

The background is a dark blue gradient with a subtle pattern of white dots. Overlaid on the left side are several concentric circles and a large circular scale with degree markings from 140 to 260. Some circles have arrows indicating a clockwise direction.

FINDING A SPOT FOR SOME BUBBLE TEA

IBM APPLIED DATA SCIENCE CAPSTONE

NEED SOME BUBBLE TEA

- Bubble Tea is a refreshing drink consumed by too little of our population.
- Well known in Asian communities, tends to be popular when spread to other social circles.
- Our stakeholder's goal is to help spread the love in NYC.

FEATURE SELECTION

- Want to narrow down best choices of location.
- Keep it simple – looking to choose initial locations by demographic and local competition.
 - Build from census and Foursquare data

FEATURE SELECTION (CONT.)

- Competition
 - Use Foursquare data to determine
 - Number of bubble tea shops in a neighborhood (Direct Competition)
 - Percentage of total venues that are refreshment shops (Indirect Competition)
- Demographic
 - Use census data to obtain Asian population percentage

MODEL

- Use k-means clustering to obtain a groupings of similar neighborhoods
- Take cluster with high Asian demographic and low competition as cluster of interest for future analyses

RESULTS

- Obtained a cluster of 8 neighborhoods

Neighborhood	Percent Asian	Bubble Tea Shop Count	Percentage Refreshment Shops	Borough
Chelsea	64.5	0.0	5.77	Manhattan
Chelsea	64.5	0.0	5.77	Staten Island
East Harlem	58.8	0.0	0.00	Manhattan
Erasmus	57.3	0.0	10.00	Brooklyn
Maspeth	57.3	0.0	3.45	Queens
Mount Hope	54.4	0.0	7.69	Bronx
Queensboro Hill	63.6	0.0	4.00	Queens
Upper East Side	50.3	0.0	3.00	Manhattan



MAP VISUALIZATION OF CLUSTER