Source

Electric Micromobility: Scooting Towards Equitable, Carbon-Free Transportation

Samantha Buck, MClinEpi, MLS(ASCP)
Galvanize Denver, 2020

Sam Buck

Ballerina
Molecular Microbiologist
Epidemiologist
Data Scientist
Scooter Enthusiast

Source

Chicago Scooter Pilot

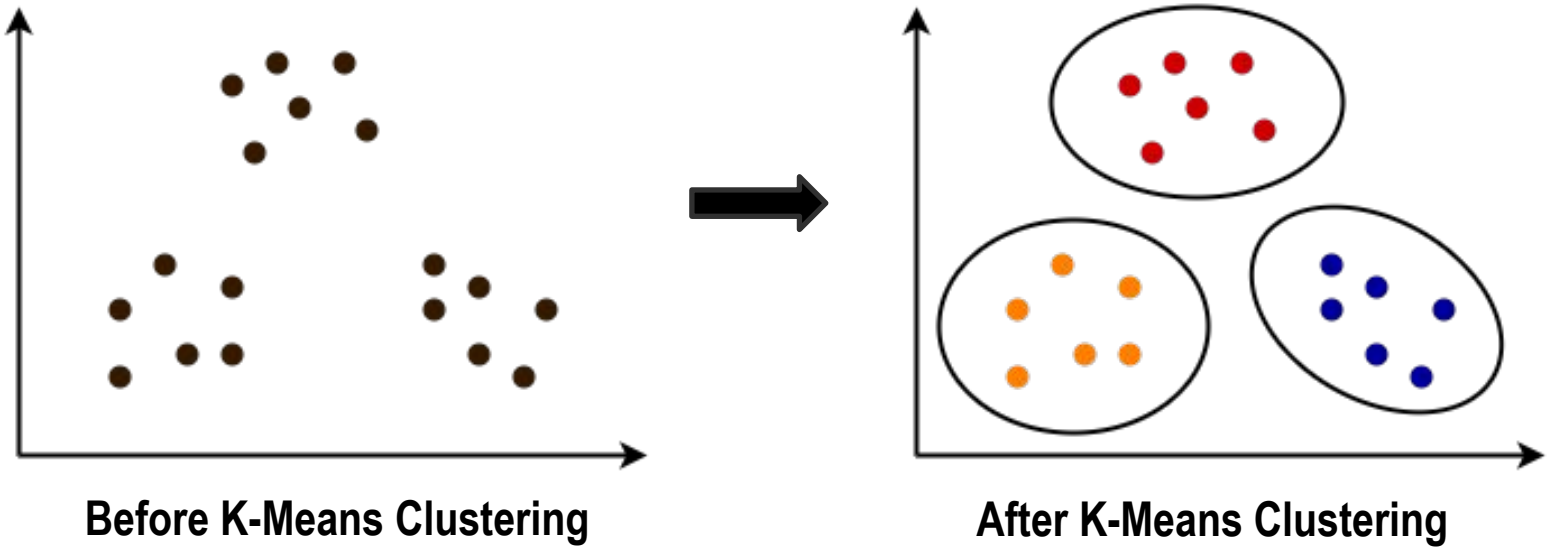
- **Goal:** How can e-scooter companies increase their profit?
- **Data:** Four month e-scooter pilot (June-October 2019); over 650,000 trips recorded
- **Why:** Sustainable + equitable transportation solution

Methodology

- **Goal:** How can e-scooter companies increase their profit?
- **Method:**
  - Apply an inferential model on the data
  - Do certain factors relate to higher profit (more rides)
  - How do these factors relate to each other
- **Model:** Unsupervised Machine Learning - Clustering

Google Slides Image

K-Means Clustering: Unsupervised Learning



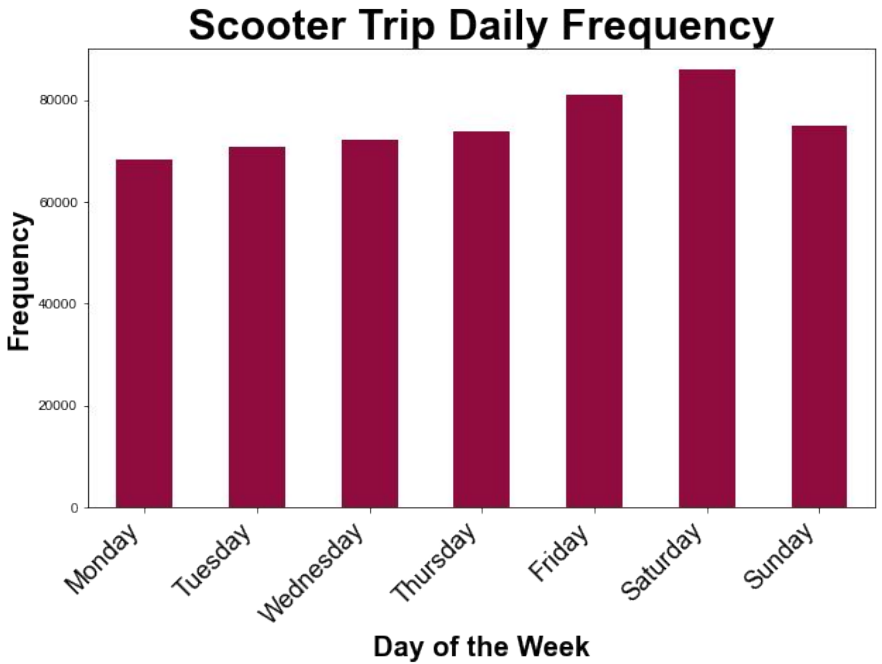
Before K-Means Clustering

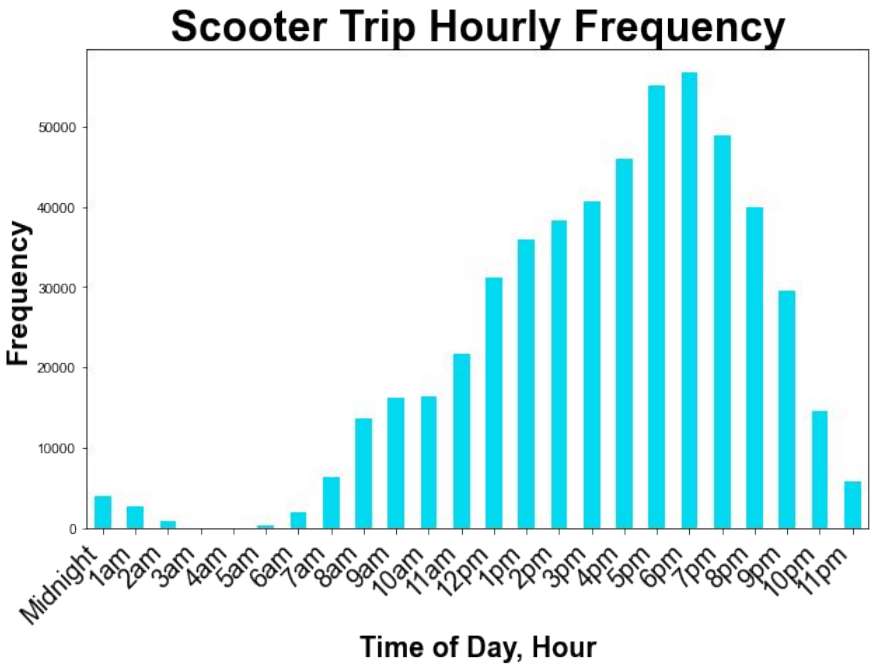
After K-Means Clustering

Discover underlying patterns by grouping similar data points together

Source:https://www.gatevidyalay.com/k-means-clustering-algorithm-example/

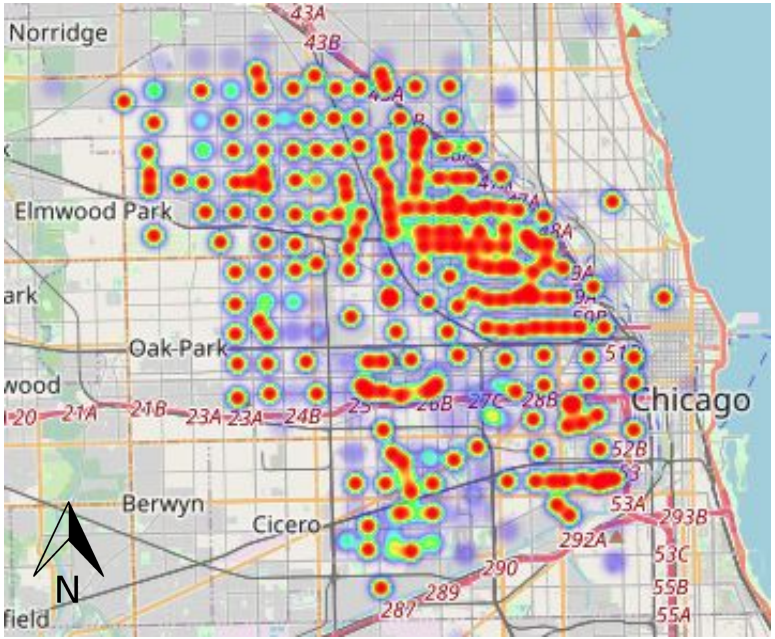
Data + Feature Analysis



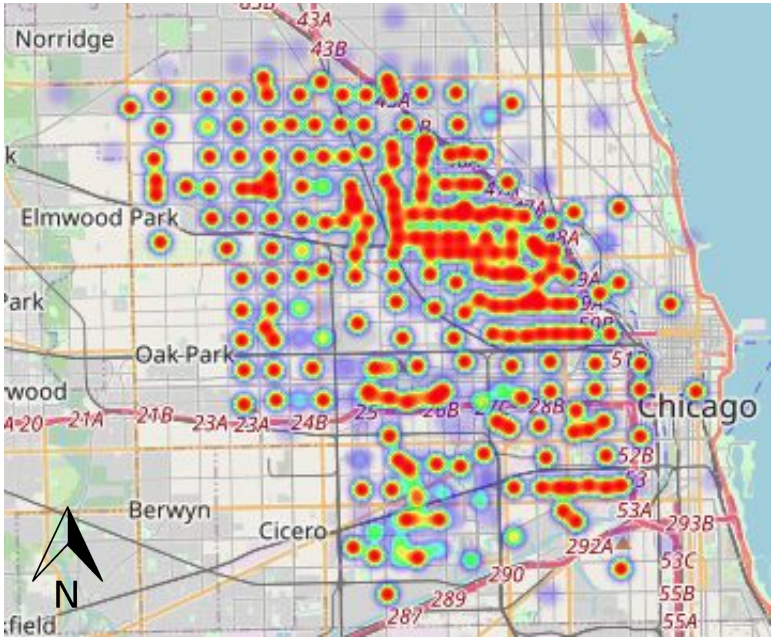


Data + Feature Analysis

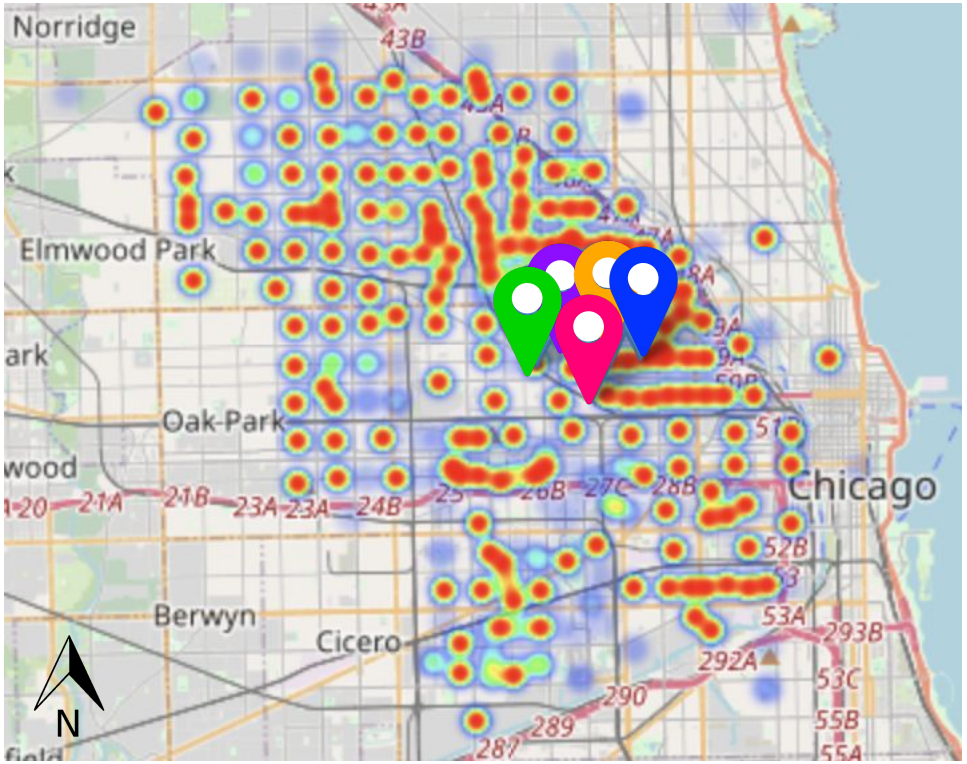
Scooter Ride Origin Heat Map



Scooter Ride Destination Heat Map

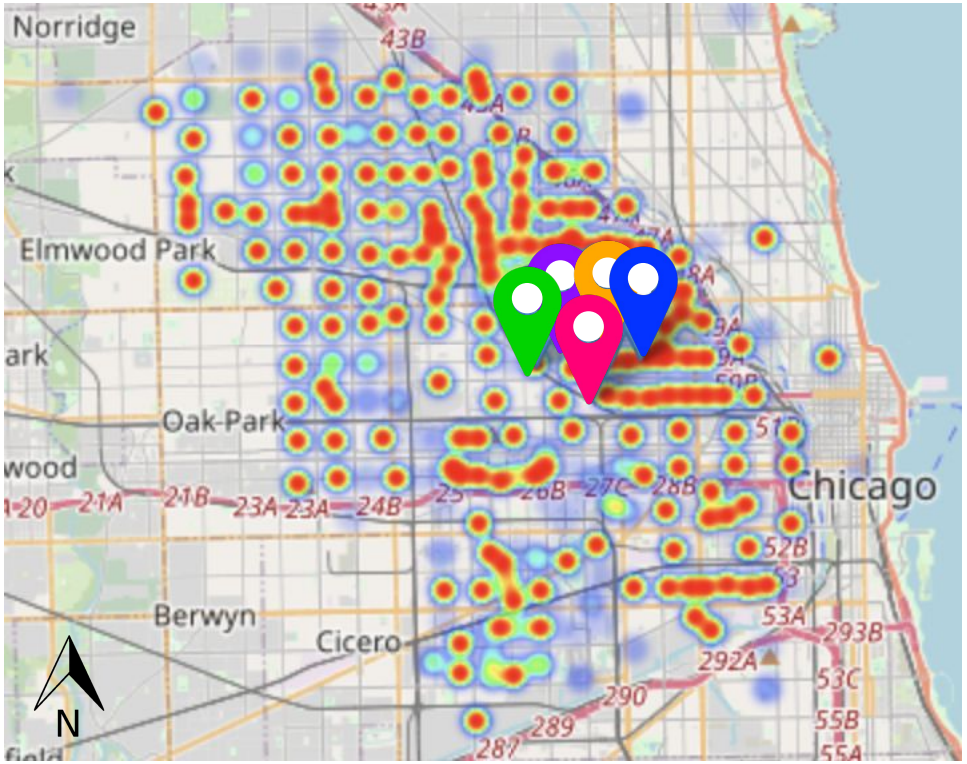


Results



Five Clusters	
3.93 miles Thursday	229.2 min 3:24pm
2.23 miles Thursday	18.8 min 4:06pm
4.37 miles Thursday	91.1 min 3:57pm
0.86 miles Thursday	6.3 min 3:18pm
3.65 miles Thursday	41.5 min 4:12pm

Results



Five Clusters		
3.93 miles Thursday	229.2 min 3:24pm
2.23 miles Thursday	18.8 min 4:06pm
4.37 miles Thursday	91.1 min 3:57pm
0.86 miles Thursday	6.3 min 3:18pm
3.65 miles Thursday	41.5 min 4:12pm

Conclusions + Recommendations

- Further analysis on the dataset
  - 1.5 - 3 hour clusters → Not enough scooters available
  - Battery life → Analyze rides under 2 minutes = dead scooters
- Focus on weekday afternoon deployment
  - Especially Thursday afternoons between 3 - 4pm
- Extend pilot outside of peak summer months

Future Work

- Model with continuous + categorical data
- Pedestrian/Bike infrastructure comparison
- Visualize most trafficked transit stops
- Compare scooter clusters against transit data
- Google Street photos of cluster centers

Source

Thanks!

Questions?

Tech Stack

matplotlib

pandas

scikit learn

python

aws

SciPy

NumPy

Folium

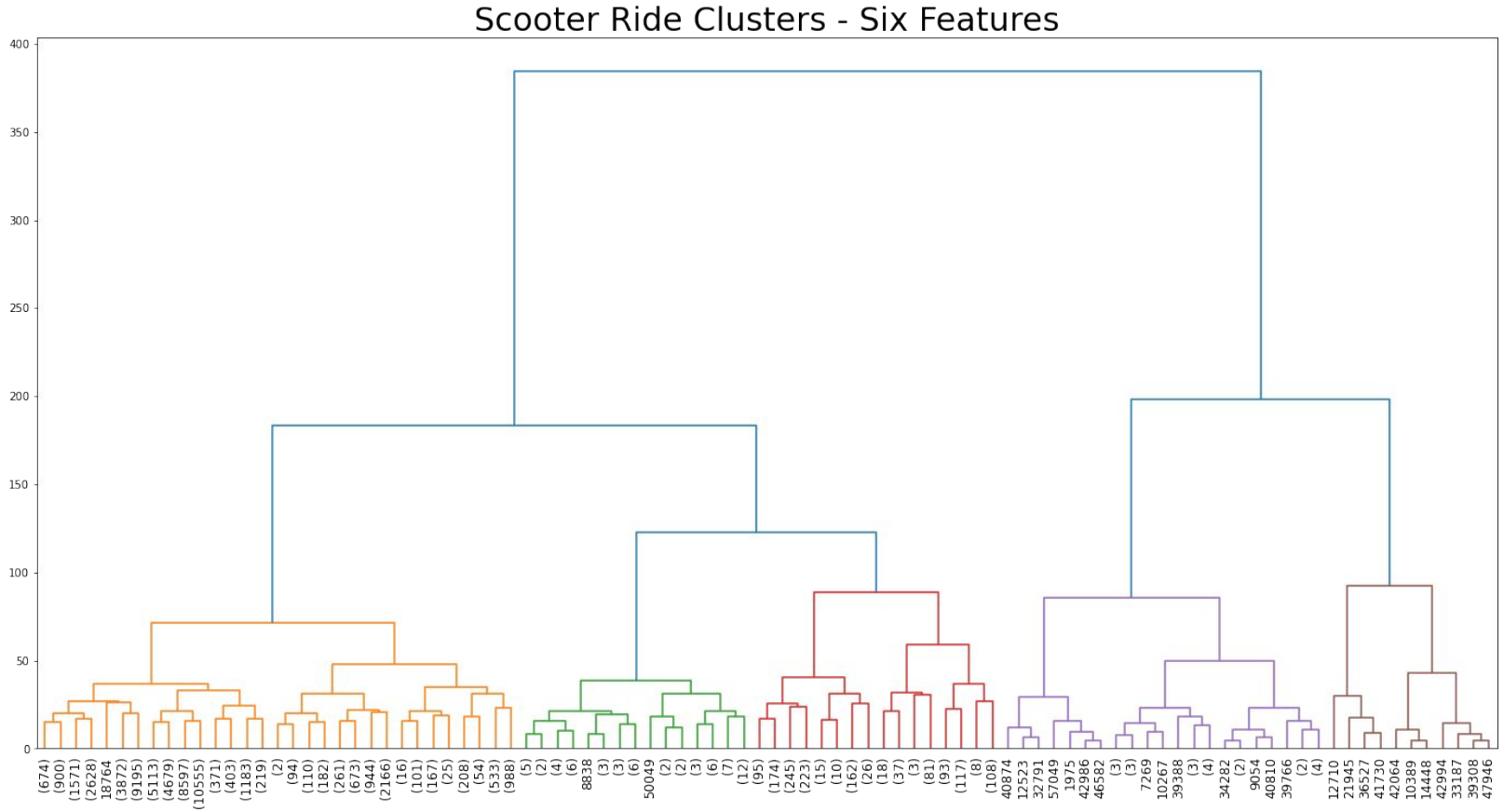
Contact Info:

sambuck2014@gmail.com

sambuck2014

sambuckwild

Appendix


Scooter Ride Clusters - Six Features
400
350
300
250
200
150
100
50
0
(674) (900) (1571) (2628) 18764 (3872) (9195) (5113) (4679) (8597) (10555) (371) (403) (1183) (219) (2) (94) (110) (182) (261) (673) (944) (2166) (16) (101) (167) (25) (208) (54) (533) (988) (5) (2) (4) (6) 8838 (3) (3) (6) 50049 (2) (2) (3) (6) (7) (12) (95) (174) (245) (223) (15) (10) (162) (26) (18) (37) (3) (81) (93) (117) (8) (108) 40874 12523 32791 57049 1975 42986 46582 (3) (3) 7269 10267 39388 (3) (4) 34282 (2) 9054 40810 39766 (2) (4) 12710 21945 36527 41730 42064 10389 14448 42994 33187 39308 47946