

Introducción y discusión del caso

Con motivo del "Record Store Day" (día mundial en el que se lanzan ediciones especiales de discos), la compañía discográfica "Eso a lo que tú llamas indie" está interesada en abrir su propia tienda de discos en la que distribuir su propio material.

Dado el auge que ha obtenido en la última década la compra/venta de discos de vinilo, su principal preocupación es definir una ubicación en Madrid adecuada, donde pueda tener una buena consolidación sin tener una excesiva competencia, ya que el mercado de compra/venta de discos de vinilo está en auge, pero aún no está lo suficientemente expandido.

EL objetivo es encontrar un barrio de Madrid donde se cumplan los siguientes requisitos:

- Alta concentración de salas de conciertos y/o locales donde se toque música en directo: La discográfica considera que su público objetivo, regenta este tipo de locales
- Baja concentración de tiendas de música: Esto limitará sustancialmente la competencia, por lo que ayudará a la consolidación de la tienda
- Lo más cercano al centro posible: El turismo en Madrid es elevado y principalmente los turistas suelen moverse por el centro de la ciudad. Muchos aficionados a la música extranjero invierten tiempo y dinero en adquirir música en sus viajes al extranjero

Descripción y uso de los datos

Primero procederemos a obtener toda la información referente a los barrios de Madrid de la siguiente web:

'https://es.wikipedia.org/wiki/Anexo:Barrios_administrativos_de_Madrid'

Con la ayuda del módulo "*Nominatim*" podremos obtener la ubicación de cada barrio con sus coordenadas de latitud y longitud, y así poder obtener su distancia al centro de la ciudad

Con la ayuda de la API de Foursquare, obtendremos los puntos de interés definidos en el apartado anterior para cada barrio.

Por último, con todos los datos necesarios, aplicaremos el algoritmo K-Means para la obtención de distintos clusters que nos permitan escoger la correcta ubicación de la nueva tienda

Recolección y preparación de los datos

Se ha obtenido un primer DataFrame con los datos de todos los barrios de Madrid, con la información referente a su ubicación (latitud y longitud), así como su distancia al centro de la ciudad:

	Distrito	Número	Nombre	Latitud	Longitud	Distancia
0	Centro	11	Palacio	40.415129	-3.715618	1.376854
1	Centro	12	Embajadores	40.409681	-3.701644	0.804495
2	Centro	13	Cortes	40.414779	-3.697584	0.675495
3	Centro	14	Justicia	40.423276	-3.696680	1.037271
4	Centro	15	Universidad	40.425426	-3.706023	1.010150

Seguidamente con la ayuda de la Api de Foursquare, hemos podido obtener los datos y ubicaciones de los lugares localizados en un radio de 1000 metros al centro de cada barrio, donde se celebren eventos relacionados con las artes escénicas:

	Barrio	Distancia Centro	Lugar	Latitud Lugar	Longitud Lugar	Categoría Lugar
0	Palacio	1.376854	Corral de la Morería	40.412760	-3.714035	Concert Hall
1	Palacio	1.376854	ContraClub	40.412533	-3.713977	Concert Hall
2	Palacio	1.376854	Las Tablas Tablao Flamenco	40.421889	-3.713589	Music Venue
3	Palacio	1.376854	Teseo Teatro	40.408551	-3.715884	Theater
4	Palacio	1.376854	Las Carboneras	40.414744	-3.709295	Music Venue
...
1237	Timón	3.396441	El Grito	40.380319	-3.707109	Rock Club
1238	Timón	3.396441	Teatro Benavente	40.393019	-3.707500	Theater
1239	Timón	3.396441	Big Step	40.393857	-3.706994	Music Venue
1240	Timón	3.396441	Producciones el Zurdo	40.393967	-3.711080	Theater

Usando la misma técnica podemos obtener las tiendas de música asociadas a cada barrio:

	Barrio	Distancia Centro	Lugar	Latitud Lugar	Longitud Lugar	Categoría Lugar
0	Palacio	1.376854	Discos la Metralleta	40.418644	-3.706733	Record Store
1	Palacio	1.376854	Escridiscos	40.419094	-3.706467	Record Store
2	Palacio	1.376854	La Gramola	40.419166	-3.705977	Record Store
3	Palacio	1.376854	La quinta de Mahler	40.417069	-3.710689	Record Store
4	Palacio	1.376854	Sun Records	40.420474	-3.709101	Record Store
...
68	Argüelles	2.109148	Nakasha Discos	40.432598	-3.714253	Record Store
69	Argüelles	2.109148	Toni Martin Discos	40.425170	-3.714396	Record Store
70	Argüelles	2.109148	La Gramola	40.424995	-3.707120	Record Store
71	Pueblo Nuevo	6.885574	Cuna Records	40.434677	-3.636429	Record Store
72	Quintana	7.025536	Cuna Records	40.434677	-3.636429	Record Store

Juntando estos tres DataFrames y aplicando el método `get_dummies`, obtenemos un único DataFrame en el que aparece toda la información que necesitamos para nuestro análisis:

	Barrio	Comedy Club	Concert Hall	Indie Theater	Jazz and Blues Venue	Music Venue	Performing Arts Venue	Record Store	Rock Club	Theater	Latitud	Longitud	Distancia
0	Abrantes	0	1	0	0	0	1	0	2	1	40.381328	-3.727455	4.756185
1	Acacias	2	9	3	0	15	1	1	0	20	40.404075	-3.705957	1.433996
2	Adelfas	0	0	0	1	4	0	0	0	3	40.401903	-3.670958	3.958680
3	Alameda de Osuna	0	1	0	0	0	0	0	0	0	40.457581	-3.587975	13.605442
4	Almagro	0	11	1	0	5	0	0	1	13	40.433663	-3.694286	2.131744
...
93	Vallehermoso	0	1	0	0	3	1	0	2	13	40.443189	-3.710823	3.051762
94	Ventas	0	3	0	0	2	1	0	1	12	40.430831	-3.663280	4.721211
95	Vinateros	0	1	0	0	0	0	0	0	0	40.405197	-3.641547	6.989153
96	Vista Alegre	1	6	0	0	3	0	0	0	2	40.387792	-3.739654	5.158987

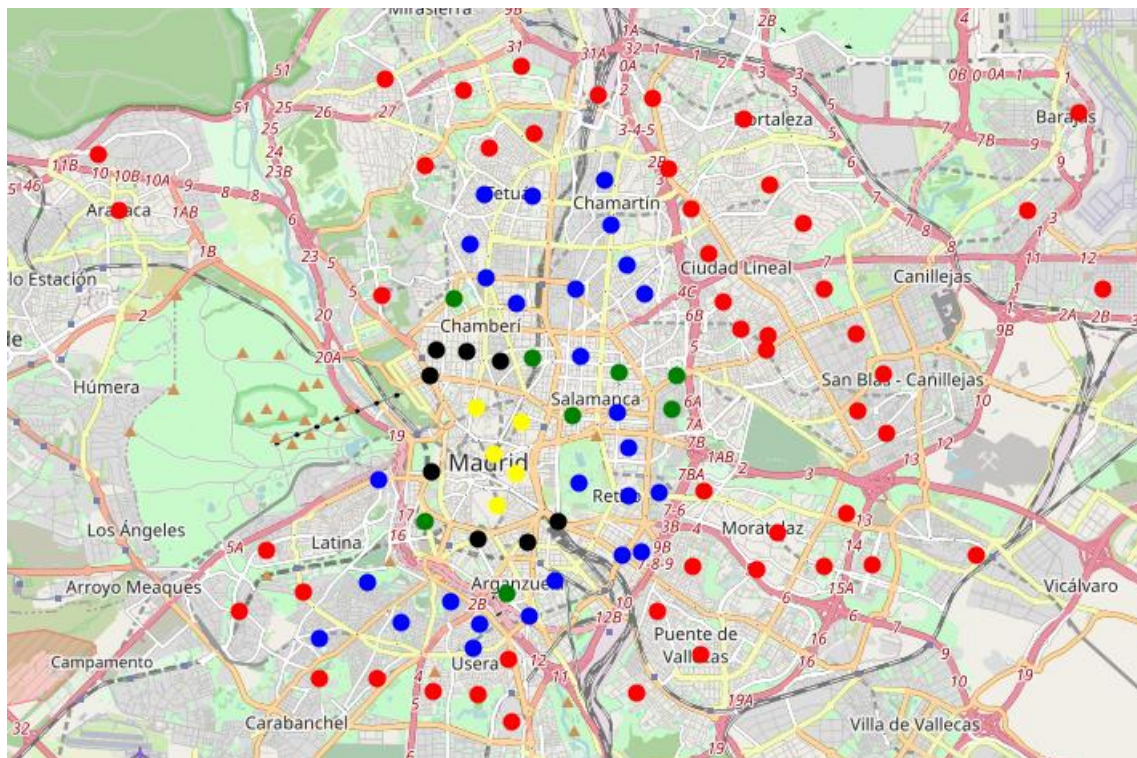
Se observa en esta parte, que no aparecen todos los barrios de la ciudad de Madrid, esto es por que durante el proceso, han aparecido barrios donde no había ni tiendas de música, ni lugares donde se celebren eventos relacionados con las artes escénicas.

Estos barrios se han descartado para el posterior análisis, ya que no son de utilidad.

Análisis

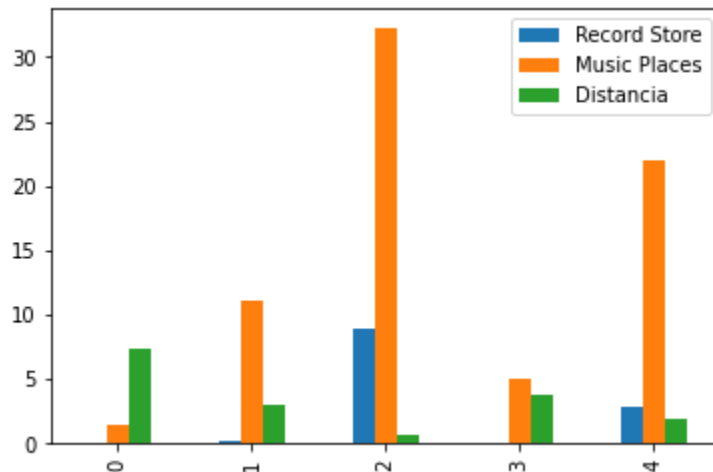
Con los datos obtenidos en el anterior apartado, se aplica el método K-Means para la agrupación de clusters. Se ha optado para la especificación de 5 clusters, considerando que estos pueden ser suficientes para una correcta agrupación de los barrios.

Este es el resultado obtenido:



Se observa claramente que la agrupación es consistente con los datos previamente observados. Los distintos clústers se diferencian por su distancia al centro de la ciudad.

	Record Store	Music Places	Distancia
0	0.040000	1.520	7.425796
1	0.250000	11.125	3.040322
2	9.000000	32.200	0.740281
3	0.037037	5.000	3.721949
4	2.875000	22.000	1.860267



Realizamos un análisis preliminar por cluster:

- **Cluster 0:** Es una zona que podemos descartar de antemano. Tiene una baja concentración de tiendas de discos, pero son barrios muy alejados del centro y con pocos lugares donde se realicen eventos de artes escénicas
- **Cluster 1:** Es una zona con muchos lugares donde se realizan espectáculos de artes escénicas y hay pocas tiendas de música. La desventaja que tiene es que la distancia al centro no es mucha, pero no es la más cercana
- **Cluster 2:** Es una zona con muchos lugares donde se realizan espectáculos de artes escénicas y muy céntrica (Esta zona se compone únicamente de los cinco barrios más céntricos de la ciudad). Como contra tiene es que también es que es la zona con mayor concentración de tiendas de música
- **Cluster 3:** Es una zona que también podríamos descartar en una primera aproximación. Es el segundo cluster más alejado del centro, y también es el segundo con menos lugares donde se celebren artes escénicas
- **Cluster 4:** Es una zona interesante, ya que se cumplen los requisitos buscados: Cercanía al centro, gran concentración de lugares con artes escénicas, y pocas tiendas de música

Con este análisis preliminar podríamos descartar los barrios de los clusters 0 y 3 por su gran distancia al centro de la ciudad. Y los barrios del cluster 2, que son los barrios más céntricos, pero también los que tienen una mayor concentración de tiendas de música

Con esto únicamente nos queda optar por algún barrio del clúster 1 o 4. Empecemos analizando el clúster 4 que es el más cercano al centro:

	Barrio	Comedy Club	Concert Hall	Indie Theater	Jazz and Blues Venue	Music Venue	Performing Arts Venue	Record Store	Rock Club	Theater	Latitud	Longitud	Distancia	Cluster	Music Places
1	Acacias	2	9	3	0	15	1	1	0	20	40.404075	-3.705957	1.433996	4	24
14	Atocha	0	5	0	1	11	0	1	0	21	40.406975	-3.688582	1.965713	4	17
64	Palos de Moguer	1	9	1	1	13	0	1	0	25	40.403638	-3.695287	1.706318	4	23

Podemos ver que el barrio más propicio para emplazar nuestra nueva tienda es Acacias, ya que tiene una única tienda de música, y 24 lugares donde se realizan eventos relacionados con música, y por último sólo está a 1.43 km del centro de la ciudad.

Veamos con detalle la información del clúster 1

	Barrio	Record Store	Music Places	Distancia
4	Almagro	0	17	2.131744
22	Chopera	0	6	2.438248
38	Fuente del Berro	0	6	4.446719
45	Imperial	1	17	1.896203
50	Lista	0	8	3.486656
75	Recoletos	1	23	2.150034
93	Vallehermoso	0	6	3.051762
94	Ventas	0	6	4.721211

Vemos que en esta zona ninguno de los barrios mejora la predisposición del barrio de Acacias que vimos en el punto anterior. Hay 3 barrios que podrían ser viables, pero no tienen tan buena predisposición como Acacias:

- Almagro
- Imperial
- Recoletos

Conclusiones

Como resultado de nuestro analisis a la hora de encontrar la mejor ubicación para una nueva tienda de música, obtenemos que el barrio de **Acacias** es el que mejor se adapta a las especificaciones que se detallan en el primer punto de este informe.

Aun así, se adjuntan al resultado final los 4 barrios con mejor predisposición a ubicar la tienda:

1. Acacias
2. Imperial
3. Almagro
4. Recoletos