Homework 2

Eric Zou

Honor code:

"I have neither given nor received unauthorized assistance on this assignment." EZ

I receive help from " " and give help to " ".

Please remember to add title and proper labels to your plot.

Problem 1

Create a vector of integers from 0 to 100 (inclusive) and assign it to the variable x. Multiple x by 2 and output the result. Extract the index of x with value larger than or equal to 90. Extract all the even numbers from x.

```
x = 0:100
x= 2*x

y = x>=90
x[y]
```

```
##  [1]  90  92  94  96  98 100 102 104 106 108 110 112 114 116 118 120 122 124 126
## [20] 128 130 132 134 136 138 140 142 144 146 148 150 152 154 156 158 160 162 164
## [39] 166 168 170 172 174 176 178 180 182 184 186 188 190 192 194 196 198 200
```

```
y = x%%2==0
x[y]
```

```
##   [1]   0   2   4   6   8  10  12  14  16  18  20  22  24  26  28  30  32  34
##  [19]  36  38  40  42  44  46  48  50  52  54  56  58  60  62  64  66  68  70
##  [37]  72  74  76  78  80  82  84  86  88  90  92  94  96  98 100 102 104 106
##  [55] 108 110 112 114 116 118 120 122 124 126 128 130 132 134 136 138 140 142
##  [73] 144 146 148 150 152 154 156 158 160 162 164 166 168 170 172 174 176 178
##  [91] 180 182 184 186 188 190 192 194 196 198 200
```

Problem 2

Draw a random sample with size 9 from a standard Gaussian distribution (`rnorm(9)`) and fill it into a 3by3 matrix X. Draw a random sample with size 12 from a standard Gaussian distribution and fill it into a 3by4 matrix Y. Calculate the matrix multiplication between X and Y. What will happen if you use * to multiply X and Y? Why will this happen?

```
sample = rnorm(9)
x=matrix(sample,nrow = 3, ncol = 3)

sample2 = rnorm(12)
y = matrix(sample2, nrow = 3, ncol = 4)
```

```
x%*%y
```

```
##            [,1]        [,2]       [,3]       [,4]
## [1,] -0.4828002 -0.51498442 -2.1791199  0.5844589
## [2,] -0.8874835  0.91016600  0.6475574 -0.3064609
## [3,] -2.7163780 -0.07360048 -4.6738860 -0.1908746
```

```
#x*y doesn't work because * is only for scalars
```

Problem 3

Look at the help document of mtcars data set using (`?mtcars`). Whether `mtcars` is a data.frame or not (hint: use `is.data.frame()` function)? Print out the first few rows of the mtcars data using `head()` function. Find the indices of rows with 6 cylinders and extract the mpg of the cars with 6 cylinders. Calculate the means and standard deviations of the mpg within each category of cylinders using mean() and sd() function.

```
?mtcars
is.data.frame(mtcars)
```

```
## [1] TRUE
```

```
head(mtcars)
```

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

```
y = mtcars$cyl==6

x = 0:32
x = x[y]

mpg = mtcars$mpg[x]

mean(mpg)
```

```
## [1] 20.24286
```

```
sd(mpg)
```

```
## [1] 2.523131
```

Problem 4

In mtcars dataset, find the indices of rows with hp large than its mean. Create a vector to store the values of hp for cars with hp larger than its mean and another vector to store the values of hp for cars with smaller than its mean. Calculate their mean and standard deviation.

```
meanHP = mean(mtcars$hp)
indicesL = mtcars$hp>meanHP
indicesS = mtcars$hp<meanHP
largerHP = mtcars$hp[indicesL]
```

```
smallerHP = mtcars$hp[indicesS]

mean(largerHP)
```

```
## [1] 206.9333
```

```
sd(largerHP)
```

```
## [1] 49.91059
```

```
mean(smallerHP)
```

```
## [1] 93.52941
```

```
sd(smallerHP)
```

```
## [1] 22.87225
```

Problem 5

In mtcars dataset, find cars with cyl equal to 8 or 6. Extract those rows from mtcars to create a new dataframe. Find cars with cyl equal to 4. Extract those rows from mtcars to create another dataframe. Compare the hp in these two newly created dataframe. Based on the comparison, make your conclusion (Try to explain the results).

```
rm(mtcars) #resetting cause i messed it up somewhere
```

```
## Warning in rm(mtcars): object 'mtcars' not found
```

```
cyls8 = mtcars$cyl==8
cyls6 = mtcars$cyl==6
cyls86 = cyls8+cyls6
indices86 = 1:32
indices86clean = indices86[as.logical(cyls86)]
rows86 = mtcars[indices86clean,]
mean(rows86$hp)
```

```
## [1] 180.2381
```

```
cyls4 = mtcars$cyl==4
indices4 = 1:32
indices4clean = indices4[as.logical(cyls4)]
rows4 = mtcars[indices4clean,]
mean(rows4$hp)
```

```
## [1] 82.63636
```

The mean of 6/8 was 180.2381 while the mean HP of cars with four cylinders was 82.63636. My conclusion is that the higher the number of cylinders, the higher the horsepower as a general trend.