2018-12-06

# Using machine learning to find optimal locations for food bank depots

Mercado, Joseph

Calgary Food Bank

# Using machine learning to find optimal locations for food bank depots

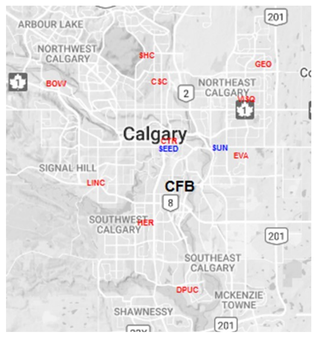
Joseph Mercado, BSc. Statistics Student

## 1 Introduction

- Calgary Food Bank (CFB) serves the city's food-insecure by providing food hampers.

| CFB Main Location | Pro | Con |
|---|---|---|
| Highfield Industrial | Storage space | Far from clients |

- To improve accessibility, CFB has satellite depots and hubs throughout the city.
- A computer algorithm clustered groups of client households and found central locations which would minimize travel distance.




**Benefits:** Less $CO^2$ emissions, easier access to services for clients, improve decision making with respect to depot planning.

**Left:** The locations of food bank depots (red), hubs (blue) and the main depot (bolded).


In the first 5 months of 2018, over **24,700** km have been saved by clients travelling to a satellite location instead of the main depot.

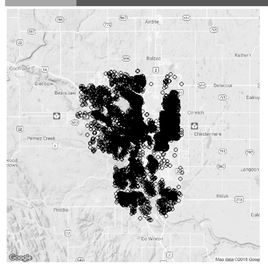
**Main Question:** Are depots located in optimal spots for clients?

**Clockwise from top:** main depot, Sunrise Community Link, Mustard Seed, Deer Park United, St. Andrew's

## 2 Depots and hubs

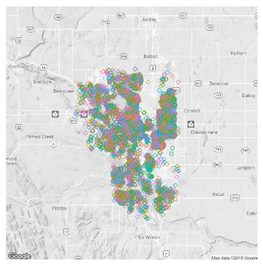
| Code | Name | Community | Code | Name | Community |
|---|---|---|---|---|---|
| BOW | Foothills United Church | Bowness | HER | St. Andrew's Presbyterian Church | Haysboro |
| CSC | Centre Street Church | Highland Park | LINC | Lincoln Park | Lincoln Park |
| CTR | Carter Place | Downtown | SEED | Mustard Seed | Beltline |
| DPUC | Deer Park United Church | Deer Ridge | SHC | Good Shepard Church | Thorncliffe |
| EVA | Evangel Christian Assembly | Forest Lawn | SUN | Sunrise Community Link | Southview |
| GEO | George-Don Hartman Sportslex | Coral Springs | VISQ | Village Square Leisure Centre | Pineridge |

## 3 Research approach

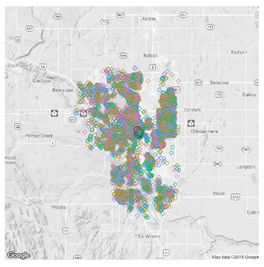


### K-Means Algorithm

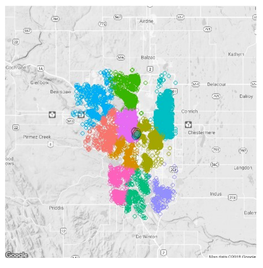
**Step 1)** Collected data pertaining to household locations of food bank clients who picked up a food hamper between January 1— May 31, 2018. In total, there were 6,815 unique locations.*



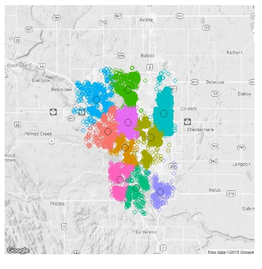
**Step 2)** Randomly assign data points to one of ten clusters. Ten clusters were chosen because it was desired that clients be grouped to one of ten depots.



**Step 3)** Calculate the *centroid,* indicated by a large circular point, of each cluster. The centroid is the central point of each cluster. Initially, centroids are grouped in the middle because households were randomly assigned a cluster.



**Step 4)** Redesignate clusters by assigning each point to the cluster of its nearest centroid.



**Step 5)** From the new clusters, recalculate the centroids.

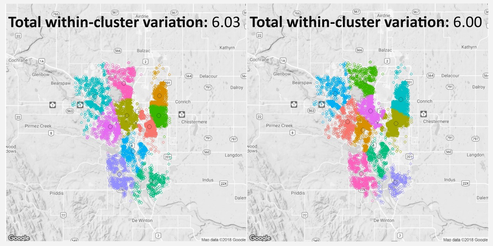
**Step 6)** Repeat steps 4 and 5 until the points no longer change clusters. This is considered a *local optimum*.

**Outlier Detection:** Since the K-means algorithm is very sensitive to outliers in the data, the "K-means minus minus" algorithm was applied to locate outliers. No outliers were present.

*multiunit dwellings with multiple client households were counted as one data point and one unique location.

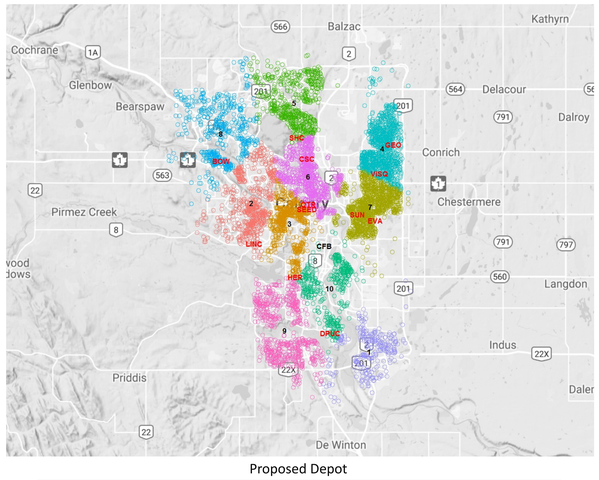
## 4 Optimizing

- Final result is dependent on how the points were initially randomized into the clusters. Local optimum may not be the best.
- Algorithm was applied a 100 different times and the randomization which yielded the smallest variance was chosen.



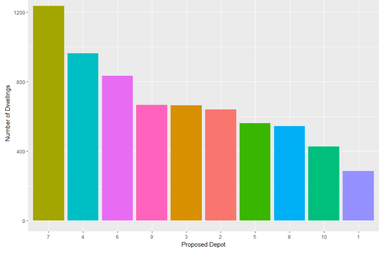

## 5 Results

**Below:** Proposed depot locations indicated by bolded numbers.




Proposed Depot

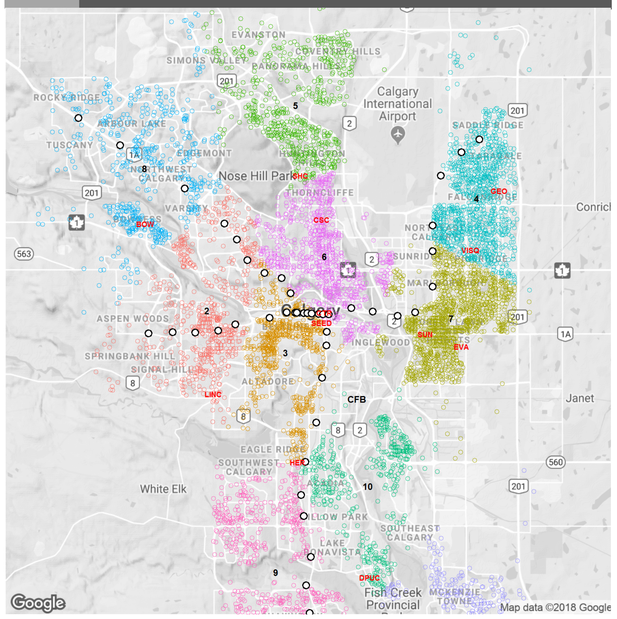
| Actual Depot | One | Two | Three | Four | Five | Six | Seven | Eight | Nine | Ten | Total |
|---|---|---|---|---|---|---|---|---|---|---|---|
| BOW | 0 | 22 | 2 | 0 | 7 | 6 | 0 | 79 | 0 | 0 | 116 |
| CSC | 0 | 8 | 4 | 5 | 49 | 91 | 2 | 13 | 0 | 0 | 172 |
| CTR | 2 | 11 | 38 | 2 | 6 | 34 | 7 | 3 | 2 | 2 | 107 |
| DPUC | 8 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 7 | 9 | 25 |
| EVA | 0 | 0 | 0 | 2 | 0 | 0 | 14 | 0 | 0 | 0 | 16 |
| GEO | 0 | 0 | 0 | 116 | 1 | 2 | 34 | 1 | 0 | 1 | 155 |
| HER | 0 | 0 | 2 | 0 | 0 | 0 | 1 | 0 | 6 | 4 | 13 |
| LINC | 0 | 8 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 11 |
| SEED | 0 | 0 | 4 | 1 | 0 | 3 | 1 | 0 | 0 | 0 | 9 |
| SHC | 0 | 3 | 2 | 1 | 14 | 10 | 3 | 4 | 0 | 0 | 37 |
| SUN-R | 1 | 1 | 1 | 2 | 0 | 1 | 57 | 0 | 3 | 2 | 68 |
| VISQ | 1 | 0 | 0 | 16 | 1 | 4 | 7 | 0 | 0 | 0 | 29 |
| Total | 12 | 53 | 56 | 145 | 78 | 151 | 127 | 100 | 18 | 18 | 758 |




**Top:** How satellite client households were distributed amongst the proposed depots.

**Left:** Number of dwellings clustered per proposed depot.

## 6 Conclusions




**Top:** Zoomed in map of Calgary showing CTrain stations in white circles.

**Right:** Locations of proposed depots.

| Proposed Location | Coordinates (lat, lng) | Community |
|---|---|---|
| 1 | (50.90934,-113.9699) | Inverness |
| 2 | (51.04807,-114.1448) | Wildwood |
| 3 | (51.02891, -114.0880) | Upper Mount Royal |
| 4 | (51.09900, -113.9501) | Falconridge |
| 5 | (51.14120, -114.0808) | Country Hills |
| 6 | (51.07269, -114.0600) | Tuxedo Park |
| 7 | (51.04448, -113.9682) | Forest Lawn |
| 8 | (51.11263, -114.1897) | Silver Springs |
| 9 | (50.92920, -114.0950) | Evergreen Estates |
| 10 | (50.96836, -114.0288) | Sue Higgins Park |

- Given the results, whether or not the proposed depots were near public transit was considered.
- However, figures from the HER depot location indicate that public transit access may not be an important aspect to depot planning.
- Many clients do not use public transit due to the bulkiness and weight of food hampers.

According to client surveys, only **11%** of clients used public transit to travel to CFB.

## 7 Citations


- Calgary Food Bank. (2018). *MySQL client database*. Unpublished data.
- Chawla, S., & Gionis, A. *K-means--: A unified approach to clustering and outlier detection* . Retrieved from: http://pmg.it.usyd.edu.au/outliers.pdf
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2017). *An Introduction to Statistical Learning.* Retrieved from https://www-bcf.usc.edu/~gareth/ISL/ISLR%20Seventh%20Printing.pdf
- Kahle, D., & Wickham, H. (2013). *ggmap: Spatial Visualization with ggplot2*. The R Journal, 5(1), 144-161. Retrieved from http://journal.r-project.org/archive/2013-1/kahle-wickham.pdf


## 8 Acknowledgements

Thank you to the Calgary Food Bank staff for always giving your full support behind my research endeavours. I would also like to thank Nexen for making it possible for students like myself to take part in non-profit sector research over the summer.

**Questions? Contact joseph.mercado@ucalgary.ca**