



Identify opportunities for new businesses

IBM Data Science Capstone Project



Table of Content

Introduction - The business problem and solution approach

Data Description

Methodology

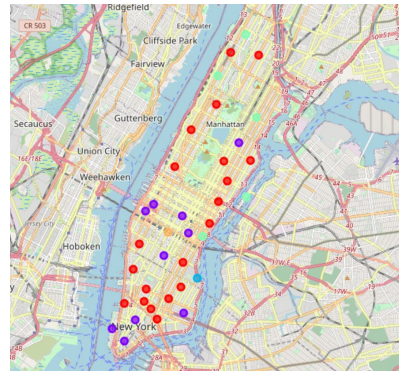
Results

Summary



Problem to solve

- For possible founders of new restaurants or other venues, one of the first and hardest choices is to pick the right place
- You need to have a solid customer base in the neighborhood you locate your business in, so you should locate your business in an area where you know people could use your venue and would appreciate its availability. On the other hand, you want that the service you provide will be in sufficient demand.
- Goal: Identify for neighborhoods what venues would be in demand
- Use case: Identify promising new venues for neighborhoods in New York City



Clustering identified similar neighborhoods in Manhattan.
Now we identify what new venues would be promising



Data Description

For each neighborhood in a cluster, we have the top 300 venues listed and ranked based on how common they are.

Here, we focus on the neighborhoods of cluster 1.

Neighborhood	Chinatown	Manhattanville	Central Harlem	Upper East Side	Yorkville	Lenox Hill
1st Most Common Venue	Bakery	Seafood Restaurant	African Restaurant	Italian Restaurant	Italian Restaurant	Italian Restaurant
2nd Most Common Venue	Chinese Restaurant	Coffee Shop	Chinese Restaurant	Exhibit	Gym	Coffee Shop
3rd Most Common Venue	Cocktail Bar	Italian Restaurant	Bar	Coffee Shop	Coffee Shop	Sushi Restaurant
4th Most Common Venue	Spa	Mexican Restaurant	Seafood Restaurant	Bakery	Sushi Restaurant	Cocktail Bar
5th Most Common Venue	Salon / Barbershop	Food & Drink Shop	French Restaurant	Gym / Fitness Center	Bar	Pizza Place

Methodology

Neighborhood	Chinatown	Manhattanville	Central Harlem	Upper East Side	Yorkville	Lenox Hill
1st Most Common Venue	Bakery	Seafood Restaurant	African Restaurant	Italian Restaurant	Italian Restaurant	Italian Restaurant
2nd Most Common Venue	Chinese Restaurant	Coffee Shop	Chinese Restaurant	Exhibit	Gym	Coffee Shop
3rd Most Common Venue	Cocktail Bar	Italian Restaurant	Bar	Coffee Shop	Coffee Shop	Sushi Restaurant
4th Most Common Venue	Spa	Mexican Restaurant	Seafood Restaurant	Bakery	Sushi Restaurant	Cocktail Bar
5th Most Common Venue	Salon / Barbershop	Food & Drink Shop	French Restaurant	Gym / Fitness Center	Bar	Pizza Place

For each neighborhood in a cluster, we have the top 300 venues listed and ranked based on how common they are

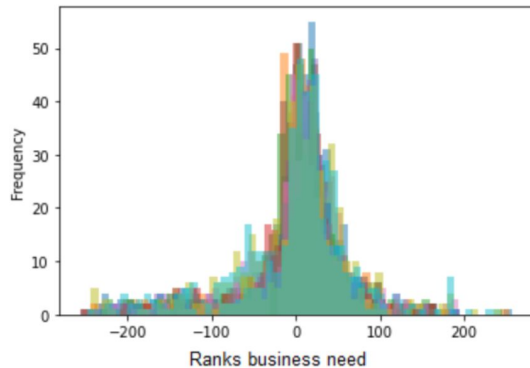
	List	Chinatown	Manhattanville	Central Harlem	Upper East Side
0	Accessories Store	300.0	300.0	300.0	300.0
1	Adult Boutique	92.0	119.0	120.0	87.0
2	Afghan Restaurant	98.0	120.0	121.0	92.0
3	African Restaurant	100.0	121.0	1.0	93.0
4	American Restaurant	100.0	29.0	6.0	11.0

For each kind of venue we calculate the average rank how common it is in the cluster of neighborhoods.

	List	Chinatown	Manhattanville	Central Harlem	Upper East Side	Yorkville	Lenox Hill
0	Accessories Store	38.35	38.35	38.35	38.35	38.35	38.35
1	Adult Boutique	-31.05	-4.05	-3.05	-36.05	6.95	20.95
2	Afghan Restaurant	-26.40	-4.40	-3.40	-32.40	6.60	-62.40
3	African Restaurant	-23.10	-2.10	-122.10	-30.10	8.90	19.90
4	American Restaurant	-31.80	-12.80	-35.80	-30.80	91.20	100.20

Based on average rank in the cluster and current rank in a neighborhood, we define a score for the business need

Results



The scores for the business needs are normal distributed. The histogram shows the score distribution for all neighborhoods. A score above 150 seems to be an outstanding business opportunity.

Here are for four neighborhoods the most-urgently needed venues:

- In the neighborhood: Manhattanville, we recommend to open:
Pizza Place, Thai Restaurant
- In the neighborhood: Central Harlem, we recommend to open:
Italian Restaurant, Mexican Restaurant, Sushi Restaurant, Thai Restaurant
- In the neighborhood: Upper East Side, we recommend to open:
Ice Cream Shop, Park
- In the neighborhood: Yorkville, we recommend to open:
Hotel



Conclusion

- The goal is to identify promising locations for new venues
- For this, we cluster neighborhoods to identify similar neighborhoods
- Based on the average availability of a venue in a cluster and current availability in a neighborhood we determine a score for business need
- If a venue has a score above 150 in a neighborhood, we label them as outstanding business opportunities
- We provide a list of outstanding opportunities for all neighborhoods



Thank you.

