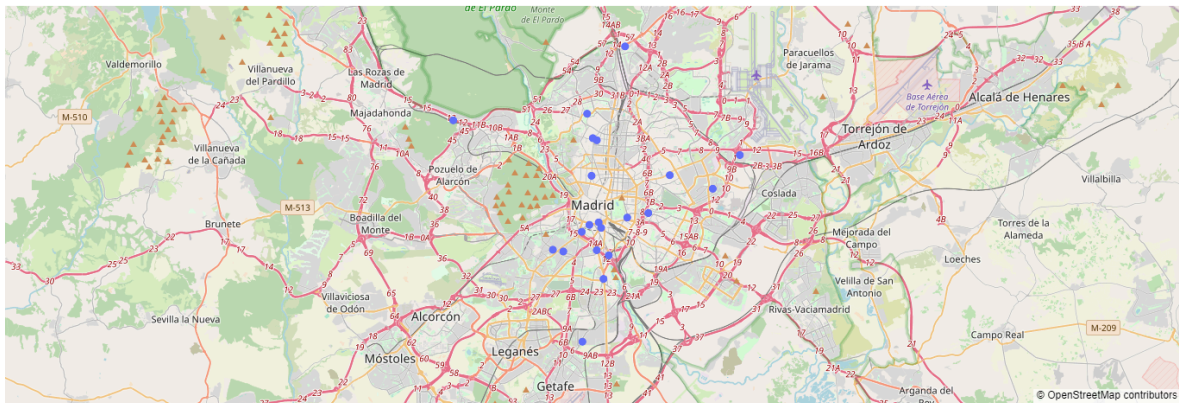


Entendimiento del negocio.

El negocio es una empresa que se dedica a la distribución de mercancías, necesitan que se calcule las rutas óptimas entre los clientes y el almacén, donde tenemos que cumplir con ciertos objetivos teniendo en cuenta las diferentes restricciones.

- **Objetivos:**
 - Minimizar el coste total.
 - Maximizar el número de pedidos entregados.
- **Restricciones:**
 - Capacidad de carga de los vehículos.
 - Autonomía limitada de los vehículos.
 - Todas las rutas empiezan y acaban en el almacén.

Aquí un mapa donde podemos ver los diferentes clientes y el almacén del negocio:



Entendimiento de los datos.

Para el proyecto número 1 nos han entregado 6 archivos excel con los datos que vamos a usar para entrenar el modelo y probarlo. Los diferentes archivos contienen varias variables que explicaré aquí abajo:

- **DF_Locations:** Archivo excel con datos de los diferentes clientes y el almacén.

Cliente	Object	Variable que da nombre a los 20 clientes y el almacén.
Latitud	Float64	Variable que marca la Latitud en un mapa de cada cliente.
Longitud	Float64	Variable que marca la Longitud en un mapa de cada cliente.

- **DF_Distance_Km:** Archivo excel que compara la distancia en km entre los 20 clientes y el almacén, los datos están en un formato de matriz de distancias.

Cliente_X	Float64	Distancias entre X cliente o almacén.
-----------	----------------	---------------------------------------

- **DF_Distance_Min:** Archivo excel que compara la distancia en minutos entre los 20 clientes y el almacén, los datos están en un formato de matriz de tiempo.

Cliente_X	Float64	Tiempo entre X cliente o almacén.
-----------	----------------	-----------------------------------

- **DF_Historic_Order_Demand:** Archivo excel con el historial de ordenes de pedidos de los diferentes clientes, cada cliente tiene un número de pedidos diferente por cada mes.

Cliente	Object	Variable que da nombre a los 20 clientes.
mes_anio	Object	Variable que da los datos de mes y año para cada cliente, el formato de la fecha dada es Mes-Año.
order_demand	int64	Variable que muestra los pedidos recibidos por cada cliente en la fecha dada.

- **DF_Orders:** Archivo excel con los datos de pedidos de los diferentes clientes durante el último mes, cada cliente tiene un número diferente de pedidos.

Cliente	Object	Variable que da nombre a los 20 clientes.
mes_anio	Object	Variable que da los datos de mes y año para cada cliente, el formato de la fecha dada es Mes-Año.
order_demand	int64	Variable que muestra los pedidos recibidos por cada cliente en la fecha dada.

- **DF_Vehicles:** Archivo excel con los datos de los vehículos que se usan para el reparto de pedidos, cada vehículo tiene un identificador, su capacidad de almacenamiento, el coste de uso por km y la autonomía en kilómetros.

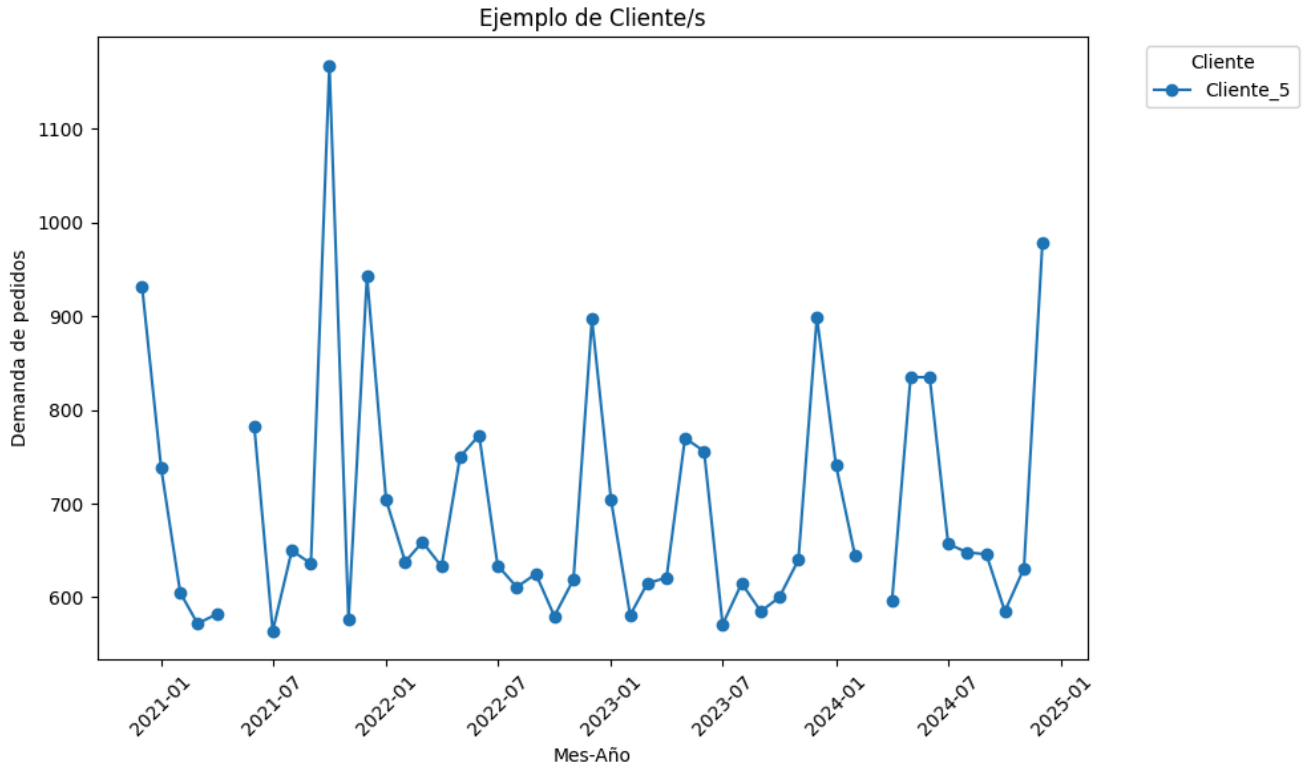
vehiculo_id	int64	Variable con el identificador de los vehículos usados para el transporte de mercancías.
capacidad_kg	int64	Variable que da los datos de peso máximo de mercancías que puede llevar cada vehículo.
costo_km	float64	Variable que da los datos de coste en euros por cada kilómetro recorrido de cada vehículo.
autonomia_km	int64	Variable que da los datos de la autonomía de cada vehículo sin necesidad de repostar.

Preparación de Datos.

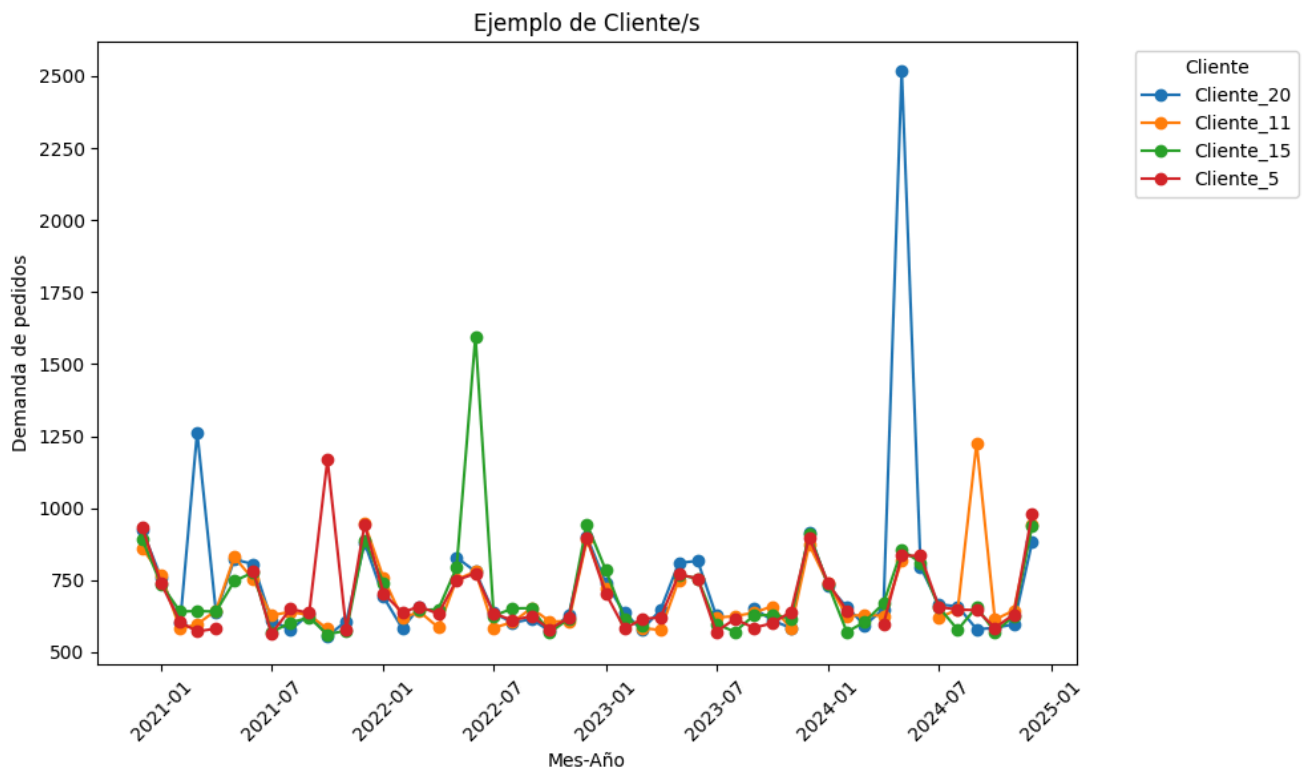
Tras la investigación de los datos, hemos encontrado que hay un DataSet (**df_historic_order_demand**) que requiere de imputación/eliminación de datos.

EJEMPLO: Cliente_5, le falta información de los siguientes meses_año:

- 05-2021
- 03-2024



Para ‘df_historic_order_demand’, hay ciertos valores nulos en diferentes datos de la columna ‘order_demand’.



Si nos fijamos, se podría imputar por una media de ese mes faltante, ya que los pedidos siguen un “patrón” de pedidos recibidos.