# Select the Ideal Community to Launch a Local Sharing Service App

Prepared By: Yao Xie
Date: 02/29/2020

# Selecting the right community is critical for Enso Street's Launch

- Enso Street, a Chicago-based startup firm, is developing a local sharing platform that focuses on family tools and equipment.

- The firm plans to debut its services in one Chicago community to test its platform, in March 2020.

- Identifying the right community help the startup to find product-market fit and gain meaningful operating metrics before their next funding event

- Besides Enso Street's founding team, future investors might be interested in the process in selecting target locations to expand geographically in the future.

# Data Acquisition

- Complete List of Communities in Chicago
  https://data.cityofchicago.org/Health-Human-Services/Uptown-Census-Data/vdfh-mxit

- Socioeconomic Data of Communities in Chicago
  https://data.cityofchicago.org/Health-Human-Services/Uptown-Census-Data/vdfh-mxit

- Geographic Data
  - Coordinates of Communities: acquired through ArcGIS geocoding API
    https://developers.arcgis.com/features/geocoding/
  - Venue Data: acquired through FourSquare API with required parameters
    https://foursquare.com/developers/apps

# Data Cleaning

- Dropped rows whose 'Community Area Number' value was NaN

- Dropped 'Community Area Number'

- Renamed 'Per Capita Income ' (with space after Income) to 'Per Capita Income'

# Methodology – Dataframe Filtering

- Our first selection criteria are socioeconomic factors.

- For each socioeconomic indicator, we only selected communities with value better than the group average value.

- The process allowed us to narrow down the list of communities that are suitable and subject to our further analysis.

# Methodology − Dataframe Filtering

| | PERCENT HOUSEHOLDS BELOW POVERTY | PERCENT AGED 25+ WITHOUT HIGH SCHOOL DIPLOMA | PERCENT OF HOUSING CROWDED | PERCENT AGED 16+ UNEMPLOYED | HARDSHIP INDEX | PERCENT AGED UNDER 18 OR OVER 64 | PER CAPITA INCOME |
|---|---|---|---|---|---|---|---|
| count | 15.000000 | 15.000000 | 15.000000 | 15.000000 | 15.000000 | 15.000000 | 15.000000 |
| mean | 13.300000 | 8.006667 | 2.173333 | 7.326667 | 11.933333 | 24.620000 | 48420.133333 |
| std | 4.554433 | 4.406726 | 1.215652 | 2.229499 | 8.250253 | 5.925152 | 17952.246520 |
| min | 3.300000 | 2.500000 | 0.300000 | 4.700000 | 1.000000 | 13.500000 | 25828.000000 |
| 25% | 11.150000 | 3.950000 | 1.200000 | 5.450000 | 5.500000 | 21.750000 | 35454.500000 |
| 50% | 13.800000 | 7.400000 | 1.900000 | 7.000000 | 10.000000 | 23.800000 | 43198.000000 |
| 75% | 16.100000 | 11.900000 | 3.300000 | 8.300000 | 18.000000 | 26.200000 | 59567.500000 |
| max | 20.600000 | 14.800000 | 4.100000 | 12.400000 | 25.000000 | 35.500000 | 88669.000000 |

# Methodology − Geocoding

- We employed the geocoding API provided by ArcGIS to retrieve latitude and longitude data for each community.

| | COMMUNITY AREA NAME | COMMUNITY LATITUDE | COMMUNITY LONGITUDE |
|---|---|---|---|
| 0 | Lincoln Square | 41.97580 | -87.68914 |
| 1 | North Center | 41.95411 | -87.68142 |
| 2 | Lake View | 41.93982 | -87.65682 |
| 3 | Lincoln Park | 41.92184 | -87.64744 |
| 4 | Near North Side | 41.90021 | -87.63433 |
| 5 | Edison Park | 42.00789 | -87.81399 |
| 6 | Jefferson Park | 41.97054 | -87.76329 |
| 7 | Logan Square | 41.92328 | -87.69810 |
| 8 | West Town | 41.89329 | -87.65743 |
| 9 | Near West Side | 41.87301 | -87.67613 |
| 10 | Loop | 41.87834 | -87.61997 |
| 11 | Near South Side | 41.85388 | -87.62711 |
| 12 | Hyde Park | 41.79388 | -87.59326 |
| 13 | O'Hare | 41.97959 | -87.90446 |
| 14 | Edgewater | 41.98571 | -87.66346 |

# Methodology – One Hot Encoding

- Convert categorical variables into a form that could be utilized by machine learning algorithms to provide optimal predictions.

| | COMMUNITY AREA NAME | Accessories Store | Airport | Airport Lounge | Airport Service | Airport Terminal | American Restaurant | Amphitheater | Antique Shop | Arcade | ... | Tunnel | Vegetarian / Vegan Restaurant | Video Game Store | Video Store | Vietnamese Restaurant | Whisky Bar | Wine Bar | Wine Shop | Wings Joint | Yoga Studio |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | Lincoln Square | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | ... | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | Lincoln Square | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | ... | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | Lincoln Square | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | ... | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3 | Lincoln Square | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | ... | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 4 | Lincoln Square | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | ... | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

5 rows × 202 columns

# Methodology – 10 Most Common Venues

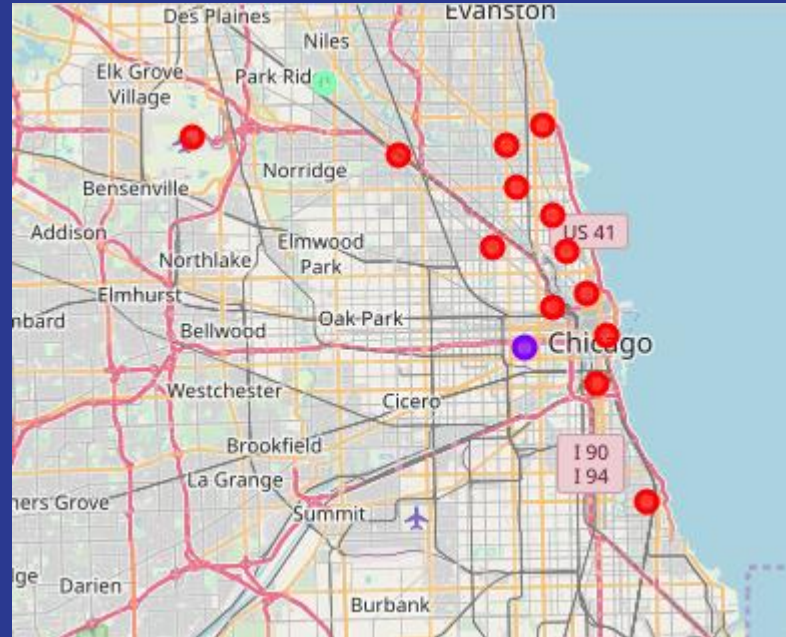- Selected the 10 most common venues ONLY for our analysis, before we conducted our K-means clustering analysis.

| | COMMUNITY AREA NAME | 1st Most Common Venue | 2nd Most Common Venue | 3rd Most Common Venue | 4th Most Common Venue | 5th Most Common Venue | 6th Most Common Venue | 7th Most Common Venue | 8th Most Common Venue | 9th Most Common Venue | 10th Most Common Venue |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | Edgewater | Mexican Restaurant | Bus Station | Indian Restaurant | Sushi Restaurant | Asian Restaurant | Antique Shop | Bakery | Deli / Bodega | Yoga Studio | Convenience Store |
| 1 | Edison Park | Theater | Neighborhood | Park | Yoga Studio | Flower Shop | Fish & Chips Shop | Filipino Restaurant | Fast Food Restaurant | Farmers Market | Exhibit |
| 2 | Hyde Park | Coffee Shop | Bookstore | Sandwich Place | Train Station | Pharmacy | Shipping Store | Bubble Tea Shop | Café | Rental Car Location | Gym |
| 3 | Jefferson Park | Bar | Pharmacy | Video Store | Coffee Shop | Ice Cream Shop | Park | Chinese Restaurant | Sushi Restaurant | Supermarket | Restaurant |
| 4 | Lake View | Café | Japanese Restaurant | Coffee Shop | Bakery | Bagel Shop | Sandwich Place | Performing Arts Venue | Gym / Fitness Center | Pizza Place | Pharmacy |

# Methodology – K-Means Clustering

- Employed unsupervised learning algorithm K-Means clustering twice in our analysis. First on our socioeconomic data and second on our venue data.

- Used cluster number of 3 in both analysis.

# Results

- The communities are categorized into three different clusters.
- We visualized our clustered communities on a Leaflet map using Folium with different colors.

# Discussion

- 13 out of 15 communities are clustered in #1, indicating a high level of homogeneity in these Chicago communities. The 10 most common venues might imply that these communities have a large population of young professionals.

| | COMMUNITY AREA NAME | 1st Most Common Venue | 2nd Most Common Venue | 3rd Most Common Venue | 4th Most Common Venue | 5th Most Common Venue | 6th Most Common Venue | 7th Most Common Venue | 8th Most Common Venue | 9th Most Common Venue | 10th Most Common Venue |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | Lincoln Square | Bar | Hot Dog Joint | Liquor Store | Bus Station | Convenience Store | Pizza Place | Korean Restaurant | Grocery Store | Football Stadium | Food Truck |
| 1 | North Center | Bar | Coffee Shop | Bank | Dive Bar | Pub | Pharmacy | Mobile Phone Shop | Boutique | American Restaurant | Latin American Restaurant |
| 2 | Lake View | Café | Japanese Restaurant | Coffee Shop | Bakery | Bagel Shop | Sandwich Place | Performing Arts Venue | Gym / Fitness Center | Pizza Place | Pharmacy |
| 3 | Lincoln Park | Pizza Place | Sandwich Place | Coffee Shop | Gym / Fitness Center | Bar | Taco Place | Breakfast Spot | Mexican Restaurant | Vietnamese Restaurant | Art Gallery |
| 4 | Near North Side | Gym / Fitness Center | Gym | Restaurant | Spa | American Restaurant | Bar | Breakfast Spot | Sandwich Place | Pub | Pool |
| 6 | Jefferson Park | Bar | Pharmacy | Video Store | Coffee Shop | Ice Cream Shop | Park | Chinese Restaurant | Sushi Restaurant | Supermarket | Restaurant |
| 7 | Logan Square | Bar | Cocktail Bar | Mexican Restaurant | Pizza Place | Restaurant | Ice Cream Shop | Discount Store | Donut Shop | Coffee Shop | Fast Food Restaurant |
| 8 | West Town | Sandwich Place | Pizza Place | Bar | Sushi Restaurant | Yoga Studio | Theater | Deli / Bodega | Pub | Burger Joint | Coffee Shop |
| 10 | Loop | Harbor / Marina | Park | Boat or Ferry | Ice Cream Shop | Concert Hall | Sushi Restaurant | Museum | Coffee Shop | Garden | Sandwich Place |
| 11 | Near South Side | Chinese Restaurant | Pizza Place | Dessert Shop | Bubble Tea Shop | Bakery | Rental Car Location | Nightclub | Grocery Store | Caribbean Restaurant | Candy Store |
| 12 | Hyde Park | Coffee Shop | Bookstore | Sandwich Place | Train Station | Pharmacy | Shipping Store | Bubble Tea Shop | Café | Rental Car Location | Gym |
| 13 | O'Hare | Airport Service | Coffee Shop | Snack Place | Airport Lounge | American Restaurant | Accessories Store | Bar | Tea Room | Tex-Mex Restaurant | Dessert Shop |
| 14 | Edgewater | Mexican Restaurant | Bus Station | Indian Restaurant | Sushi Restaurant | Asian Restaurant | Antique Shop | Bakery | Deli / Bodega | Yoga Studio | Convenience Store |

# Discussion

- Cluster No. 2 has only one community. From the venue type information, this community Near West Side might have a large population of students.

| | COMMUNITY AREA NAME | 1st Most Common Venue | 2nd Most Common Venue | 3rd Most Common Venue | 4th Most Common Venue | 5th Most Common Venue | 6th Most Common Venue | 7th Most Common Venue | 8th Most Common Venue | 9th Most Common Venue | 10th Most Common Venue |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 9 | Near West Side | Coffee Shop | Mexican Restaurant | Sandwich Place | College Gym | Baseball Field | Fast Food Restaurant | Train Station | Flower Shop | Fish & Chips Shop | Filipino Restaurant |

# Discussion

- There is one community, namely Edison Park, in Cluster #3. The three most venues are theaters, neighborhoods, and parks, all of which are highly popular among families.

| | COMMUNITY AREA NAME | 1st Most Common Venue | 2nd Most Common Venue | 3rd Most Common Venue | 4th Most Common Venue | 5th Most Common Venue | 6th Most Common Venue | 7th Most Common Venue | 8th Most Common Venue | 9th Most Common Venue | 10th Most Common Venue |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 5 | Edison Park | Theater | Neighborhood | Park | Yoga Studio | Flower Shop | Fish & Chips Shop | Filipino Restaurant | Fast Food Restaurant | Farmers Market | Exhibit |

# Conclusion

- After analyzing the socioeconomic indicators and venue types of each community in Chicago, we strongly recommend that Enso Street select Edison Park as the community to launch their services.

- As actual customer data becomes available, we will finetune our methodologies and analysis for future operation expansion.

| | COMMUNITY AREA NAME | PERCENT HOUSEHOLDS BELOW POVERTY | PERCENT AGED 25+ WITHOUT HIGH SCHOOL DIPLOMA | PERCENT OF HOUSING CROWDED | PERCENT AGED 16+ UNEMPLOYED | HARDSHIP INDEX | PERCENT AGED UNDER 18 OR OVER 64 | PER CAPITA INCOME |
|---|---|---|---|---|---|---|---|---|
| 8 | Edison Park | 3.3 | 7.4 | 1.1 | 6.5 | 8.0 | 35.3 | 40959 |