

Análisis Google Places

En este segmento se utilizó la API de Google Places

(<https://developers.google.com/places/?hl=es-419>) para obtener información adicional sobre las propiedades provistas por Properati. Esta API permite entre otras cosas buscar sitios (definidos en esta API como establecimientos, ubicaciones geográficas o puntos de interés destacados) dentro de un área definida, como los límites de un mapa o alrededor de un punto fijo. Para poder utilizarla sin embargo, se necesitan de coordenadas del tipo latitud y longitud.

El set de datos provistos contiene algunas ubicaciones en este formato pero no todas. Por suerte dentro del set de datos está la información denominada como "geoname_id". Este tipo de identificación pertenece a una base de datos geográfica con más de 10 millones de entradas únicas. Mediante la base de datos completa descargada de Kaggle (<https://www.kaggle.com/geonames/geonames-database>) se completaron las latitudes y longitudes faltantes para así tener una muestra completa mayor.

Para realizar pedidos a Google Places es necesario tener una API Key que esencialmente controla el tráfico diario de pedidos a la API. Esta limitación, utilizando una cuenta prioritaria, es de 150000 pedidos por día. A la vez el tiempo entre que se realiza un pedido y se obtiene su respuesta es muy alto. Es por esto que se limitaron las búsquedas a las siguientes categorías (cantidades por propiedad en un radio de 400 metros):

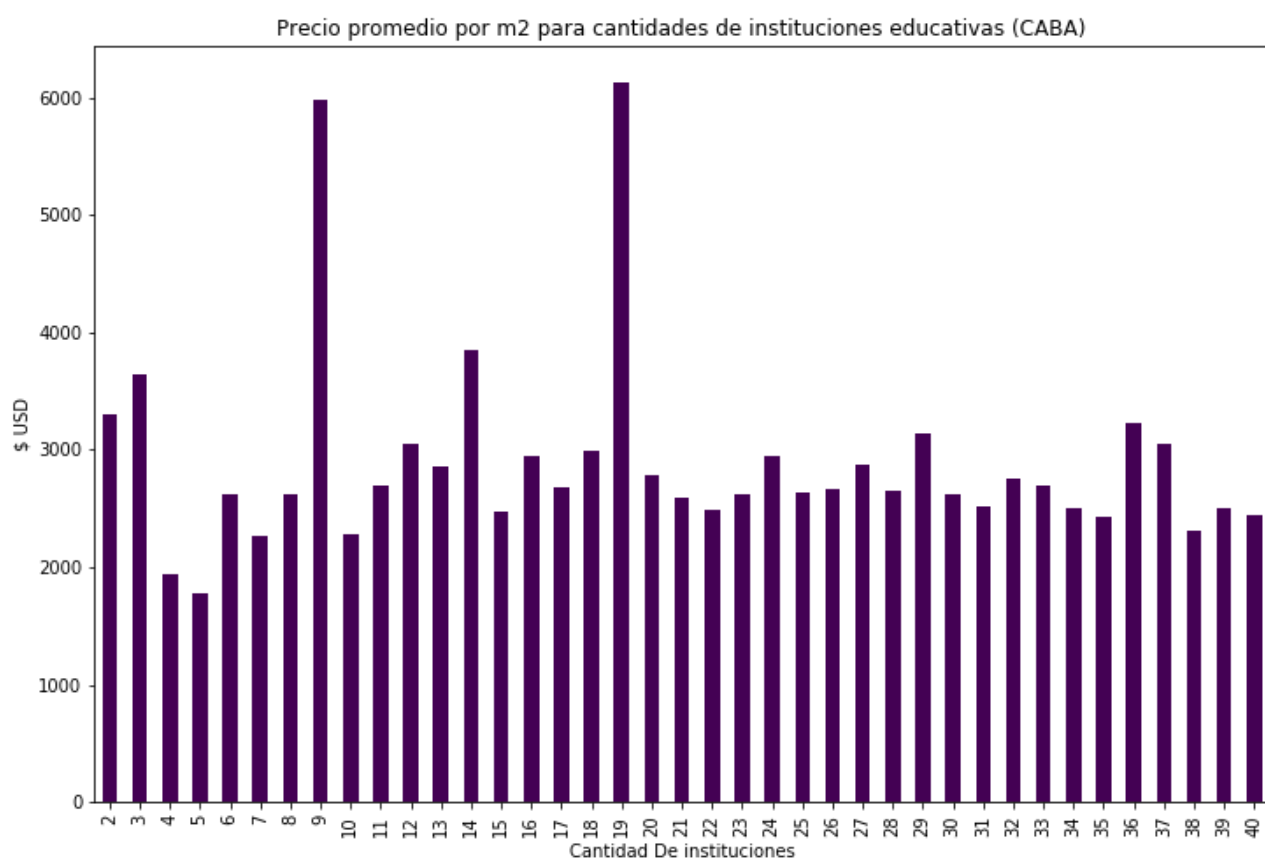
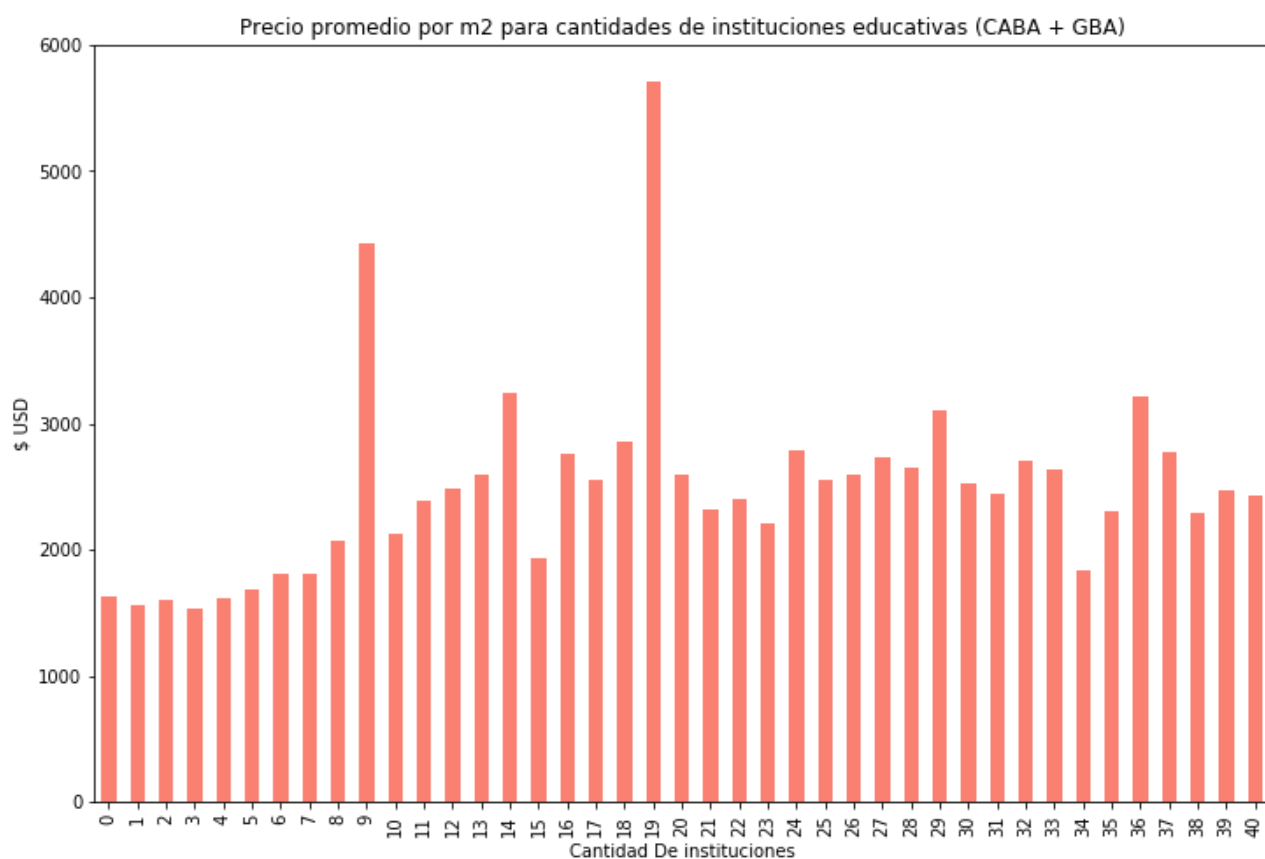
- Locales de tipo gastronómico. Categorías de Google Places: FOOD, BAKERY, BAR, MEAL_TAKEAWAY, MEAL_DELIVERY, RESTAURANT.
- Instituciones educativas. Categorías de Google Places: SCHOOL, UNIVERSITY.
- Puntos de interés cultural. Categorías de Google Places: ART_GALLERY, MUSEUM, PLACE_OF_WORSHIP.
- Espacios verdes. Categorías de Google Places: PARK, NATURAL_FEATURE.
- Estaciones de transporte público. Categorías de Google Places: BUS_STATION, SUBWAY_STATION.

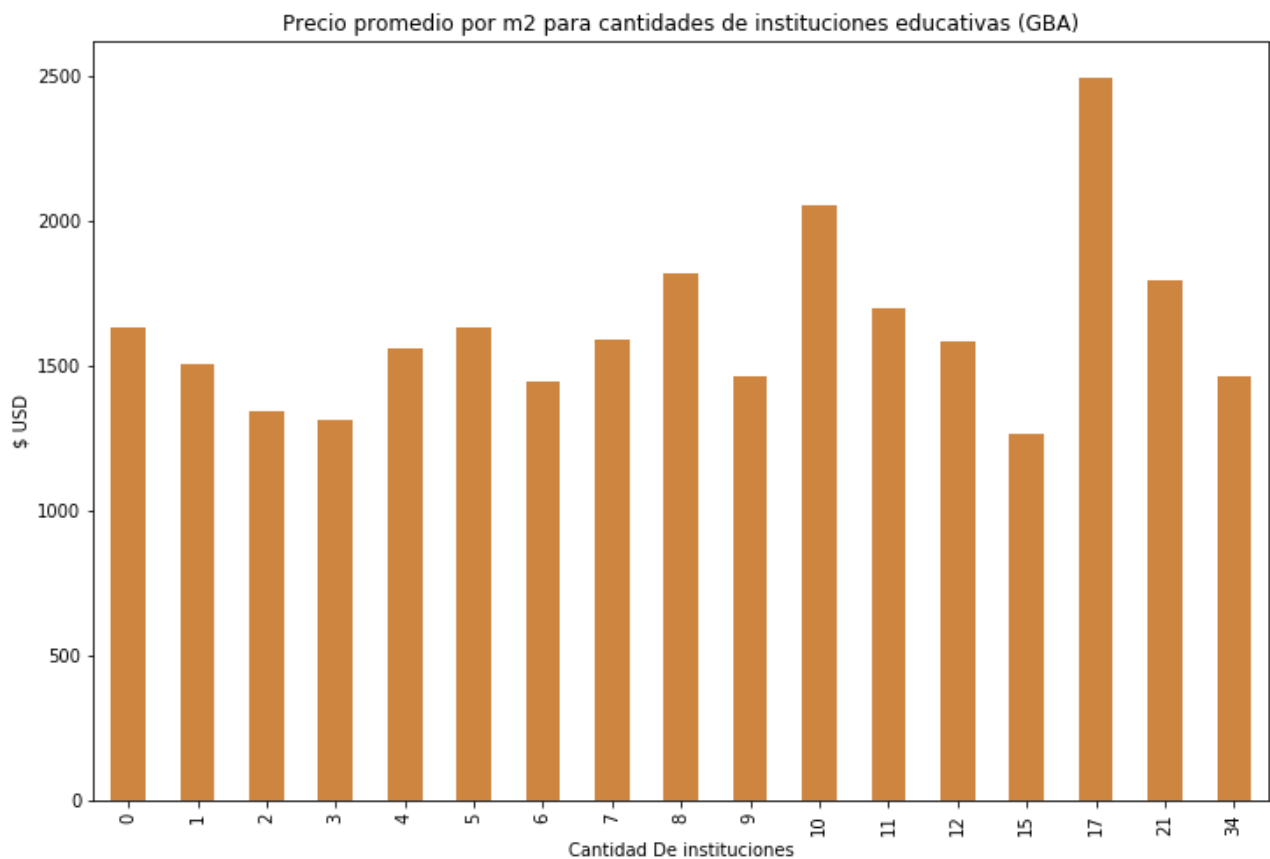
Análisis de Instituciones Educativas:

A continuación se encuentra el análisis de las instituciones educativas para un set de datos completo de 72474 entradas.

Análisis del precio por metro cuadrado:

Para los siguientes gráficos se agruparon las propiedades por cantidad de instituciones cercanas y sacando el promedio del precio por metro cuadrado para ellas. Para que los resultados sean significativos, se filtró a todos los conjuntos de propiedades con menos de 300 entradas para las cantidades de instituciones cercanas. Por ejemplo: Si solo hay 15 propiedades con 60 instituciones cercanas, estas no se tuvieron en cuenta para el análisis.

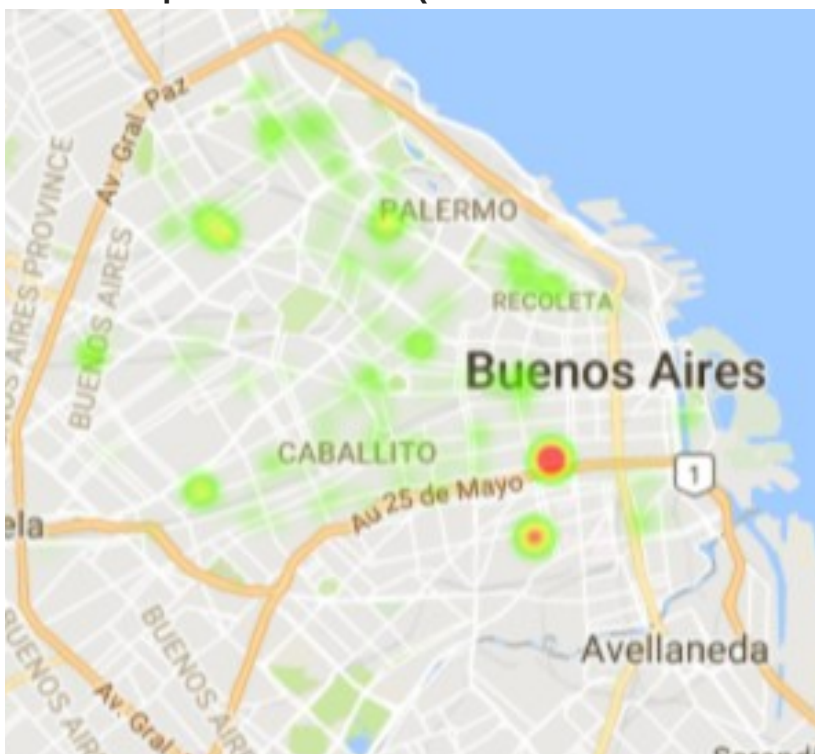




Como se puede observar en los tres gráficos, hay una tendencia de un pico máximo en todos para una cantidad de instituciones cercana a 20 instituciones. Es importante notar sin embargo, que dichos máximos son muy distintos siendo el de CABA casi el triple que el de GBA (6000 contra 2400 respectivamente).

Se procede a aislarlas y observar en que zonas se concentran.

HEATMAP por ubicaciones (CABA con 19 instituciones)



HEATMAP por precio del metro cuadrado (CABA con 19 instituciones)



HEATMAP por ubicaciones(CABA con 9 instituciones)



HEATMAP por precio del metro cuadrado (CABA con 9 instituciones)



HEATMAP por ubicaciones(GBA con 17 instituciones)



HEATMAP por precio del metro cuadrado (GBA con 17 instituciones)



¿Que conclusiones se obtienen de estos heatmaps?

Se pueden ubicar polos en donde hay una cantidad importante de instituciones educativas (entre 17 y 20 en un rango de 400 metros de radio) y a la vez los precios por metro cuadrado son altos en comparación a la totalidad de las propiedades. En ellos se genera una tendencia que propia marcar una relación directa entre estos dos factores. Estos lugares son:

Olivos
San Cristóbal
Villa Urquiza/General Urquiza
Parque Patricios
Colegiales

A menor escala, zonas céntricas de los siguientes barrios:

Merlo
San Miguel
Quilmes
Recoleta
Belgrano
Puerto Madero
Almagro

Resulta interesante notar una situación particular que ocurre en lugares como Villa Crespo y Boedo en donde el promedio de instituciones cercanas es 9 pero sin embargo hay precios muy altos en las propiedades. Esto genera una problemática en cuanto a la validez del análisis por instituciones educativas y sugiere quizás realizar un filtro mas estricto en los datos para futuros análisis. Sin embargo 10 instituciones educativas en un rango de 400 metros sigue siendo un numero importante.

¿Cuales son las instituciones educativas en estos polos?

A continuación se muestran los nombres de dichas instituciones por barrio(Solo de los 3 barrios mas significativos) para tenerlas en cuenta en futuros análisis cuando se busque predecir precios por propiedades. Cabe aclarar que son mas de 20 ya que se toman el conjunto de todas las instituciones del barrio y no solo las de una ubicación en particular.

Olivos:

Northlands
Colegio San Andrés secondary
Centro Cultural Italiano - Colegio Alessandro Manzoni
St. Andrew's Scots School
Instituto Jesús en el Huerto de los Olivos
Escuela Montessori Olivos SRL
Colegio San Ignacio
Colegio Nuestra Señora de la Paz
St. Luke's College
Action Integral Institute of Performing Arts
Escuela Municipal Paula Albarracín de Sarmiento
San Andres Secundario Olivos
St. Nicholas College
Instituto Superior De Musica Jose Hernandez
Escuela EPB N° 2 "Benemérito Teniente Gral. Bartolomé Mitre"

UCES OLIVOS
Colegio Tarbut
Estudio Cambrée Tatiana Flaker
UCES UNIVERSITY OF BUSINESS AND SOCIAL SCIENCES
Fundacion Universidad de San Isidro
Colegio San Nicolas Primario
COLEGIO SAN NICOLAS JARDIN
Escuela Superior De Informatica De La Prefectura Naval Argentina
CENTRO PAMPA / escuela de diseño
Colegio Santa Magdalena
Colegio Eidep
De Los O Colegio Jesus En El Huerto
Jardin San Ignacio
E.M.P.A.S
Ganesha YOGA
Colegio Feli
Escuela Hija St Andrews
Auditorio niño Jesus De Praga
Escuela Municipal De Musica
Niño Jesús Del Praga
Jardín Maternal Niño Jesús de Praga
Auditorio Northlands School Olivos
Escuela De Tomas, Francisco Borges Y Rosales
Jardin Jho
Jardin de infantes CCI - Centro Cultural Italiano -
Jardin centro cultural italiano
Colegio Centro Cultural Italiano
SCUOLE CCI
Jardin Jesus en el Huerto de los Olivos
Jardin Dante
ArtBA
ITBA
Jardin Maternal Osecac
Instituto San Migue
St. Nicholas College
English Boutique
Scout Huerto De Los Olivos
Grupo Capoeira Brasil Buenos Aires GCB - La Lucila
CID vicente lopez
Colegio Nuestra Sra De La Paz
Toefl
Centro Cultural y Político Micaela García
Centro de Instrucción Aeronáutica C.I.A.
Escuela N 16 Marcelino Ugarte
Escuela EST Nro 3

San Cristobal:

DE LAS VICTORIAS
La Aldea del Buen Ayre
Instit Salesiana - Colegio San Antonio
Escuela Generación del Futuro
Danza Árabe Escuela Aldana Arguello

Sol de America
Crema y Chocolate
Crema y Chocolate
Fundacion tomas eloy martinez
Curso de cerrajería presencial e intensivo
Cenedi
AUDITORIO NAMUNCURA
Colegio San José de Calasanz
Colegio Calasanz
San Antonio
ILEC - Instituto Laico de Estudios Contemporaneos
Special Education Institute OUR LADY OF LUJAN
Curso Sublimacion
Escuela Domiciliaria N° 2
JIC N°4 DE 6° MARIANO BOEDO
Ciber Pibes
Escuela Infantil Cyberpibes
Instructorado De KIZOMBA
Esc de Com N° 22 DE 6 "G. M. Zubiria "
Espacio De Creacion Yapeyu
ESCUELA N. 6 D.E. 8 SAN JOSÉ DE CALASANZ
Escuela N 25 Paula A De Sarmiento
Pasillo al fondo "Centro Cultural"
Instituto Calazans
Escuela Lucia
ESCUELA PAULA ALBARRACIN DE SARMIENTO
Escuela Infantil La Torrecita
"Puente Azul" Jardín de Infantes
Taller De Arte Hilodearbol
San Antonio Salesian house
Curso Calidad
SANTA MARIA INSTITUTE
Instituto San Antonio-A.226
Escuela De TANTRACLASICO
CFP No.30
Yoga
Escuela N° 9 D.E. 8 - Florentino Ameghino
Supervision D E 8 Primaria

Villa Urquiza/General Urquiza

School No. 24 Francisco Morazan
San Patricio Secondary Institute
Nuevos Aires SRL
Sir Thomas Malory
Mad Escuela
Sir Thomas Malory School
Estudio Joya
Instituto Superior del Profesorado en Educación Especial
INA - Instituto Nuevos Aires
St. Patrick's School
Escuela Infantil Chiquilines
Instituto Junín

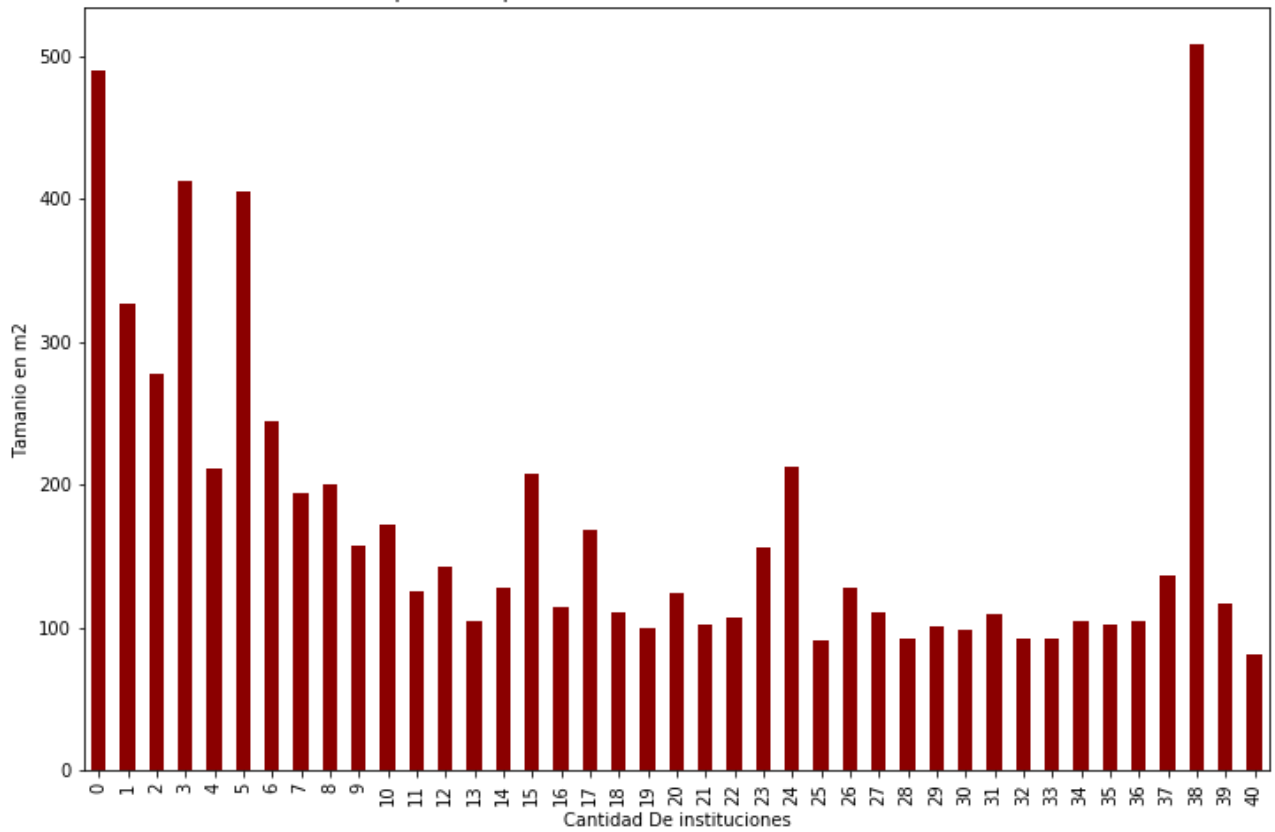
The Garden of the Fund
Special Education School 11
Burdel de maderas
Clases de Guitarra en Villa Urquiza - Música y creatividad
St. Patrick's School instituto San Patricio
Saint Patrick
Acha Club
Caebt 56 - Parroquia Jesús Misericordioso
St patrick's Kinder
Naranon grupo
Escuela Nro. 4 D.E. 15
Escuela Nro 24 D.E.15 - Escuela Nro 8
Ispee
Facu Aye
Facultad Moron
ESc Infantil N 8 DE 15
San Pablo
Island of My Dreams
escuela republica de costa rica
Escuela N° 24 SIGLO XXI
Escuela n°15 acevedo
Universidad -ciclo basico
Drago Uba
CBC Drago
Colegio Franco
UBA - Drago
UBA Sede Drago
Cbc
CBC UBA - Sede Drago
Sede Drago

Es importante notar, como se puede ver en las ultimas instituciones de Villa Urquiza, la sede Drago del CBC aparece subida repetidas veces pero escrita de distinta manera. Esto muestra que mas allá del gran poder que tiene Google Places, los datos pueden no ser del todo fehacientes.

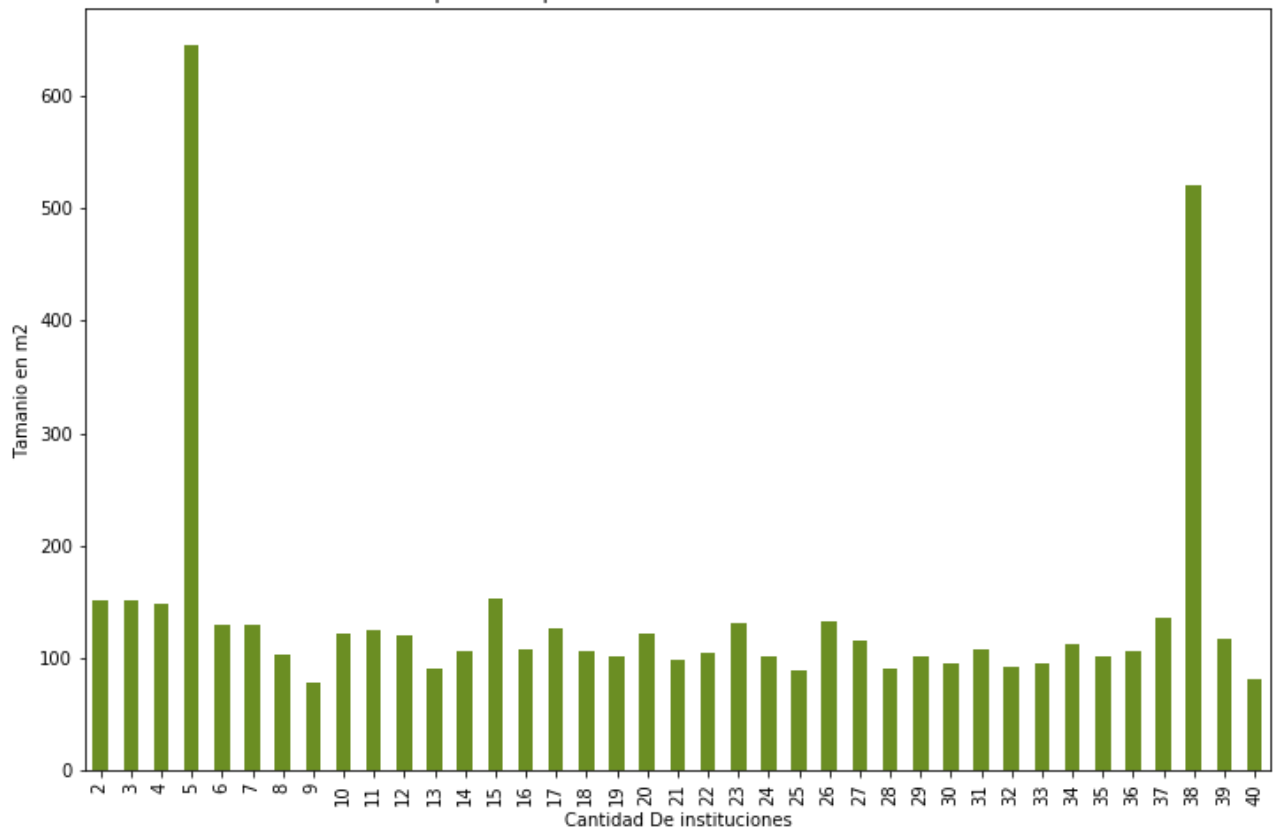
Relación entre el tamaño de una propiedad y la cantidad de instituciones educativas

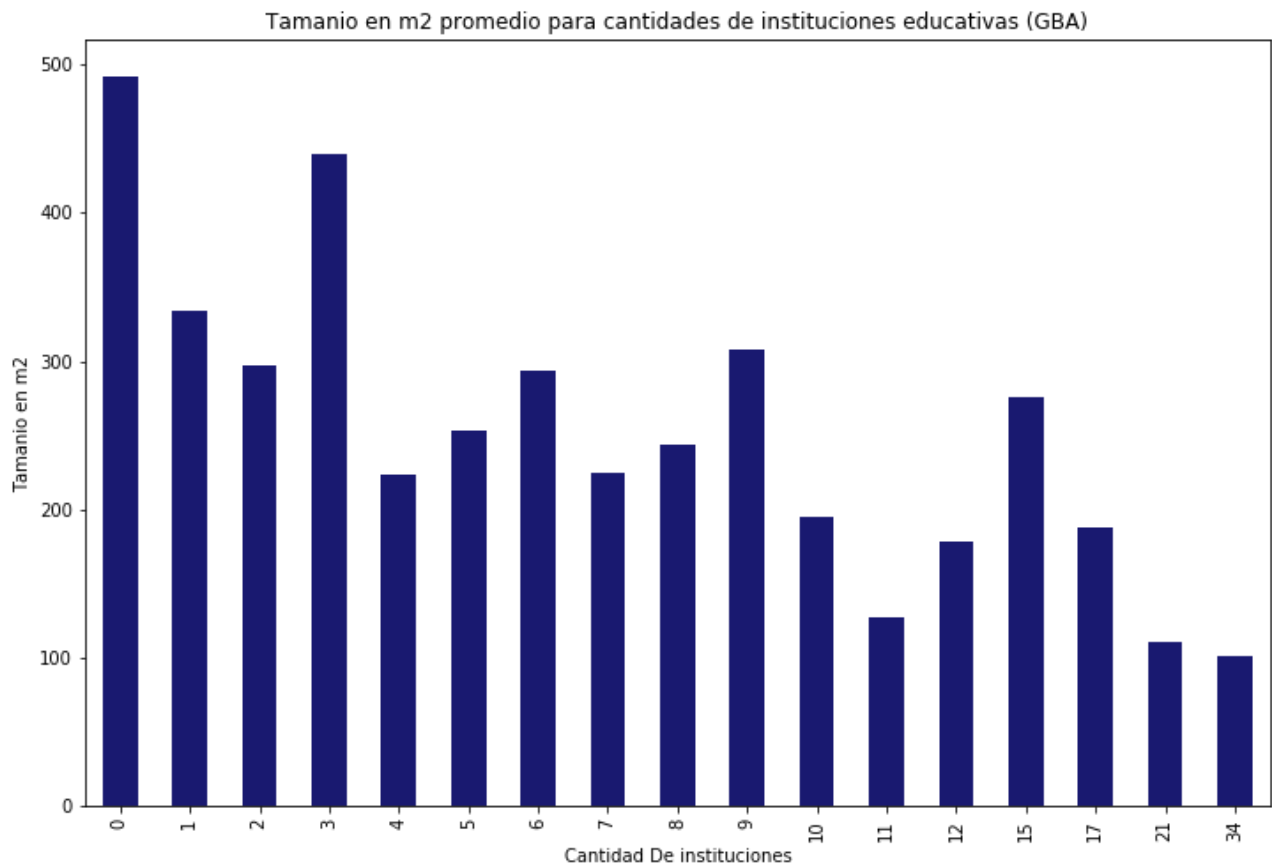
A partir de este análisis se busca encontrar alguna relación entre el tamaño de la propiedad y la cantidad de instituciones en su cercanía. Previo a los resultados se supone que puede llegar a haber una relación teniendo en cuenta que mientras mas grande sea, es mas probable que mas personas vivan allí y por consiguiente necesiten de variadas instituciones educativas. A la vez se podría dar también que pequeñas propiedades estén en zonas donde la demanda de instituciones educativas se muy alta y por esta razón priorizar la cercanía a las instituciones dejando de lado otras comodidades como puede ser un mayor espacio.}

Tamaño en m2 promedio para cantidades de instituciones educativas (CABA + GBA)



Tamaño en m2 promedio para cantidades de instituciones educativas (CABA)





Al darle una vista rápida a los gráficos se ven que los comportamientos de CABA Y GBA son muy distintos. Pero en este caso los picos máximos tienen valores similares rondando entre 500 y 600 metros cuadrados. Se procede a analizar los datos por separado para obtener conclusiones mas precisas.

¿Que conclusiones se obtienen de GBA?

En GBA se ve una tendencia a la baja de tamaños, siendo que a mayor cantidad de instituciones, las propiedades tienen un tamaño menor. Una posible razón puede adjudicarse a que muchas de las propiedades de GBA pertenecientes al Dataframe son de barrios cerrados en donde los terrenos suelen ser particularmente grandes y a la vez las distancias a, no solo instituciones educativas si no también a locales o zonas con mayor población, son superiores.

Relación entre el tamaño total de la superficie y la cantidad de instituciones educativas para CABA:

Se distinguen dos claros picos máximos en 5 instituciones, y en 38. Mas allá de esto no se ven otras relaciones que resulten de interés para el análisis. Para ubicar las propiedades en cuestión se realiza un heapmap con ellas.

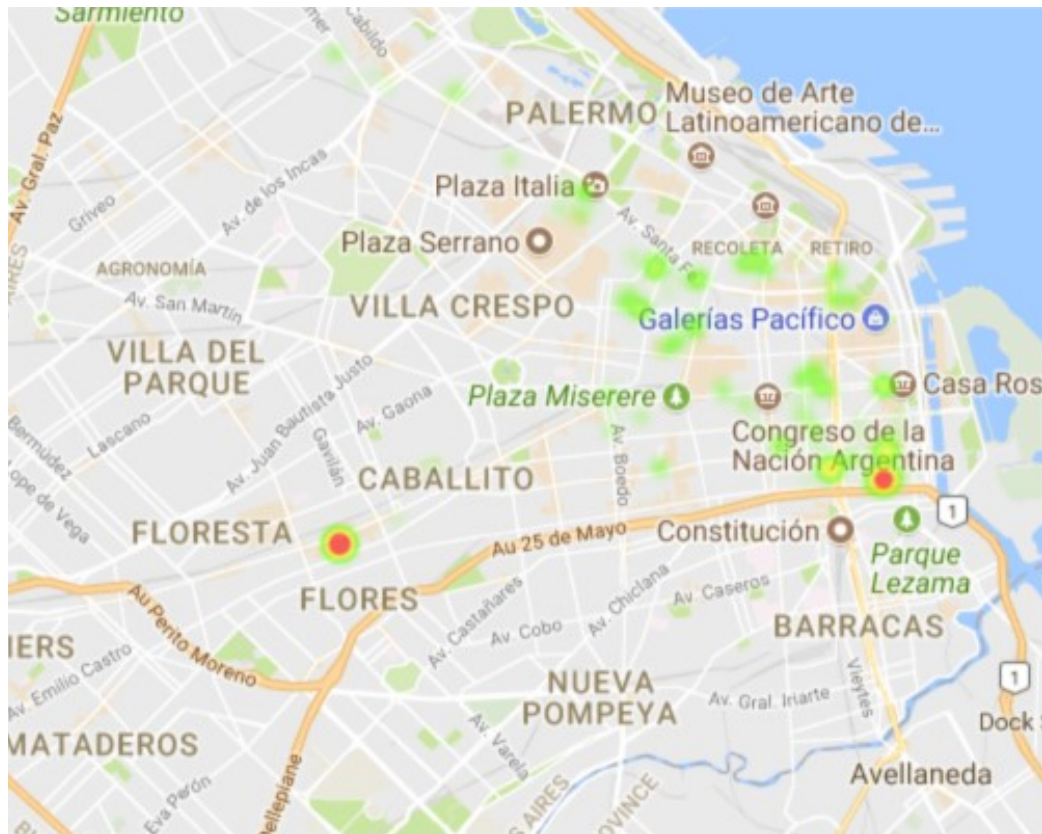
HEATMAP por ubicaciones (CABA con 5 instituciones)



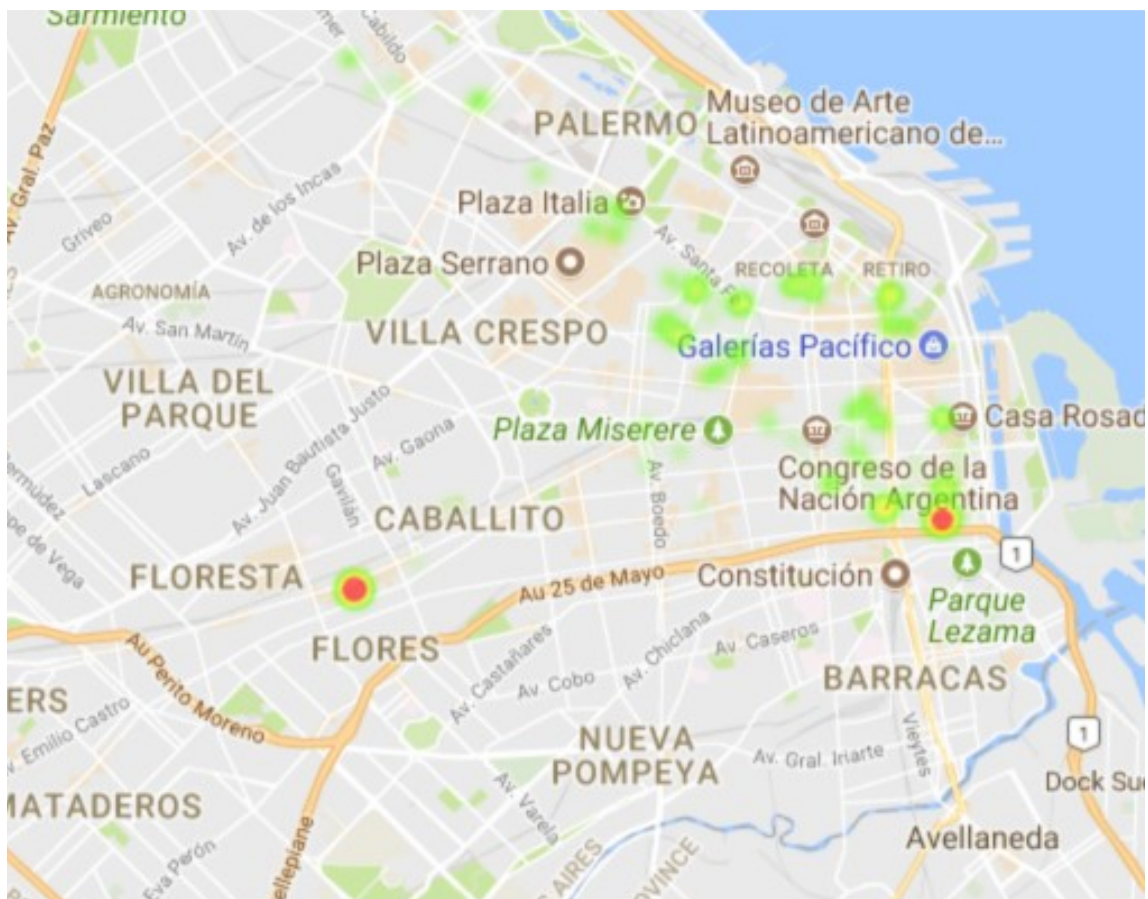
HEATMAP por tamaño de superficie (CABA con 5 instituciones)



HEATMAP por ubicaciones (CABA con 38 instituciones)



HEATMAP por tamaño de superficie (CABA con 38 instituciones)



¿Que conclusiones se obtienen de estos heatmaps?

En los heatmaps con 5 propiedades en CABA se marca la tendencia que ubicaba a Villa Crespo como un polo. En donde hay una base solida pero no muy grande de instituciones educativas y a la vez hay precios altos por metro cuadrado y propiedades de tamaños importantes. A esto se agrega la zona de puerto madero en donde se observan precios muy altos, grandes propiedades, y un buen numero de instituciones educativas, particularmente universidades.

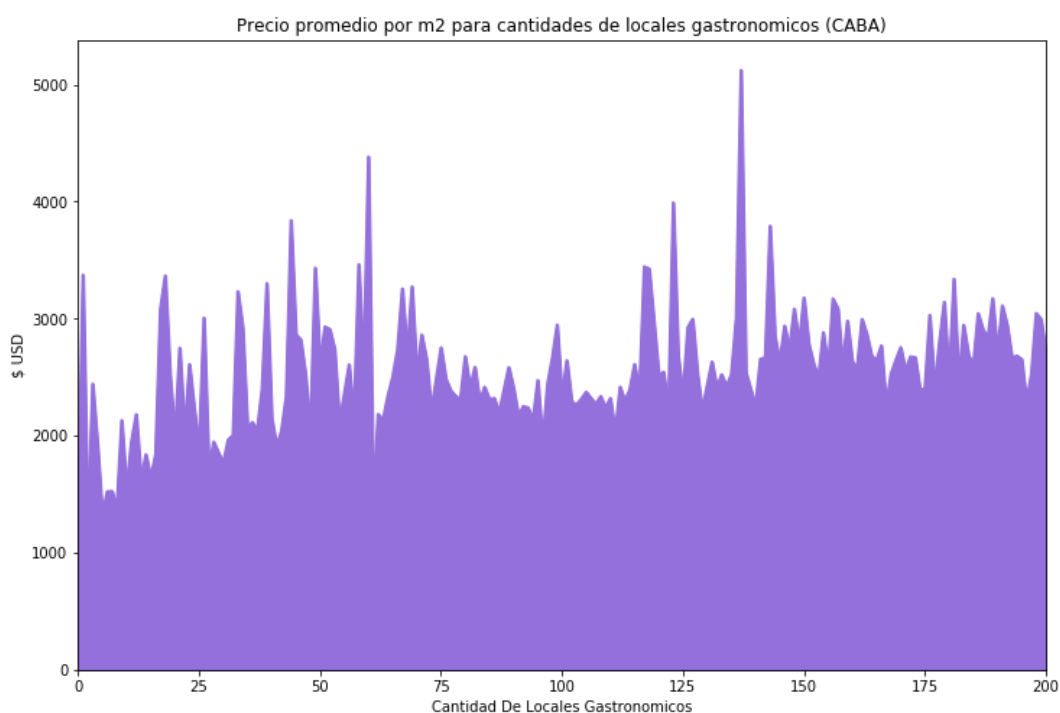
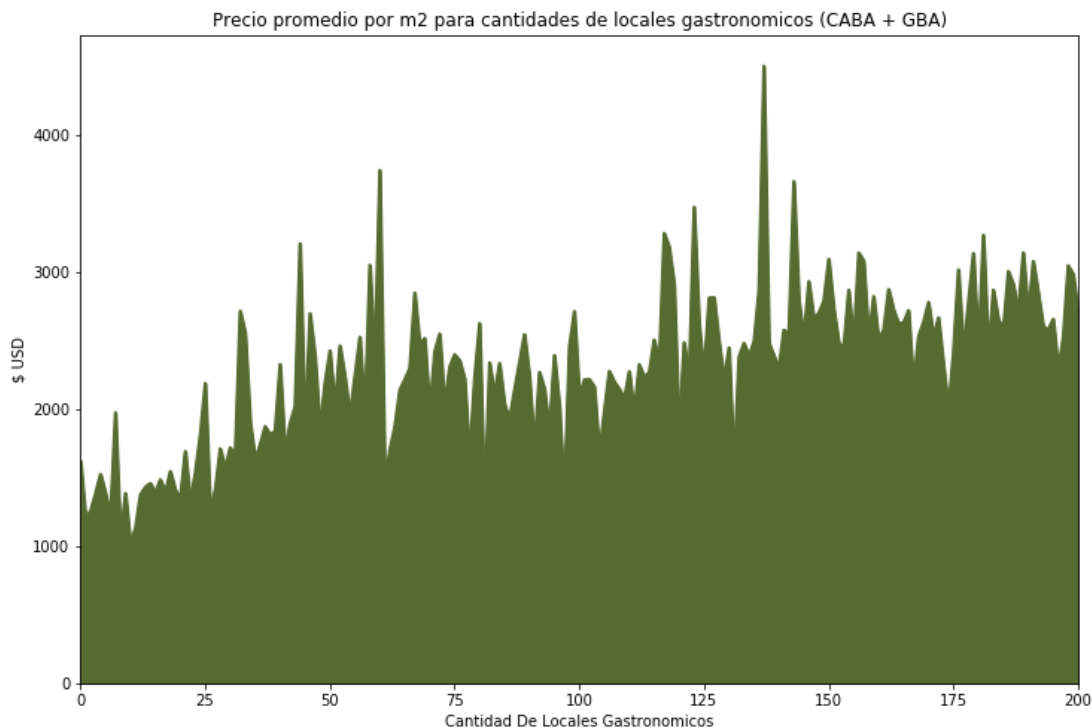
Finalmente se ubican dos nuevos polos con propiedades muy grandes y realmente un numero enorme de instituciones (38 instituciones) siendo estos Plaza Dorrego y en Flores cerca de la Av. Rivadavia entre Av.Nazca y Av.Carabobo.

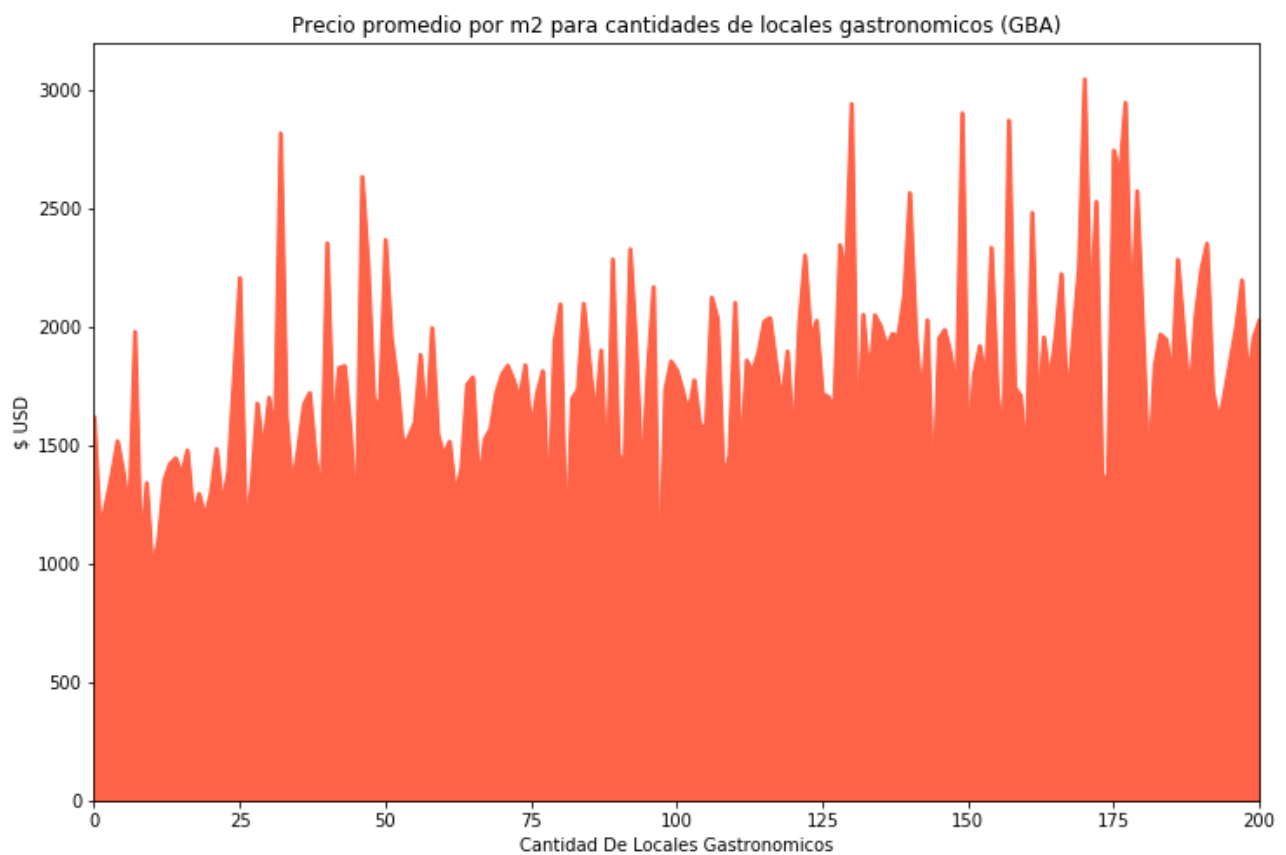
Análisis de Locales Gastronómicos:

A continuación se encuentra el análisis de las instituciones de tipo gastronómicas para un set de datos completo de 72474 entradas.

Análisis del precio por metro cuadrado:

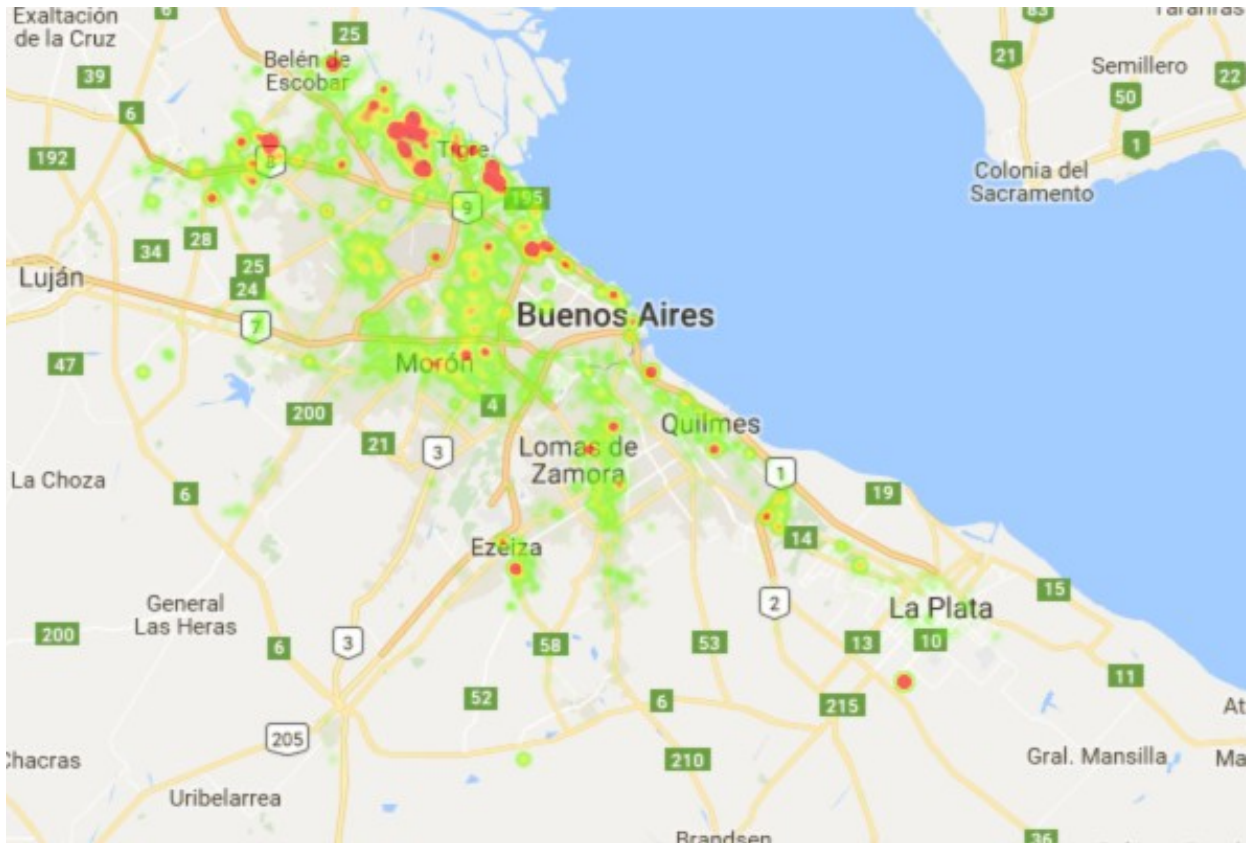
Para los siguientes gráficos se agruparon las propiedades por cantidad de instituciones cercanas y sacando el promedio del precio por metro cuadrado para ellas. Para que los resultados sean significativos, se filtro a todas los conjuntos de propiedades con un precio menor a \$usd8000 teniendo en cuenta el análisis inicial.



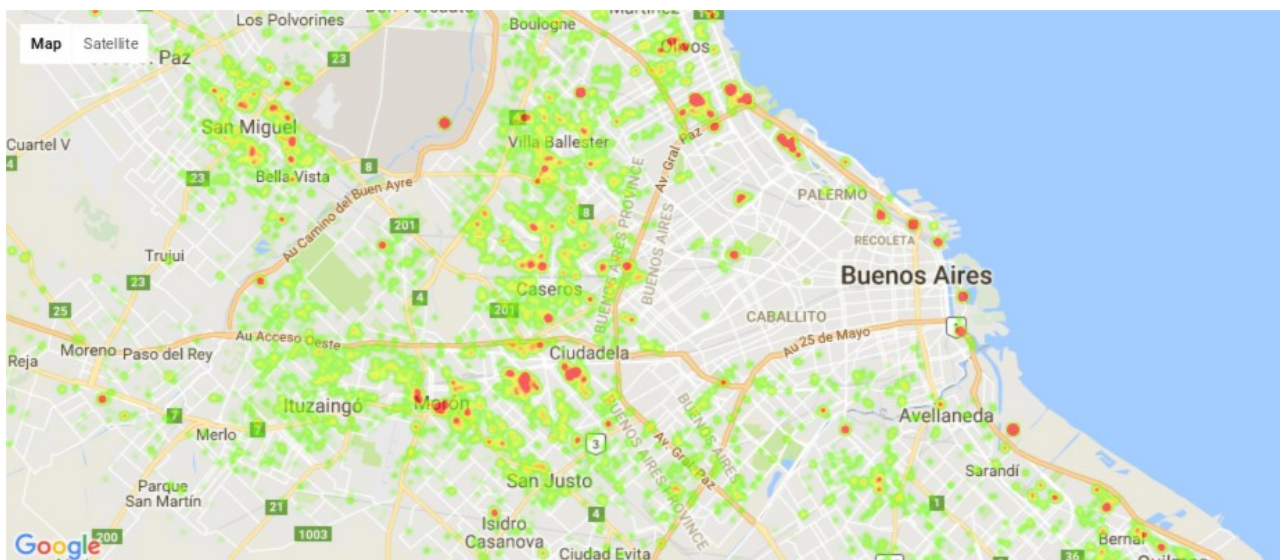


A simple vista los gráficos no dicen mucho. Es observable una leve tendencia de crecimiento en cuanto a mayor cantidad de locales el precio de las propiedades seria mayor y lo mismo para menor cantidad de locales. Dicha tendencia se puede apreciar mejor en el gráfico de GBA + CABA. En los otros gráficos, la tendencia de punta a punta (mas allá de abruptas variaciones) se asemeja a constante. Se procede a ubicar donde están los lugares con menores locales y donde están los que mas tienen, es decir, los extremos.

Heatmap CABA + GBA de 0 a 20 locales con el peso en el precio por metro cuadrado de locales gastronómicos

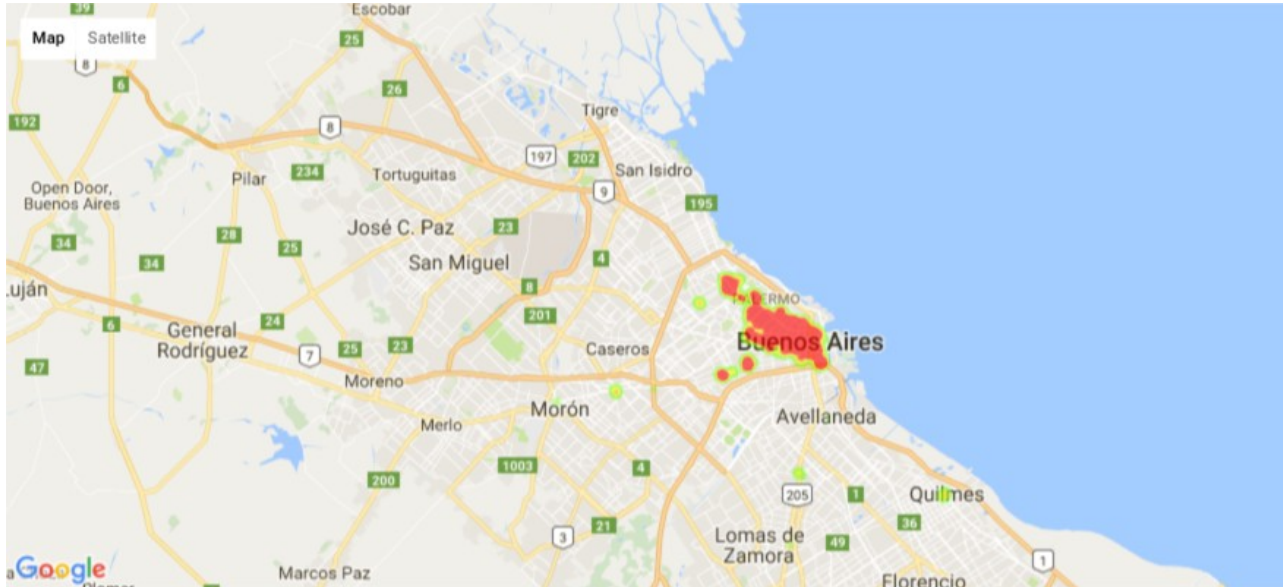


Mismo Heatmap con ZOOM



De estos últimos heatmaps se puede extraer la interesante conclusión de que las propiedades con pocos locales gastronómicos a su alrededor pertenecen a GBA y prácticamente excluyen a CABA.

Heatmap CABA + GBA de 180 a 200 locales con el peso en el precio por metro cuadrado de locales gastronómicos



Continuando con el análisis previo, las propiedades con entre 180 y 200 locales gastronómicos en sus cercanías están en su totalidad ubicados en CABA, mas particularmente su zona céntrica.

Principales polos gastronómicos:

Analizando mas detalladamente las posiciones de las propiedades en el mapa se encuentran los siguientes sectores como principales polos gastronómicos:

- Plaza Serrano
- Recoleta
- Plaza Dorrego
- Belgrano

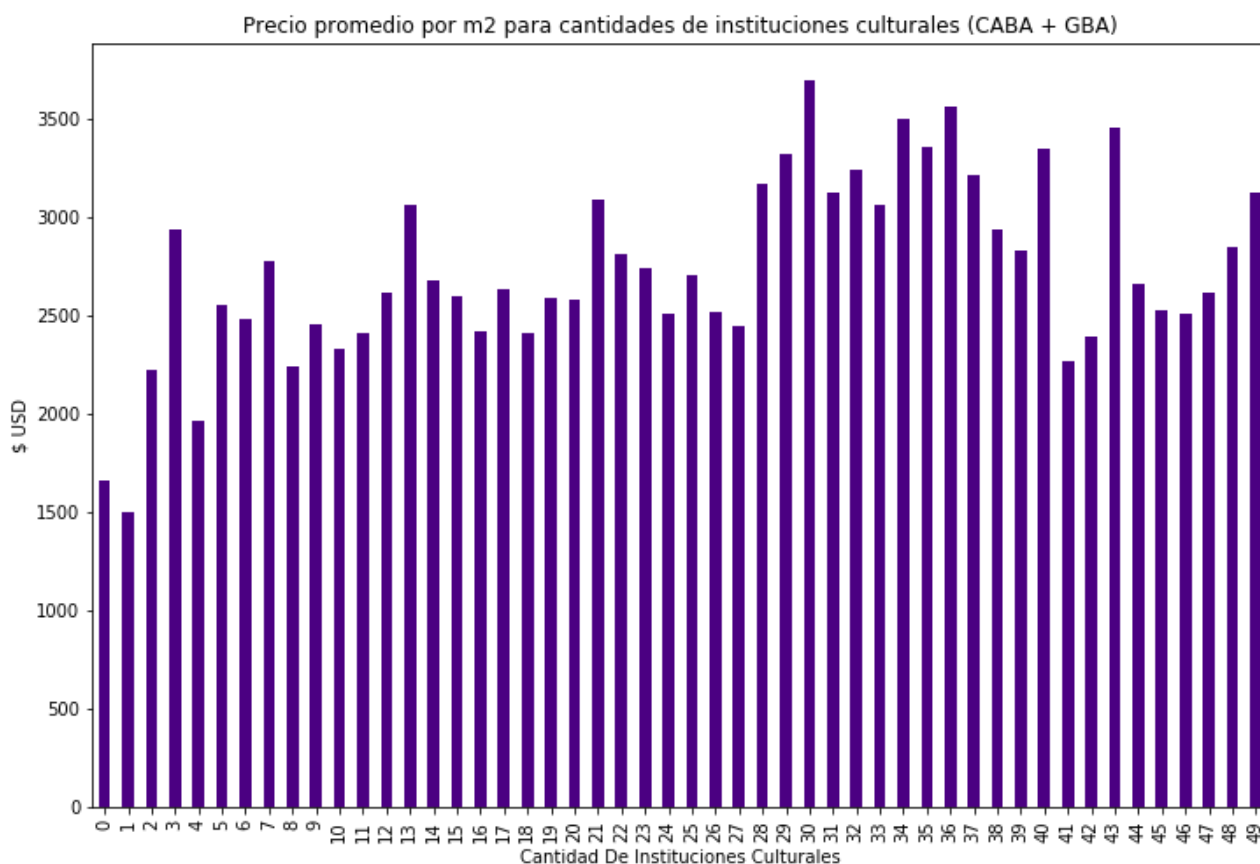
Cabe aclarar que en GBA las mayores concentraciones están ubicadas en la cercanía de centros comerciales (shopping centers).

Análisis de Puntos de Interés Cultural:

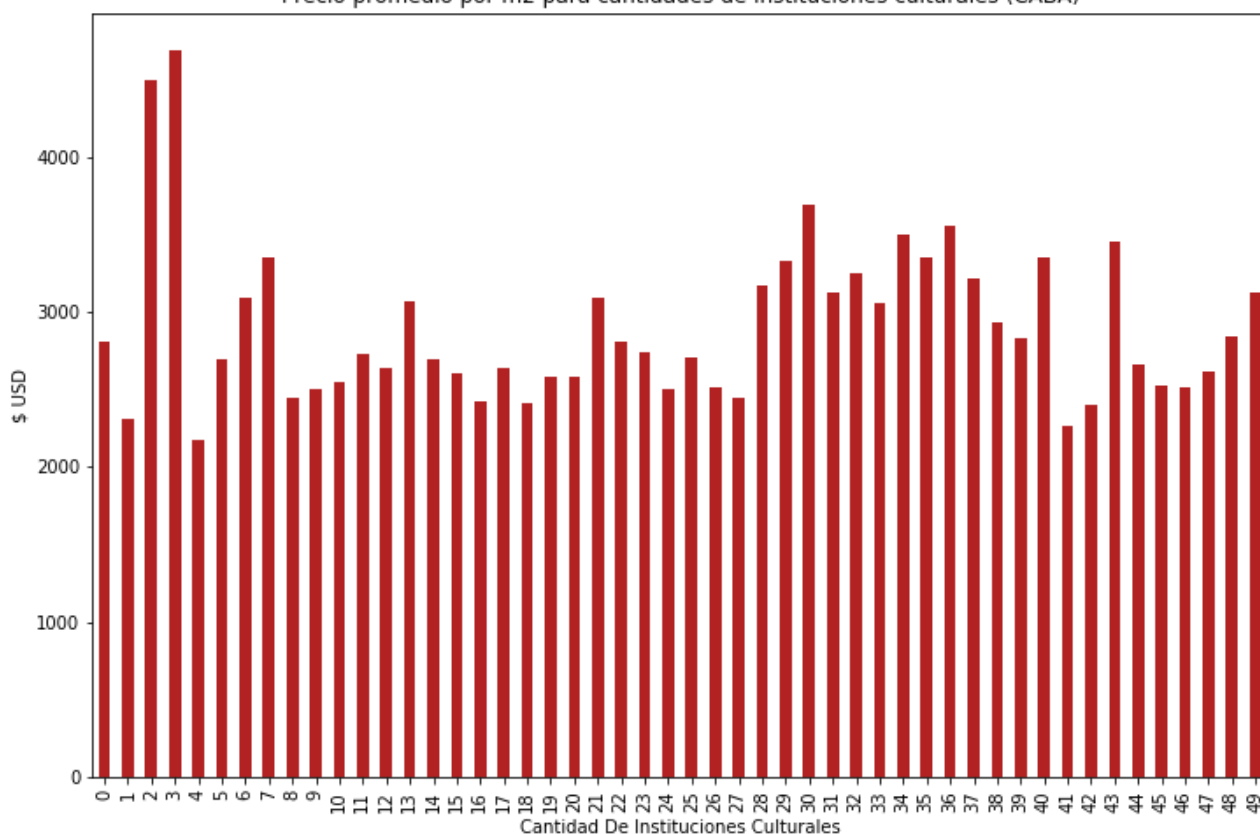
A continuación se encuentra el análisis de los puntos de interés cultural para un set de datos completo de 72474 entradas.

Análisis del precio por metro cuadrado:

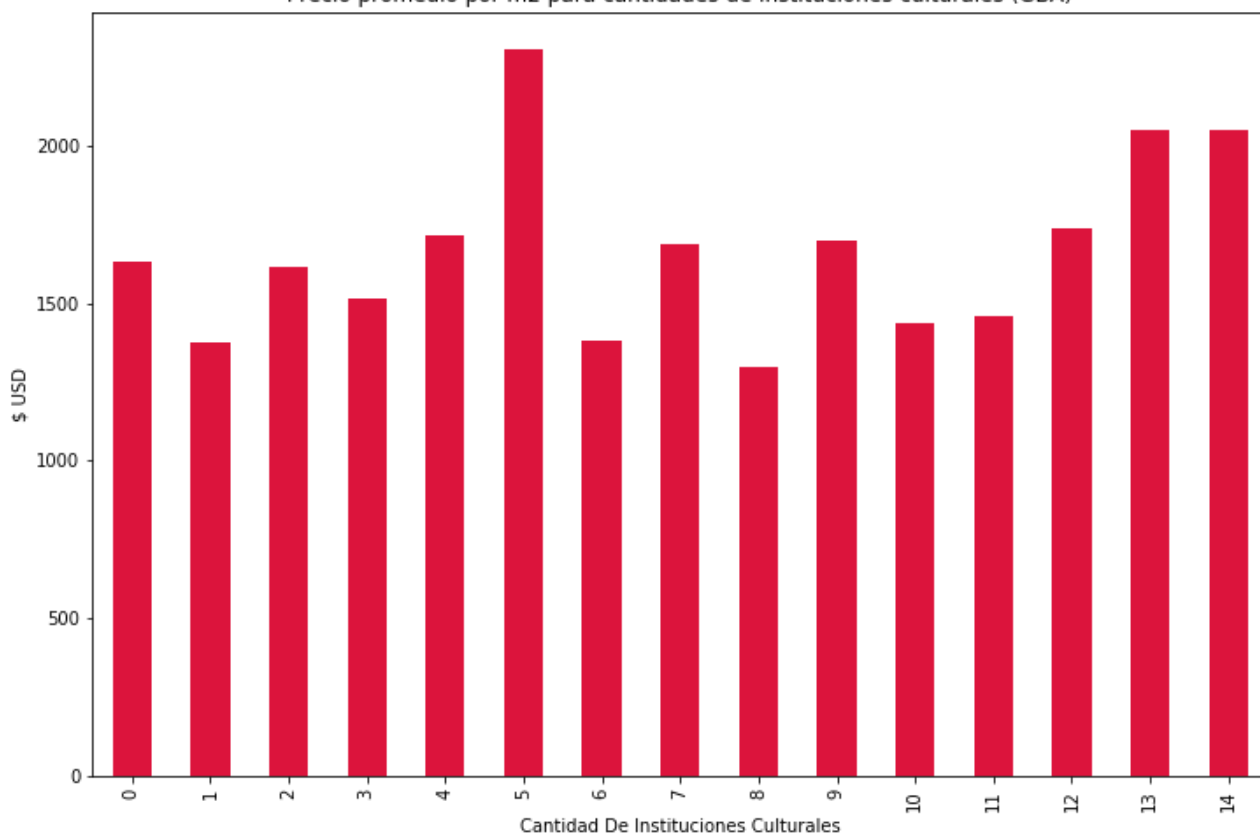
Primero se comienza sin agregar ningún recorte extra al set de datos y se analiza una descripción del set de datos. Utilizando la función describe, es posible notar que menos del %50 de los datos tienen una cantidad de entradas menor a 60. Esto quiero decir que puede haber por ejemplo solo una propiedad con un numero muy alto de entradas que distorsiones a todo el resto. Como lo que se esta tratando de buscar son tendencias. Se procede a cortar la cantidad de instituciones por el porcentual del %99 para achicar posibles disrupciones. Finalmente recortando por cantidad de puntos culturales menores a 65 y que estos tengan por lo menos 10 repeticiones se obtienen los siguientes gráficos:



Precio promedio por m2 para cantidades de instituciones culturales (CABA)



Precio promedio por m2 para cantidades de instituciones culturales (GBA)



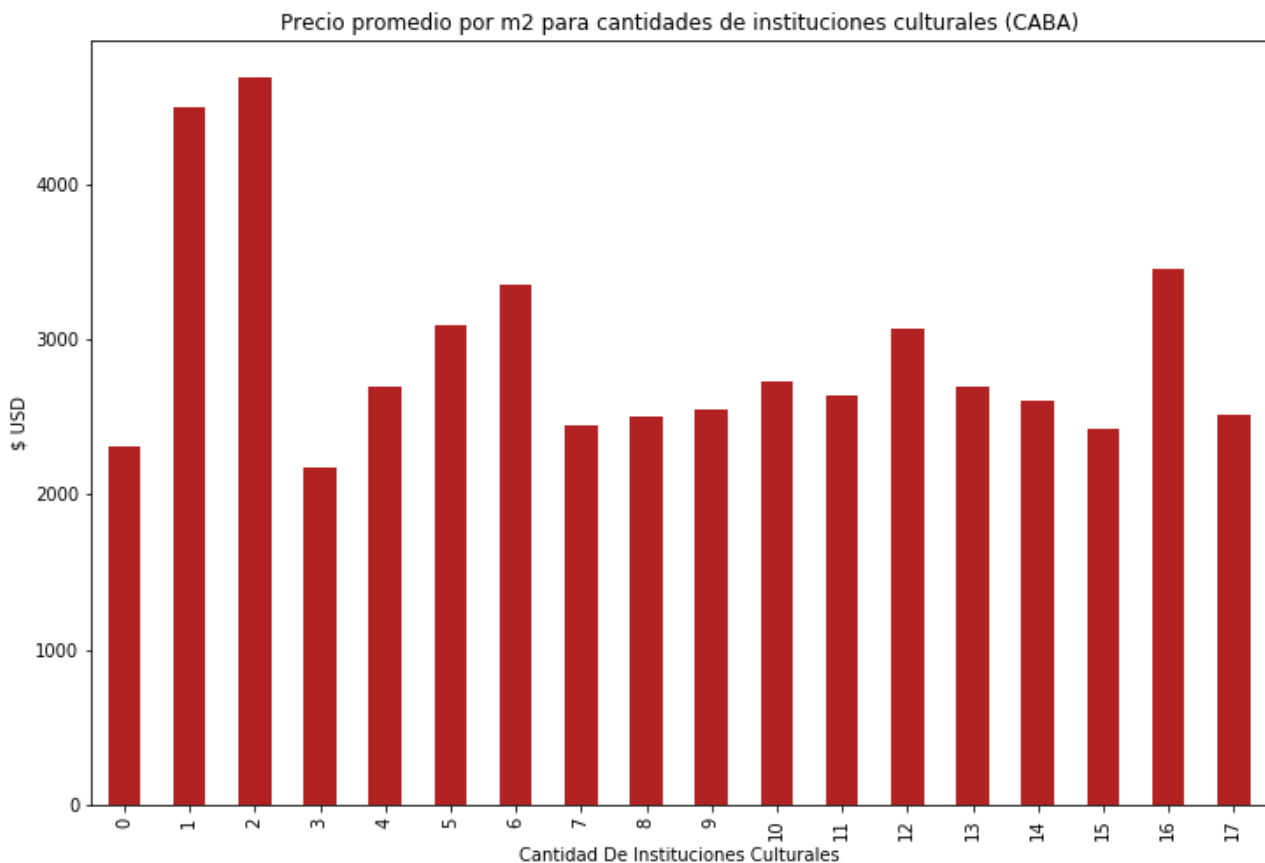
Instantáneamente se percibe lo que conociendo el territorio sería sospechable. Esto es en primer medida que la cantidad de puntos de interés cultural en CABA contra los de GBA es considerablemente mayor siendo 49 contra 14 las respectivas cantidades máximas. A la vez es notable resaltar que en gráfico de CABA + GBA hay una clara disminución en el precio de la propiedad para lugares con 0 o 1 puntos cercanos. Mas allá de estos conceptos, los tendencias sugieren ser a simple vista constantes. Se trata de ubicar a las propiedades con pocos puntos cercanos en zonas particulares.

Heatmap CABA + GBA con 0 o 1 puntos de interés cultural cerca



Como los gráficos anticipaban, casi en su totalidad estas propiedades están ubicadas en GBA. Sin embargo surge un punto interesante y es que se concentran mas que nada en la zona norte de GBA. Esto podría adjudicarse a que un gran porcentaje de las propiedades utilizadas pertenecen efectivamente a esa zona.

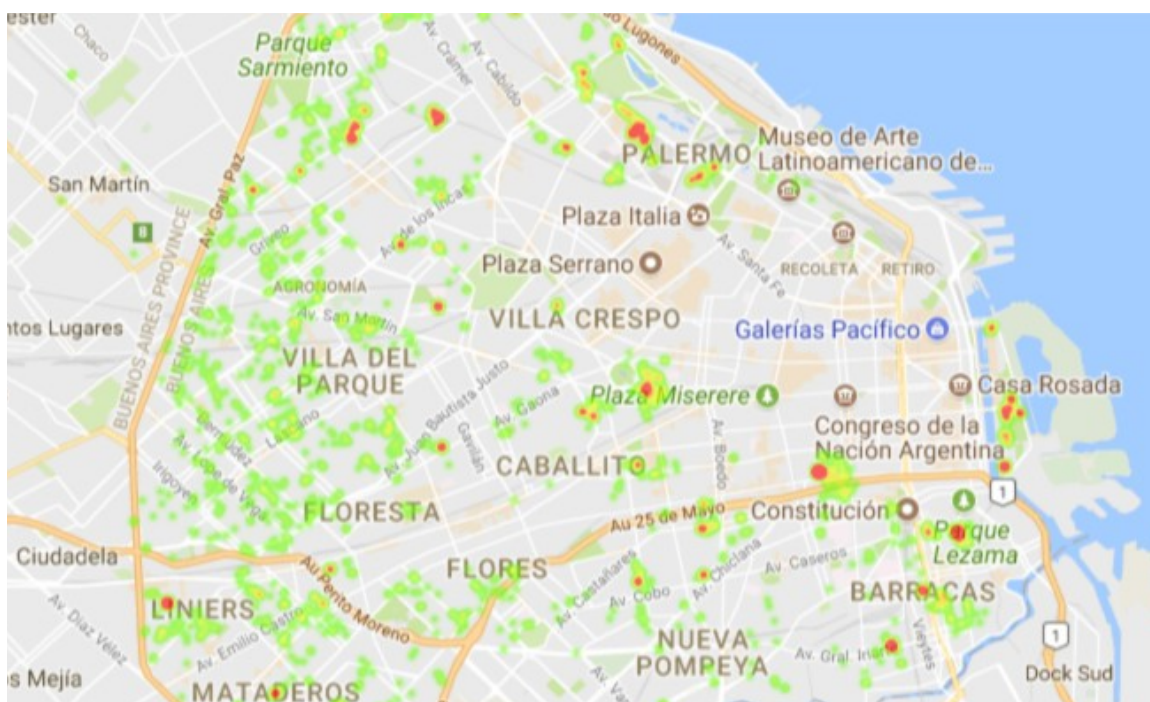
Al tratar de analizar los gráficos previos, los diferenciados por CABA muestran indicios de que podría aplicarse una análisis mas particular para lograr obtener sus polos mas importantes como se realizo con el resto de las investigaciones. Es así que se realiza un nuevo gráfico pero esta vez recortando a las cantidades de puntos de interés con bajas frecuencias, para así encontrar concentraciones mas importantes.



Los picos en 1 y 2 instituciones se siguen sosteniendo mas allá de la filtración de los datos, esto podría sugerir o que la gran mayoría de las propiedades en CABA tienen esa cantidad de instituciones cerca, o algo mas difícil de demostrar que quizás tener pocas instituciones cerca realmente mejora el precio de la propiedad.

Se asilan los puntos en CABA con una o dos instituciones cerca y se las ubica en un heatmap.

Heatmap CABA con 2 o 3 puntos de interés cultural cerca



Las propiedades en cuestión están ubicadas en las periferias de la zona céntrica de CABA (salvo excepciones). Esto muestra que hay un numero de propiedades con un valor alto por metro cuadrado que se ubican lejos de las zonas con mas afluencia de personas, o hasta turistas, ya que los puntos de interés culturales suelen ser visitados por ellos. Esto podría sugerir que al estar “aislado” de los lugares mas transitados pero seguir perteneciendo a la capital le da un valor elevado a la propiedad, mas que nada en el sentido de lugares de residencia.

Principales polos de puntos de interés cultural:

El siguiente heatmap ubica a las propiedades con mayor cantidad puntos de interés cultural (utilizando esta cantidad como el peso del heatmap):



Analizando mas detalladamente las posiciones de las propiedades en el mapa se encuentran los siguientes sectores como principales polos de lugares culturales:

- Plaza Serrano
- Plaza Dorrego
- Belgrano

Es interesante comparar estos resultados con los polos gastronómicos. Parecería que al menos a simple vista hay una correlación entre los ambos puntos con mayor concentración de puntos de interés cultural y de locales gastronómicos.

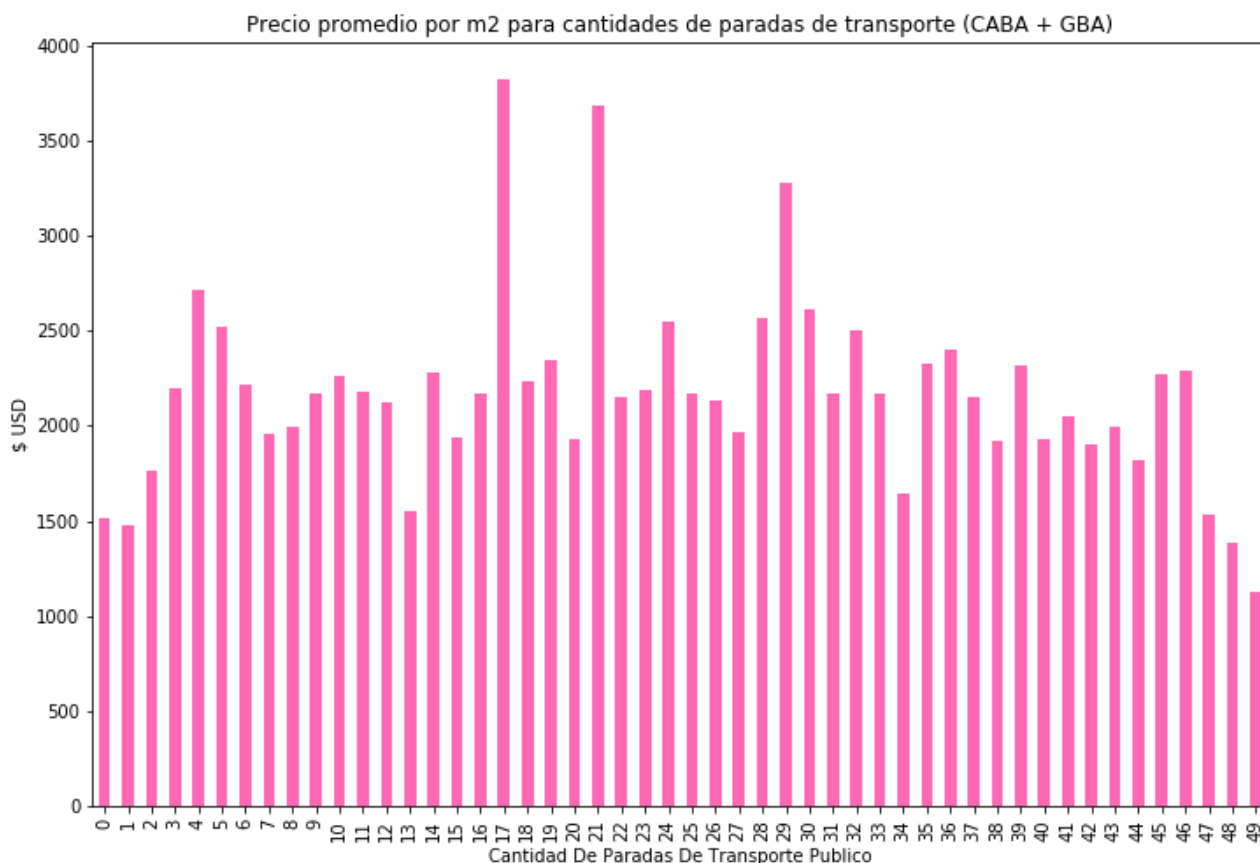
Análisis de Transporte Publico:

A continuación se encuentra el análisis de los servicios de transporte publico para un set de datos completo de 72474 entradas.

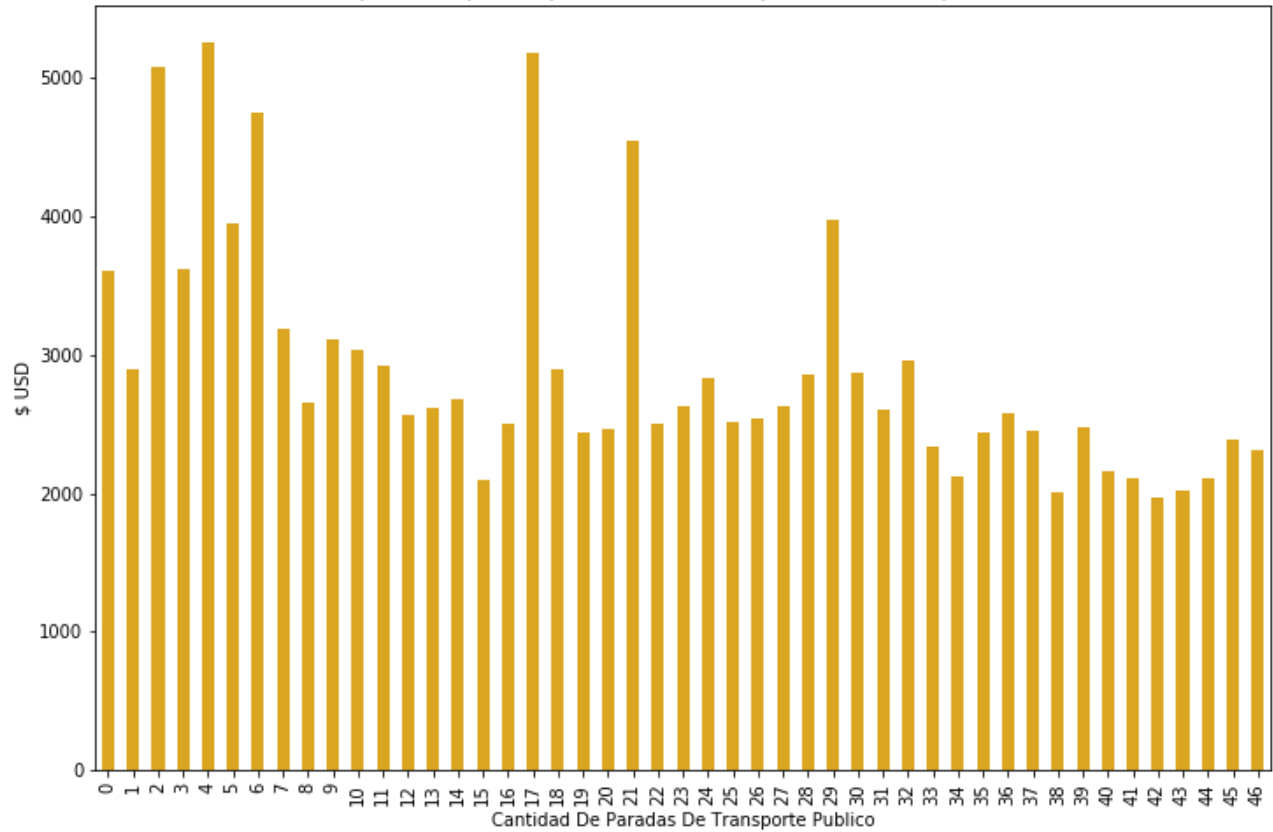
Con transporte publico se refiere a las paradas o estaciones de colectivos o subterráneos en un radio de 400m de la propiedad.

Análisis del precio por metro cuadrado:

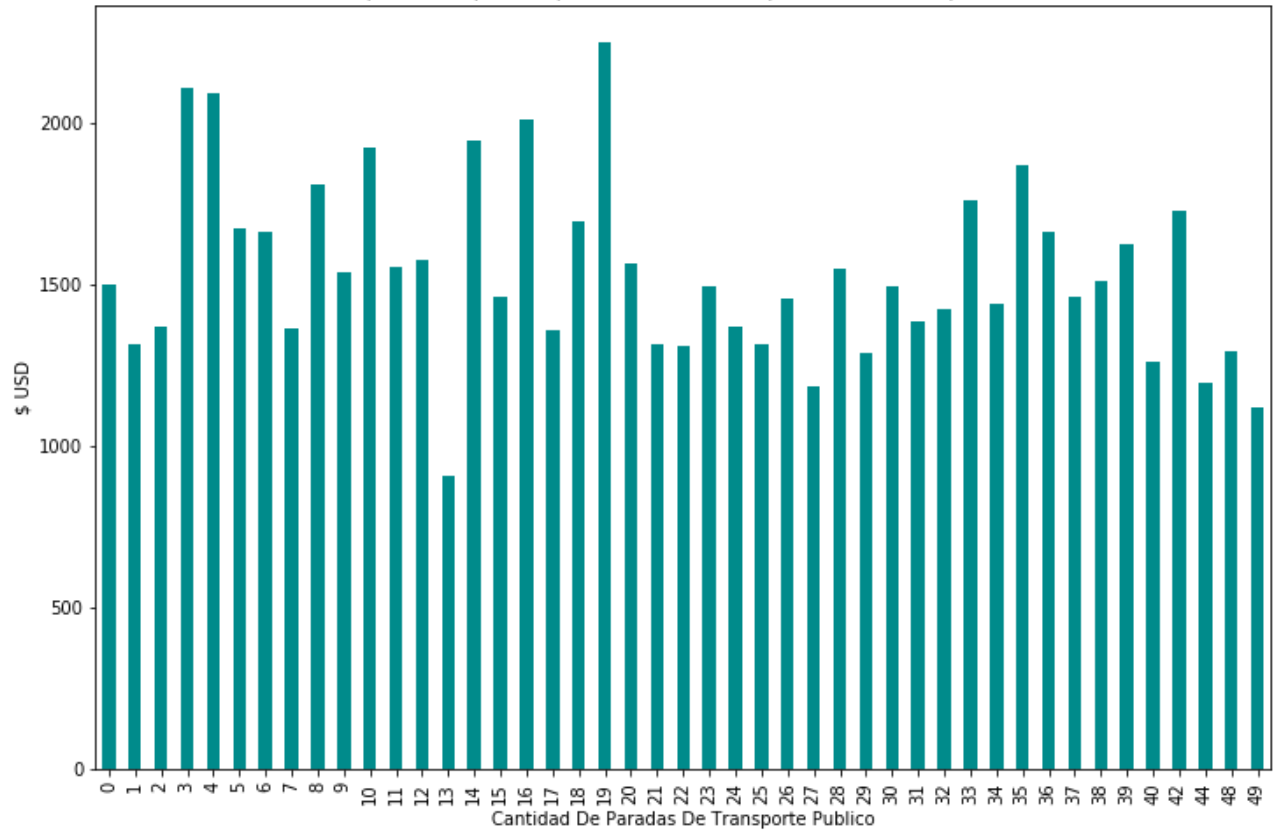
Para los siguientes gráficos se agruparon las propiedades por cantidad de paradas de transporte publico cercanas y sacando el promedio del precio por metro cuadrado para ellas. Al ser el máximo de 57 paradas relativamente bajo para el radio tomado, se filtro a todas los conjuntos de propiedades con menos de 10 entradas para las cantidades de paradas cercanas y así levemente limpiar de distorsión al set de datos.



Precio promedio por m2 para cantidades de paradas de transporte (CABA)



Precio promedio por m2 para cantidades de paradas de transporte (GBA)

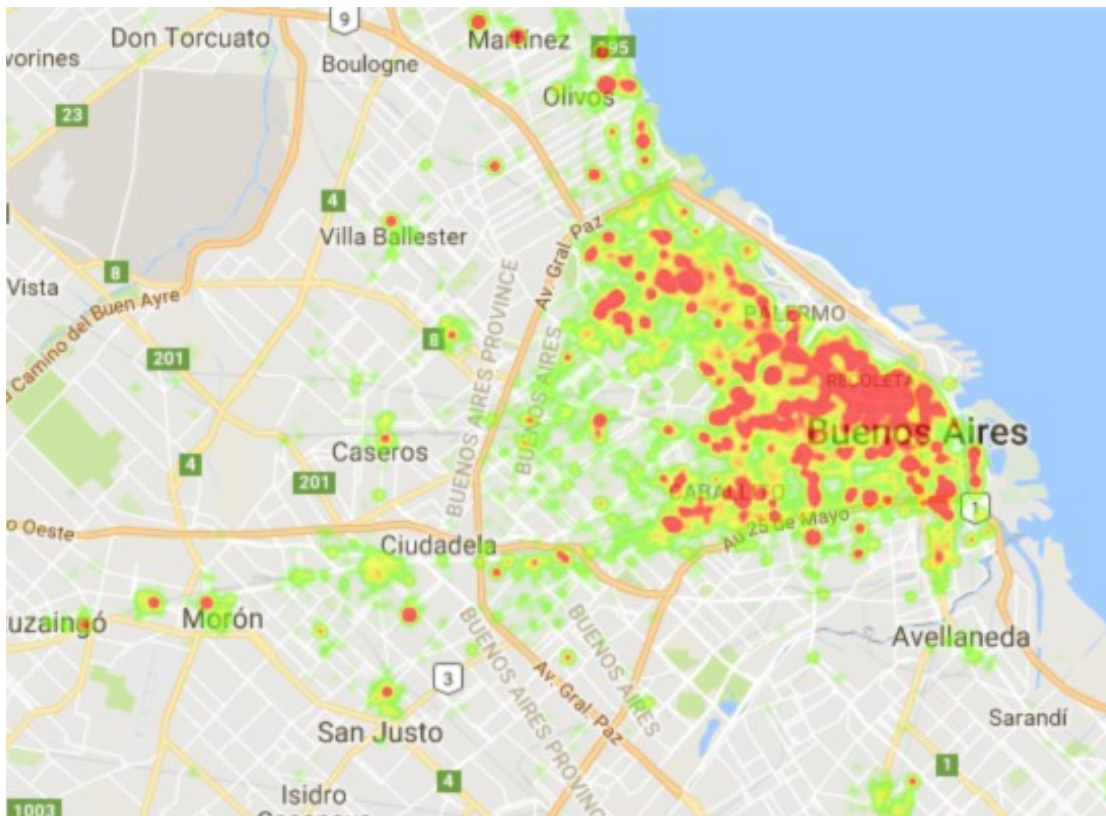


Separación de CABA + GBA por precio:

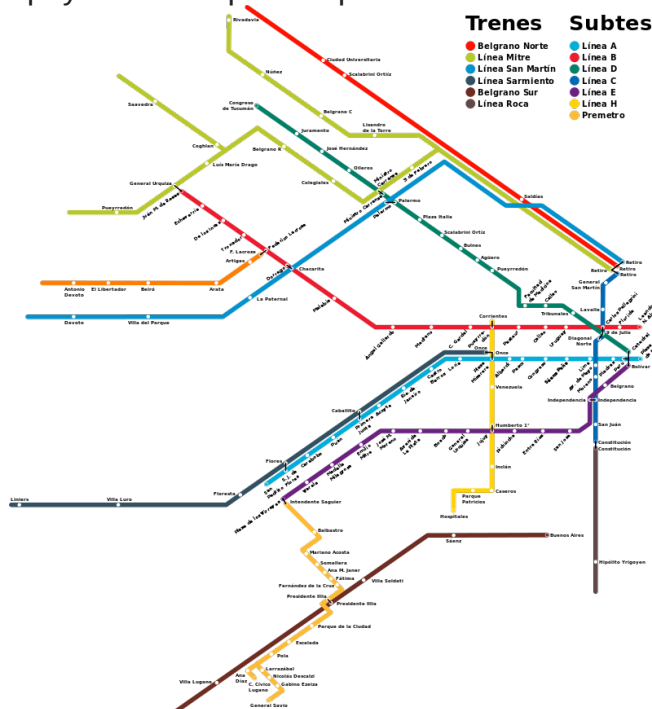
Heatmap CABA + GBA propiedades con precio por metro cuadrado mayor a USD\$2000 y el peso del heatmap puesto en la cantidad de paradas



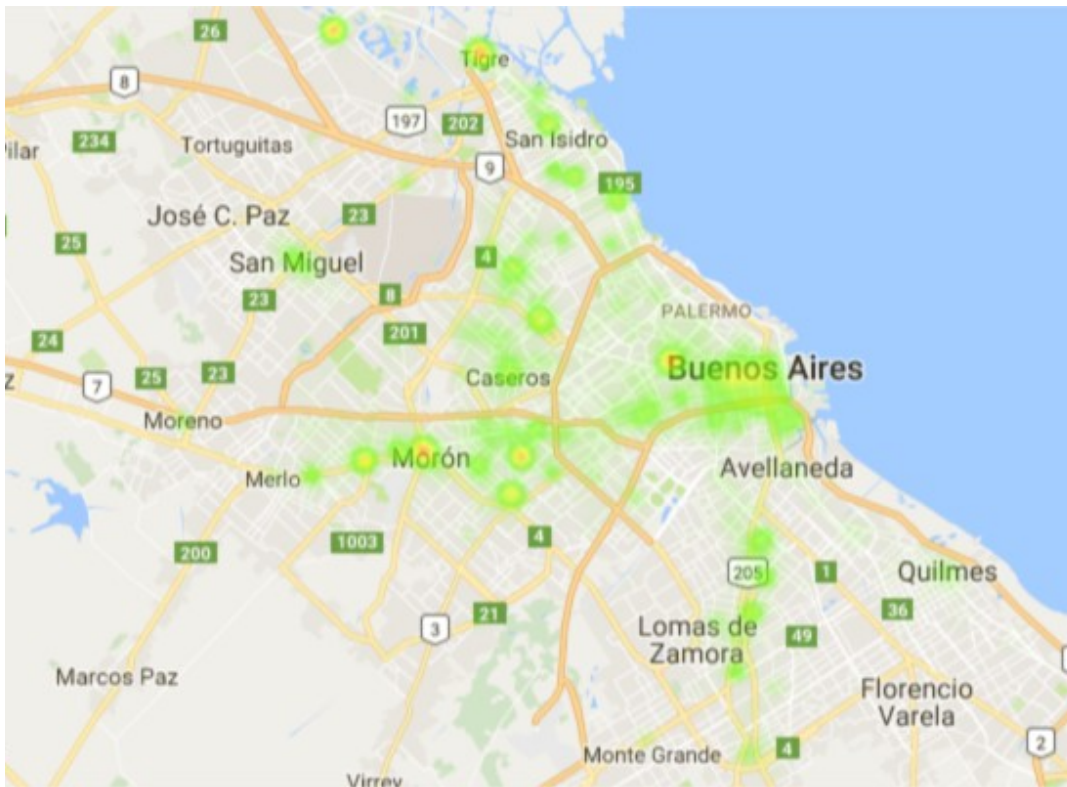
ZOOM al HEATMAP anterior



Hay una clara concentración en la capital, con la excepción de centros de los cordones mas inmediatos de GBA. Viendo el siguiente mapa de trenes y subterráneos, este muestra como el tendido de redes se expande a medida que uno se aleja del centro de CABA. El heatmap parecería tener una forma muy similar sugiriendo que las propiedades de mayor precio podrían estar situadas cerca de paradas de subterráneo. Teniendo en cuenta que la distribución de paradas de colectivos por la capital es mas pareja, esto apoyaría a la hipótesis planteada.



Heatmap CABA + GBA propiedades con precio por metro cuadrado menor a USD\$2000 y el peso del heatmap puesto en la cantidad de paradas



La cantidad de propiedades menores a USD\$2000 parecerían ser las de menor cantidad de paradas de transporte publico cercanas. Eso se denota ya que los heatmaps están armados utilizando como peso la cantidad de paradas cercanas. Es decir que mientras mas paradas cercanas tenga, mas roja se vera la ubicación en el mapa. A la vez se ve que las concentraciones de lugares con muchas paradas se ubican en las principales zonas de GBA (a nivel de población) siendo estas Morón, San Miguel, Tigre, Lomas de Zamora, entre otras.

Zonas de CABA + GBA por mayor cantidad de paradas:

Se ubican las propiedades con mas de 35 paradas de transporte publico cercanas.

Heatmap CABA + GBA propiedades con mas de 35 paradas cerca y el peso del heatmap puesto en la cantidad de paradas



Gran parte de estos pertenecen al microcentro de CABA, mientras que a la vez se encuentran ciertos puntos con una alto concentracion:

Lanus

San Justo

Constitucion

Quilmes

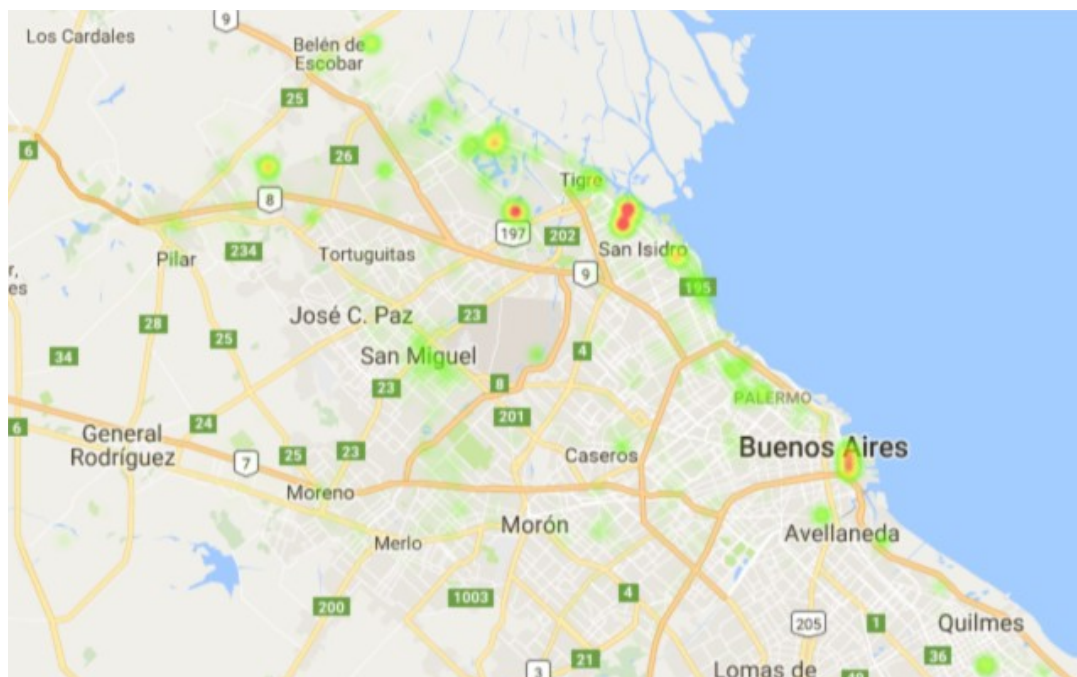
Zoom en CABA



Lógicamente, las propiedades con mayor cantidad de paradas cercanas se encuentran en el micro-centro porteño, de donde salen todas las líneas de subterráneo y una abundante cantidad de colectivos. A la vez Plaza Italia, Plaza Miserere y la zona de la Facultad de Medicina funcionan como puntos principales.

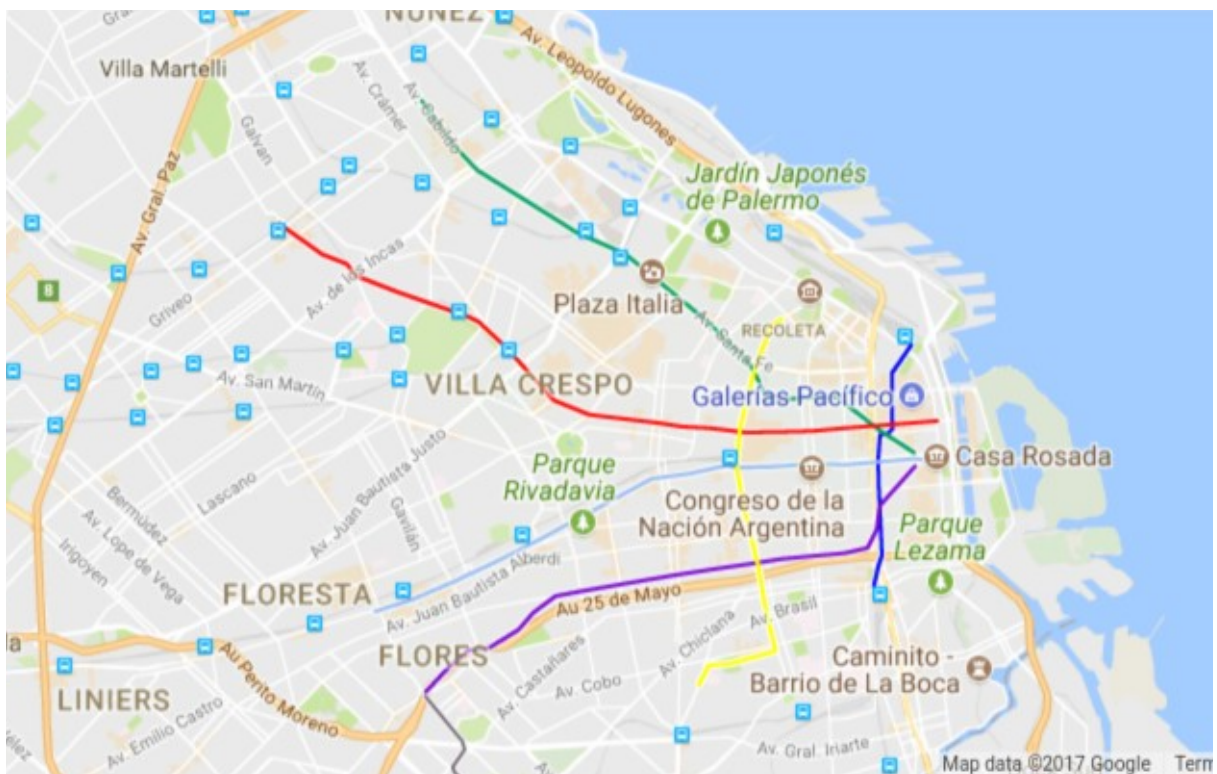
Zonas de CABA + GBA por menor cantidad de paradas:

Se ubican las propiedades con 7 o menos paradas de transporte público cercanas. (Heatmap con el peso en la cantidad de paradas)

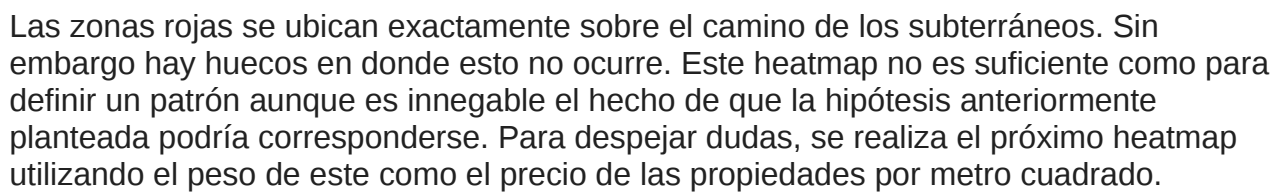


Distribución de las cantidad de paradas contra el precio por metro cuadrado en CABA (desarrollo de hipótesis de red de transporte):

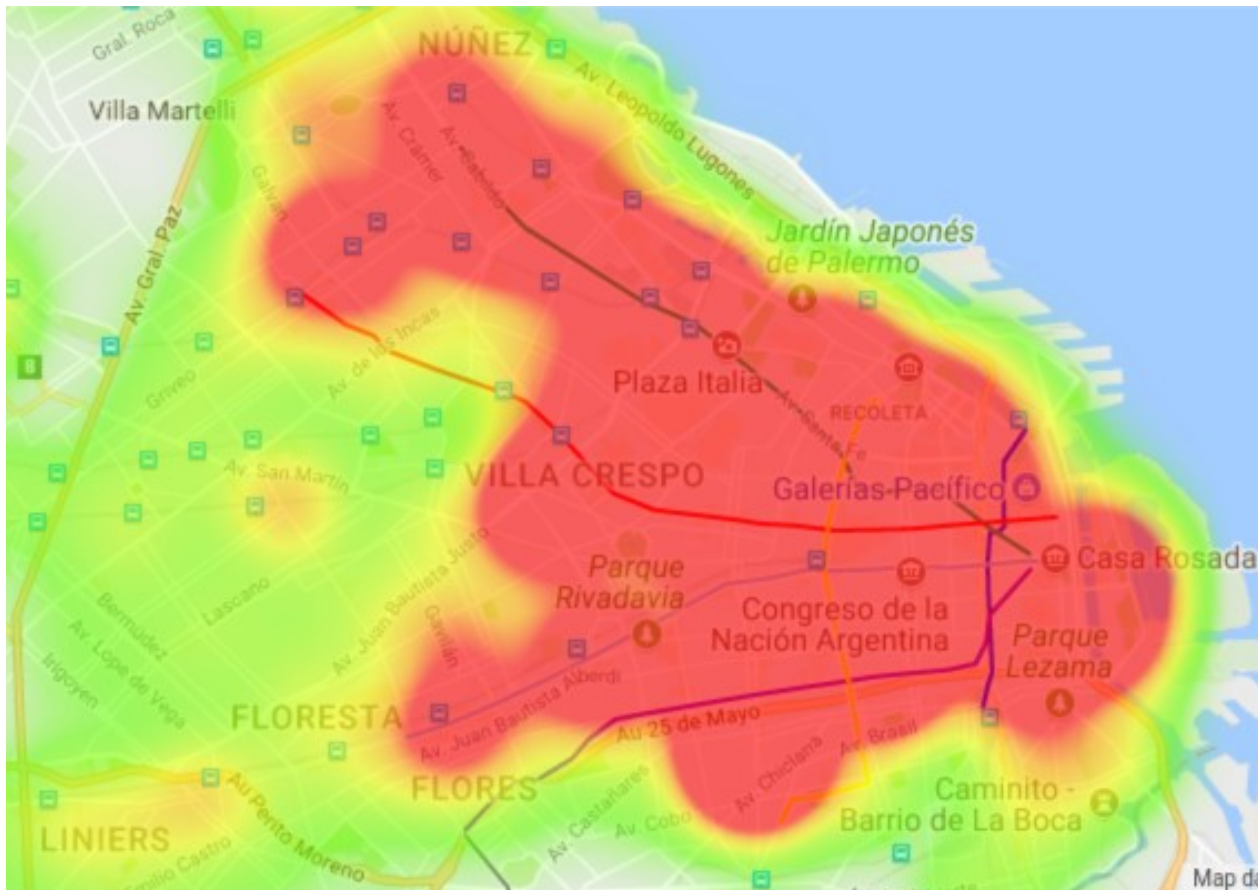
- Línea A (Celeste)
- Línea B (Roja)
- Línea C (Azul)
- Línea D (Verde)
- Línea E (Violeta)
- Línea H (Amarilla)



Heatmap con el peso en la cantidad de paradas cercanas:



Heatmap con el peso en el precio por metro cuadrado:



Casi en su totalidad, las propiedades cubren las líneas de subterráneo. Con la salvedad de un pequeño tramo de la línea B. Esto marca una clara tendencia en la relación directa entre el precio de una propiedad con la cercanía a las paradas de subterráneo en la Capital Federal.