

ANALISIS DE LOS DATOS DE BROADNET S.A.

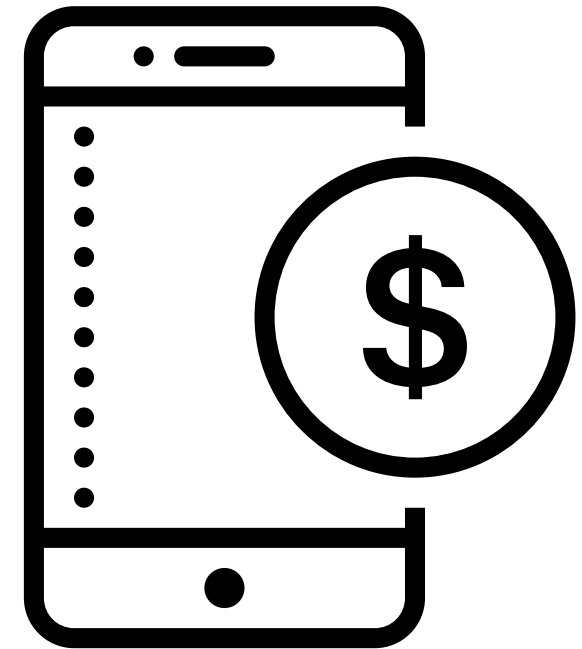
Grupo #9

10/09/2022

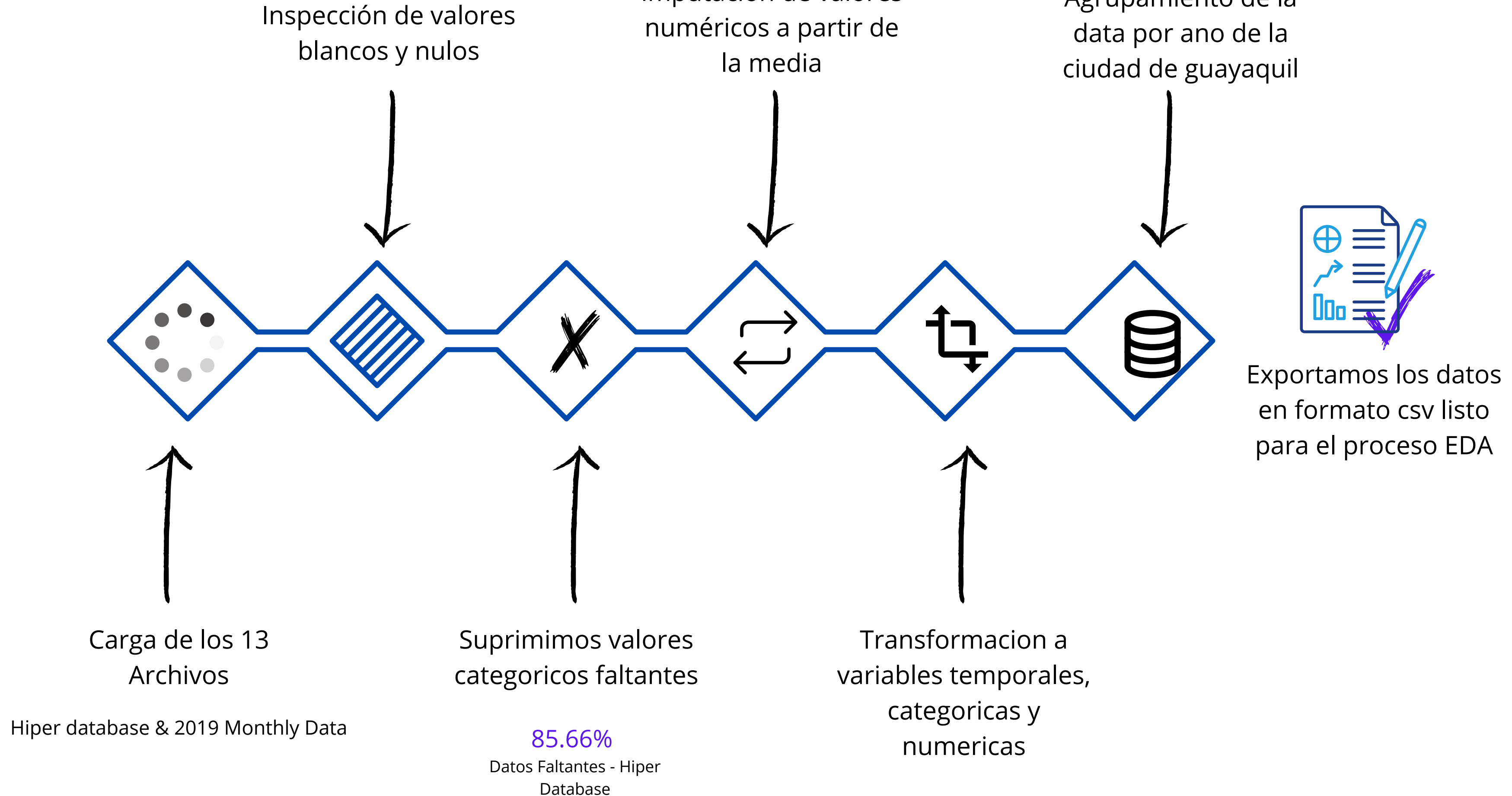


Antecedentes del DataSet:

- Los DATASETS propuestos para este proyecto corresponden a una empresa Ecuatoriana pioneros en Red de Servicios Independiente del Ecuador dedicada al desarrollo y comercialización de productos digitales y al procesamiento de transacciones.
- La muestra original de datos consta de 2 tipos de archivos:
 - Pdvs_hiper (detalla los puntos de ventas por ubicación 2008 hasta 2020).
 - Ventas (detalla transacciones en pdvs período 2019. Contiene:
 - 13 columnas (features)
 - 189.324 registros)
- La muestra de datos de las ventas consta de 12 archivos que contienen el detalle de las recargas mensualizadas de enero a diciembre del 2019, donde aproximadamente cada una tiene 4 millones de registros.



Proceso ETL



Detalles de campos del dataset

Característica	Tipo de dato	Descripción
fecha	Temporal	Fecha ingreso de cada transacción.
hora	Temporal	Hora de cada transacción.
Name	Categórico	Identificador del grupo donde se realiza la transacción.
comercio	Categórico	Identificador del comercio donde se realiza la transacción.
proveedor	Categórico	Identificador del proveedor donde se realiza la transacción.
producto	Categórico	Identificador del producto de la transacción.
cliente_final		

(70410, 16)

Detalles de campos del dataset:



Característica	Tipo de dato	Descripción
medio	Categórico	Identificador del medio de la transacción.
total	Numérico	Representa el total de la transacción .
cliente_final	Object	Identificador del cliente que realiza la transacción (Por seguridad están encriptados).
saldo_despues	Numérico	Representa el saldo despues de la transacción .

Features Seleccionados

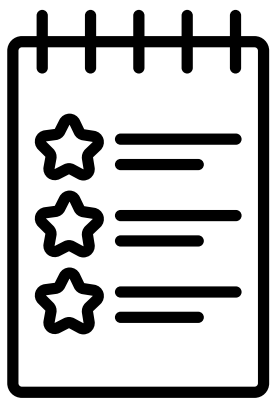
Ciudad de Guayaquil



Característica	Tipo de dato	Descripción
DateTime	Object	Fecha de ejecucion de la transaccion.
Hora	Object	Hora de ejecucion de la transaccion.
Name	Object	Nombre del barrio en que se realiza la transaccion .
comercio	Float64	Identificador del comercio al que pertenece el punto de venta.
Latitud	Float64	Latitud a la que pertenece el punto de venta.
Longitud	Float64	Longitud a la que pertenece el punto de venta.
Cliente_Final	Object	Cliente final que realiza transaccion encriptado

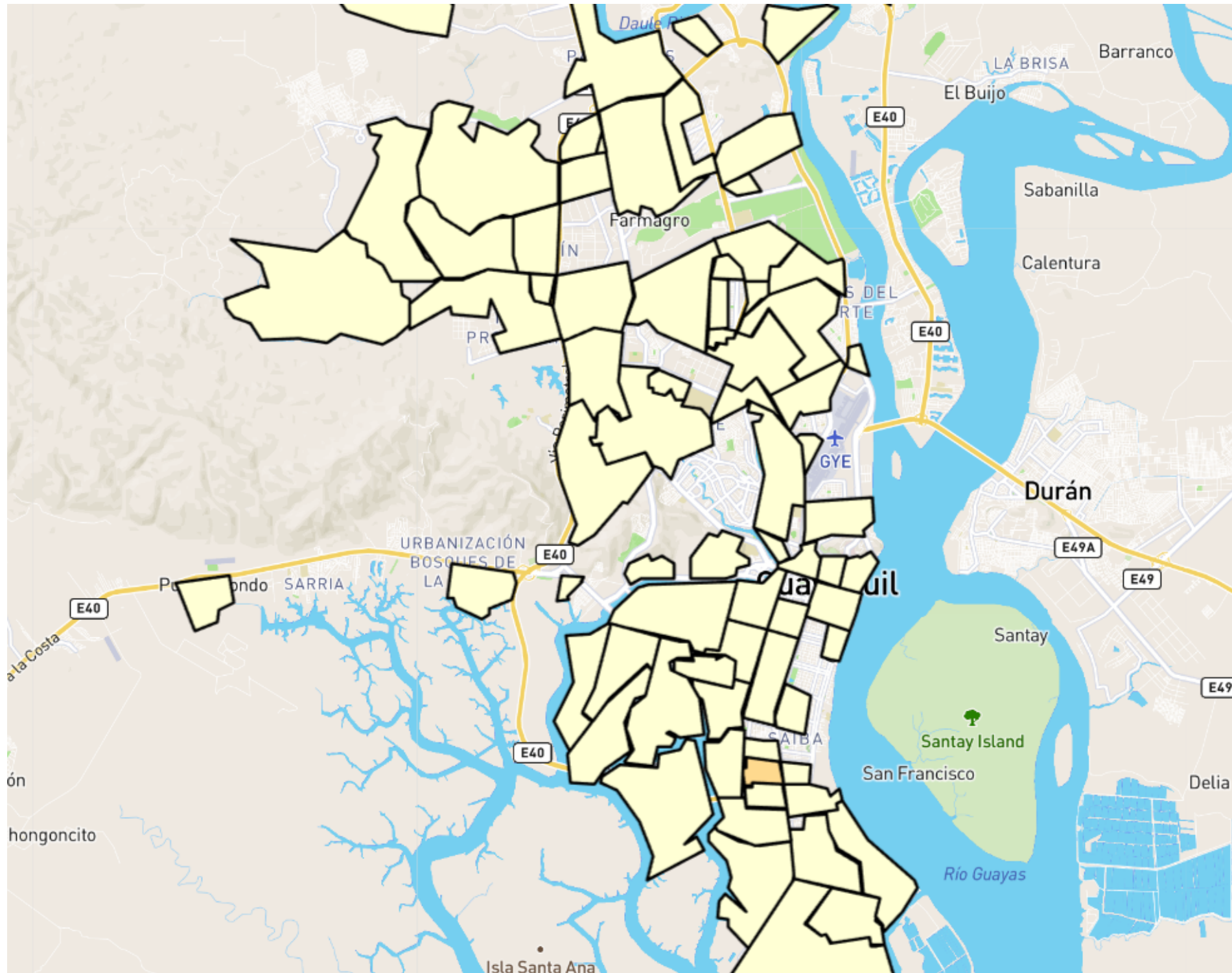
Features Seleccionados

Ciudad de Guayaquil



Característica	Tipo de dato	Descripción
Distance_Straight_line	Float64	Distancia de la línea recorrida desde Home hasta comercio .
Number_Of_Locations	int64	Numero de comercios.
N_visits	int64	Numero de visitas al comercio.
location_frequency	int64	Frecuencia de visitas al comercio.
Home_Ing	Float64	Latitud a la que pertenece el punto home del cliente.
Home_lat	Float64	Longitud a la que pertenece el punto home del cliente.

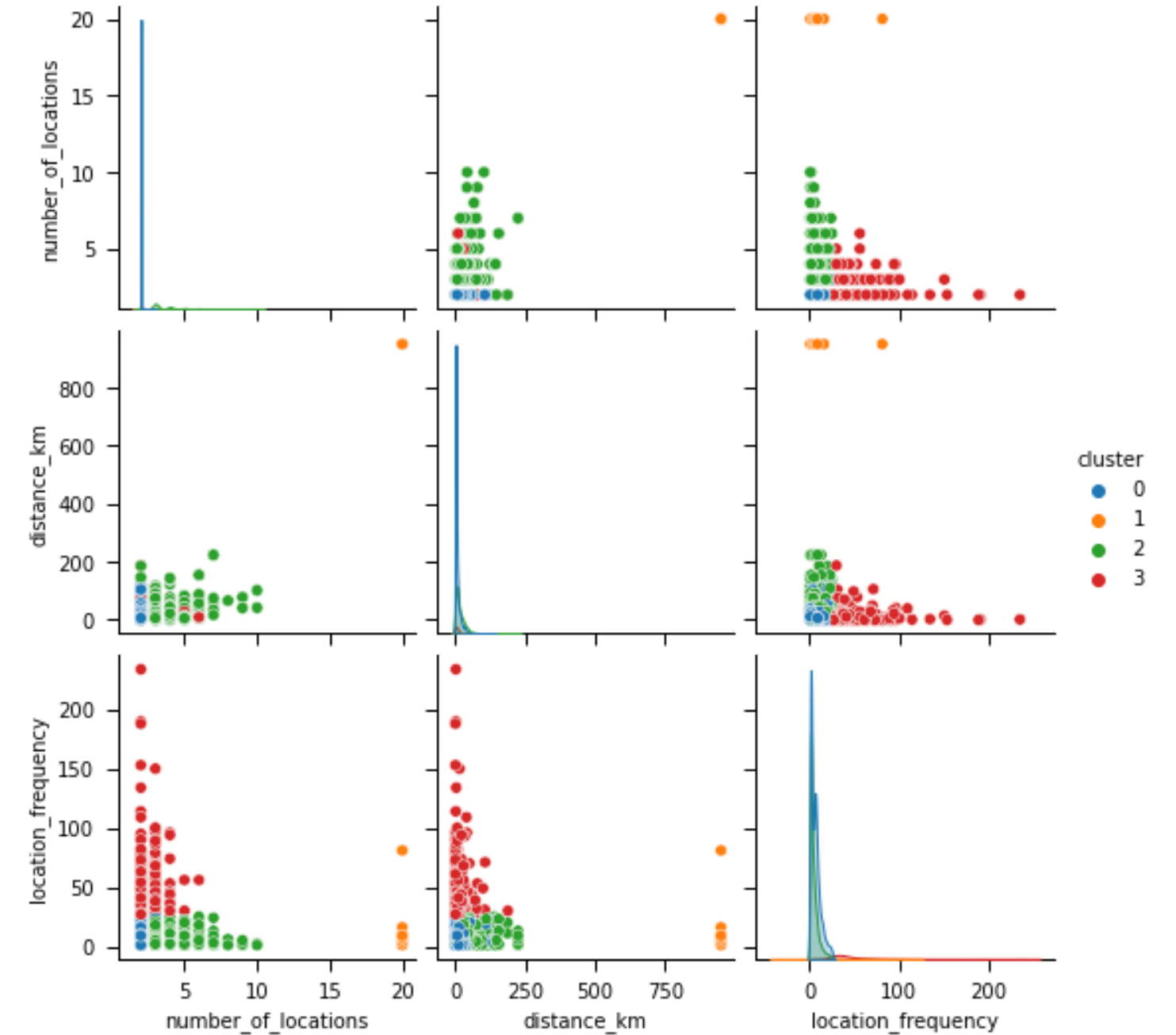
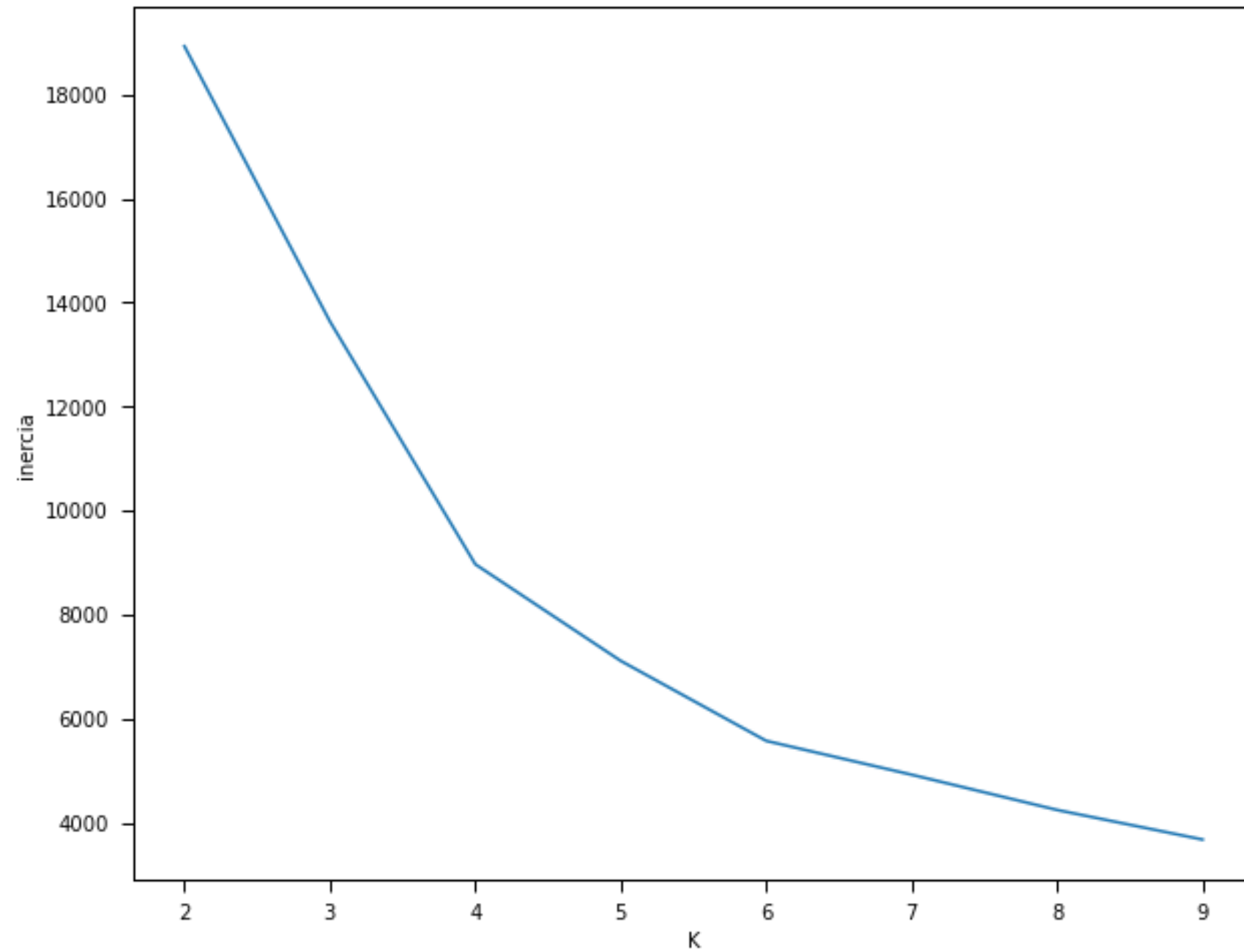
¿Cuáles son los comercios que mas transacciones registraron de usuarios de otros barrios en el 2019?



Comercio Concurrido

En el área de Huancavilca Sur se localiza el comercio más visitado en el 2019, que realizó 2.504 transacciones de usuarios de otros barrios, según el análisis de la data de ventas de Recarga de BROADDNET.

Grupos descriptivos respecto a los features: number_of_locations, distance_km, location_frequency



Predicción de la distancia recorrida por los usuarios en el 2019

Correlacion

	comercio	proveedor	latitud	longitud	distance_km	number_of_locations	n_visits	location_frequency
comercio	1.000000	0.013893	-0.219703	0.116640	0.017399	0.047731	-0.464006	-0.196978
proveedor	0.013893	1.000000	0.020112	0.019334	0.318512	0.256882	-0.022783	0.060457
latitud	-0.219703	0.020112	1.000000	-0.572423	-0.028373	-0.061221	-0.115333	0.025682
longitud	0.116640	0.019334	-0.572423	1.000000	0.060062	0.058475	-0.173077	-0.011742
distance_km	0.017399	0.318512	-0.028373	0.060062	1.000000	0.785092	-0.113099	0.026298
number_of_locations	0.047731	0.256882	-0.061221	0.058475	0.785092	1.000000	-0.041946	-0.014025
n_visits	-0.464006	-0.022783	-0.115333	-0.173077	-0.113099	-0.041946	1.000000	0.120575
location_frequency	-0.196978	0.060457	0.025682	-0.011742	0.026298	-0.014025	0.120575	1.000000

Predicción de la distancia recorrida por los usuarios en el 2019

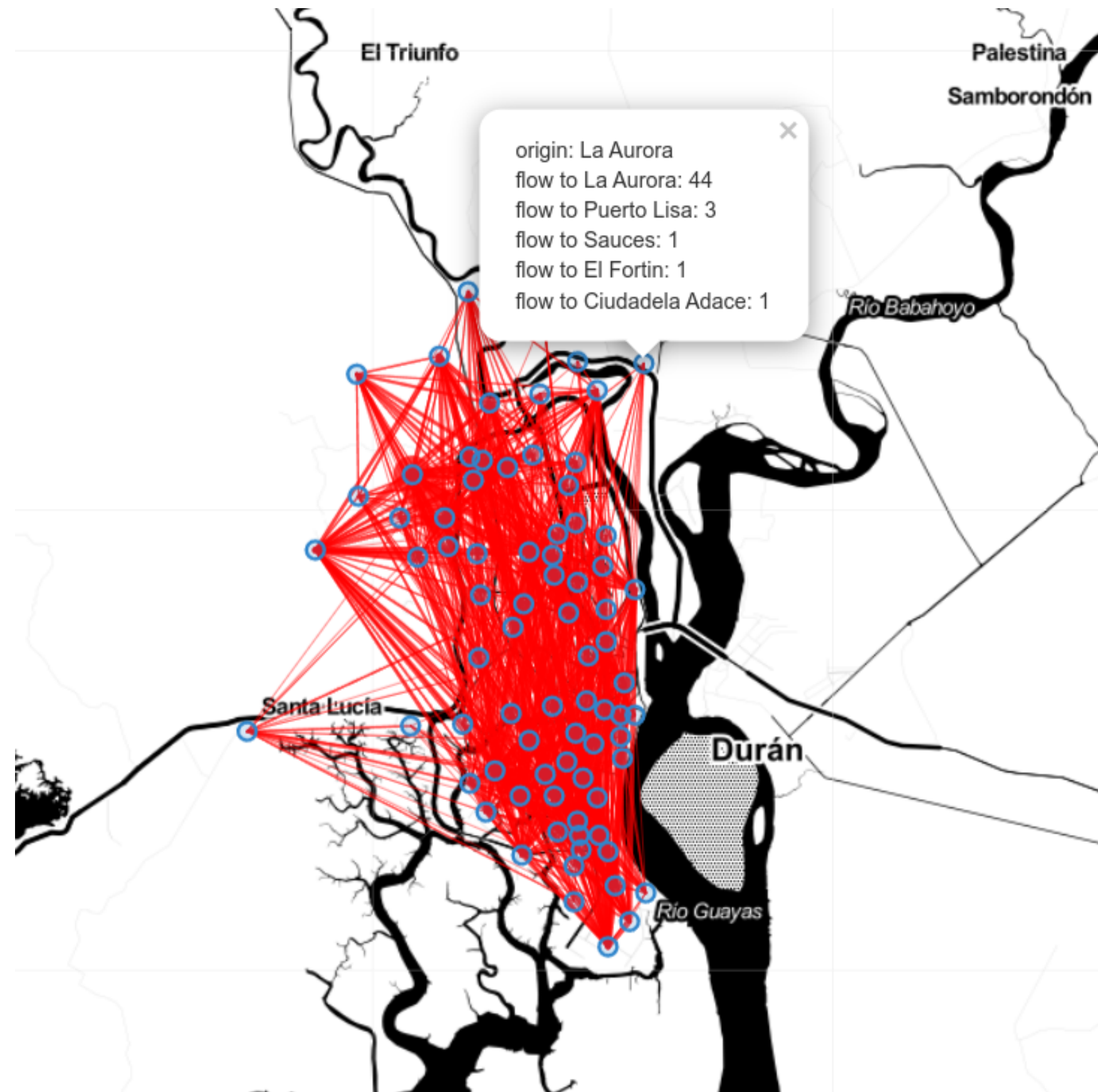
Regresión Lineal

El entrenamiento de este modelo supervisado se consideraron los siguientes features:

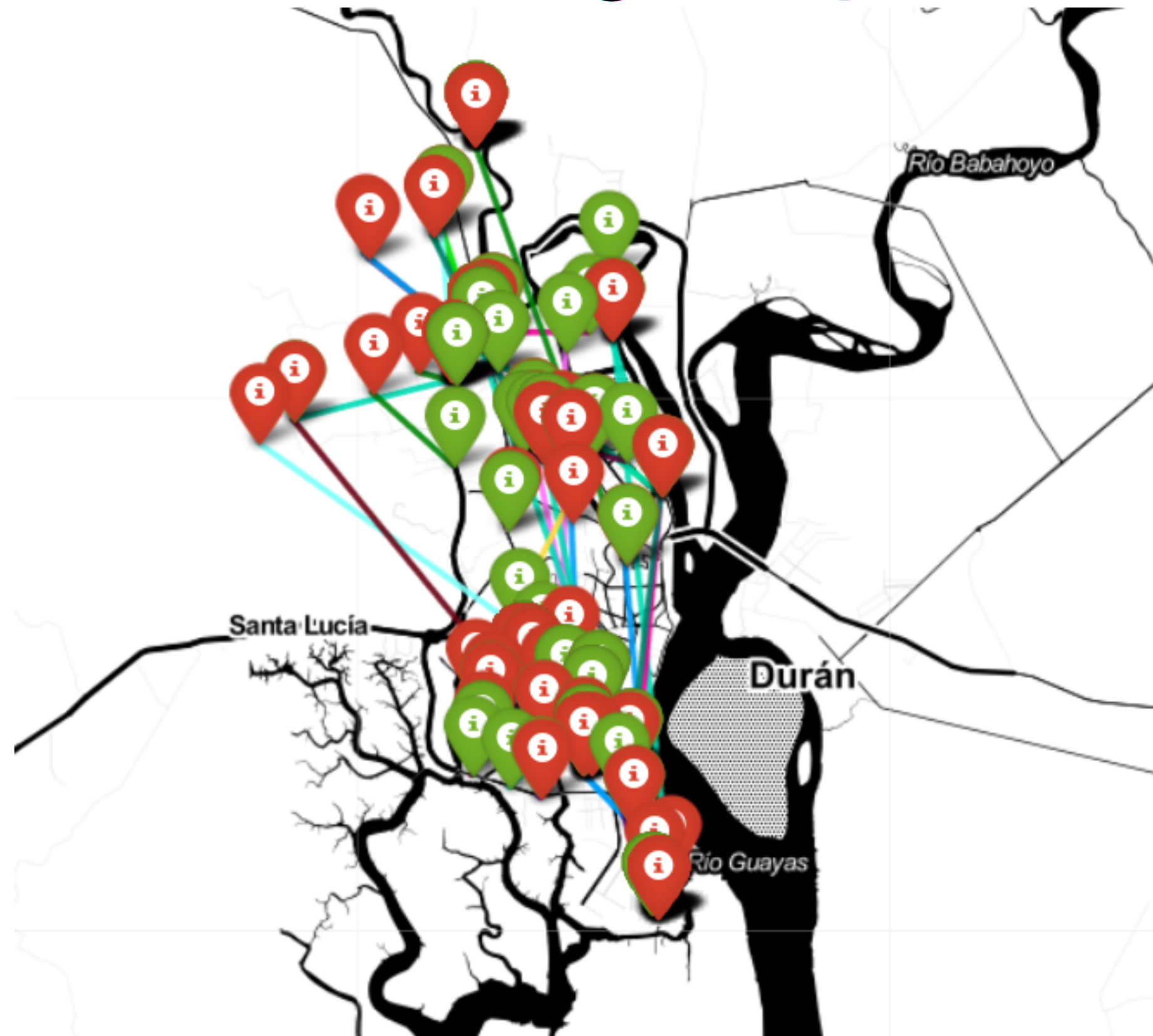
- "latitud","longitud"
- number_of_locations"
- "n_visits"
- "location_frequency"

Metricas de evaluacion	
MAE	17.68
Regression Score	0.636

¿Que relación tiene la movilidad de los barrios?



Cual es la movilidad de los primeros 200 usuarios de la ciudad de guayaquil?

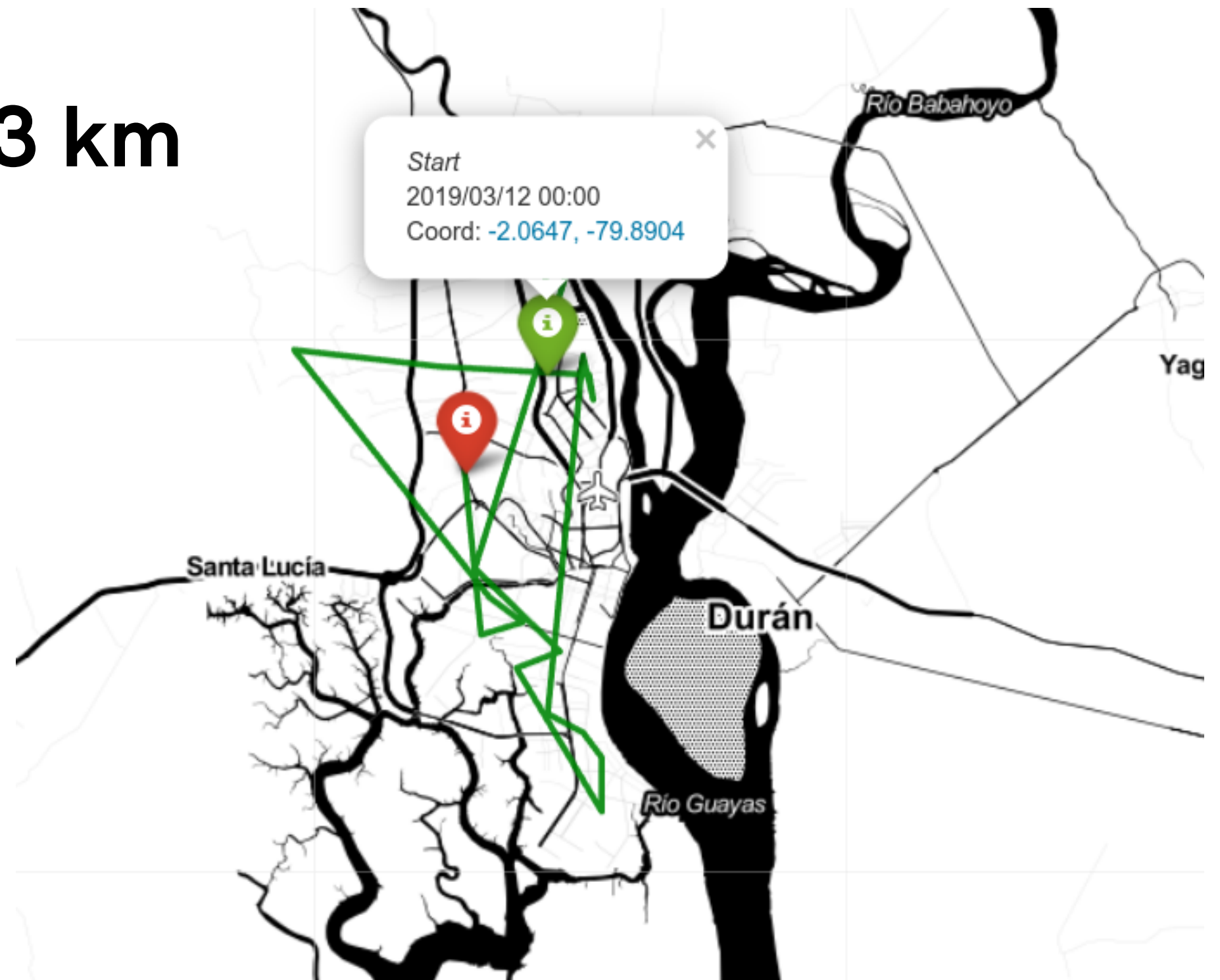


Cual fue la trayectoria de movilidad del usuario que estuvo en mas barrios de Guayaquil?

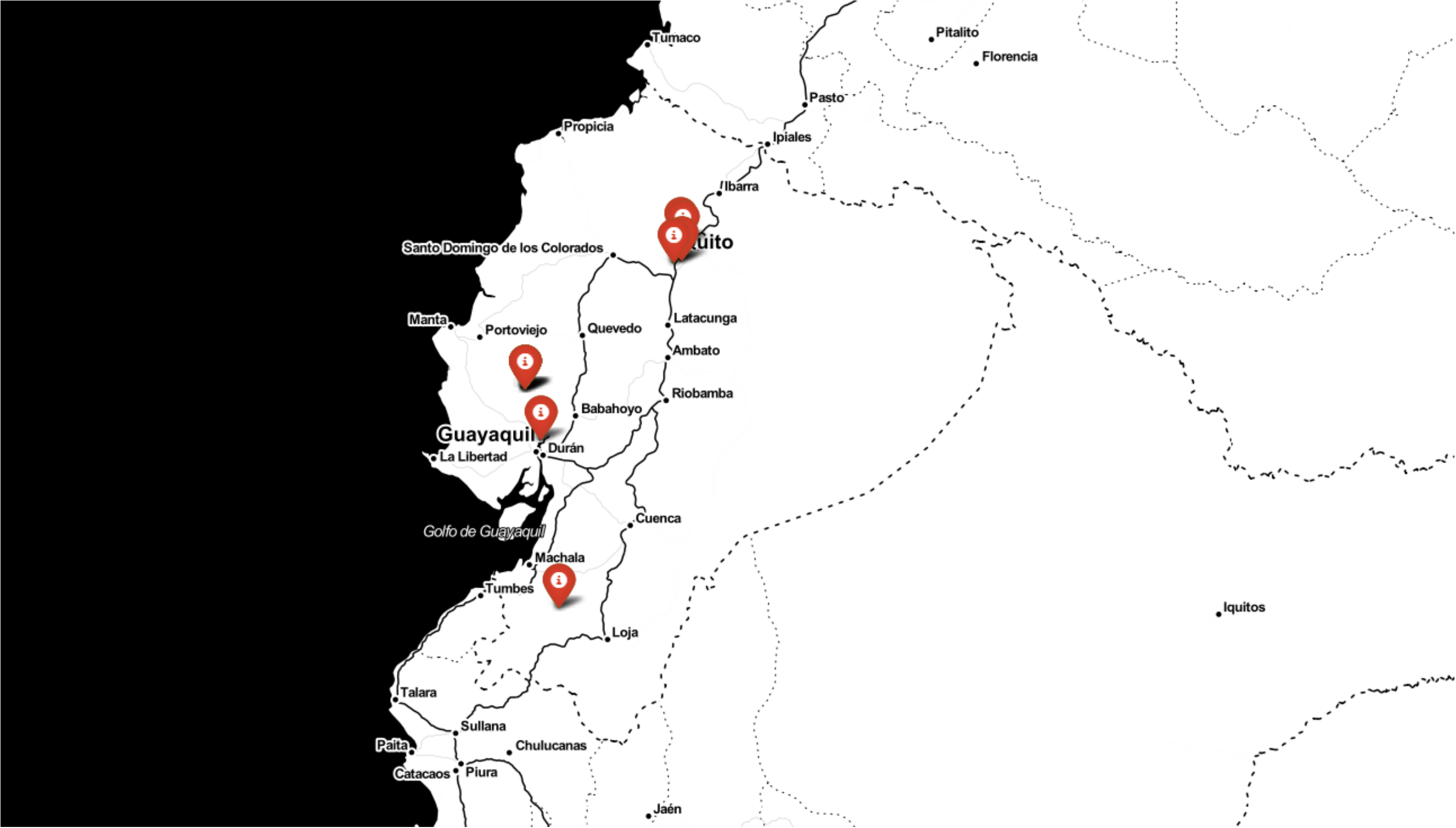
user_id = %%(

distancia recorrida = 950.053 km

'Febres Cordero', 'Guayacanes',
'Samanes', 'Sergio Toral',
'Mapasingue', 'Guasmo Sur - La Union',
'Parroquia Garcia Moreno',
'Sauces', 'Vergeles',
'Cristo del Consuelo', 'Guasmo Norte - Rio
Guayas', 'Puerto Lisa',
'Huancavilca Sur', 'Florida', 'Mucho Lote 2',
'San Eduardo', 'Pradera'



¿ Cuales son las ubicaciones mas visitadas por los clientes en el año 2019?



Ubicacion
Piñas
La Aurora
Colimes
Guamani
Conocoto
San Isidro

Conclusiones



Se trabajo con los usuarios que realizaron mas de 6 transacciones y para establecer su punto home, se considero el horario que realizo la transaccion, y el tipo del local donde hizo la transaccion.

Es posible medir la movilidad de los clientes mediante los distintos puntos de los comercios que visitaron en el año.

Se descartaron los usuarios que solo hicieron transacciones en un solo barrio, ya que no aportan valor al analisis de movilidad.

Bibliografia

[1] Human mobility models reveal the underlying mechanism of seasonal movements across China. https://www.researchgate.net/publication/356108123_Human_mobility_models_reveal_the_underlying_mechanism_of_seasonal_movements_across_China?enrichId=rgreq-667703fdc4ed5f5568822e4017eb8314-XXX&enrichSource=Y292ZXJQYWdlOzM1NjEwODEyMztBUzoxMTA1ODI0NTI1MDkwODE2QDE2NDA2NjA1MjAwMTQ%3D&el=1_x_3&_esc=publicationCoverPdf.

[2] Exploración y análisis de datos de telefonía celular para estudiar comportamientos de movilidad en la ciudad de buenos aires. https://www.researchgate.net/publication/319732756_Exploracion_y_analisis_de_datos_de_telefonia_celular_para_estudiar_comportamientos_de_movilidad_en_la_Ciudad_de_Buenos_Aires.

[3] Analysis of mobile phone call data to determine user characteristics and to enhance user experience. https://www.researchgate.net/publication/260553625_Analysis_of_Mobile_Phone_Call_Data_to_Determine_User_Characteristics_and_to_Enhance_User_Experience.