EXPERIMENT no. 5 and 6 [R PROGRAMMING LANGUAGE]

- Asst. Prof. Ashwini Mathur

Question 1. What happens when the following code is run? gender <- c("M" , "M" , "F" , "F" , "F") whereF<- (gender == "F") gender[whereF] <- "Female"

Question 2. In R, evaluate the expressions

```
2^52 + k - 2^52
2^53 + k - 2^53
2^54 + k - 2^54,
```

for the cases where k = 1, 2, 3, 4. Explain what you observe. What could be done to obtain results in R which are mathematically correct?

Question 3. Construct the data frame char-num in the following way: char <- c("2" , "1" , "0") num<- 0:2 charnum<- data.frame(char, num);

Question 4. The following are a sample of observations on incoming solar radiation at a greenhouse: 11.1 10.6 6.3 8.8 10.7 11.2 8.9 12.2

```
     a.    Assign the data to an object called solar.radiation.
     b.    Find the mean, median, range, and variance of the radiation o
bservations.
     c.    Add 10 to each observation of solar.radiation, and assign the
result to sr10. Find the mean, median, range,                     and va
riance of sr10. Which statistics change, and by how much?
     d.    Plot a histogram of the solar.radiation, sr10, and srm2.
```

Question 4. Modify the code to generate the Fibonacci sequence in the following ways.

```
1. Change the first two elements to 2 and 2.
2. Change the first two elements to 3 and 2.
```

Question 5. Use the inbuilt data car and uses the possible graphical plots using **ggplot2** graphical packages.

Question 1. What happens when the following code is run? gender <- c("M" , "M" , "F" , "F" , "F") whereF<- (gender == "F") gender[whereF] <- "Female"

In [45]:
```
gender <- c("M" , "M" , "F" , "F" , "F" )
                    whereF <- (gender == "F" )
                    gender[whereF] <- "Female"
#Print gender
gender
```

'M' · 'M' · 'Female' · 'Female' · 'Female'

Question 2. In R, evaluate the expressions

```
2^52 + k - 2^52
2^53 + k - 2^53
2^54 + k - 2^54,
```

for the cases where k = 1, 2, 3, 4. Explain what you observe. What could be done to obtain results in R which are mathematically correct?

In [46]:

```r
k <- 1:4

#1
2^52 + k - 2^52
#2
2^53 + k - 2^53
#3
2^54 + k - 2^54
```

1 · 2 · 3 · 4

0 · 2 · 4 · 4

0 · 0 · 4 · 4

We are right at the borderline of accuracy for double precision foating point. **The first expression is done accurately, the other two suffer rounding error.** The result would be correct if done in a different order, e.g.

In [47]:

```r
2^52 - 2^52 + k

2^53 + k - 2^53

2^54 - 2^54 + k
```

1 · 2 · 3 · 4

0 · 2 · 4 · 4

1 · 2 · 3 · 4

Question 2. Construct the data frame char-num in the following way: char <- c("2" , "1" , "0") num<- 0:2 charnum<- data.frame(char, num);

In [48]:
```
char <- c("2", "1", "0")
num  <- 0:2

charnum <- data.frame(char, num)

#as.numeric(char)


charnum$char

charnum$num
```

2 · 1 · 0

▶ **Levels**:

0 · 1 · 2

Question 3. Modify the code to generate the Fibonacci sequence in the following ways.

1. Change the first two elements to 2 and 2.
2. Change the first two elements to 3 and 2.

In [49]:
```
#   Implement the code to generate the Fibonacci sequence.

Fibonacci <- numeric(12)
Fibonacci
Fibonacci[1] <- 2
Fibonacci[2] <- 2
Fibonacci
for (i in 3:12) Fibonacci[i] <- Fibonacci[i-2] + Fibonacci[i-1]
Fibonacci
```

0 · 0 · 0 · 0 · 0 · 0 · 0 · 0 · 0 · 0 · 0 · 0

2 · 2 · 0 · 0 · 0 · 0 · 0 · 0 · 0 · 0 · 0 · 0

2 · 2 · 4 · 6 · 10 · 16 · 26 · 42 · 68 · 110 · 178 · 288

In [50]:
```
Fibonacci <- numeric(12)
Fibonacci[1] <- Fibonacci[2] <- 1
for (i in 3:12) Fibonacci[i] <- Fibonacci[i-1] - Fibonacci[i-2]
Fibonacci
```

1 · 1 · 0 · -1 · -1 · 0 · 1 · 1 · 0 · -1 · -1 · 0

Question 4. The following are a sample of observations on incoming solar radiation at a greenhouse:

```
    11.1 10.6 6.3 8.8 10.7 11.2 8.9 12.2

      a.    Assign the data to an object called solar.radiation.
      b.    Find the mean, median, range, and variance of the radiation o
bservations.
      c.    Add 10 to each observation of solar.radiation, and assign the
result to sr10. Find the     mean, median, range,and variance of sr10. W
hich statistics change, and by how much?
      d.    Plot a histogram of the solar.radiation, sr10, and srm2.
```

In [51]:
```
#   Assign the data to an object called solar.radiation.

solar.radiation <- c(11.1,10.6,6.3,8.8,10.7,11.2,8.9,12.2)
solar.radiation
```

11.1 · 10.6 · 6.3 · 8.8 · 10.7 · 11.2 · 8.9 · 12.2

In [52]:
```
#b. Find the mean, median, range, and variance of the radiation observations.

mean(solar.radiation)
median(solar.radiation)
```

9.975

10.65

In [53]:
```
range(solar.radiation)
var(solar.radiation)
```

6.3 · 12.2

3.525

In [54]:
```
#Add 10 to each observation of solar.radiation, and assign the result to sr10. F

sr10 <- solar.radiation + 10

mean(sr10)

median(sr10)

range(sr10)

var(sr10)
```

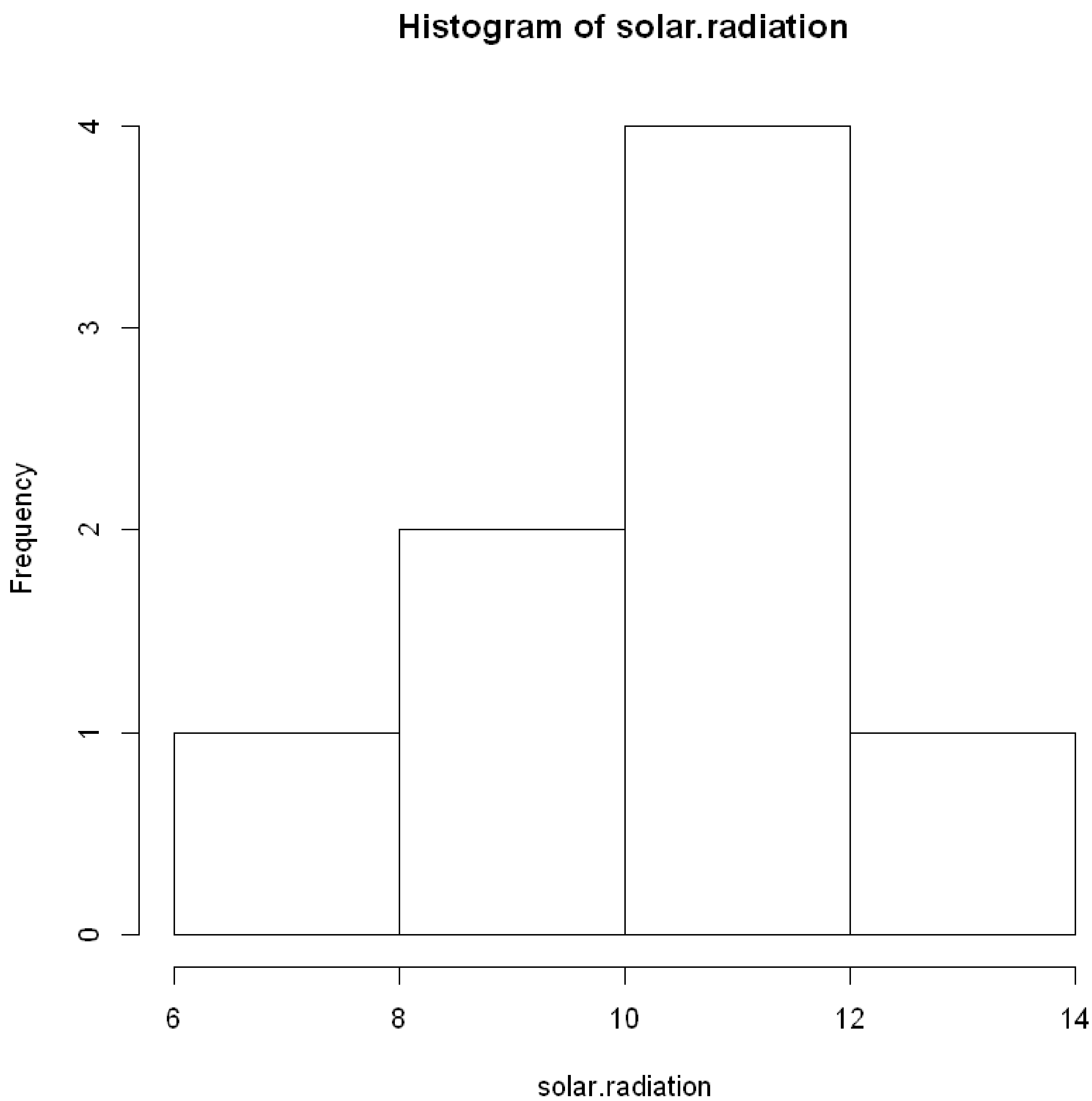
19.975

20.65

16.3 · 22.2

3.525

In [55]:
```
#d. Plot a histogram of the solar.radiation, sr10.

hist(solar.radiation)
```



Question 5. Use the inbuilt data car and uses the possible graphical plots using **ggplot2** graphical packages.

In [56]:
```
data(mtcars)
```

In [57]:
```
?mtcars
```

In [58]:
```
mtcars
```

A data.frame: 32 × 11

	mpg	cyl	disp	hp	drat	wt	qsec	vs	am	gear	carb
	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
Mazda RX4	21.0	6	160.0	110	3.90	2.620	16.46	0	1	4	4
Mazda RX4 Wag	21.0	6	160.0	110	3.90	2.875	17.02	0	1	4	4
Datsun 710	22.8	4	108.0	93	3.85	2.320	18.61	1	1	4	1
Hornet 4 Drive	21.4	6	258.0	110	3.08	3.215	19.44	1	0	3	1
Hornet Sportabout	18.7	8	360.0	175	3.15	3.440	17.02	0	0	3	2
Valiant	18.1	6	225.0	105	2.76	3.460	20.22	1	0	3	1
Duster 360	14.3	8	360.0	245	3.21	3.570	15.84	0	0	3	4
Merc 240D	24.4	4	146.7	62	3.69	3.190	20.00	1	0	4	2
Merc 230	22.8	4	140.8	95	3.92	3.150	22.90	1	0	4	2
Merc 280	19.2	6	167.6	123	3.92	3.440	18.30	1	0	4	4
Merc 280C	17.8	6	167.6	123	3.92	3.440	18.90	1	0	4	4
Merc 450SE	16.4	8	275.8	180	3.07	4.070	17.40	0	0	3	3
Merc 450SL	17.3	8	275.8	180	3.07	3.730	17.60	0	0	3	3
Merc 450SLC	15.2	8	275.8	180	3.07	3.780	18.00	0	0	3	3
Cadillac Fleetwood	10.4	8	472.0	205	2.93	5.250	17.98	0	0	3	4
Lincoln Continental	10.4	8	460.0	215	3.00	5.424	17.82	0	0	3	4
Chrysler Imperial	14.7	8	440.0	230	3.23	5.345	17.42	0	0	3	4
Fiat 128	32.4	4	78.7	66	4.08	2.200	19.47	1	1	4	1
Honda Civic	30.4	4	75.7	52	4.93	1.615	18.52	1	1	4	2
Toyota Corolla	33.9	4	71.1	65	4.22	1.835	19.90	1	1	4	1
Toyota Corona	21.5	4	120.1	97	3.70	2.465	20.01	1	0	3	1
Dodge Challenger	15.5	8	318.0	150	2.76	3.520	16.87	0	0	3	2
AMC Javelin	15.2	8	304.0	150	3.15	3.435	17.30	0	0	3	2
Camaro Z28	13.3	8	350.0	245	3.73	3.840	15.41	0	0	3	4
Pontiac Firebird	19.2	8	400.0	175	3.08	3.845	17.05	0	0	3	2
Fiat X1-9	27.3	4	79.0	66	4.08	1.935	18.90	1	1	4	1
Porsche 914-2	26.0	4	120.3	91	4.43	2.140	16.70	0	1	5	2
Lotus Europa	30.4	4	95.1	113	3.77	1.513	16.90	1	1	5	2
Ford Pantera L	15.8	8	351.0	264	4.22	3.170	14.50	0	1	5	4
Ferrari Dino	19.7	6	145.0	175	3.62	2.770	15.50	0	1	5	6
Maserati Bora	15.0	8	301.0	335	3.54	3.570	14.60	0	1	5	8

	mpg	cyl	disp	hp	drat	wt	qsec	vs	am	gear	carb
	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
Volvo 142E	21.4	4	121.0	109	4.11	2.780	18.60	1	1	4	2

In [59]:
```
tail(mtcars)
```

A data.frame: 6 × 11

	mpg	cyl	disp	hp	drat	wt	qsec	vs	am	gear	carb
	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
Porsche 914-2	26.0	4	120.3	91	4.43	2.140	16.7	0	1	5	2
Lotus Europa	30.4	4	95.1	113	3.77	1.513	16.9	1	1	5	2
Ford Pantera L	15.8	8	351.0	264	4.22	3.170	14.5	0	1	5	4
Ferrari Dino	19.7	6	145.0	175	3.62	2.770	15.5	0	1	5	6
Maserati Bora	15.0	8	301.0	335	3.54	3.570	14.6	0	1	5	8
Volvo 142E	21.4	4	121.0	109	4.11	2.780	18.6	1	1	4	2

In [60]:
```
ncol(mtcars)
nrow(mtcars)
```

11

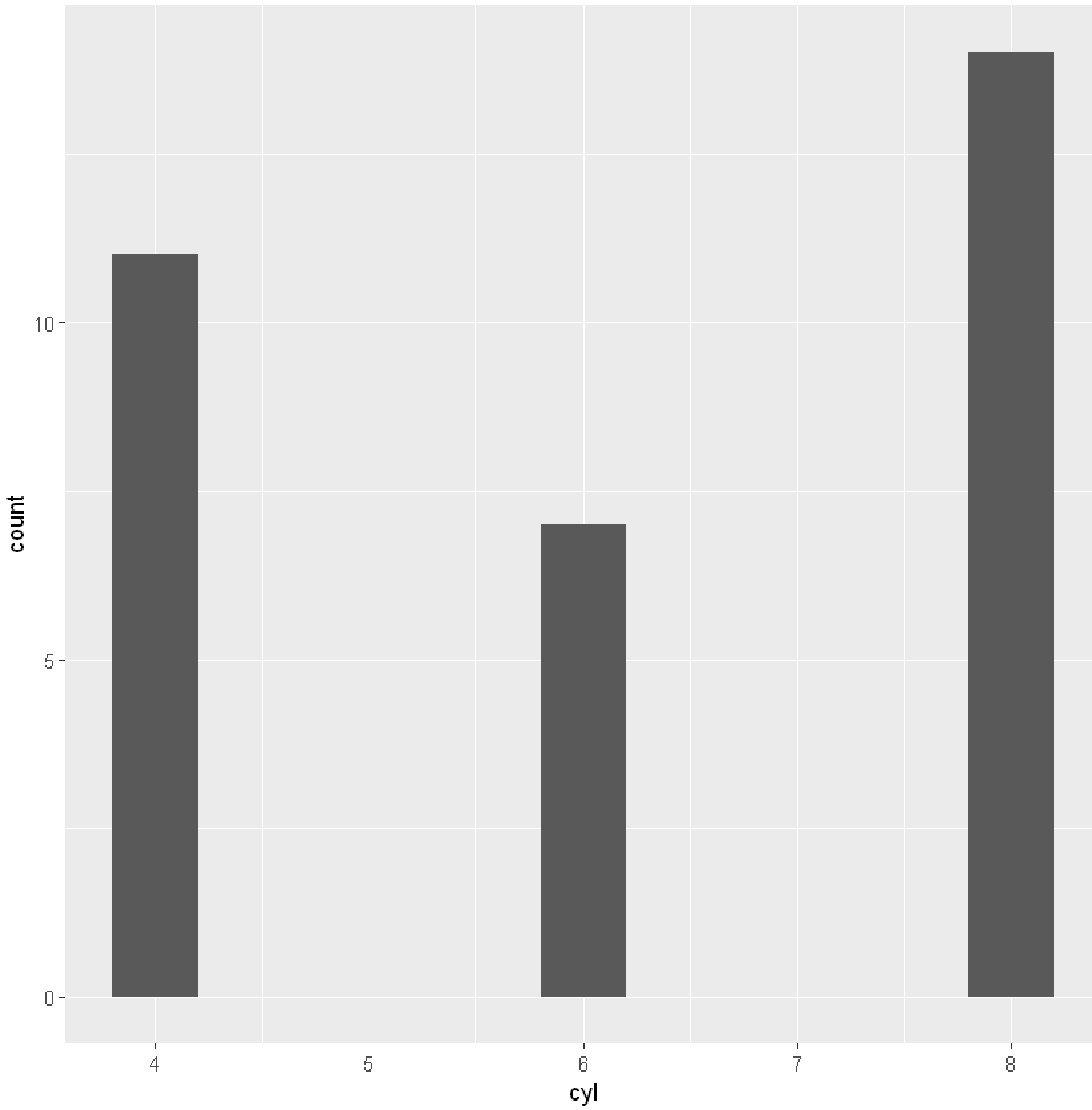
32

In [61]:
```
dim(mtcars)
```

32 · 11

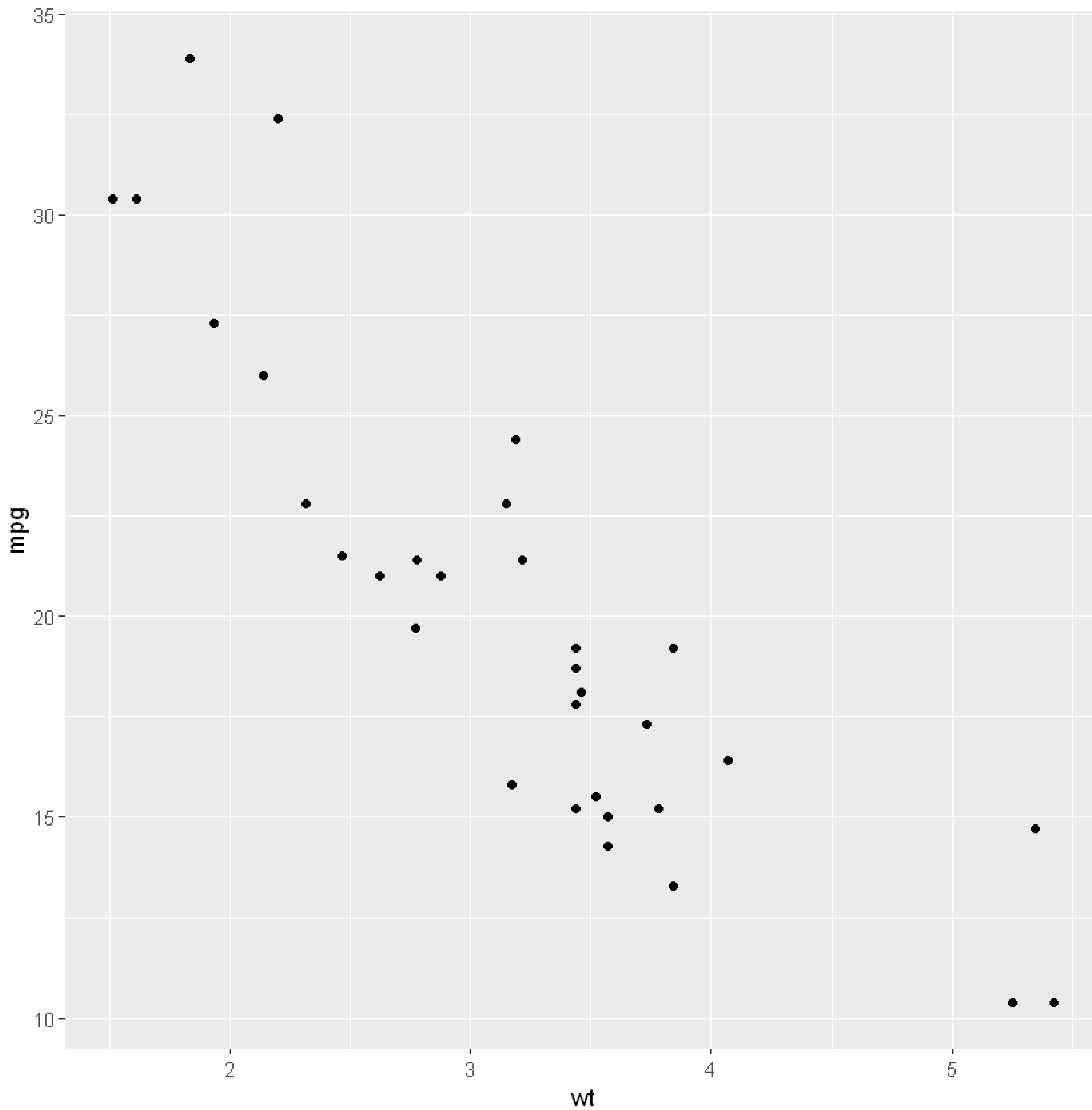
In [62]:
```
library(ggplot2)
```

In [63]:
```
my_plot = ggplot(mtcars,aes(x=cyl)) + geom_histogram(binwidth=0.4)
my_plot
```



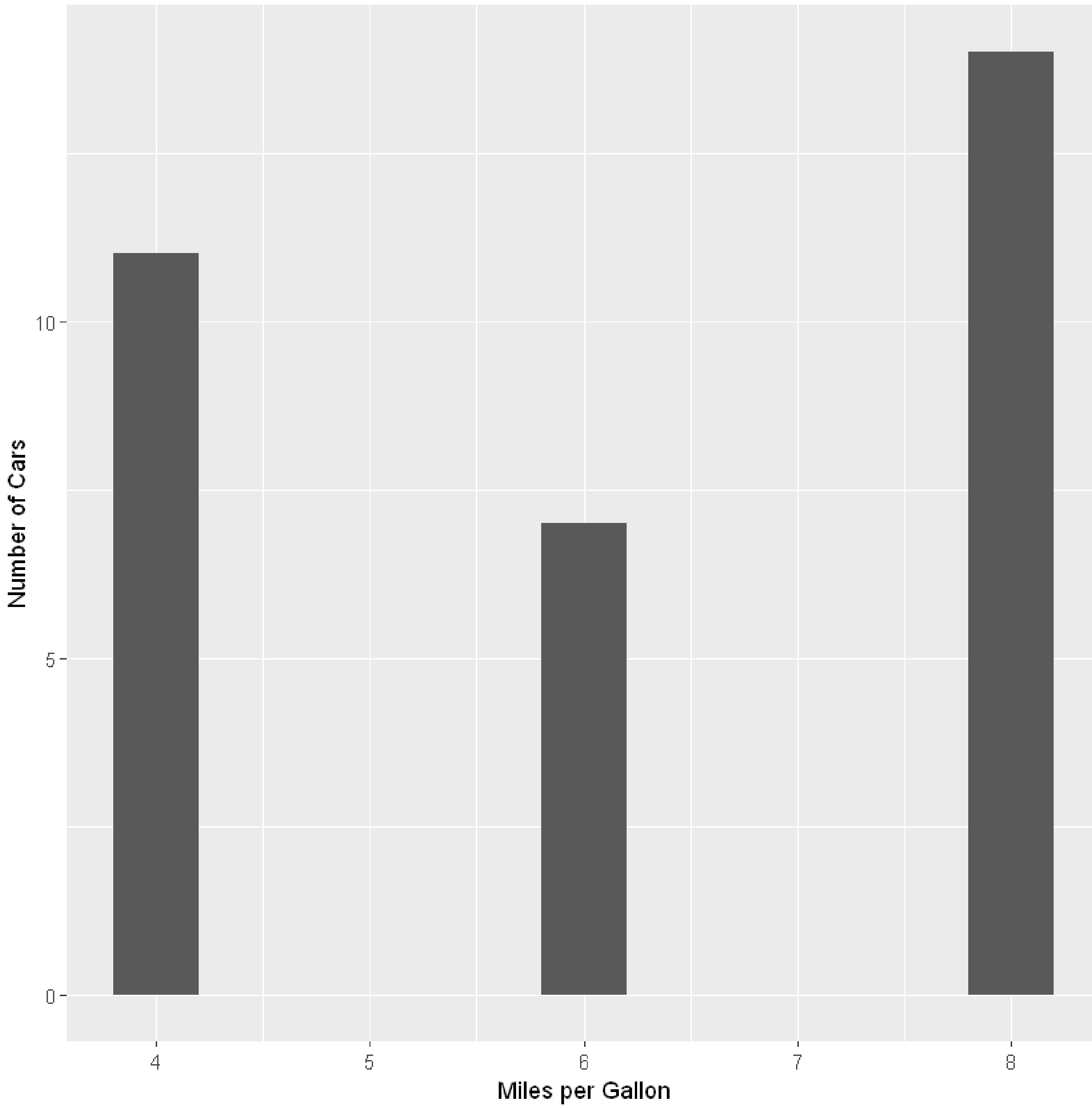
aes = aesthetic here meaning a mapping between a variable and a characteristic of the plot
x relates to the x-axis (ie wt links to x)

In [64]:
```
# using full ggplot syntax
ggplot(data = mtcars, aes(x = wt, y = mpg)) +
       geom_point()
```

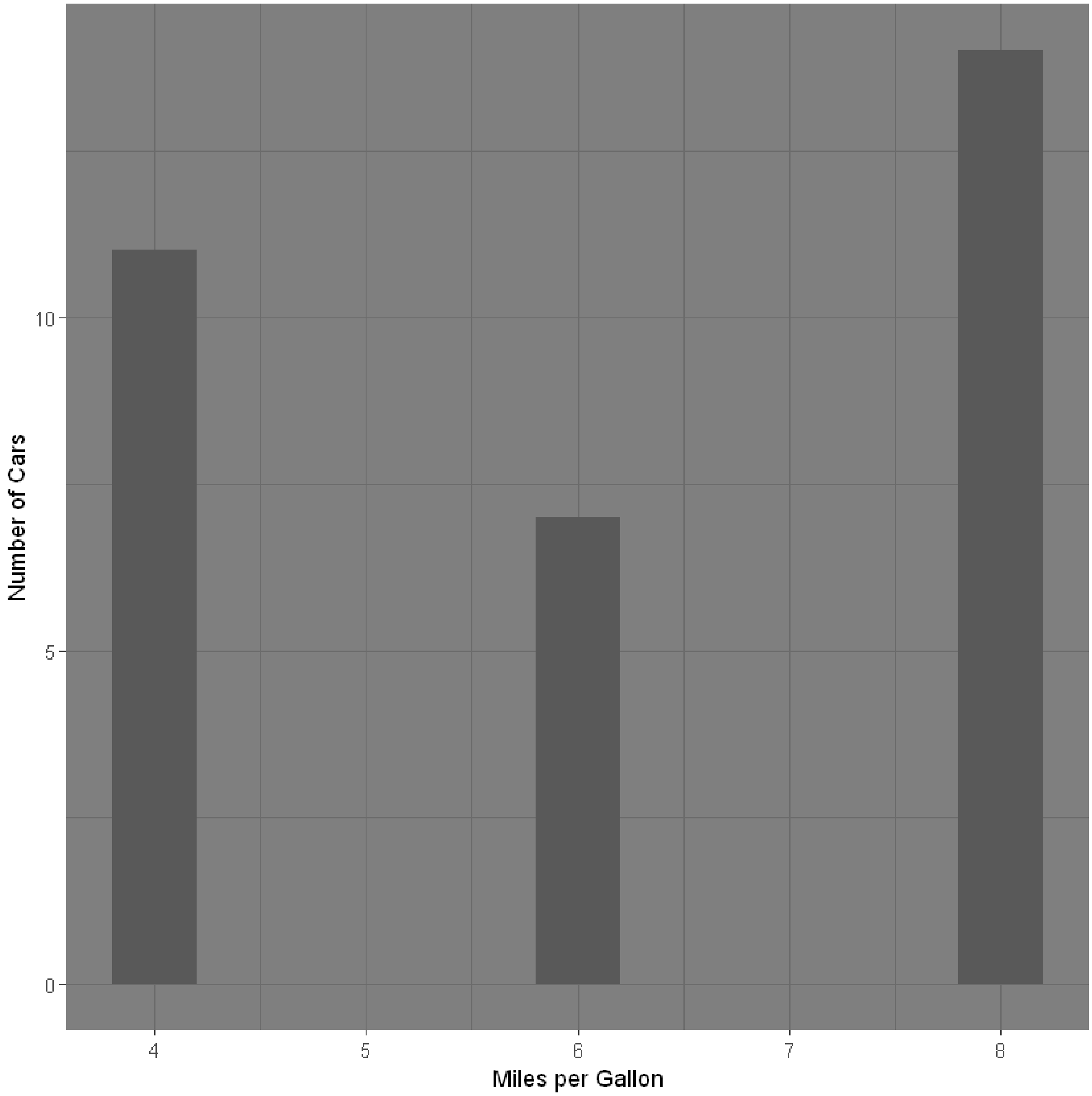


labs adds labels to a plot

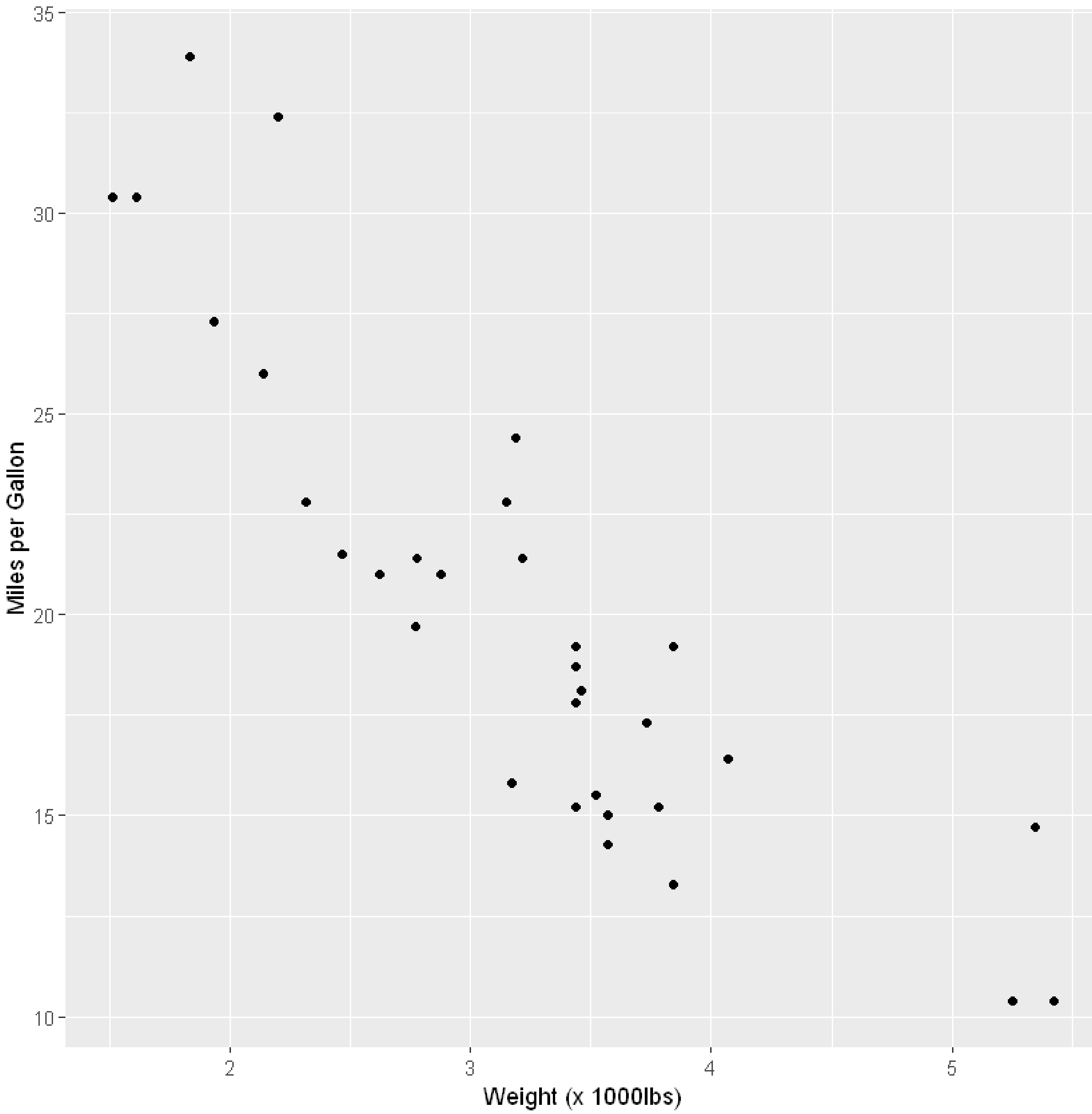
In [65]:
```
my_plot + xlab('Miles per Gallon')+ylab('Number of Cars')
```



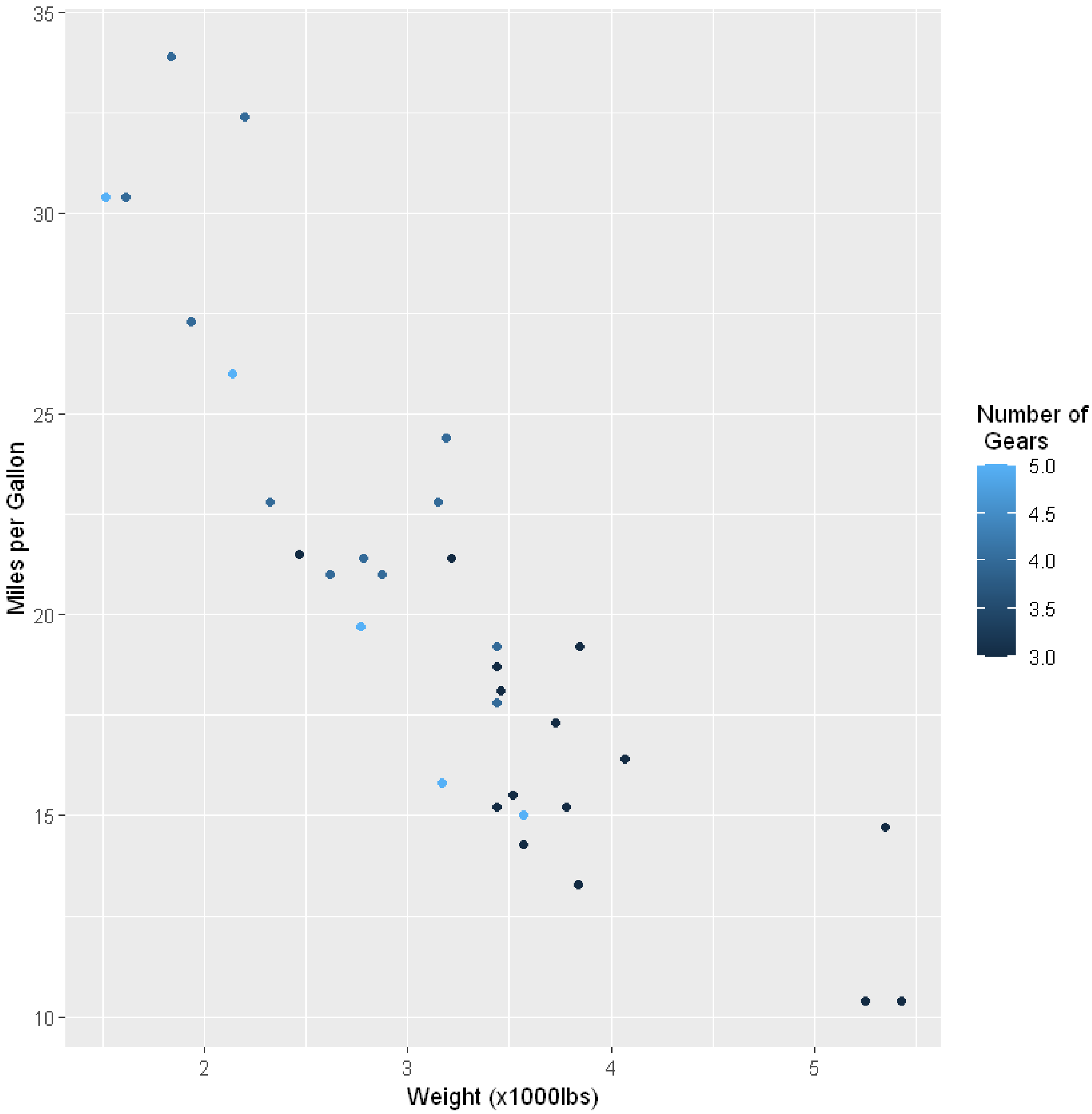
In [66]:
```
#For changing the theme

my_plot + xlab('Miles per Gallon')+ylab('Number of Cars')  + theme_dark()
```



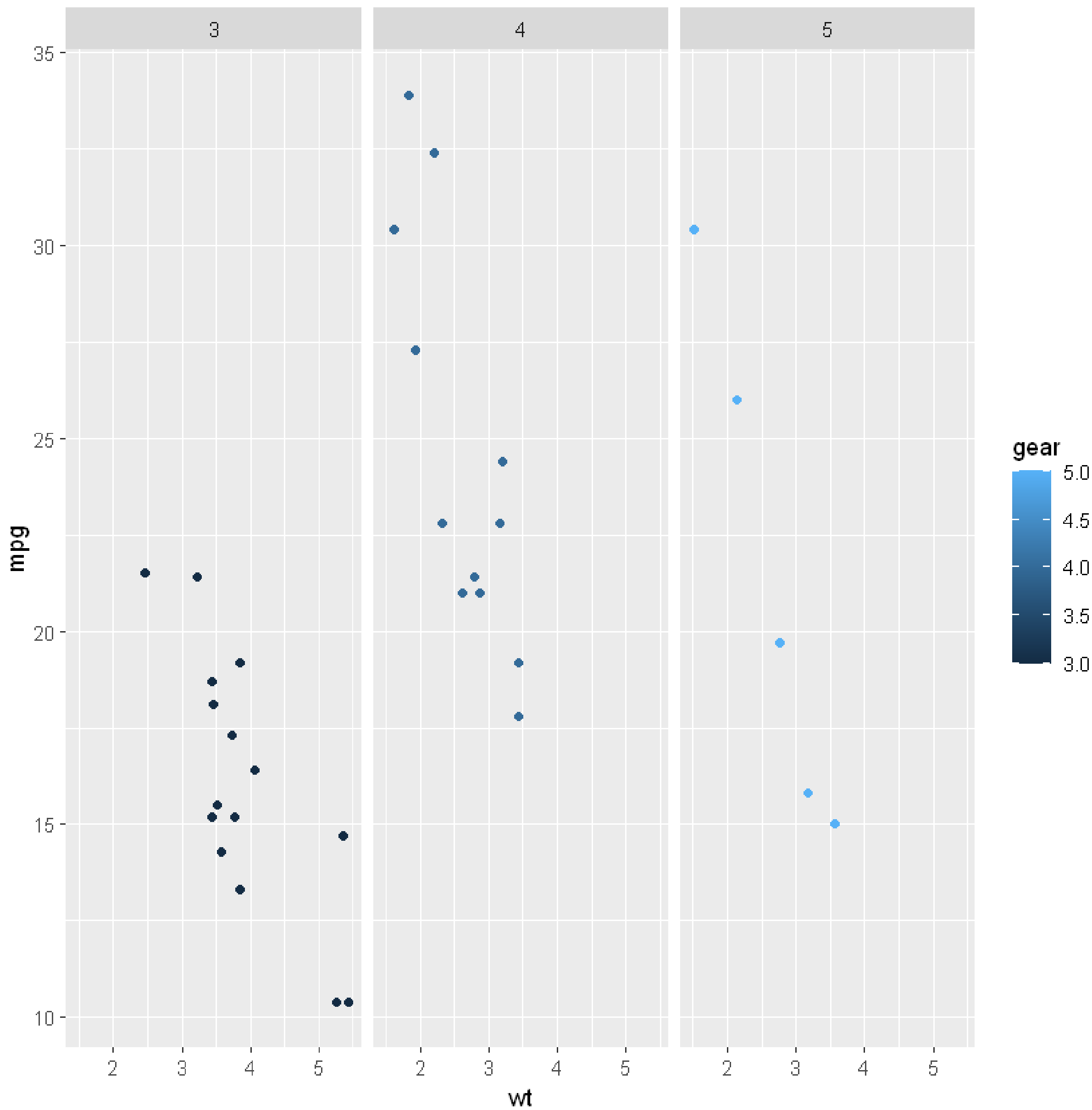
In [67]:
```
my_scatplot <- ggplot(mtcars,aes(x=wt,y=mpg)) + geom_point()
my_scatplot + xlab('Weight (x 1000lbs)') + ylab('Miles per Gallon')
```



In [68]:
```
my_scatplot <- ggplot(mtcars,aes(x=wt,y=mpg,col=gear)) + geom_point()
my_scatplot + labs(x='Weight (x1000lbs)',y='Miles per Gallon',colour='Number of\
```



In [69]:
```
my_scatplot <- ggplot(mtcars,aes(x=wt,y=mpg,col=gear)) + geom_point()
my_scatplot + facet_grid(~gear)
```



In [70]:
```
my_scatplot <- ggplot(mtcars,aes(x=wt,y=mpg)) + geom_point()
my_scatplot + xlab('Weight (x 1000lbs)') + ylab('Miles per Gallon') + geom_smoot
```

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