Capstone Project

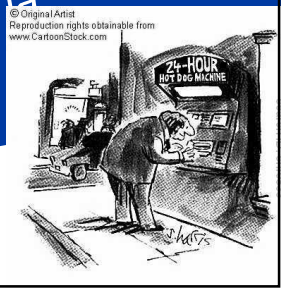
DATA ANALYSIS: SNACKS ON WHEELS - A MOBILE SNACK SHOP

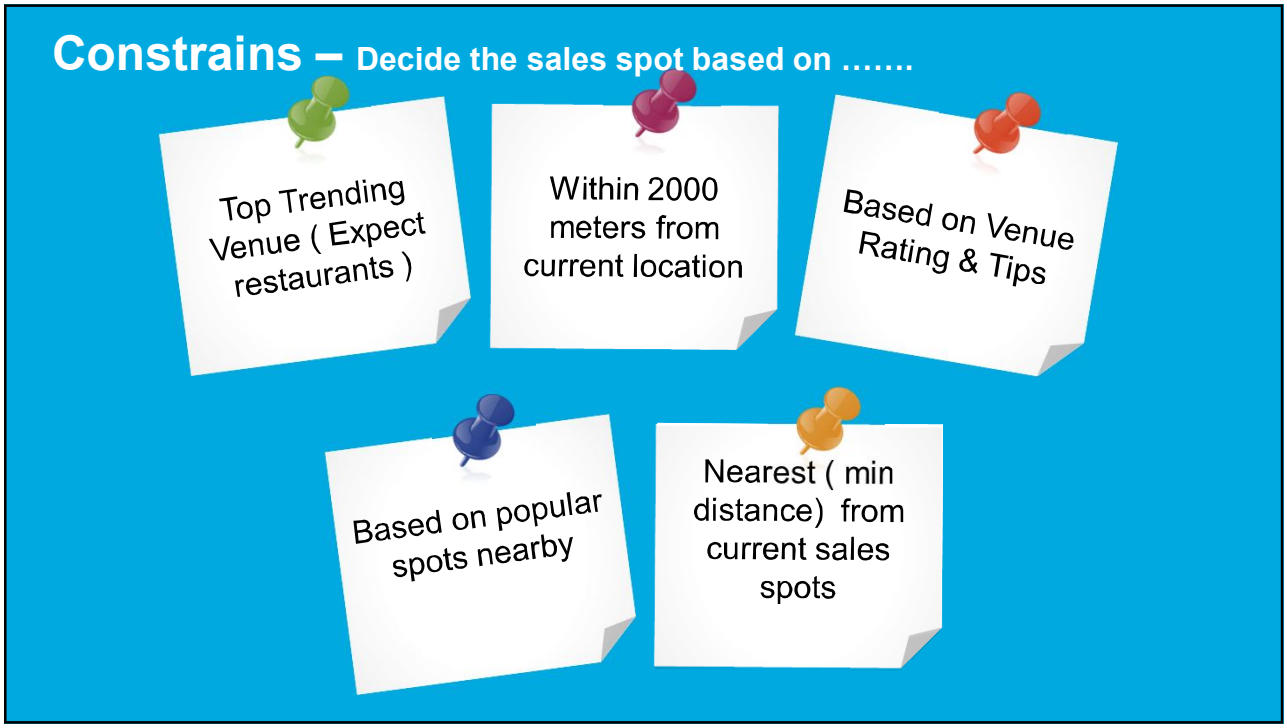
Find Best Sales Spot

INTRODUCTION

An investor in New York , USA would like to start a new busin
"Snacks On Wheels", a mobile snack shop.

This project provide data analysis support for "Snacks on wh
business to decide the NEXT sales spot/location within
neighborhood of city of New York based some criteria




Constrains – Decide the sales spot based on …….
Top Trending Venue (Expect restaurants)
Within 2000 meters from current location
Based on Venue Rating & Tips
Based on popular spots nearby
Nearest (min distance) from current sales spots

Data Used

- New York neighborhood data from New York University spatial data repository
- https://geo.nyu.edu/catalog/nyu_2451_34572

- Foursquare data on Trending Venue, Venue Rating , Tips , Users , Nearby Location & Explore

Methodology Used

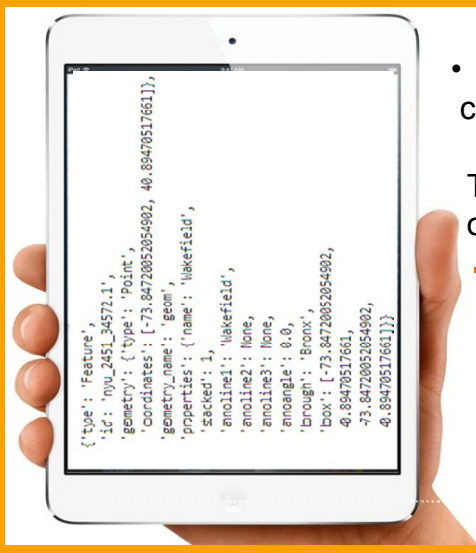
Statistical min() distance
Mean & Standard deviation

Haversine formula

FourSquare Trending Venue

Mean Venue Rating , Tips , User Details & Popular spots

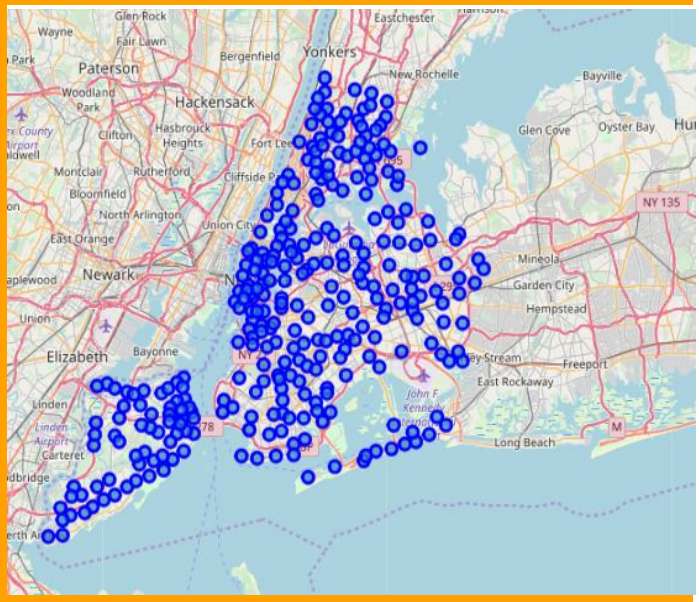
Results



- Scrape the Wikipedia page , wrangle the data, cleaned it, and then read it into a *pandas* dataframe

This data frame combined with Geospatial Data & obtained the final data set.

	Borough	Neighborhood	Latitude	Longitude
0	Bronx	Wakefield	40.894705	-73.847201
1	Bronx	Co-op City	40.874294	-73.829939
2	Bronx	Eastchester	40.887556	-73.827806
3	Bronx	Fieldston	40.895437	-73.905643
4	Bronx	Riverdale	40.890834	-73.912585



Results cont..

- For testing, we randomly selected a current location Geospatial Data from the final data set and send to foursquare API.

```
cur_loc=manhattan_data ['Neighborhood'].sample()
latitude=manhattan_data['Latitude'].loc[manhattan_data['Neighborhood'].isin(cur_loc)] .values[0]
longitude=manhattan_data['Longitude'].loc[manhattan_data['Neighborhood'].isin(cur_loc)].values[0]
cur_Borough=manhattan_data['Borough'].loc[manhattan_data['Neighborhood'].isin(cur_loc)].values[0]
print(latitude, longitude)
print (cur_Borough)
```

```
40.77352888942166 -73.98533777001262
Manhattan
```

- The return result from Foursquare further filtered to satisfy the criteria of exclude “restaurants” from trending venue data set

Results cont..

- The Trending Venues after apply filter

********** TRENDING VENUS **********

	name	categories	address	city	distance	lat	lng
0	sweetgreen	Salad Place	100 Kenmare St	New York	1940	40.721184	-73.997111
1	Brooklyn Bridge	Bridge	Brooklyn Bridge	New York	1184	40.705967	-73.996707
2	Whole Foods Market	Grocery Store	270 Greenwich Street	New York	988	40.715877	-74.012514

Discussion

The statistical function min () is used to find the minimum distance trending venue and then it assigned as initial preferred venue for venue rating and tips analysis . This method shows positive results

The Venue Rating and Tips analysis provide very useful insight about the trending venues and user details gave an opportunity to chat with top rated user for more location information.

Filtering “Restaurants” helps to avoid false positives.
Exploring the venue further for ‘other popular spots nearby’ may help for risk management in the event of trending is incident based.

Trending Venues Nearby

	text	agreeCount	disagreeCount
0	If you've forgotten what a normal grocery experience is like, come here and spread out in this spacious and well-organized location; best in NY. Full of all the delicious items your heart could desire	9	0

Venue Rating

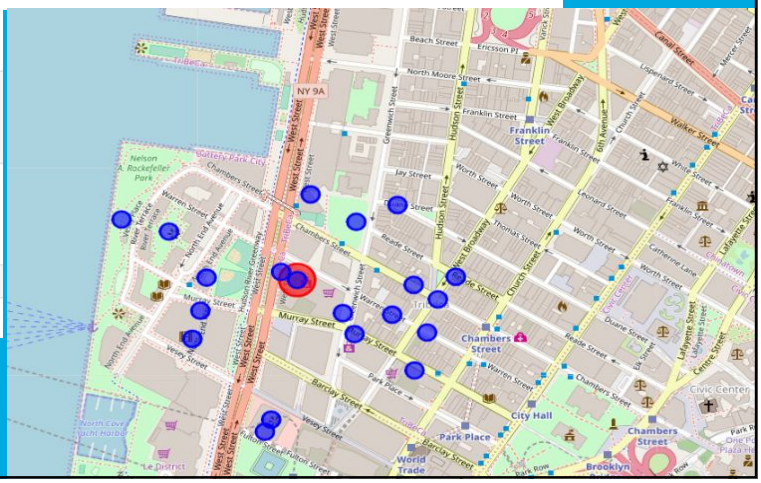
```
try:
    print(result['response']['venue']['rating'])
except:
    print('This venue has not been rated yet.')
```

```
9.3
```

Number of tips

```
result['response']['venue']['tips']['count']
```

```
247
```



Conclusion

- Developed a data analysis platform which helps 'Snacks on wheel' mobile snack shop to find a best sales spot based on current location and FourSquare trending venue details
- Venue Details analyzed further using venue rating , no of tips , tips details , etc.
- Result presented to business owner alone with other 'Popular spots nearby' data.

LIMITATIONS

References

- Y. Li, M. Steiner, L. Wang, Z. Zhang and J. Bao, "Exploring venue popularity in foursquare," *2013 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, Turin, 2013, pp. 205-210. doi: 10.1109/INFCOMW.2013.6562896

- X. Long, L. Jin and J. Joshi, "Understanding venue popularity in Foursquare," *9th IEEE International Conference on Collaborative Computing: Networking, Applications and Worksharing*, Austin, TX, 2013, pp. 409-418.


Questions ???
Email us : binejose@gmail.com