

Tienda de Café en Buenos Aires

Estudio de ubicación a partir de datos de FourSquare

Jorge Palacio

Marzo de 2003

1. Introducción

Nuestro cliente es un inversor del área gastronómica, que desea abrir una tienda de café de especialidad en la ciudad de Buenos Aires. La persona en cuestión es un sommelier experto en café, y el público que espera atraer a su negocio es aquel que esté interesado en tomar café de primera calidad en un ambiente relajado. Por lo tanto, desea establecerse en un barrio donde ya haya cafeterías, para captar al público habituado a consumir en ese tipo de locales. El cliente desea que analicemos cada barrio de la ciudad, buscando la cantidad de cafeterías que tiene cada uno y las características de la mismas (popularidad, precio y opiniones de los usuarios).

2. Datos a utilizar

2.1.Fuentes de datos

Se utilizarán las siguientes fuentes de datos para confeccionar el informe:

- **Dataset con los barrios de la ciudad de Buenos Aires:** Se utilizará el archivo disponible en <https://cdn.buenosaires.gob.ar/datosabiertos/datasets/ministerio-de-educacion/barrios/barrios.geojson>. Dado que cada barrio en este archivo es un polígono, se calculará un centroide para cada uno, que será el dato de entrada para las consultas de FourSquare.
- **FourSquare Place Search:** Se buscarán todas las cafeterías presentes en cada barrio
- **FourSquare Get Place Details:** Se buscarán datos de popularidad, precio y rating para cada una de las cafeterías encontradas en el punto anterior.

2.2.Selección y limpieza de los datos

Los datos geográficos de los barrios de la ciudad de Buenos Aires se presentan en formato GeoJson, que es apto para ser procesado directamente en Python, utilizando las librerías Folium y Shapely. No se requiere tareas de post-procesamiento ni de limpieza.

Respecto a los datos de FourSquare, las consultas del endpoint *Place Search* tienen un límite de 50 respuestas por consulta. Por lo tanto, para poder consultar todos los cafés de la ciudad de Buenos Aires, se realizó una consulta por cada barrio, filtrando por la categoría *13032 (Dining and Drinking > Cafes, Coffee, and Tea Houses)*. Se comprobó que ninguna de ellas excedía el límite de repuestas, y solo utilizamos los datos básicos que devuelve la consulta (ID de FourSquare de cada ubicación, nombre, coordenadas y categoría).

El endpoint *Place Details* se utilizó para enriquecer la consulta anterior, agregándole a cada resultado un valor de popularidad, precio y rating. Algunos cafés no tienen todos estos datos disponibles, en esos casos se decidió dejar esos valores en *None*, de forma tal que sean ignorados al calcular los valores promedio en las estadísticas agrupadas.

El resultado de correr todas las consultas concluyó con 949 cafés dentro de la ciudad. A partir de este dataset, se analizaron los barrios para determinar cuáles son los que mejor cumplen con los requisitos del cliente.

3.2. Análisis estadístico por barrio

Si bien el análisis anterior nos da una primera idea de la ubicación de los cafés, nuestro cliente está interesado en zonas con cafés de alta calidad. Para determinar estas zonas, realizamos un análisis estadístico de las características de los cafés de cada barrio, que incluye la cantidad de cafés, y los valores promedio de popularidad, precio y rating (Tabla 1).

Tabla 1: Cafés por barrio

BARRIO	#CAFES	AVG POPULARITY	AVG PRICE	AVG RATING
RETIRO	19	0.99	1.00	7.74
VILLA RIACHUELO	1	0.98	1.00	NaN
RECOLETA	33	0.98	1.21	7.46
COGHLAN	5	0.97	1.00	6.83
PALERMO	71	0.96	1.26	7.51
SAN NICOLAS	42	0.96	1.29	7.54
PARQUE AVELLANEDA	5	0.95	1.00	NaN
VELEZ SARSFIELD	4	0.95	1.00	6.80
BALVANERA	23	0.93	1.15	6.86
ALMAGRO	19	0.93	1.18	7.03
PARQUE CHAS	8	0.92	1.17	6.33
COLEGIALES	19	0.91	1.16	7.31
VILLA CRESPO	28	0.91	1.13	7.25
PARQUE CHACABUCO	10	0.91	1.13	7.50
BELGRANO	54	0.91	1.19	7.11
MONSERRAT	41	0.88	1.23	6.98
NUÑEZ	21	0.87	1.14	6.60
FLORESTA	13	0.86	1.00	6.30
CHACARITA	11	0.86	1.20	7.37
BOEDO	12	0.85	1.11	7.34
VILLA GRAL. MITRE	9	0.85	1.38	6.23
SAN TELMO	19	0.81	1.16	7.55
LINIERS	24	0.81	1.07	7.23
VILLA ORTUZAR	11	0.81	1.29	6.88
VILLA LURO	11	0.80	1.22	7.50
CABALLITO	66	0.80	1.08	6.79
PUERTO MADERO	13	0.77	1.00	7.43
VILLA DEVOTO	37	0.76	1.09	6.73
VERSALLES	2	0.75	2.00	NaN

BARRIO	#CAFES	AVG POPULARITY	AVG PRICE	AVG RATING
BARRACAS	28	0.73	1.00	6.50
SAAVEDRA	30	0.72	1.18	6.73
MATADEROS	20	0.71	1.18	6.44
FLORES	46	0.70	1.03	6.85
VILLA DEL PARQUE	25	0.69	1.15	6.64
VILLA URQUIZA	34	0.69	1.24	6.99
AGRONOMIA	7	0.67	1.17	7.57
PATERNAL	10	0.66	1.00	7.40
BOCA	17	0.66	1.13	6.11
CONSTITUCION	14	0.65	1.14	5.80
VILLA PUEYRREDON	10	0.65	1.00	8.20
VILLA SANTA RITA	11	0.64	1.33	7.20
MONTE CASTRO	12	0.62	1.13	6.40
PARQUE PATRICIOS	22	0.62	1.15	6.70
VILLA LUGANO	8	0.58	1.00	7.40
SAN CRISTOBAL	10	0.55	1.00	5.90
NUEVA POMPEYA	9	0.12	1.00	NaN
VILLA REAL	4	NaN	NaN	NaN
VILLA SOLDATI	1	NaN	NaN	NaN

3.3.Ranking de los barrios

El factor fundamental que utilizaremos para rankear a los barrios será la popularidad, ya que el principal interés de nuestro cliente es establecerse en una zona gastronómica. La Tabla 1 ya está rankeada por popularidad, pero se observa que en los primeros puestos hay cafés con bajo promedio de rating. Por ejemplo, Coghlan, tiene 6.83 de promedio de rating, y Villa Riachuelo tiene un solo café, sin datos de rating.

Por lo tanto, se decidió definir dos umbrales para seleccionar nuestros barrios objetivo: un valor de popularidad promedio mayor a 0.95 y un valor de rating promedio

mayor a 7. De aquí se obtuvieron cuatro barrios, que se listan en la Tabla 2 y se grafican en la Ilustracion 2.

Tabla 2: Barrios seleccionados

BARRIO	#CAFES	AVG POPULARITY	AVG PRICE	AVG RATING
RETIRO	19	0.99	1.00	7.74
RECOLETA	33	0.98	1.21	7.46
PALERMO	71	0.96	1.26	7.51
SAN NICOLAS	42	0.96	1.29	7.54

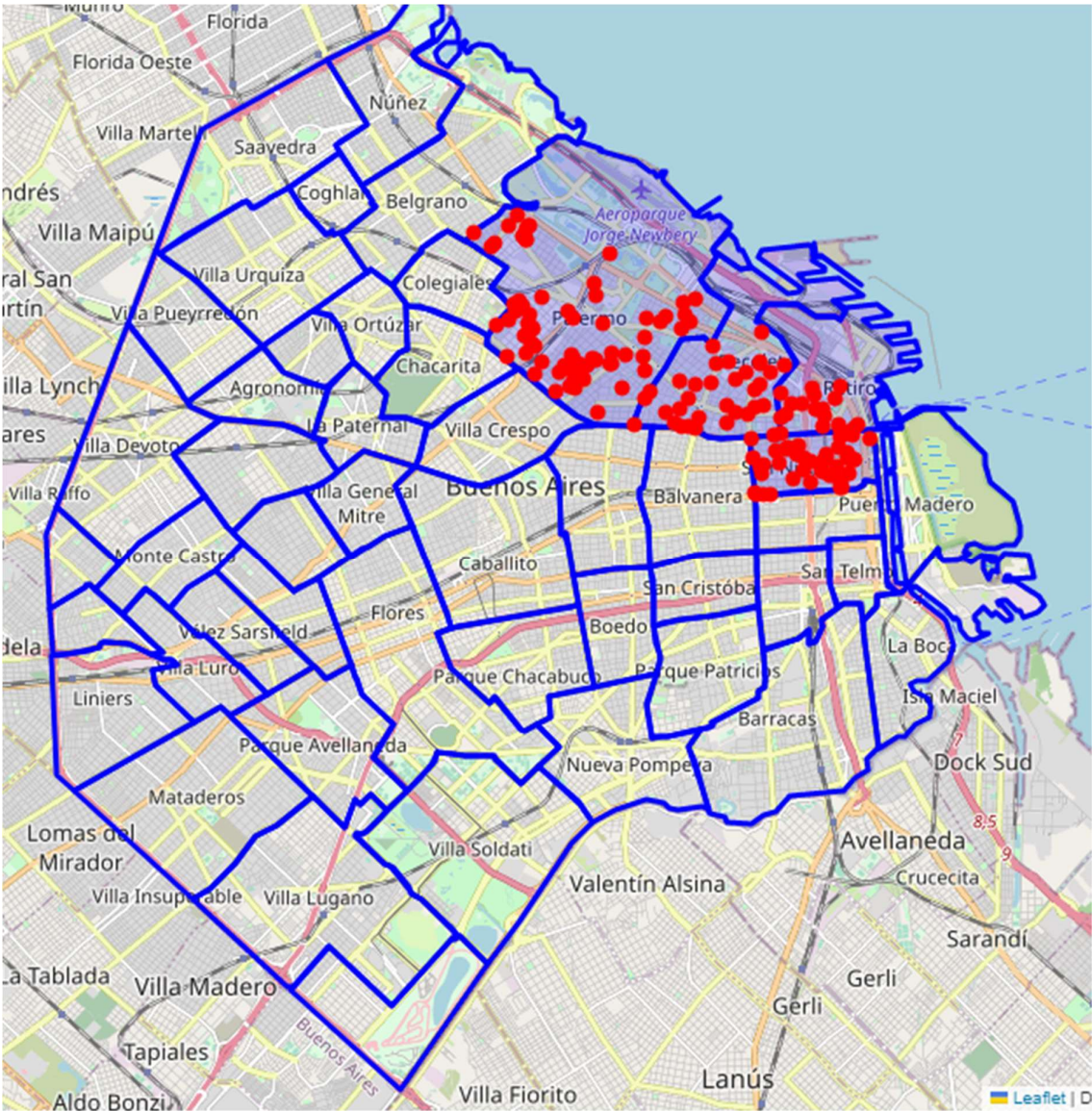


Ilustración 2: Barrios seleccionados

4. Resultados

A través del estudio de los datos de FourSquare, se pudieron detectar cuatro barrios que cumplen con los requisitos planteados inicialmente por el cliente, y se generaron mapas que servirán como guía para seleccionar la ubicación final del café que se desea abrir. Este trabajo es el primer paso de un análisis más profundo, que deberá realizarse luego de presentar esta información al cliente y teniendo en cuenta el feedback que este nos provea.

5. Próximos pasos

Se deberá coordinar una reunión con el cliente, para mostrar los resultados obtenidos y obtener feedback sobre los mismos. Luego, se puede proseguir de la siguiente manera:

- Obtener información sobre locales en alquiler o en venta, con las características que el cliente requiera, y ubicarlos sobre los mapas generados en este informe.
- Realizar análisis de distancias a los cafés más cercanos, para evitar ubicarse en un sitio demasiado cercano a un competidor directo. En este análisis se tomará en cuenta el factor precio, para determinar cuáles cafés son competidores directos.
- Estudiar con el cliente cualquier otro requisito que nos presente y agregar el análisis correspondiente en una versión futura del informe.

6. Conclusiones

A partir del análisis de los datos geográficos disponibles, se pudo establecer un área objetivo para la ubicación del negocio del cliente, basándonos en sus requisitos. Este informe es el punto de partida de un análisis más detallado, que al agregar más información y feedback del cliente culminará con la propuesta de locales específicos para que el cliente pueda evaluar.