



Guerra de Vecindarios

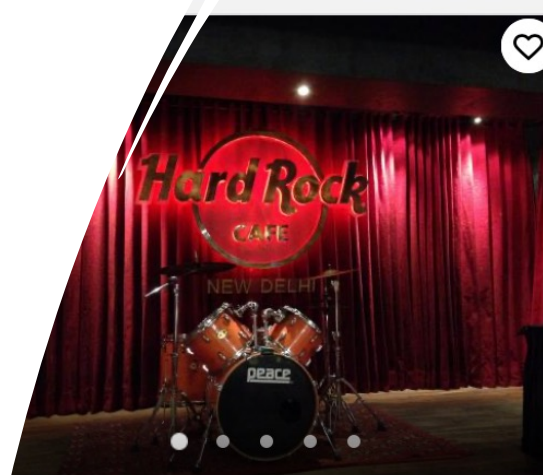
Data Science III



• Introducción/Problema

El actual trabajo correspondiente al entrenamiento de ciencia de datos, consistirá en realizar un ejercicio geoespacial para ubicar las mejores zonas geográficas de un restaurantes, o café de comida Americana.

Puntualmente este proyecto está orientado a responder a las personas interesadas en invertir y desarrollar un modelo de negocio de mayor afluencia de público, basado en sus preferencias. Dada la complejidad del caso de estudio, buscaremos la forma de agrupar zonas que contengan la mayor cantidad de negocios del rubo de comidas u ocio (según sea el caso). Una de las principales prioridades es ubicar este tipo de comercios lo más cerca del centro de la ciudad. Buscaremos en lo posible presentar las mejores zonas de comidas las cuales servirán como propuesta para un grupo exclusivo de inversores .



1. Hard Rock Cafe

●●●●○ 524 opiniones · Cerrado ahora
Estadounidense, Bar · \$\$ - \$\$\$ · 🍴 Menú ↗

- ” “Lindo como todos los HRC!!”
- ” “excelente para unas burgers”



2. The All American Diner

●●●●○ 464 opiniones · Abierto ahora
Estadounidense, Fuente de soda · \$\$ - \$\$\$

- ” “Diferente y acogedor”
- ” “Desayuno Increible”



3. Johnny Rockets

●●●●○ 148 opiniones · Cerrado ahora
Estadounidense, Fuente de soda · \$\$ - \$\$\$

- ” “Muy buen lugar para tomar té y postres.”
- ” “Grandes batidos, no vayas por una ensalada”

- _____



• Exploración

```
[17] 1 #Foursquare Credentials
      2 CLIENT_ID = 'JTB4R2ZERJU1QIVN1L4DXTEHZZS3ALDRVPDITI5KSV45D0DG '
      3 CLIENT_SECRET = 'ICQ5C1WJ0IFWHALH01K3XKDN4UFX3Q5PT3I4ZBNVW3P1SVKD '
      4 VERSION = '20180605' # Foursquare API version
      5
      6 print('Your credentails:')
      7 print('CLIENT_ID: ' + "CLIENT_ID")
      8 print('CLIENT_SECRET:' + "CLIENT_SECRET")
```

Your credentails:
CLIENT_ID: CLIENT_ID
CLIENT_SECRET:CLIENT_SECRET

```
1 #Convert the venue list into dataframe
2 venues_df = pd.DataFrame(venues)
3 venues_df.columns = ['Locality', 'Latitude', 'Longitude', 'Venue name', 'Venue Lat', 'Venue Lng', 'Venue Category']
4 venues_df.head()
```

	Locality	Latitude	Longitude	Venue name	Venue Lat	Venue Lng	Venue Category
0	Agara, Bangalore	12.84292	77.48582	Aagara Lake	12.849695	77.489070	Lake
1	Arekere	12.88567	77.59673	Decathlon Sports India Pvt Ltd	12.887513	77.597712	Sporting Goods Shop
2	Arekere	12.88567	77.59673	Natural Ice Cream	12.892188	77.598222	Ice Cream Shop
3	Arekere	12.88567	77.59673	Swensens	12.876071	77.595542	Ice Cream Shop
4	Arekere	12.88567	77.59673	Cinepolis	12.876119	77.595455	Multiplex

```
[21] 1 #Number of venues for each Locality
      2 venues_df.groupby(['Locality']).count()
```

	Latitude	Longitude	Venue name	Venue Lat	Venue Lng	Venue Category
Locality						
Agara, Bangalore	1	1	1	1	1	1
Arekere	81	81	81	81	81	81
Banashankari	100	100	100	100	100	100
Banaswadi	67	67	67	67	67	67
Basavanagudi	100	100	100	100	100	100
Basaveshwaranagar	57	57	57	57	57	57
Begur, Bangalore	12	12	12	12	12	12
Bellandur	78	78	78	78	78	78
Bengaluru Pete	100	100	100	100	100	100
Bidadi	100	100	100	100	100	100
Bommasandra	8	8	8	8	8	8

- **Algoritmo KNN**

Ejecutamos el algoritmo Agrupamiento de K-medias sobre un marco de datos para el agrupamiento de los datos. Al ejecutar el K-means obtenemos las etiquetas para las primeras 10 localidades, Aquí es donde creamos un marco de datos de blr_american, procedemos a agregar etiqueta a blr_labels y logramos una fusión de los marcos de datos blr_labels y neig_df para obtener la latitud y la longitud de cada una de las localidades. Agrupando las localidades de acuerdo a sus Etiquetas de cada Segmentación/Clúster, procedemos a Trazar el clúster en el mapa, establecemos el marcador para Bangalore y Obtener los colores para los grupos con sus respectivos marcadores de localidades. El algoritmo KNN, es un clasificador de aprendizaje supervisado, y se basa en establecer la cercanía de las clasificaciones o predicciones sobre la agrupación de un punto en el plano.



Conclusiones

- Del análisis anterior podemos inferir que el grupo 1 (mostrado en color rojo) casi no tiene restaurantes americanos con los números más altos del mismo en el grupo 2 (mostrado en color verde) y un número moderado de restaurantes americanos están presentes en el grupo 3 (mostrado con color azul) ubicado en la parte central de la ciudad.
- Este análisis presenta una gran oportunidad para que los empresarios aprovechen el potencial no utilizado de las zonas exteriores de la ciudad de Bangalore mediante la apertura de restaurantes americanos.
- También es evidente que el grupo 2 (alrededor de la parte central de la ciudad) está sufriendo de alta competencia y exceso de oferta, por lo que los desarrolladores deben evitar invertir mucho en esta área.
- Los desarrolladores con propuestas de venta únicas que pueden destacarse de la competencia moderada en el clúster 3 pueden asumir un riesgo moderado y atraer a los clientes que ya visitan la localidad de este clúster debido al restaurante americano existente.

