

EVENT LOCATOR FOR THE DATA SCIENTIST MEET - 2021 IN LOS ANGELES

Applied Data science Capstone Project

-Sathya Narayanan Muralidaran

PROJECT DESCRIPTION

- A company in California was planning to set up a Data Science Annual meet-2020 in Los Angeles
- Due to Covid 19,data science meet was proposed to happen online.
- New proposal was made for the year 2021 with an idea of 5-day event.
- It needs to be planned accordingly where the people can spend quality time during off hours .
- Challenge is to choose the right spot where people can engage their leisure times in Restaurants, Parks and shopping malls.
- The venue should give some idea of choosing the right places in Los Angeles

DATA SELECTION

- Use Four Square API,
- we need to identify the below places in LA
 - Hotel
 - Restaurant
 - Parks
 - Shopping malls
- We need to convert Json file to data frame for the above searches
- Apply Data cleansing activity to all the frames

PROJECT ACTIVITIES

- Libraries Import
- Using Foursquare credentials publishing connections with Foursquare API by passing the necessary parameters
- Collecting data for Hotels
- Creating Data frames for Hotels, Restaurants, Parks and Shopping Malls
- Removing unwanted Columns for Hotels, Restaurants, Parks and Shopping Malls
- Deleting Blank or NAN rows and also ensure no duplicates
- Repeat the above 4 steps for Restaurants, Parks and Shopping Malls
- Consolidate the all the data and visualize the same in folium map.

SAMPLE DATA BEFORE CLEANING

Hotel data frame

	name	categories	address	cc	city	country	crossStreet	distance	formatted Address	labeledLatLngs	lat	lng	neighborhood	postalCode	state	id
0	The Palm Woods Hotel	Residential Building (Apartment / Condo)	The Palm Woods Hotel	US	Los Angeles	United States	Hollywood	86	[The Palm Woods Hotel (Hollywood), Los Angeles...	[{'label': 'display', 'lat': 34.054037, 'lng': 21919709...}]	34.054037	-118.241923	NaN	90012	CA	505db930e0e2e3d789c50aeb
1	Kyoto Grand Hotel Parking & Gardens Garage	Parking	120 S Los Angeles St	US	Los Angeles	United States	NaN	352	[120 S Los Angeles St, Los Angeles, CA 90012, ...]	[{'label': 'display', 'lat': 34.050526, 'lng': ...}]	34.050526	-118.242682	NaN	90012	CA	4cd50b6876ab721ebce76a6a
2	Miyako Hotel	Hotel	328 E 1st St	US	Los Angeles	United States	NaN	502	[328 E 1st St, Los Angeles, CA 90012, United S...	[{'label': 'display', 'lat': 34.049749, 'lng': 09864657...}]	34.049749	-118.240111	NaN	90012	CA	4a47ab0ff964a520eca91fe3
3	Hotel Cortez	Hotel	NaN	US	Los Angeles	United States	NaN	107	[Los Angeles, CA, United States]	[{'label': 'display', 'lat': 34.053043, 'lng': 36547851...}]	34.053043	-118.241898	NaN	NaN	CA	56292e5f498efdb6cec7a40a
4	Wyndham Garden Hotel	None	5757 Telegraph Rd	US	Commerce	United States	NaN	74	[5757 Telegraph Rd, Commerce, CA 90040, United...	[{'label': 'display', 'lat': 34.053335, 'lng': 36922002...}]	34.053335	-118.243448	NaN	90040	CA	4c95fb98efecb713c816e89d

DATA CLEANSING ACTIVITIES

- Removing unwanted columns
- Removing Duplicates
- Removing data with NAN values
- Removing invalid data
- Removing data that belongs to other category

we need to do data cleansing activity for data frame such as Hotels, Restaurants ,Parks and Shopping malls and consolidate the data to neighbourhood frame. This is the base data frame for clustering.

SAMPLE DATA AFTER CLEANING

Hotel data frame

	name	categories	address	lat	lng	postalCode	state
2	Miyako Hotel	Hotel	328 E 1st St	34.049749	-118.240111	90012	CA
5	Roes hotel	Hotel	2683 Euclid Avenue	34.052979	-118.243178	90001	CA
8	Kawada Hotel	Hotel	200 S Hill St	34.05287	-118.247909	90012	CA
11	Little Toyko Hotel	Hotel	327 1/2 E 1st St	34.050409	-118.240011	90012	CA
15	DoubleTree by Hilton	Hotel	120 South Los Angeles Street	34.050352	-118.242768	90012	CA

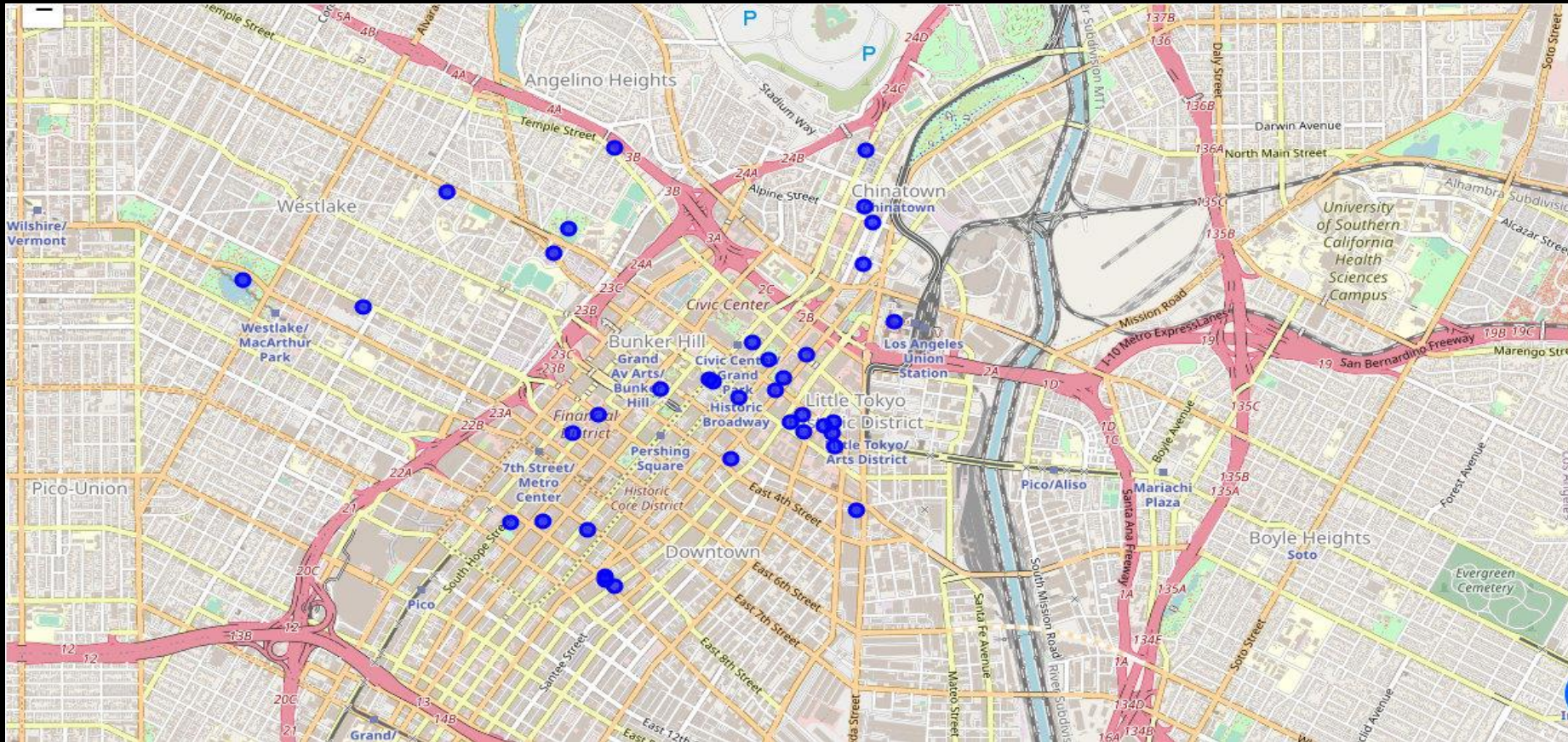
DATA CLUSTERING

During data clustering all our data will be combined into a single neighbourhood data frame.

This data is the base data for generating maps using Folium

	name	categories	address	lat	lng	postalCode	state
0	Miyako Hotel	Hotel	328 E 1st St	34.049749	-118.240111	90012	CA
2	Kawada Hotel	Hotel	200 S Hill St	34.05287	-118.247909	90012	CA
3	Little Toyko Hotel	Hotel	327 1/2 E 1st St	34.050409	-118.240011	90012	CA
4	DoubleTree by Hilton	Hotel	120 South Los Angeles Street	34.050352	-118.242768	90012	CA
5	Grand Park	Park	200 N Grand Ave	34.055034	-118.245179	90012	CA
6	Elysian Park	Park	1880 Academy Dr	34.083197	-118.234477	90012	CA
11	Grand Park Farmer's Market	Park	210 W Temple St	34.054035	-118.244134	90012	CA
13	Mayflower Seafood Restaurant	Restaurant	679 N Spring St	34.059582	-118.238129	90012	CA
14	Noe Restaurant & Bar	Restaurant	251 S Olive St	34.05229	-118.250963	90012	CA

DATA VISUALIZATION - FOLIUM MAP



CONCLUSION

Accordingly, we finalize that best place is southern California area of Los Angeles with the postal code 90012. This is the best place to conduct an event where all the Restaurants, Parks and Shopping malls are near to hotels.



THANK YOU