Where to Build Food Banks and Pantries: A Two-Level Machine Learning Approach

Gavin Ruan, West Lafayette Jr./Sr. High School

1. Introduction

Motivation and Background:
- In the United States, over **44 million** people suffer from food insecurity, **13 million** of whom are children.
- Food banks are critical for these people's nutrition and health (**49 million** turned to food programs in 2022)
- Last year, I created an application using the K-Means clustering algorithm. While this improved food bank locations, it didn't factor in important practical considerations such as roads, and also didn't factor in resources such as food pantries.

Project Goal:
- Improve the initial machine learning approach for finding optimal locations by **considering roads** and creating a **two-level system that has food banks and pantries**

2. Methodology

Optimization:
- Find the food bank locations with the shortest total **road distance** to all served households

Datasets:
- Datasets consisted of Indiana houses from 2020 GIS data

6 houses, 2 pantries, 1 bank

Food bank
Food pantry

OSRM
OPEN SOURCE ROUTING MACHINE

M	a	b	c	d	e
a	0	11	10	9	15
b	11	0	3	12	18
c	10	3	0	11	17
d	9	12	11	0	8
e	15	18	17	8	0

Distance/Dissimilarity Matrix

OSRM:
- Open source geospatial data
- Requires data to be scaled down in size

K-Medoids Method:
- Unsupervised clustering algorithm.
- Heuristic approach to solve facility location problems quickly using a distance matrix

Example flow chart for 12 houses and 2 food pantries

Given households
Choose Initial Food Pantries
For each household, find closest food pantry, forming two clusters
Update pantry location
If new location is better, keep it
Form new clusters

Comparisons:
- Generated 11 food banks and 176 total pantries to compare to an equal number of real ones

Performance:
- Jupyter notebook's time function for computational cost
- OSRM road distance converted to miles for distance measurements
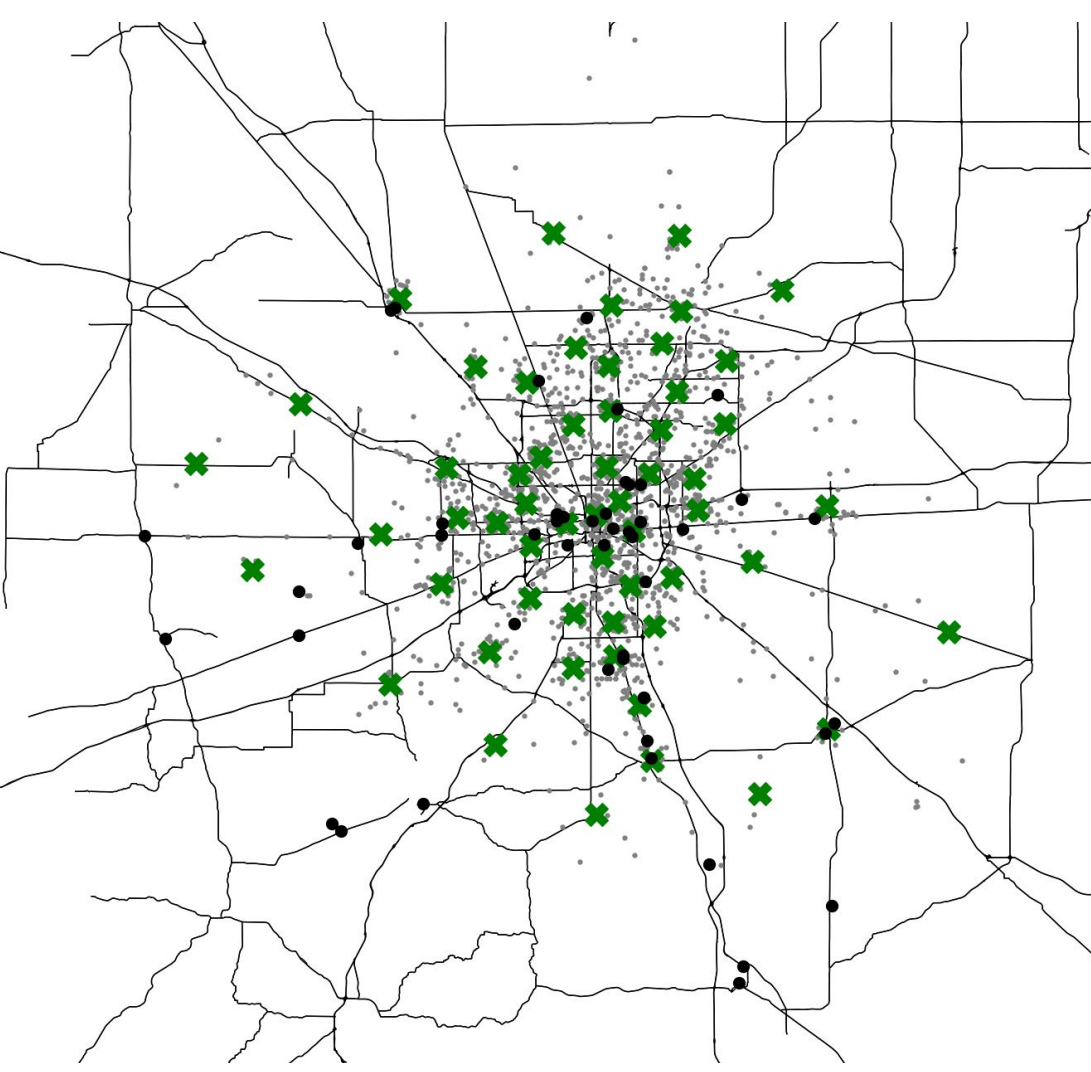
3. Results

Pantry-Household Level

● Real Food Pantries   ✖ AI Generated Pantries
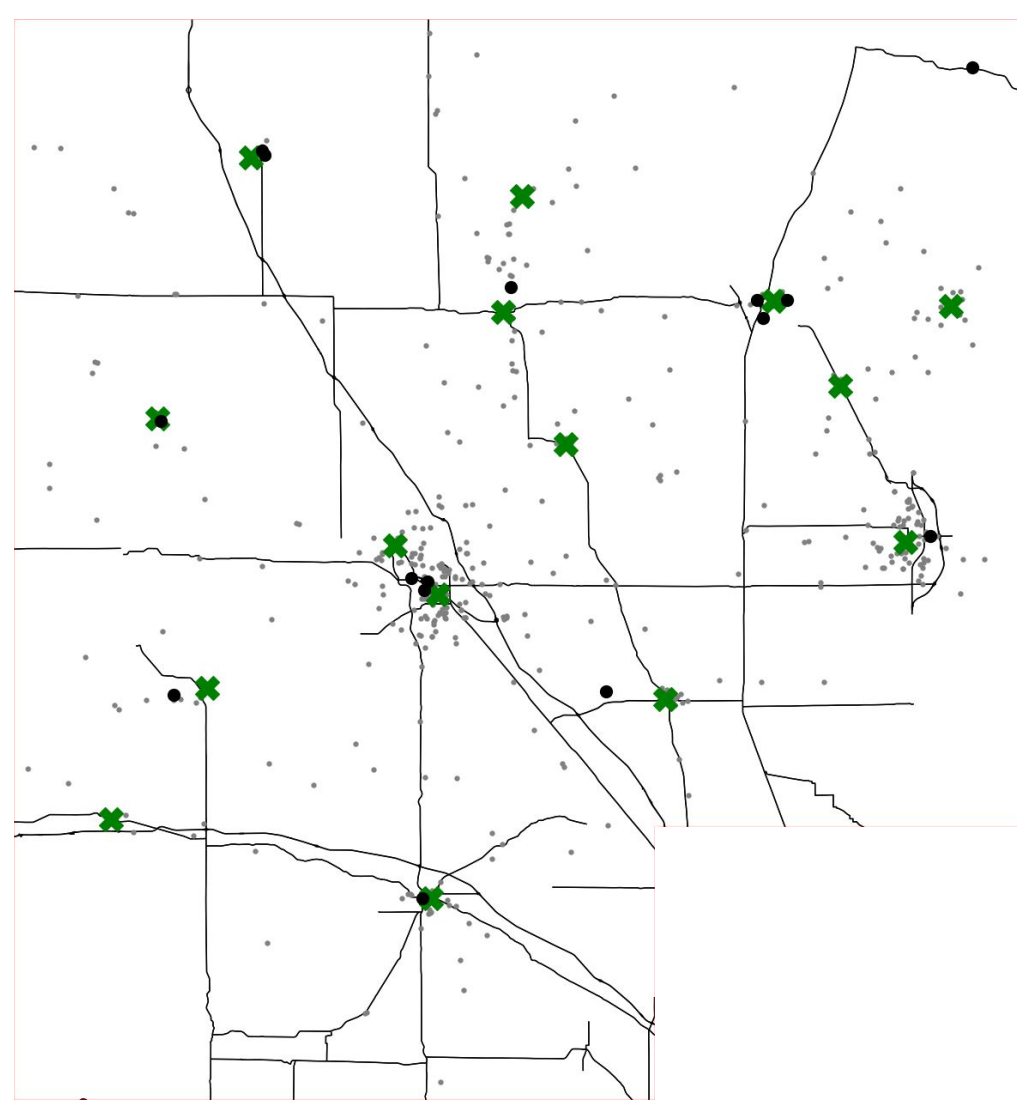
Lafayette
Original distance: 9.34 mi.
AI distance: 6.37 mi.
Distance saving: 2.97 mi.
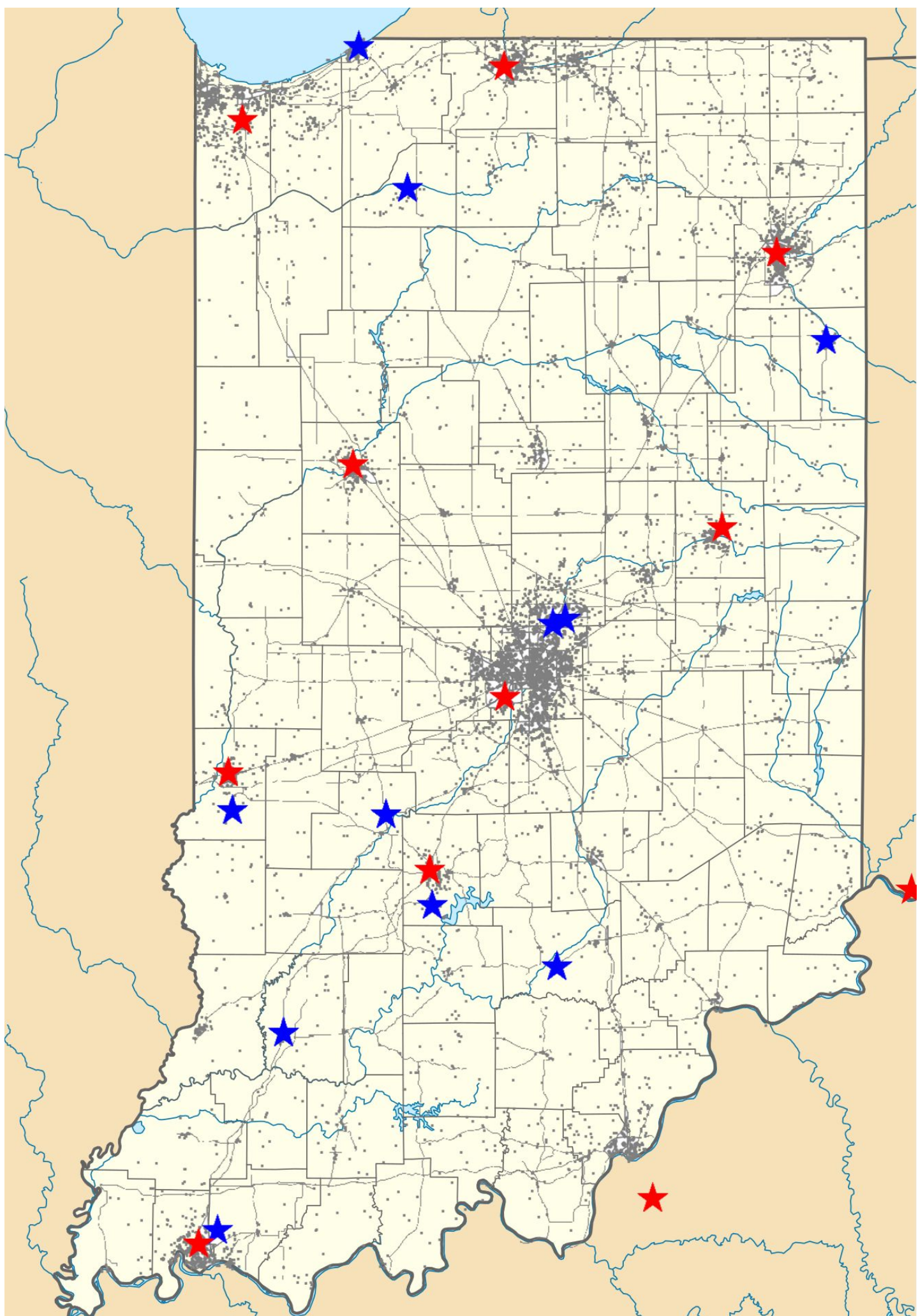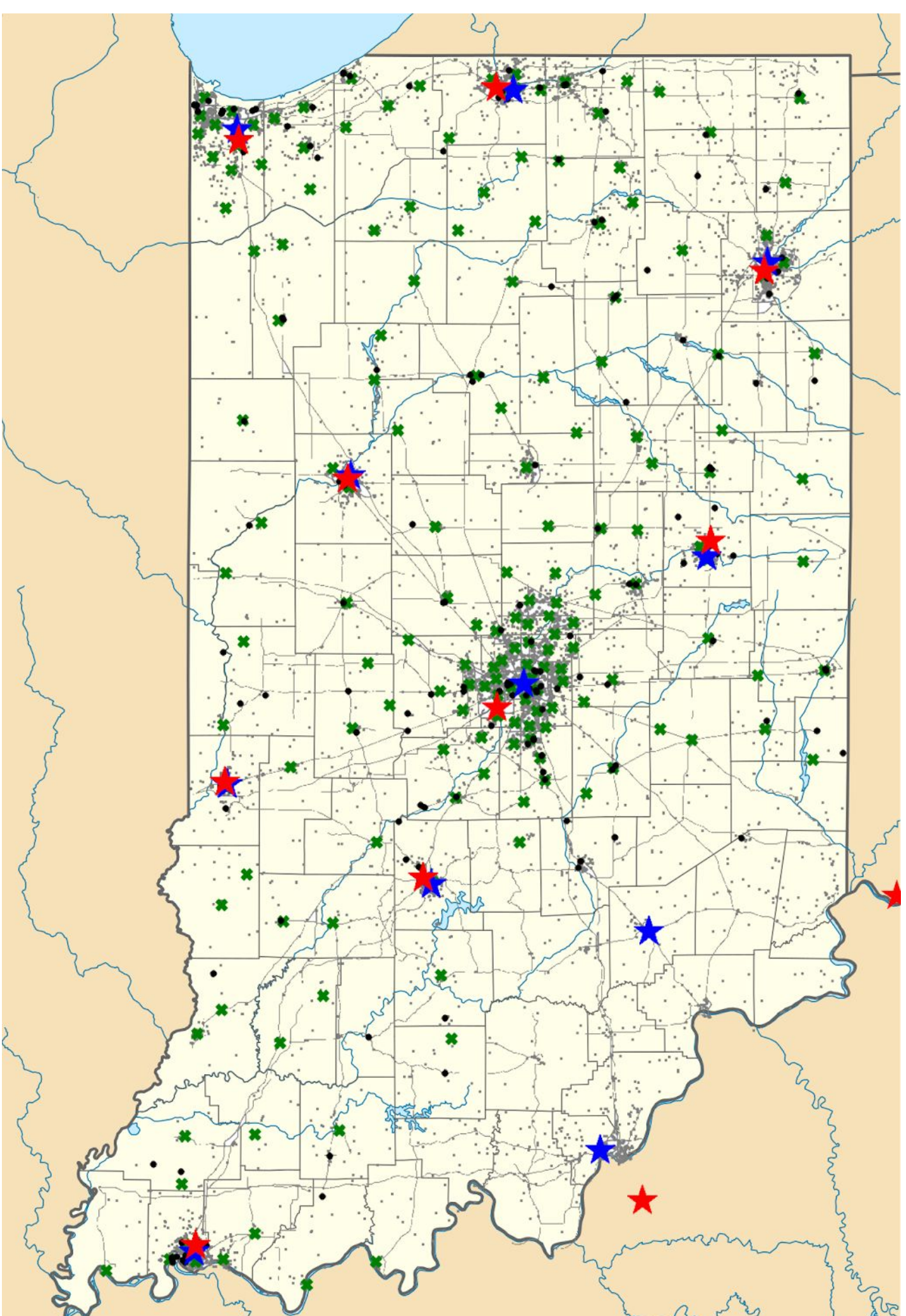Houses: 456
Number of pantries: 15

Indianapolis
Original distance: 4.41 mi.
AI distance: 2.25 mi.
Distance saving: 2.16 mi.
Houses: 1718
Number of pantries: 57

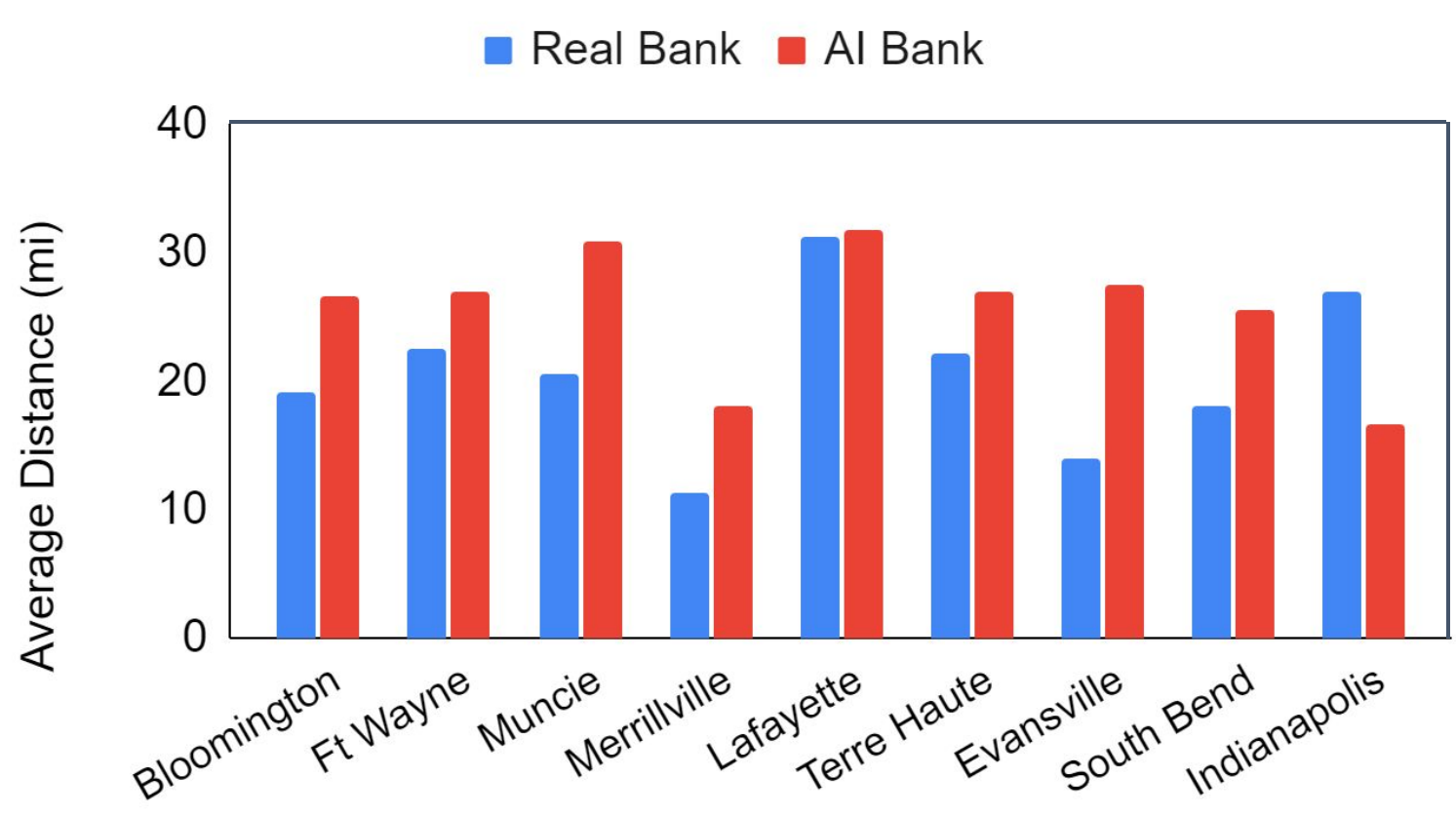**Large Distance Savings between pantries and households:**
- 6293 Houses
- Clusters ranged from 270 to 1700 houses
- **3.52 average miles saved per household**
- **22,181.423 total miles saved**
- **Average saved miles ranged from 2.17 to 8 miles in different cities**

■ Real Pantries ■ AI Pantries

(Bloomington, Ft Wayne, Muncie, Merrillville, Lafayette, Terre Haute, Evansville, South Bend, Indianapolis)

Food Bank-Pantry Level

★ Real Food Banks   ★ AI-Generated Food Banks

Initialization   Converged Results

(Computation Cost (s) vs Number of Houses chart)
18.4 years!
1.8 weeks
98 seconds
15.3 hours
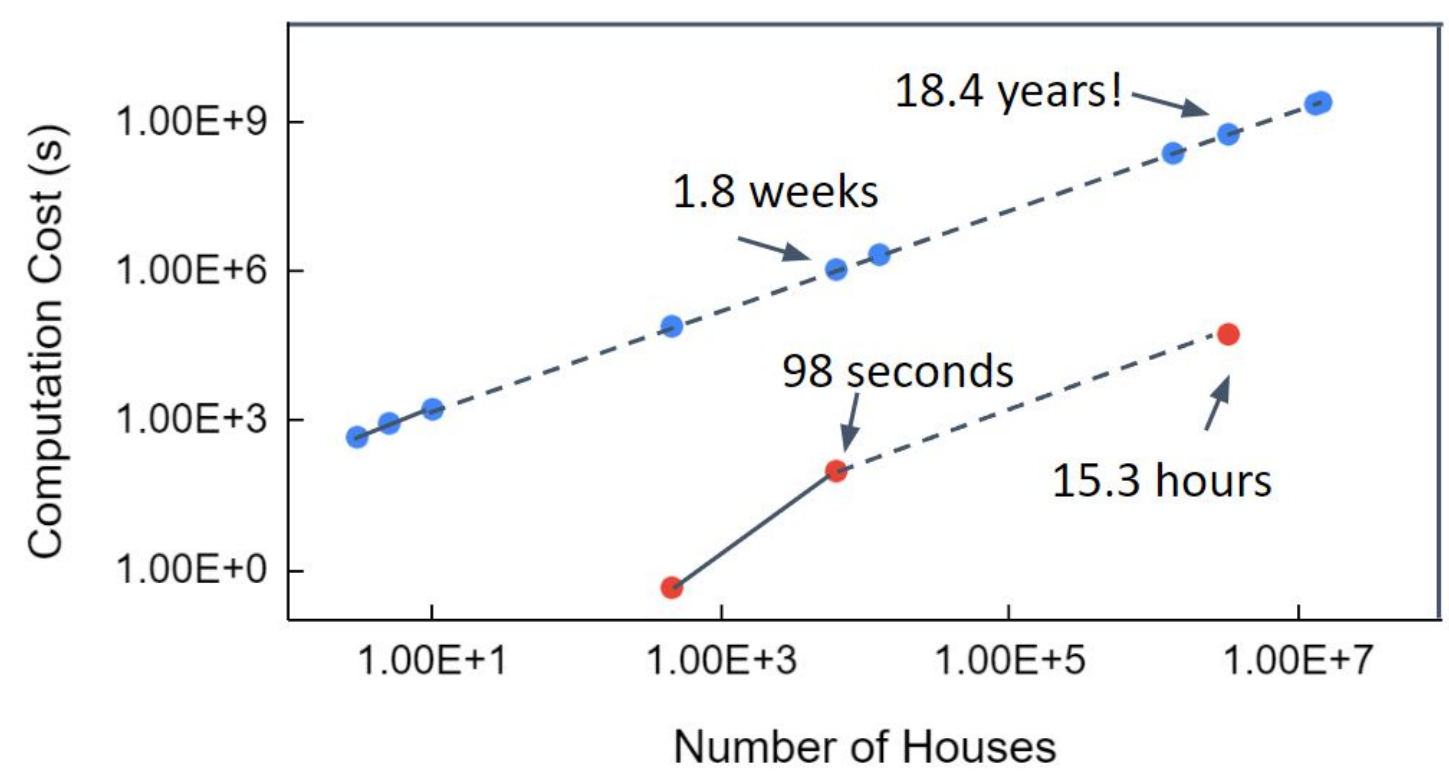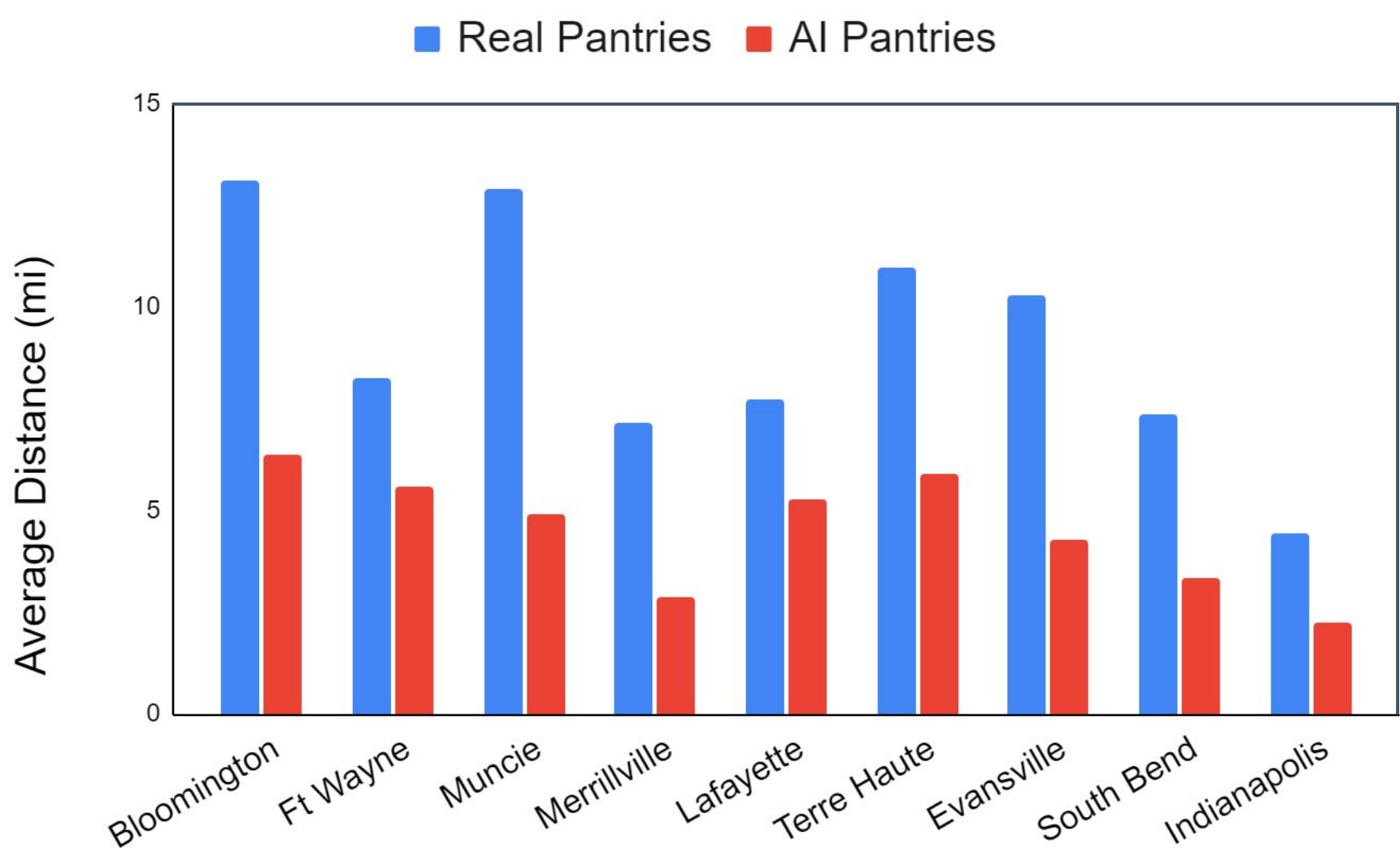
# Houses	Bruteforce	AI
6293 houses	1.8 weeks	98 seconds
~3.3 million houses	18.4 years	15.3 hours

Small Distance Penalty on Bank Level:
- 176 pantries
- 1.56 miles penalty per pantry
- 273.75 miles penalty

■ Real Bank ■ AI Bank

(Bloomington, Ft Wayne, Muncie, Merrillville, Lafayette, Terre Haute, Evansville, South Bend, Indianapolis)

4. Conclusion

- Results show that the two-level approach is able to consider real roads and generate a set of food banks and pantries both extremely quickly, and with more optimized locations than current existing locations.
- Current layouts prioritize food pantry proximities with food banks
- Contrary to the status quo, AI has showed that the planning strategy needs to be changed to prioritize households

5. Future Work

- Consider capacity of food banks
- Try a bottom-up approach with the food bank placements
- Expand dataset size and include different weights for houses (income range, socio-economic data, etc.)

6. References

Celik Turkoglu, D., Erol Genevois, M. A comparative survey of service facility location problems. Ann Oper Res 292, 399–468 (2020). https://doi.org/10.1007/s10479-019-03385-x

Héctor J. Carlo, Francisco Aldarondo, Priscilla M. Saavedra & Silmarie N. Torres (2012) Capacitated Continuous Facility Location Problem With Unknown number of Facilities, Engineering Management Journal, 24:3, 24-31, DOI: 10.1080/10429247.2012.11431944

Shih, H. (2015) Facility Location Decisions Based on Driving Distances on Spherical Surface. American Journal of Operations Research, 5, 450-492. doi: 10.4236/ajor.2015.55037.

Varghese, S.; Gladston Raj, S. Clustering Based Model For Facility Location In Logistic Network Using K-Means. Int. J. Sci. Invent. Innov. 2016, 1, 26–32.

Unless otherwise noted, all charts, graphs, and images are work of the author