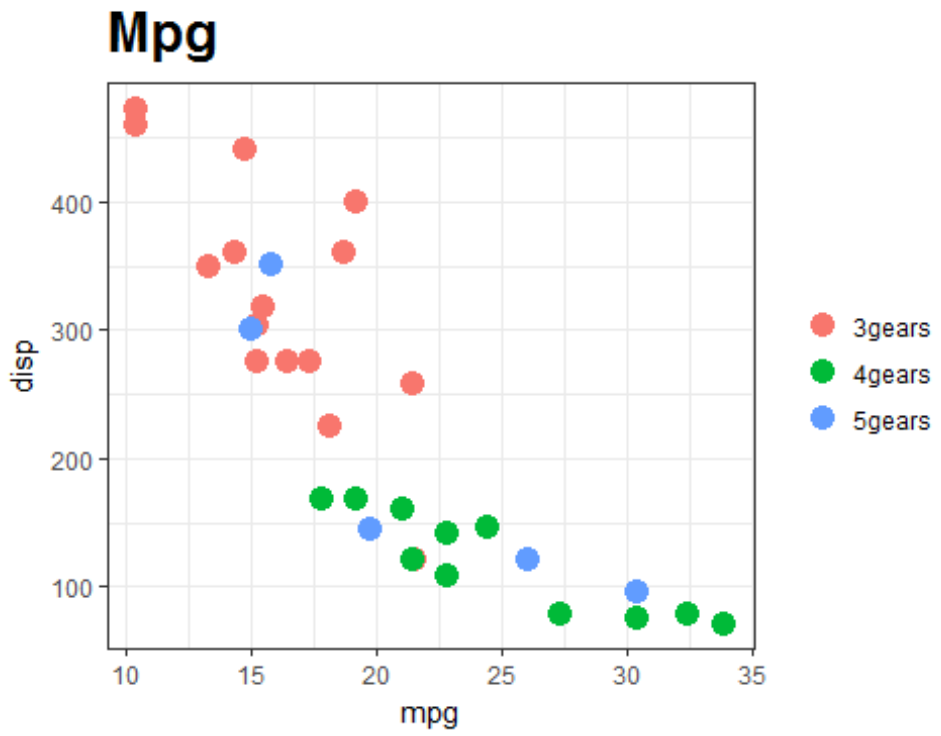
```
library(Matrix)

##
## Attaching package: 'Matrix'

## The following object is masked from 'package:reshape':
##
##     expand

scatterplotMatrix(~mpg+disp+drat+wt+gear|cyl,data=mtcars, main="gear type")
```



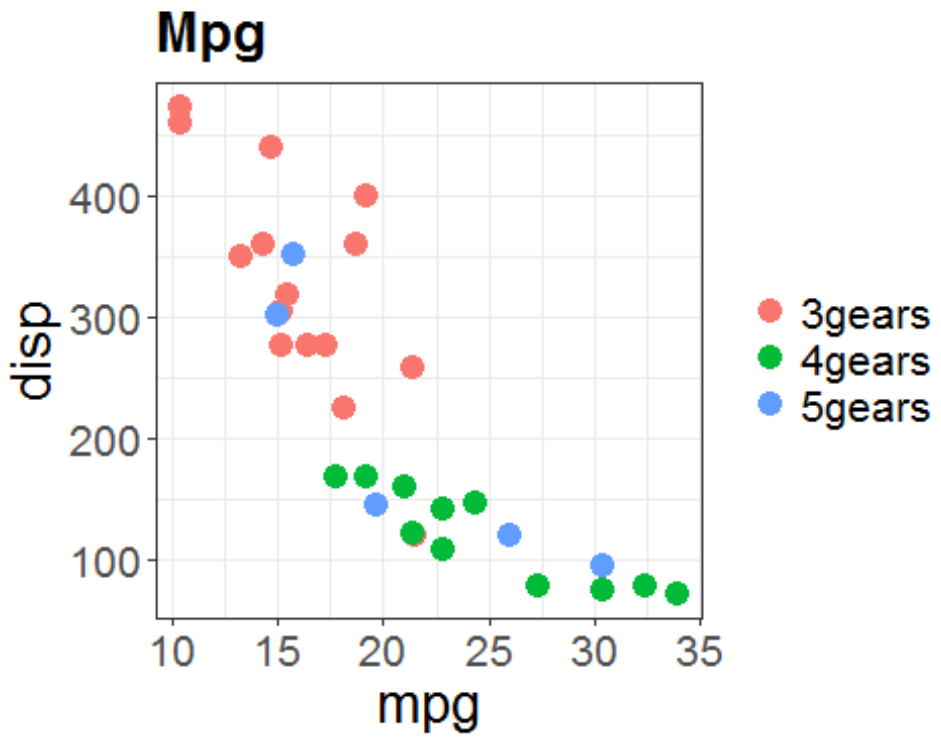

```
library(dplyr)
data(mtcars)
mtcars$gear <-factor(mtcars$gear,levels=c(3,4,5),
    labels=c("3gears","4gears","5gears"))
mtcars$am <- factor(mtcars$am,levels=c(0,1),
    labels=c("Automatic","Manual"))
mtcars$cyl <- factor(mtcars$cyl,levels=c(4,6,8),
   labels=c("4cyl","6cyl","8cyl"))
g <- ggplot()
g <- g + theme_bw()
g <- g + geom_point(data=mtcars, aes(x=mpg, y=disp,colour=gear))
+theme(legend.title=element_blank())
g
```



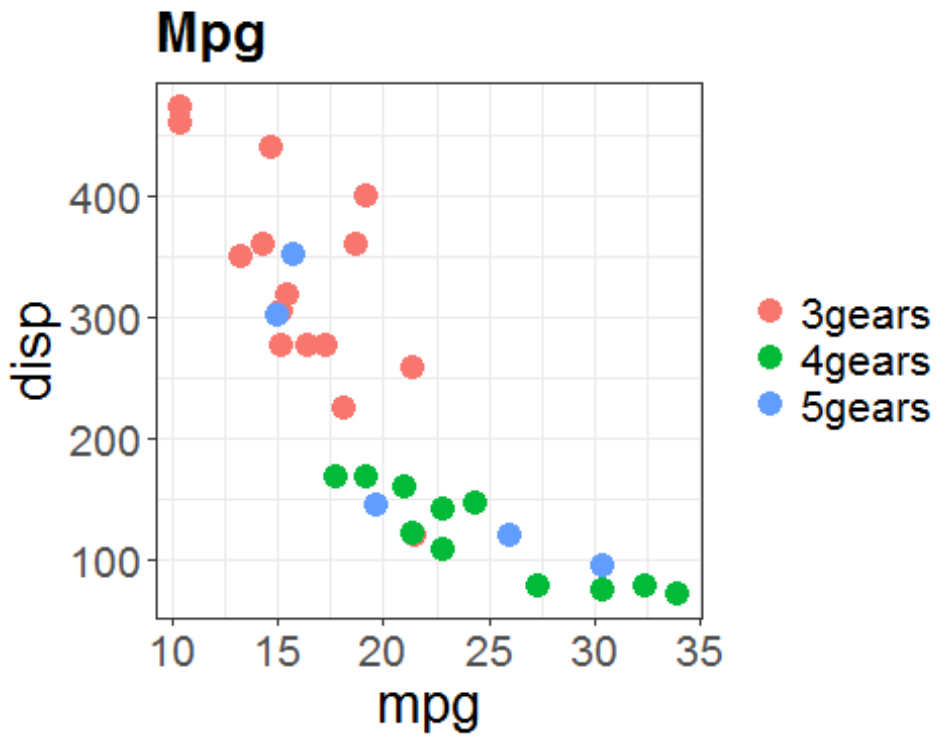

```
g <- ggplot()
g <- g + theme_bw()
g <- g + geom_point(data=mtcars, aes(x=mpg, y=disp,colour=gear),size = 4.0)
+theme(legend.title=element_blank())+
     guides(colour = guide_legend(override.aes = list(size=4)))+
     ggtitle('Mpg')+theme(plot.title = element_text(size=20, face="bold",
margin = margin(10, 0, 10, 0)))
g
```



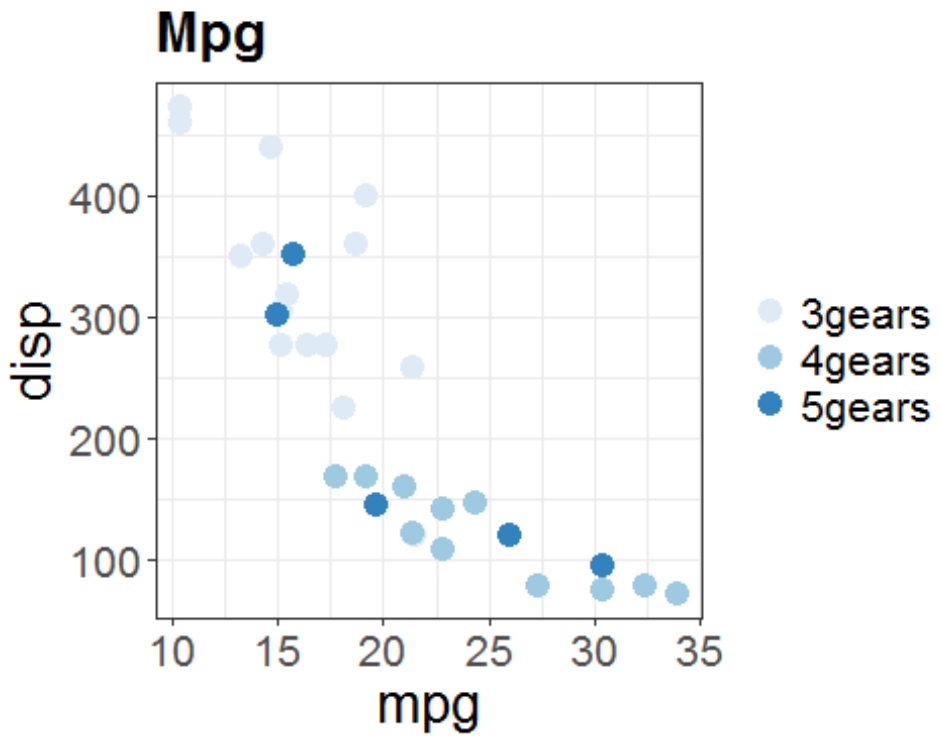

```
g <- ggplot()
g <- g + theme_bw()
g <- g + geom_point(data=mtcars, aes(x=mpg, y=disp,colour=gear),size = 4.0)
+theme(legend.title=element_blank())+
     guides(colour = guide_legend(override.aes = list(size=4)))+
     ggtitle('Mpg')+theme(plot.title = element_text(size=20, face="bold",
margin = margin(10, 0, 10, 0)))
g<- g + labs(x = "mpg", y="disp") +theme(text = element_text(size=20))
g
```




```
g <- ggplot()
g <- g + theme_bw()
g <- g + geom_point(data=mtcars, aes(x=mpg, y=disp,colour=gear),size = 4.0)
+theme(legend.title=element_blank())+
     guides(colour = guide_legend(override.aes = list(size=4)))+
     ggtitle('Mpg')+theme(plot.title = element_text(size=20, face="bold",
margin = margin(10, 0, 10, 0)))
g<- g + labs(x = "mpg", y="disp") +theme(text = element_text(size=20))
g
```

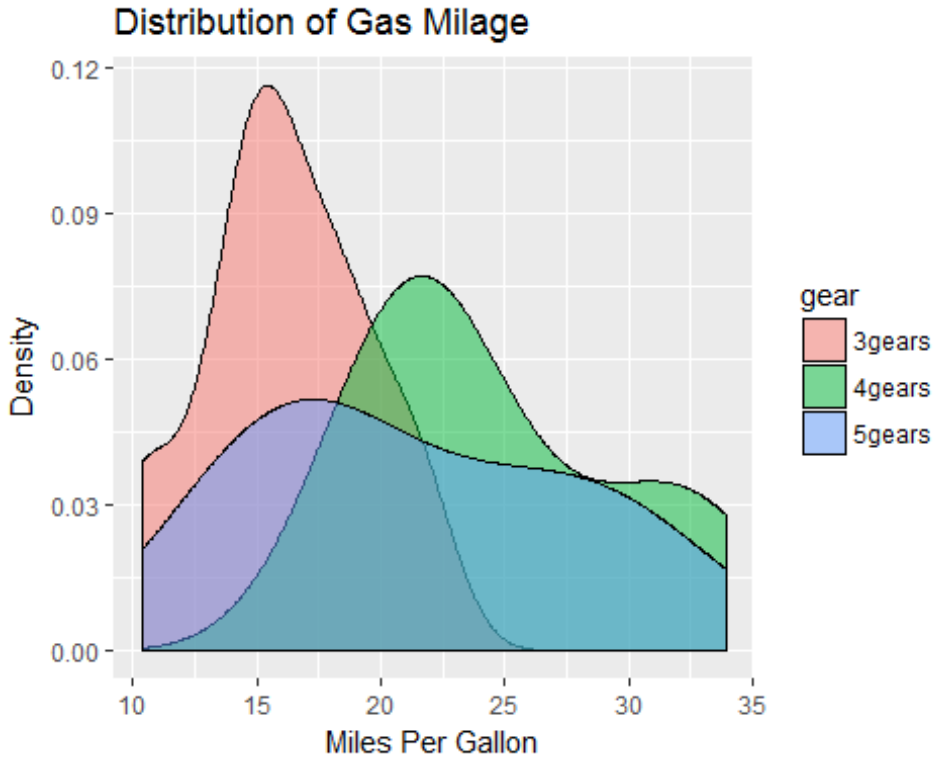



```
g <- ggplot()
g <- g + theme_bw()
g <- g + geom_point(data=mtcars, aes(x=mpg, y=disp,colour=gear),size = 4.0)
+theme(legend.title=element_blank())+
     guides(colour = guide_legend(override.aes = list(size=4)))+
     ggtitle("Mpg")+theme(plot.title = element_text(size=20, face="bold",
margin = margin(10, 0, 10, 0)))
g<- g + labs(x = "mpg", y="disp") +theme(text = element_text(size=20))+
scale_colour_brewer("Diamond\nclarity")
g
```


Mpg
disp
400
300
200
100
10
15
20
25
30
35
mpg
3gears
4gears
5gears

Write a program to create a plot density by class variable

```
qplot(mpg, data=mtcars, geom="density", fill=gear, alpha=I(.5),
   main="Distribution of Gas Milage", xlab="Miles Per Gallon",
   ylab="Density")
```



R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see http://rmarkdown.rstudio.com.

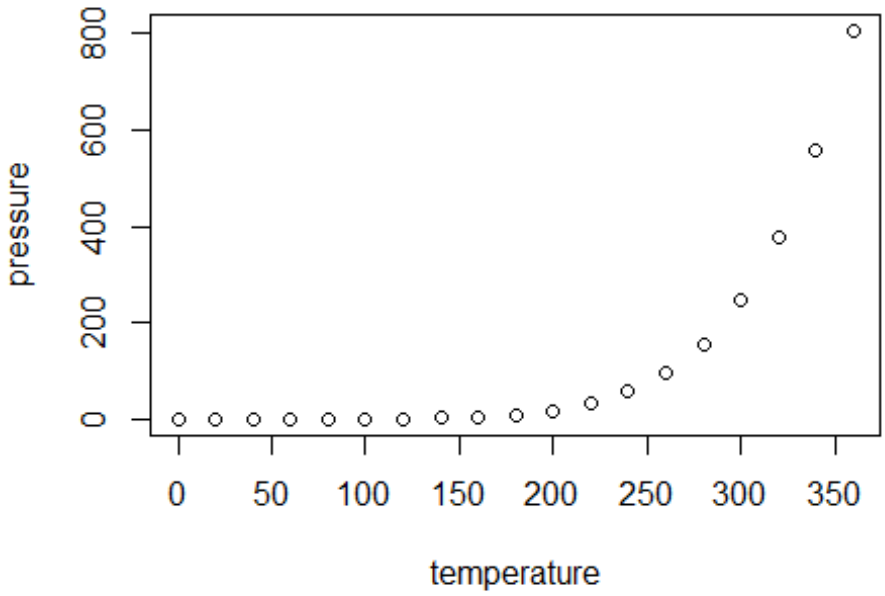
When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

```
summary(cars)
```

```
##      speed           dist       
##  Min.   : 4.0   Min.   :  2.00  
##  1st Qu.:12.0   1st Qu.: 26.00  
##  Median :15.0   Median : 36.00  
##  Mean   :15.4   Mean   : 42.98  
##  3rd Qu.:19.0   3rd Qu.: 56.00  
##  Max.   :25.0   Max.   :120.00
```

Including Plots

You can also embed plots, for example:



Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.