

# Assignment #01

## **I. Introduction:**

My understanding of this problem is that, according to the required data the area where facility(A) lies need more staff and resources to give health services to the required patients.

## **II. Assumptions:**

My assumption is if the couple of peoples from Group 'B' could to transfer to Group 'A', or the peoples near to the Group 'A' could arrange some staff for the peoples near Group 'A' the problem for curing the people there would be solved.

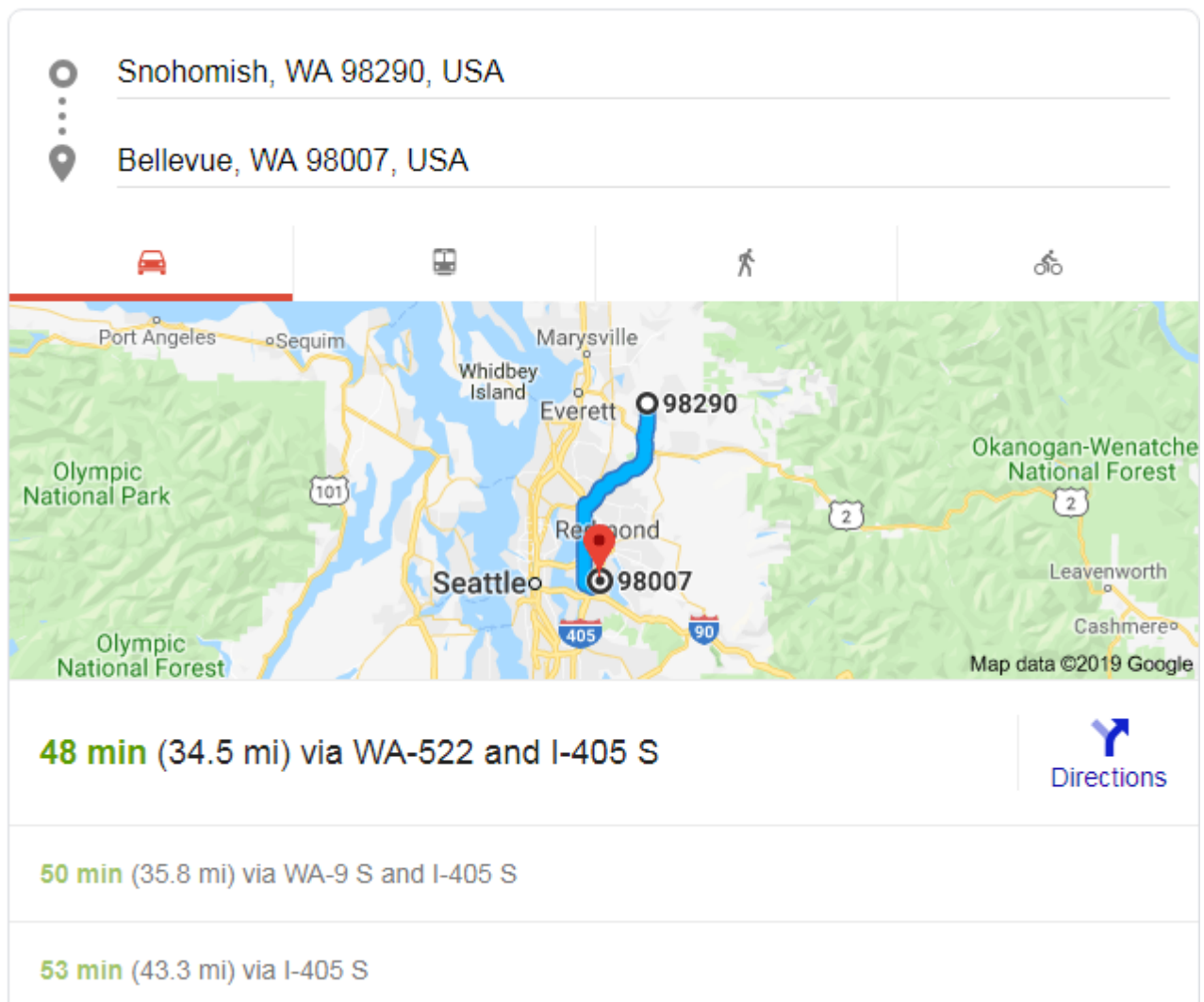
## **III. Data gathering, handling, cleaning, processing:**

Looking at the given data, I've analyzed the data and gather some information like first getting Zip codes for all the US dept. and getting their locations.

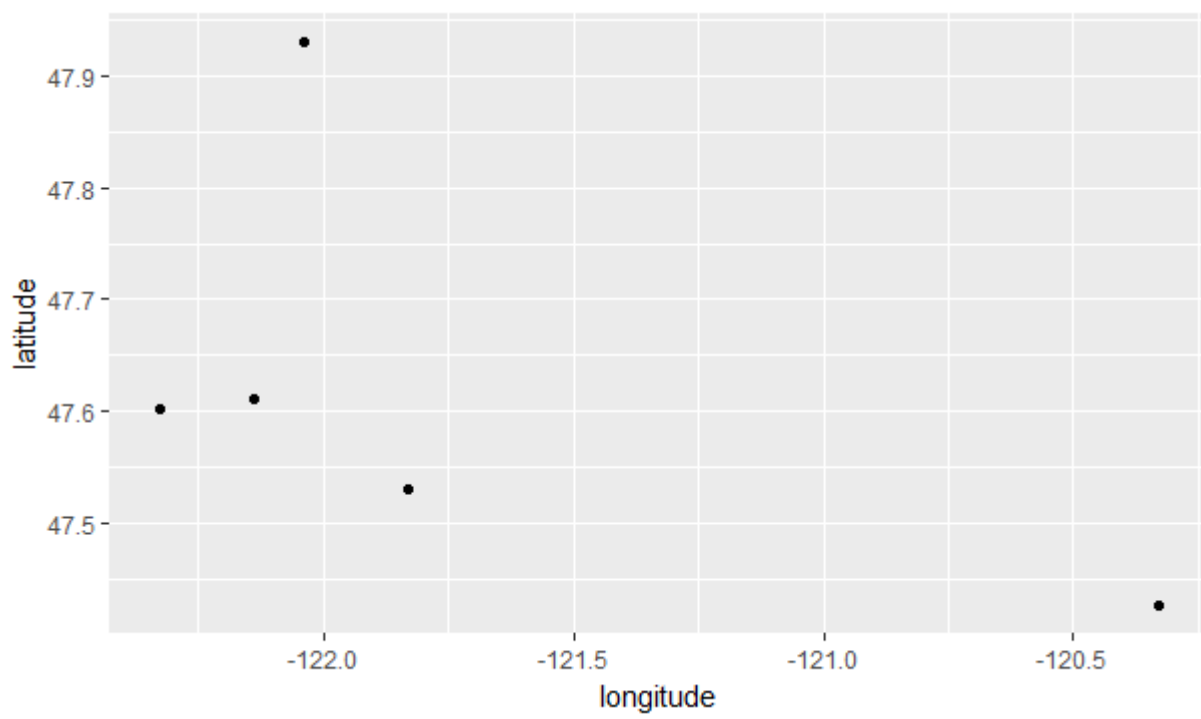
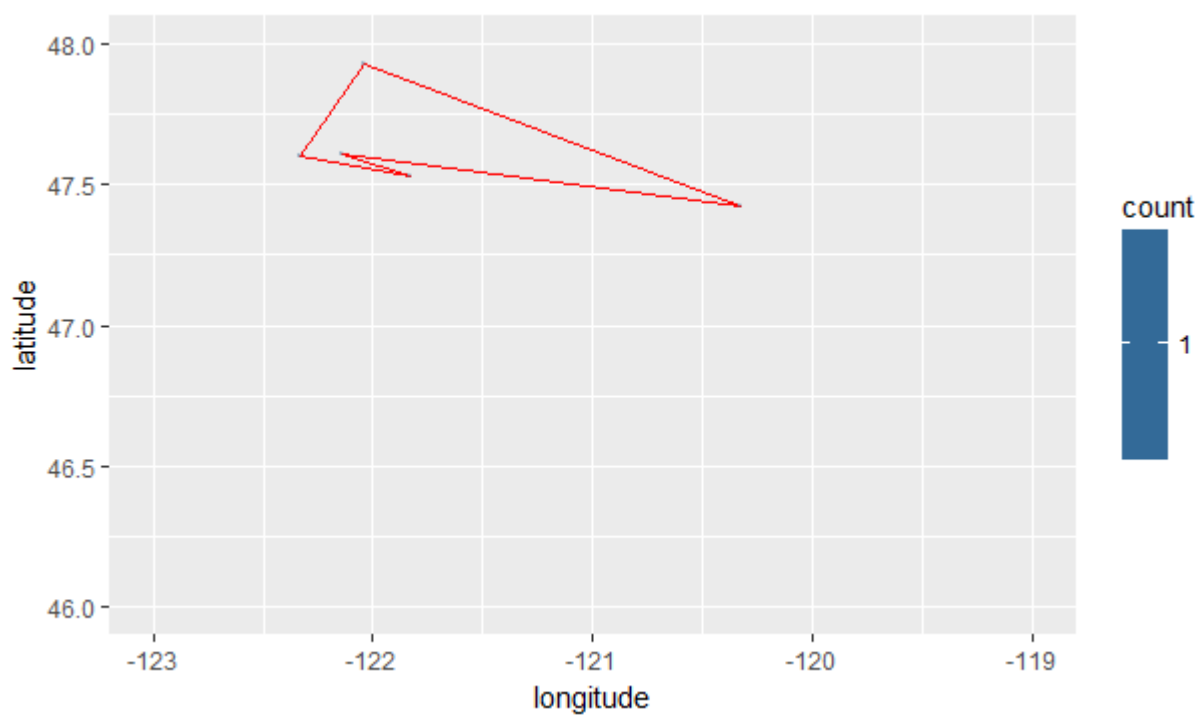
	zip	postal	city	state	latitude	longitude
1	98007	98007	Bellevue	WA	47.61093	-122.1410
2	98065	98065	Snoqualmie	WA	47.53064	-121.8305
3	98104	98104	Seattle	WA	47.60252	-122.3286
4	98290	98290	Snohomish	WA	47.93090	-122.0398
5	98801	98801	Wenatchee	WA	47.42550	-120.3288

Cleaned the data and got this figure. Also done some research from where the distance from Group 'A' to Group 'B' lies.

Fig:



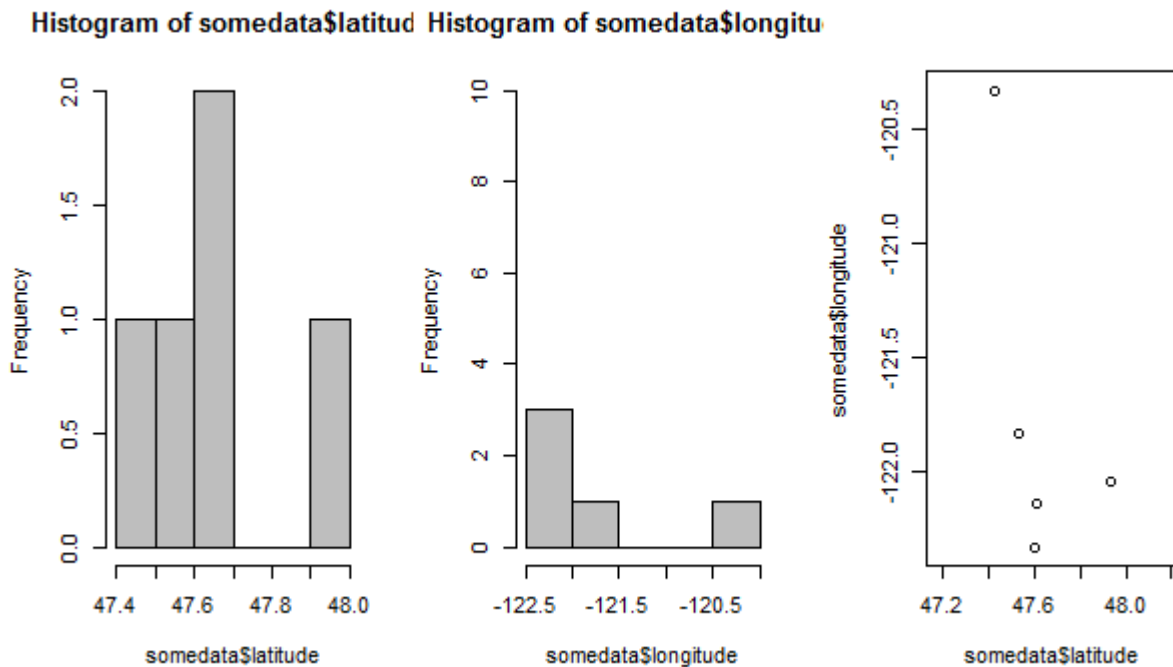
After this, I've done some processing for area connections by plotting.  
Fig:



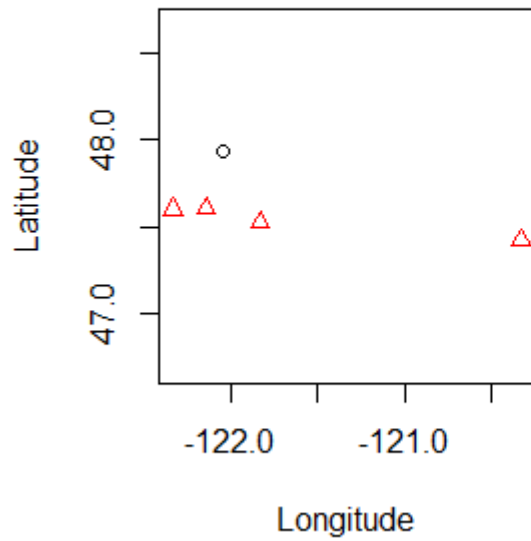
#### IV. Proposed Solution:

I've proposed is this the peoples where the population is less, they should get more staff for the health services. For which I've use 'Clustering'. Which will calculate the distance between the groups and make our process easier.

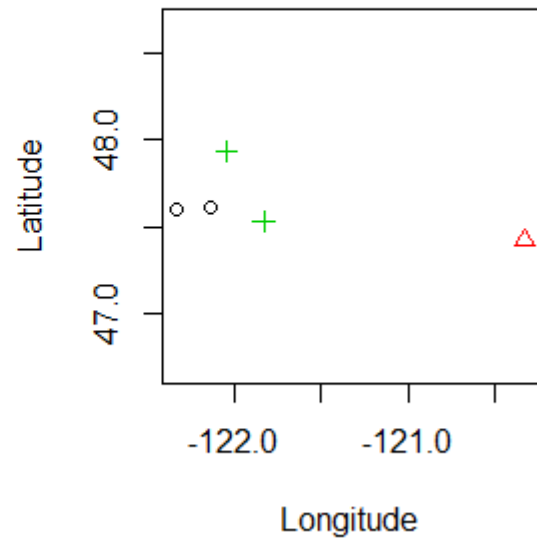
Fig:



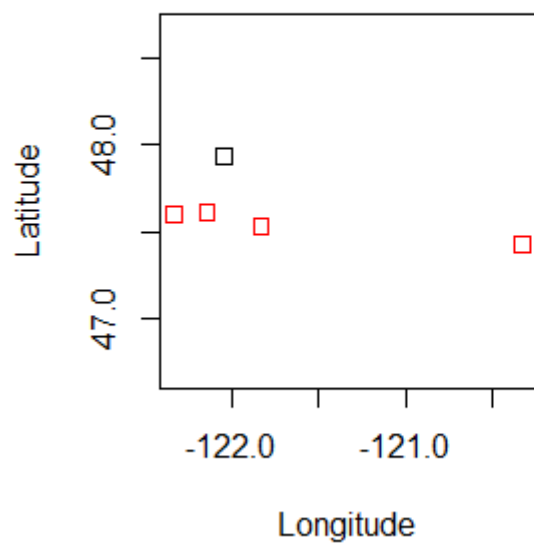
**Sites for two kiosks**



**Sites for three kiosks**



**Two-cluster kiosks in three-clust**



## V. References:

I've use many libraries while making these analysis possible. And ,I want to donate that, I wanted to give the graphical display on analyzing data but, I don't have API so couldn't use "get\_map" function.

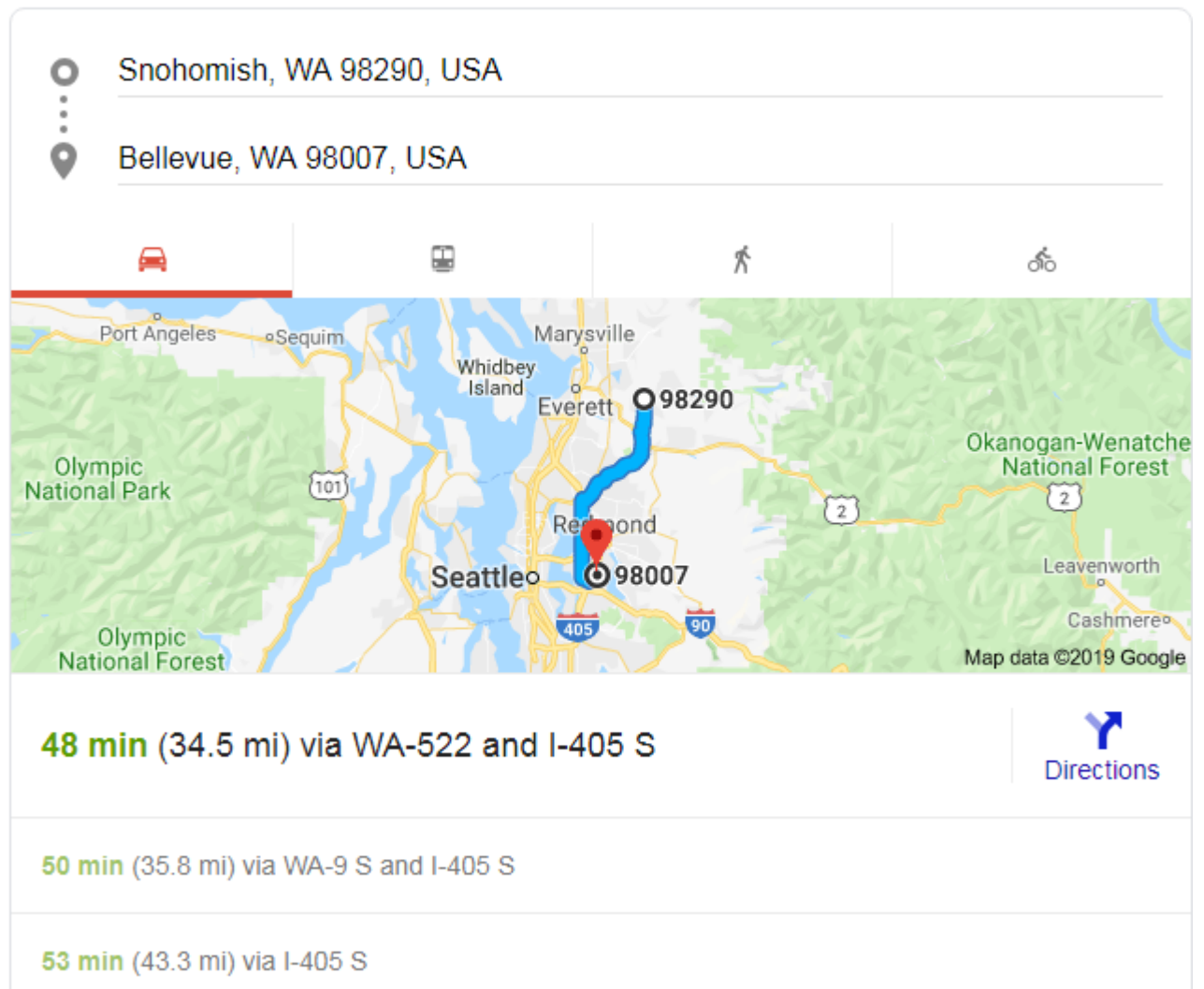
Number of libraries use are in the "service\_delv.R" file.

## VI. Average Time Taking:

Didn't have access to API it isn't free so just use google to map the distance from the required problem.

If I've had access at API ,I would be excited to use "get\_map", "ggmap", and "geosphere".

In my opinion get\_map is easy, you could call the particular location and map it.



I was able to use shape as to point the groups using Latitude and Longitude. But for mapping, I need access to API ☹.

Heres the shape,I got from “cecus.org”

Fig:



## **VII. Population Density:**

I would consider using “Kaggle” where, I get most of my data for practicing on my Data Science skills.

The answer from where, I would extract the information to build-up my analysis, “Google” , “Cencus”, “Kaggle” and also using “Zipcode” libraries to locate the latitude and longitude of the given zipcodes.

## **VIII. Optimization:**

With my experience at this current period, I think clustering, k-mean is an effective way to solve this problem. Because clustering connects with the nearest cluster and points distance using API and using this technique we can optimize our problem for selecting branches for our nearest healthcare hospital.