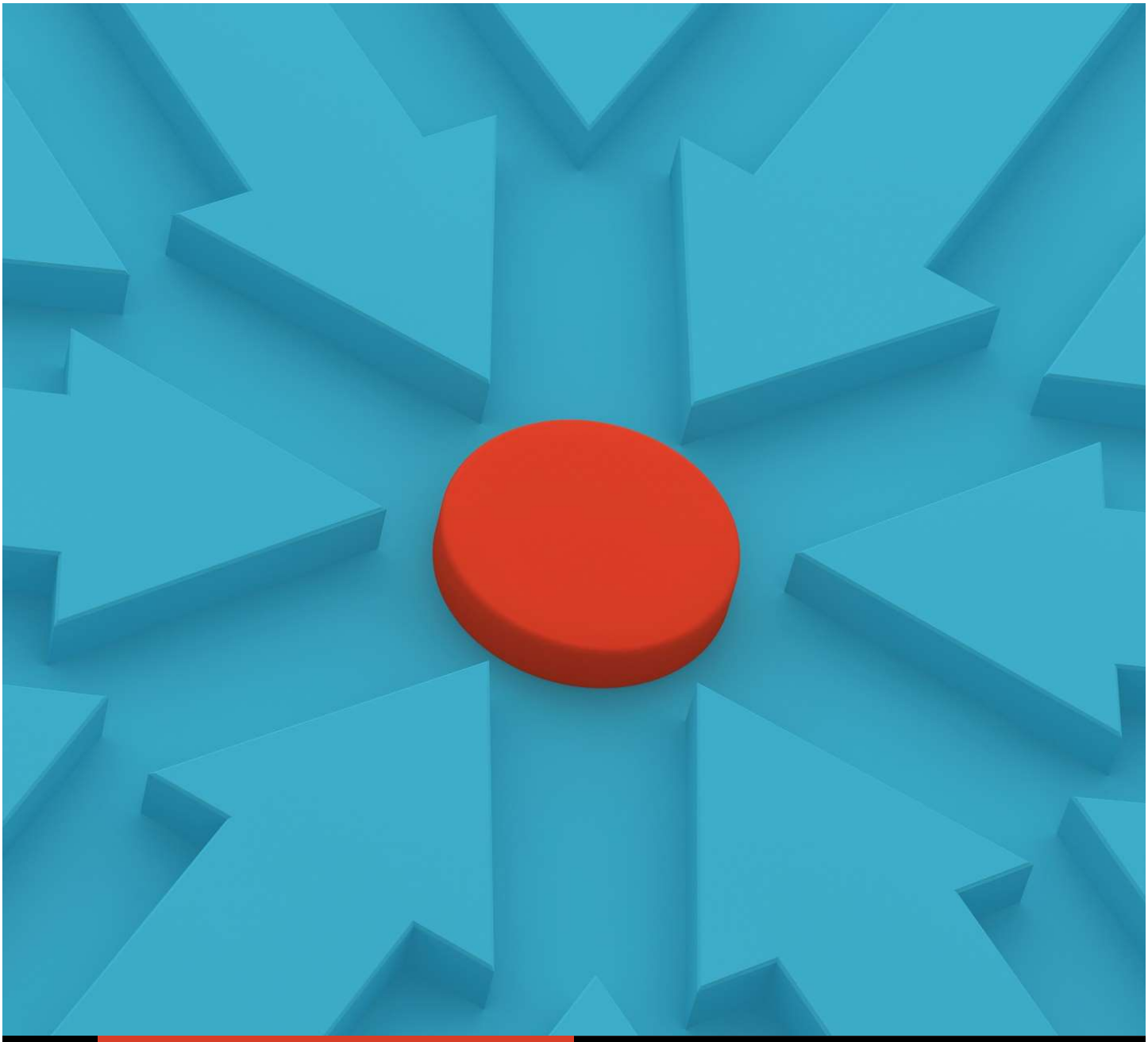
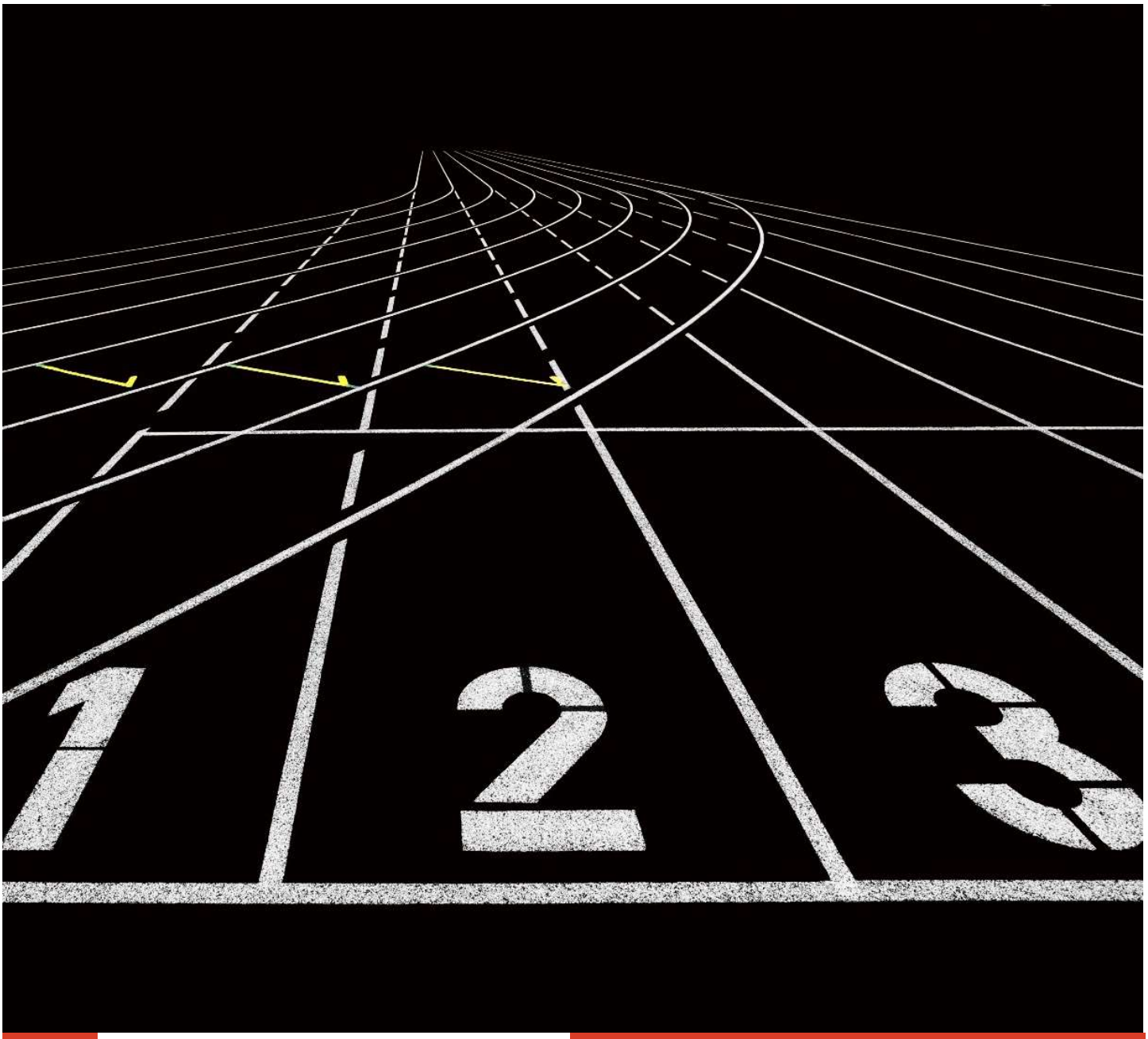


NEW BRANCH LOCATION

SELECTING BEST LOCATION IN MIAMI, FL

GYM SERVICES AND SPORT WEAR CO

Info@gymnextgen.com



FORWARD LOOKING

As societies evolves to be more fitness and health conscious, major cities witness a major change in the number of gyms and sport stores in the heart of their more populous places. Avid for lifestyle change, people look for ways to be more active after long hours sitting in an office, usually the most common position for city slickers living and working within tower buildings. Our company is committed to help people and service providers in their search for better life. Our next step, opening a branch in the paradisiac Miami, but where?

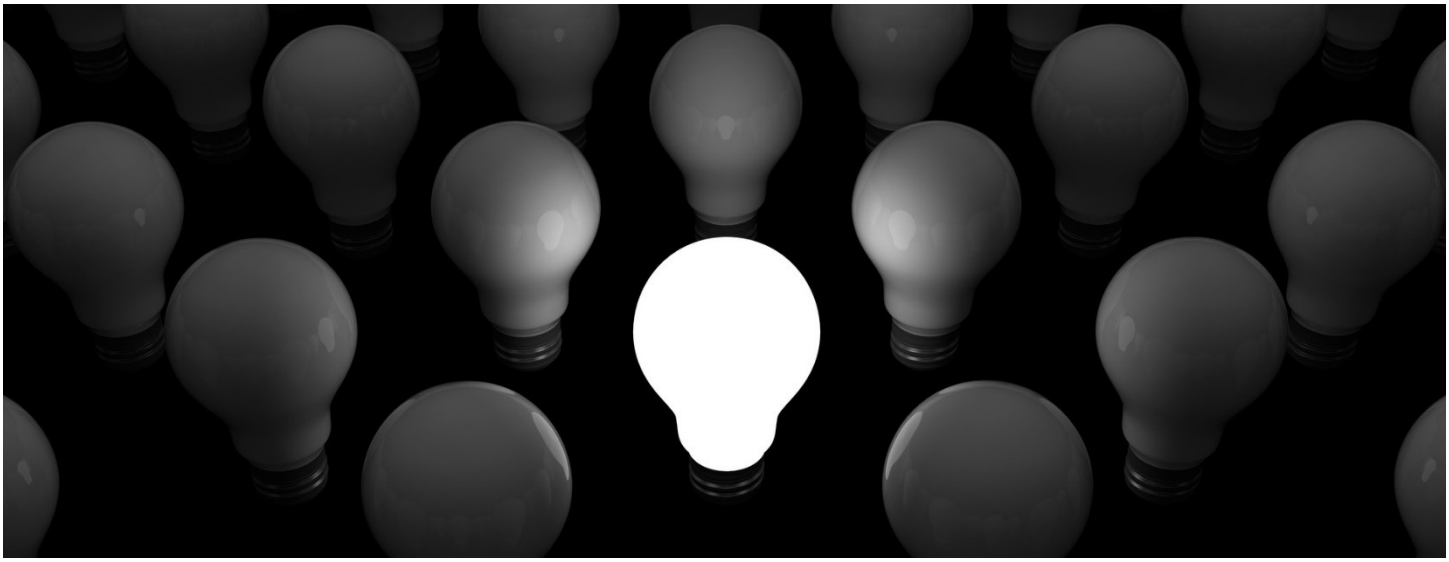
NEW METHODS FOR OLD PROBLEMS

Data Search

Presumably the most efficient and optimum location for any given store providing any kind of goods or service can be selected looking at the major geographic concentrations of potential customers. New available geospatial analytics data provide pertinent information regarding business and customers locations.

FourSquare

FourSquare is a service company specialized in providing many sorts of geospatial data oriented to business and customers. Most common data provided by FourSquare are tables containing information of businesses name, type, geographical coordinates and addresses which are located within a certain radius from a given location. The most popular businesses at current time located around a certain location. Business users reviews and users' activity information.



Capturing Relevant Data from the Internet

Interrogating FourSquare

FourSquare requires as input coordinates and radii to give back a result. In this case, a public webpage from Wikipedia.org containing geographical coordinates of the different neighborhoods within Miami, Fl is used. The data was capture using a technique called webscraping which involves using two specific python libraries in a Jupyter Notebook: requests and BeautifulSoup.

The webscraped webpage is located in following address:
https://en.wikipedia.org/wiki/List_of_neighborhoods_in_Miami

DATA DESCRIPTION

In the following lines is presented a brief description of the two main datasets collected.

Web scraped Miami neighborhood data

The data available in Wikipedia locates 24 neighborhoods belonging to city of Miami and covering a total area of 92 square kilometers. The table includes each neighborhood population and population density. Here below a snapshot of the five (5) most populous areas.

Neighborhood	Demonym	Pop 2010	Pop/km2	Lat	Lon
Brickell	Brickellite	31,759	14541	25.758	-80.193
Downtown	Downtowner	71,000 (13,635 CBD only)	10613	25.774	-80.193
Little Havana		76,163	8423	25.773	-80.215
Arts & Entertainment District		11,033	7948	25.799	-80.190
Edgewater		15,005	6675	25.802	-80.190

Business Indexes and Locations from FourSquare Service

The word ‘gym’ was used to request a data query to FourSquare Service within a radius of 1,000m for each neighborhood location coordinates to ensure all targeted businesses were considered. The resultant data was parsed and converted into a “pandas” dataframe. Next, several data rows belonging to business which were not of interest were filtered out (e.g. school gyms, hospitals, hotels, etc.). Finally, all duplicates were removed. The resulting dataset includes 129 gym venues, their names, address and geographical location.

Wikipedia Locations

Miami Neighborhood Location and Population data “scraped” from Wikipedia webpage

FourSquare Venus List

FourSquare Service query listed all gym venues across Miami Neighborhoods.

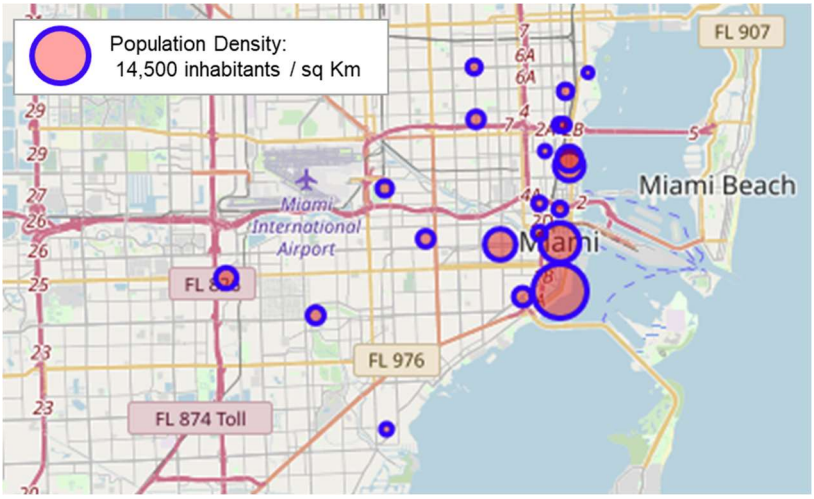
DataFrame

Venues List into DataFrame, filtered and cleaned ready to use

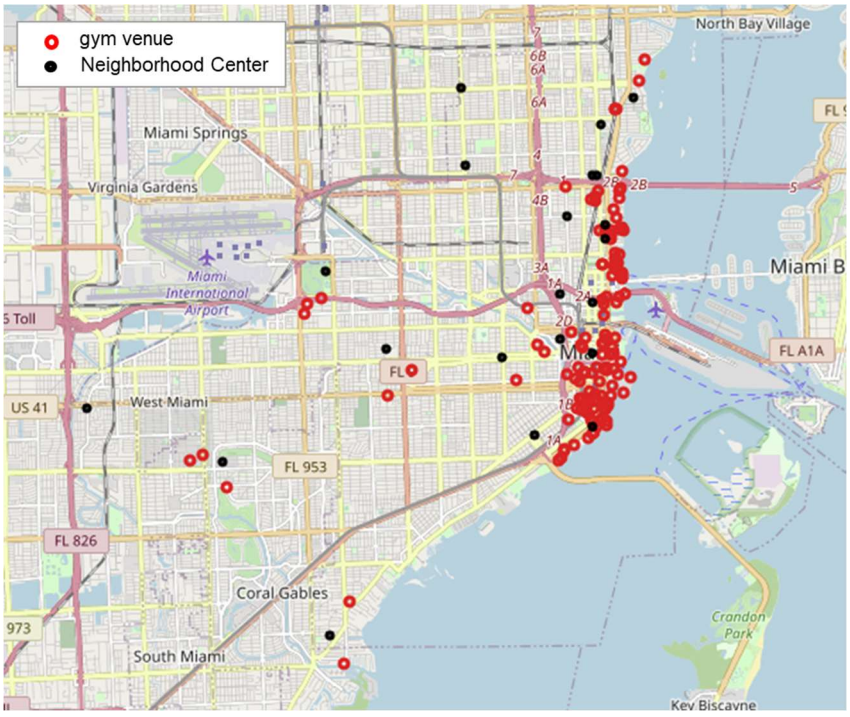
EXPLORING BEST BRANCH LOCATION

Gym Services and Sport Wear Company next branch will be necessarily located within the most “live” parts of the city, where aims to be inserted in its ecosystem and evolve as the city continues growing along with the need of its inhabitants for fitness services.

The following map clearly shows that the city is more densely populated near and along its beach.



Not surprisingly, the found gym locations display also a very strong clustering along the Beach while only few were found within more interior areas.



THE BEST CLUSTER

The available data was loaded into an algorithm called K Means, which is a simple method used in data analysis to split a population into groups or clusters around centers that tend to represent those groups. The method decides to which cluster each sample belongs based on its “distance” to each cluster center.

The method was used to highlight split the Gym locations geographically into clusters. Three metrics were calculated for each cluster:

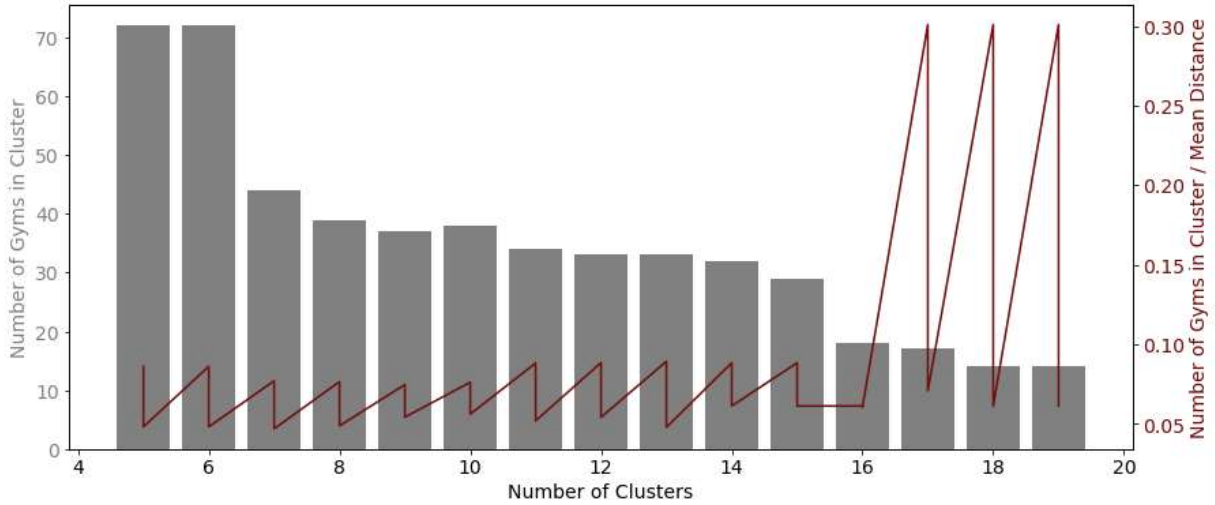
1. The number of gyms within the cluster.

2. The mean distance between the cluster center and each of the gym locations. The distance was estimated as Euclidean distance in degrees and multiplied by a constant 111,000 to obtain an approximate distance in meters.

3. A density proxy metric estimated as the number of gyms divided by the mean distance to cluster center.

The cluster centers are good options for a branch store as by design they aim to be in the center of the nearby population. The only input required prior running the algorithm is the number of clusters which is unavoidable subjective in its nature. To avoid the possibility of choosing wrong, ten (10) scenarios with different number of clusters from five (5) to nineteen (19) were run and the centers of the two (2) clusters with the highest density of gyms were selected (using the number of gyms divided by the mean distance parameter. Forcing division by more than 19 clusters results in the appearance of clusters without Gyms.

The results indicate that more than fifteen clusters result in meaningless clusters with too small population although concentrated in a small area. However, between eleven (11) and fifteen (15) clusters results in stable best clusters with similar number of stores around 32 to 35 gyms and similar density of Gyms.



Clusters Number	Cluster Centroid		Number of Gyms	Mean Distance	Gyms/MeanDist	ranking
	Latitude	Longitude				
5	25.76586	-80.1932	72	836	0.086	best place
5	25.7934	-80.1885	32	663	0.048	second best
6	25.76586	-80.1932	72	836	0.086	best place
6	25.7934	-80.1885	32	663	0.048	second best
7	25.76117	-80.1938	44	571	0.077	best place
7	25.79954	-80.1883	35	745	0.047	second best
8	25.76033	-80.1934	39	509	0.077	best place
8	25.79033	-80.1877	21	430	0.049	second best
9	25.76007	-80.1938	37	495	0.075	best place
9	25.77128	-80.1902	27	497	0.054	second best
10	25.76019	-80.1936	38	499	0.076	best place
10	25.77192	-80.1911	30	534	0.056	second best
11	**25.76278**	**-80.1923**	**34**	**385**	**0.088**	**best place**
11	**25.77301**	**-80.1906**	**24**	**462**	**0.052**	**second best**
12	**25.76193**	**-80.1921**	**33**	**373**	**0.089**	**best place**
12	**25.77255**	**-80.1916**	**29**	**535**	**0.054**	**second best**
13	**25.76141**	**-80.1927**	**33**	**370**	**0.089**	**best place**
13	**25.77217**	**-80.1887**	**20**	**418**	**0.048**	**second best**
14	**25.76245**	**-80.1921**	**32**	**362**	**0.088**	**best place**
14	**25.7928**	**-80.1873**	**14**	**228**	**0.061**	**second best**
15	**25.76175**	**-80.1926**	**29**	**328**	**0.088**	**best place**
15	**25.7928**	**-80.1873**	**14**	**228**	**0.061**	**second best**
16	25.7928	-80.1873	14	228	0.061	best place
16	25.7612	-80.1944	18	297	0.061	second best
17	25.82734	-80.1873	3	10	0.301	best place
17	25.7629	-80.1943	17	240	0.071	second best
18	25.82734	-80.1873	3	10	0.301	best place
18	25.7928	-80.1873	14	228	0.061	second best
19	25.82734	-80.1873	3	10	0.301	best place
19	25.7928	-80.1873	14	228	0.061	second best

CONCLUSION

The best locations resulting from dividing the city between eleven (11) and fifteen (15) clusters clearly indicates a concentration of good locations in an area bounded by the Southeast 1st Avenue, the Brickell Avenue, the Southwest 14th Street and Sourtheast 10th Street.

As result of this evaluation, is recommended to explore options to get a suitable location for the next company branch within the aforementioned area.