**MS5108 Applied Customer Analytics**

**Assignment 1 (10%)**

**Name**: Jayakarthi Boovendran

**Student ID**: 19230487

**Manipulation of Numbers and Vectors Exercises**

**Question 1**

You keep track of the minutes you spend on your smartphone. The last 5 times you checked, your total minutes were: 361 667 826 923 1,178

* Enter these numbers into a vector using R. Use the function diff() on the data. What does it indicate?

**diff( )**

* An R function for computing the difference of a series
* **Syntax** : diff(vector,lag,differences)
* The function takes in three parameters, the vector, lag and differences. Where, ‘lag’ is the no. of periods to be considered for diff() and the ‘differences’ indicates the function is to be invoked. The ‘lag’ and ‘differences’ take on default value 1, if not explicitly specified.

**R Code**

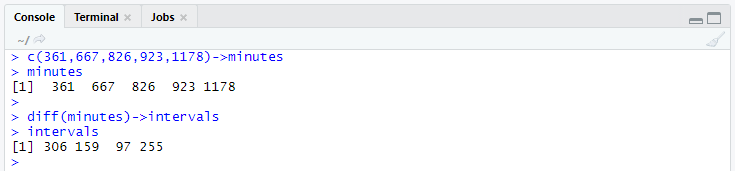
c(361,667,826,923,1178)->minutes

minutes

diff(minutes)->intervals

intervals

**Output**



* Use the max() function to find the maximum number of minutes between any two times you checked, the mean() function to find the average number of minutes between checks, and the min() function to get the minimum number of minutes between checks.

**R Code**

max(intervals)

mean(intervals)

min(intervals)

**Output**



**Question 2**

Suppose you track your commute times for two weeks (10 days) and you find the following times in minutes: 32 45 28 35 38 51 27 28 48 41

* Enter this into R. Use the function max() to find the longest commute time, the function mean() to find the average and the function min() to find the minimum.

**R Code**

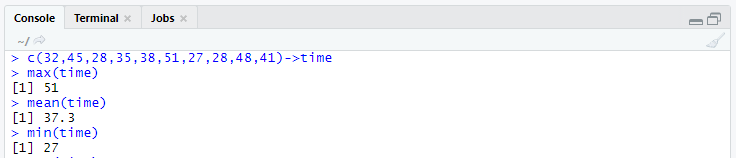
c(32,45,28,35,38,51,27,28,48,41)->time

max(time)

mean(time)

min(time)

**Output**



* What is the variance of this vector of commute times? The standard deviation?

**R Code**

var(time)

sd(time)

**Output**



* Oops, the 28 minute entry was a mistake. It should have been 38 minutes. Fix this mistaken entry, and then find the new average.

**R Code**

replace(time,time==28,38)->time

mean(time)

**Output**



* How many times was your commute 40 minutes or more? What do you get?

**R Code**

sum(time>=40)

**Output**



* What percent of your commutes are less than 35 minutes? How can you answer with R?

**R Code**

sum(time<35)->count

(count/length(time))\*100

**Output**



**Question #3**

Your cell phone bill varies from month to month. Suppose that is one year your phone bill has the following monthly amounts: 46 33 39 37 46 30 48 32 49 35 30 48

* Enter this data into a variable called bill. Use the sum() function to find the amount you spent this year on the cell phone. What is the smallest amount you spent in a month? What is the largest? How many months was the amount greater than €40? What percentage was of total was this?

**R Code**

c(46,33,39,37,46,30,48,32,49,35,30,48)->bills

#Sum of Bills

sum(bills)->TotalSum

TotalSum

#Minimum Bill

min(bills)

#Maximum Bill

max(bills)

#Bills > $40

sum(bills>40)->count

count

#Percentage of Bills > $40

(count/length(bills))\*100

**Output**

