OPR 624 Project

Group Members: Alex Fu, Jiameizi Yao

Sites Selection for Dockless Vehicle Charging Stations

Background

Louisville government has been promoting the dockless vehicles in the city since 2018. These find-and-ride vehicles offer a fun and easy option for getting around the city. We noticed that these vehicles are electronic and users usually park wherever they want. Louisville government has released a series of policies indicating how companies should operate their dockless vehicles business in the city. Louisville government has also published the data about individual dockless vehicle trips from July 2018 to January 2020.

Problem Description

Although dockless vehicles are supported by the Louisville government, companies are responsible to move or collect any vehicle that is parked improperly, and to provide necessary charging and maintenance services to vehicles. Companies should also pay attention to the restricted areas (Figure 1) and available distribution zones (Figure 2).

We would like to set charging stations in each distribution zones in order to provide services to the vehicles that need to be collected or charged while also follow the policies that Louisville government released.

Our goal of this project is to find the optimal station locations in each distribution zones which will minimize the total distance between each vehicle that need services and the charging stations, saving time and reducing labor costs. We will mainly utilize the geographic data (longitude and latitude) of both vehicle trips and restricted areas as coordinates. Model will be trained based on 333,512 trip observations from July 2018 to January 2020.

Model

- Variables and Notations:

  A set of coordinates of charging stations in each distribution zones.

$$(x_j, y_j)$$

$$j = 1, 2, \ldots n \ (n = number\ of\ charging\ stations)$$

  The ending geographic positions of each vehicles in the sample data.

$$(a_i, b_i)$$

$$i = 1, 2, \ldots m \ (m = number\ of\ observations\ in\ the\ data)$$

Binary variables to determine whether vehicle i is assigned to station (x_j, y_j)

$$w_{ij} = \begin{cases} 1 & if\ vehicle\ i\ is\ assigned\ to\ station\ (x_j, y_j) \\ 0 & otherwise \end{cases}$$

- Objective Function:

Minimize the overall euclidean distance between all dockless vehicles, which need to be collected and charged, and their belonging charging stations located in different distribution zones.

$$\text{Minimize} \sum_{j=1}^{n} \sum_{i=1}^{m} w_{ij} \sqrt{(a_i - x_j)^2 + (b_i - y_j)^2}$$

$$i = 1, 2, \ldots m\ (m = number\ of\ observations\ in\ the\ data),$$

$$j = 1, 2, \ldots n\ (n = number\ of\ charging\ stations)$$

- Constraints:

1. Each vehicle should only be assigned to one charging station.

$$\sum_{i=1}^{m} w_{ij} = 1\ for\ each\ vehicle\ i,$$

$$w_{ij} \in \{0,1\}\ for\ each\ vehicle\ i, station\ j$$

2. Each distribution zone should at most have one charging station, and all stations should be built in the specific distribution zones.

$$(x_j, y_j) \in set\ of\ cordinates\ in\ zone\ j$$

3. The stations must not locate in any restricted areas in the city.

$$(x_j, y_j) \notin (p, q)$$

$$where\ (p, q) \in set\ of\ cordinates\ in\ restricted\ areas$$

Data Overview and Manipulation

Louisville government has established various data and resources about the dockless vehicle. We mainly utilize part of the resources based on different purpose. The details about the resources we use for the project are shown below:

- Dockless Vehicle Trips

This data is about individual dockless vehicle trips across the city from August 2018 through July 2020. Vehicle ID has been removed for privacy concerns. A detailed data variable description is shown below:

Variable Names	Description
TripID	A unique ID created by Louisville Metro
StartDate	In YYYY-MM-DD format
StartTime	Rounded to the nearest 15 minutes in HH:MM format
EndDate	In YYYY-MM-DD format
EndTime	Rounded to the nearest 15 minutes in HH:MM format
StartLatitude	Rounded to nearest 3 decimal places
StartLongitude	Rounded to nearest 3 decimal places
EndLatitude	Rounded to nearest 3 decimal places
EndLongitude	Rounded to nearest 3 decimal places

We mainly utilize “EndLatitude” and “EndLongitude” as coordinates (a_i, b_i) to test on the model. These are the m observations we use in the model that we find the optimal station locations where the sum of all the distance between the ending geographic locations of these observations and the belonging stations should be minimized. However, we also use “StartTime”, “EndTime”, “StartLatitude” and “StartLongitude” to help identify the vehicles that needed to be charged, and group them by “TripID” and “Date”. Since vehicle ID has been removed for privacy concerns, we use above variables to restore trips into their belonging vehicles. Based on the start and end time, start and end positions of the trips, we are able to roughly figure out which trips belong to one vehicle. We also find some extreme cases where the trips end in couple seconds and end without moving, and assume most of these vehicles in the extreme cases are either broken or out of power, thus they need to be collected.

- Dockless Vehicle Distribution Zones

This data shows the operating area distribution zones (Figure 2) for the dockless vehicle. Zones are divided and used for equitable distribution of vehicles across the city. The data includes variables such as “Dist_Zone” which is the zone number and “PI2040Area” which is the area description for the zone (e.g., East Core). One main variable from the data is the geometry information, which tells the set of coordinates on the edge of each distribution zones. It ends up with a polygon area if all the coordinates in one set is connected one by one followed by a ring-like sequence. With the “Shapely” package in Python, we are able to determine or restrict if a point is within a polygon area. Besides, we decide to only open one charging station in each distribution zones.

- Dockless Regulated and Restricted Areas

Similar to the distribution zones, data of regulated and restricted areas (Figure 1) shows the set of coordinates on the edge of each restricted areas which ends up with a polygon shape on the map. The goal of using this data is to make sure that none of the charging stations locate within any of these restricted areas.

- Dockless Vehicle Policy

We refer to the official dockless vehicle policy from Louisville Metro for a lot of information in terms of operating requirements, especially policies about the operational zone and distribution zones.

As mentioned above, the data that Louisville government establish about the dockless vehicle trips has removed vehicle ID for privacy concerns. In order to find every single vehicle with aggregated trips, we use start/end time and start/end location to roughly restore trips into their belonging vehicles.

Typically, we assume that trips with start/end time and start/end location exactly matched are all belong to one vehicle. And we end up with 8,655 vehicles as our observations.

Approach Description

The objective function of the model can be seen as two parts. The first part is the distance matrix between each vehicles and the unknown charging stations, which is represented as:

$$d_{ij} = \sqrt{(a_i - x_j)^2 + (b_i - y_j)^2} \; for\ each\ vehicle\ i\ to\ each\ station\ j$$

Assuming we construct 9 charging stations (1 for each zone) in total and there are 300 vehicles in the observations, it will end up with a 300 x 9 matrix expressing all the possible distance between the vehicles and stations.

However, it's realistic that a vehicle could only be assigned to one station, then the second part of the objective function is to introduce a binary variable w_{ij} which expresses whether vehicle i is assigned to station j or not. This ends up with another 300 x 9 matrix but with all inputs to be binary. Then multiplying the binary matrix w_{ij} with the distance matrix d_{ij} gives out the final distance matrix that vehicles travel to their belonging stations. By summing up all the distances in the final matrix, we are able to find the total distance that dockless vehicles travels in the city. And the objective is to minimize the total distance in order to provide convenient service to vehicles that need to be collected and charged.

As mentioned in the distribution zones description, we set at most one charging station in each distribution zones in order to better serve each area, also because the zones are divided based on the equitable distribution purpose, thus, the stations are also equitably distributed in the city and the total number of stations should be at most nine.

In terms of the constraint regarding to the fact that one vehicle could only be assigned to one station, we make a constraint that the sum of w_{ij} for each vehicle j equals 1. Essentially, this constraint sums up all the assignment situation for each vehicle j to all possible stations and force it to be 1, which expresses that only one station should be selected for vehicle j to assign.

The last two constraints are both area-related constraints. We implement a great package called “Shapely” with its “Point” and “Polygon” modules that allow us to force a specific coordinate to be in or out of a desired area, which solves the problems related to distributions zones and restricted areas.

Implementation and Result

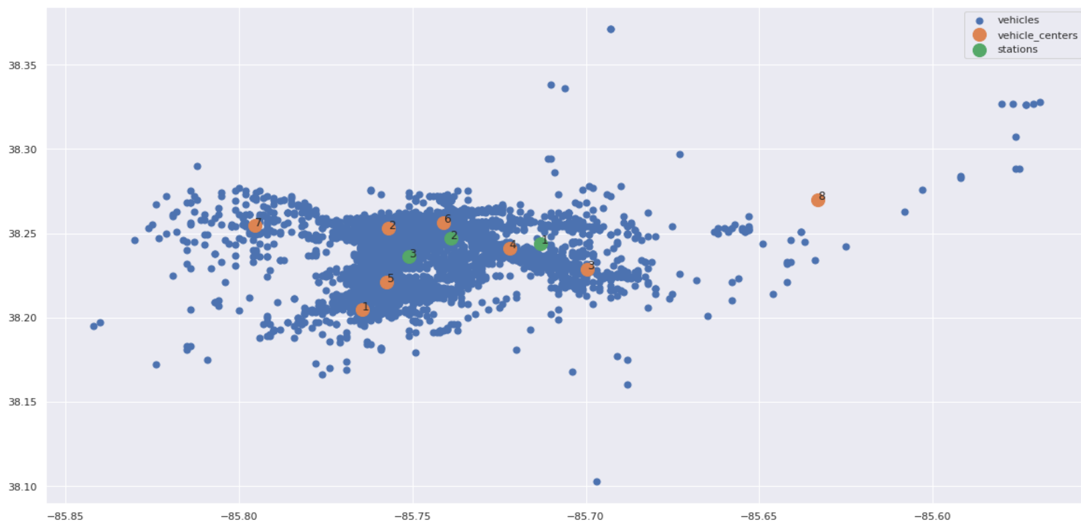
To simplify the implementation process of the model, we use K-means model to help reduce the obeservations when implement the model. We aggregate all the oberservations into 8 clusters, as shown below, and figure out the center of each clusters, then we use these center points as our new observations (a_i, b_i) and assign them to 3 different charging stations, which ends up with a 8 x 3 matrix of the assignment solution and 3 coordinates solutions of the charging stations to demonstrate our objective, which then is to minimize the total distance between the 8 cluster centers and the 3 possible charging stations.

Also, for the formula of Euclidean distance, we only use the square of the difference between longtitude and latitude, instead of calculating the square root additionally, to reduce the calculation burden.

After using the LSO solver in SAS, we got the result of optimized stations, whose coordinates are:

[1]	X	Y
1	-85.713	38.244
2	-85.739	38.247
3	-85.751	38.235

Then we plot the original eight thousands died scooters, eight centers and optimized stations, together. As we can see, the three stations are located in the center of the all eight thousands vehicles but still separated to each other. Also, their latitudes are similar, longitudes are scattered along with the overall shape of the vehicles' location.



Assignment Result Map

Model result:

Solution Summary	
Solver	LSO
Objective Function	TotalDistance
Solution Status	Function Convergence
Objective Value	0.0181298862
Infeasibility	0
Random Seed Used	1
Evaluations	9675
Cached Evaluations	2743
Iterations	64
Presolve Time	0.00
Solution Time	0.53

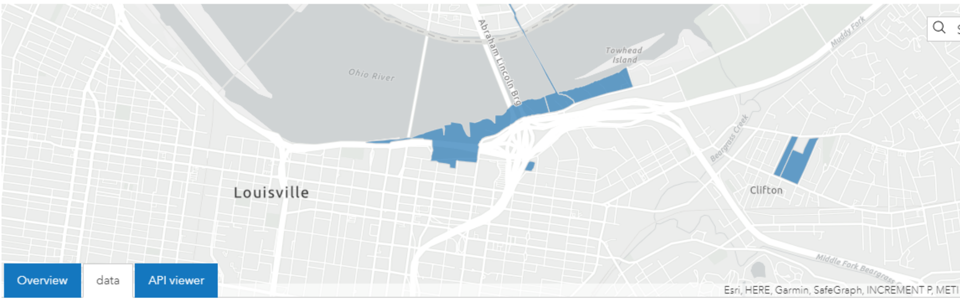
W	1	2	3	4	5	6	7	8
1	1	0	0	0	0	1	0	1
2	0	1	0	1	0	0	0	0
3	0	0	1	0	1	0	1	0

The process takes 64 iterations and the total distance optimized is 0.01813, which is based on coordinates. The matrix W shows the assignments result for the stations and cluster centers. For example, center 8 is assigned to station 1, and center 2 is assigned to station 2. To clarify, the index in tabel W might not be consistent with the label in the assignment result map.

Appendix

Louisville KY Dockless Regulated and Restricted Areas - No Parking

Last update last year | 5 records



2019/9/20 feature layer customization license

download ▾ FIRE ▾

Display 1-5/5

Tip: Used to filter columns

OBJECTID	NAME	Shape__Area	Shape__Length
1	American Printing House for the Blind	38337.84765625	1154.22723386537
2	KY School for the Blind	98311.2890625	1592.63027970463
3	Louisville Extreme Park	9221.09375	398.095466295219
4	Big Four Bridge	76774.484375	3558.13900357813
6	Waterfront Park	534465.1171875	7763.22154163145

Figure 1: Restricted Areas

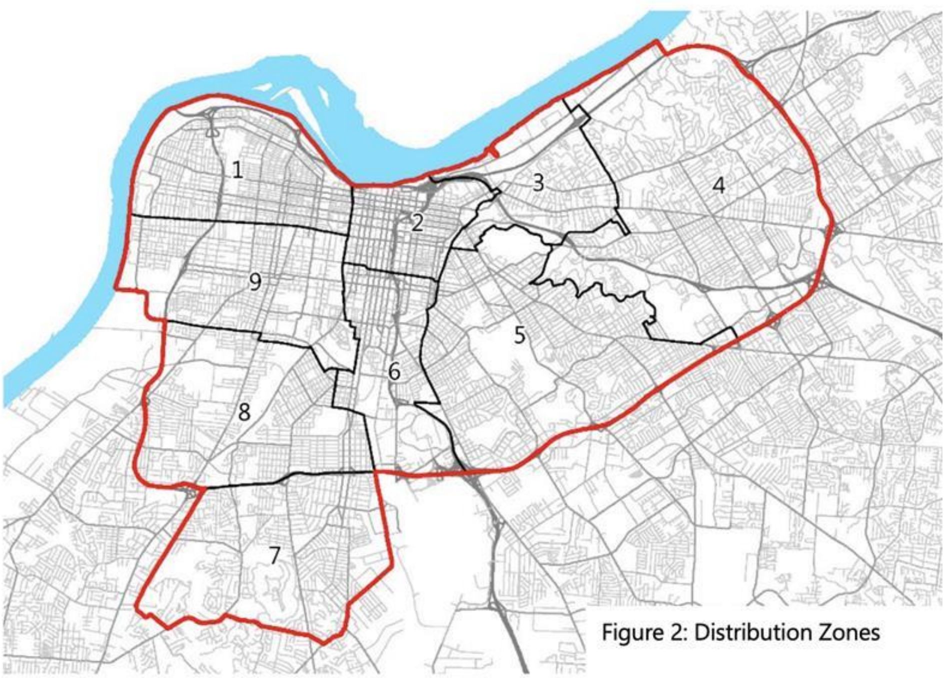


Figure 2: Distribution Zones

Process Result in SAS:

The OPTMODEL Procedure

Problem Summary	
Objective Sense	Minimization
Objective Function	TotalDistance
Objective Type	Nonlinear
Number of Variables	30
Bounded Above	0
Bounded Below	0
Bounded Below and Above	30
Free	0
Fixed	0
Binary	24
Integer	0
Number of Constraints	9
Linear LE (<=)	0
Linear EQ (=)	9
Linear GE (>=)	0
Linear Range	0
Constraint Coefficients	48

References:

[1] https://www.theatlantic.com/technology/archive/2018/05/charging-electric-scooters-is-a-cutthroat-business/560747/

[2] Dataset https://data.louisvilleky.gov/dataset/dockless-vehicles

[3] Dockless vehicle policy - 2020-09

https://data.louisvilleky.gov/sites/default/files/METRO%20DOCKLESS%20VEHICLE%20POLICY_9_14_20.pdf

[4] Programming realization https://automating-gis-processes.github.io/CSC/notebooks/L2/geopandas-basics.html

[5] Python model reference https://python-mip.readthedocs.io/en/latest/examples.html

[6] SAS code reference
https://www.sas.com/content/dam/SAS/en_ca/User%20Group%20Presentations/Edmonton-User-Group/LetsexploreSASPROCOPTMODEL_v2.pdf

[7] SAS model reference
https://support.sas.com/documentation/onlinedoc/or/141/optmodel.pdf