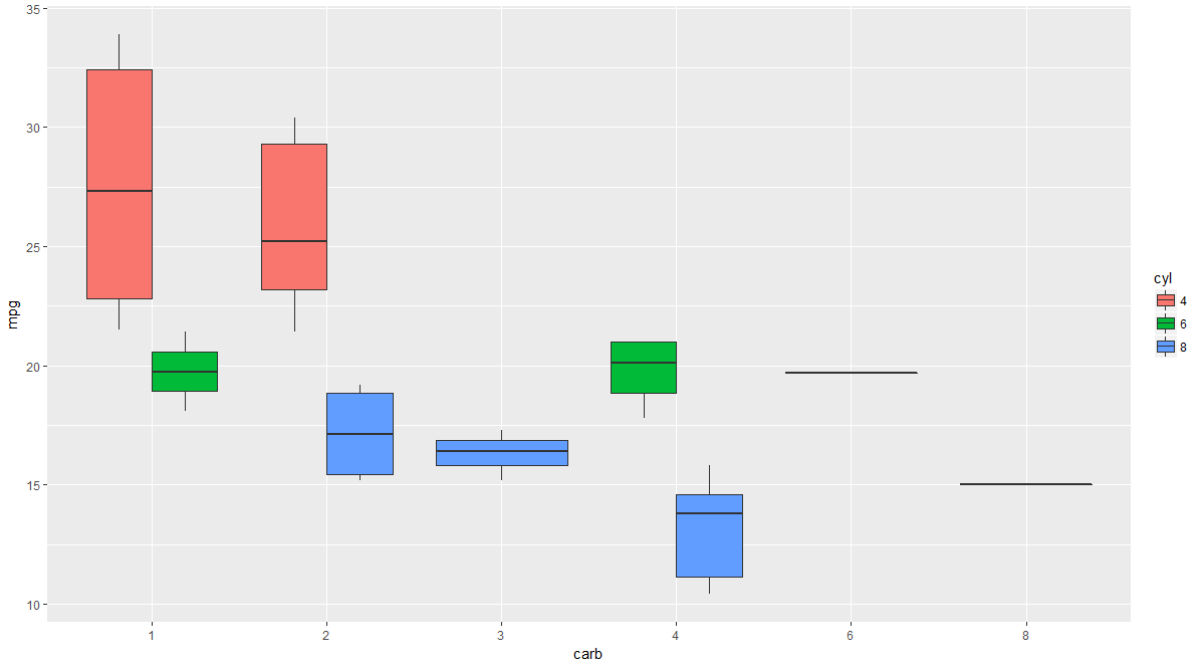
Assignment - 3

Session 7 – Basic Statistics

1. Create a box and whisker plot by class using mtcars dataset.

Ans:

```
> str(mtcars)
'data.frame':  32 obs. of  11 variables:
 $ mpg : num  21 21 22.8 21.4 18.7 18.1 14.3 24.4 22.8 19.2 ...
 $ cyl : Factor w/ 3 levels "4","6","8": 2 2 1 2 3 2 3 1 1 2 ...
 $ disp: Factor w/ 27 levels "71.1","75.7",..: 13 13 6 16 23 15 23 12 10 1
4 ...
 $ hp  : num  110 110 93 110 175 105 245 62 95 123 ...
 $ drat: num  3.9 3.9 3.85 3.08 3.15 2.76 3.21 3.69 3.92 3.92 ...
 $ wt  : num  2.62 2.88 2.32 3.21 3.44 ...
 $ qsec: num  16.5 17 18.6 19.4 17 ...
 $ vs  : Factor w/ 2 levels "0","1": 1 1 2 2 1 2 1 2 2 2 ...
 $ am  : Factor w/ 2 levels "0","1": 2 2 2 1 1 1 1 1 1 1 ...
 $ gear: Factor w/ 3 levels "3","4","5": 2 2 2 1 1 1 1 2 2 2 ...
 $ carb: Factor w/ 6 levels "1","2","3","4",..: 4 4 1 1 2 1 4 2 2 4 ...
> View(mtcars)
> library(ggplot2)
> library(dplyr)
> mtcars1<- mutate(mtcars,
+                  cyl=as.factor(cyl),
+                  disp=as.factor(disp),
+                  vs=as.factor(vs),
+                  am=as.factor(am),
+                  gear=as.factor(gear),
+                  carb=as.factor(carb),
+                  mpg=mpg, hp=hp, drat=drat, qsec=qsec)
> str(mtcars1)
'data.frame':  32 obs. of  11 variables:
 $ mpg : num  21 21 22.8 21.4 18.7 18.1 14.3 24.4 22.8 19.2 ...
 $ cyl : Factor w/ 3 levels "4","6","8": 2 2 1 2 3 2 3 1 1 2 ...
 $ disp: Factor w/ 27 levels "71.1","75.7",..: 13 13 6 16 23 15 23 12 10 1
4 ...
 $ hp  : num  110 110 93 110 175 105 245 62 95 123 ...
 $ drat: num  3.9 3.9 3.85 3.08 3.15 2.76 3.21 3.69 3.92 3.92 ...
 $ wt  : num  2.62 2.88 2.32 3.21 3.44 ...
 $ qsec: num  16.5 17 18.6 19.4 17 ...
 $ vs  : Factor w/ 2 levels "0","1": 1 1 2 2 1 2 1 2 2 2 ...
 $ am  : Factor w/ 2 levels "0","1": 2 2 2 1 1 1 1 1 1 1 ...
 $ gear: Factor w/ 3 levels "3","4","5": 2 2 2 1 1 1 1 2 2 2 ...
 $ carb: Factor w/ 6 levels "1","2","3","4",..: 4 4 1 1 2 1 4 2 2 4 ...
> # distribution of mpg for each carb
> boxplot(mpg~carb, data=mtcars1, col=heat.colors(5))
> # distribution of mpg for each carb per cycle
> ggplot(mtcars1, aes(x=carb, y=mpg, fill=cyl))+geom_boxplot()
```


35
30
25
20
15
10
mpg
1
2
3
4
6
8
carb
cyl
4
6
8