

Tópicos de Analítica de Datos con SQL Avanzado

Trabajo práctico N°2

Ejercicio 1

- Hay 12 puntos de interés.
- Hay 27 líneas.
- El único día disponible es el 15 de mayo de 2019.
- Hay un promedio de 5,14 ascensos por línea, un mínimo de 1 y un máximo de 109..

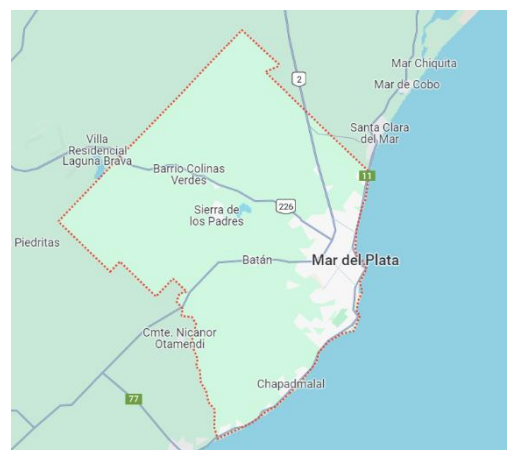
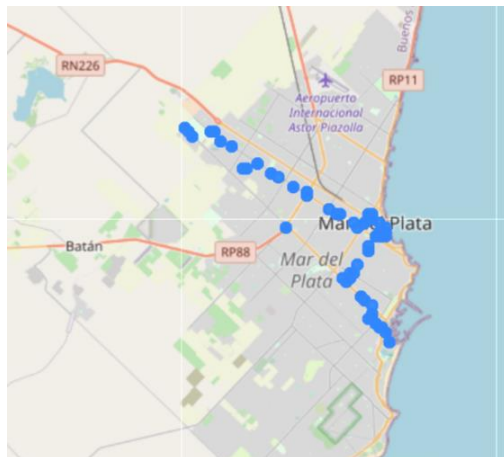
Ejercicio 2

Código de la vista en el archivo.

De los 12 puntos de interés, solo 3 están a menos de 500m de los puntos de transacciones.

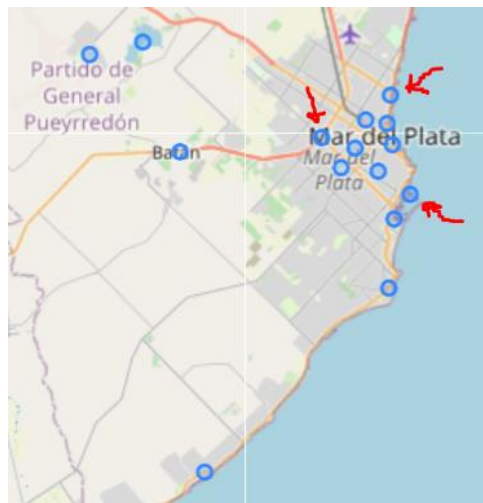
Ejercicio 3

Haciendo una inspección visual entre los puntos de las transacciones (izquierda) y la región del municipio obtenida de Google Maps (derecha), se puede verificar fácilmente que todos los puntos están en General Pueyrredón.



Ejercicio 4

Añado el **Museo Municipal de Ciencias Naturales**, el **Museo de la Fuerza de Submarinos** y el **Monumento al Gaucho**. Confirmando que no están a menos de 1200m de distancia del resto de los puntos de interés (ver código). Estos nuevos puntos están apuntados con flechas rojas (ver imagen).



Ejercicio 5

Las horas de **mayor demanda** (elijo 3 entre las 6 y 20 hs) en General Pueyrredón son las **12, 17 y 13 hs**, en ese orden. En cambio, las horas de **menor demanda** son las **6, 20 y 19 hs**.

El día considerado es un miércoles, por lo que esto puede deberse a que esos horarios suelen ser donde los chicos salen de la escuela (12-13hs) o algunos trabajadores terminan su jornada (17hs).



Por otro lado, si consideramos aquellos lugares cercanos al Casino Central, la Terminal de Ómnibus, la FCEyN (UNMdP), el Museo Municipal de Ciencias Naturales, el Museo de la Fuerza de Submarinos y el Monumento al Gaucho, las horas de **mayor demanda** (entre las 6 y 20 hs) son las **17, 16 y 18 hs**, en ese orden. Las de **menor demanda** son las **6, 8 y 10 hs**.

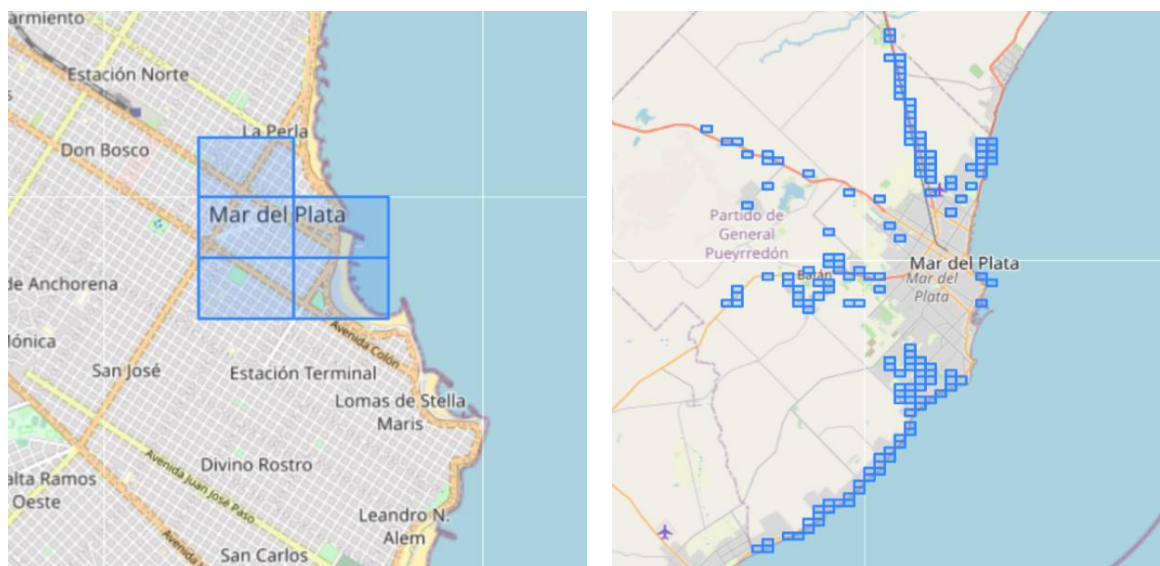
Debido a que los puntos de interés están más que nada en zonas céntricas, es de suponer que estos horarios de mayor demanda se deben a gente que sale de trabajar.



Cantidad de pasajes por hora cerca de los puntos de interés mencionados

Grafico las 5 zonas de mayor cantidad de transacciones económicas registradas en el día (**izquierda**) y las zonas en las que pasa 1 sola línea (**derecha**).

En el primer caso, las zonas están en pleno centro de Mar del Plata, mientras que en el segundo caso las zonas están en la periferia de la ciudad. Estos resultados son consistentes con la densidad de personas que típicamente hay en el centro de una ciudad (mucho) y en sus afueras (poco).



Se cumple que la línea con mayor número de traslados tiene mayor máximo de ascensos (7), agrupado por geohash- tiempo, que el máximo de ascensos (6) de la línea de menor número de traslados.

Además, la línea que pasa por más puntos de interés es la LINEA_511.