Lab Report 1

Bingying Liu

January 13, 2019

Exercise 1

Create a code chunk and set the header parameter to TRUE and print out the top rows of the table with head() as above.

```
data <- read.table(file = url("http://www2.stat.duke.edu/~pdh10/FCBS/Exercises/azdiabetes.dat"), header = TRUE)
head(data)
```

```
##   npreg glu bp skin  bmi   ped age diabetes
## 1     5  86 68   28 30.2 0.364  24       No
## 2     7 195 70   33 25.1 0.163  55      Yes
## 3     5  77 82   41 35.8 0.156  35       No
## 4     0 165 76   43 47.9 0.259  26       No
## 5     0 107 60   25 26.4 0.133  23       No
## 6     5  97 76   27 35.6 0.378  52      Yes
```

Exercise 2

Generate a sequence of 100 equispaced real numbers from 0 to 1 and store it in a variable called seq2.

```
seq2 <- seq(from = 0, to = 1, length.out = 101)
seq2
```

```
##   [1] 0.00 0.01 0.02 0.03 0.04 0.05 0.06 0.07 0.08 0.09 0.10 0.11 0.12 0.13 0.14
##  [16] 0.15 0.16 0.17 0.18 0.19 0.20 0.21 0.22 0.23 0.24 0.25 0.26 0.27 0.28 0.29
##  [31] 0.30 0.31 0.32 0.33 0.34 0.35 0.36 0.37 0.38 0.39 0.40 0.41 0.42 0.43 0.44
##  [46] 0.45 0.46 0.47 0.48 0.49 0.50 0.51 0.52 0.53 0.54 0.55 0.56 0.57 0.58 0.59
##  [61] 0.60 0.61 0.62 0.63 0.64 0.65 0.66 0.67 0.68 0.69 0.70 0.71 0.72 0.73 0.74
##  [76] 0.75 0.76 0.77 0.78 0.79 0.80 0.81 0.82 0.83 0.84 0.85 0.86 0.87 0.88 0.89
##  [91] 0.90 0.91 0.92 0.93 0.94 0.95 0.96 0.97 0.98 0.99 1.00
```

Exercise 3

Sort the entries in seq3 from greatest to least.s

```
seq3 <- seq(from = -3, to = 3, by = .5)
seq3 <- sort(seq3, decreasing = TRUE)
seq3
```

```
##  [1]  3.0  2.5  2.0  1.5  1.0  0.5  0.0 -0.5 -1.0 -1.5 -2.0 -2.5 -3.0
```

Exercise 4

Find the variance of each row of mat5.

```
# generate large matrix
mat5 <- matrix(seq(1,100,1), nrow = 4, ncol = 25, byrow = T)
apply(X= mat5, MARGIN = 1, FUN = var)
```

```
## [1] 54.16667 54.16667 54.16667 54.16667
```

Exercise 5

Generate 500 samples from a Beta distribution with shape parameter [a,b]=[0.5,0.5] and store the samples in a variable called W.

```
w <- rbeta(500, shape1 = 0.5, shape2 = 0.5)
```

Exercise 6

Browse online resources (some below), or use code from above to make a few plots of your own.

I used the mpg dataset (which is the fuel economy data from 1999 and 2008) to create some plots using ggplot darkmode.

```
p <- ggplot(mpg) +
  geom_point(aes(cty, hwy, color = class)) +
  scale_y_continuous() +
  guides(color = guide_legend(reverse = TRUE)) +
  labs(title = "Highway fuel efficiency by City fuel efficiency and Type of car",
       x = "City fuel efficiency (miles per gallon)",
       y = "Highway fuel Efficiency (miles per gallon)",
       color = "Class")

p + dark_theme_gray()
```
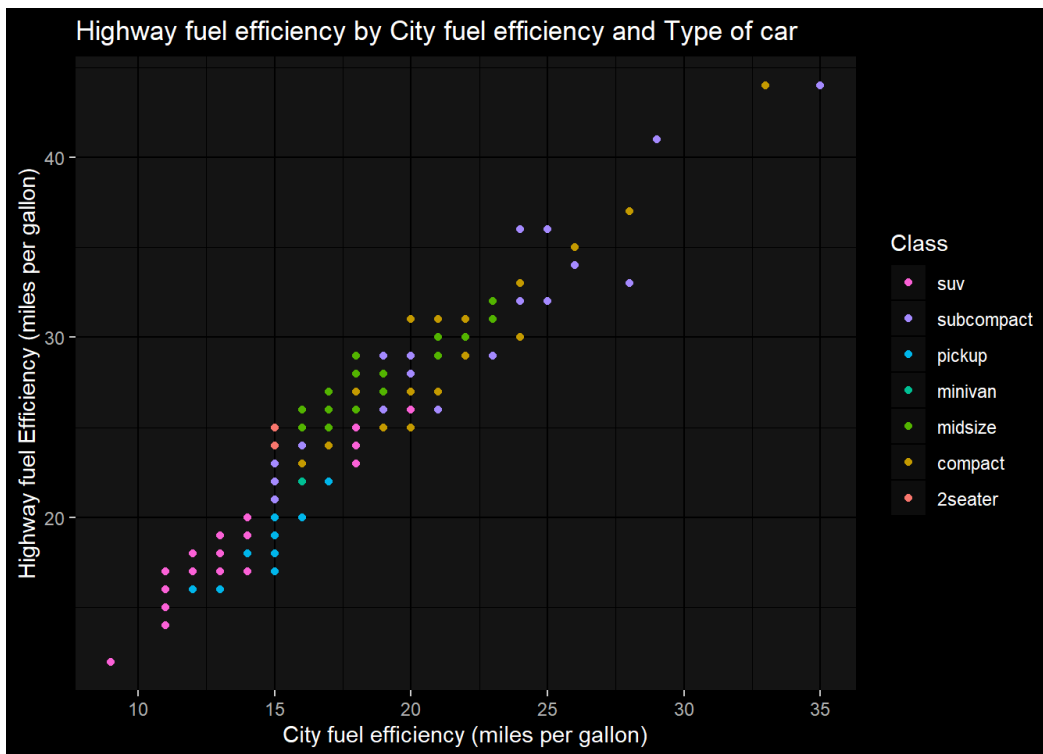


```
p <- ggplot(mpg) +
  geom_boxplot(aes(trans, hwy, color = trans)) +
  coord_flip() +
  guides(color = guide_legend(reverse = TRUE)) +
  labs(title = "Highway fuel efficiency by type of transmission",
       x = "Type of transmission",
       y = "Highway fuel Efficiency (miles per gallon)")

p + dark_theme_gray()
```

Highway fuel efficiency by type of transmission

***