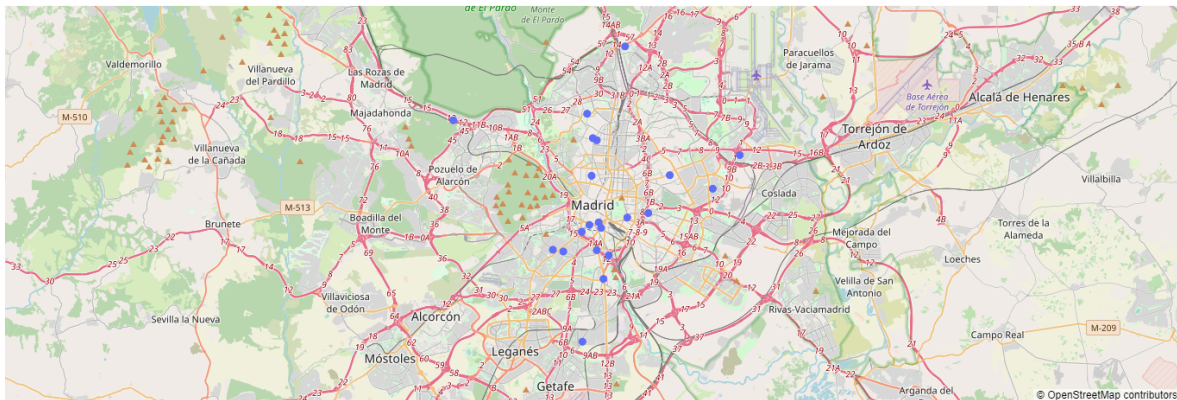


# Entendimiento del negocio.

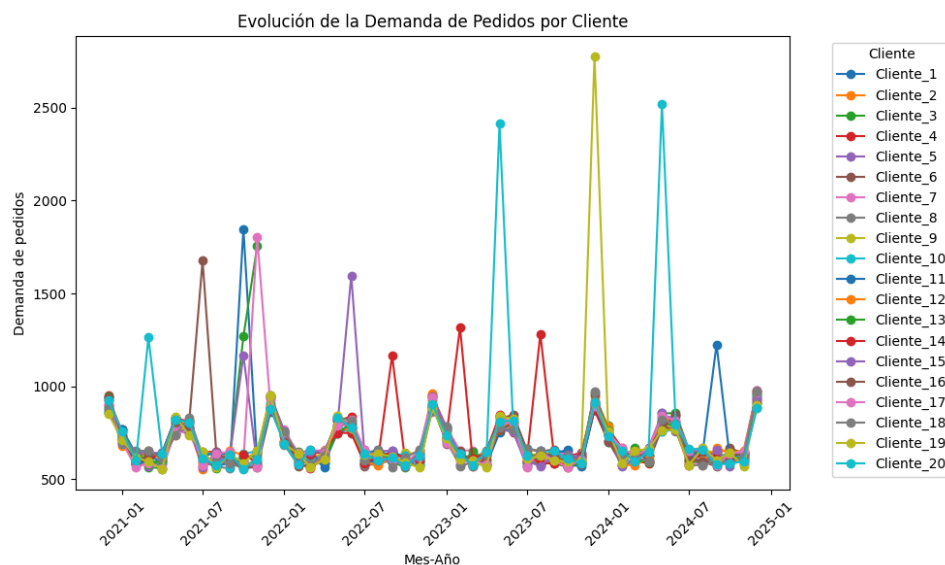
El negocio es una empresa que se dedica a la distribución de mercancías, necesitan que se calcule las rutas óptimas entre los clientes y el almacén, donde tenemos que cumplir con ciertos objetivos teniendo en cuenta las diferentes restricciones.

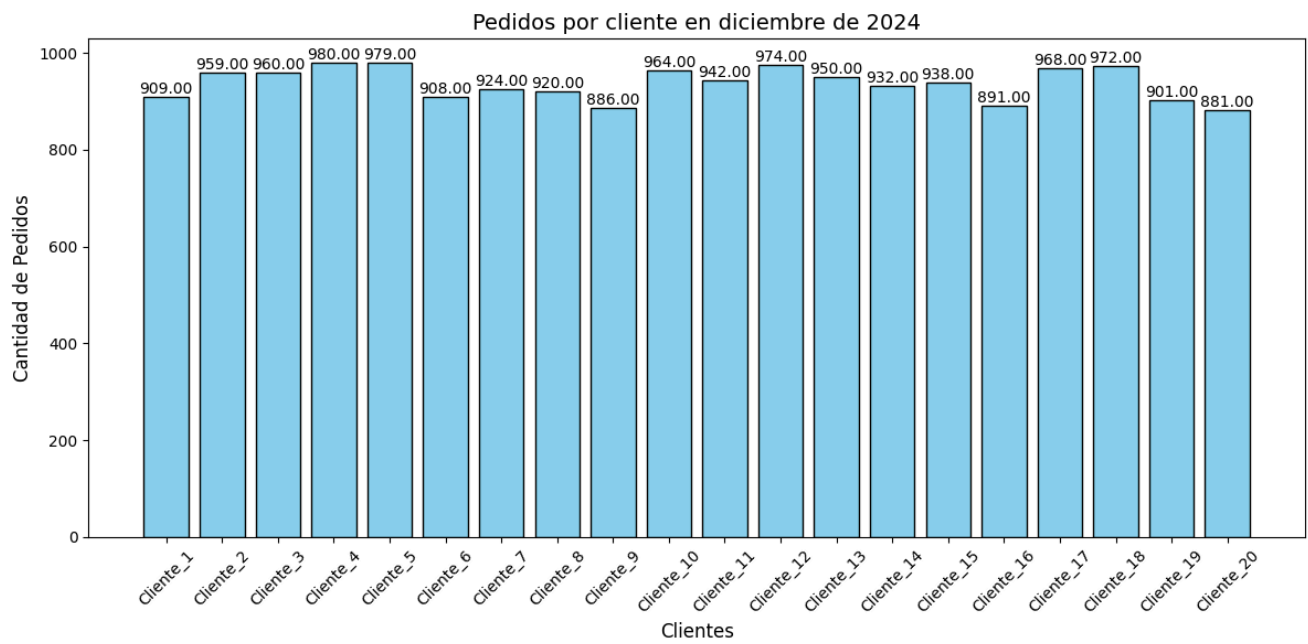
- **Objetivos:**
  - Minimizar el coste total.
  - Maximizar el número de pedidos entregados.
- **Restricciones:**
  - Capacidad de carga de los vehículos.
  - Autonomía limitada de los vehículos.
  - Todas las rutas empiezan y acaban en el almacén.

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

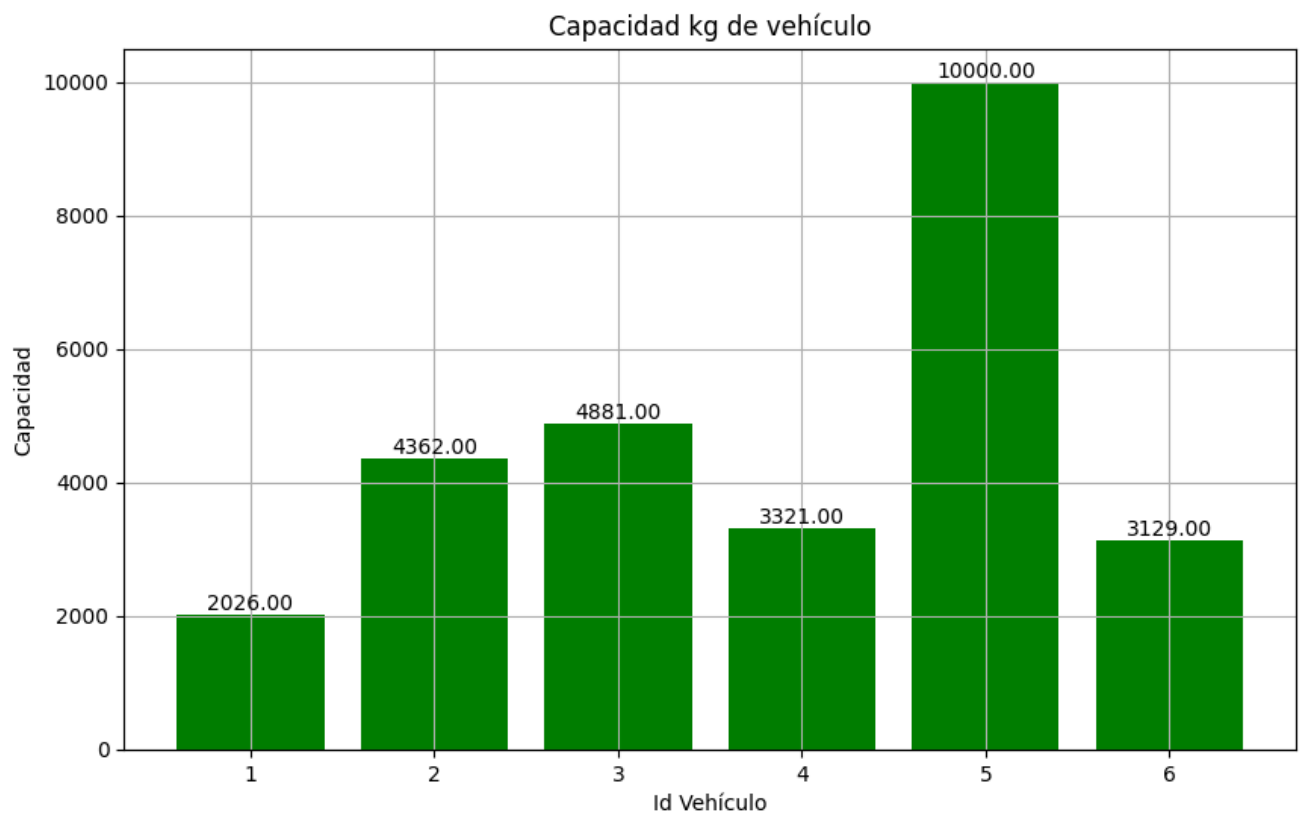


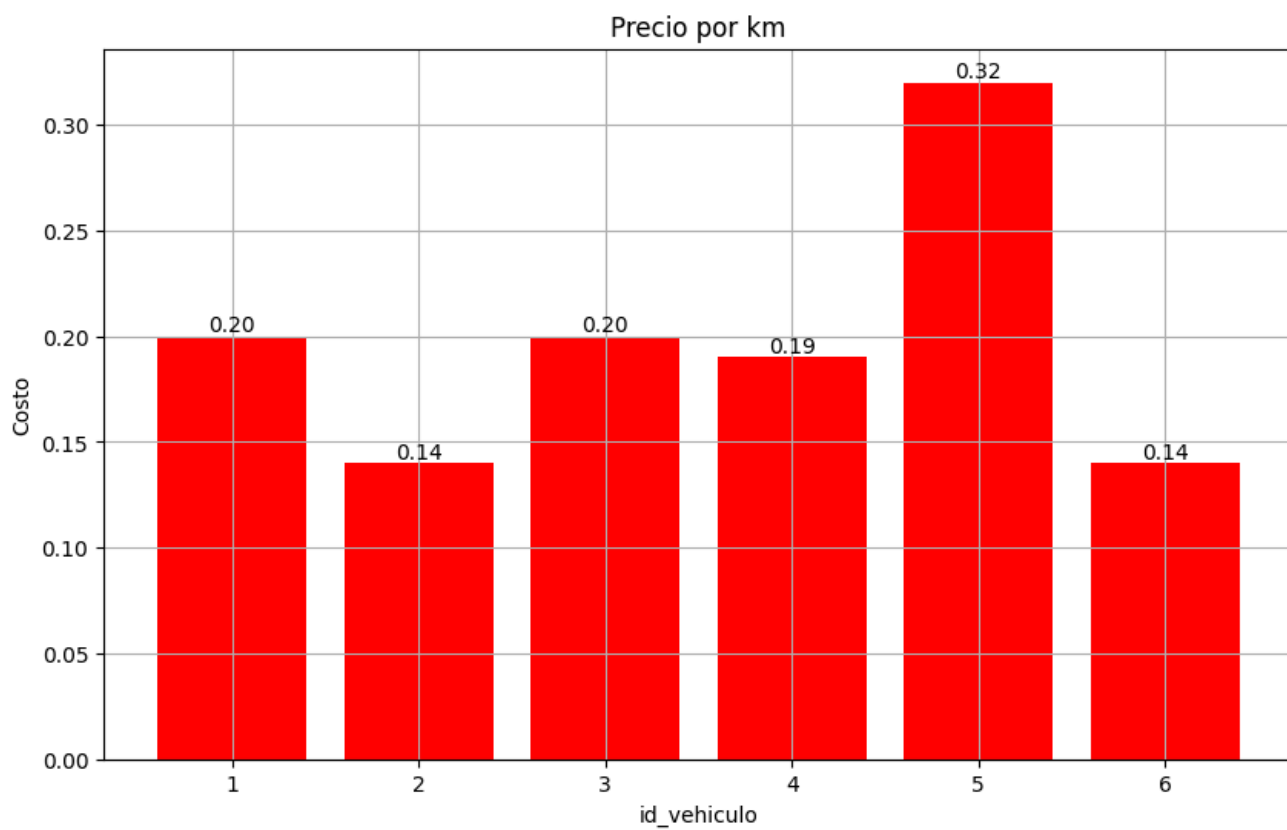
Contamos con 20 clientes, cada uno requiere de una demanda diferente dependiendo del mes.





Además, contamos con 6 camiones de diferentes características que deben ser respetadas, como son tanto la autonomía como el peso de carga permitido.





## 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  | <b>Object</b>  | Variable que da nombre a los 20 clientes y el almacén.     |
| Latitud  | <b>Float64</b> | Variable que marca la Latitud en un mapa de cada cliente.  |
| Longitud | <b>Float64</b> | 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 | <b>Float64</b> | 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 | <b>Float64</b> | 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      | <b>Object</b> | Variable que da nombre a los 20 clientes.   |
| mes_anio     | <b>Object</b> | Variable que da los datos de mes y año para cada cliente, el formato de la fecha dada es Mes-Año. |
| order_demand | <b>int64</b>  | 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      | <b>Object</b> | Variable que da nombre a los 20 clientes.   |
| mes_anio     | <b>Object</b> | Variable que da los datos de mes y año para cada cliente, el formato de la fecha dada es Mes-Año. |
| order_demand | <b>int64</b>  | 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.

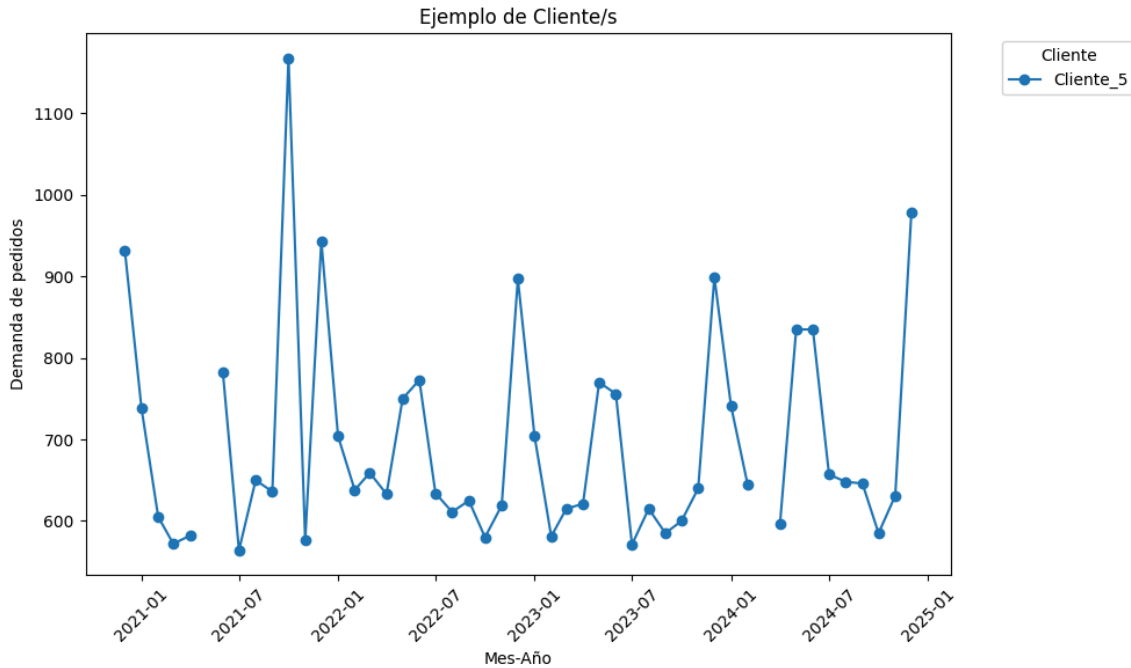
|              |                |  |
|--------------|----------------|--|
| vehiculo_id  | <b>int64</b>   | Variable con el identificador de los vehículos usados para el transporte de mercancías.    |
| capacidad_kg | <b>int64</b>   | Variable que da los datos de peso máximo de mercancías que puede llevar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

## Preparación de Datos.

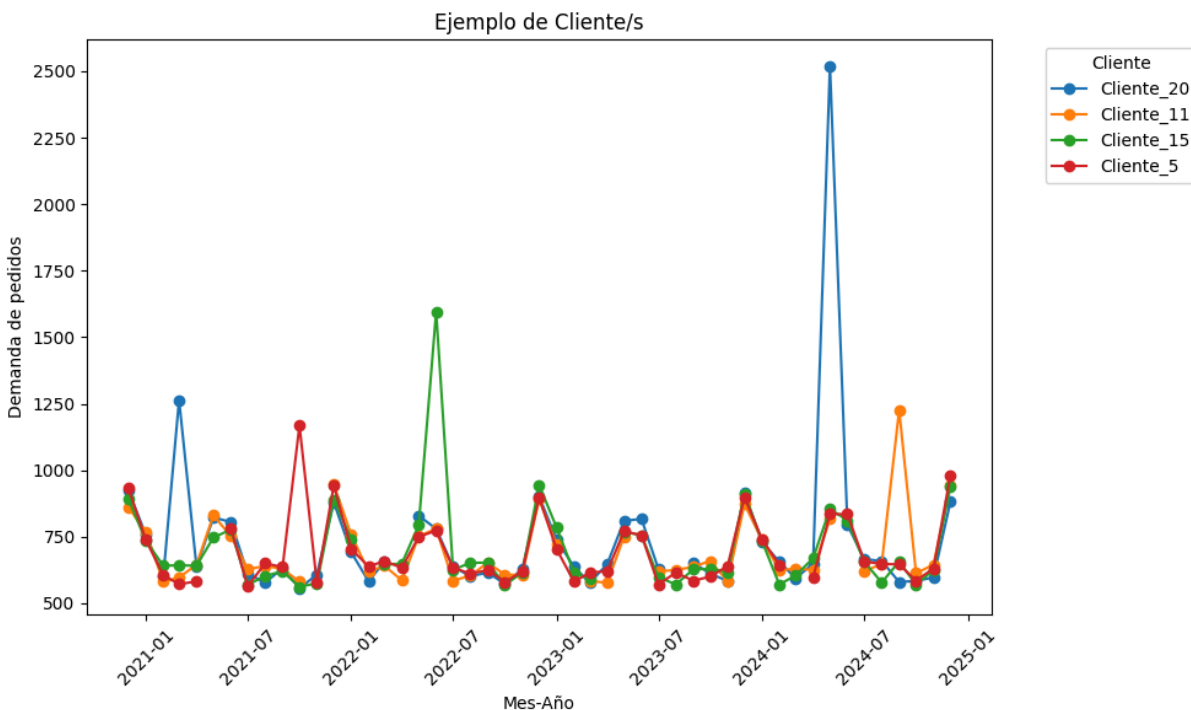
Tras la investigación de los datos, hemos encontrado que hay un DataSet (**df\_historic\_order\_demand**) que requiere de imputación/eliminación de datos.

**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.