Assignment #2

Sarah Kate Shore

1/20/2016

```
data(iris)
```

Question 1: Object "output" gives the mean value for each column of data according to each iris species. For example, the mean sepal length of setosa is "5.006"

Question 2: Output is defined as a matrix with number of rows

```
#Loop through each row of iris record values using iterator k, add up the record values of the column wi
 #Loop through each column, defines variables x and y
    #Loop through each subset, subset defined by species
```

Question 3: Output could be renamed "species_mean", because it is the mean of the values for each species per column. "x" is the sum of each subset row per column, so it should be renamed "sum_row_values". "y" is the sum of the rows in each subset, so should be renamed "sum_row_number".

Question 4: Got rid of "k" level loop by using the function "mean." The mean function uses each row of each subset for its argument. Still retained the loop that creates the subset and the loop that goes through each column.

```
sp_ids = unique(iris$Species)

species_mean = matrix(0, nrow=length(sp_ids), ncol=ncol(iris)-1)
rownames(species_mean) = sp_ids
colnames(species_mean) = names(iris[ , -ncol(iris)])

for(i in seq_along(sp_ids)){
iris_sp = subset(iris, subset=Species == sp_ids[i], select=-Species)
    for(j in 1:(ncol(iris_sp))) {
    species_mean[i,j]=mean(iris_sp[,j])
    }
 }
(species_mean)
```

```
##            Sepal.Length Sepal.Width Petal.Length Petal.Width
## setosa            5.006       3.428        1.462       0.246
## versicolor        5.936       2.770        4.260       1.326
## virginica         6.588       2.974        5.552       2.026
```

Question 5: The vector y is the sum of vector x up to that value of x. For the loop, I defined the x and y vector dimensions, then did one loop where a given value of y is the sum of x from 1 to that value.

```
y = vector("double", 10)
x = c(1:10)
for(i in 1:10 ) {
  y[i] = sum(x[1:i])
   }
y
```

```
##  [1]  1  3  6 10 15 21 28 36 45 55
```

Question 6: Used an "if" statement to produce the response "NA" for y values greater than 10.

```
y = vector("double", 10)
x = c(1:10)
for(i in 1:10) {
  y[i] = sum(x[1:i])
  if(y[i] > 10)
    y[i] = NA
}
y
```

```
##  [1]  1  3  6 10 NA NA NA NA NA NA
```

Question 7: Used "x" instead of "1:10" in for loop so that any vector x could work.

```
y = vector("double", 10)
x = c(1:10)
for(i in x) {
  y[i] = sum(x[1:i])
  if(y[i] > 10)
  y[i] = NA
  }
y
```

```
##  [1]  1  3  6 10 NA NA NA NA NA NA
```