**Curso Ciencia de Datos IBM - Coursera**

**Tarea final – Proyecto Coursera Capstone**

**Titulo : La Batalla de los barrios (vecindarios) en Montevideo.**

**Tarea Semana 1 ( semana 4 del curso)**

1. **Introducción y descripción del caso.**

El caso a analizar refiere a una empresa multinacional dedicada al rubro gimnasios, centros deportivos y comercialización de artículos deportivos, accesorios y vinculada a una cadena de comida saludable, que desea conocer la mejor ubicación para instalar una nueva sucursal en la ciudad de Montevideo, Uruguay.

Dicha multinacional ya posee ubicaciones en otras principales ciudades de Lationamérica.

La ubicación deberá estar en alguna de las zonas o barrios de Montevideo con mayor población y mejor nivel socioeconómico, así como estar cercana a espacios abiertos, como zonas costeras, playas y parques.

Montevideo cuenta con una franja costera conocida como la Rambla, de más de 15 km de extensión, con playas de arena fina sobre el Rio de la Plata.

El objetivo del caso es determinar cuáles zonas, municipios o barrios son los más aptos para la instalación de la sucursal, y que estén dentro cercanos a la zona de playas y parques. Montevideo cuenta con 62 barrios agrupados en 8 municipios.

Fuente :<https://es.wikipedia.org/wiki/Anexo:Barrios_de_Montevideo>, <https://es.wikipedia.org/wiki/Anexo:Municipios_de_Montevideo>

1. **Forma de abordaje y metodología de análisis**

En una primera parte vamos a simplificar el marco de datos para el análisis, seleccionando aquellos municipios que cuentan con las características del público objetivo.

Montevideo posee 8 municipios, y si bien en todos la población está integrada y es heterogénea en cuanto a aspectos socioeconómicos, en algunos municipios predominan los estratos socioeconómicos más altos ( de clase media a clase media alta y clase alta) y en otros predominan los de estratos socioeconómico más bajo. Por lo cual buscaremos preseleccionar aquellos municipios, y dentro de ellos aquellos barrios o vecindarios , con mayor nivel socioeconómico.

Por otro lado, hay municipios con mayor densidad demográfica que otros, generalmente los que tienen mayor acceso a servicios, transporte, y cuentan con mayor historia en la ciudad, por lo cual dese el punto de vista demográfico preseleccionaremos aquellos municipios, y dentro de ellos aquellos barrios, con mayor población.

Finalmente, dado que se pretende contar con espacios al aire libre cercanos, asi como cercanía a la costa, preseleccionaremos aquellos municipios y dentro de ellos aquellos barrios, con mayor cantidad de espacios abiertos como plazas y parques y que visualmente en el mapa se encuentren más cerca de la costa.

Para la preselección socioeconómica tomaremos en cuenta la Encuesta de Hogares del año 2017, del Instituto Nacional de Estadísticas del Uruguay, que cuenta con indicadores para Montevideo, y deglosados por barrios, del nivel socioeconómico de los hogares encuestados.

Para la preselección demográfica tomaremos en cuenta los datos del último Censo Oficial y completo de Población, del año 2011, realizado por el mismo Instituto antes mencionado, que trae la población de Montevideo desglosada por barrios.

Para la preselección demográfica construiremos una tabla de coordenadas geoespaciales usando la librerìa Geopy y y el módulo Nominatim de Python , dado que es el que hemos probado antes que proporciona resultados confiables.

Para asociar los 62 barrios a los 8 municipios usaremos la información proporcionada por el Gobierno departamental de Montevideo y que está publicada tanto en su sitio web como en Wikipedia, de esta última importaremos los datos usando métodos panda de Python. Cabe decir que solamente tomaremos los datos del área urbana, no así del área rural del departamento mencionado.

Luego de completada esta primera parte, y preseleccionados los municipios y barrios, usaremos la API de Foursquare sugerida por el curso, para determinar que sitios de interés, tanto deportivo como gastronómico, posee cada barrio No consideraremos otros sitios de interés comercial, cultural o político.

Determinados los sitos de interés, procederemos a aplicar la metodología vista en el curso para los barrios de Manhathan , Nueva York, para el armado y depuración de los dataframes de Python-panda, y para luego aplicar clustering usando K-means, como se vió en el curso. De los clústeres obtenidos, seleccionaremos aquellos con mayor población (no procederemos a seleccionar por nivel socioeconómico, dado que como señalamos, los municipios son bastante heterogéneos en ese aspecto), y mayor cantidad de sitios vinculados a deportes y espacios libres, así como vinculados al rubro gastronómico. Entendemos que, dado que ya hay otras propuestas o sitios vinculados al rubro de negocios que pretendemos desplegar, la población objetivo es más propensa a aceptar una nueva propuesta integral de deportes, recreación y gastronomía como se proyecta.

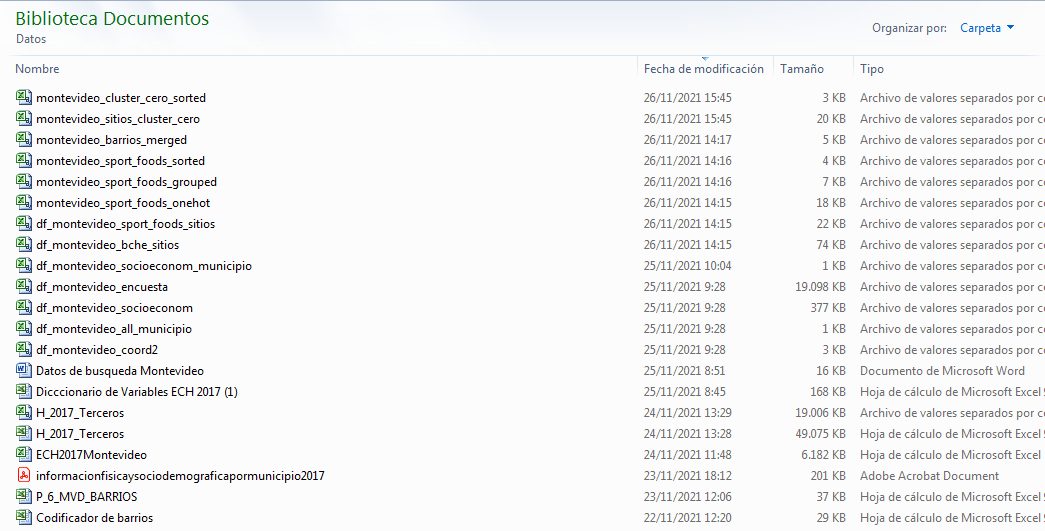
Si fuera necesario hacer una subdivisión de alguno de los clústeres seleccionados, repitiríamos los pasos indicados previamente, construyendo los dataframes y usando K-means.

Finalmente, seleccionaríamos aquel barrio que cuente con mayor cantidad de sitios similares al que queremos proyectar, y con mayor población y cercanía a la costa.

Por lo cual el trabajo propuesto cuenta de dos partes: una de preseleccioón usando dataframes y pandas, y otra de análisis de los municipios preseleccionados, usando Foursquare, y clustering con K-means.

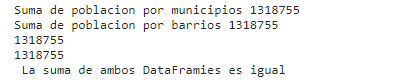
1. **Datos utilizados.**

Detallamos los archivos de datos utilizados que se encuentran en el repositorio de GitHub compartido.

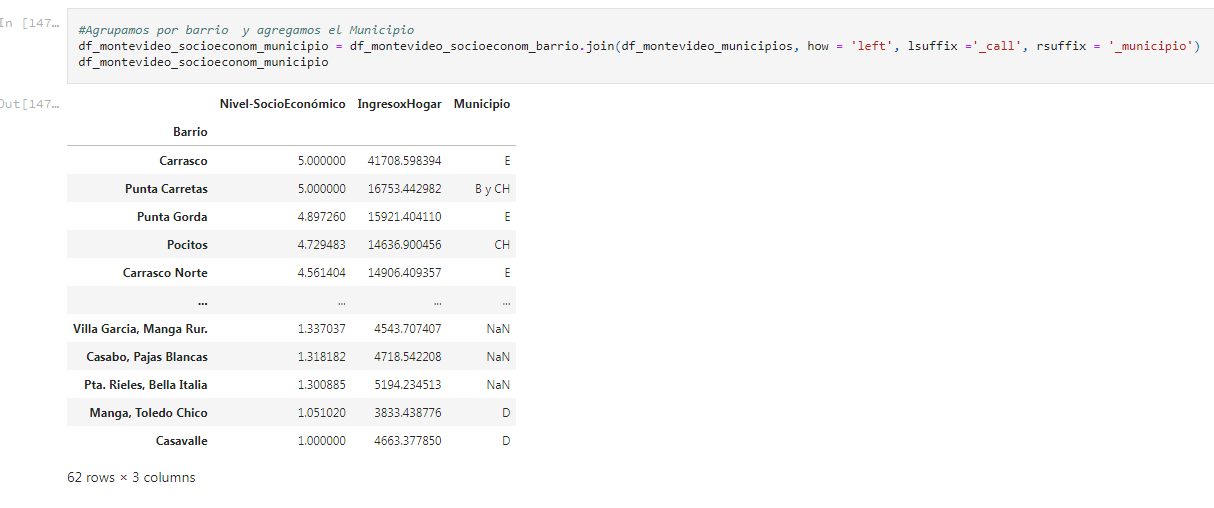


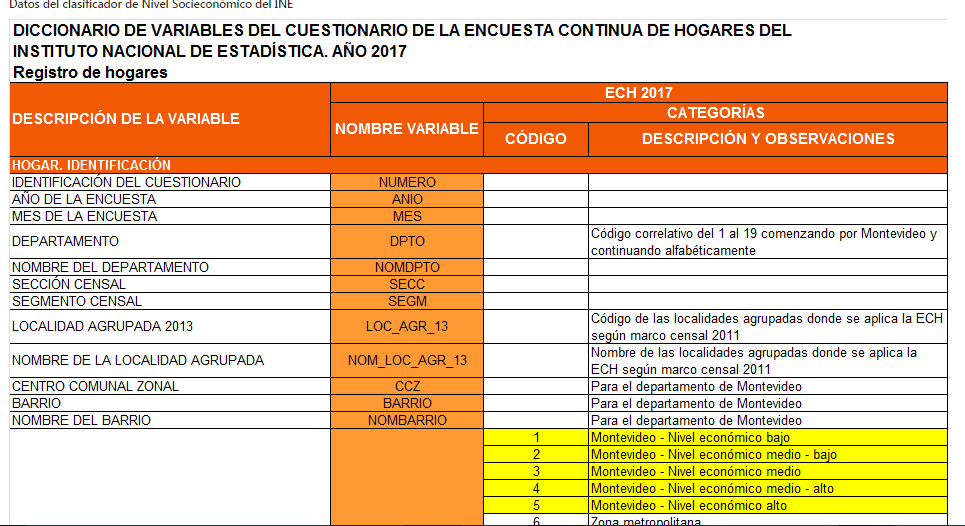
**3.1- Datos de población**



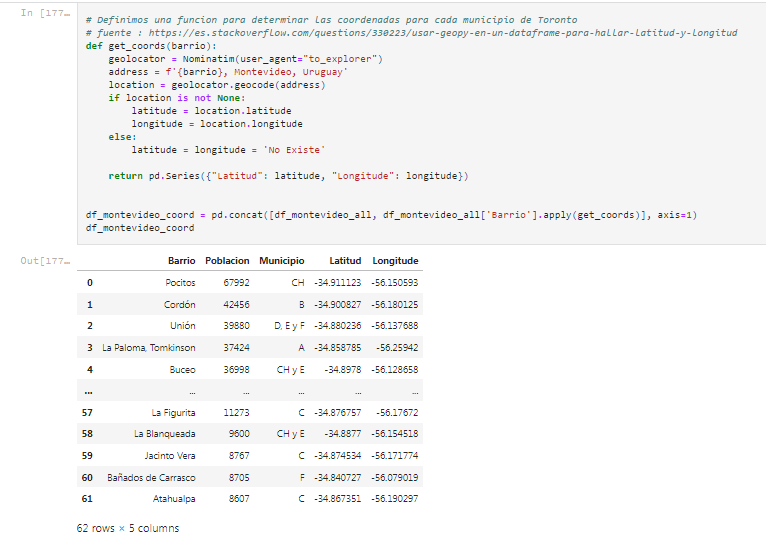


**3.2 – Datos Socioeconómicos**





**3.3 Datos Geográficos y coordenada.**





27-11-2021

HB.