

Assignment_2(Intermediate)

Ashley Jensen

1/21/2019

Provided Information:

```
data(iris)
head(iris)
```

```
## Sepal.Length Sepal.Width Petal.Length Petal.Width Species
## 1          5.1          3.5          1.4          0.2 setosa
## 2          4.9          3.0          1.4          0.2 setosa
## 3          4.7          3.2          1.3          0.2 setosa
## 4          4.6          3.1          1.5          0.2 setosa
## 5          5.0          3.6          1.4          0.2 setosa
## 6          5.4          3.9          1.7          0.4 setosa
```

```
sp_ids = unique(iris$Species)
```

```
output = matrix(0, nrow=length(sp_ids), ncol=ncol(iris)-1)
rownames(output) = sp_ids
colnames(output) = 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))) {
    x = 0
    y = 0
    if (nrow(iris_sp) > 0) {
      for(k in 1:nrow(iris_sp)) {
        x = x + iris_sp[k, j]
        y = y + 1
      }
      output[i, j] = x / y
    }
  }
}
output
```

```
## 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
```

Exercises:

Iris Loops

1. Describe the values stored in the object “output”.

The values stored in the object ‘output’ are the average data values of each species per column. For example, 5.006 is the average sepal length for the setosa species.

2. Describe using pseudo-code how “output” was calculated.

Loop from 1 to the length of the unique species #Create a subset from Iris data (iris_sp) #Loop from 1 to the number of columns of Iris data (iris_sp) #Assign zero as a starting point for x and y #If number of rows of iris_sp is greater than zero, #then loop from 1 to number of rows #Assign x= sum of the values of each row per column #Assign y= y + 1 (this is the total count/iterations run in the code) #Assign Output as a matrix of the average values

3. The variables in the loop were named to be vague. How would you rename them to clarify what is occurring in the loop?

Output = avg_iris_characteristic, x = total_value, y = total_counts

4. Please suggest another way to calculate output by decreasing the number of loops by one.

SUM OF A SEQUENCE

5. You have a vector x with numbers 1:10. Write a for loop that will produce the vector y that contains the sum of x up to that index.

```
x<-c(1:10)
y<-NULL
for(i in x) {
  y[i]=sum(x[1:i])
}
y
```

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

6. Modify your for loop so that if the sum is greater than 10 the value of y is set to NA

```
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"
```

7. Place your for loop into a function that accepts as its argument any vector of arbitrary length and it will return y.

```
x<-c(1:10)
vector_addition<-function (m) {
  y<-NULL
  for(i in m) {
    y[i]=sum(m[1:i])
    if(y[i]>10) {
      y[i]='NA'
    }
  }
  return(y)
}
vector_addition(x)
```

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