Optimizing a Healthcare Network for Improved Service Delivery

1.1 Introduction:

The Washington State Health Ministry is planning to upgrade staff or resources in existing facilities across Washington so that they can allocate resources to where they are most needed, based on demand for services in different geographic regions. To address this problem, the health ministry needs to:

• Make services more accessible by allocating resources to where they are most needed

• Minimize people's travel time so that they only travel to their nearby facility.

1.2 Problem Statement:

The Washington State Health Ministry want to allocate resources in the manner that everyone can use healthcare facility when needed. For this time, let's suppose that healthcare facility staff in some areas is much less than its population and not enough to facilitate everyone. And in some areas the staff is more than people needs.

For example, Majority of the population living in Area A need access to health services, but would need to travel on average 2 hours per return trip to get their needs met in Facility B, instead of travelling 0.5 hours to their nearby Facility A. This is because Facility A has not enough staff to fulfill needs of every patient.

1.3 Assumptions:

To solve the above stated problem, I have made some assumptions that are given below:

- Sum the population of all areas to get total population and then find percentage of the population of each area, also get total staff by taking sum of all staff and then allocate staff according to the percentage of the population
  For example if there are 3 areas A, B & C. The population of Area A, B, C are 10, 20 &30 simultaneously and Health care facility staff allocated to facilities A, B & C are 4, 5 & 2 simultaneously.
  To solve this 1st we calculate total population which is 10+20+30 = 60 and then find population percentage of each area which is 16.67% for A, 33.33% for B and 50% for C. Now we calculate total staff which is 4+5+2 = 11. And now allocate staff to every region according to its population percentage.

  Staff allocated to area 'A': $\frac{11\times16.67}{100}$ = 1.8 = 2
  Staff allocated to area 'B': $\frac{11\times33.33}{100}$ = 3.7 = 4

Staff allocated to area 'C': $\frac{11\times50}{100}$ = 5.5 = 5

- The 2nd assumption is according to a rule of thumb is that the healthcare worker to patient ratio should be approximately 1:2808 in order to not overwork staff members. So 1 healthcare worker can reasonably serve 2808 standard patient needs a year. We assume at least more than half would seek services for a variety of problems.
  So I thought to according to the healthcare worker to patient ratio. If we assume half of the population use healthcare facility, so half of 2808 is 1404 and healthcare worker to patient ratio becomes 1:1404.
  Now the solution become simple that first we calculate total staff & then allocate it according to healthcare worker to patient ratio. And in the end, if the some staff members are left to allocate then allocate them according to percentage of area population
  For example if there are 3 areas A, B & C. The population of Area A, B, C are 10, 20 &30 simultaneously and Health care facility staff allocated to facilities A, B & C are 4, 5 & 2 simultaneously & healthcare worker to patient ratio should be 1:18.
  To solve this, 1st we sum staff of all area which is 4+5+2 = 11. Now allocate staff according to healthcare worker to patient ratio. So we have,

  ### Staff allocated to each area in terms of ratio:
  Staff allocated to area 'A': 1 (healthcare worker to patient ratio: 1:10 < 1:18)
  Staff allocated to area 'B': 2 (healthcare worker to patient ratio: 1:10 < 1:18)
  Staff allocated to area 'C': 2 (healthcare worker to patient ratio: 1:15 < 1:18)

  After allocating staff according to healthcare worker to patient ratio 6 staff members are left out of 11. So we allocate remaining 6 staff members according to the percentage of area population.

  ### Staff allocated to each area in terms of percentage of area population:
  Staff allocated to area 'A': $\frac{6\times16.67}{100}$ = 1.0 = 1
  Staff allocated to area 'B': $\frac{6\times33.33}{100}$ = 1.9 = 2

  Staff allocated to area 'C': $\frac{6\times50}{100}$ = 3.0 = 3
  And now we calculate total staff allocated to each area.

  Staff allocated to area 'A': 1+1 = 2
  Staff allocated to area 'B': 2+2 = 4
  Staff allocated to area 'C': 2+3 = 5

  It gives same result as 1st assumption. So According to me 1st assumption is more appropriate result than 2nd one because it requires less calculation.

1.4 Data gathering, cleaning & processing:
Population Data is scraped by a website "Washington Demographics by Cubit" that contains current US census data. From the website we extracted the data in which information of the population of each area with its zip code exist.

After gathering the data, we faced task to clean it. The data contained 3 columns.

1- Washington Zip Codes by Population Rank 2- Zip Code 3- Population

We have a task to find population of each area with its zip code. So the 1st column is useless for us & we dropped it. Then we make a table (data frame in pandas) that contain population of each area according to its zip code. The table consist only those zip codes that were given in the document.

1.5 Proposed Solution:

My proposed solution of the given problem is the first assumption that I have made in the assumption section.
Sum the population of all areas to get total population and then find percentage of the population of each area, also get total staff by taking sum of all staff and then allocate staff according to the percentage of the population
For example if there are 3 areas A, B & C. The population of Area A, B, C are 10, 20 &30 simultaneously and Health care facility staff allocated to facilities A, B & C are 4, 5 & 2 simultaneously.
To solve this 1st we calculate total population which is 10+20+30 = 60 and then find population percentage of each area which is 16.67% for A, 33.33% for B and 50% for
C. Now we calculate total staff which is 4+5+2 = 11. And now allocate staff to every region according to its population percentage.

Staff allocated to area 'A': $\frac{11\times16.67}{100}$ = 1.8 = 2

Staff allocated to area 'B': $\frac{11\times33.33}{100}$ = 3.7 = 4

Staff allocated to area 'C': $\frac{11\times50}{100}$ = 5.5 = 5

References:

- Google Maps API
- Bing Maps API
- pandas
- geopy
- folium
- Matplotlib

2.1 Average Travel Time:

For extracting average travel time that each person would need to travel to each facility, we use "Google Maps Distance Matrix API". It is a service that provides travel distance and time taken to reach destination. It returns the recommended route between origin and destination, consists of duration and distance values for each pair.

Coded Example:

```
import requests, json

api_key ='Your_api_key'

source = input()

dest = input()

url ='https://maps.googleapis.com/maps/api/distancematrix/json?'

r = requests.get(url + 'origins = ' + source +
                       '&destinations = ' + dest +
                       '&key = ' + api_key)

x = r.json()

print(x)
```

API to find distance & Travel time other than Google Maps API:

There is a python library ‘geopy’ that finds distance between two locations but it is unable to find travel duration between them. So ‘Google Maps API’ and ‘Bing Maps API’ are best for finding distance and travel time between 2 locations but the problem is they are not free.

2.2 Population Density:

Here is a coded example that demonstrates how I extract the information from population data source.

In [1]:
```
# read table from Washington-Demographic data
import pandas as pd

df = pd.read_html('https://www.washington-demographics.com/zip_codes_by_population',header=0)

population_df = df[0]
population_df.head()
```

Out[1]:

	Washington Zip Codes by Population Rank	Zip Code	Population
0	1	99301	77226.0
1	2	98052	65251.0
2	3	98012	62039.0
3	4	98208	58030.0
4	5	98682	56411.0

We use a website “Washington Demographics by Cubit” that is a freely, publicly available data source contains current US census data. We use pandas (python library for data manipulation) to read & manipulate data.

2.3 Optimization:

The optimization technique that will solve the problem is given below:
Sum the population of all areas to get total population and then find percentage of the population of each area, also get total staff by taking sum of all staff and then allocate staff according to the percentage of the population
For example if there are 3 areas A, B & C. The population of Area A, B, C are 10, 20 &30 simultaneously and Health care facility staff allocated to facilities A, B & C are 4, 5 & 2 simultaneously.
To solve this 1st we calculate total population which is 10+20+30 = 60 and then find population percentage of each area which is 16.67% for A, 33.33% for B and 50% for
C. Now we calculate total staff which is 4+5+2 = 11. And now allocate staff to every region according to its population percentage.

Staff allocated to area 'A': $\frac{11\times16.67}{100}$ = 1.8 = 2

Staff allocated to area 'B': $\frac{11\times33.33}{100}$ = 3.7 = 4

Staff allocated to area 'C': $\frac{11\times50}{100}$ = 5.5 = 5

Why I think that this optimization technique will solve the problem?

I think that this optimization technique will solve the problem because it equally distributes staff according to the population percentage of each area and minimize healthcare worker to patient ratio.

Some features that are needed to solve the given problem:

- Zip Code
- Population
- Facility Staff Count

~~Coded Example:~~

~~You can see the coded example in the following github repository.~~
~~https://github.com/ABDULSABOOR1995/Data-Science-Dojo-Project/blob/master/Optimizing%20a%20Healthcare%20Network%20for_Improved%20Service%20Delivery.ipynb~~