Usage, equity, and safety of shared electric scooters: a geospatial analysis of trips taken in Minneapolis, Minnesota

Carston Hernke

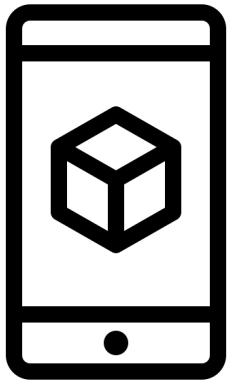
hernko09@umn.edu

UNIVERSITY OF MINNESOTA

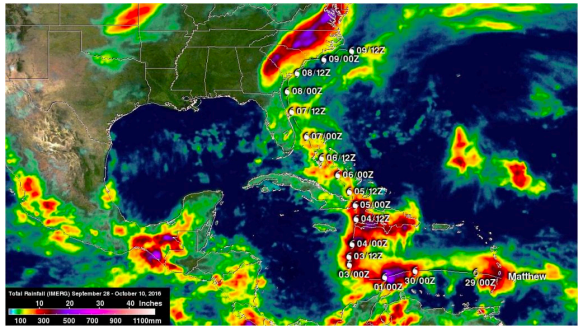
About me

- Recently completed: Bachelor's of Science in Business, University of Minnesota (magna cum laude)
  - Major: Management Information Systems, Minor: Computer Science
- Internship experience in agribusiness, startups, and technology consulting
- Director at the only completely student-run VC fund in the United States (Atland Ventures)

Personal Motivation for Thesis



Technology-enabled business models



Geospatial analysis

Urban design/ smart cities

Data processing pipeline

Scooter Providers

BIRD

Lime

- Data collected from vehicles
- Proprietary data stores

Flat File (.csv)

City of Minneapolis Government

Minneapolis
City of Lakes

- City law requires data sharing from scooter providers (public data)
- Anonymization process bins rides to nearest street centerline and times to nearest 15 minutes

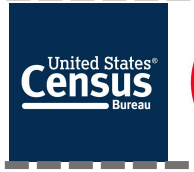
Flat File (.csv)

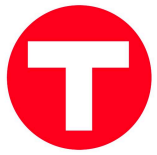
This Research

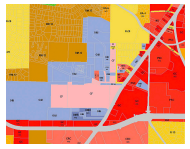
Join with street centerline data → Bin into hex zones → Enrich with other spatial data → Tests for correlation → Model building

Spatial data (.shp)

Secondary Data Sources







- Demographic data (population, income, students)
- Transit location data
- Zoning/ Land Use data

Spatial data (.shp)

Technologies used

The scripting language R was used for the analysis because of its popularity in academia and its straightforward support for geospatial data and visualization

The ‘sf’ library for R implements *simple feature access,* a formal standard for describing how objects in the physical world can be represented in computers (ISO 19125-1:2004)

Flat spatial data files, including *shapefiles (.shp)* were used to store spatial data. A geodatabase was not used because of the limited scale of the data.

Version control software (VCS) was used to track changes to the R files used for analysis as the research progressed. All the code for this thesis can be found online: https://github.com/carstonhernke/dockless-scooter-analysis

Research Questions

- Usage
  - **RQ-A1** What types of trips are users using scooters for?
  - **RQ-A2** Are scooters being used as a compliment to public transit?
  - **RQ-A3** Are scooters being used for commuting to jobs?
- Equity
  - **RQ-B1** Is there a relationship between the household income in an area and the number of scooter rides from that area?
- Safety
  - **RQ-C1** Is there a relationship between the presence of enhanced bike infrastructure and per-capita scooter ride originations?

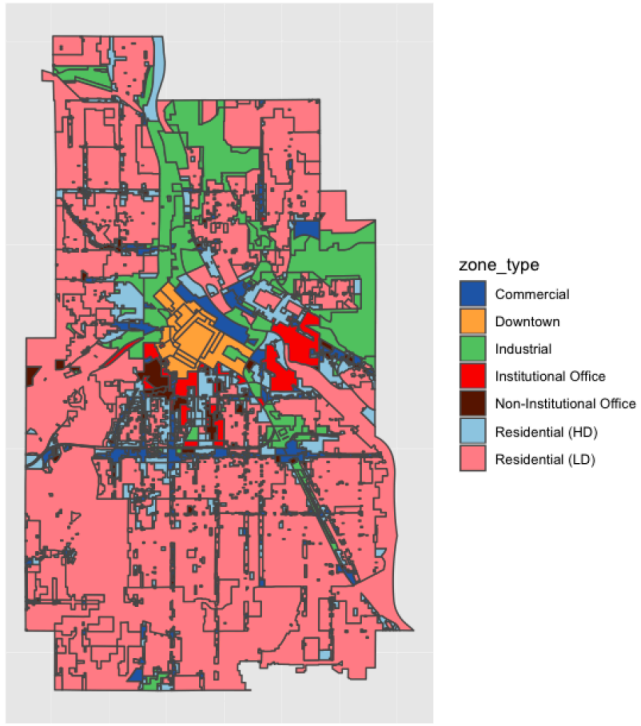
Usage

RQ-A1 What types of trips are users using scooters for?

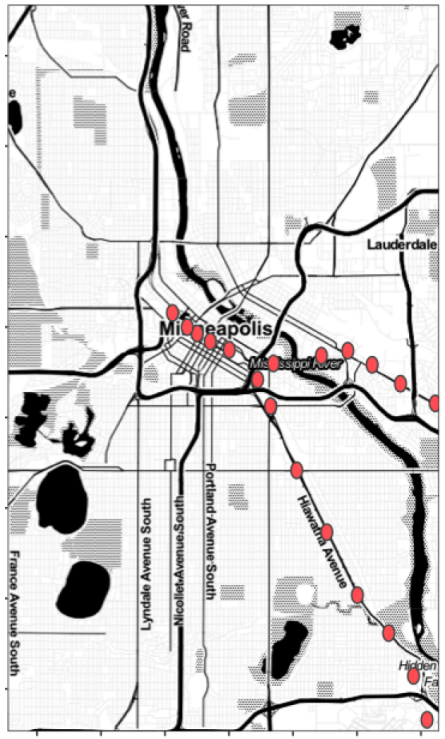
RQ-A2 Are scooters being used as a compliment to public transit?

RQ-A3 Are scooters being used for commuting to jobs?

Usage Methodology



Zoning in Minneapolis



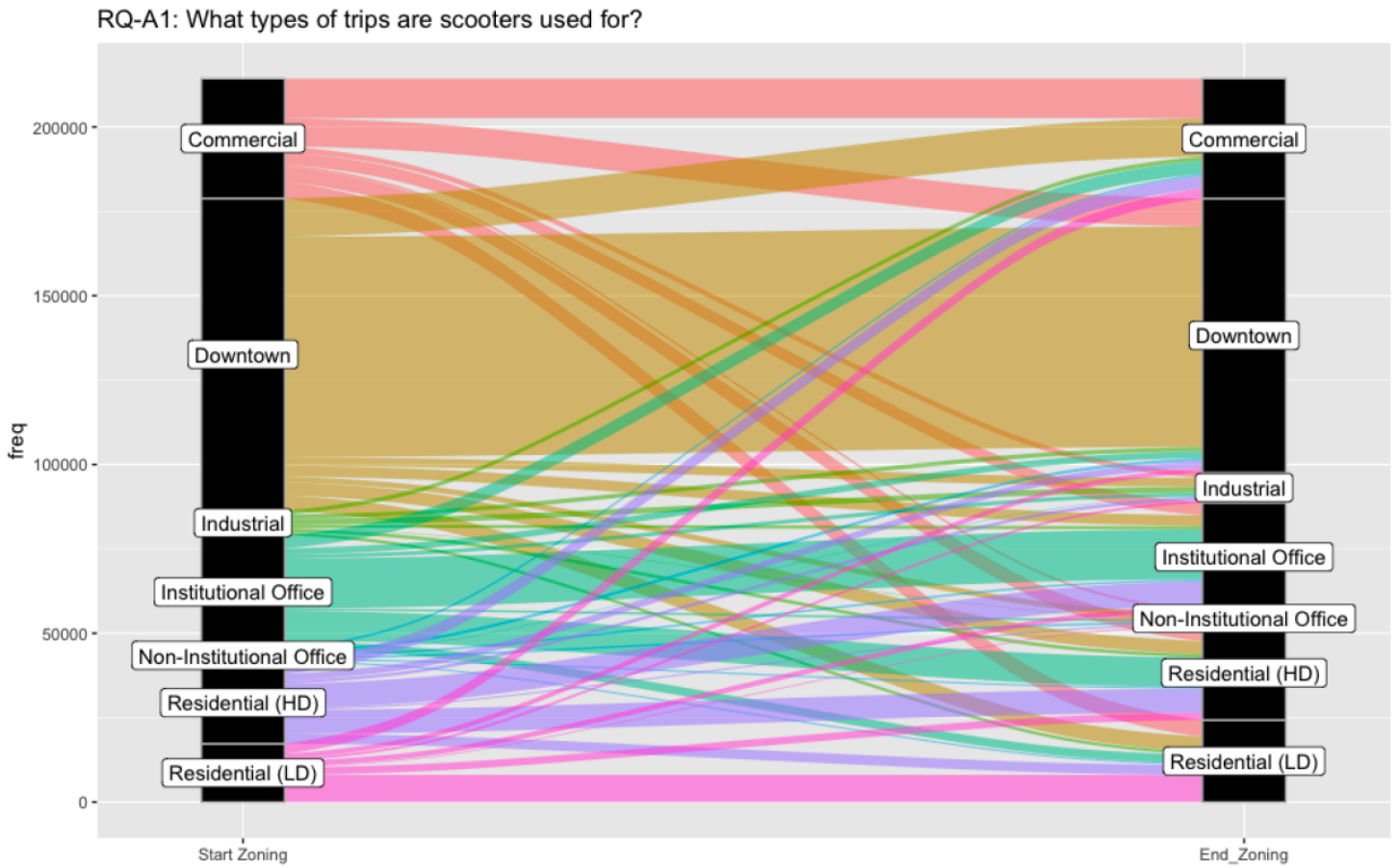
Transit (Light Rail) Stations in Minneapolis

Using zoning data, each trip was enriched with an origin type and destination type

Using transit station location data, each trip was tagged with its proximity to a light rail station

This enriched trip data could then be used to answer the research questions

RQ-A1 What types of trips are users using scooters for?



Most rides were taken between downtown destinations and between institutional office locations, including college campuses. There is a limited amount of inter-area travel

RQ-A2 Are scooters being used as a compliment to public transit?

Transit Usage Category	Frequency	Proportion
1. No Transit	123301	0.80
2. Start Near Transit	14560	0.09
3. End Near Transit	12957	0.08
4. Start and End Near Transit	3341	0.02

Proportion of trips starting or ending near transit

19% of scooter rides had some relationship to the light rail network in Minneapolis, suggesting mode-sharing between scooters and trains

RQ-A3 Are scooters being used for commuting to jobs?

Start Zone	End Zone	Derived Commute Type
Residential (LD) OR Residential (HD)	Downtown OR Industrial OR Institutional Office OR Non-Institutional Office	Journey to Work
Downtown OR Industrial OR Institutional Office OR Non-Institutional Office	Residential (LD) OR Residential (HD)	Journey from Work
All Others	All Others	FALSE

Table 5: Logic used to determine whether a trip is a commute

Day Type	Non Commute Proportion	Commute Proportion
Weekend	0.84	0.16
Weekday	0.82	0.18

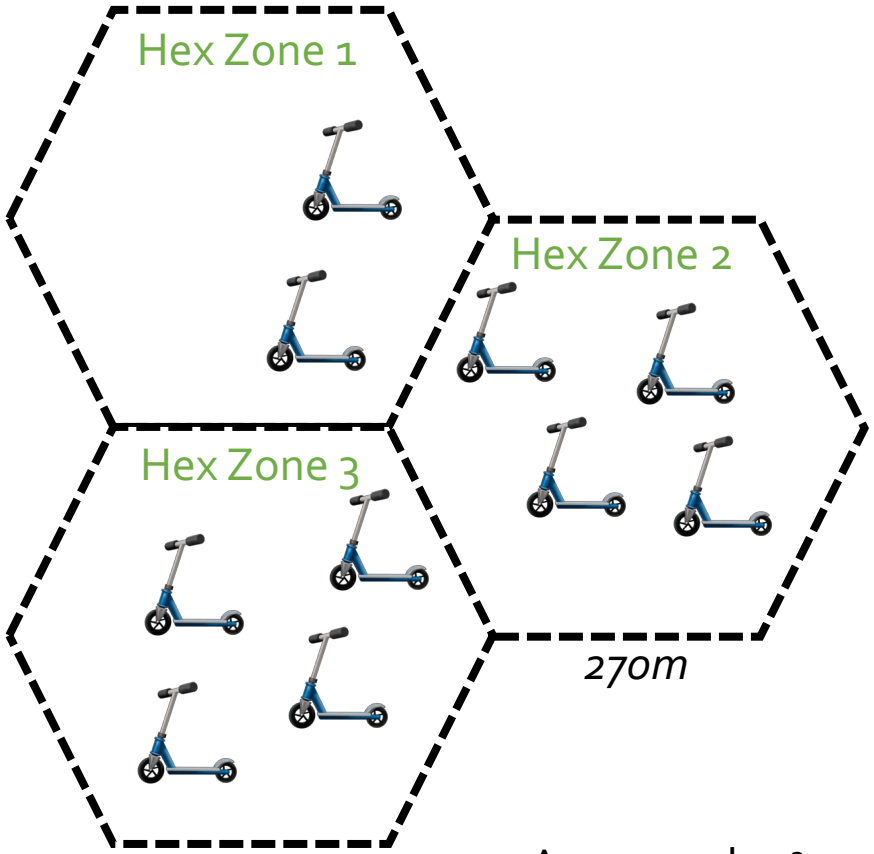
Table 6: Proportion of commuting trips, by type of day

Scooters are not commonly used for commuting, as shown by the very limited 2% increase in commute-type trips on weekdays vs. weekends

Equity

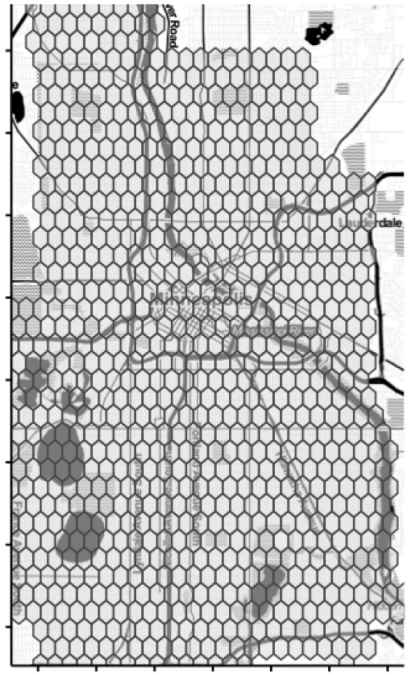
RQ-B1 Is there a relationship between the household income in an area and the number of scooter rides from that area?

Hex Zone Methodology

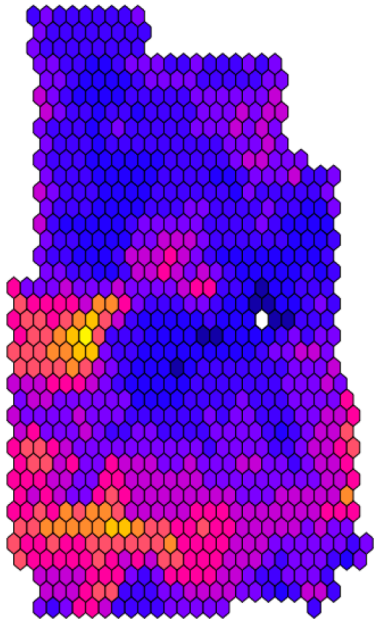


- Discretize the continuous spatial data using hexagonal zones
  - Hex zones are superior to legacy TAZ (Transportation Analysis Zones) for micromobility
  - Represent a realistic walking distance
  - 854 hex zones in Minneapolis
- Hex zones allow for joining other spatial data
  - Used in this research: Population, Number of Students, Household Income

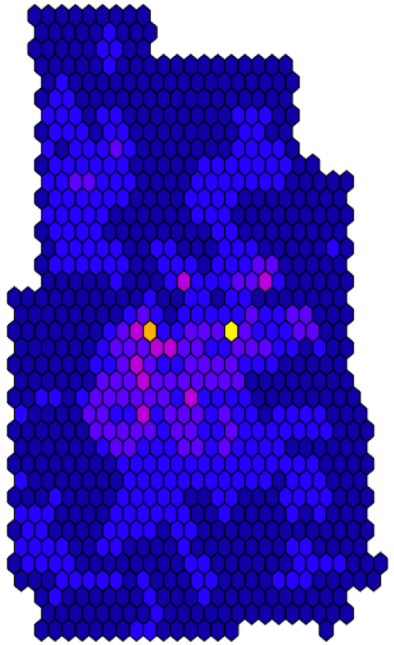
Demographics by Hex Zone



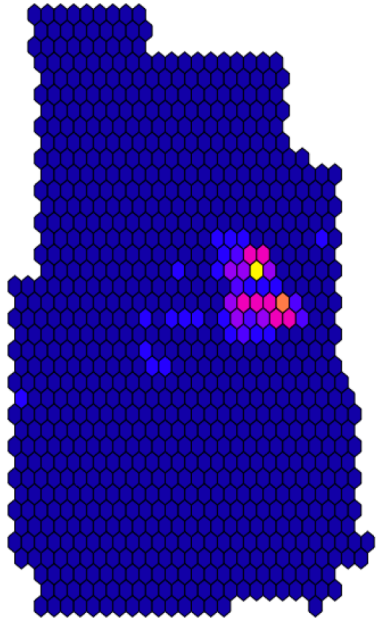
Hex Zones on Map



Average Income



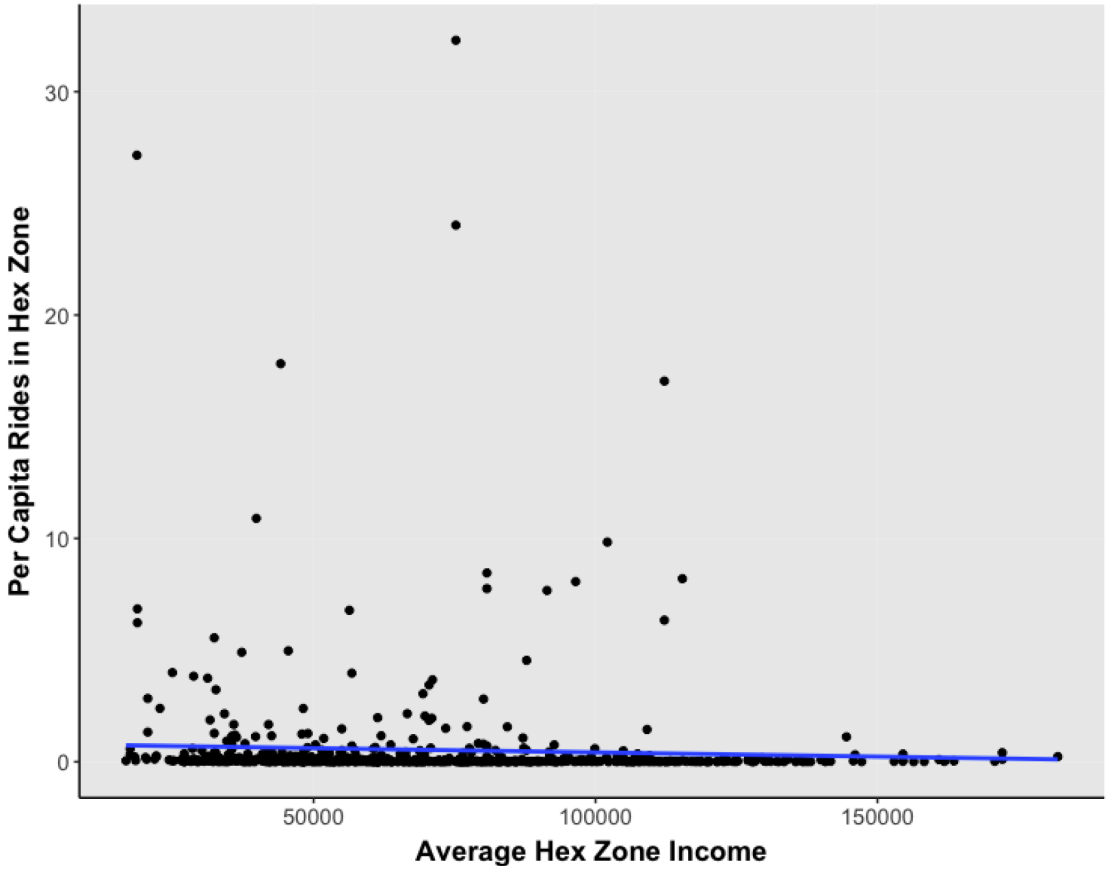
Population Density



Student Population

** Demographic data is not collected in hex zones, instead it was aggregated to these areas using extensive areal interpolation*

RQ-B1 Is there a relationship between the household income in an area and the number of scooter rides from that area?



Scooter trips are slightly more likely to begin in lower-income areas of the city. There is not evidence of widespread inequities in scooter availability.

Safety

RQ-C1 Is there a relationship between the presence of enhanced bike infrastructure and per-capita scooter ride originations?

Safety Methodology

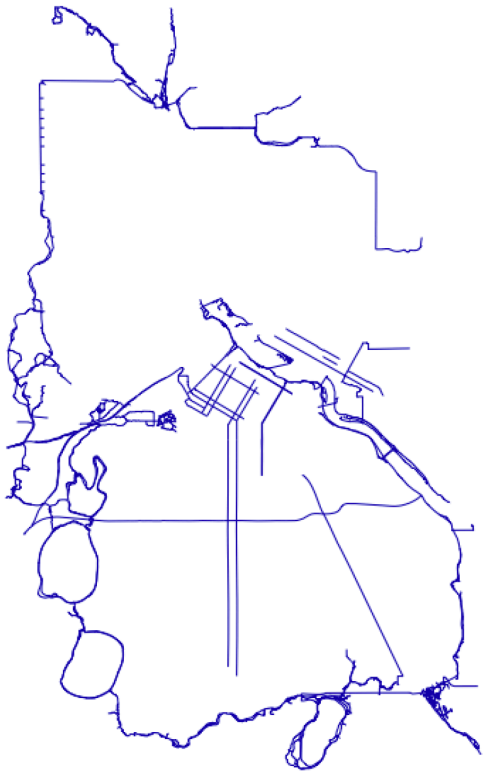


Figure 6: Bike infrastructure in Minneapolis

- The amount of enhanced bike infrastructure (protected lanes, off-street paths/trails, etc.) was calculated for each hex zone
- The relationship between kilometers of enhanced infrastructure and the number of scooter ride originations was examined

There is a very weak relationship between the amount of enhanced bike infrastructure within a hex zone and the average daily number of scooter rides in the zone.

for every additional kilometer of enhanced bike infrastructure per square kilometer, the average daily number of scooter rides increases by 1.489

Beyond my thesis research

How can the technologies and methods used in my thesis be extended to improve scooter companies?

Unit economics are an open question for scooter companies

"Right now, the unit economics don't add up, and with so crowded a field, consolidation is inevitable. So is greater regulation."

BCG

"2019 May Bring The Scootergeddon"

crunchbase news

- My take: because deployment and overnight charging are all fixed expenses, **scooter companies should focus on maximizing the utilization of their fleet**
- Based on an internal Atland Ventures analysis, the breakeven point for a scooter to be profitable is 6+ rides per day
- Scooter companies can increase their utilization and move towards profitability by **using data to place scooters where they will be ridden most**

Using machine learning and geospatial data to deploy scooters to optimize profitability & unit economics at TIER

Big idea: scooter companies are struggling to make the unit economics profitable. High usage (6+ rides per day) on each scooter is necessary for profitability. Only deploying scooters in places that will be profitable can help solve this problem.

Expressed demand data

Scooter trip data

Latent demand data

App-open data

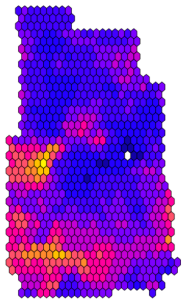
Other transit demand data

Tertiary data

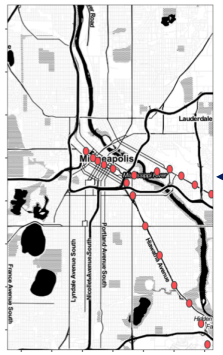
Weather

Special Events

Aggregate demand model for next day *(unique to each city)*



Proposed scooter deployment locations



Predicted trips (based on historic trip data)

Projected profitability for each scooter

Iteratively adjust scooter placement to optimize for projected profitability

Potential optimization algorithms: gradient descent, simulated annealing, genetic algorithms, CSP solvers

Scooter ID	Trips	Revenue
ABC1	4	$10.10
ABC2	2	$3.53
ABC3	10	$26.39

After optimization is finished

Final scooter deployment plan

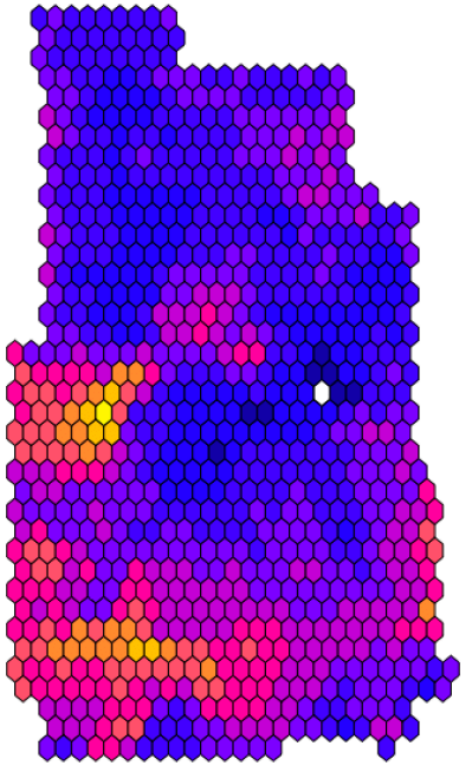
Plan is communicated to chargers via mobile app

Scooters are deployed

Usage data is continually collected, further improving the system

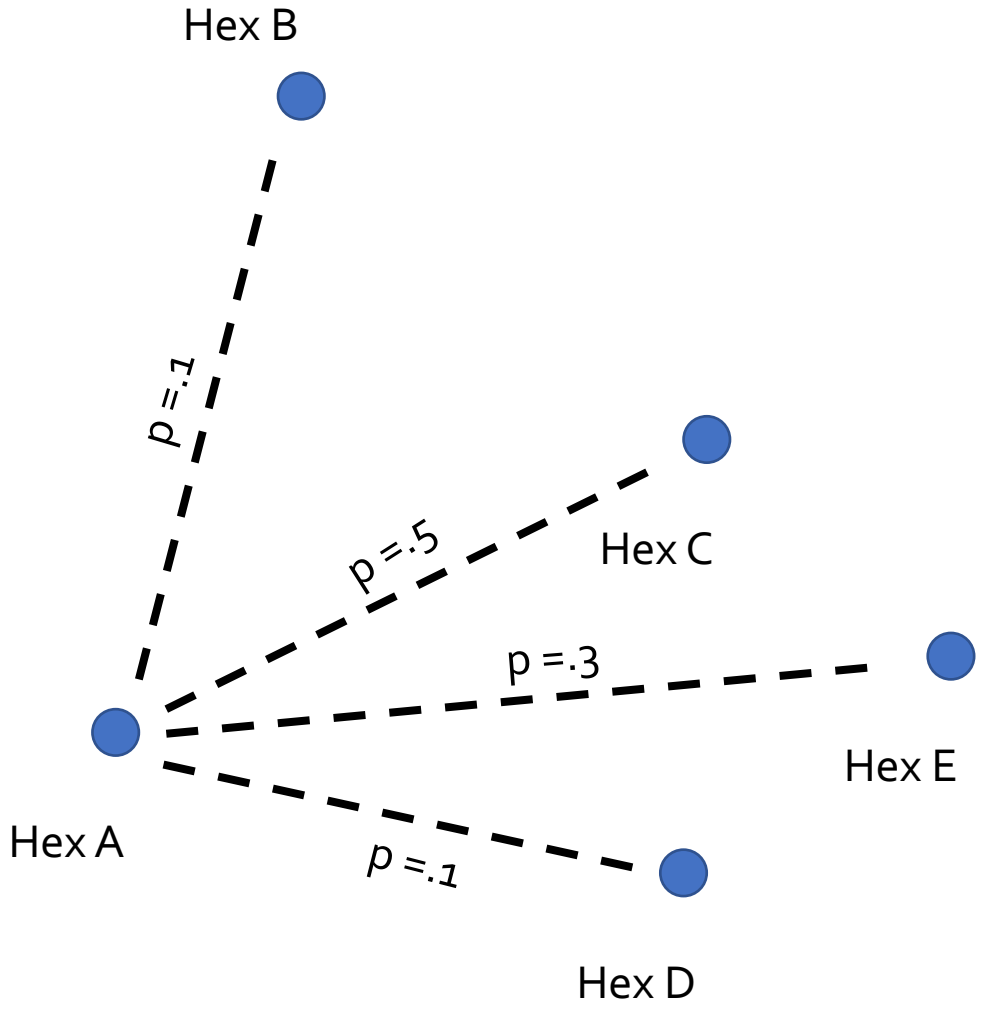
Modeling how scooter placement affects revenue generated

Assumption 1: There is a differing amount of demand for scooters in different areas of the city. If we could determine this exactly, we could represent it as a scalar field.



Modeling how scooter placement affects revenue generated

Assumption 2: for each location in the city, we can use previously collected data to build a probability distribution of scooter destinations



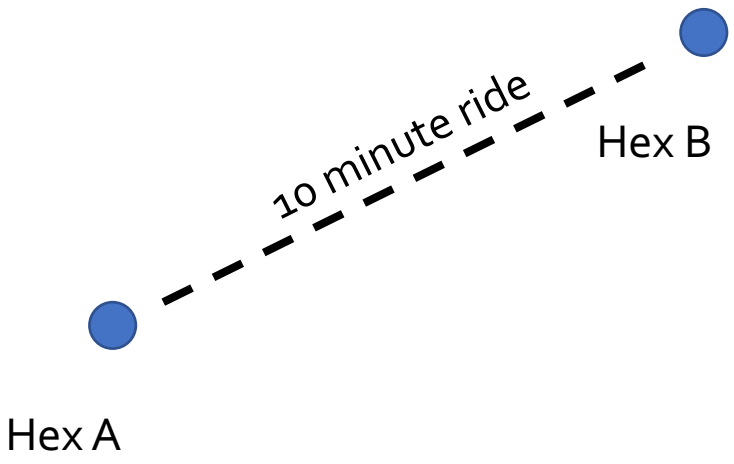
Modeling how scooter placement affects revenue generated

The revenue generated by a single scooter trip is straightforward to calculate:

Revenue = Flat Cost + (Duration * Variable Cost)

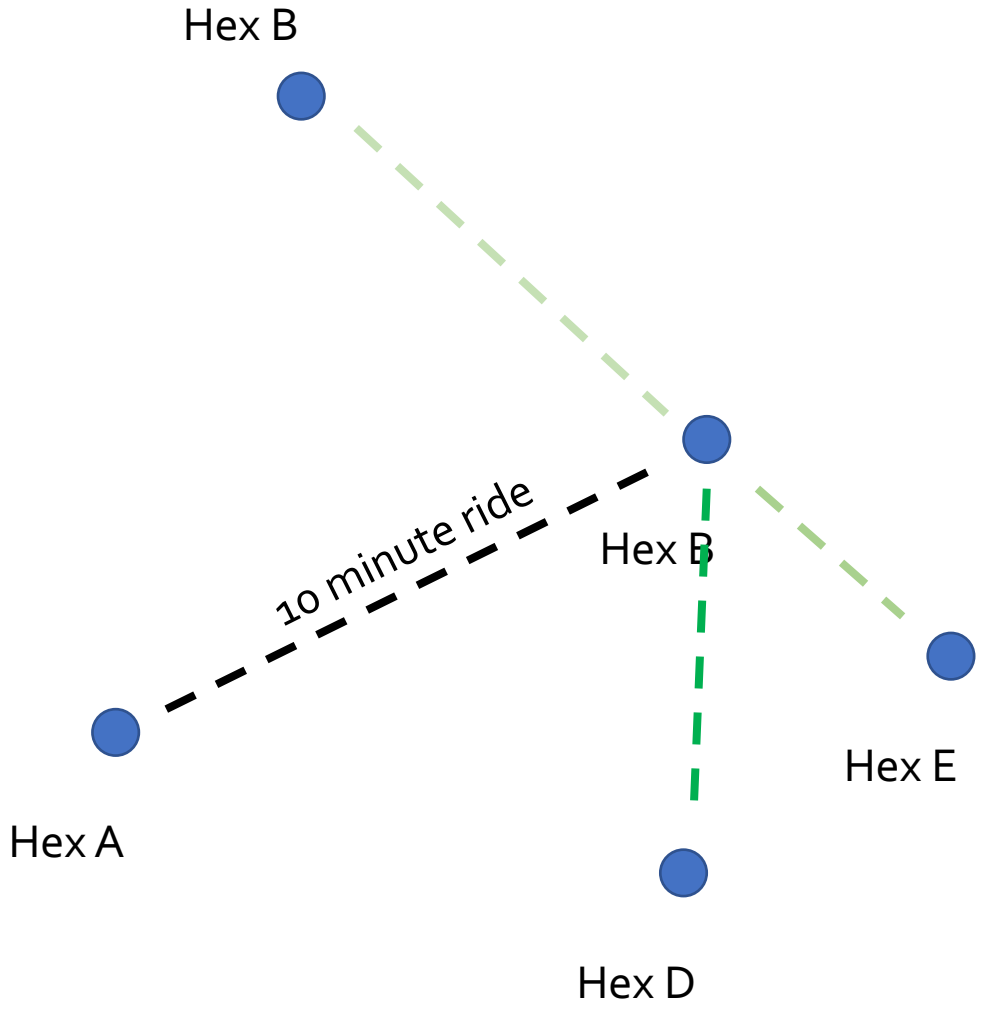
$Revenue_{A->B}$ = $1 + (.15 * 10)

$Revenue_{A->B}$ = $2.50



Modeling how scooter placement affects revenue generated

However, finding the simple value of the trip between A and C **doesn't capture the whole story**. After the trip, the scooter is now located at Location C, which **changes its revenue potential for the rest of the day.**

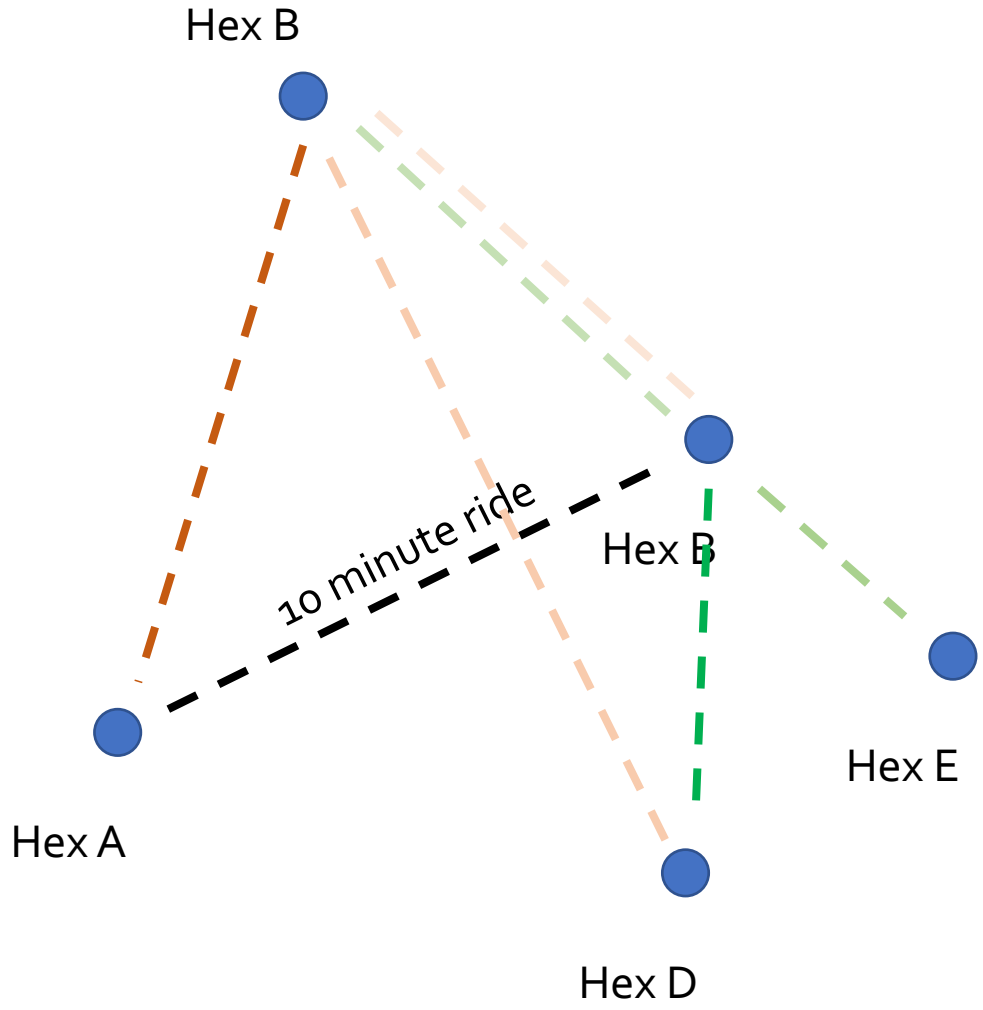


Modeling how scooter placement affects revenue generated

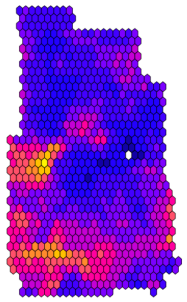
Proposed solution: **recursively explore all potential combinations of routes, using the trip length and probability to calculate an expected value of the initial placement.**

Pseudocode:

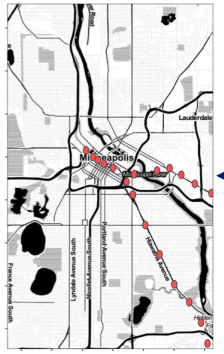
```
function expected_revenue(scooter_origin):
    probability_vector = get_vector(scooter_origin)
    revenue = 0 for i in probability_vector where p > .01:
        expected_trip_revenue = length * probability
        revenue = expected_trip_revenue + expected_revenue(next_origin)
    return revenue
```



Once we have a method for finding the value of placing a scooter in a certain location, we can use common optimization algorithms to find the placement of scooters that maximized expected revenue *(taking interactions into account)*



Aggregate demand model for next day *(unique to each city)*



Proposed scooter deployment locations

Predicted trips (based on historic trip data)

Projected profitability for each scooter

Iteratively adjust scooter placement to optimize for projected profitability

Potential optimization algorithms: gradient descent, simulated annealing, genetic algorithms, CSP solvers

Scooter ID	Trips	Revenue
ABC1	4	$10.10
ABC2	2	$3.53
ABC3	10	$26.39

Questions?

Feel free to email me with additional questions: hernk009@umn.edu