MSDS 6306: Introduction to Data Science - Live session Unit 01 Assignment

### 1)

1. log(10) = 2.3025851
2. ?log(): description says that log() default is the natural log. Using base 10: log(10, base = 10) = 1
3. Log(-10) = NaN Cannot take the log of a negative number. The log curve is asymptotic to negative infinity as x approaches zero(log(0) = -\infty{}
4. sqrt(4) = 2

### 2)

rNormVector <- rnorm(15)  
mean(rNormVector)

## [1] 0.09415648

sd(rNormVector)

## [1] 1.346474

rNormVector <- rnorm(15,mean = 10,sd = 2)  
mean(rNormVector)

## [1] 9.275331

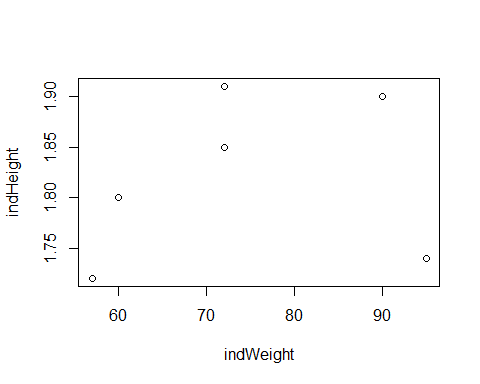
sd(rNormVector)

## [1] 1.343167

1. The values are randomly sampled fromt he distribution so the exact values will not match. As the sample size increases the mean will get closer to the actual mean.

### 3)

indWeight <- c(60, 72, 57, 90, 95, 72)  
indHeight <- c(1.80, 1.85, 1.72, 1.90, 1.74, 1.91)   
plot(indWeight, indHeight) # Weight and height would seem to be positivly related

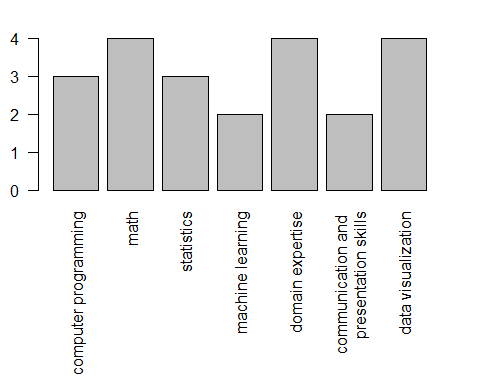


indBMI <- indWeight / sqrt(indHeight)  
sum(mean(indWeight) - indWeight)

## [1] -2.842171e-14

### 4)

brychan <- data.frame(  
 area = c("computer programming", "math", "statistics", "machine learning", "domain expertise",   
 "communication and \npresentation skills", "data visualization"),  
 rank = c(3, 4, 3, 2, 4, 2, 4)  
 )  
par(mar = c(10,2,2,2))  
barplot(brychan$rank, names.arg = brychan$area,las=2)



### 5)