Determining optimal locations for Indian restaurants in New York City

RAKESH SATPUTE

Business Problem

- The objective of this project is to find an optimal location for opening an **Indian restaurant** in **Manhattan, New York City**.

- Since there are lots of restaurants in Manahattan the idea is to detect locations that are not already very crowded with restaurants. We are also particularly interested in areas with no Indian restaurants in vicinity.

- Data science is used to generate a few most promising neighborhoods based on this criteria. Advantages of each area will then be clearly expressed so that best possible final location can be chosen by stakeholders.

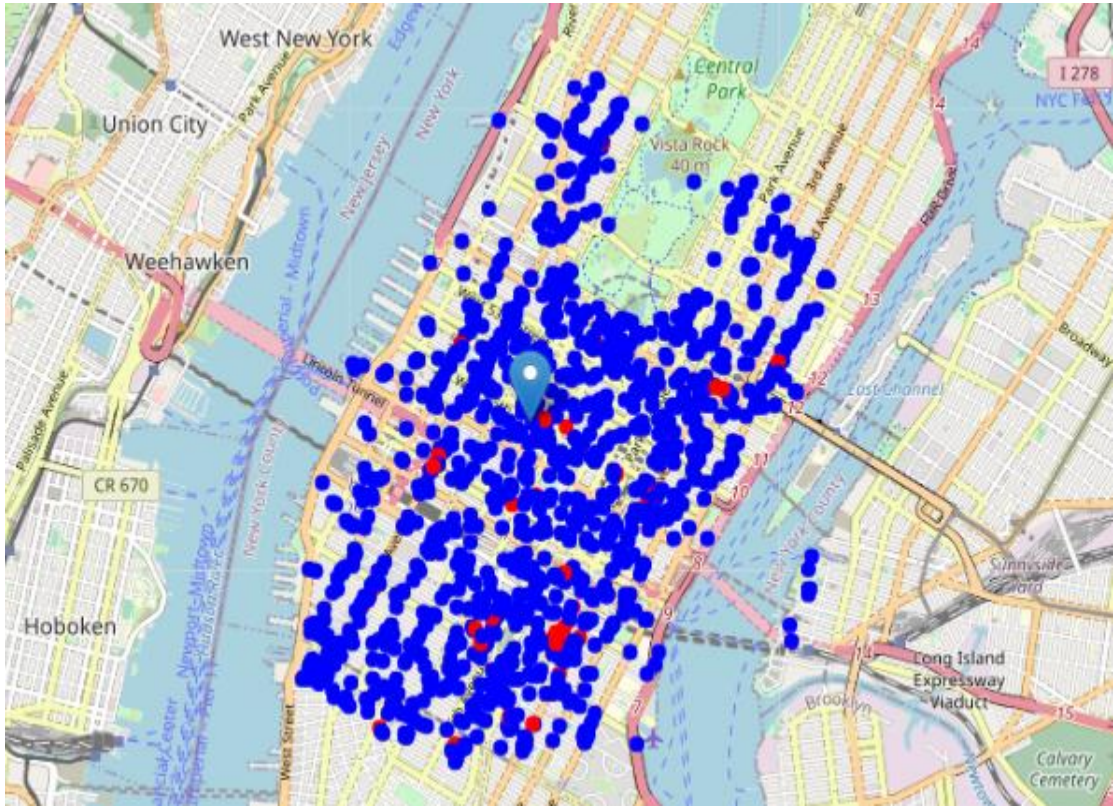
Data

Decision Points

- number of existing restaurants in the neighborhood (any type of restaurant)
- number of and distance to Indian restaurants in the neighborhood, if any

Regularly spaced grid of locations are used, centered around Times Square which is like heart of Manhattan NYC, to define our neighborhoods.

Data..



Density of restaurants in general and specifically Indian restaurants is determined around 4km of Times Square

- Total number of restaurants: 1370 (**Blue**)
- Total number of Indian restaurants: 48 (**Red**)

- Percentage of Indian restaurants: 3.50%
- Average number of restaurants in neighborhood: 14.65

Methodology

Objective is to detect areas of Manhattan that have low restaurant density, particularly those with low number of Indian restaurants. Analysis is limited to area 4km around Times Square. There are 3 main steps.

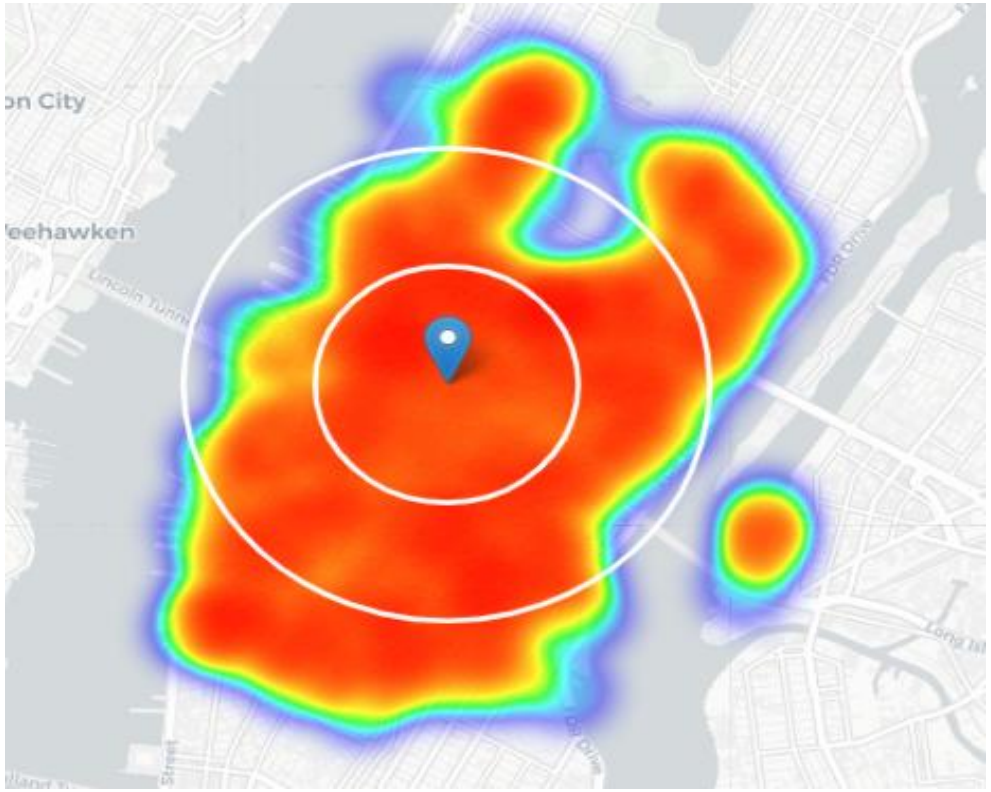
- Data collection of location and type (category) of every restaurant within 4km from Times Square and identification of Indian restaurants (according to Foursquare categorization).

- Calculation and exploration of '**restaurant density**' across different areas of Manhattan. **Heatmaps** to identify a few promising areas close to center with low number of restaurants in general (*and* no Indian restaurants in vicinity) and focus our attention on those areas.

- Within those create clusters of locations with **no more than ten restaurants in radius of 250 meters** and **without Indian restaurants in radius of 400 meters**. Clusters of these locations are created (using **k-means clustering**) to identify general zones / neighborhoods / addresses.

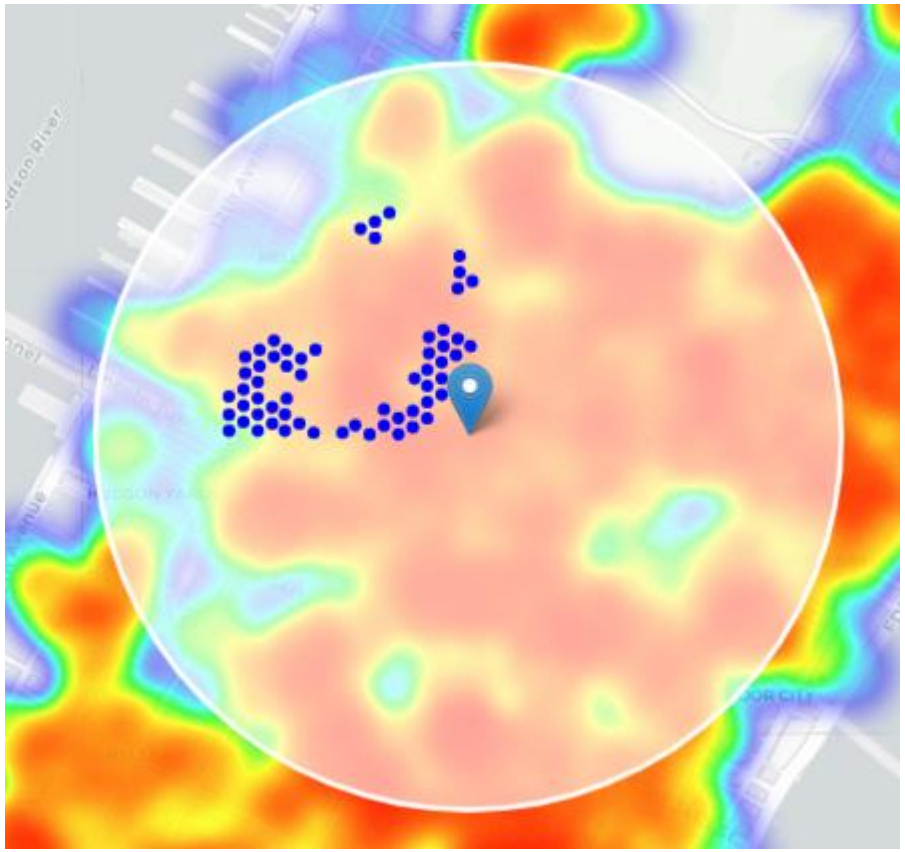
Analysis

Heat Map of Restaurants



Heat Map of Indian Restaurants

Analysis..



Further analysis is focused on **west** of Times Square which has very low density of Indian restaurants and within 1.5 km radius.

Locations with no more than ten restaurants in radius of 250 meters, and no Indian restaurants in radius of 400 meters.

- Locations with no more than ten restaurants nearby: 91
- Locations with no Indian restaurants within 400m: 154
- Locations with both conditions met: 60

Analysis..

- Clear indication of zones with low number of restaurants in vicinity, and *no* Indian restaurants at all nearby is determined.
- These locations are clustered using **K-means** clustering create centers of zones containing good locations.
- Those zones, their centers and addresses are the final result of analysis.

Results

The candidate area centers are reverse geocoded to get the addresses which can be presented to stakeholders.

- 560 10th Ave, New York, NY 10036 => 1.3km from Times Square
- 825 8th Ave, New York, NY 10019 => 0.7km from Times Square
- 711 8th Ave, New York, NY 10036 => 0.4km from Times Square
- 444 W 52nd St, New York, NY 10019 => 1.4km from Times Square
- 605 10th Ave, New York, NY 10036 => 1.3km from Times Square

Conclusion

- Determined best locations in Manhattan to open Indian restaurants using Data Science approach.

- Final decision of optimal location of Indian restaurants by stakeholders in every zone dependent on specific factors like
  - proximity to office space or park etc.
  - levels of noise / proximity to major roads
  - real estate availability
  - prices
  - social and economic dynamics of every neighborhood