

Proyecto Final

Batalla de vecindarios

Análisis de la ciudad de Mérida, Yucatán

3 de enero del 2022

EMILIO BALAN JUNCO



Introducción

Contexto empresarial

Actualmente la ciudad de Mérida en el estado de Yucatán esta presentando un crecimiento exponencial en el número de habitantes. Factores que han incentivando a este aumento en su población son las siguientes:

- Ser unas de las ciudades menos violentas de México. En el 2019 fue catalogada por la revista CEO World como la segunda ciudad más segura de América. Únicamente detrás de Québec, Canadá.
- Atracción de inversión extranjera. En el año 2022 la entidad estableció un récord de plazas laborales generadas este año, arriba de 400 mil, así como en inversión extranjera, la cual rebasó los 600 millones de dólares.

Por estos razones una empresa de tecnología esta interesada en traer sus operaciones a la Ciudad de Mérida. Pero sabe que eso implica que muchos de sus empleados de mayor confianza tenga que mudarse. Por lo que están interesados en buscar la mejor colonia o Fraccionamiento para que puedan apoyar a dichos empleados en este proceso de mudanza como muestra de su gratitud a ellos.

Lo que busca dicha empresa, es que las nuevas residencias de sus empleados cuenten con restaurantes, con tiendas de servicios, parques, etc. Lo mas cercano posible, para que estos no tengan un proceso de adaptación mas complicado de lo esperado y ocasione una baja en su rendimiento.

De igual forma esto lo aprovechara la empresa, para rentar sus oficinas para que los empleados no tarden tanto en transportarse.

Datos

Descripción de los datos a emplear

Para poder resolver esta situación, necesitaremos en primer lugar los datos geoespaciales de las distintas Colonias y Fraccionamientos que hay por hoy en la ciudad de Mérida. Los datos geoespaciales a emplear son los siguientes:

- Latitud
- Longitud

Que nos será de extrema utilidad, dado que con el apoyo de la API de Foursquare, obtendremos las distintas tiendas, restaurantes, etc. Que hay en un radio de 500 metros por cada Colonia/Fraccionamiento.

Los datos antes mencionados lo obtendremos en el siguiente enlace:

https://raw.githubusercontent.com/FernandoCastro29/Coursera_Capstone/main/ZipCod_lat_lon_Neigh_Merida.csv

Metodología

Análisis Exploratorio

Una vez que hemos obtenido nuestros datos de interés procedemos a echarle un vistazo muy rápido. Lo primero que podemos percatarnos de la nuestra información es lo siguiente:

Contamos con 658 colonias/fraccionamientos en la ciudad de Mérida.

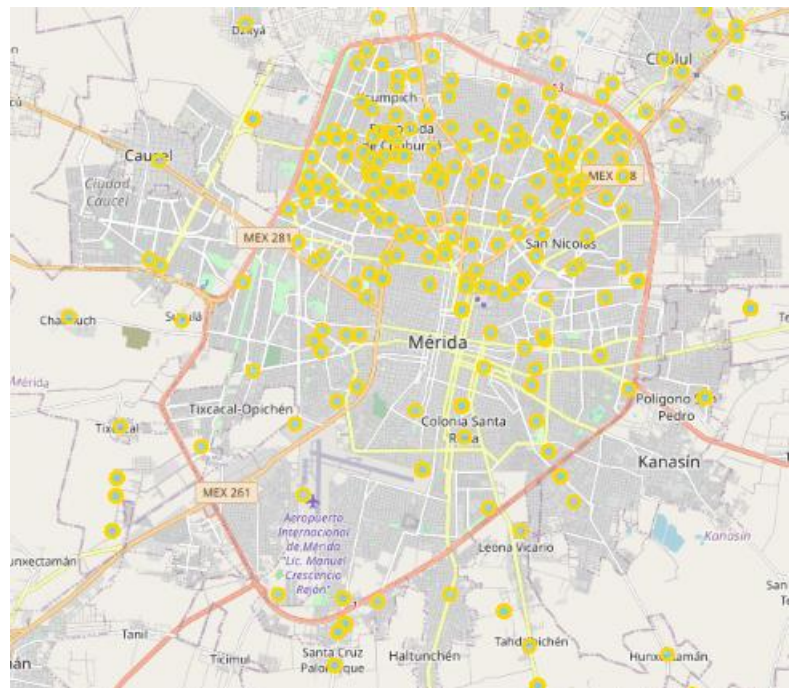
El código postal que las colonias mas comparten es el 97203, este código lo poseen las colonias del Norte de la ciudad, así como de las pobladas como son los 5 fraccionamientos de Francisco de Montejo.

CP	Frecuencia
97203	46
97314	40
97205	26
97246	26
97219	25

CP	Frecuencia
97203	Francisco de Montejo
97203	Francisco de Montejo II
97203	Francisco de Montejo III
97203	Francisco de Montejo IV
97203	Francisco de Montejo V

Mapa de Mérida

Para tener una mejor visión de los datos, empleamos un mapa de Mérida con las colonias superpuestos por encima ¿Qué podemos observar de manera mas precisa? Que la mayoría de las colonias y fraccionamientos de nuestros datos se encuentran en el Norte de la Ciudad. Esto no es mera coincidencia, dado que el crecimiento de la ciudad ha exigido un numero impresionante de Fraccionamientos nuevos, así como la expansión de las colonias. Y estás en su gran mayoría se han dado en el norte de la ciudad.



Exploración de las colonias y fraccionamientos

Una vez que hemos visto, como se encuentran distribuidos las colonias y Fraccionamientos, en toda la ciudad. Procedemos a explorar los servicios con los que cuentan cada una de estas. Considerando como servicios los supermercados, las tiendas de servicio, restaurante, etc. Esta exploración se realizara con apoyo de la API de Foursquare.

De todos datos, obtenemos que contamos con 285 servicios únicos.

Exploramos los 5 sitios mas comunes por cada barrio, adjuntamos algunos ejemplos:

```
----60 norte----
venue freq
0 Health and Beauty Service 0.11
1 Beer Garden 0.11
2 Mexican Restaurant 0.11
3 Automotive Service 0.11
4 Retail 0.11

----aguilas chuburna----
venue freq
0 Convenience Store 0.33
1 Business and Professional Services 0.33
2 Garden 0.33
3 Music Store 0.00
4 Office Supply Store 0.00
```

Creamos una nueva base de datos y mostramos los 10 lugares mas comunes para cada colonia y fraccionamiento:

Colonia	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
60 norte	Health and Beauty Service	Housewares Store	Mexican Restaurant	Beer Store	Beer Garden	Retail	Automotive Service	Office Supply Store	Restaurant	Construction Supplies Store
aguilas chuburna	Business and Professional Services	Garden	Convenience Store	Department Store	Design Studio	Dessert Shop	Diner	Dining and Drinking	Discount Store	Yucatecan Restaurant
alcal martin	Mexican Restaurant	Food and Beverage Retail	Coffee Shop	Electric Vehicle Charging Station	Juice Bar	Dance Studio	Seafood Restaurant	Car Dealership	Attorney / Law Office	Dining and Drinking
algarrobos desarrollo residencial	Arts and Entertainment	Real Estate Service	Bar	Yucatecan Restaurant	Design Studio	Dessert Shop	Diner	Dining and Drinking	Discount Store	Dive Bar
almendros	Restaurant	Wine Bar	Café	Mexican Restaurant	German Restaurant	Bar	Yucatecan Restaurant	Dining and Drinking	Department Store	Design Studio

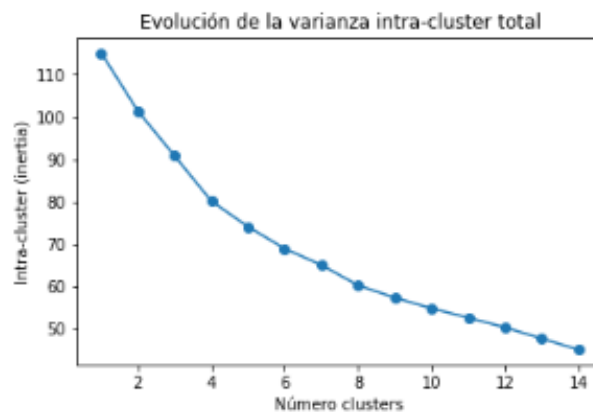
Análisis de Clusters

¿Por qué un análisis de Clusters?

Recordemos que el análisis cluster es un conjunto de técnicas multivariantes utilizadas para clasificar a un conjunto de individuos en grupos homogéneos. Cuya finalidad es encontrar patrones o grupos (clusters) dentro de un conjunto de observaciones. Las particiones se establecen de forma que, las observaciones que este dentro de un mismo grupo, son similares entre ellas y distintas a las observaciones de otros grupos. Se trata de un método de aprendizaje no supervisado (unsupervised), ya que el proceso no tiene en cuenta a qué grupo pertenece realmente cada observación (si es que existe tal información). Y justamente lo que buscamos son las mejores colonias y fraccionamientos para mudarse. Por lo que al realizar un análisis de cluster nosotros agruparemos las colonias y fraccionamientos de tal manera que las que se encuentren en un mismo cluster compartan características que en este caso serían los lugares mas comunes que tiene cada uno de estos barrios. Por lo que se podría proponer no solo una colonia o fraccionamiento, sino una cantidad mayor a ellos, permitiéndole a la empresa tener alternativas en caso de que alguna residencia tenga obstáculos, así como también permitiría reducir los datos, lo que permitiría realizar un análisis más detallado trabajando con otras variables de interés que podría surgir después de este primer análisis.

Valor k

Para elegir el valor k a emplear en el análisis de agrupamiento, emplearemos el modelo de k-means por medio de la curva del codo.



A partir de 8 clusters la reducción en la suma total de cuadrados internos parece estabilizarse, indicando que $K = 8$ es una buena opción.

Resultados

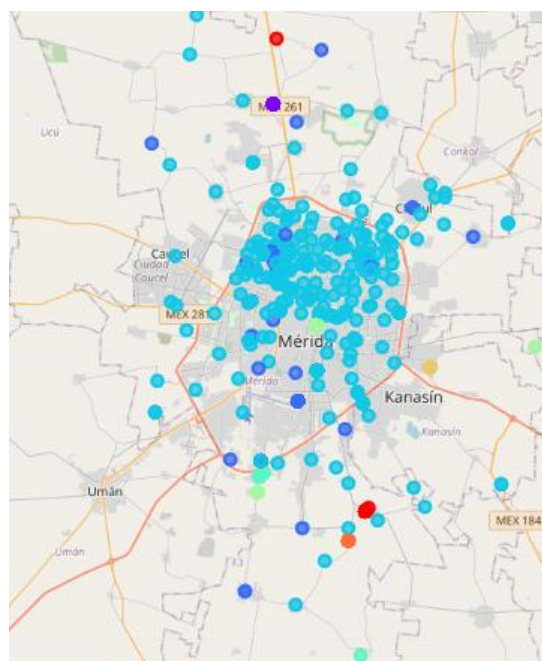
Observaciones

1. De las 658 colonias/fraccionamientos con la que comenzamos nuestro análisis, al final por temas de sesgo de la información solo se pudieron agrupar 437 de ellas, teniendo una pérdida 33.58% de la información. Esto se debe en gran medida a que las colonias no contaban con servicios cercanos a un radio de 500 metros y la mayoría de ellas eran del sur de la ciudad. Por lo que podemos decir que nuestros resultados finales no resultaron tan perjudicados.
2. En la tabla de abajo podemos observar, que el cluster que alberga más colonias es el numero 3 y el que menos alberga es el cluster 7.

Cluster	Frecuencia
0	12
1	14
2	53
3	285
4	31
5	23
6	13
7	6

Mapa de Mérida agrupados

En el siguiente mapa, podemos visualizar de mejor forma los resultados que vimos en la tabla de la diapositiva anterior



Observaciones de los cluster obtenidos

A comparación de otros análisis que se pueden hacer, el análisis de cluster no nos permite sacar resultados a partir de la frecuencia de cada uno de estos, sino tendremos en este caso ver como se comporta cada cluster, es decir cuales son los 10 lugares mas frecuentes con los que cuenta cada uno de ellos y ver que es lo que caracteriza estos clusters.

Lugares mas comunes para los 6 cluster con menos barrios

Cluster/Lugares	0	1	4	5	6	7
1	Otros grandes espacios al aire libre	Club de deportes	Servicios empresariales y profesionales	Restaurante	Gran Almacen	Piscina
2	Restaurante Yucateco	Restaurante Chino	Jardines	Bar de vinos	Tiendas de abarrotes / Supermercados	Restaurante Yucateco
3	Guardería	Grandes Almacenes	Tienda de Conveniencia	Cafe	Tienda de buceo	Antro
4	Grandes Almacenes	Estudio de diseño	Grandes Almacenes	Restaurante Mexicano	Estudio de Diseño	Grandes Almacenes
5	Estudio de Diseño	Tienda de Postres	Estudio de Diseño	Restaurante Alemán	Tiendas de Postres	Estudio de Diseño
6	Tiendas de Postre	Lugares para Cenar	Tienda de Postres	Bar	Lugares para Cenar	Tiendas de Postres
7	Lugares para Cenar	Comer y beber	Lugares para Cenar	Restaurante Yucateco	Comer y Beber	Lugares para Cenar
8	Comer y beber	Tienda de descuento	Comer y Beber	Comer y Beber	Tiendas de descuento	Comer y Beber
9	Tienda de descuento	Antro	Tienda de descuento	Grandes Almacenes	Antro	Tiendas de descuento
10	Antro	Tienda de buceo	Restaurante Yucateco	Estudio de Diseño	Parque para perros	Tienda de buceo

Lugares mas comunes para los cluster con mas barrios

- Cluster 2: Este cluster se caracteriza por tener como 10 lugares mas frecuentes, muchos restaurante o sitios relacionados a comidas; como pueden ser Pizzería, restaurante de comida mexicana, yucateca,japonesa; taquería, Buffet, de comida saludable, camiones de comida, de donas, etc.
- Cluster 3: Este sobresale en servicios no relacionados a comida, como pueden ser farmacias, servicios de salud, tiendas de electrónicos, bank. De igual forma si cuenta con restaurantes, pero no tanto como el cluster 3.

Debate

Múltiples Perfiles

Como podemos observar cada cluster, parece inclinarse como a ciertos lugares en específico. Por ejemplo si el perfil de la empresa fuera buscar un lugar que contara con servicios adicionales a los relacionados con las comidas, el mejor cluster sería el 3. Dado que ahí las colonias y fraccionamientos parecen residir en un equilibrio entre servicios de comida y los ajenos a estos.

Ahora bien, por otra parte si estos estuvieran interesados, en que el sitio donde residan tanto las oficinas como los empleados, tuviera una cantidad considerable de opciones para ir a echarse un bocado, las colonias y fraccionamientos de los cluster 2 y 5 deben ser adecuados para saciar ese perfil.

Si lo que buscara fueran lugares donde se pueda entretener como puede ser un antro o sacar a pasear a tu perro en un parque, entonces les es conveniente las colonias del cluster 6.

Y para el caso de la renta de las oficinas, los cluster que podrían ser muy preferentes sería el 3 como se menciono anteriormente y el 4 ya que este cuenta con servicios empresariales.

Conclusiones

Todo depende de la empresa

Todo va depender de que perfil busque y le convenga mas a la empresa, como a sus empleados. Por lo que será conveniente analizar que quiere la empresa, para en este caso proceder a realizar un segundo análisis para poder definir adecuadamente el plan a ejecutar.

De igual forma podemos observar, como el análisis de cluster, nos hizo este análisis más sencillo, ya que nos permite agrupar a diferente colonias con un perfil similar en cuestión de lugares. Por lo que hasta cierto grado, esto hace que la decisiones sean menos difíciles de tomar, ya que el número de elecciones se ven más reducidas.

Referencias:

- <https://inmobiliare.com/la-ciudad-de-merida-y-sus-centros-comerciales-en-2021/>
- https://www.yucatan.gob.mx/saladeprensa/ver_notas.php?id=4156
- <https://datamexico.org/es/profile/geo/merida#:~:text=Acerca%20de%20M%C3%A9rida&text=En%202020%2C%20la%20poblaci%C3%B3n%20en,85%25%20respecto%20al%20a%C3%B1o%20anterior>
- <https://www.linkedin.com/pulse/analysing-clustering-neighborhoods-m%C3%A9rida-yucat%C3%A1n-castro-trejo>
- https://stackoverflow-com.translate.goog/questions/11816844/object-of-type-nonetype-has-no-len-error?_x_tr_sl=en&_x_tr_tl=es&_x_tr_hl=es-419&_x_tr_pto=sc
- <https://www.cienciadedatos.net/documentos/py20-clustering-con-python.html>

