EPP 8273 Empirical Research in Theory and Practice: Sabet Activity 1

R Activity 1

Afsoon Sabet

1. Question 1

```
iris_df_csv <-read.table(file="assignments/Activity_2/Iris_data_text_file.txt",
header=TRUE, sep="\t", na.strings = ".")
```

2. Question 2

```
summary(iris_df_csv)
```

```
##   Sepal.Length     Sepal.Width      Petal.Length     Petal.Width
##  Min.   :4.300   Min.   :2.000   Min.   :1.000   Min.   :0.100
##  1st Qu.:5.175   1st Qu.:2.800   1st Qu.:1.675   1st Qu.:0.400
##  Median :5.900   Median :3.000   Median :4.500   Median :1.400
##  Mean   :5.909   Mean   :3.037   Mean   :3.906   Mean   :1.282
##  3rd Qu.:6.500   3rd Qu.:3.300   3rd Qu.:5.100   3rd Qu.:1.900
##  Max.   :7.700   Max.   :4.000   Max.   :6.900   Max.   :2.500
##  NA's   :2       NA's   :4       NA's   :2       NA's   :2
##    Species
##  Length:118
##  Class :character
##  Mode  :character
##
##
##
##
```

```
nrow(iris_df_csv)
```

```
## [1] 118
```

There are 118 rows in this dataframe. I used both summary and nrow to find this number.

3. Question 3

```
summary(iris_df_csv)
```

```
##    Sepal.Length    Sepal.Width     Petal.Length     Petal.Width
##  Min.   :4.300   Min.   :2.000   Min.   :1.000   Min.   :0.100
##  1st Qu.:5.175   1st Qu.:2.800   1st Qu.:1.675   1st Qu.:0.400
##  Median :5.900   Median :3.000   Median :4.500   Median :1.400
##  Mean   :5.909   Mean   :3.037   Mean   :3.906   Mean   :1.282
##  3rd Qu.:6.500   3rd Qu.:3.300   3rd Qu.:5.100   3rd Qu.:1.900
##  Max.   :7.700   Max.   :4.000   Max.   :6.900   Max.   :2.500
##  NA's   :2       NA's   :4       NA's   :2       NA's   :2
##    Species
##  Length:118
##  Class :character
##  Mode  :character
##
##
##
##
```

```
sum(is.na(iris_df_csv$Species))
```

```
## [1] 0
```

Using summary(), I can see that there are 2 NAs in the Sepal.Length column, 4 in Sepal.Width, 2 in Petal.Length, and 2 in Petal.Width. I confirmed that there were 0 NAs in Species using the is.na() function.

4. Question 4

```
print(head(iris_df_csv, n=10))
```

```
##    Sepal.Length Sepal.Width Petal.Length Petal.Width Species
## 1           5.1          NA           NA         0.2  setosa
## 2           4.9         3.0          1.4         0.2  setosa
## 3           4.7         3.2          1.3         0.2  setosa
## 4            NA         3.1          1.5         0.2  setosa
## 5           5.0         3.6          1.4         0.2  setosa
## 6           5.4         3.9           NA         0.4  setosa
## 7           4.6         3.4          1.4         0.3  setosa
## 8           5.0         3.4          1.5         0.2  setosa
## 9            NA         2.9          1.4         0.2  setosa
## 10          4.9          NA          1.5         0.1  setosa
```

5. Question 5

```
print(tail(iris_df_csv, n=3))
```

```
##     Sepal.Length Sepal.Width Petal.Length Petal.Width   Species
## 116          6.5         3.0          5.2         2.0 virginica
## 117          6.2         3.4          5.4         2.3 virginica
## 118          5.9         3.0          5.1         1.8 virginica
```