Intro to Data Science - HW 1

Copyright 2022, Jeffrey Stanton, Jeffrey Saltz, and Jasmina Tacheva

```
# Enter your name here: Bhavya Shah
```

Attribution statement: (choose only one and delete the rest)

```
# 1. I did this homework by myself, with help from the book and the professor.
```

Define a variable:

```
value <- 150
```

Print the content of this new variable:

```
value
```

```
## [1] 150
```

Define the following vectors, which represent the **hospital capacity** (measured in number of hospital beds) and **number of colleges** in each of the five counties in Central New York (CNY) **Cayuga**, **Cortland**, **Madison**, **Onondaga**, and **Oswego**, in this order (e.g. the first number in the **capacity** vector suggests there are 124 hospital beds in Cayuga county):

```
capacity <- c(124, 70, 51, 1403, 86)
colleges <- c(2, 2, 3, 9, 2)
```

Part 1: Calculating statistics using R

A. Show the number of observations in the **capacity** vector with the length() function:

```
length(capacity)
```

```
## [1] 5
```

B. Show the number of observations in the **colleges** vector with the length() function:

```
length(colleges)
```

```
## [1] 5
```

C. Calculate the average number of colleges in CNY using the mean() function:

```
mean(colleges)
```

```
## [1] 3.6
```

D. Calculate the total number of colleges in CNY using the sum() function:

```
sum(colleges)
```

```
## [1] 18
```

E. Calculate the average CNY hospital capacity using the mean() function:

```
mean(capacity)
```

```
## [1] 346.8
```

F. Calculate the total hospital capacity in CNY using the sum() function:

```
sum(capacity)
```

```
## [1] 1734
```

G. Calculate the average number of colleges in CNY again, this time using **the results from steps B & D**:

```
sum(colleges) / length(colleges)
```

```
## [1] 3.6
```

H. Calculate the average CNY hospital capacity again, this time using **the results from steps A & F**:

```
sum(capacity) / length(capacity)
```

```
## [1] 346.8
```

Part 2: Using the max/min and range functions in R

I. How many colleges does the county with most colleges have? Hint: Use the max() function:

```
max(colleges)
```

```
## [1] 9
```

J. What is the hospital capacity of the county with the smallest number of hospital beds in CNY? **Hint:** Use the min() function:

```
min(capacity)
```

```
## [1] 51
```

K. Display the number of beds in the county with the highest hospital capacity and the county with the lowest hospital capacity together, using a single command. **Hint:** Use the range() function:

```
range(capacity)
```

```
## [1]   51 1403
```

Part 3: Vector Math

L. Create a new vector called **lowerCap**, which is the current hospital capacity of each county **- 10** (each county has 10 less hospital beds):

```
lowerCap <- capacity -10
lowerCap
```

```
## [1]  114   60   41 1393   76
```

M. Calculate the average of **lowerCap**:

```
mean(lowerCap)
```

```
## [1] 336.8
```

N. Create another vector called **megaCap**, which is the original hospital capacity in each county (stored in **capacity**) plus **value** (the variable we declared in the beginning of this assignment).

```
megaCap <- capacity + value
```

O. Print the content of **megaCap**:

```
megaCap
```

```
## [1]  274  220  201 1553  236
```

P. In a variable called **bigCap**, store all the hospital numbers from the original **capacity** vector which are **greater than 80** (using **subsetting** in R):

```
bigCap <- capacity[capacity > 80]
bigCap
```

```
## [1]  124 1403   86
```

Q. Print the length of **bigCap** and explain in a comment how many counties have more than 80 hospital beds:

```
length(bigCap)
```

```
## [1] 3
```

#as per the above data, 3 counties have more than 80 hospitals beds.