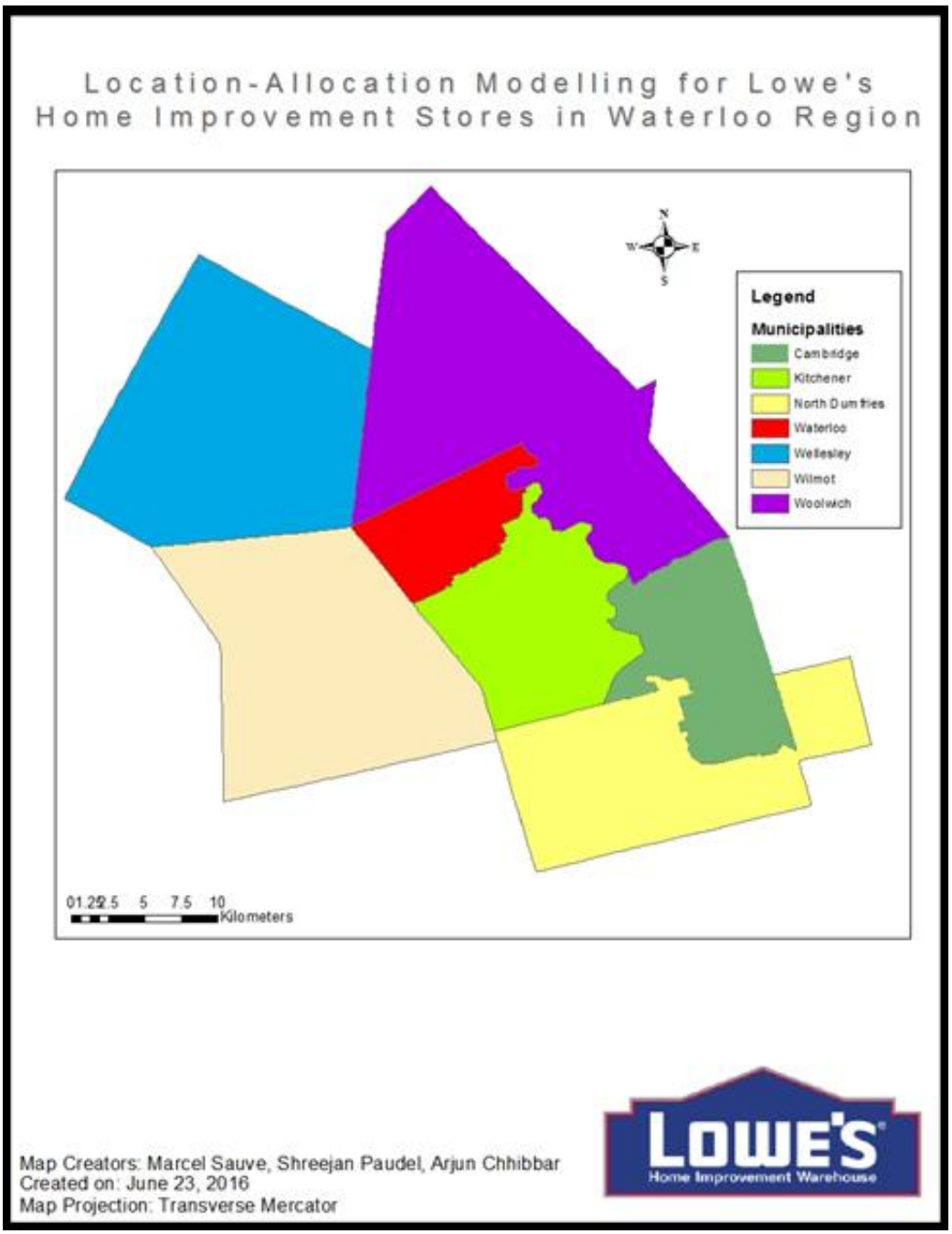
Network Analysis

May 16, 2016



Arjun Chhibbar
Shreejan Paudel
Marcel Sauvé

Table of Contents

- Table of Contents
- Introduction
- Methods
  - Creating a Network Dataset
  - Estimating Demand
- Results
  - Existing Service Areas
  - Candidate Area Locations
  - Single Optimal Site Selection
  - Dual Optimal Site Selection
- Discussion
  - Limitations
  - Long Term Planning
- References

Introduction

Choosing where to build a new retail location is a challenging task for any business. There are many factors that need to be considered before a location is selected. Some of these factors include: competing locations, transportation networks and the demand for products in an area. Lowe's is currently facing this problem, as it is seeking to expand its operations in Canada by establishing a new store in the city of Waterloo. Furthermore, to help them in their decision making process, Lowe's has hired our group to compile a report analysing the viability of potential store locations around Waterloo. This report details our methodology and the results we received using ArcGis to perform network analytics on the given list of store location candidates.

Network analysis is the process of working with networks of interconnected vertices and edges. A vertex is a point in a network, and edges are used to connect these vertices together. These edges are typically assigned traversal costs, which represent the work required to move from one vertex to another using a particular edge. Costs can be measured in units such as time or distance. Some common problems in network analysis are finding the shortest route from one vertex to another or showing all of the vertices within a certain cost of a particular vertex. Network analysis can be applied to many real world scenarios, for example service areas can be calculated for police stations to make sure that police can be dispatched to any location within a certain amount of time.. Location-allocation modelling, a type of network analysis problem, deals with how effective a new installation will be based on factors such as distance to customers and competing locations.

Predicting which customers will shop at the new location is vital to choosing the correct site location. Network analysis was the most appropriate approach to finding this essential information. However, before we could utilize the network analyst toolset, some additional preparation was required. Creating a networked dataset and estimating demand both needed to be done before any advanced network analysis could take place.

Methods

Creating a Network Dataset

Our first step in solving this problem was to create a network dataset. To be able to do any kind of network analysis, a network must be present to analyze. In its initial state, the data that we were given was not formatted as a network; therefore, we had to use ArcGis to turn our street map into a network.

The street map contained all of the important elements required for a network. The streets were used as edges and the intersections were used as vertices. The length of the streets was used as the cost, or impedance, of the edges. In addition to the length of the route, another important consideration is travel time. The street map did not contain travel time

estimates, but it contained the speed limits for each street. Using the assumption that cars would always travel at the speed limit, estimated travel time was calculated using the street length and speed limit data. Lastly, the final thing to consider was one way streets. The street map had a field which indicated if a street was a one way street and, if it was, the direction of traffic. By inputting all of this data into the new network dataset tool, we created a new network dataset which could simulate the movement of vehicles along the streets of Waterloo.

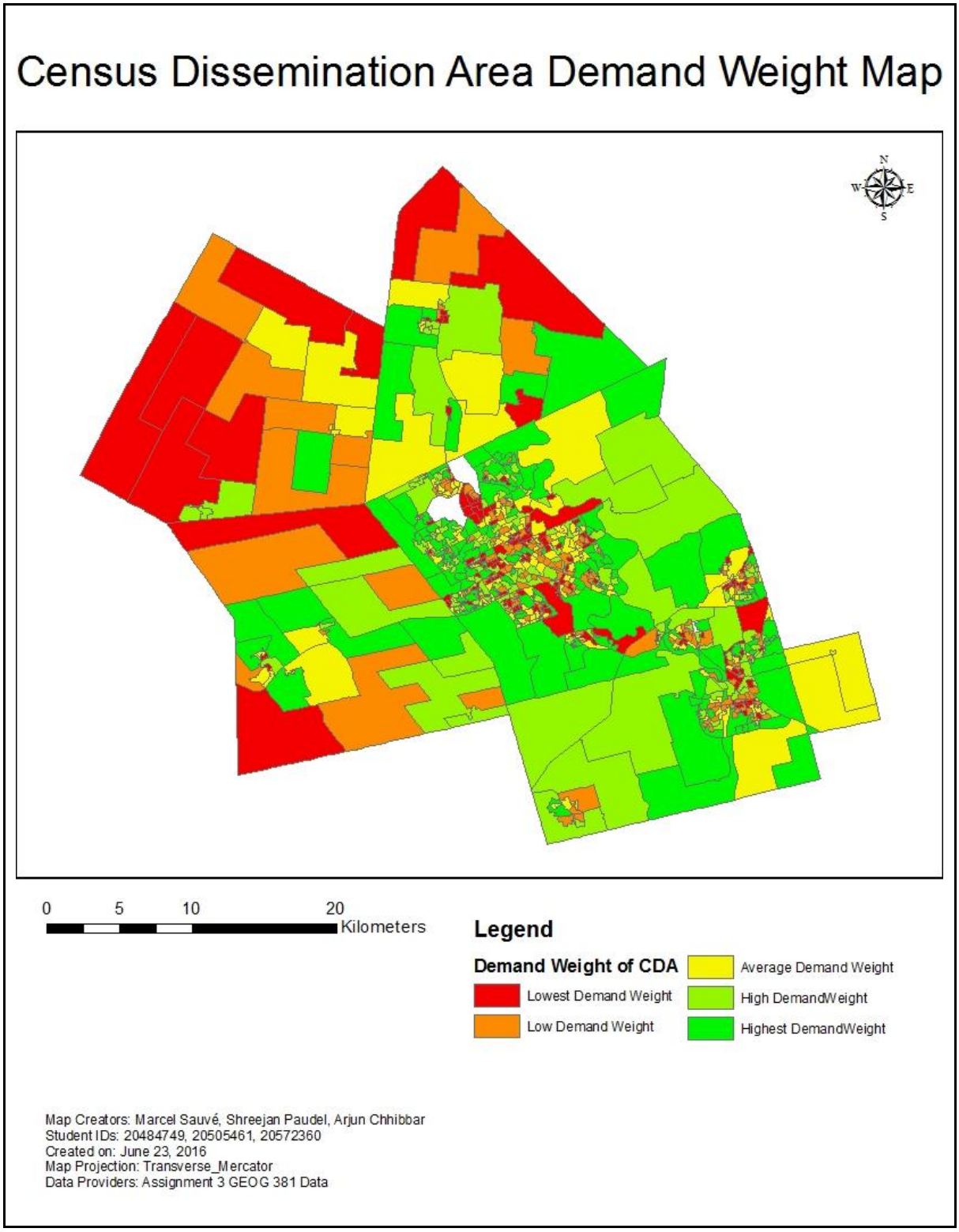
Estimating Demand

Our next task was estimating the demand for home furnishing and improvement products. The data we received from Lowe's contained census data about the average household income of the Kitchener/Waterloo area in 2005; by itself, this data was not enough to estimate the demand for Lowe's products. We used Ontario household expenditure data from the same year to compliment the income data. The expenditure data was broken down by category; the "Household furniture and equipment" category was the most representative of Lowe's products so it was used in our demand estimate. Roughly 2.9% of average household expenditure was used on household furniture and equipment. Average household expenditure was about 98% of average household income. We assumed that these ratios held true for the Waterloo area. The census data was divided into many census dissemination areas (CDAs). To estimate the total demand from each CDA, we multiplied its average household income by 0.98 to receive an estimate of household expenditure. The household expenditure estimate was multiplied by 0.029 to get an average household furniture and equipment expenditure estimate. This was then multiplied by the amount of households in the CDA to obtain a total demand estimate for the CDA. Figure 1 shows the raw expenditure and income numbers.

Total Expenditure of 2005: 66,857.
Household income before tax: 68,102.
Household furniture and equipment: 1,969
Approximate Income Ratio to Expenditure: 1:0.98
Approximate Expenditure ratio to Household Operations: 1:0.029

Figure 1 (Average Household Expenditure. (2008))

The reason we used 2005 expenditure data was to remain consistent with the provided 2005 income data from Lowe's. If these numbers were adjusted for inflation, they should produce a decent estimate for 2016. We choose not to adjust these numbers for inflation before running the calculations because the raw numbers were not. What is important for this analysis is demand of one CDA relative to other CDAs. Being close to many CDAs that have high estimated demand relative to other CDAs will be better for the store regardless of the raw numbers behind the relative data.



The CDA Demand weight map shows the relative estimated demand of the CDAs. The CDAs were split into 5 equally sized categories based on their estimate demand. Ideally, stores should be placed near as many "Highest DemandWeight" CDAs as possible while avoiding "Lowest DemandWeight" CDAs.

Now that a network dataset was in place and the demand for stores was estimated, we could proceed with network analysis. Firstly, we examined the existing competition that was

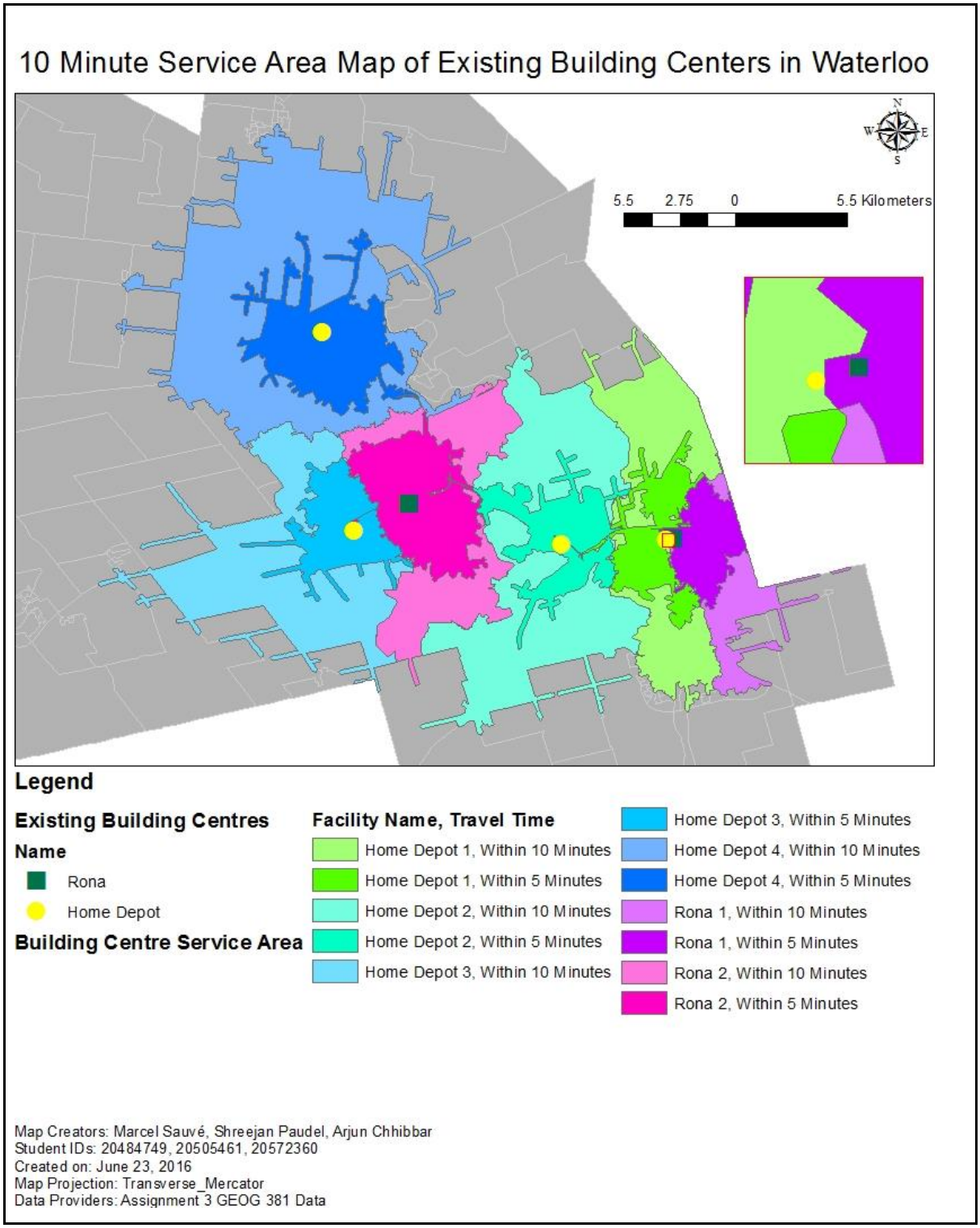
present in the Waterloo region. In the data package from Lowe's there was a feature class which held the position data of 4 Home Depot and 2 Rona stores. To examine what areas these existing stores served, we used the service area tool to overlay the areas that are less than 10 minutes away from the stores by car.

It was finally time to use the location-allocation tools that are present in the network analyst tool kit. These tools are commonly used to determine the placement of new stores and services, thus, they were the the most relevant to the task. The location-allocation tools required demand points to be created. Demand points are point features which symbolises where the demand for services are in relation to the network. These points can be weighted to make some points represent higher demand than others. Previously, we had estimated the demand in each CDA using income and expenditure data. CDAs are represented by polygon objects, where as demand points must be point objects. To comply with these requirements, we found the centroid of each CDA and assigned our estimated demand to the new centroid point object. These centroid points would serve as demand points for the the location-allocation simulation.

The location-allocation tool has many different settings which can be used to fine tune the analysis. For this scenario, we used the maximize market share function. The maximize market share function takes a list of candidate location and a list of competing locations. It then chooses the requested number of candidate locations in such a way to obtain the highest possible demand by servicing nearby demand points. It needs the competing locations because it factors the proximity of competitors to calculate how much demand will be serviced at each possible candidate location. It also has multiple functions to model distance decay. Distance decay describes how the further away a service or store is, the less likely a customer is to use that service or buy from that store. We selected the power function of distance decay, which means that the likelihood of going to the store decreases at an increasing rate the further away the demand point is from the store. This is opposed to a linear function where the rate stays constant. The location-allocation tool was used twice to model two different scenarios; the first selecting one candidate site, and the second selecting two candidate sites.

Results

Existing Service Areas

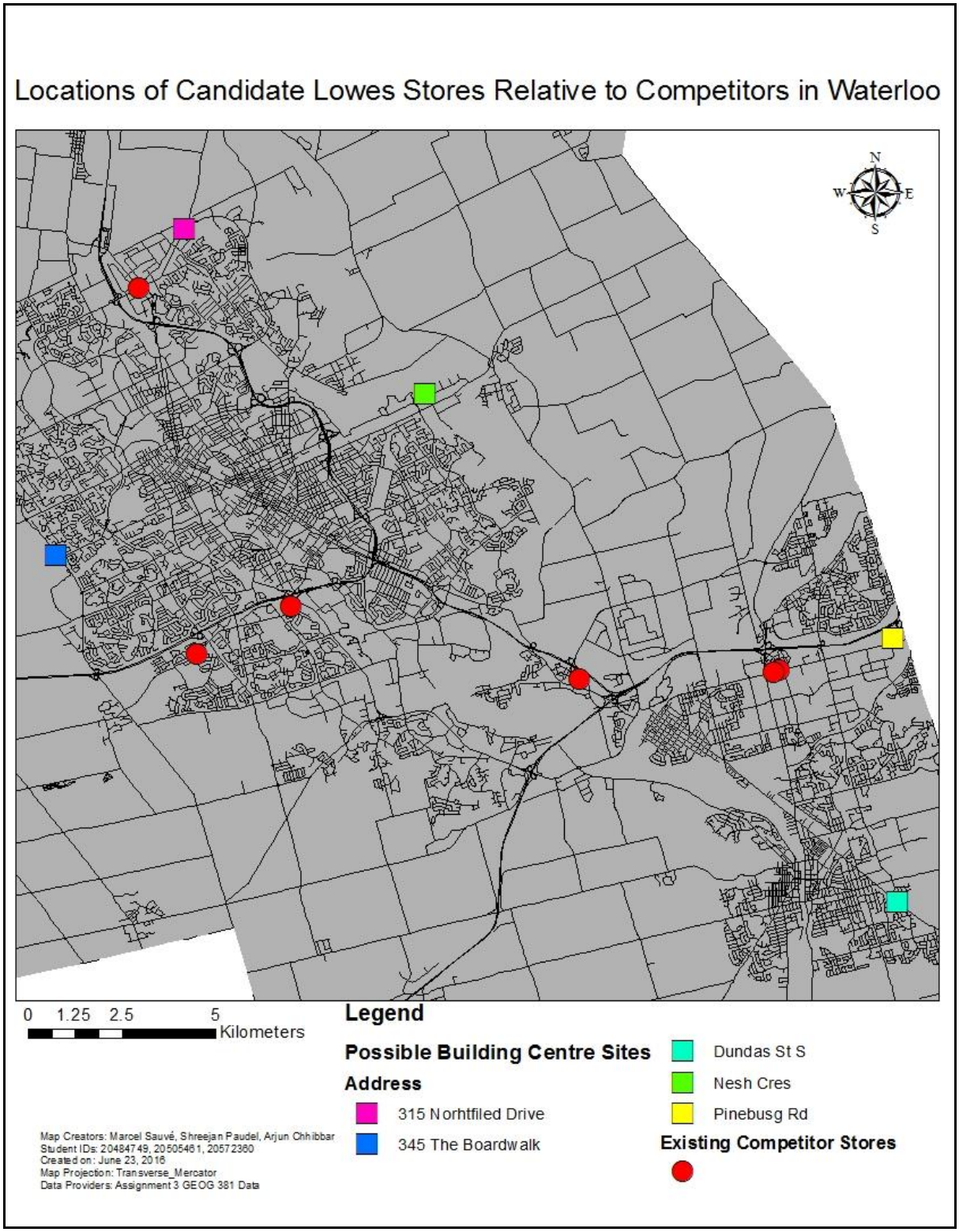


The above map shows the services areas of the existing building centres in Waterloo. There are currently 2 Rona and 4 Home Depot stores which will end up competing with any

Lowe's locations that are built. 2 Overlays were drawn around each existing competitor, one representing locations less than 10 minutes away and the other representing locations less than 5 minutes away. Rona 1 and Home Depot 1 are located so close to each other that their symbology touches on the map, because of this another frame was added to show that they are in fact two separate locations. This map is useful for visualizing which areas of the city already have easy access to similar types of stores.

It is already apparent that there are some populated areas which are underserved by the existing stores. An example of this is the area northwest of Home Depot 3 and southwest of Home Depot 4. The "345 The Boardwalk" Lowe's location is located within that area and would serve customers that do not have any nearby home building centres as well as divert some business from Home Depot 3 and 4.

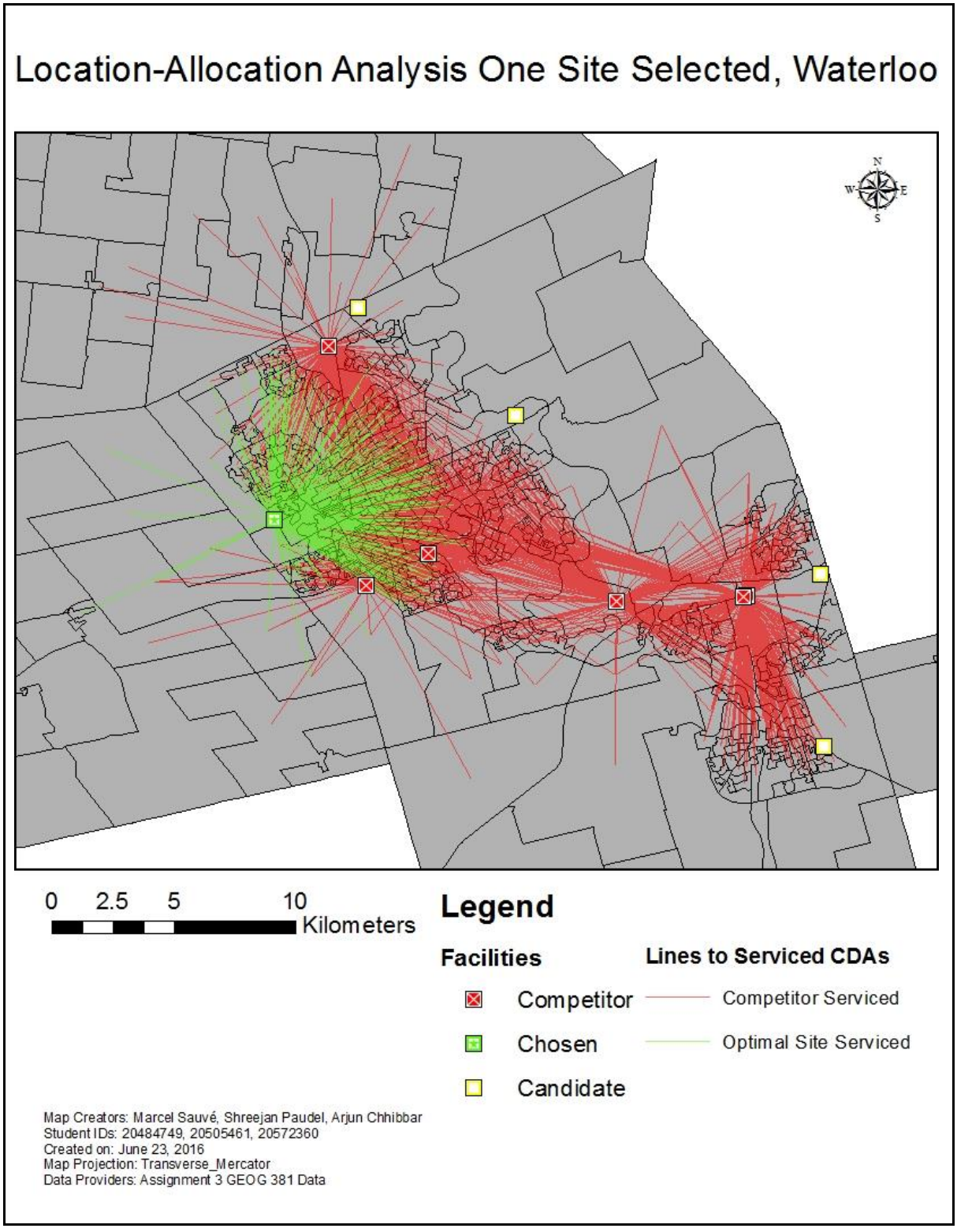
Candidate Area Locations



This map shows the location of candidate Lowe's stores. The Lowe's store candidates are named based on their addresses or nearby street names. They are coloured differently to better distinguish them. All of the candidate locations are relatively far away from

competitors except for the “315 Northfield Drive” location which is just northeast of a competitor.

Single Optimal Site Selection



This map displays competitors, candidates and the chosen Lowe's location. The lines that stem from the chosen and competitor locations led to the centroids of CDAs that will be serviced by that particular location. The candidates that were not chosen do not have any lines stemming from them because the lines are based on location and competition. Since the chosen location is the only location that impacts the scenario, it is the only candidate that serves any demand points. This is because modeling the other candidates would make them act as competition, since only one Lowe's is going to be built, taking them into account will give false results.

One demand point can be served by multiple locations. The total demand that is allocated to a location is based on the location's proximity as well as any competitors proximity to a demand point. For example, if two locations are near point A, location 1 can receive 40% of point A's demand while location 2 receives 60%.

As predicted by the "10 Minute Service Area of Existing Competitors, Waterloo" map, the "345 The Boardwalk" Lowe's was selected. Its location is excellent to cover the gaps in its competitor's service areas.

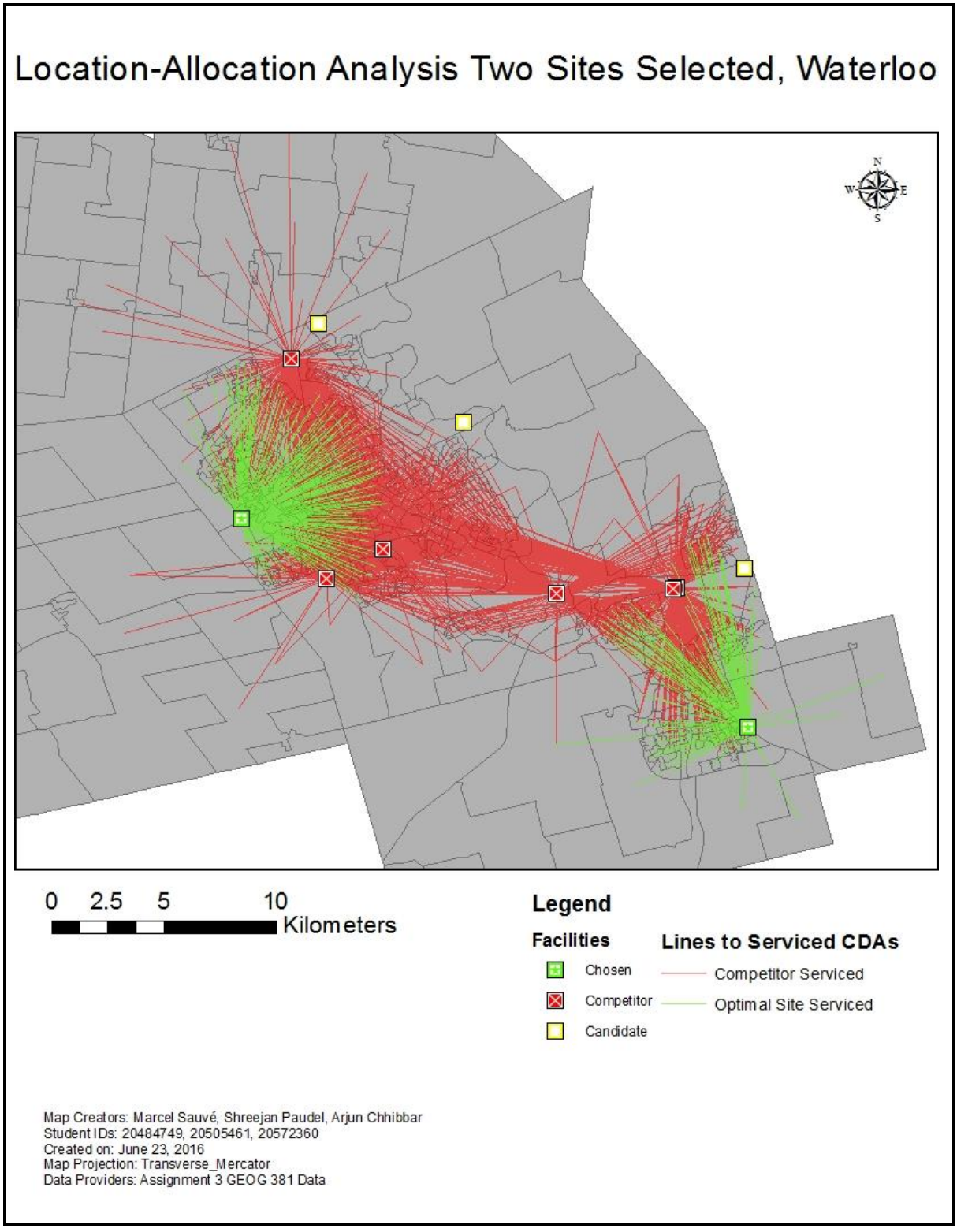
Name	FacilityType	DemandCount	DemandWeight
Lowes (345 The Boardwalk)	Chosen	252	42151708.44
Rona	Competitor	364	73416989.96
Rona	Competitor	209	39457767.83
Home Depot	Competitor	274	69889960.25
Home Depot	Competitor	335	50415585.62
Home Depot	Competitor	292	35062556.39
Home Depot	Competitor	213	39819749.39

Table 1

The above is a summary table of the building centre locations when a single site is picked for a new Lowes. The DemandCount column represents the amount of census dissemination areas that will be serviced by each facility. The DemandWeight column represents how much demand will be generated by all of the serviced census dissemination areas. The potential Lowes should be viable and compete well with the existing stores. Based on the location analysis, the Lowes will be in a location with more demand than half of the competing stores after being built. There are only 724 demand points in the network, but the

sum of the DemandCount column is far above 724. This is because one point can be serviced by multiple locations.

Dual Optimal Site Selection



This map is similar to the “Location-Allocation One Site Selected, Waterloo” map, but it has two selected locations instead of just one. The “345 The Boardwalk” location remains selected as the second selected candidate was far enough away to not affect the performance of the “345 The Boardwalk” store.

Name	FacilityType	DemandCount	DemandWeight
Lowes (345 The Boardwalk)	Chosen	252	42151708.44
Lowes (Dundas Street S)	Chosen	156	40762440.35
Rona	Competitor	364	73416989.96
Rona	Competitor	209	26320535.12
Home Depot	Competitor	274	69889960.25
Home Depot	Competitor	335	50415585.62
Home Depot	Competitor	292	33564186.40
Home Depot	Competitor	213	26907885.20

Table 2

The above is a summary table of the building centre locations when two sites are picked for new lowes location. The two locations are far enough to not affect each other. However, the new site will affect the surrounding competitors. The new Lowes gains 156 DemandCount and has high demand weight. It affects two Home Depots and one Rona. This placement will be viable for another Lowes to open and has the biggest demand compared to its neighbouring competitors.

Discussion

Limitations

Our approach has a few key limitations. Firstly, we only took one method of transportation into account in our network dataset. The network dataset only models the movement of cars along roads. Therefore, other methods of transportation such as; walking, public transit, biking, could have been implemented to better our results. Walking, in

particular, would have been interesting to model because pedestrians are not as restricted to streets as vehicles.

A second limitation is how we calculated the travel time of our network dataset. The only factors we considered were the length of the street segment and the speed limit. Cars do not constantly travel the speed limit, even in perfect driving conditions. Other factors, such as weather, construction, and traffic data can also impact travel time. More data could have been collected and analyzed to obtain a more accurate transport network simulation.

We also assumed that the only factor influencing a person's likelihood to shop at a store is the store's proximity. While proximity is a significant factor, this approach ignores differences in inventory between stores, how large a store is, and consumer brand loyalty. Of the previously mentioned factors and more can influence where a person shops and additional data and research could be used to model this more accurately.

Another serious limitation is the lack of data for areas near the "Pinesburg Rd" Lowe's candidate location. This candidate store is located right at the edge of the CDAs for which data was provided. Since they are right at the edge, it is unknown whether there is any demand directly east of the store. If there was significant demand in that area it may have been chosen as an optimal site.

A final limitation is that we assumed that areas with low income and expenditure would spend the same portion of their expenditure on Lowe's and similar products as areas which higher income and expenditure. This assumption may not hold true in reality. We speculate that households with more disposable income are more likely to spend a larger portion of their expenditures on home improvement compared to poorer families.

Long Term Planning

Since Lowe's plans for building centres to be active for at least 25 years, many long term factors can be considered when plotting the location for a new store. One key factor is land use plans for nearby land. If land next to a potential Lowe's location is going to be used for residential purposes it can be assumed that new customers will appear in the service area as the housing is built. Also, knowing where competitors plan to build future stores would greatly aid in long term planning. Those areas can be avoided so as not to serve a split customer base.

The economic development of an area can also be considered for long term planning. If the income of the residents around building centre improves, they may spend more money on home improvement and will generate more sales. Inversely, areas that are stagnating or declining economically should be avoided even if they seem economically viable at the present moment.

Infrastructure development is another key factor. If road networks are improved or expanded customers from further away will come to the store. Similarly, public transit improvements could increase attendance at store locations.