* **Question 1:** Read your WIFI csv file

Build a function using R.

* + The purpose of function is to calculate the difference of two Access Points in same frequency.
  + The signature of the function should be (AP1, AP2, frequency);
  + The output is an array with difference of signal strength.
  + The basic algorithm of the function
    1. Retrieve two dataset of two APs with the input frequency
    2. Pair those two datasets based on time and fingerprint;
    3. filter out the NA data;
    4. Calculate and output the difference array / column.

**RCode:**

**setwd('C:/Users/Neha/Desktop/RData')**

**data = read.csv('wifi.csv')**

**attach(data)**

**z335\_CISACL = data[AP\_Name=='CISACL',]**

**z335\_CISATV = data[AP\_Name=='CISATV',]**

**"calculateDiff" <- function(AP1, AP2, Freq)**

**{**

**var1 = subset(AP1, AP1[,4]==Freq)**

**var2 = subset(AP2, AP2[,4]==Freq)**

**merging <- merge(var1,var2, by=c("TimeStamp", "Fingerprint\_ID"))**

**differences = merging[,7] - merging[,17]**

**return(differences)**

**}**

**calculateDiff(z335\_CISACL,z335\_CISATV, 2412)**

**Question 2:**

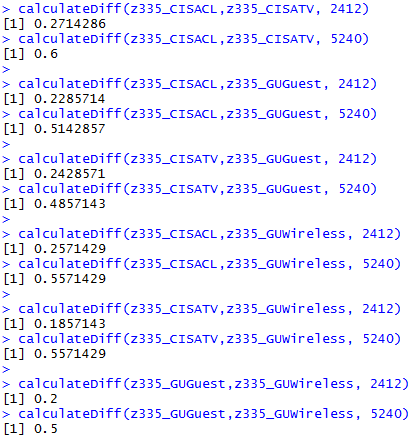
* Test your function with Access Points: CISACL and CISATV with frequency 2412 or 5240
* Statistical test
  + H0: CISACL and CISATV output same signal strength with same frequency and fingerprint.
  + HA: CISACL and CISATV output different signal strengths with same frequency and fingerprint.
  + Put your conclusion in your report (word file)

Answer : The P-Value is small. It is 2e^-16. The evidence against the Null-hypothesis is very strong. It’s rejected.

**Question 3:** Calculate the **distance** table of different APs with different frequencies.

The definition of **distance** =  ***sum(ABS(difference array)) /length(difference array)***

The table should be similar as follow

**RCode:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| vs | **CISACL** | **CISATV** | **GUGuest** | **GUWireless** |
| **CISACL** |  | f2412 = 0.2714  f5240 = 0.6 | f2412 = 0.2285  f5240 = 0.51428 | f2412 = 0.25714  f5240 = 0.55714 |
| **CISATV** |  |  | f2412 = 0.242857  f5240 = 0.48571 | f2412 = 0.18571  f5240 = 0.55714 |
| **GUGuest** |  |  |  | f2412 = 0.2  f5240 = 0.5 |
| **GUWireless** |  |  |  |  |

**Challenge question:**

Sometimes, using time and fingerprint as primary key will is not unique. For example,

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| CISACL Frequency | CISACL Signal Strength | CISACL  Time | CISACL Fingerprint |  | CISATV Frequency | CISATV Signal Strength | CISATV  Time | CISATV  Fingerprint |
| 2437 | -75 | 9/17/2014 11:59 | 101504391 |  | 2437 | -75 | 9/17/2014 11:59 | 101504391 |
|  |  |  |  |  | 2437 | -82 | 9/17/2014 11:59 | 101504391 |

How can you solve the problem? Can you integrate your solution into your function?

We can merge two datasets by creating the dataset called x which merges the CISACL’s data depends on the frequency 2437 as below:-

**x = subset(z335\_CISACL, z335\_CISACL[,4]==2437)**

Also, the same for Access point CISATV

**y = subset(z335\_CISATV, z335\_CISATV[,4]==2437)**

Merging of timestamp and the finger point as follows:-

**g <- merge(x,y, by=c("TimeStamp", "Fingerprint\_ID"))**

Here, we can build the function to make the required join between these datasets

**joinFun <- function(AP1, AP2, Freq)**

**{**

**x = subset(AP1, AP1[,4]==Freq)**

**y = subset(AP2, AP2[,4]==Freq)**

**g <- merge(x,y, by=c("TimeStamp", "Fingerprint\_ID"))**

**z=g$AP\_SignalStrength.x-g$AP\_SignalStrength.y**

**return(z)**

**}**

**t=wf(z335\_CISACL,z335\_CISATV, 2437)**

**t**