

hw2

Anyesha Ray, Git Hub: anyeshar

January 30, 2020

R Command Questions

Question 1

```
fahrenheit <- c(24,15)
celcius <- c()
for (temp in fahrenheit) {
  celcius <- c(celcius, (temp-32)*5/9)
}
print(celcius)
```

```
## [1] -4.444444 -9.444444
```

Question 2

```
vec200 <- c()
for (x in c(1:200)) {
  if (x %% 2 == 0) {
    vec200 <- c(vec200, 1)
  }
  else {
    vec200 <- c(vec200, 3)
  }
}
print(vec200)
```

[illegible]

Question 3

```
numPerfect <- 0
for (x in c(1:2001)) {
  sqrt <- x^0.5
  if (sqrt == round(sqrt)) {
    numPerfect = numPerfect + 1
  }
}
print(numPerfect)
```

```
## [1] 44
```

MPG Questions

Question 1

```
library(ggplot2)
summary(mpg)
```

```
## manufacturer      model      displ      year
## Length:234      Length:234      Min.   :1.600      Min.   :1999
## Class :character  Class :character  1st Qu.:2.400      1st Qu.:1999
## Mode  :character  Mode  :character  Median :3.300      Median :2004
##                                     Mean   :3.472      Mean   :2004
##                                     3rd Qu.:4.600      3rd Qu.:2008
##                                     Max.   :7.000      Max.   :2008
##      cyl      trans      drv      cty
## Min.   :4.000      Length:234      Length:234      Min.   : 9.00
## 1st Qu.:4.000      Class :character  Class :character  1st Qu.:14.00
## Median :6.000      Mode  :character  Mode  :character  Median :17.00
## Mean   :5.889                                     Mean   :16.86
## 3rd Qu.:8.000                                     3rd Qu.:19.00
## Max.   :8.000                                     Max.   :35.00
##      hwy      fl      class
## Min.   :12.00      Length:234      Length:234
## 1st Qu.:18.00      Class :character  Class :character
## Median :24.00      Mode  :character  Mode  :character
## Mean   :23.44
## 3rd Qu.:27.00
## Max.   :44.00
```

```
head(mpg)
```

```
## # A tibble: 6 x 11
##   manufacturer model displ year   cyl trans  drv      cty   hwy fl      class
##   <chr>         <chr> <dbl> <int> <int> <chr> <chr> <int> <int> <chr> <chr>
## 1 audi         a4      1.8  1999     4 auto(~ f      18    29 p      comp~
## 2 audi         a4      1.8  1999     4 manua~ f      21    29 p      comp~
## 3 audi         a4      2    2008     4 manua~ f      20    31 p      comp~
## 4 audi         a4      2    2008     4 auto(~ f      21    30 p      comp~
## 5 audi         a4      2.8  1999     6 auto(~ f      16    26 p      comp~
## 6 audi         a4      2.8  1999     6 manua~ f      18    26 p      comp~
```

```
#top 3
sort <- mpg[order(-mpg$hwy), ]
top3 <- sort[1:3, ]
print(top3)
```

```
## # A tibble: 3 x 11
##   manufacturer model displ year   cyl trans  drv      cty   hwy fl      class
##   <chr>         <chr> <dbl> <int> <int> <chr> <chr> <int> <int> <chr> <chr>
## 1 volkswagen   jetta   1.9  1999     4 manu~ f      33    44 d      comp~
## 2 volkswagen   new b~   1.9  1999     4 manu~ f      35    44 d      subc~
## 3 volkswagen   new b~   1.9  1999     4 auto~ f      29    41 d      subc~
```

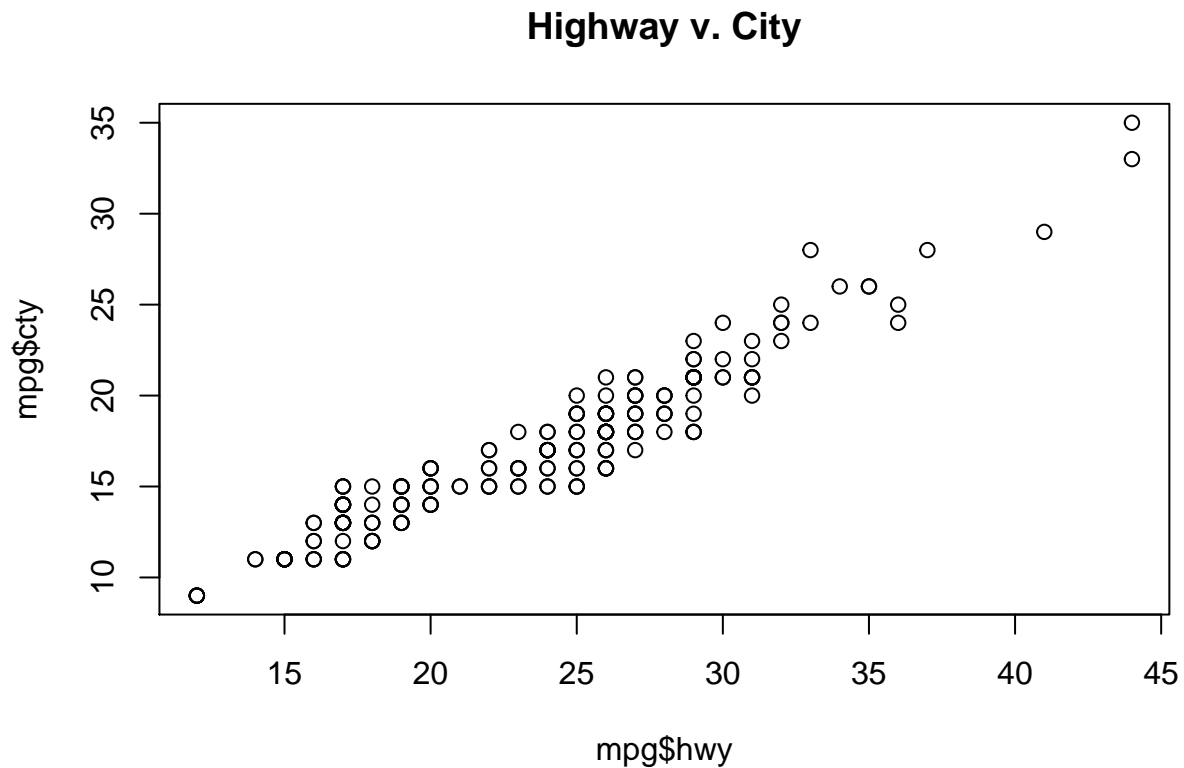
Question 2

```
Compact <- mpg[mpg$class == 'compact', ]
numCompact <- length(unique(Compact$model))
print(numCompact)
```

```
## [1] 8
```

Question 3

```
plot(mpg$hwy,mpg$cty, main = "Highway v. City")
```



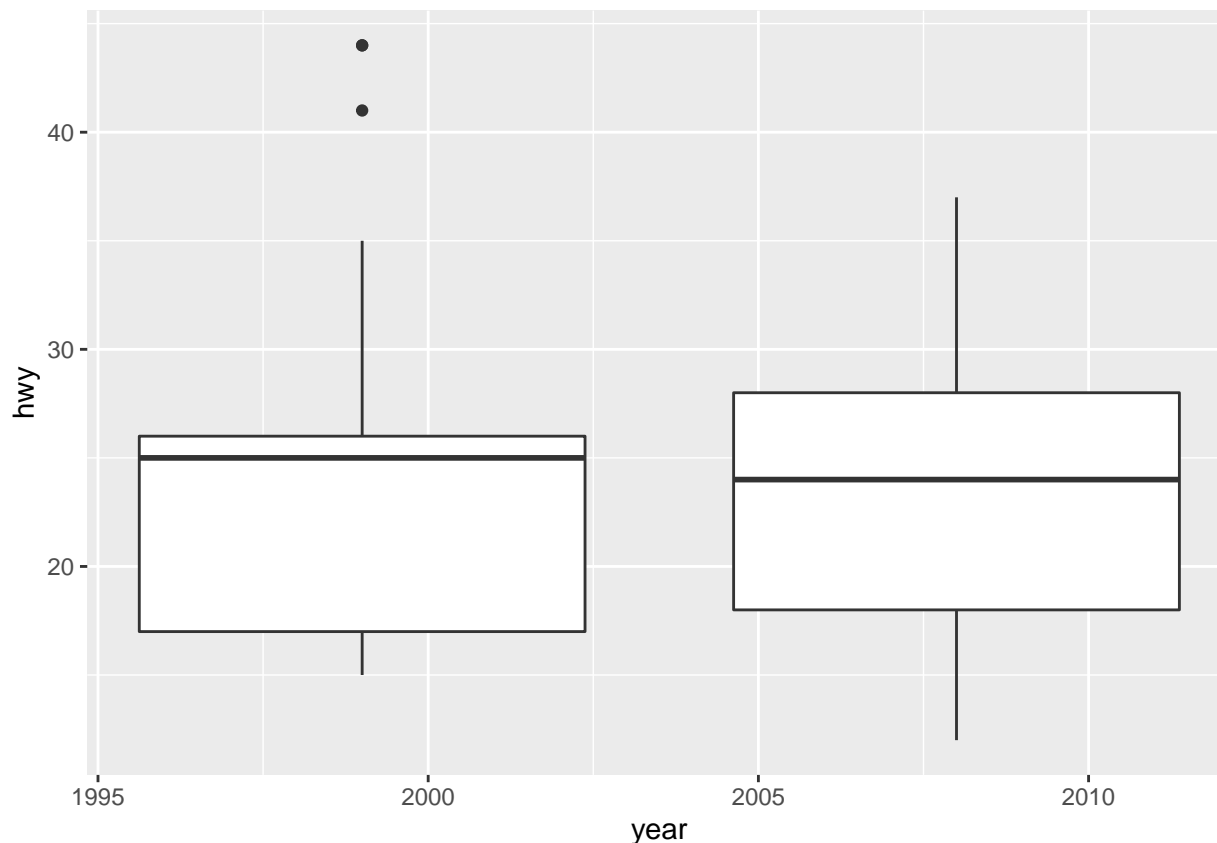
There is a positive, linear, and moderately strong to strong relationship between the two variables. Not all the data points may be shown because of overlapping.

Question 4 Here boxplots make the most sense because they are a visual representation of the summary statistics.

```
#using ggplot
```

```
#hwy
```

```
hwy_boxplot <- ggplot(mpg, aes(group=year, x=year, y=hwy)) + geom_boxplot()  
hwy_boxplot
```



```
summary(mpg[mpg$year == '1999', ])
```

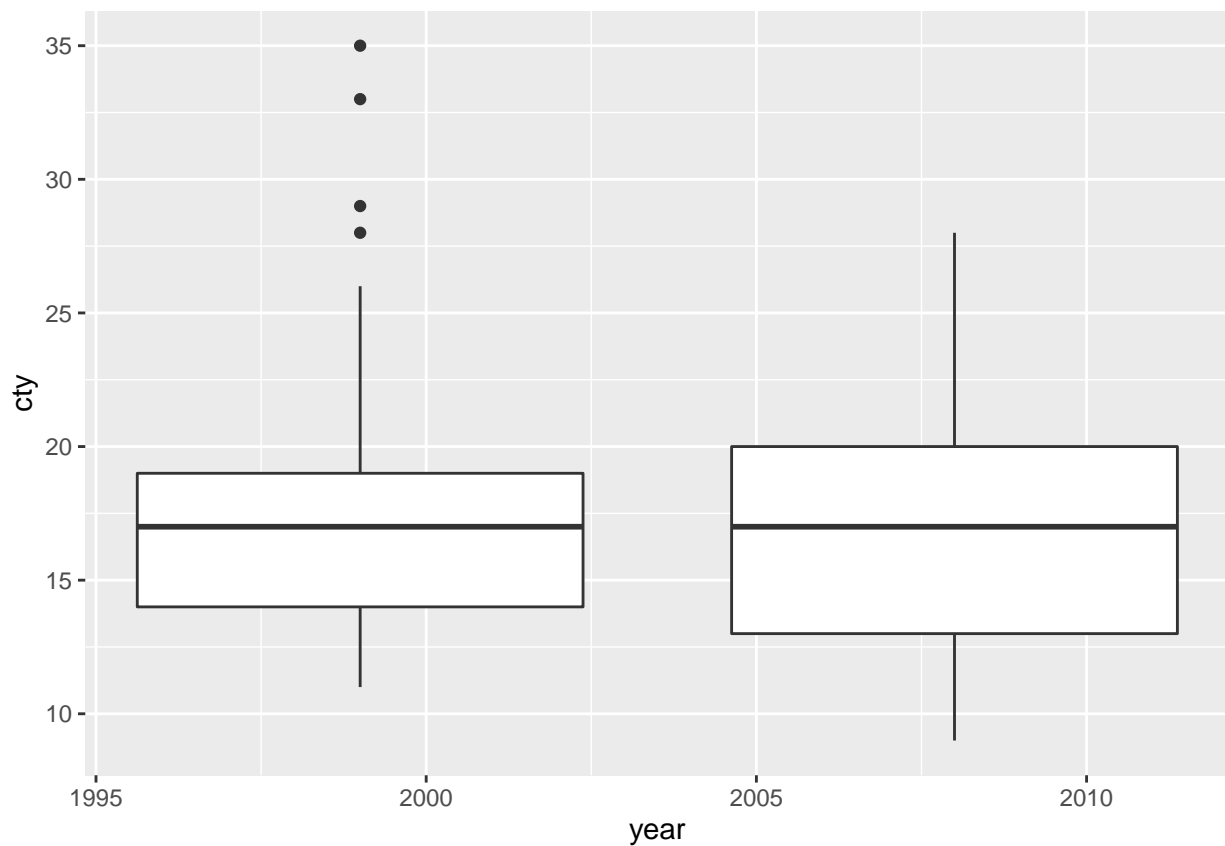
```
## manufacturer      model      displ      year
## Length:117      Length:117      Min.   :1.600      Min.   :1999
## Class :character Class :character 1st Qu.:2.200      1st Qu.:1999
## Mode  :character Mode  :character Median :3.000      Median :1999
##                                     Mean  :3.281      Mean  :1999
##                                     3rd Qu.:4.000      3rd Qu.:1999
##                                     Max.   :6.500      Max.   :1999
##      cyl      trans      drv      cty
## Min.   :4.000      Length:117      Length:117      Min.   :11.00
## 1st Qu.:4.000      Class :character      Class :character 1st Qu.:14.00
## Median :6.000      Mode  :character      Mode  :character Median :17.00
## Mean   :5.692                                     Mean  :17.02
## 3rd Qu.:6.000                                     3rd Qu.:19.00
## Max.   :8.000                                     Max.   :35.00
##      hwy      fl      class
## Min.   :15.00      Length:117      Length:117
## 1st Qu.:17.00      Class :character      Class :character
## Median :25.00      Mode  :character      Mode  :character
## Mean   :23.43
## 3rd Qu.:26.00
## Max.   :44.00
```

```
summary(mpg[mpg$year == '2008', ])
```

```
## manufacturer      model      displ      year
```

```
## Length:117      Length:117      Min.   :1.800  Min.   :2008
## Class :character Class :character 1st Qu.:2.500  1st Qu.:2008
## Mode  :character Mode  :character Median :3.600  Median :2008
##                                     Mean  :3.662  Mean  :2008
##                                     3rd Qu.:4.700  3rd Qu.:2008
##                                     Max.   :7.000  Max.   :2008
##      cyl      trans      drv      cty
## Min.   :4.000  Length:117  Length:117  Min.   : 9.0
## 1st Qu.:4.000  Class :character Class :character 1st Qu.:13.0
## Median :6.000  Mode  :character Mode  :character Median :17.0
## Mean   :6.085                                     Mean  :16.7
## 3rd Qu.:8.000                                     3rd Qu.:20.0
## Max.   :8.000                                     Max.   :28.0
##      hwy      fl      class
## Min.   :12.00  Length:117  Length:117
## 1st Qu.:18.00  Class :character Class :character
## Median :24.00  Mode  :character Mode  :character
## Mean   :23.45
## 3rd Qu.:28.00
## Max.   :37.00
```

```
#cty
ctyBox <- ggplot(mpg, aes(group=year, x=year, y=cty)) + geom_boxplot()
ctyBox
```



```
summary(mpg[mpg$year == '1999', ])
```

```
## manufacturer      model      displ      year
## Length:117      Length:117      Min.    :1.600      Min.    :1999
## Class :character Class :character 1st Qu.:2.200      1st Qu.:1999
## Mode  :character Mode  :character Median :3.000      Median :1999
##                                     Mean  :3.281      Mean   :1999
##                                     3rd Qu.:4.000      3rd Qu.:1999
##                                     Max.   :6.500      Max.   :1999
##      cyl      trans      drv      cty
## Min.    :4.000      Length:117      Length:117      Min.    :11.00
## 1st Qu.:4.000      Class :character Class :character 1st Qu.:14.00
## Median :6.000      Mode  :character Mode  :character Median :17.00
## Mean    :5.692                                     Mean   :17.02
## 3rd Qu.:6.000                                     3rd Qu.:19.00
## Max.    :8.000                                     Max.   :35.00
##      hwy      fl      class
## Min.    :15.00      Length:117      Length:117
## 1st Qu.:17.00      Class :character Class :character
## Median :25.00      Mode  :character Mode  :character
## Mean    :23.43
## 3rd Qu.:26.00
## Max.    :44.00
```

```
summary(mpg[mpg$year == '2008', ])
```

```
## manufacturer      model      displ      year
## Length:117      Length:117      Min.    :1.800      Min.    :2008
## Class :character Class :character 1st Qu.:2.500      1st Qu.:2008
## Mode  :character Mode  :character Median :3.600      Median :2008
##                                     Mean   :3.662      Mean   :2008
##                                     3rd Qu.:4.700      3rd Qu.:2008
##                                     Max.    :7.000      Max.   :2008
##      cyl      trans      drv      cty
## Min.    :4.000      Length:117      Length:117      Min.    : 9.0
## 1st Qu.:4.000      Class :character Class :character 1st Qu.:13.0
## Median :6.000      Mode  :character Mode  :character Median :17.0
## Mean    :6.085                                     Mean   :16.7
## 3rd Qu.:8.000                                     3rd Qu.:20.0
## Max.    :8.000                                     Max.   :28.0
##      hwy      fl      class
## Min.    :12.00      Length:117      Length:117
## 1st Qu.:18.00      Class :character Class :character
## Median :24.00      Mode  :character Mode  :character
## Mean    :23.45
## 3rd Qu.:28.00
## Max.    :37.00
```

For both the city and highway the mean mpg appears to be about the same. It is different in the city and highway but within the two they are very similar graphs and numbers.