CEE 6625 GIS in Transportation

Homework 4:

Demand Analysis using Location-Allocation Analysis

11.18.2016

—

Han Gyol Kim, Sage Roberts, and Somdut Roy

School of Civil and Environmental Engineering
College of Engineering
Georgia Institute of Technology

Introduction

It has been said that the three most important decisions a retail business owner will make are "location, location, location" (Spaeder 2016). While many factors contribute to the success of a business, perhaps the most critical is the careful determination of locations. In addition to the spatial location, the number of locations must be considered. For the retail chain planning to expand its business in San Francisco, this report analyzes data for the area in order to recommend where to open stores, and how many to open. This is accomplished using location-allocation, the goal of which is to locate facilities in a way that supplies the demand most efficiently (ESRI 2016). For the retail chain, the specific goal analyzed is to minimize the number of stores while still serving the most customers.

Methodology and Assumptions

Location-allocation analyses simultaneously find the best locations for a facility and allocate demand points to those facilities (ESRI 2016). Because the best location varies for different types of facilities and goals, there are many different problems that can be solved with location-allocation analyses, including minimize impedance, maximize coverage, maximize capacitated coverage, minimize facilities, maximize attendance, maximize market share, and target market share (ESRI 2016). The ArcGIS Network Analyst tool was used to determine options for where the retail chain should open stores and how many to open, and then cost analysis was performed to determine the specific locations.

Location-Allocation Analysis

In order to choose the locations that will generate the most business for the chain, the objective of the analysis is to find stores that are close to population centers, and the premise of this is that people tend to shop more at nearby stores than ones farther away. For retail store locations, the three problem types that are typically solved are maximize attendance, maximize market share, and target market share (ArcGIS 2016). The maximize market share problem type was selected over the other two as the best tool for this analyses for the following reasons, along with short descriptions of the problem types (ESRI 2016).

- Maximize Attendance
    - Chooses facilities where as much demand weight as possible is allocated while assuming that demand weight decreases with distance.
    - Typically used for specialty stores that have little or no competition, or general retailers/restaurants without necessary data on competitors that is used for market share problems.
- Maximize Market Share
    - A specific number of facilities are selected so that the allocated demand is maximized in the presence of competitors, with the goal being to capture as much of the total market share as possible.

    - This problem requires the most data because you need comprehensive information about your facilities as well as competitor's data.
    - More specific information about this problem type is included below.
- <u>Target Market Share</u>
    - The minimum number of facilities are chosen that capture a specific percentage of the market, given that there are competitors.
    - Typically used to find how much expansion is necessary to reach a certain market share or keep up with competitors, but results often represent what stores would do if budgets weren't a concern. But if budgets are of concern, like in this analysis, stores just use the Maximize Market Share problem.

Location-allocation analysis was performed using the Network Analyst tool with the Maximize Market Share problem type. The overall methodology was to use the maximize market share problem, with an assumed impedance, and a specified number of facilities. Because the optimum number of facilities was not known, numerous options were tested, from 2 new stores to 7. For each of those, two analysis were run: one with the existing store being required and one where it was not required. Then the outputs from the model were used in conjunction with cost estimates for opening and closing stores, to find the optimum number of facilities to open, and where they should be located.

Inputs into the analyst were facilities (existing, candidate and competitor stores) and demand points. The specific inputs are outlined below, along with reasons and italicized *assumptions* that were made. All data for San Francisco used in the Network Analyst step was obtained from tutorial Exercise 9 (ArcGIS 2016).

Model Inputs:

- Create new "Location-Allocation" analysis layer
- Add point features to "Facilities" class
    - Required
        - These are facilities that must be included in the solution, and this setting can be adjusted once the facility is loaded to the map.
        - The Existing Store was added first as "required" but throughout the analysis, it was set as candidate to determine if it should be kept open or not (discussed later).
    - Candidate
        - These are potential places that a store can be opened, and a subset of them will be the solution to the problem.
        - There are an unlimited number of factors that can be considered when determining suitability of a site, including parcels with existing buildings that are large enough, are within budget, and is properly zoned. The location-allocation problem will only be able to output useable, appropriate locations if the inputs are actually suitable locations, and for that reason the already established candidate store layer was used (ESRI 2016).

        - *Assume the "CandidateStores" layer was generated based on a suitability analysis that was performed in order to find the locations that are suitable for the new retail store.*
    - Competitor
        - These facilities are specific to maximize/target market share problems, and represent facilities of rival businesses that you will have to compete against for the same clientele (ESRI 2016).
        - Each facility can be assigned a weight, which is only applicable in market share problems, and a higher weight represents a more important or attractive facility compared to others, such as larger square footage etc.
        - *Assume that candidate stores and competitor stores all have equal weights of 1, so facility attractiveness is just based on location with more demand is assigned to near facilities than far facilities. This is because of limited data and noting that distance is more important to customers of retail stores than other facility factors such as square footage which would be important for warehouses (ESRI 2016).*
- Add point features to "Demand Points" class
    - These points typically represent people or things that require the goods/services your facilites provide.
    - The demand points used are census tract centroids, weighted by the existing populations, POP2000, so that the shop can best serve those populations (ArcGIS 2016).
    - *Assume that people shop at nearby stores more than those that are further away, as previously noted.*
    - *Assume that the expected business generated by a census tract is proportional to the population there. This is because we are assuming that a retail shop is equally likely to be visited by anyone in the population, meaning that it is a shop such as a general store that appeals to everyone regardless of demographic.*

The figure below shows the initial input map we used for the analysis, with all of the data mentioned above.

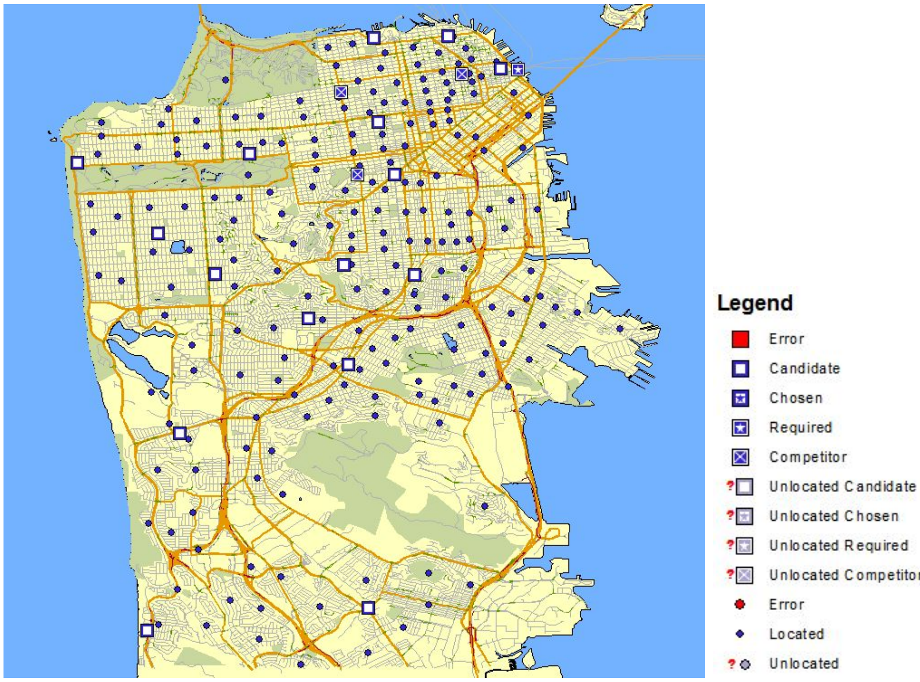


Figure 1: The study Area

Once the input data is added to the location-allocation layer, specific parameters need to be set on the Layer Properties dialog box, specifically for the Analysis Settings Tab and Advanced Settings Tab. These inputs specify different properties of the network and preferences, so that the solution obtained reflects the goals of the analysis. Figures of each are found below, with descriptions of why certain inputs were chosen and *assumptions* made.

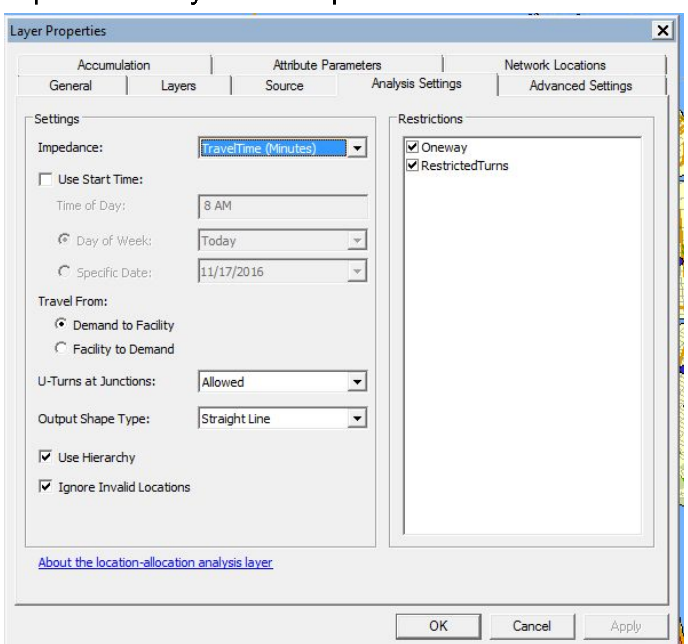


Figure 2: Analysis Settings

- Impedance

    - This property specifies the network cost attribute that is used to define the traversal cost along the elements of the network, this analysis used TravelTime (minutes).
        - *Assume that people are more likely to shop at a store closer to them than further, so travel time is in question.*
- Start Time/ Day
    - Specific travel start times and days were not selected
        - *Assume people will go shopping at all different times and days, it is dependent on various demographics and without access to that data it would not benefit the analysis*
- Travel From
    - Can calculate from demand points to facilities or opposite, and restrictions like one-way streets and impedances will affect travel times, thereby affecting the facilities chosen.
    - Used the “Demand to Facility” option
        - *Assume that store is more concerned with the time it takes people to reach the store, which is common for retail stores, instead of like a fire station where they travel to the demand (ESRI 2016).*
- Restrictions
    - Usually these cause roads to be prohibited or avoided, more useful for emergency vehicle routing or trucks with heavy loads.
    - Check the one-way restriction and restricted turns
        - *Assume that there will be a fair amount of one-way streets or restricted turns in a dense downtown area like San Francisco, so they could have a substantial impact on travel time.*

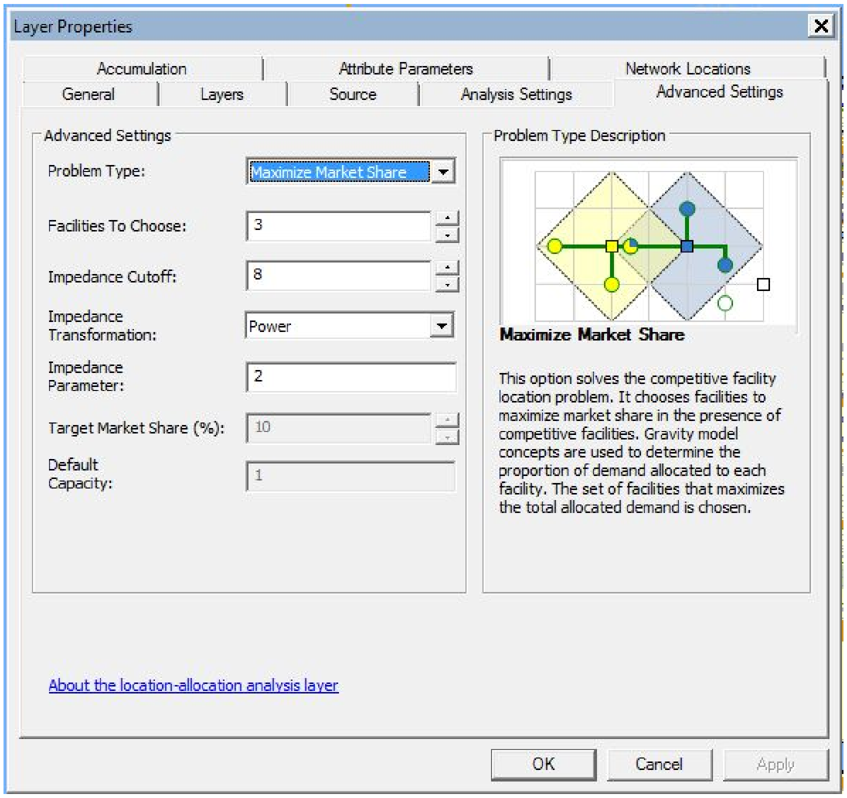


Figure 3: tab showing use of “Maximize Market Share” tool

The advanced settings tab is where the problem type is selected and its properties are set. It allows you to change constraints and influence the solver’s priorities for facilities.

- Problem Type
  - Previously described, and assumptions explained for reasoning of choosing Maximize Market Share
- Facilities to Choose
  - Started with 2 facilities, and ran analysis for up to 7 facilities, each for there being a required store, and no required store.
- Impedance Cutoff
  - This represents how far people are willing to travel to the facility, which in this case is the retail store.
  - *Assume people are willing to travel 8 minutes to a store in this location. This is because Americans, on average, are willing to travel 17 minutes to a local business, and 16 minutes to a store in particular (Marchant 2014). Because this study was for the whole US, where people live in a range of rural and urban areas, we assumed that people living in a densely populated area such as San Francisco would only be willing to travel half as far for a store. This assumption should be valid, unless the store is very specialized, because people are willing to travel farther for that.*
- Impedance Transformation

Model Outputs:

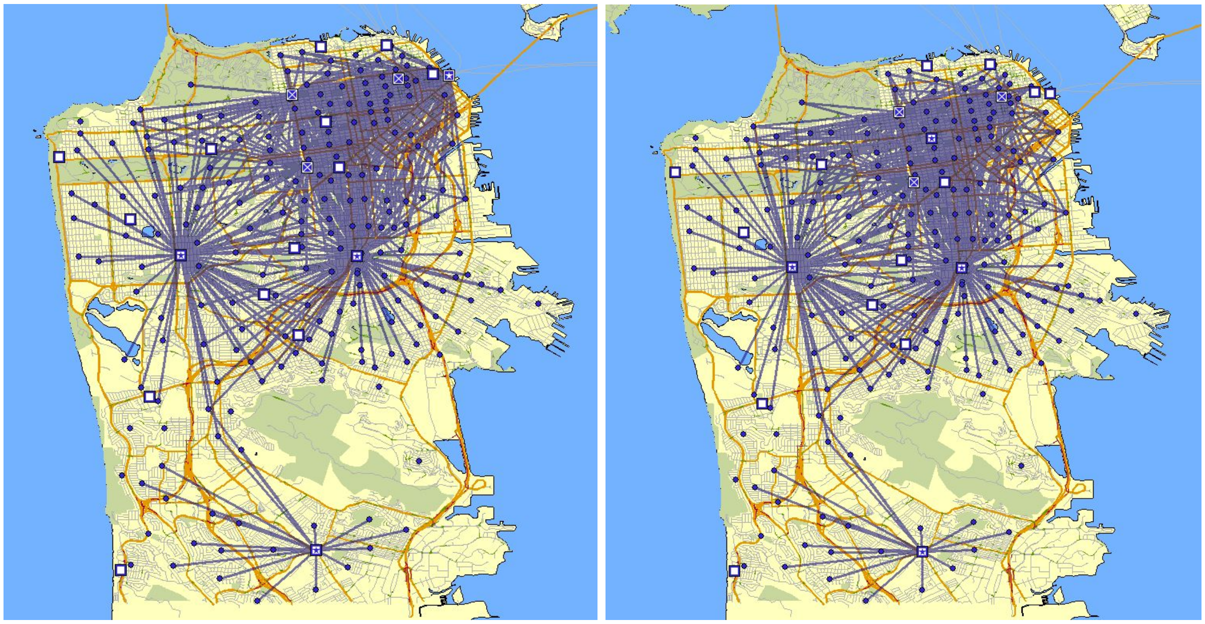
After analysis, following fields are outputs of the attribute tables for the network (ArcGIS 2016)

- Demand count
    - Count of demand points allocated to that facility
    - Did not use this for our analysis because the demand points overlap for several facilities, so we are more concerned with how much of that demand point is going to each facility, i.e. demand weight.
- Demand weight
    - This is the sum of the effective weight from all the demand points that are allocated to a specific facility.
    - For maximizing market share, the value is an apportioned sum of the weight field values, because they tend to follow decay models.
    - Specifically for maximize market share, the weight of a demand point is split among the facilities proportionally to the facility's attractiveness (weight) and has an inverse relationship with the distance between the facility and demand point (ESRI 2016).
    - *Assumed that the competitor facilities and existing/candidate facilities all had the same weight, because of a lack of data on other attractiveness factors. This means that more demand is assigned to near facilities than far.*
- Total Market Share
    - This is not directly output into the table, but it is the sum of the weight of all the demand points located on a network. It can be used to calculate the captured market share.
    - Total market share for our analysis included the sum of the demand weights for competitor facilities and chosen facilities. Used the weight assigned to our client's store, compared to the total/other competitors, to see if adding more facilities would increase the customer base enough to make it worth all the costs.

The table below summarizes the values for all the different combinations of number of facilities and required or not. As it is evident that the market share keeps on increasing as we keep adding stores. However, for every total number of facilities selected, the existing store was not chosen as one of the best locations unless it was required. But apart from that, there is a need to consider the costs associated with opening and closing all of these various locations.

Table 1: Market penetration variation with changing number of stores

Total Stores Open	**New Store**	**Existing Store**	**Total Demand Weight**	**Market Share**
2	1	1	749875	45.5%
	2	0	769061	60.0%
3	2	1	803020	60.5%
	3	0	864673	64.4%
4	3	1	898632	64.7%
	4	0	898632	70.1%
5	4	1	898632	72.2%
	5	0	924107	73.7%
6	5	1	924107	75.6%
	6	0	924107	78.0%
7	6	1	924107	79.3%
	7	0	924107	80.3%



4 Locations (3 New, 1 Existing) 4 Locations (4 New, 0 Existing)

Figure 4: Stores opened on the basis of maximum market share.

Attempts have been made to look into stores - competing, existing and potential on the basis of maximized market share. The figure above shows two such scenarios, the left one being where the analysis was done considering that the existing store needs to be there and the right one being for the existing store did not necessarily had to be in function if the retail had expanding the market share in mind.

Cost Analysis

This section is dedicated to explaining the breakdown of cost for opening and closing a store and the formula to analyze the overall index for comparison between different scenarios of opening of store.

Cost of Opening a store (small scale):

Table 2: Opening cost breakdown

Rent and operating Cost ($)		
Security Deposit	6300	
Advanced First Month rent + Others	6300	2800
Utilities security deposit	1000	
Local Improvement Costs ($)		
Customize	13000	
Fixing	10000	
Supplies	500	

Miscellaneous Operating Cost ($)		
Hangers	3000	
Liability Insurance (quarterly)	600	
Accounting setup	80	
Association fees and directories	800	
Office and cleaning	120	
IT ($)		
Computer and other electronics	900	
Tech Support	600	
Telecom service	100	
Domain registration	15	
Google services	50	
Web development	3500	
POS Software	1000	
Marketing ($)		
Signs	7000	
Décor	1000	
Trademark Registration	1000	
Public Relations ($)		
Grand opening event	1000	
Media	1500	
Advertising	200	
Total Cost ($)	62365	

The above is an estimate based on a report about opening consignment retail shop in Austin and comparing the cost of living of Austin and San Francisco. The cost of closing a store could include in it, costs such a remaining year rent, the advanced rent payment returns and the infrastructure cost due to closing, which is estimated to be around 20 percent of the opening infrastructure costs.

The cost breakdown is as mentioned below.

Table 3: Closing Cost breakdown

Rent ($)	
4 months rent	25200
Return advance	-1000
Infrastructure cost ($)	
20 percent of total	4700

Total Cost ($)	28900

After getting the cost sorted, the next step involves analysing different scenarios and their comparison based on certain principles of econometrics.

(a) The favorability of opening a store will have direct relation with the market share of stores of the retail chain and inverse.
(b) The favorability of opening a store will also have an inverse relation with the cost incurred due to opening of the stores as well as closing of the store (in case the existing store is closed).

To summarize, we take a sample equation:

f = (a x Net demand of Retail chain stores - b x Net demand of competitor store) / (total cost)

Where f = favorability quotient
a, b = unknown positive quotients

It is worth mentioning the factors considered for analysis are independent of each other. So, our result is expected to devoid of biases. The following twelve scenarios have been studied: (1) 2 stores including existing, (2) 2 stores excluding existing, (3) 3 stores including existing, (4) 3 stores excluding existing, (5) 4 stores including existing, (6) 4 stores excluding existing, (7) 5 stores including existing, (8) 5 stores excluding existing, (9) 6 stores including existing, (10) 6 stores excluding existing, (11) 7 stores including existing and (12) 7 stores excluding existing. A sample table for the case of 2 stores to be opened (existing store not being a necessity) is shown below.

Table 4: Sample summary table for spatial “maximum market share” analysis

2 TOTAL (2 CANDIDATES 0 EXISTING)								
FID	Shape *	ObjectID	Name	FacilityTy	Weight	Capacity	DemandCoun	DemandWeig
16	Point	17	Competitor_1	2	1	0	76	84984.9663
18	Point	19	Competitor_3	2	1	0	94	98736.48775
17	Point	18	Competitor_2	2	1	0	114	123926.0533
12	Point	13	Store_16	3	1	0	124	187708.6222
3	Point	4	Store_4	3	1	0	77	273704.8705
						Total		769061
						Our Market Share		0.5999699538

Trial-and-error method is used with different combinations of a and b (within reason) and the trend of f values in different scenario as a response to different values of a and b have been studied. The figure below shows a plot where a case with b=0.5 and different values of a across different scenarios yielded different trends of f.

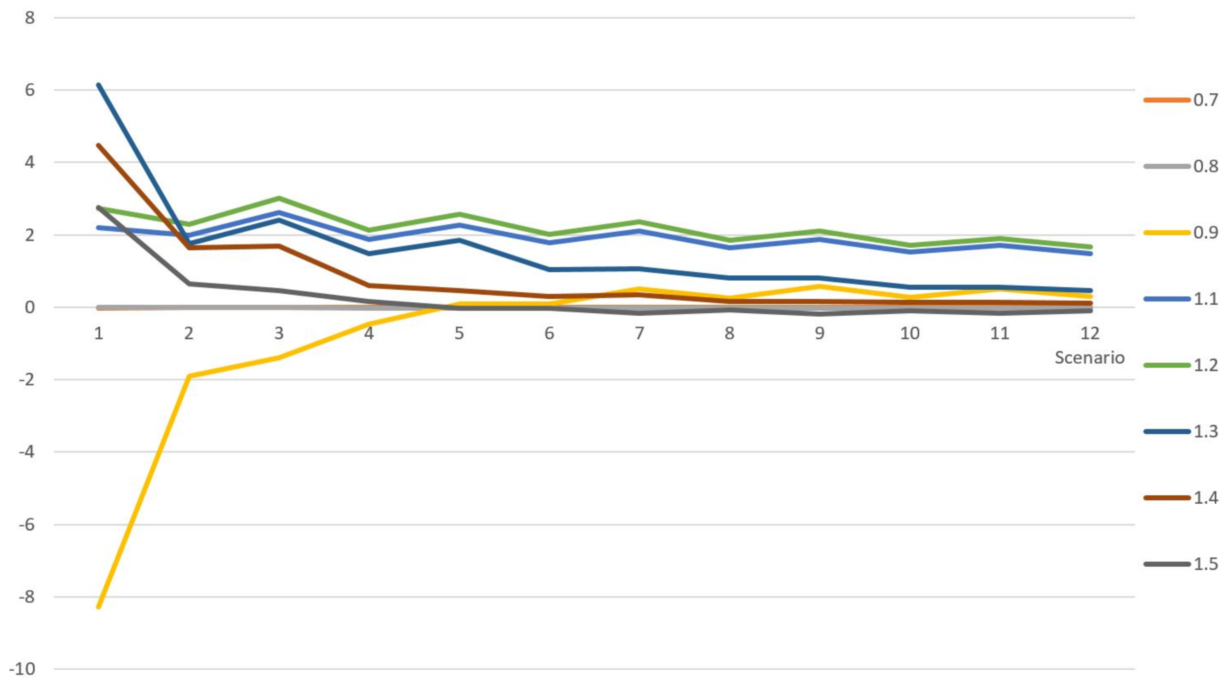


Figure 5: Different testing values of “a” with fixed value of “b”

Using several trails we came up with 3 trends which gave us definitive trends and all the trends surprisingly gave similar result regarding which scenario to go with for maximizing market share with minimum store and optimum cost.

(1) a= 1.1, b= 0.5;
(2) a= 1.2, b= 0.5 and
(3) a= 1.1, b= 0.7.

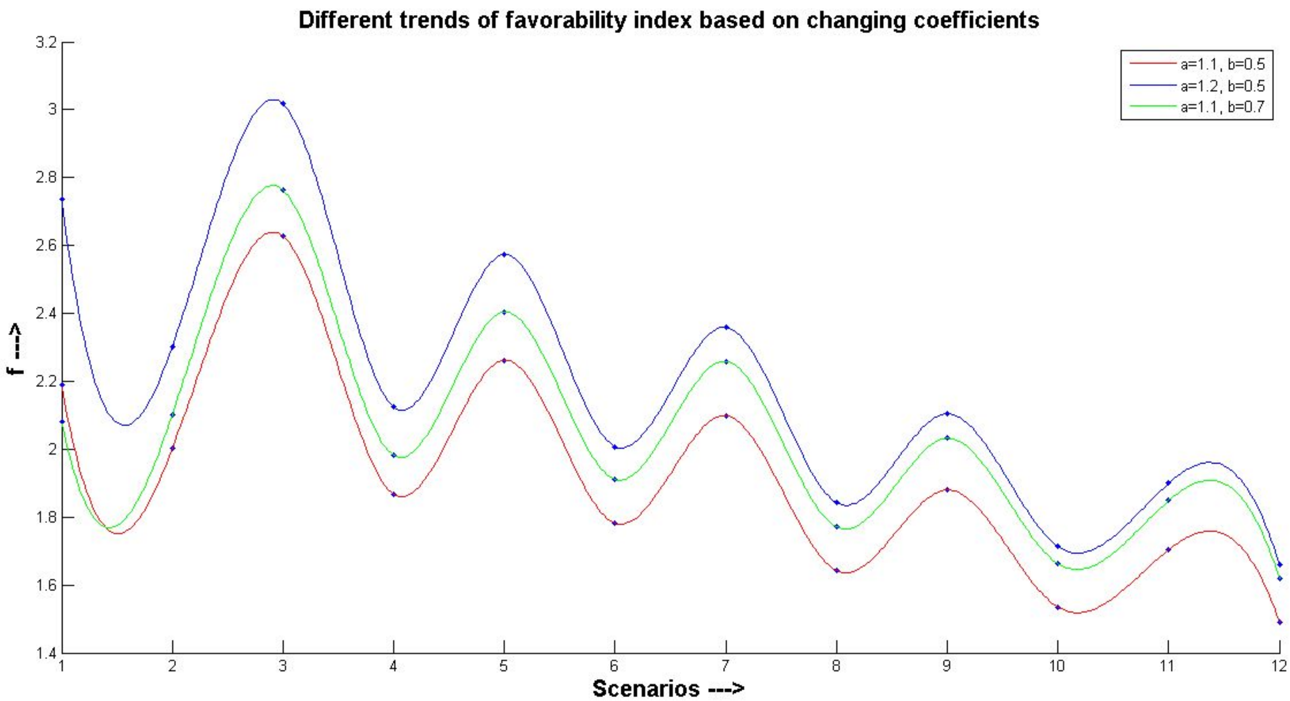


Figure 6: Accepting formula and their trends with changing scenarios.

Based on the figure above, it is evident that all the equations lead us to a conclusion that the third scenario (i.e. having 2 new stores and keeping the existing store) is the most favorable strategy for the retail chain as f seems to peak at 3 and has a constantly decreasing upper bound.

Addressing Existing Store

The analysis steps outlined above were performed for various numbers of total stores that the retail chain could have locations, including and excluding the existing store. It looks intuitive that the existing store needs to be closed considering that it never featured when we did not imposed the condition of mandatory existence of that store. But as we take the cost factor into consideration is becomes clear that closing any store needs lot of factors to be considered and presence of a store depends on factors other than just market share.

Addressing Competitor Stores

To locate our store, we used max market share, but just based on distance because detailed information about stores that could predict its attractiveness to general public was not provided. As Spaeder 2016 suggested, "Quite simply, the best place to be is as close to your biggest competitor as you can be," our final result also gave us store close to the competitor stores which further consolidates the idea.

Recommendations

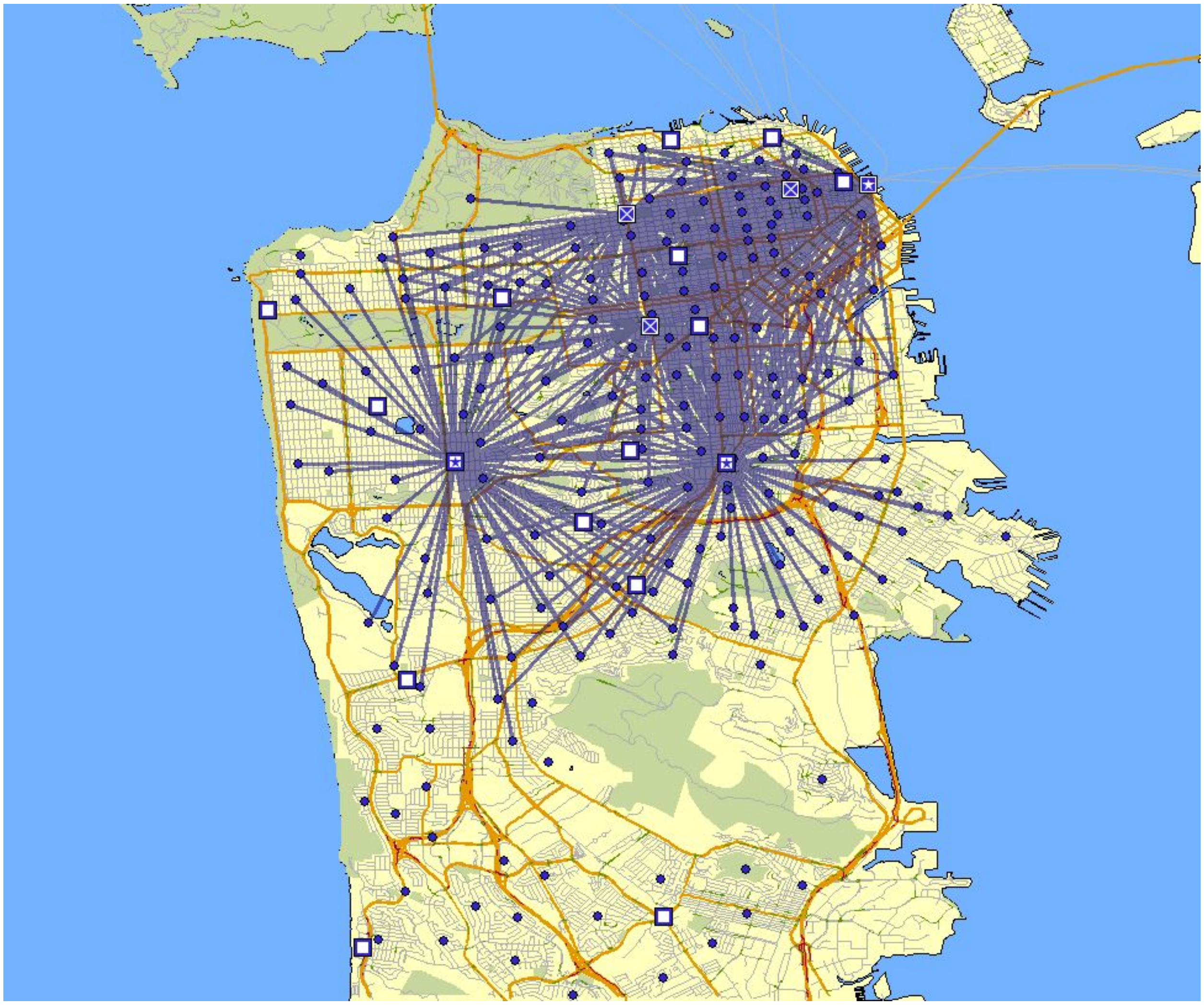


Figure 7: Map of Recommended Store Locations

Based on our analysis and further processing of the analysed spatial data, we have come to the conclusion that we keep the existing store and open two new stores at the locations mentioned in the map above. We may not be covering most of the southern part but as it is evident, a chunk of population lives up north justifying the location of all competitor stores being to the north.

Conclusion

Our work here is not free of limitations. Instead of going with the more mechanical trial and error method, proper demand analysis could have been done but due to lack of resources and prior knowledge this analysis was beyond our scope. Specific cost of the area could have been factored into our analysis to make it more reasonable. Long term cost analysis could have been done based on the operating costs derived from different available demand models.

As a continuation of this work, one could devise an index that consider more factors like age group, demographics which are integral part of demand analysis.

To conclude, it must be mentioned that in spite of limitations, the work can be a helpful tool to derive an idea necessary to make such business decisions on the basis of limited spatial data.

References

ArcGIS. (2016). “Exercise 9: Choosing optimal store locations using location-allocation”. http://desktop.arcgis.com/en/arcmap/latest/extensions/network-analyst/exercise-9-choosing-optimal-store-sites-using-location-allocation.htm

ESRI. (2016). “Location-allocation analysis”. http://desktop.arcgis.com/en/arcmap/latest/extensions/network-analyst/location-allocation.htm

Marchant, R. (2014). “Customers will travel 17 mins to reach a local business”. https://www.brightlocal.com/2014/05/01/local-business-travel-times/

Spaeder, K. (2016). “How to Find the Best Location”. https://www.entrepreneur.com/article/73784

Dahl, D. (2011). “The Cost of Starting Up a Retail Shop”. http://www.inc.com/articles/201108/business-start-up-costs-retail-store.html

Cost of Living Calculator | San Francisco, CA vs Austin, TX | NerdWallet - Personal Finance. (n.d.). https://www.nerdwallet.com/cost-of-living-calculator/compare/san-francisco-vs-austin