The NGO path of events on Social Life Awareness Cincinnati, Ohio

By Putrawan

10.06.2020

The challenge

A Non-Government Organisation (NGO) plans to create a new path of events to grab donors to increase awareness of social life.

The NGO that stands in the city of Cincinnati, Ohio will partnering with venues where people are present to have social life such as bar and café to lift visibility and increase sponsorship interest. The criterias as below:

- The selected venues proximity must consist at least 2 bar and 2 cafés. The cluster of venues as availabe planned budget and location distance is max 7 venues.
- All the venues candidate must have open hours around 3pm –10pm on every saturdays as the day is considered the right time for having optimum participants.
- The venues must also have a well-known reputation, and the candidates for venues must not classified expensive.

Data Usage

Credible data sources that will be used to meet the need by using the following data source:

- List of Neighbourhoods in Cincinnati, Ohio. from the “Cincinnati Area Geogrpahic Information System”
  https://data-cagisportal.opendata.arcgis.com/datasets/cincinnati-sna-boundary
  It provides land area in acres and boundary coordinates to parse for the center point of the business district.

- Use the Foursquare API to get list of neighborhood venues with hours, reviews and approximate prices. Then, set the “venues-explore” endpoint with the parameters: latitude, longitude, radius = 1000 meters, limit = 100, section = drinks and coffee. We will explore the geo-location, name, and category.

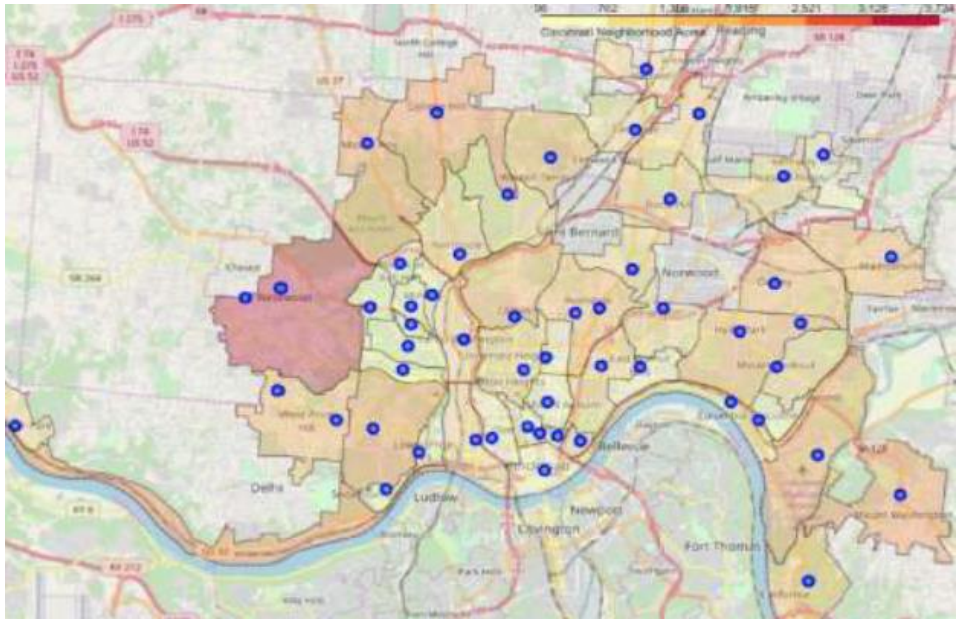
Data Parameters

- **VENUE_PRIME = ['bar', 'pub', 'brewery', 'lounge']**, bar patterns categories. **VENUE_SECONDARY = ['caf', 'coffee', 'tea', 'desert', 'ice cream', 'donut']**, café patterns categories.
- **MAX_VENUES = 7**, maximum of venues per event.
- **MAX_WALK = 0.8**, around 0.5 miles
- **MAX_PRICE = 3**, the Foursquare ranks for prices range 1 to 4.
- **MIN_PRIME = 2**, minimum number of venues that match bars. **MIN_SECONDARY = 2**, minimum number of cafes.
- **WEEK_DAY = 6**, Saturday.
- **START_TIME = 1500**, 3 PM. **END_TIME = 2200**, 10 PM.
- **PRIORITY_ORDER = {'Rating': 4, 'Count': 2, 'Likes': 1}**, provides weighting scale.

Methodology (1/2)

Exploring geodata from city APIs and display Choropleth map

To explore two geographic API datasets that are important



Define venues that match criteria and clean outliers from DBSCAN

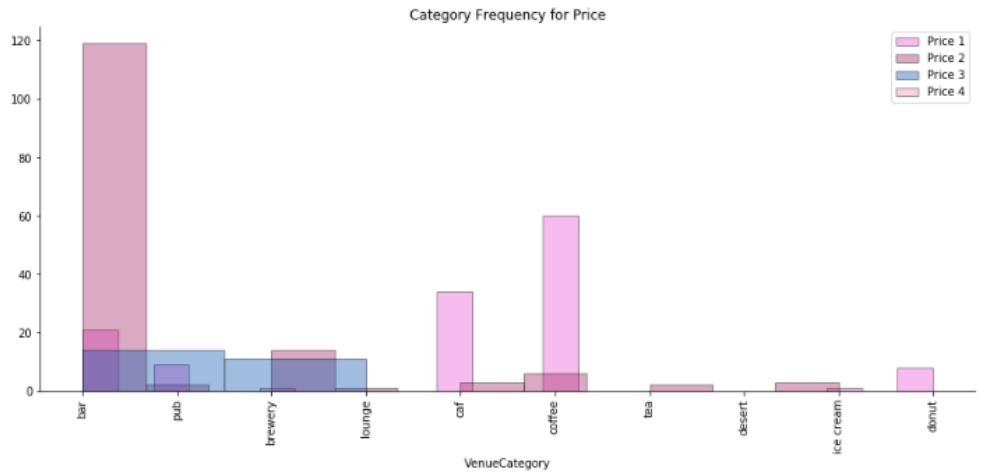
Find venues that match criteria, set center points, and clean outliers

	Neighborhood	BusinessDistrict	NeighborhoodLatitude	NeighborhoodLongitude	VenueName	VenueId	VenueLatitude	VenueLongitude	V
0	East Walnut Hills	1	39.128889	-84.476823	The Woodburn Brewery & Taproom	55461bf6498eac118325e62e	39.129030	-84.476892	
1	East Walnut Hills	1	39.128889	-84.476823	Myrtle's Punch House	5473d783498ec0bbca9021d6	39.124276	-84.476130	
2	East Walnut Hills	1	39.128889	-84.476823	The Growler House	545d54ab498ea427d9af9d2d	39.129763	-84.477778	
3	East Walnut Hills	1	39.128889	-84.476823	The Skunk Lounge	5182cdbd498e1c1b38b47f1c	39.124213	-84.476246	
4	East Walnut Hills	1	39.128889	-84.476823	Cliche	5d6459abca17630008abf539	39.123820	-84.477040	
5	Queensgate	0	39.108472	-84.533758	City West Brewing Company	580ceb4a38faa26bf32db135	39.108208	-84.525736	
6	Queensgate	0	39.108472	-84.533758	Royal Imports	4f3246f419836c91c7c7cd1e	39.102755	-84.526398	
7	Northside	1	39.160795	-84.538946	Listing Loon	51e5e417498efc97bd300341	39.161013	-84.539239	
8	Northside	1	39.160795	-84.538946	Northside Tavern	4b0220f4f964a520eb4722e3	39.162420	-84.539773	
9	Northside	1	39.160795	-84.538946	The Littlefield	53b74740498e8cb5722a80a2	39.157769	-84.539931	

Methodology (2/2)

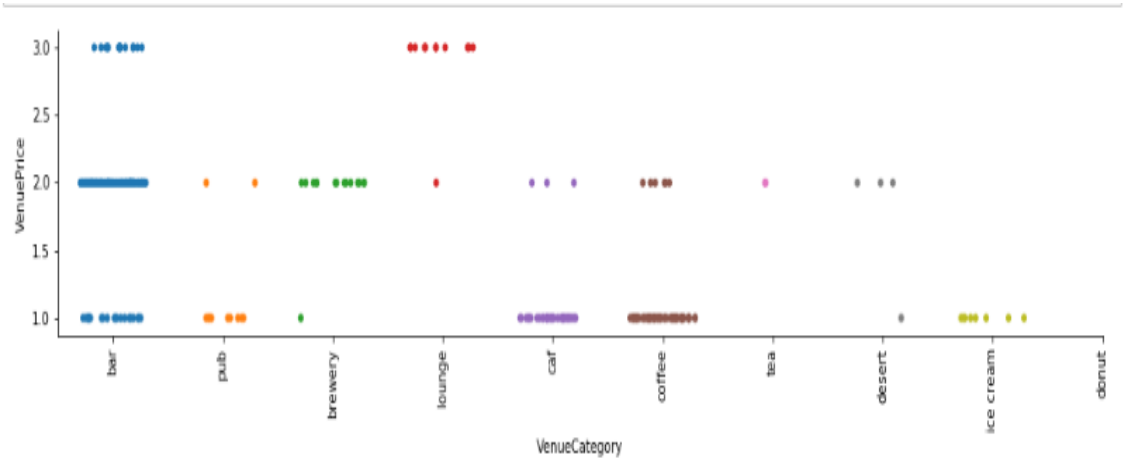
Define venues hours and use K-Nearest Neighbours machine learning

Check and exercise more venue details and patch the gap



Define venue features correlation, categorical relations with Scatterplot

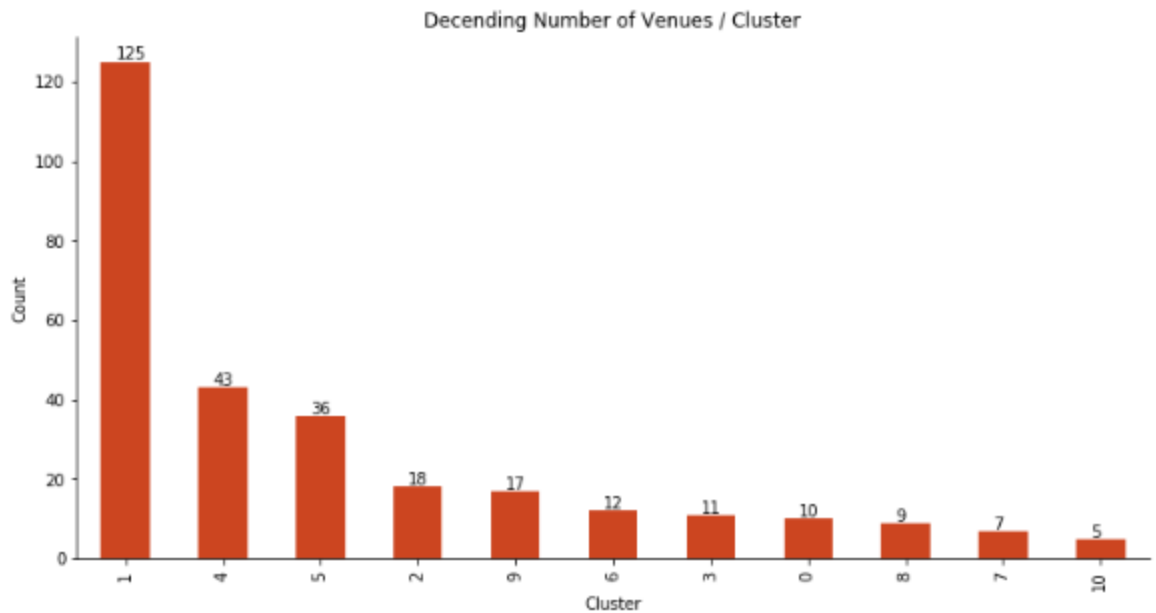
Get details, fill the missing and exercise the relations



Result Analysis

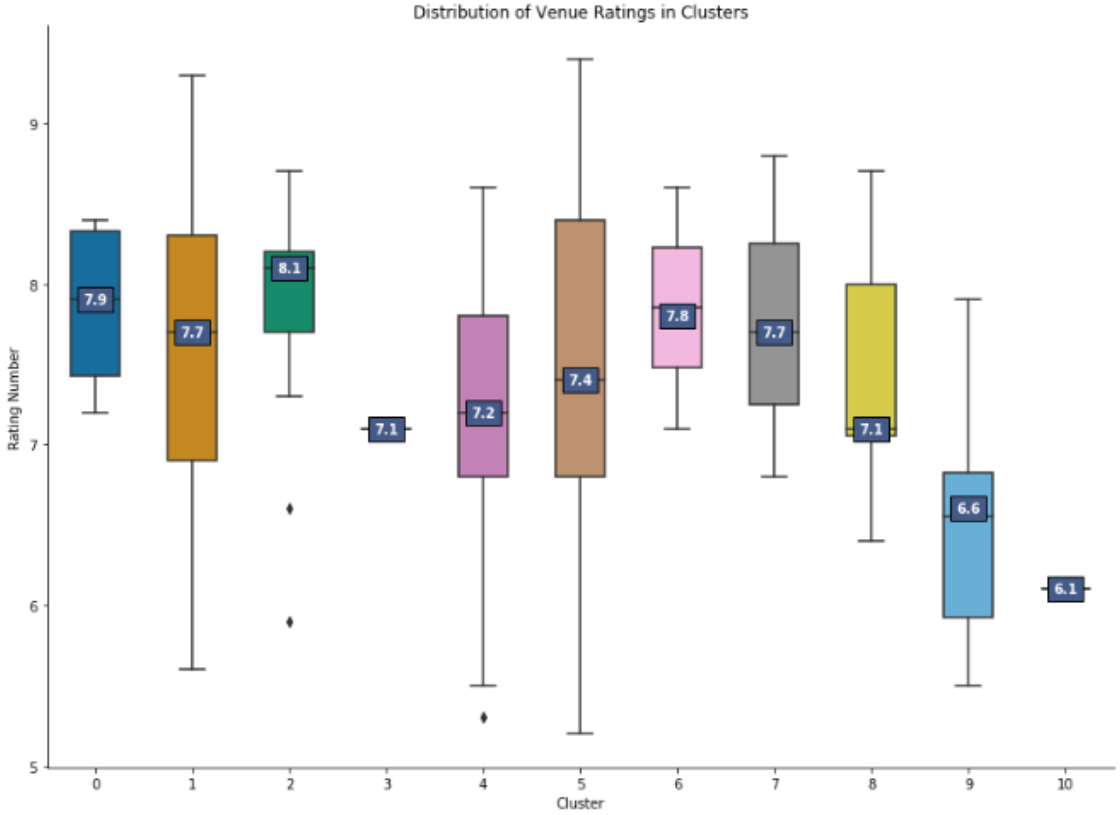
- The most qualified venues for the events are the downtown and its adjacents.
- Several clusters bleed into outside neighborhoods that are not incorporated in the city of Cincinnati.
- Cluster 4 (clifton), and cluster 5 (Hyde Park) are significant number of venues.

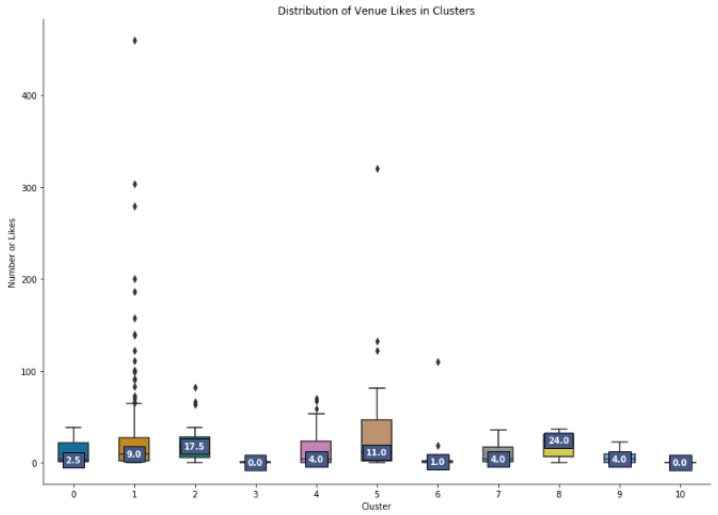
	DbCluster	Neighborhoods
0	0	[East Walnut Hills]
1	1	[Queensgate, West End, Mt. Auburn, Downtown, Over-the-Rhine, Pendleton]
2	2	[Northside]
3	3	[Roselawn]
4	4	[Mt. Auburn, Avondale, Clifton, Clifton Heights, Corryville, Walnut Hills]
5	5	[Hyde Park, Oakley]
6	6	[Evanston, North Avondale]
7	7	[Pleasant Ridge, Kennedy Heights]
8	8	[East End, Columbia Tusculum, Mt. Lookout]
9	9	[Westwood]
10	10	[West Price Hill]

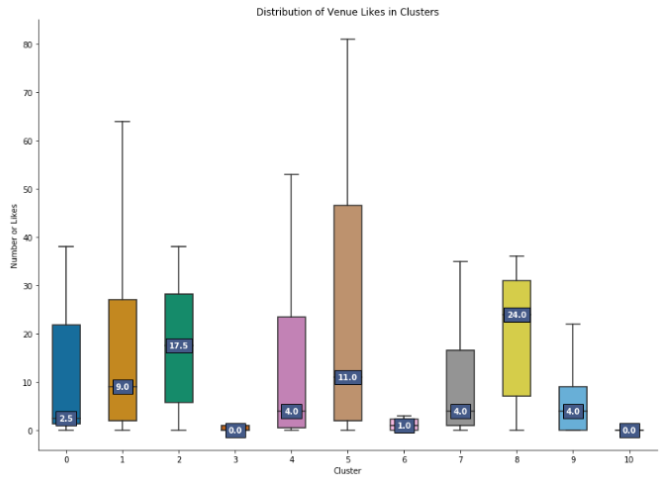


Result Analysis Boxplot

- Boxplot revealed that distribution of ratings in the clusters since quartiles, and the median are the most important values.
- Distribution of ratings in the clusters since quartiles, and the median are the most important values.
- The gimmic plan in the business-like discount package and other marketing tools influence the venue counts of like. Two boxplots below show outliers and not outliers.





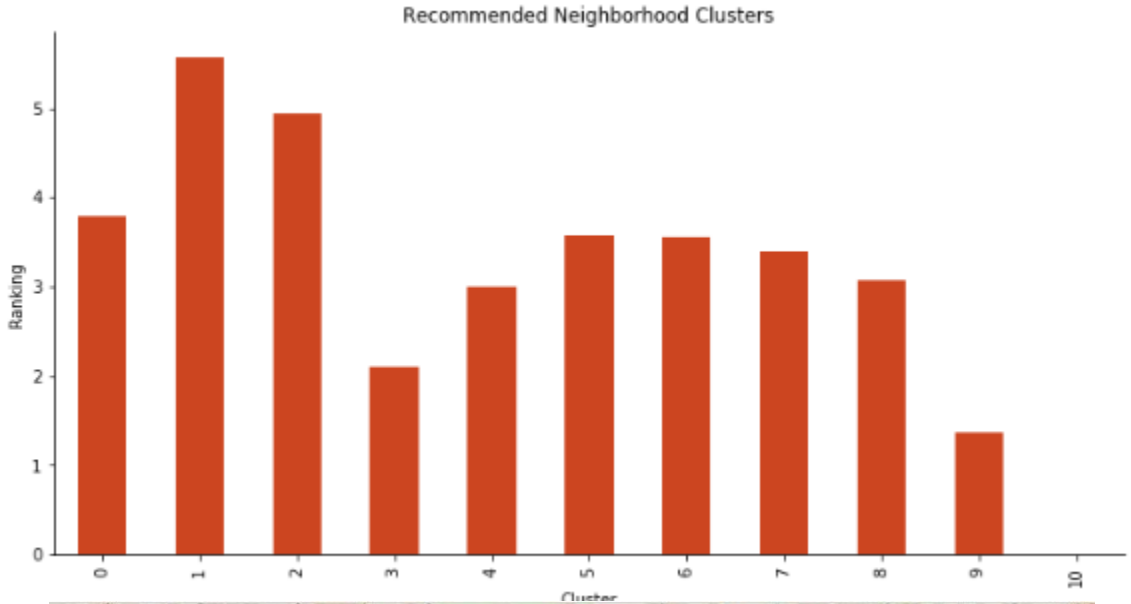


Recommendation

The cluster of venues show that the higher the ranking the better the venues. Downtown wins primarily based on its number of venues, with Northside is the top-rated venues

	VenueName	Neighborhood	VenueLatitude	VenueLongitude	VenueRating	VenueLikes	VenuePrice	PrimaryCategory	DbCluster
0	Rhinegeist Brewery	Over-the-Rhine	39.117221	-84.520129	9.3	460.0	1.0	1	0
1	Taft's Ale House	West End	39.111378	-84.517476	9.3	304.0	2.0	1	0
2	Coffee Emporium	Downtown	39.107498	-84.512390	9.1	279.0	1.0	0	0
3	Graeter's Ice Cream	Over-the-Rhine	39.110662	-84.515525	9.0	51.0	2.0	0	0
4	Cheapside Cafe	Downtown	39.105442	-84.507739	8.9	91.0	1.0	0	0
5	Longfellow	Over-the-Rhine	39.109734	-84.512704	8.9	25.0	2.0	1	0
6	1215 Wine Bar & Coffee Lab	Over-the-Rhine	39.108851	-84.515014	8.8	101.0	2.0	0	0

The Downtown group is fairly evenly distributed. This map show that the best group of venues based on ratings and much less likes and it is safe to pick any area in the downtown to hold our event.



Conclusion

The downtown neighborhoods won the top recommendations as this study analyzed the distribution of venues. The challenge found when collecting and cleaning data. External sources such as Foursquare API has limitation and must be combined with .csv data file stored for days run.

This work can be optimized by other APIs location as comparation, and enrich the criteria such as hours of venue operational.