

An aerial photograph of downtown Los Angeles, showing various skyscrapers and buildings. A large, semi-transparent blue rectangle is overlaid on the center of the image, serving as a background for the text.

# **OPENING NEW RESTAURANTS IN DOWNTOWN LOS ANGELES**

**Coursera Capstone  
IBM Applied Data Science Capstone Project**

**By: Saeedeh Farivar**

# Introduction - Business Problem

---

Construction works are heating up in some grungy neighborhoods in downtown Los Angeles to keep pace with the high demand for office space. Some of the biggest projects already complete and a lot of projects planned for the area. If ultimately built, they would bring more than 5,000 new apartments and condos, along with hundreds of hotel rooms and more offices, shops and restaurants.

Building restaurants of any type in an accessible walking distance to the new developed office buildings would provide the opportunity for office staff to find a convenient place to eat at their lunch time and avoid driving in downtown traffic.

**Business question:** If developers are looking to open new restaurants, where would you recommend that they open it?

**Objective:** To analyze and select the best locations in downtown Los Angeles to open new restaurants

# Data

---

## **Data Required:**

- List of all neighborhoods in downtown Los Angeles.
- Latitude and longitude coordinates of those neighborhoods.
- Venue data, particularly data related to restaurants.

## **Source of Data:**

- This Wikipedia page ([https://en.wikipedia.org/wiki/Downtown\\_Los\\_Angeles#Districts](https://en.wikipedia.org/wiki/Downtown_Los_Angeles#Districts)) contains a list of neighborhoods.
- Python Geocoder package for latitude and longitude coordinates of the neighborhoods.
- Foursquare API (<https://foursquare.com/>) to get the venue data for those neighborhoods.

# Methodology

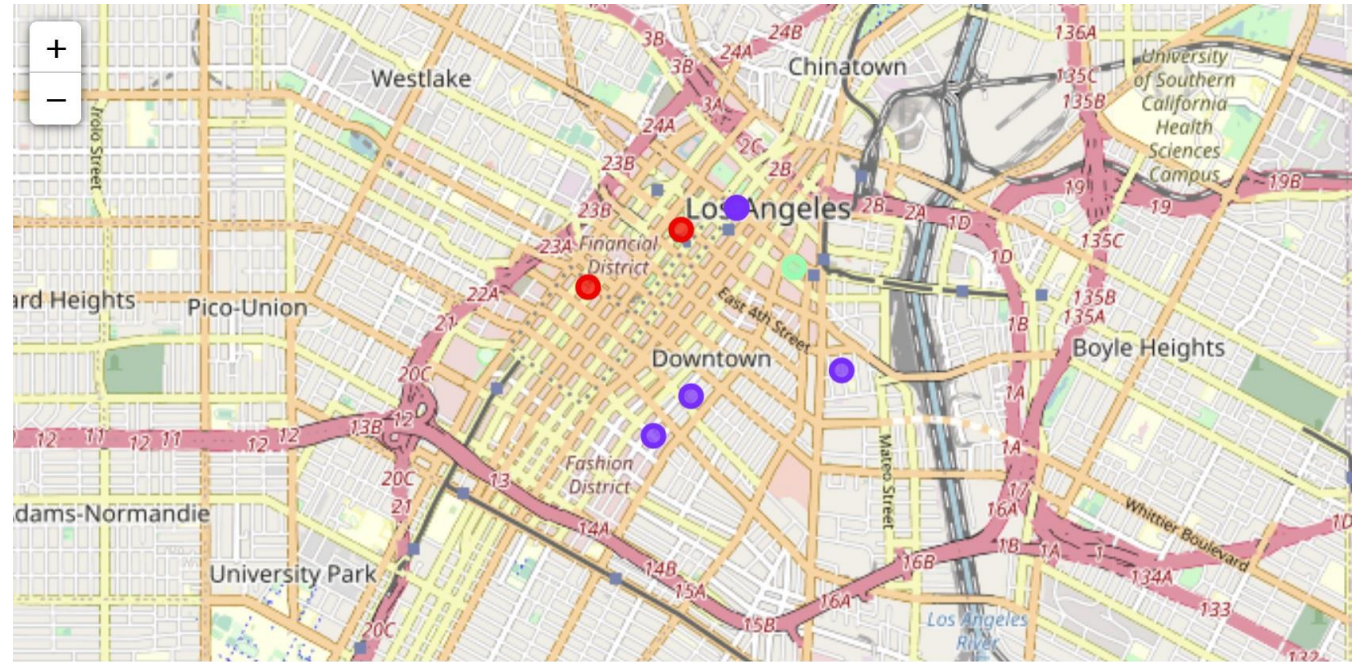
---

- Web scraping using Python requests and beautifulsoup packages to extract the list of neighborhoods
- Get the geographical coordinates in the form of latitude and longitude using Python Geocoder package
- Get the venue data using Foursquare API (<https://foursquare.com/>)
- Filter the data by extracting the venue categories containing *Restaurant*
- Analyze each neighborhood by grouping the rows by neighborhood and taking the total frequency of occurrence of restaurants.
- Perform clustering on the data by using k-means clustering

# Results

Categorize the neighborhoods into 3 clusters based on the frequency of occurrence for “Restaurant”:

- Cluster 0 – Red:  
Neighborhoods with moderate number of restaurants
- Cluster 1 – Purple:  
Neighborhoods with low number of restaurants
- Cluster 2 – Mint Green:  
Neighborhoods with high concentration of restaurants



# Discussion

---

- The highest number of restaurants in cluster 2.
- Moderate number of restaurants in cluster 0 .
- Low number of restaurants in cluster 1.

## **Recommendations:**

The nine neighborhoods in cluster 1 are the most preferred locations to open new restaurants.

# Conclusion

---

- The neighborhoods in cluster 1 are the most preferred locations to open new restaurants.
- The findings of this project will help the relevant stakeholders to capitalize on the opportunities on high potential locations while avoiding overcrowded areas in their decisions to open new restaurants.

# Resources

---

- GitHub Repository Link: <https://github.com/sfarivar/Coursera-Capstone-project/blob/master/Applied%20Data%20Science%20Capstone%20-%20Final%20Week%20Project.ipynb>
- Downtown Los Angeles, Geography, Districts in Wikipedia. Retrieved from [https://en.wikipedia.org/wiki/Downtown\\_Los\\_Angeles#Districts](https://en.wikipedia.org/wiki/Downtown_Los_Angeles#Districts)