Module 3 In Class Activity

Author: Andres Felipe Alba Hernández **Department:** Electrical Engineering **Date:** September 13, 2018
Course: ISYE670 Data Science for Engineers **Professor Northern Illinois University**

PART A

1) Load the islands dataset and obtain the total number of observations.

```
rm(list=ls())
data("islands") #load the dataset islands
#help("islands")
summary(islands)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    12.0    20.5    41.0  1253.0   183.2 16990.0
```

Description: The areas in thousands of square miles of the landmasses which exceed 10,000 square miles.
Usage: islands **Answer: The total number of observation is:**

2) Calculation of mean and media:

```
mean(islands)
```

```
## [1] 1252.729
```

```
median(islands)
```

```
## [1] 41
```

3) Using range calculate max and min value size of the islands: it will be display as min max in the output of the following command.

```
range(islands)
```

```
## [1]    12 16988
```

4) Standard desviation and Range.

```
sd(islands) #standard desviation
```

```
## [1] 3371.146
```

```
range(islands) #range output = (min,max)
```

```
## [1]    12 16988
```

5) Quantile Function:

a) Find the quantiles for: 0%, 25%, 50%, 75%, 100%

```
quantile(islands)
```

```
##       0%      25%      50%      75%     100%
##    12.00    20.50    41.00   183.25 16988.00
```

b) Find the quantiles for: .05%, 95%

```
quantile(islands,probs = c(0.005,0.95))
```

```
##     0.5%      95%
##   12.235 8481.750
```

c) What does the parameter na.rm do?

```
data_test <- c(0.5,10,NaN)
quantile(data_test,na.rm = TRUE)
```

```
##     0%    25%    50%    75%   100%
##  0.500  2.875  5.250  7.625 10.000
```

```
try(quantile(data_test))
try(quantile(data_test,na.rm = FALSE))
```

If the na.rm flag is set as FALSE the NaN values are not removed before the computation, therefore they are not allow to be in the data set. As may be observed above.

6)