Assignment 2
Ju Ho Yoon
07/09/15

1) Scores: 90, 95, 89, 71, 73, 96, 87, 95, 107, 89, 96, 80, 97, 95, 102, 97, 93, 101, 82, 83, 74, 91, 83, 98, 95, 111, 99, 120, 93, 84

   A) Sample mean score: 92.2. Sample mean score estimates the population mean.
   B) Because we are sampling, there is a variation to the next. And therefore, it will not be exactly equal to the population parameter. It may get close to the population parameter, but not exactly.
   C) Standard error: 1.976529.
   D) The standard error measures the variability of a population parameter: the spread of the sampling distribution of the sample mean.
   E) 95% confidence interval: (88.15754, 96.24246)
   F) The confidence interval is the range of values that you can be 95% certain contains the true mean of the population.

2) Null hypothesis: Male and females have the same mean cholesterol concentrations.
   Male: 220.1, 218.6, 229.6, 228.8, 222.0, 224.1, 226.5
   Female: 223.4, 221.5, 230.2, 224.3, 223.8, 230.8

   maleMean: 224.2429
   femaleMean: 225.6667
   sd(male): 1.608143
   sd(female): 1.578537
   nMale: 7
   nFemale: 6
   maleMedian: 224.1
   femaleMedian: 224.05
   95% CI Male: (220.3079, 228.1778)
   95% CI Female: (221.6089, 229.7244)
   Two Sample t-test
   p-value = 0.5405

   *Because the p-value is 0.5405, which is greater than 0.05, we cannot reject the null hypothesis that the male and female have the same mean cholesterol concentrations.

3) A) False, because p-value doesn't tell anything about size of the effect.
   B) True.
   C) False, because type I error is determined by alpha, .05, beforehand.
   D) False, because type II error is determined by the effect size.
   E) True.

4) Person A: 248, 236, 269, 254, 249, 251, 260, 245, 239, 255
Person B: 380, 391, 377, 392, 398, 374

```
personA <- c(248, 236, 269, 254, 249, 251, 260, 245, 239, 255)
personB <- c(380, 391, 377, 392, 398, 374)
personA2 <- 1.5*personA

layout(matrix(2:1, ncol =1))
hist(personA2, xlim=c(350,410))
hist(personB, xlim=c(350, 410))
```

data: personB and personA2
W = 44, **p-value = 0.1471**
alternative hypothesis: true location shift is not equal to 0

*Because the p-value is 0.1471, we cannot reject the null hypothesis.

5) Standard deviation quantifies how much the values vary from one another, while the standard error of a mean quantifies how precisely you know the true mean of the population: it takes into account both the standard deviation and the sample size. Also, the standard error of the mean is always smaller than the standard deviation. Lastly, standard error of a mean gets smaller as sample size gets larger; yet, standard deviation does not predictably change as more data is acquired.