Optimizing a Healthcare Network for Improved Service Delivery

1 INTRODUCTION

The Washington State Health Ministry would like to optimally 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/areas.

To address this problem, the health ministry needs to:

- Make services more accessible by allocating resources to where they are most needed; increasing capacity in some facilities and decreasing capacity in others.
- Minimize people's travel time so that they only travel to their nearby facility.

1.1 SAMPLE PROBLEM

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 only 21 staff/resources as opposed to 52 staff in Facility B. People living in Area A cannot travel this far, so they need to get their needs met by Facility A, otherwise they go without healthcare.

1.2 DESIRED SOLUTION

To model how many staff members should be shifted from Facility B to Facility A, so that the population living in Area A can travel to their nearby Facility A instead of Facility B, improving travel cost/access to healthcare.

Depending on the level of care, 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. Not the entire population living nearby a facility would need healthcare services, but we assume at least more than half would seek services for a variety of problems.

1.3 DATA SOURCES

1.3.1 Facilities

Facility ID	Facility Area-Zipcode	Facility Staff Count
Facility A	98007	21
Facility B	98290	52
Facility C	98065	43
Facility D	98801	9
Facility E	98104	64

1.3.2 Satellite images and Maps

Example sources: Google Maps, Google Earth, OpenStreetMap, USGS Earth Explorer/Landsat, etc.

These sources can be used for estimating resident population counts in the geographic area of the facility, as well as nearby areas to the facility. Why not grab population estimates from the Census with numeric data instead of using images? The population estimates in the Census are not reflective of the real population, making it an unreliable data source. Population estimates of geographic regions need to be obtained another way. Therefore, computer vision using satellite images or maps data extraction is a necessary task.

These sources can also be used to estimate average time travel for people traveling to a facility in their area or nearby their area, and the average time to all other facilities. Roads, types of terrain, elevation/slope all can affect time travel.

2 DELIVERABLE

Write a short technical document addressing the following:

- **Introduction:** What is your understanding of the problem? Can you write the introduction and the problem statement in formal conference paper-like format?
- **Assumptions:** What assumptions would you make to simplify the problem?
- **Data gathering, handling, cleaning, processing:** How would you acquire the data? What steps would you perform to process or clean the data? How would you extract useful data inputs for this task?
- **Proposed Solution:** Share your optimization technique and any necessary details.
- **References:** Cite any libraries, APIs, publications. We believe in standing on the shoulders of giants. Please reuse any existing research papers, source code, libraries but make sure to cite them.
- **Source code:** Share your R or Python source code.

2.1 POPULATION DENSITY

- Describe how you would use the satellite image or maps data sources to extract household/ population estimates of the facility areas and nearby areas.

- Give a short outline or coded example demonstrating the concept of how you would extract this information from the satellite or maps data sources.
- Are there any other freely, publicly available satellite image or maps data sources you would consider using and how would you extract information from these sources?

2.2 AVERAGE TRAVEL TIME

- Describe how you would use the satellite image or maps data sources to extract the average travel time each person would need to travel to each facility.
- How would you extract average travel time per trip, taking into account mode of transport and terrain?
- Give a short outline or coded example demonstrating the concept of how you would extract this information from this data source.
- Are there any other freely, publicly available satellite image data sources you would consider using and how would you extract information from these sources?

2.3 OPTIMIZATION

- Describe the optimization technique you think would be suitable for this problem scenario and why you think it would solve this problem.
- List some of the inputs/features/variables that you think would be needed to solve this problem.
- Give a short outline or coded example demonstrating the concept of how this optimization technique solves this problem.

3 EVALUATION

Your submission will be evaluated based on the following aspects.

Note: Don't worry if you think any of the aspects is missing. We are not looking for a top-notch PhD dissertation. Rather, we are looking for a deliverables that demonstrates genuine interest in data science; motivation and curiosity; good writing and communication skills; coding skills; dealing with ambiguity; and ability to work under tight deadlines.

- **Understanding of the problem:** Data Science Dojo consultants often work on cutting-edge data science, machine learning, optimization, and mathematical modelling problems. Your ability to understand problem is critical to your success at Data Science Dojo.
- **Rigor:** How detailed your research was and how hard you tried matters more than the actual results.
- **Attention to detail:** Make sure that your document is not missing any details. Typos are never fun.
- **Demonstration of skills:** We are not looking for a unicorn. We don't expect you to be outstanding in all the areas of the problem. We do, however, want you to be good at a few of the skills. Your coding skills; data extraction and manipulation; investigate and research data sources and publications; ability to model problems; and ability to communicate complex concepts.