E-SCOOTER SHARE PILOT IN CHICAGO

Ziyi Yang

CPLN 680- Advanced Topics in GIS

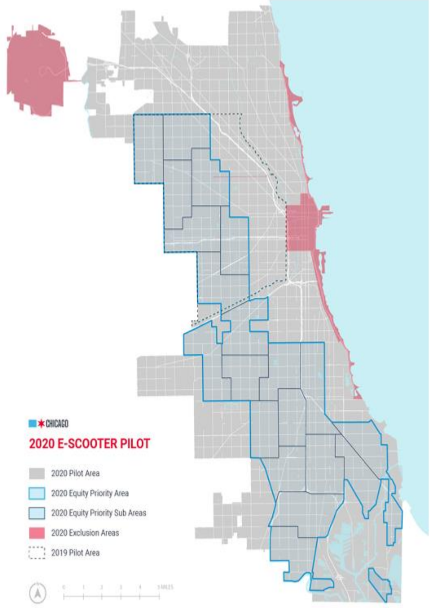
- ## Introduction:

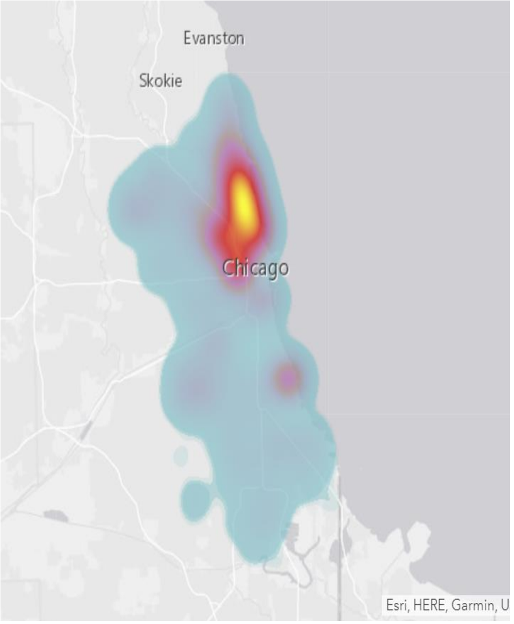
Shared e-scooter services, such as Lime and Bird, first appeared in the US in 2017. Since then, it has been introduced in many cities around the world to reduce traffic congestion and air pollution and enhance community relations as well as the resilience of urban transport, which are important assets for sustainable urban transport. In addition, the use of shared electric scooters as an alternative to public transport for short trips is increasing during the COVID-19 pandemic. E-scooters are often "unparked" compared to shared bikes, meaning they don't have fixed racks but drop off and pick up from certain locations in the service area. Scooter sharing system aims to provide a fast and convenient means of transportation for the public.

- ## Motivation:

In Chicago, people are constrained by a lack of transportation options other than bicycles. However, there are clear ways to meet the city's need for improved transportation and ways to ensure that every Chicagoan benefits. Job one for expanding transportation choices is shoring up current transit system, and then helping the communities connect to transit stops. Therefore, we need to improve local, block-by-block transit access for folks who are most in need or who are less likely to live within walking distance of a Reach bus or train option. Every neighbor deserves the same access to our entire city -- a neighborhood network. Shared e-scooters can provide that missing link, giving people the means to access transit stops that were previously accessible only by a long walk or expensive car ride. It is importance to know whether there is an imbalance in demand for Scooter and whether it will give residents an unfair right to use the scooter.

Who currently has access to E-shared scooters?





In 2019 and 2020, the city of Chicago conducted a four-month scooter sharing pilot to evaluate the feasibility of Chicago scooter sharing, test technology and rules, and better understand how to form a more durable program. Although these vehicles have been rapidly integrated into everyday urban life during the experimental phase, it seems that the adoption of electric scooters in different parts of the city is not equal.

If users don't have a scooter available within walking distance when they want to use one, they're likely to give up looking, and sharing a scooter is useless. In order to improve service levels, it is worth thinking about how we should properly and efficiently redistribute scooters to ensure they are available for everyone. It is also important to predict the demand for their use within a given area. whether there is an imbalance in scooter demand, which leads to unfair scooter use rights for residents. On the map, most travel takes place in the northeast of the city. If suppliers do not actively rebalance, vehicles will be concentrated in a few areas and users in other areas will not be able to have scooters in time. To encourage more people to use shared scooters, Lime has launched pioneer areas, where the price of scooters is much lower than other areas. We need to understand how to determine this region and whether its existence has resulted in a single region being deliberately over-served while others are under-served.

Using the 2020 scooter rider data available on Chicago's Open Data, machine learning methods were used to create a model that predicted the total number of scooter trips per census tract in the city during the second quarter of this year. The model uses features including the city's demographic, socio-economic and built environment characteristics to make predictions. The model forecasts reflect both the potential demand for scooters in census tract and the impact of the scooter companies' fleet management and distribution choices. The model used existing passenger data to predict how scooter usage would have been if a business-as-usual approach had been followed.

Literature Review:

Kim's article studied a demand prediction model for the use and development of shared electric scooters in Seoul, South Korea, in October 2020. The space unit of analysis is set as a grid of 200 square meters, and the hourly demand of each grid is summarized according to the electric scooter travel data. The community structure approach was used to cluster the spatial areas into five communities before predicting the needs. They built a demand prediction model based on long and short-term memory (LSTM) and compared the prediction results based on activation function.

Ursaki et al. compared the social and economic characteristics of U.S. Census Bureau block groups within and outside bike-sharing service areas in seven cities based on the American Community Survey. By creating a 500-meter buffer zone around each site in ArcGIS, the location of the bike-sharing site was used to define the bike-sharing service area. Using the student T-test to compare socioeconomic characteristics within and outside bike-sharing service areas, significant differences in visits based on race and income variables were found in Boston, Chicago, Denver, Seattle, and New York City.

Works Cited:

Kim S, Choo S, Lee G, Kim S. Predicting Demand for Shared E-Scooter Using Community Structure and Deep Learning Method. *Sustainability*. 2022; 14(5):2564. https://doi.org/10.3390/su14052564

Ursaki, Julia, Aultman-Hall Lisa. Quantifying the Equity of Bikeshare Access in U.S. Cities. TRC Report 15-011. 2015. https://rosap.ntl.bts.gov/view/dot/36739

The Pilot program in Chicago only runs for four months, from August to December, not a whole year dataset. As I chose September here, when the weather is not color or hot and scooter use peaks.
Census tract were chosen as units of spatial analysis because they represent the highest level of geographic aggregation in the scooter passenger data set.

- **Demographic**

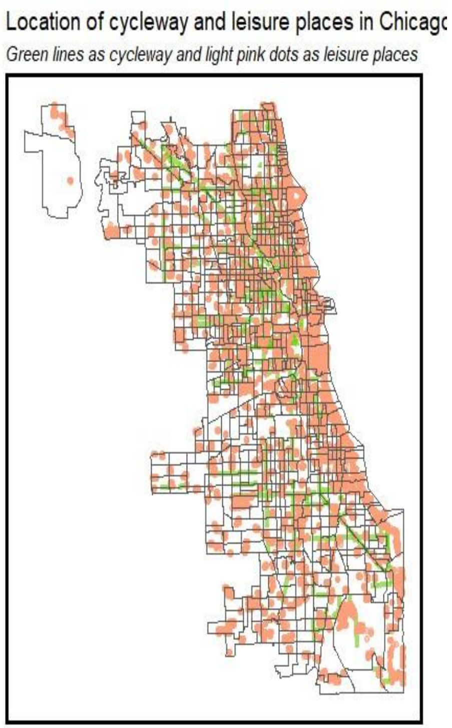
Total Population
Median Age
Percentage White Population
Percentage Female Population

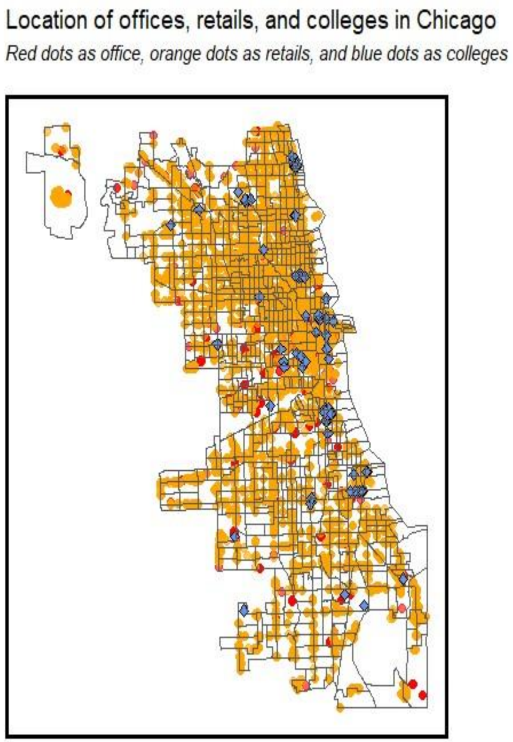
- **Socio-economic**

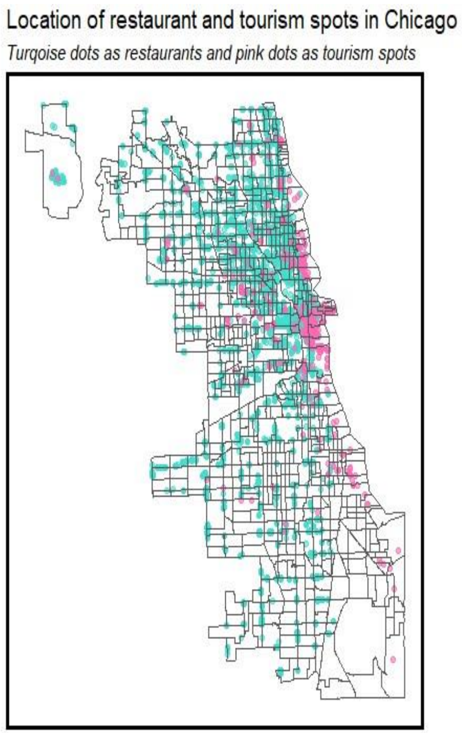
Household Income
Home Values and Rental Prices
Commute Modeshare (transit v driving)
Commute Distance (30+ minutes)
Housing Units and Occupancy Rates
Vehicle Ownership

- **Built Environment**

Retail Stores
Restaurants
Leisure Activities and Tourism Destinations
Transportation Infrastructure
Offices
College







Below, is a correlation matrix with our features. We see that some variables, median rent and median home value, median household income, are strong correlated, also for the most part, our explanatory variables show collinearity.

- *Built Environment:*

Density_college	% of the city's college in this tract
Density_restaurant	% of the city's restaurant in this tract
Density_pubtran	% of the city's public transportation in this tract
Density_tourism	% of the city's tourist attractions in this tract
Density_leisure	% of the city's places for leisure activity in this tract
Density_office	% of the city's offices in this tract
Density_retail	% of the city's retail in this tract

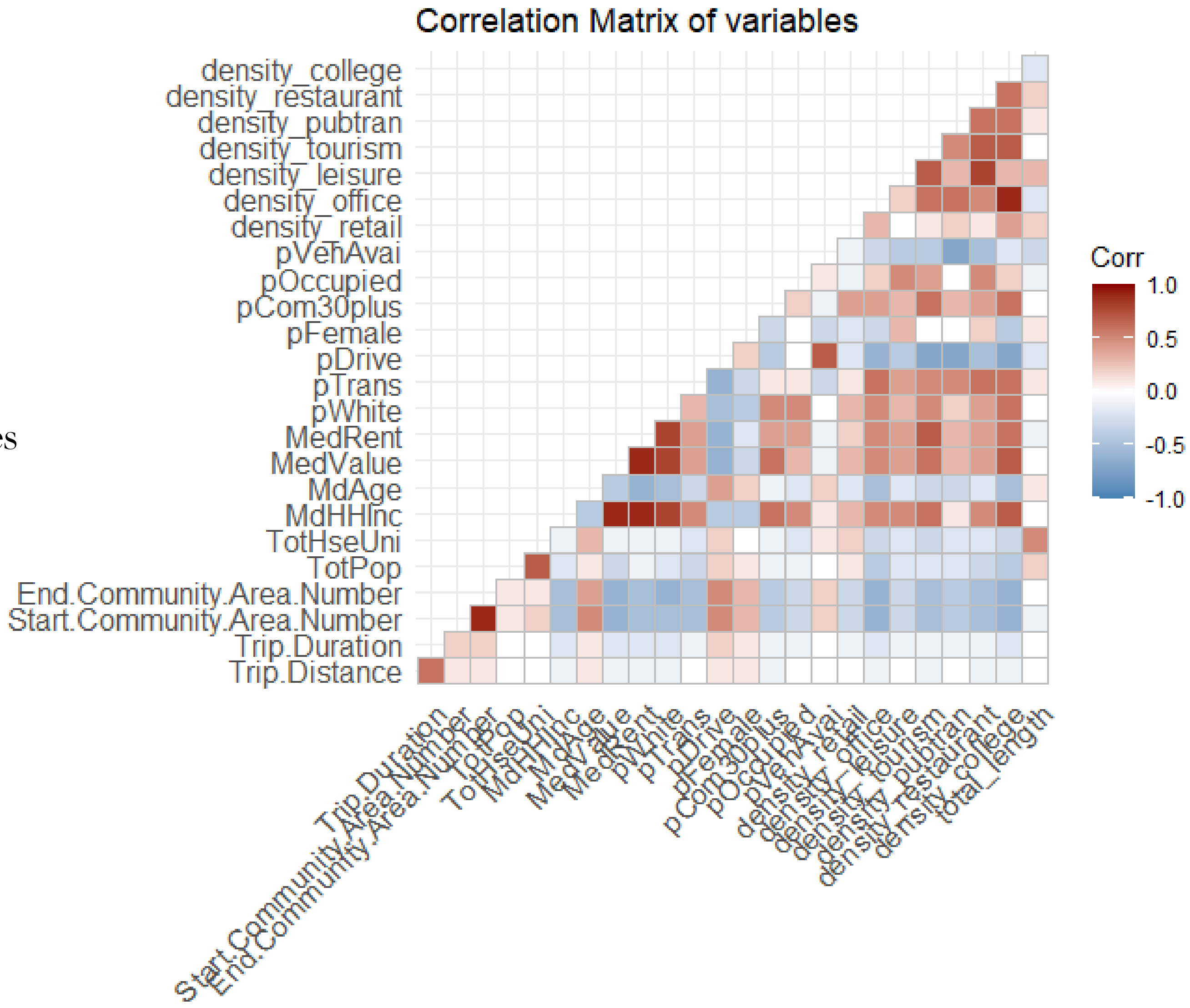
- *Socio-Economic:*

pVehAvai	% of the population that owns a vehicle
pOccupied	Housing occupancy rate
pCom30plus	% of the population with a commute >30 minutes
pDrive	% of the population that drives to work
pTrans	% of the population that takes transit to work
MedRent	Median rent
MedValue	Median Home Value
MdHHInc	Median household income
TotHseUni	Total Housing Units

- *Demographic:*

pWhite	% of the white population
pFemale	% of the female population
MdAge	Median Age
TotPop	Total Population

End.Community.Area.Number
Start.Community.Area.Number
Trip.Duration
Trip.Distance



Future works:

- Run Model --

Using the final set of features below to construct several models that predict raw trip counts in each census tract for the city in this study.

1. Linear Model: an OLS linear regression. Linear models are a way of describing a response variable in terms of a linear combination of predictor variables.

2. A random forest or random decision forest is an ensemble learning method for classification, regression, and other tasks that operates by building a large number of decision trees at training time. Random decision forests correct the habit of decision trees to overfit their training set.

3. Another model?

- Prediction
- Priority Area/Rebalance

split the data into 75%/25% train and test sets
use random k-fold cross-validation
test models on the 25% test set and select the framework with a lower MAE

Filter Cols: « ‹ 1 - 50 › »

	Trip.ID	Start.Time	End.Time	Trip.Distance	Trip.Duration	Vendor	Start.Community.Area.Number	End.Community.Area.Nu
1	265a6159-201a-4271-9737-46fd4d323462	2020-09-01 00:00:00	2020-09-01 00:00:00	3411	1140	spin	66	
2	febd7482-ca04-4f6a-9717-b5623e4a6d82	2020-09-01 00:00:00	2020-09-01 00:00:00	2096	946	spin	66	
3	79f1c1a0-c753-4ebb-8f0c-bdfefbefe884	2020-09-01 00:00:00	2020-09-01 00:00:00	4162	1966	spin	69	
4	923659bb-7205-4383-b123-856b0d873a01	2020-09-01 00:00:00	2020-09-01 00:00:00	4261	2050	spin	69	
5	2f624202-9da8-46f6-ac97-1863858d4382	2020-09-01 00:00:00	2020-09-01 00:00:00	14	25	spin	66	
6	4a12f06e-172b-4fe0-8c1a-8075c77dcfc3	2020-09-01 00:00:00	2020-09-01 00:00:00	14	36	spin	67	
7	f2d9d785-e4f0-40be-9e18-6bb91d720a04	2020-09-01 01:00:00	2020-09-01 02:00:00	7010	2423	spin	69	
8	a6af0ab2-333c-4952-81f8-6e61af776093	2020-09-01 01:00:00	2020-09-01 02:00:00	6520	2535	spin	69	
9	1a8f23b8-ad3a-430d-a19a-50056c9ee64c	2020-09-01 04:00:00	2020-09-01 04:00:00	727	592	spin	48	
10	7b6accb3-7b75-47f3-81eb-a4bb4e4691ef	2020-09-01 04:00:00	2020-09-01 04:00:00	6519	1402	spin	30	
11	ecab6238-8e1b-493e-aee0-331b5979f9c0	2020-09-01 05:00:00	2020-09-01 05:00:00	853	163	bird	21	
12	8e88c45e-bce1-4ed8-b1f9-20dad4c66bc3	2020-09-01 05:00:00	2020-09-01 06:00:00	2386	5561	lime	6	
13	5bc60e16-0842-40e7-b8ed-69c78f2374ae	2020-09-01 05:00:00	2020-09-01 05:00:00	3685	1463	lime	28	
14	cefa4ef1-263b-4944-969f-03dbefc55554	2020-09-01 05:00:00	2020-09-01 05:00:00	4685	718	bird	6	
15	c4e6c2f5-bde7-4459-bd98-1f7d6a59ae89	2020-09-01 05:00:00	2020-09-01 05:00:00	1556	290	bird	24	
16	5a3cd626-4519-4023-9d86-fd72584078cf	2020-09-01 05:00:00	2020-09-01 05:00:00	721	186	bird	16	
17	63c5e04f-3bb5-4cb3-9c7d-e72dab5ae09c	2020-09-01 05:00:00	2020-09-01 05:00:00	48	68	bird	19	
18	a23e1ab0-4a63-4544-9e7c-177ee4d9ffc3	2020-09-01 05:00:00	2020-09-01 05:00:00	4571	951	bird	19	
19	ff978928-d446-4600-b2db-05f81eb701bd	2020-09-01 05:00:00	2020-09-01 05:00:00	3298	686	bird	68	

Showing 1 to 19 of 232,534 entries, 56 total columns