

Proyecto Capstone - La Batalla de los vecindarios

Ciencia de Datos IBM - Coursera

Andreina Peraza

04/05/2022

1. Antecedentes

La industria cosmética como se desarrolla a nivel mundial es un contaminante ambiental, pues conlleva al uso de productos químicos como sulfatos, alcoholes, parabenos, siliconas, entre otros que a la larga son dañinos tanto para el cuerpo como para el medio ambiente es por ello que se hace imprescindible desarrollar alternativas sustentables y crear conciencia de que el planeta que habitamos merece tanto amor como nosotros.

Por otro lado, la mala alimentación es uno de los grandes problemas de la vida moderna, debido a las prisas y a los horarios complicados se ha generado que la comida chatarra gane terreno y su exceso en grasas saturadas, sal, azúcar, etc., puedan dañar nuestra salud. Según datos y cifras de la Organización Mundial de la Salud publicadas en su página web, desde 1975 la obesidad se ha casi triplicado en todo el mundo, en 2016 más de 1900 millones de adultos de 18 o más años tenían sobrepeso, de los cuales más de 650 millones eran obesos y el 39% de las personas adultas de 18 o más años tenían sobrepeso. La mayoría de la población mundial vive en países donde el sobrepeso y la obesidad se cobran más vidas de personas que la insuficiencia ponderal

Por lo tanto, es beneficioso tanto para la población como para el planeta migrar hacia alternativas cosméticas eco-amigables y hacia un estilo de alimentación saludable que nos aporte una mejor calidad de vida integral. Es por ello que el nacimiento de emprendimientos y empresas con modelos económicos que persiguen otros fines además de la rentabilidad financiera, como lo son generar un impacto social a través de acciones en determinadas comunidades y prácticas medioambientales que contribuyen a la mitigación del cambio climático son necesarias para enfrentar los problemas medioambientales y sanitarios que vivimos actualmente y que amenaza nuestra existencia de cara al futuro

1.2 Problema

La empresa Novaterra desea iniciar sus operaciones abriendo una tienda tanto física como virtual, dicha empresa desea realizar varias actividades económicas entre las que destacan: extracción de aceites esenciales, fabricación y venta de productos

cosméticos naturales y compra y venta de productos orgánicos saludables. El tren ejecutivo de la empresa desea conocer cuál es la mejor ubicación para establecer su tienda física y a qué tipo de personas de la ciudad de Montevideo, Uruguay puede dirigir su campaña publicitaria en redes sociales para captar sus primeros clientes.

La tienda física deberá estar ubicada en las adyacencias de los vecindarios donde la población tenga un nivel adquisitivo medio-alto, cercana a centros de bellezas, spa, restaurantes de comida saludable y si es posible cerca de espacios donde se pueda disfrutar de un ambiente natural.

Este proyecto tiene como objetivo seleccionar la mejor ubicación para establecer la tienda física y determinar a qué tipo de personas se debe dirigir la campaña publicitaria.

2. Fuente de los datos.

Los datos utilizados para la elaboración del proyecto se pueden encontrar en las siguientes fuentes:

- División geográfica de Uruguay de acuerdo a sus municipios y vecindarios:
https://es.wikipedia.org/wiki/Anexo:Barrios_de_Montevideo
- Datos públicos del último censo poblacional realizado por el Instituto Nacional de Estadísticas del Uruguay en el año 2011.
https://www.ine.gub.uy/c/document_library/get_file?uuid=4718ab8a-c64e-439c-ba8e-7d56371b6c89&groupId=10181
- Datos de ingresos percibidos por las personas en la ciudad de Montevideo, así como el nivel socioeconómico de cada hogar extraído del Instituto Nacional de Estadísticas del Uruguay en el año 2017:
https://www.ine.gub.uy/c/document_library/get_file?uuid=75b6cb00-387e-40ee-a694-eb64314e81db&groupId=10181

Con estos conjuntos de datos se pretende extraer la información de los municipios y barrios que constituyen la ciudad de Montevideo, así como también conocer como se distribuye la población y los ingresos económicos por hogar en estos vecindarios. Luego se realizará una búsqueda de los sitios de interés en estas zonas utilizando la API de Foursquare.

3- Metodología de análisis

Para el desarrollo del presente proyecto se procedió a elaborar la limpieza, unión y análisis de los diferentes conjuntos de datos utilizados para posteriormente con las respectivas visualizaciones adecuadas llegar a las conclusiones que se mostrarán más adelante en este informe.

Una vez extraída la tabla con la información de los barrios y municipios de la ciudad de Montevideo de la página de Wikipedia del primer conjunto de datos que se menciona en el ítem anterior, se eliminaron las columnas que no eran de interés para el análisis y se les dio nombres apropiados a las columnas de interés. Seguidamente se ubicaron las coordenadas geográficas de cada barrio (longitud, latitud) con la ayuda de la librería Geopy y el módulo Nominatim de Python para luego eliminar las filas de aquellos barrios que no se pudieron geolocalizar.

Por otro lado, del segundo conjunto de datos se extrajo la cantidad de personas que habitan cada barrio de la ciudad de Montevideo. Del mismo modo se eliminaron filas y columnas que no se necesitaba para el proyecto, se eliminaron duplicados y filas con valores nulos y se renombraron las columnas para luego consolidar esta información en un solo dataframe uniéndolo con la información del primer conjunto de datos. En este punto nos interesa saber cuáles son los vecindarios en donde se concentra la mayor cantidad de población Tomando en cuenta la cantidad de hombres y mujeres que habitan en dichos vecindarios.

Por último, se extrajeron los datos de los ingresos económicos y nivel socioeconómico de las personas que habitan en cada barrio de Montevideo del tercer conjunto de datos, el proceso de limpieza de este conjunto de datos se realizó de forma similar a los anteriores. Cabe resaltar que de acuerdo con la información suministrada por la fuente estos datos públicos, los niveles socioeconómicos se clasifican de acuerdo con lo siguiente:

- 1 Montevideo - Nivel económico bajo
- 2 Montevideo - Nivel económico medio – bajo
- 3 Montevideo - Nivel económico medio,
- 4 Montevideo - Nivel económico medio - alto,

5 Montevideo - Nivel económico alto

En esta etapa del proceso se busca Conocer cuáles son los municipios, y por ende los vecindarios, en donde se encuentre la mayor concentración de personas con mayor nivel socioeconómico o poder adquisitivo

Por último, una vez conocidos los barrios y municipios que se ajustan a los criterios requeridos, se ubicaron aquellos barrios que contaban con sitios asociados a la actividad económica que la empresa Novaterra desea llevar a cabo, es decir aquellos vecindarios en donde se ubicaran espacios como peluquerías, salones de belleza, spa, parques con zonas verdes, restaurantes de comida saludable y/o vegetariana/vegana otros tipos de restaurantes. Lo cual se ejecutó realizando llamadas a la API de FourSquare

Una vez seleccionados los sitios de interés, procederemos a realizar clustering usando K-means tal y como se nos fue enseñado en el curso. Se elegirá aquel barrio q posea la mayor población con buenos ingresos económicos y con sitios asociados al cuidado y belleza integral.

De esta manera, este informe contiene dos etapas claras y definidas, a saber: La selección primaria de los municipios y vecindarios usando datos socioeconómicos y de población, y el análisis de dichas zonas utilizando Foursquare y clustering aplicando el método de agrupamiento K-means.

4- Resultados

De acuerdo a la fig. N° 1 vemos que el barrio con mayor población es Pocitos (municipio CH), seguido por Cóndor(B), Unión (D, E y F), La Paloma (A) y Buceo (CH y E). De igual forma la fig. N° 2 nos muestra la distribución poblacional en cuanto a género de los barrios mencionados anteriormente, pudiéndose notar que en todos estos vecindarios predomina una mayor concentración del sexo femenino sobre todo en Pocitos, lo cual es ideal para los objetivos comerciales de la compañía.

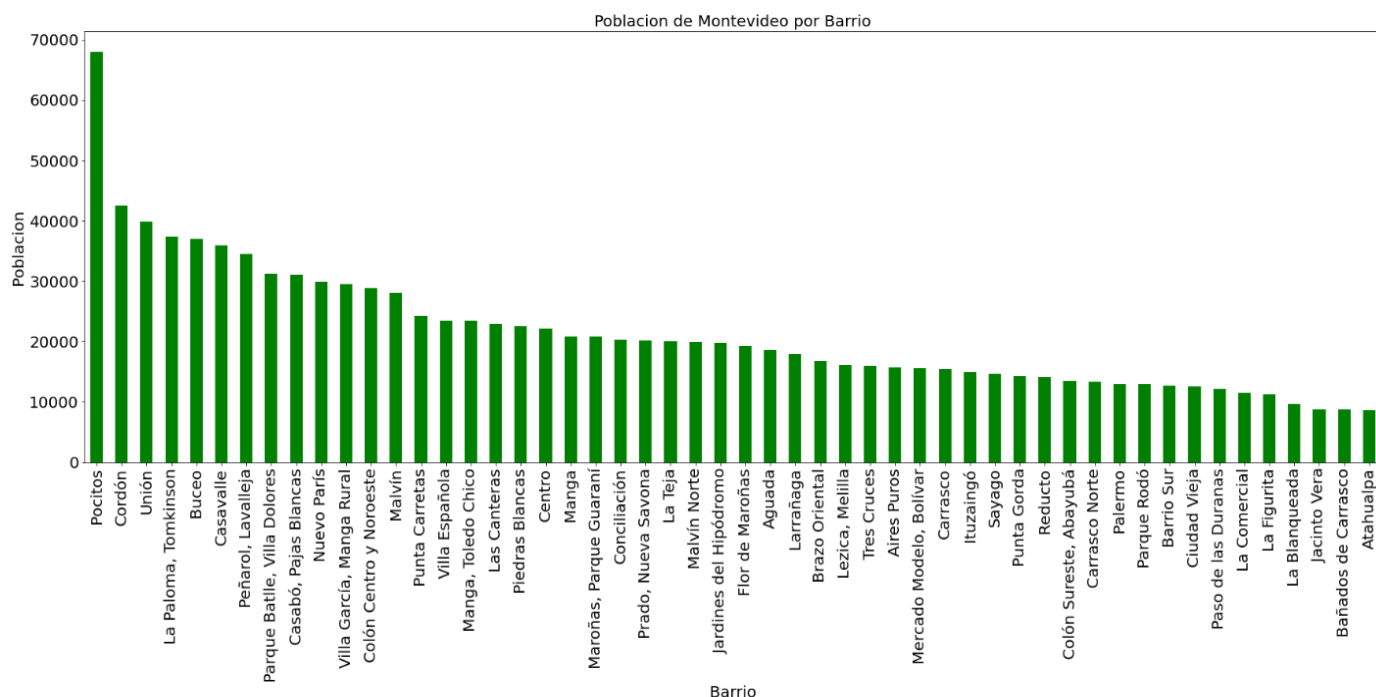


Fig. Nº 1: Población de Montevideo organizada por barrios

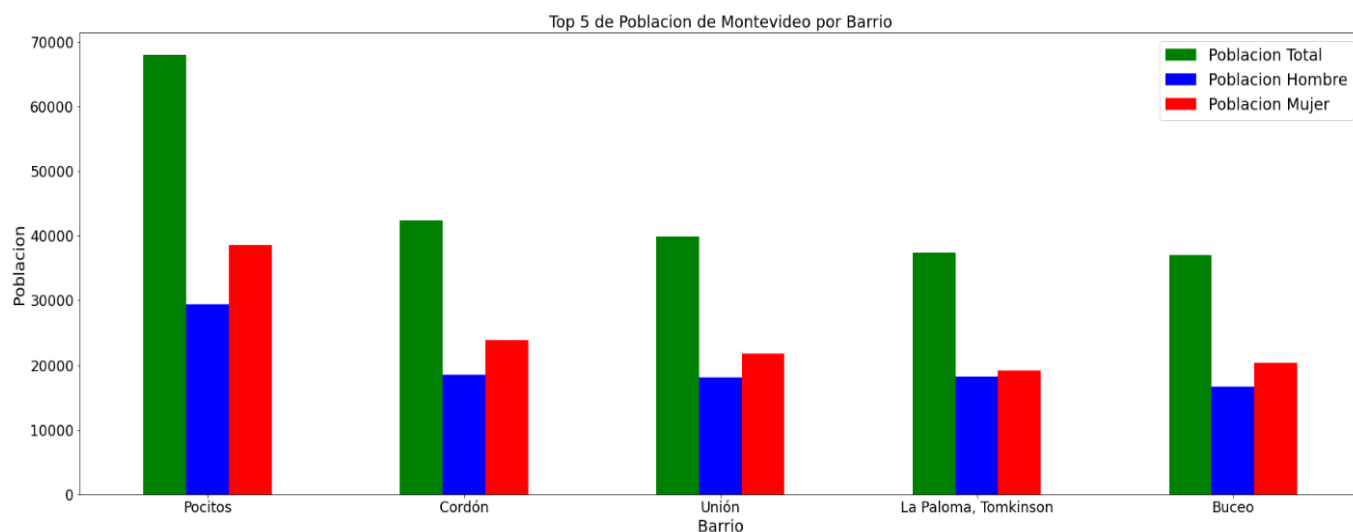


Fig. Nº 2: top 5 de población de Montevideo por barrio distribuida por género

Ahora bien, los municipios en esta ciudad están compuestos de diferentes barrios por lo que se hizo interesante analizar la cantidad de personas que habitan en Montevideo por municipios. La fig. Nº 3 contempla la población distribuida por municipios, en la cual se puede apreciar que los municipios F, G, B, CH y E son los más poblados. En la fig. Nº 4 se observa la cantidad de personas que habitan en los

municipios anteriormente mencionados desglosados por género y en donde de forma similar a la Fig. N° 2 se puede notar que hay una mayor presencia del sexo femenino, Por otro lado, también podemos notar que los municipios B, CH y E están dentro de los 5 municipios más habitados dentro de Montevideo los cuales contienen algunos de los barrios más poblados de la ciudad.

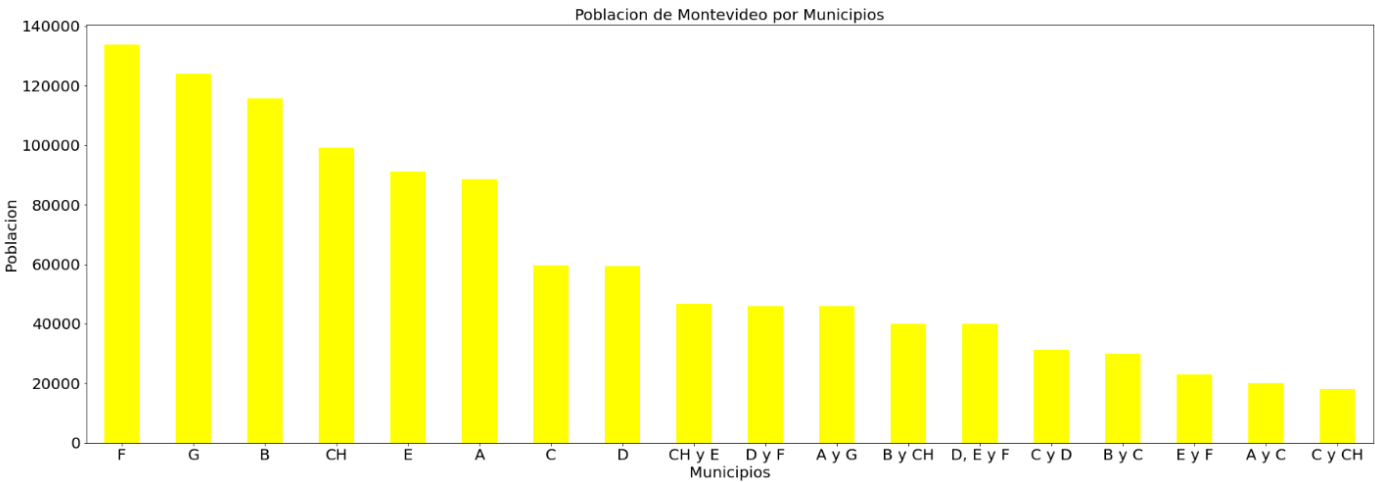


Fig. N° 3: Población de Montevideo por municipios

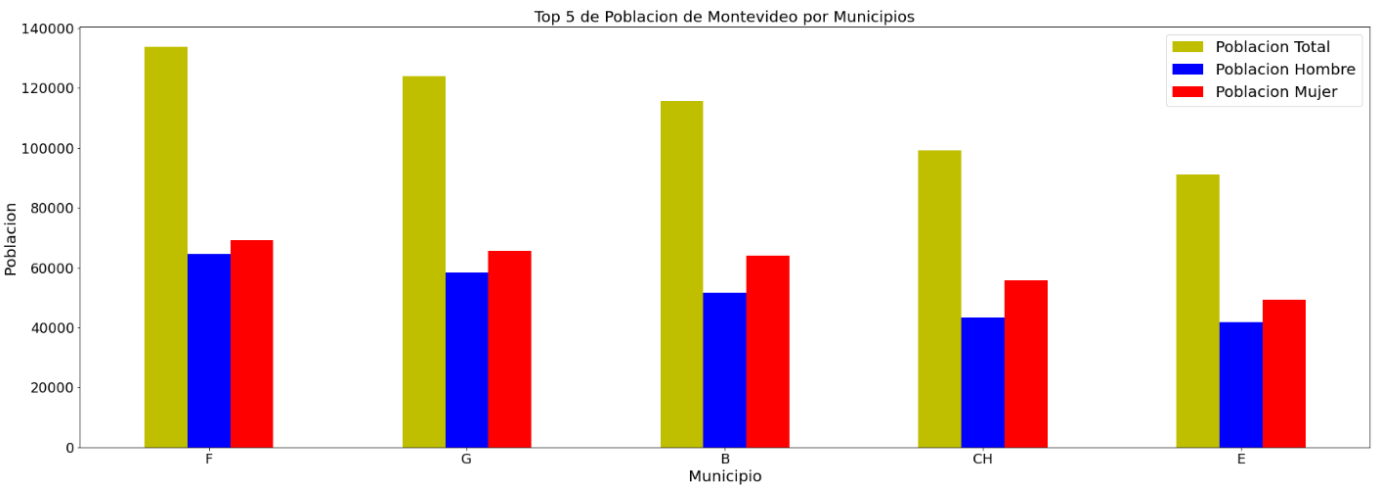


Fig. N° 4: Top 5 de la población de Montevideo por municipios distribuida por género

A continuación, se muestra una serie de diagramas de barras en las que puede notarse el nivel socioeconómico e ingresos percibidos por los habitantes de Montevideo, tanto a nivel general es decir en toda la ciudad, como cuales son las zonas en donde se concentra los mayores ingresos por hogar y mejor estrato social.

Evidentemente es de esperarse que exista una relación directamente proporcional entre la cantidad de ingresos percibidos por las personas y su nivel socio económico

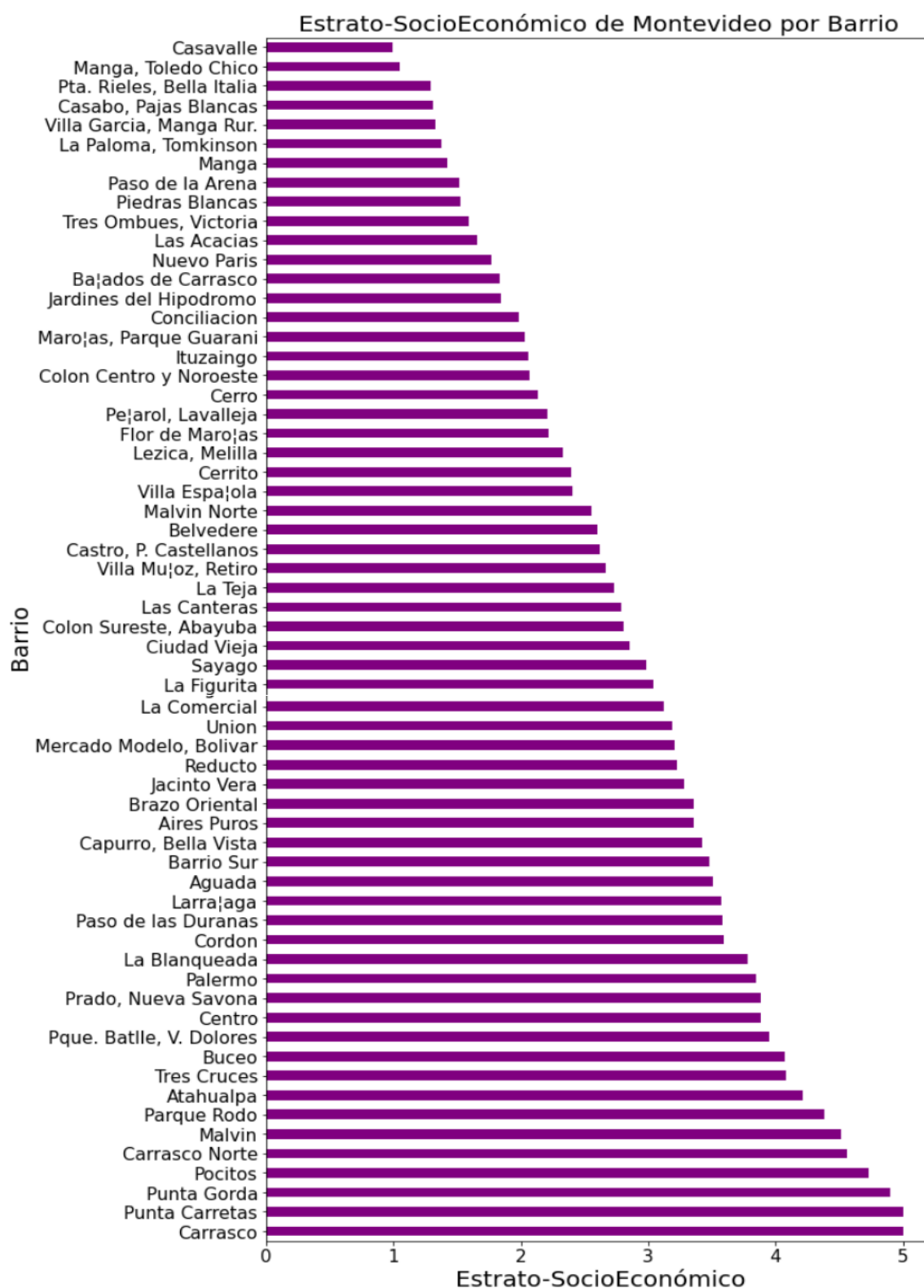


Fig. N° 5: estrato socioeconómico de Montevideo distribuido por barrios

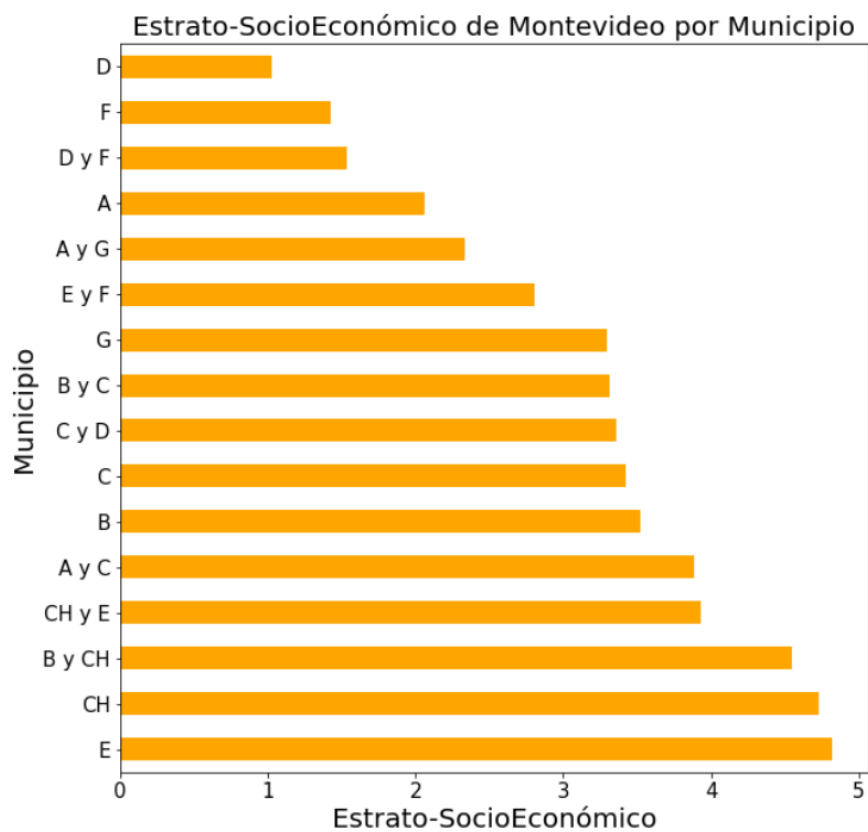


Fig. Nº 6: estrato socioeconómico de Montevideo distribuido por municipios

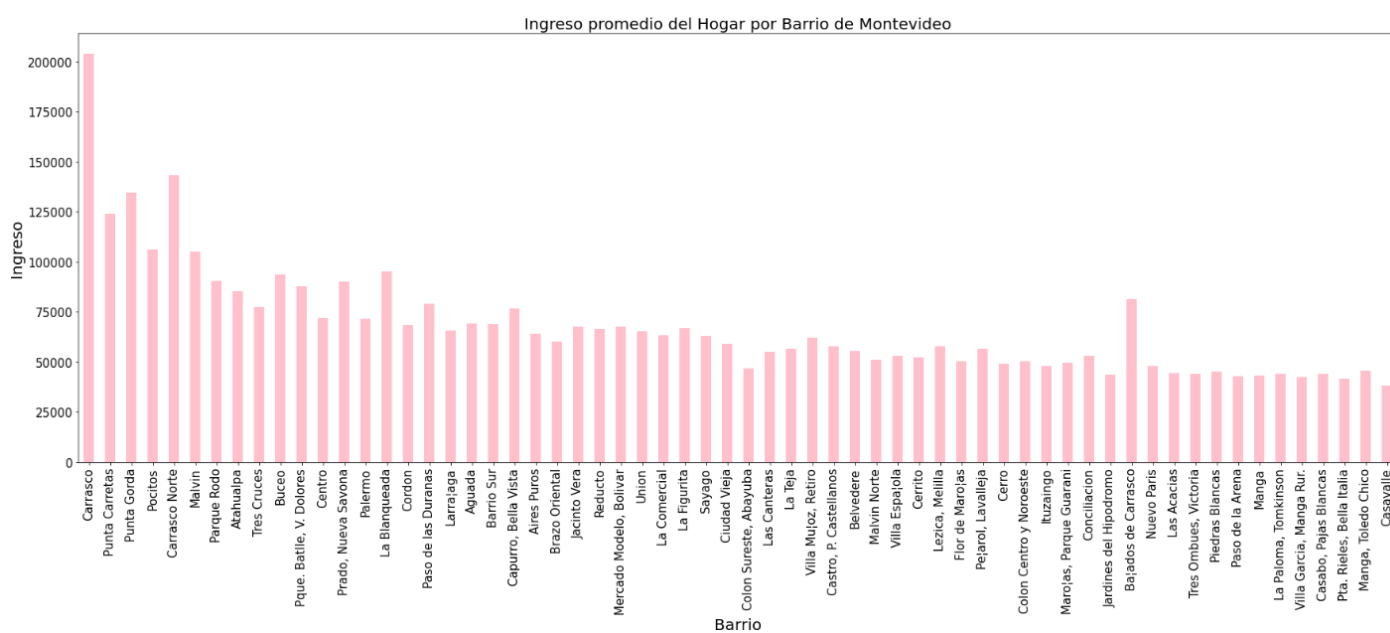


Fig. Nº 7: Ingresos económicos promedio del hogar de Montevideo distribuido por barrios

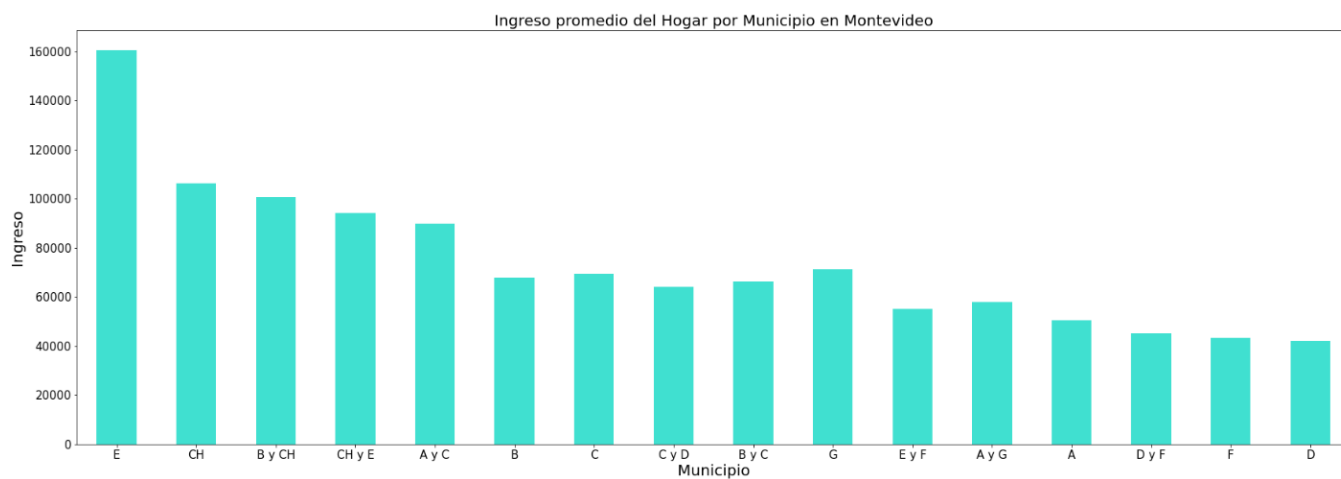


Fig. N° 8: Ingresos económicos promedio del hogar de Montevideo distribuido por municipios

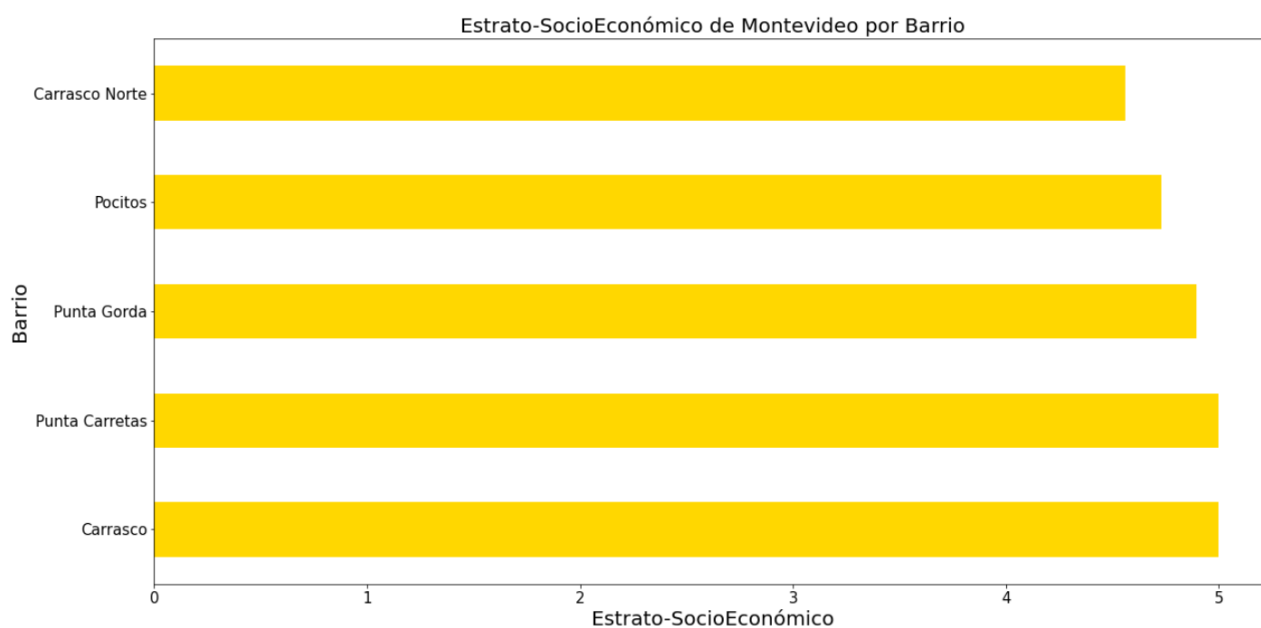


Fig. N° 9: Ingresos económicos promedio del hogar de Montevideo distribuido por barrios



Fig. N° 10: Top 5 de nivel socioeconómico de los habitantes de Montevideo distribuido por municipios

Podemos concluir en base a los diagramas anteriores que los vecindarios de Carrasco (E), Punta Carretas (B y CH), Punta Gorda (E), Pocitos (CH) y Carrasco Norte (E) son los barrios con mayor nivel socioeconómico e ingresos percibidos, mientras que los municipios con mejor posicionamiento económico son: E, CH, B y CH, CH y E, A y C.

Por otro lado, de acuerdo a la metodología aplicada para agrupar o segmentar los barrios de acuerdo al método de k-means, se puede notar que los vecindarios más apropiados para instalar una tienda física son los que pertenecen al grupo o cluster 0 debido a que poseen una amplia variedad de sitios de interés como parques y plazas con áreas verdes, restaurantes tanto de comida saludable como de comida no tan saludable (teniendo esta mayor presencia), supermercados y tiendas para el público femenino. En la figura N° 11 se muestran los resultados obtenidos de la agrupación, en donde se observan los barrios sugeridos de acuerdo a lo comentado anteriormente, sin embargo, los más recomendados teniendo en cuenta los análisis anteriores de población y nivel socioeconómico son Condor, Punta Carretas, Pocitos, Carrasco

Finalmente, tomando en cuenta todo el análisis realizado en este proyecto lo cual involucró tomar en cuenta aspectos como factores demográficos, económicos y geográficas podemos concluir que el mejor barrio para la ubicación de la tienda física de NovaTerra es Pocitos debido a que posee la mayor población de la ciudad de Montevideo representada mayormente por el sexo femenino, se ubica en la posición número cuatro de los barrios con mayor nivel socioeconómico y pertenece al municipio CH (segundo municipio con mayores ingresos económicos y el cuarto con mayor población). Además, posee lugares donde las personas de otros barrios y municipios tienden a concurrir como restaurantes, supermercados, parques y plazas.

[65]: mv_barrios_merged.loc[mv_barrios_merged['Cluster Labels'] == 0, mv_barrios_merged.columns[[0] + list(range(5, mv_barrios_merged.shape[1]))]]

[65]:

	Barrios	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
0	Ciudad Vieja	Restaurant	Plaza	South American Restaurant	Comfort Food Restaurant	Vegetarian / Vegan Restaurant	Tapas Restaurant	Swiss Restaurant	Sushi Restaurant	Southern / Soul Food Restaurant	Latin American Restaurant
1	Centro	Restaurant	Plaza	Chinese Restaurant	Market	Latin American Restaurant	Supermarket	Spanish Restaurant	Kebab Restaurant	Italian Restaurant	American Restaurant
2	Barrio Sur	American Restaurant	Smoothie Shop	Vegetarian / Vegan Restaurant	Restaurant	Modern European Restaurant	Night Market	Park	Plaza	Seafood Restaurant	South American Restaurant
3	Cordón	South American Restaurant	Plaza	Fast Food Restaurant	Flea Market	Restaurant	American Restaurant	Night Market	Park	Seafood Restaurant	Smoothie Shop
4	Palermo	Restaurant	South American Restaurant	Flea Market	Italian Restaurant	American Restaurant	Night Market	Park	Plaza	Seafood Restaurant	Smoothie Shop
5	Parque Rodó	Restaurant	American Restaurant	Flea Market	Italian Restaurant	Supermarket	Food	Grocery Store	Tapas Restaurant	Empanada Restaurant	Swiss Restaurant
6	Punta Carretas	Restaurant	Mediterranean Restaurant	Italian Restaurant	Women's Store	Night Market	Flea Market	Modern European Restaurant	Fast Food Restaurant	Mexican Restaurant	Sushi Restaurant
7	Pocitos	Italian Restaurant	Supermarket	South American Restaurant	Restaurant	Empanada Restaurant	Falafel Restaurant	Middle Eastern Restaurant	Mexican Restaurant	Chinese Restaurant	Market
9	Malvin	Park	Restaurant	American Restaurant	Mexican Restaurant	Modern European Restaurant	Night Market	Plaza	Seafood Restaurant	Smoothie Shop	South American Restaurant
12	Carrasco	Restaurant	American Restaurant	Plaza	Comfort Food Restaurant	Park	Chinese Restaurant	Market	German Restaurant	Gourmet Shop	South American Restaurant

Fig. Nº 11: Barrios del grupo Nº 0

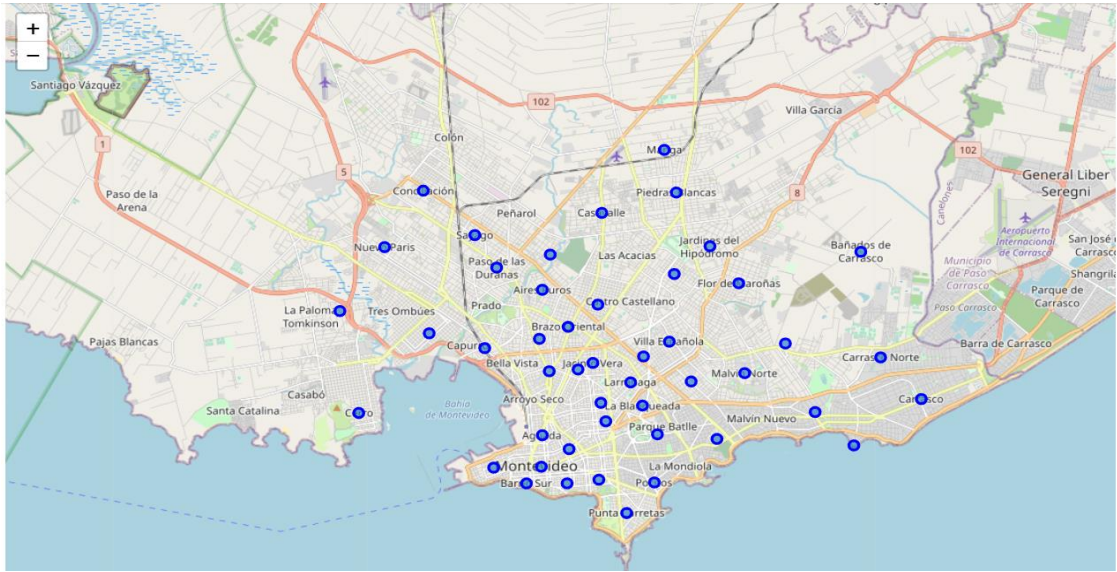


Fig. Nº 12. Ubicación geográfica de los barrios de Montevideo

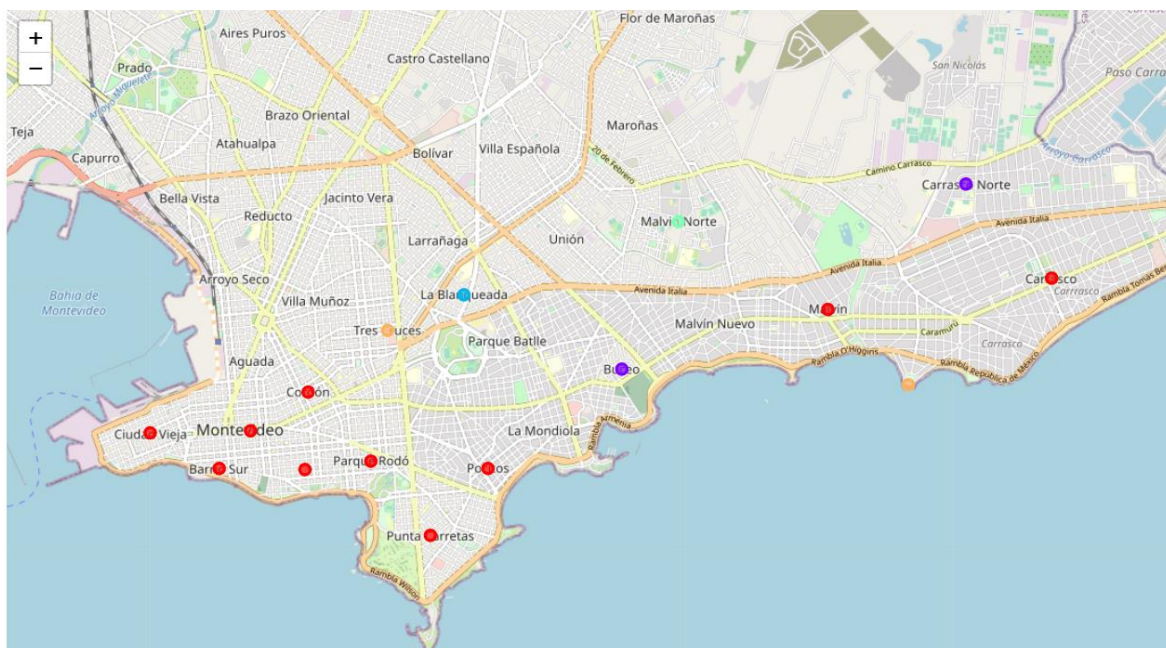


Fig N° 13. Agrupaciones obtenidas luego de aplicar clustering

5- Recomendaciones

- De acuerdo a fuentes oficiales se tiene conocimiento de que para el año 2023 se tendrá un nuevo censo de población del Uruguay, por lo cual se recomienda realizar de nuevo este estudio y/o procedimiento con datos actualizados en caso de que todavía no se abra la tienda física o en caso de que se quiera abrir otra sucursal.
- De igual forma es recomendable buscar fuentes más actualizadas de los ingresos y estado socioeconómico de las personas que habitan en la ciudad de Montevideo.
- Se recomienda aplicar otros algoritmos de clasificación de aprendizaje automático para validar el resultado obtenido con el k-means o para tener otros resultados para decidirse por el más conveniente según sea el caso.
- Se recomienda en primera instancia dirigir la publicidad a los barrios de los primeros 5 municipios con mayor nivel socioeconómico y población a través las redes sociales para ir captando la mayor cantidad de clientes posibles en

estos sectores y ver cómo responde la población ante la propuesta antes de pensar en abrir una nueva sucursal física y luego seguir con las zonas de niveles socioeconómicos más bajos, procurando para estos últimos ofrecer descuentos y ofertas en los productos comercializados.

6- Conclusiones

El vecindario más apropiados y con mayor potencial para instalar la tienda física de la empresa NovaTerra dedicada a la comercialización de productos cosméticos elaborados a base de plantas y con materia prima de origen natural, así como productos de alimentación saludables y orgánico es Pocitos, ubicado dentro del grupo cero del método de clasificación K-means, y está dentro del municipio CH con buena parte de la población concentrada allí, con buen nivel socioeconómico y cercanía a lugares como parques, plazas, supermercados, restaurantes, etc.

Los municipios y vecindarios más apropiados para realizar publicidad a través de redes sociales son todos aquellos que posean una buena población con estrato socioeconómico medio-alto, haciendo énfasis a la población femenina por ser aquellas que por lo general se preocupan más del cuidado, belleza y salud integral tanto de sí mismas como de sus familiares. Por lo que sería buena estrategia segmentar estas personas de acuerdo a sus actividades en las redes e intereses mostrados en ellas.