

Optimización de rutas de transporte en la logística para e-Commerce

Diana Paola Cadena Nito A01197399

Enrique García Varela A01705747

Alexis Hernández Spínola A01275180

Paola Sofia Reyes Mancheno A00831314

María Fernanda Torres Alcubilla A01285041

Abstract—El e-Commerce ha generado que las empresas tengan un aumento de demanda que tienen que cumplir dentro de un rango determinado de tiempo para satisfacer a los clientes. Uno de los enfoques principales es optimizar la logística de las entregas, asegurando que los costos de distribución sean mínimos. En el presente trabajo se buscó solucionar esta problemática dentro de la empresa Coppel, utilizando los datos de entrega de un día en específico. En primer lugar se realizó la agrupación de distintos domicilios con base a la distancia entre sí. Posterior, se hizo un proceso de asignación de camiones, teniendo en cuenta que se cumpliera la demanda volumétrica por agrupación y que se minimizara el costo. Por último, se muestran los resultados de los costos así como posibles áreas para estudios o mejoras futuras. Proyecto disponible en el repositorio git: <https://github.com/enriquegv001/Coople-Route-Model>

Index Terms—optimización, TSP, k-means, programación lineal, e-Commerce, rutas

I. INTRODUCCIÓN

En un mercado tan competitivo como el actual, la administración de logística dentro de una empresa es imprescindible para cumplir con las expectativas de sus clientes y al mismo tiempo minimizar sus costos. Esta necesidad se vio aumentada después de la pandemia, donde la cultura de compras en línea se volvió cada vez más común. Este fenómeno ha obligado a que las empresas que cuenten con compras en línea, presten especial atención a su logística de entrega. [1]

El problema se crea para empresas grandes, donde tienen tantas y hacen que las estrategias de entrega ya no sean suficiente para minimizar sus costos. Por lo mismo se plantea [2] como problema de optimización de rutas, al estar mostrando la necesidad de encontrar la mejor ruta para satisfacer las demandas que los clientes le están haciendo a la compañía.

Dentro de esta problemática, es importante recalcar que esta puede ser observada, y por ende, optimizada desde diferentes enfoques. Por ejemplo, se podría trabajar en el ámbito de minimizar los costos de envío (tomando en cuenta la menor distancia recorrida) u otra sería minimizando el tiempo de entrega. Esta última, es muy importante dentro de las empresas ya que significa el aumento de clientes debido a un corto tiempo de espera. No obstante, este último enfoque requiere más variables a considerar para la toma de

decisiones; volviéndolo un problema más complejo e incluso estocástico.

Partiendo de esta problemática que cientos de empresas sufren en la actualidad, se puede buscar una posible solución en la rama de las matemáticas de la optimización. Donde cada día en cada una de las zonas en las que se encuentre una empresa, habrá una solución única para lograr su objetivo mediante programación lineal. En este caso particular, se trabajará con la empresa Coppel la cual se caracteriza por ser una empresa de e-Commerce a lo largo de México.

Coppel tuvo su origen en 1941 y comenzó con la venta de regalos. Poco después, se empezaron a vender muebles a crédito debido a las repercusiones económicas de la Segunda Guerra Mundial. Con el paso de los años, la empresa expandió las líneas de producto y métodos de pago disponibles, llegando a la diversidad de productos y pagos que tienen hoy en día. [3]. En el estado de Nuevo León, se cuentan con 53 tiendas físicas, de las cuales 5 se encuentran en el municipio de San Nicolás de los Garza. [4] Además, la empresa brinda el servicio de compra de productos de manera digital. Es decir, por medio de la página web de Coppel, se puede hacer la orden de compra y la transacción correspondiente, aumentando la demanda de entregas. A este servicio digital se le puede denominar como *e-Commerce*.

II. MARCO TEÓRICO

A. e-Commerce

El *e-Commerce* se puede definir como transacciones comerciales digitalmente autorizadas entre organizaciones y clientes. Realmente existen distintas definiciones del *e-Commerce* ya que algunas toman en consideración unos aspectos y otros. Otra definición es proporcionada por el autor Andrews (2002): "...el comercio electrónico se refiere a las actividades de negocios que involucran clientes, productores, proveedores de servicios e intermediarios usando redes de computadoras tales como la Internet". [5] En general, se entiende que el comercio electrónico involucra la venta de productos por medio de herramientas y procesos digitales además de que puede ser considerado como un campo multidisciplinario. El *e-Commerce* incluye áreas técnicas (telecomunicaciones, seguridad), de negocios (mercadotecnia, facturación, pagos) y aspectos legales (protección de la información personal, impuestos, entre otros). [5]

A comparación de los métodos tradicionales de venta, el comercio electrónico ofrece distintas ventajas. Entre estas, destacan la disponibilidad, el alcance y la comodidad. [6] Por disponibilidad, se hace referencia a que, a menos que se esté haciendo mantenimiento a la página, esta se encuentra disponible para su navegación las 24 horas del día. Lx clientx puede realizar un pedido a una hora donde el local físico no se encuentre abierto y se va procesando la orden de acuerdo a las políticas de cada empresa.

El alcance se refiere al acceso tanto nacional como internacional que se puede tener a la página web y a los productos que se ofrecen. No obstante, cabe mencionar que, por lo general, se pide que se reduzca esta área geográfica a la misma de lx clientx para hacer el proceso más eficiente. Además, dependiendo de la empresa o tienda, es que se pueden ofrecer distintos servicios de envío: locales o nacionales. Para una empresa cuya sede se encuentra en un estado de la República A y tiene la posibilidad de ofrecer envíos nacionales, puede recibir pedidos del estado B, aumentando sus ingresos.

Por último, la comodidad puede implicar distintos aspectos del proceso de venta por parte de lx clientx. Puede que la persona tenga un horario laboral que se le dificulte ir físicamente a una de las tiendas y requiere comprar algo. En el contexto de la pandemia ocasionada por el Covid-19, muchas tiendas tenían aforos limitados, estaban cerradas o las personas sentían inseguridad al salir de sus casas. El ofrecer un servicio de venta en línea mitiga estas mismas situaciones planteadas, aunque pueden ser más factores por los que una persona prefiera este servicio. No obstante, cabe mencionar que este servicio no puede ser accesado por todxs debido a la brecha tecnológica propia de cada país, por lo que también podría considerarse como un cierto lujo o privilegio.

En México, el *e-Commerce* comenzó a ser implementado por Walmart en 2012, dando a conocer la relevancia de este medio de comercio tanto de manera nacional como internacional. [7] En el 2016, México se posicionó en el primer puesto, a nivel América Latina, en ventas minoristas por medio del comercio electrónico. En el 2018, se reportó un aumento en el porcentaje poblacional que realizaba búsquedas relacionadas con la compra de productos en línea a comparación del 2016. El 56% de lxs mexicanxs realizaba estas búsquedas, el 58% visitaba propiamente la tienda en línea mientras que el 43% efectuaba la compra. Sin embargo, ese mismo año se reportó que únicamente el 39% de las personas mayores a 15 años contaban con una cuenta bancaria y el 6% hace pagos en línea. [7] Estas cifras ponen en perspectiva la realidad y sugieren una mala calidad del servicio bancario y financiero, lo que influye en la compra de productos en línea.

Asimismo, la Asociación mexicana de venta online [8] describe los resultados obtenidos sobre las empresas que piensan en general lxs mexicanos, durante el 2021, al comprar online. Estando de esta manera Coppel entre los 10 primeros,

además de que este tuvo un incremento de ventas del 22% respecto al 2022. (Figura 1)

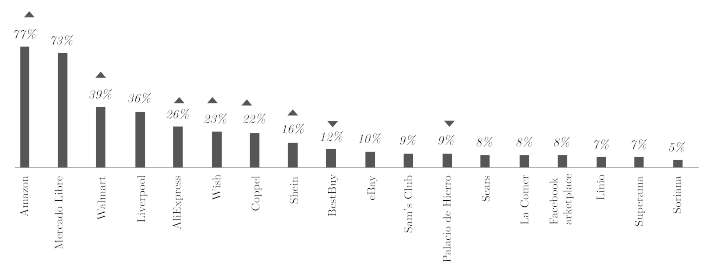


Fig. 1. Empresas e-commerce.

Además podemos observar en las mismas estadísticas que [8] del porcentaje total de las malas experiencias en la compra en línea se estima que el 70% de la gente piensa que se tardan mucho en el tiempo de entrega, este porcentaje incrementó un 50% respecto a lo que se tenía en el 2020, así como se sigue teniendo en consideración la opinión de que es muy alto el costo de envío.

Las cifras mencionadas con anterioridad sirven para poner al *e-Commerce* en contexto nacional, más la accesibilidad a estos servicios no es el enfoque del reporte. El implementar un servicio de ventas puede representar un gran reto para la empresa si no se cuenta con los medios necesarios para su manejo correcto. Se han implementado términos como las innovaciones negativas, las cuales tienen el propósito de potencializar operaciones a largo plazo pero poniendo en riesgo el desempeño. [5] Esto sobre todo en la parte de logística y entrega de los pedidos, pues estos se deben de llevar a cabo en tiempo y forma con los recursos que tengan disponibles. Posteriormente, se hablará sobre el problema del agente viajero, TSP por sus siglas en inglés, el cual aborda situaciones de esta índole.

B. Problema del agente viajero

El problema del agente viajero es un tipo de problema que plantea un número delimitado de "ciudades" a las que una persona o entidad tiene que llegar, dado un costo de viaje predeterminado para cada ciudad. El objetivo del problema es que el viajero sea capaz de recorrer todas las "ciudades" o puntos destino y regresar al lugar de origen de la forma más barata posible. (cite) Dicho problema tiene una complejidad de $(n - 1)!$, donde n es el número de ciudades a visitar. Haciendo que un problema de $n \geq 50$ tenga una complejidad computacional que rebasa las capacidades tecnológicas actuales cuando se trata de resolver de forma directa. Es decir, revisando cada una de las posibles formas de solucionarlo es prácticamente imposible. [9]

Debido a la complejidad de resolver este problema, existen múltiples planteamientos que ayudan a aproximar una solución factible cuando el número de puntos que se tienen que visitar incrementa. Cuando se lleva este problema a un contexto tradicional, donde se parte de la hipótesis que una empresa cuenta con un número limitado de vehículos y

tiene como objetivo minimizar la distancia de entrega de sus productos a sus clientes, Gu w. et al proponen un método de solución a través del uso de un Sistema de Información Geográfica (GIS). El cual puede ser adaptado al problema teniendo en cuenta que si se tienen la geolocalización de los clientes, junto con la del almacén o punto de inicio de la ruta, se puede utilizar un método de planeación de rutas para entrega basado en GIS. [10]

Este método consiste en modelar el camino a recorrer como una grafo indirecto, donde la ruta trazada está compuesta por el punto de partida, el punto final del recorrido, la localización de los clientes y los caminos relacionados. A lo cual se le suma la capacidad de cada camino, generando un escenario más realista como el mostrado en la Figura 2.

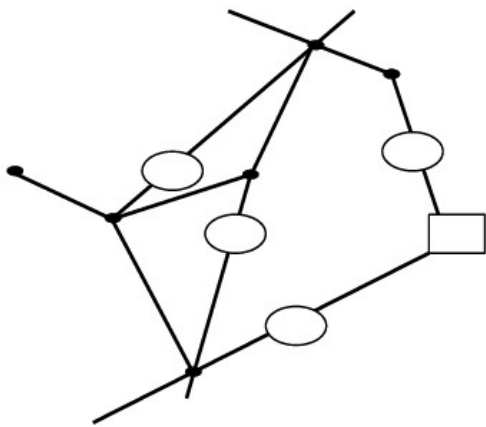


Fig. 2. Ruta de flujo para TSP usando GIS.

Para construir el algoritmo de este tipo primero es necesario que se recopile la información necesaria para que el algoritmo funcione correctamente. El primer paso es construir la base de datos donde se almacene toda la información relacionada al camino, como tráfico, nombre e identificador. Después se requiere recopilar información de los cruces de los caminos y la de los clientes. A través de un sistema de posicionamiento global (GPS) es vital recolectar la latitud y longitud de los cruces de camino y de los clientes objetivo. Para los segundos se requiere almacenar su identificador, localización (latitud y longitud), volumen histórico de la cantidad distribuida en un periodo de 3 meses y su nombre. Una vez que se tiene todo esto, la información debe ser mapeada con el objetivo de empezar a construir la red. Dicha red debe ser optimizada y en el procedimiento de [10], se propone que la red completa del problema sea dividida en un número k de regiones para así aplicar un modelo heurístico que resuelva y optimice la ruta de transporte dentro de las mismas por separado. Generando una solución factible al unir las regiones y así requerir de menor complejidad computacional.

Dentro del paper [10], los autores usaron un algoritmo híbrido, que combinará la fuerte capacidad de búsqueda global de un algoritmo genético y la solidez de la búsqueda

local del recocido simulado. Para la construcción del algoritmo genético se crean y se definen 4 partes. La primera parte es el esquema de codificación, dentro del cual se define el camino y el orden de los nodos a recorrer. Por ejemplo, si se tienen 9 nodos a visitar y se parte de un centro de distribución etiquetado como 1 el camino a recorrer sería:

$$1ABCDEFGHI1 \quad (1)$$

Esto nos indica que la ruta iniciaría en el punto inicial 1, pasando por los puntos A-H en orden y regresando al punto 1.

Después, se necesita crear y definir la función de evaluación, la cual es prácticamente la transformación de la función objetivo. La cual en el caso de ser aplicada a la distribución y logística, es preferible que se defina con un valor pequeño, cuando se define como la suma de la longitud de todos los caminos. Está dada por:

$$fitness = K * \frac{1}{\sum_{i=1}^{n-1} c_{ij}} \quad (2)$$

En tercer lugar se define al operador de cruce (OX), el cual indica que para generar a la cadena genética hija, se selecciona una parte de la cadena padre A, a la vez que se preservan el orden de los nodos de la cadena padre B. Se representa el operador como:

$$p_c(i, j) = \frac{maxf - max(f_i, f_j)}{maxf} \quad (3)$$

donde i e j son dos individuos con valores y funciones de ajuste diferentes.

Por último se define al operador de mutación, el cual se encarga de asegurar y mantener la diversidad en la población. Tratando de generar que su valor sea pequeño con el objetivo de que se pierda la menor cantidad de información genética posible. Dada por:

$$p_m(i) = k_1 \frac{maxf - maxf_i}{maxf} \quad (4)$$

Una vez que se termina la creación del algoritmo genético, se procede a definir y crear el algoritmo de recocido simulado, el cual en base a lo ya dicho se centrará en la búsqueda local. En primera instancia se tiene que tener en cuenta la expresión del estado, donde un estado es una solución al problema y su función de energía representa la función objetivo del problema.

Otro elemento importante es el movimiento, donde el estado actual se representa como i y su vecino como j . Ambos estados tienen sus respectivas funciones de evaluación dadas por f_i y f_j . El incremento en la función está dada por $\Delta f = f_i - f_j$.

Luego, se debe tener en cuenta el equilibrio térmico, el cual es el proceso de alcanzar un equilibrio dada una temperatura específica mediante un proceso de ciclado

interno. El cual debe ser lo suficientemente grande como para alcanzar un punto de equilibrio teórico, puesto que en la práctica esto no es factible.

Por último, se tiene la función de enfriado que ayuda a disminuir la temperatura y se define como:

$$T_{k+1} = T_k * r \quad (5)$$

donde r es un parámetro que [10] variaron entre 0.95 y 0.99. Entre más pequeño es el valor de este parámetro, más lento disminuye la temperatura.

Para aplicar el algoritmo híbrido dentro del paper citado, se seleccionó una temperatura T_0 , junto con su respectivo coeficiente de temperatura r y en base a esto el tamaño de la población p y se definieron las condiciones para terminar el proceso de búsqueda. Dadas estas condiciones se evalúa cada cromosoma hallado por el algoritmo genético, se clasifica en base a la función de evaluación y se retiene el mejor de estos. Al cual se le aplican los operadores de mutación y cruce. Después se generan nuevos grupos en base a la estrategia de replicación. A estos nuevos grupos se les aplica el algoritmo de recocido simulado y se producen nuevos grupos. Estos también son evaluados y de alcanzar una mejor calificación, reemplazan al mejor cromosoma encontrado previamente, siempre y cuando cumpla con las condiciones de fin de búsqueda. Reteniendo este valor como resultado del algoritmo híbrido. [10]

III. DESCRIPCIÓN DE LA PROBLEMÁTICA

Cabe destacar que, en conjunto con la creación del *e-Commerce* y los últimos años de pandemia donde la gente se veía orientada a implementar dicho tipo de compra, la distribución de los productos a cada cliente crea una tendencia positiva mucho mayor de lo que ya se tenía. De tal modo que el conjunto de todas las estadísticas mencionadas a través del marco teórico sea una justificación para entender la necesidad de la optimización de logística de la empresa.

Para el caso de Coppel, se busca como objetivo una solución factible en la que cada cliente pueda obtener su entrega de la forma más rápida, así como para que el transporte pueda ser de mínimo impacto para los precios finales. Ahora, este tipo de situaciones y la metodología de resolución se ha demostrado que se logra a través del Problema del Agente Viajero (TSP). El cual se plantea debido a la necesidad de las empresas y servicios de logística para poder determinar una ruta en la que se logren minimizar, ya sea el costo, la distancia o el tiempo entre cada lugar a visitar y en el cual se requiere visitar todos los puntos.

A. Puntos de entrega

En nuestro problema seleccionamos la creación de la ruta óptima para las entregas del 21 de enero del 2022 en el municipio de San Nicolás de los Garza, Nuevo León. Después de obtener las coordenadas de las direcciones mediante una

API utilizando Geocode de Google Maps [11] se pudo hacer una limpieza de datos eliminando los puntos de entrega que correspondían a otros municipios, contando con un total de 90 puntos de entrega y 1 Centro de Distribución (CEDIS) ubicado en el municipio de Monterrey. En la Figura 3 se pueden ver los diferentes domicilios de entrega y la ubicación del CEDIS y en el anexo se encuentra el link donde se encuentran marcados en la herramienta de MyMaps.

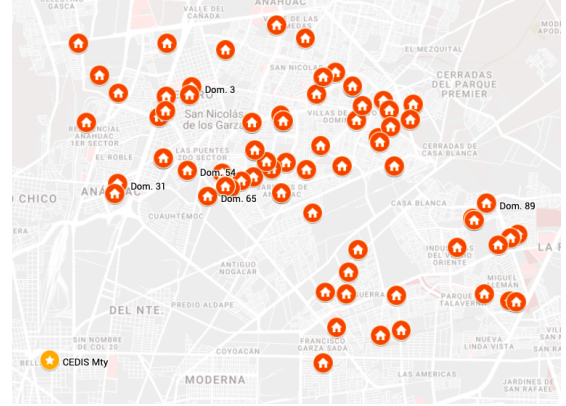


Fig. 3. Puntos de entrega y CEDIS.

B. Variables y parámetros

En la definición de variables y parámetros actual, se plantea un problema de programación lineal como habitualmente se resuelven estos problemas de rutas. No obstante, cabe recalcar que se muestra solamente por fines de definición del problema en un lenguaje matemático, más no es el método escogido dentro de este trabajo.

Sea nuestra variable de decisión:

$$x_{ij} \begin{cases} 1: \text{se asigna el } A(i, j) \\ 0: \text{otro caso} \end{cases} \quad (6)$$

siendo el i –ésimo valor el lugar de origen del Centro de Distribución de Monterrey y j –ésimo valor el lugar de destino individual de los diferentes domicilios.

Además, se tendría como parámetro el costo del origen al destino:

$$c_{ij}; i = \{1, 2, 3...n\}; j = \{1, 2, 3...n\} \quad (7)$$

Los parámetros a considerar serían:

- Del paquete: volumen.
- De la ruta: distancia.
- De lxs empleadxs: salario por hora, jornada laboral
- Del camión: kilómetros recorridos por litro de gasolina, capacidad máxima de volumen.
- Otros: precio de la gasolina.

C. Función objetivo y restricciones

De esta manera podemos establecer como función objetivo (8) el costo total de la ruta. Siendo el objetivo que se minimice Z . Donde (c) es el coste de viajar del origen (i) al destino (j) .

En la parte del coste se incluiría el salario por horas trabajadas de lxs repartidores y la gasolina consumida de acuerdo a la distancia recorrida y la medida del gasto promedio del camión.

$$\min z = \sum_{i=1}^m \sum_{j \neq i, j=1}^m c_{ij} x_{ij} \quad (8)$$

Dando así que este se vea restringido porque no se puede repetir una visita a lxs clientes, de modo que para cada conexión sólo puede haber una entrega, así como que todxs lxs clientes deben de ser visitadxs. Esto se observa en las ecuaciones 10 y 11

$$x_{ij} \in 0, 1 \quad (9)$$

$$\sum_{i=1, j \neq i}^m x_{ij} = 1, \forall j \quad (10)$$

$$\sum_{j=1, i \neq j}^m x_{ij} = 1, \forall i \quad (11)$$

Finalmente una restricción en la que no se generen rutas más pequeñas, sino que se mantenga como una general. De modo que si vemos a las posibles rutas como una matriz cuadrada de $m \times n$, siendo m los orígenes y n los destinos, que contienen los costos entre cada par de puntos, la ruta no debería contener alguna sub-ruta donde no regrese al CEDIS, esto se expresa en la siguiente ecuación:

$$\sum_{i \in Q} \sum_{j \neq i, j \in Q} x_{ij} \leq |Q| - 1 \quad \forall Q \not\subseteq 1, \dots, n, |Q| \geq 2 \quad (12)$$

Dentro de las restricciones, también se aplicaría un límite por las horas trabajadas para que estas no exceda la jornada laboral, que la suma de todos los volúmenes de los paquetes no exceda la capacidad del camión y en cuanto al cuidado del ambiente, se podría implementar una restricción de gasolina consumida de acuerdo a la distancia para que las emisiones estén dentro de la norma.

D. Diagrama

En el siguiente diagrama se muestra las relaciones entre las variables con los costos de la ruta, los cuales incluyen varios parámetros como se mencionó anteriormente:

donde m es el número de lugares marcados en la ruta (los domicilios y el CEDIS) como se pueden apreciar en la Figura 3, i es el origen y j el destino correspondiente, donde estas últimas dos van de 0 a m .

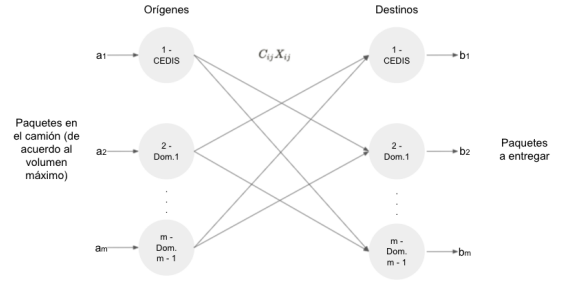


Fig. 4. Diagrama de relación de variables y acciones.

IV. MODELO DE RESOLUCIÓN

Para la resolución de la problemática, se decidió no hacer uso solamente de la programación lineal, sino también de Python. Esto con el fin de visualizar de mejor manera el cumplimiento de todos los puntos con los que se debe cumplir; como por ejemplo el uso de los diferentes camiones que tiene Coppel, el regreso de los camiones a CETIS cuando ya no tenga paquetes por entregar, las diferentes capacidades de cada camión, la ruta más corta en cada caso, etc. A continuación, se describe paso a paso cómo se definió la propuesta de optimización.

A. Definición

Antes de proceder a la especificación de los parámetros y métodos que se utilizaron para su definición, cabe aclarar lo siguiente. Se busca tomar en consideración distintos camiones asignados a distintas rutas, esto con base a los volúmenes por agrupación de direcciones. Es decir, por agrupación se busca asignar un camión que tenga la suficiente capacidad volumétrica y que a su vez optimice el tiempo de la ruta en términos de costos.

1) *De conjuntos:* Debido a la gran cantidad de puntos de entrega y de la dificultad que esto involucra al modelarlo junto con la restricción del volumen máximo a cargar en cada viaje, se decidió utilizar el método de *k-means* como primer paso. El método fue de ayuda para agrupar puntos de entrega cercanos entre sí y así simplificar la aplicación de programación lineal al tener que optimizar cada una de las rutas pero con menor cantidad de restricciones. A continuación se encuentra un ejemplo de agrupación de los puntos de entrega en 4 clusters.

Para esto, se decidió hacer la división de los puntos de entrega entre 3 y 6 *clusters* diferentes para posteriormente realizar la sumatoria del volumen total, la demanda, para cada cluster de entrega en cada caso. Estos cuatro escenarios diferentes serán procesados por separado a través de la API y la programación lineal para asignación de camiones, para posteriormente poder comparar sus resultados entre sí y escoger al mejor escenario para minimizar los costos de la empresa.

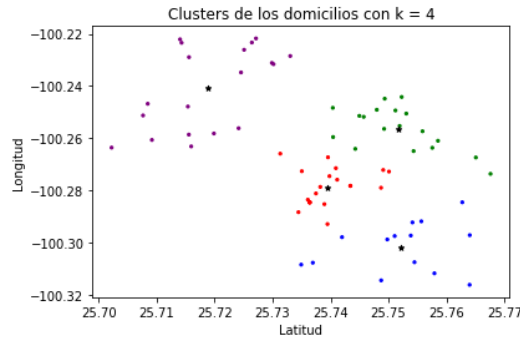


Fig. 5. Puntos de entrega segmentados en 4 zonas

2) *API*: Para resolver un problema de ruteo de vehículos y ser optimizado, se puede definir al mismo como un problema del agente viajero. Dicho problema necesita como base una matriz de costos entre los orígenes y destinos posibles. Sin embargo; esta tarea puede ser automatizada si se ocupa la interfaz de programación de aplicaciones (API) llamada Openrouteservice. La cual está conectada directamente con Openstreetmap y le permite recabar información geográfica sobre los puntos de entrega y origen. Esta información después es procesada con VROOM (Vehicle Routing Open-source Optimization Machine). El cual es un motor de optimización basado en el lenguaje de programación de C++ que toma como restricciones los puntos de inicio y fin de los vehículos, descansos para conductores, horarios de trabajo, capacidad en base a distintas métricas, etc. VROOM puede resolver este tipo de problemas de manera rápida y con un bajo costo computacional.

3) *De parámetros*: Para establecer la distancia total a recorrer así como el volumen se realizó por medio del programa; los clusters definidos junto con sus distancias y volúmenes totales se encuentran en la sección V. Para definir específicamente los volúmenes por cluster, se utilizó la base de datos proporcionada por la empresa y un diccionario de las claves de los productos. Una vez se hayan establecido los volúmenes correspondientes, se hizo una sumatoria de estos para definir la demanda.

Para establecer el parámetro de la distancia en términos de costo, se debe tomar en consideración el costo de la gasolina a utilizarse así como la distancia por agrupación. Para la distancia, se tomó el costo de la gasolina en \$23.35 MXN por litro; esto de acuerdo a la fecha de consulta del 12 de junio de 2022 y al promedio municipal. [12] No obstante, para plantear la matriz de costos por camión, se debe tomar en consideración el rendimiento de cada camión y su capacidad máxima como la oferta. A continuación, se muestran las características propias de cada camión; la especificación del camión perteneciente a la unidad n se encuentra en la sección de anexos. Además, estos datos se obtuvieron de las páginas oficiales de las marcas de los vehículos utilizados. [13] [14] [15] [16] [17] [18]

Unidad	Capacidad (m^3)	Rendimiento (km/l)
01	04.20	15.20
02	07.10	14.70
03	03.90	20.00
04	08.45	12.00
05	09.31	11.60
06	15.00	13.30
07	04.39	11.68
08	10.96	10.90
09	29.06	06.40
10	06.10	16.70
11	04.39	12.04
12	07.12	10.00
13	24.28	08.20
14	08.45	09.00
15	10.82	10.20
16	10.82	10.20
17	24.15	08.20
18	11.03	10.40
19	06.70	13.20
20	07.14	15.00
21	11.50	11.90
22	07.10	14.70

TABLE I
UNIDADES DISPONIBLES

Otro parámetro a establecerse es el tiempo en términos de costo, esto tomando en consideración el salario por hora de los choferes. Este costo se estableció como \$34.46 por hora, de acuerdo al salario medio de un chofer y a la fecha de consulta del 12 de junio de 2022. [19] Además, cabe mencionar que la hora de entregas de paquetes de Coppel, entre semana, es desde las 11:00 hasta las 17:00 horas; este dato es importante al momento de plantear las restricciones del modelo.

En la tabla II, se muestran los costos a manera de resumen.

Parámetro	Costo
Gasolina	\$23.35 por litro
Salario	\$34.46 por hora

TABLE II
COSTOS

4) *De variables*: De acuerdo a lo planteado en el marco teórico, el problema del agente viajero toma como variables la cantidad de paquetes a transportar de un origen i a un destino j . No obstante, en este caso, la variable de decisión es referente a la asignación de camiones. Como se menciona con anterioridad, la programación del algoritmo de k -means minimiza la distancia entre los puntos y los agrupa con base a eso. La cantidad de paquetes a transportar se toma en términos de volumen ya que la capacidad de los camiones está en estos mismos términos volumétricos; y el uso de la API crea la ruta óptima para cada cluster.

Por esta razón solamente se necesitaría aplicar programación lineal para la asignación de camiones para cada cluster en los

diferentes escenarios ya que ahí ya se estarían optimizando los costos de transporte. Esto pasa a ser un problema de asignación, lo cual se realiza por medio de variables binarias. No obstante, es importante señalar que aunque el resultado óptimo también dependa de la cantidad de clusters a utilizar, este número no se decidirá a través de programación lineal, sino que se escogerá la menor z entre los cinco escenarios existentes.

Por otro lado, se debe tener en consideración que Coppel utiliza externos al momento de transportar artículos de volumen relativamente pequeño y sus camiones propios para transportar artículos de volúmenes altos. No obstante, para el planteamiento del modelo, se está trabajando bajo el supuesto de que todos los paquetes serán entregados por camiones de Coppel y no externos.

B. Planteamiento del código

Para comenzar, el código presenta la sección de preparación de la base de datos. A través de ella, se sigue el procedimiento de importación de datos, manejo de los datasets para crear una relación entre ellos, de modo que tengan un mismo índice y poder relacionar el volumen correspondiente a los pedidos. El siguiente paso consistió en la limpieza de datos, enfocándonos en el volumen faltante y datos repetidos. Además, fue necesario la creación de un nuevo atributo para hacer la suma de los pedidos de un mismo domicilio.

Habiendo definido lo anterior, se siguió una modulación de minado de datos, para clasificarlos en los clusters mencionados anteriormente. Obteniendo con esto seis bases de datos finales, en la que se tuvieron como atributos de latitud, longitud, volumen, cluster, fecha de partida, fecha de posible entrega y cantidad de paquetes. Estas bases de datos variarían sobre la cantidad de clusters en los que se realizó el clasificado, siendo $k = 3, 4, 5, 6$.

Más adelante se realizaron cuatro funciones, con base en la API de Openrouteservice. La primer función consiste en mostrar todos los registros que se tienen en un mapa interactivo, con el que se muestra la cantidad de pedidos por domicilio y el punto de partida. La segunda función permite la creación del primer modelo de optimización que se estará ocupando, así como la visualización del resultado de la ruta factible creada. En la tercer función, permite obtener el recorrido total de la ruta con los tiempos y el último lo mismo pero para cada registro.

Finalmente, se manda a llamar a las funciones para cada uno de las bases de datos creadas y se muestran sus resultados individuales, así como se genera un diccionario para guardar la información de cada uno, el cual se ocupa para dar una tabla comparativa final entre los restados totales.

Con estos resultados se realiza nuevamente un modelado de optimización pero para un problema de asignación. En que se busque minimizar el costo de entrega, respecto a si se le

asigna una de las rutas divididas en los clusters para cada uno de los camiones o no. Dicho modelo se describe en la siguiente sección.

C. Planteamiento del modelo de asignación de camiones

Sean las siguientes variables de decisión:

x_{ij} : el i - ésimo camión a utilizar en el j - ésimo cluster, donde $i = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22\}$ y $j = \{1, 2, 3, 4, 5, 6\}$, $j \in k$ - ésimo escenario (13)

Minimizar

$$Z = \sum_i^{22} \sum_j^k x_{ij} c_{ij} \quad (14)$$

s.a.

$$d_j \geq \sum_{i=1}^{22} x_{ij} o_i, \forall j \quad (15)$$

donde d_j es el volumen total del j cluster y o_i es la capacidad volumétrica del i camión.

$$\sum_i^{22} x_{ij} = 1, \forall j \quad (16)$$

$$\sum_j^k x_{ij} \leq 1, \forall i \quad (17)$$

V. RESULTADOS

Para definir la cantidad de clusters a establecerse, se realizó la optimización de los hiperparámetros del algoritmo k -means. Se planteó la posibilidad de 3, 4, 5 y 6 clusters; de todos estos, se encontró que el modelo trabaja mejor con 3 clusters. Esto debido a que el tiempo total de los recorridos y la distancia es menor que en el resto de escenarios.

	3	4	5	6
distancia	110989	125880	143090	160912
duracion	12426	13500	14887	16233
dur total	0 days 15:37:06	0 days 15:55:00	0 days 16:18:07	0 days 17:00:33

Fig. 6. Resultados por clasificación de rutas

Para cada cluster establecido, se obtuvo la distancia total en metros, el tiempo total en segundos y el volumen total en metros cúbicos. En la siguiente tabla V, se muestran los datos en las unidades de kilómetros, horas y metros cúbicos respectivamente.

Agrupación	Distancia (km)	Tiempo (hrs)	Volumen (m^3)
1	37.95	5:50	11.18
2	32.98	4:06	14.59
3	40.06	5:40	10.23

TABLE III
DATOS POR AGRUPACIÓN

Para plantear la matriz de costos, se toma en consideración la

distancia total a recorrer por *cluster*, obtenida por el código del algoritmo, y el rendimiento de cada camión disponible multiplicado por el costo de la gasolina. Dicha matriz se muestra en la siguiente figura 7.

```

table
c(i,j)
1 1 2 3
1 58.30279934 50.65874671 61.53800329
2 60.28588776 52.38183333 63.63113265
3 44.3101275 38.5006475 46.7688825
4 73.8502125 64.16774583 77.9481375
5 76.39677155 66.38042672 80.63600431
6 66.63177068 57.89571053 70.32914662
7 75.87350599 65.92576627 80.08370291
8 81.30298624 70.64338991 85.8144633
9 138.4691484 120.3145234 146.1527578
10 53.06602096 46.10855988 56.01063772
11 73.60486296 63.95456395 77.68917359
12 88.620255 77.001295 93.537765
13 108.0734817 93.90401829 114.0704451
14 98.46695 85.55699444 103.93085
15 86.88260294 75.49146569 91.70369118
16 86.88260294 75.49146569 91.70369118
17 108.0734817 93.90401829 114.0704451
18 85.21178365 74.03970673 89.94015865
19 67.13655682 58.33431439 70.86194318
20 59.08017 51.33419667 62.35851
21 74.47080252 64.70697059 78.60316387
22 60.28588776 52.38183333 63.63113265;

```

Fig. 7. Costo por transporte

Por medio de GAMS, se obtuvieron los resultados encontrados en la figura 12.

	LEVEL		LEVEL
1	.1	.	12.2
1	.2	.	12.3
1	.3	.	13.1
2	.1	.	13.2
2	.2	.	13.3
2	.3	.	14.1
3	.1	.	14.2
3	.2	.	14.3
3	.3	.	15.1
4	.1	.	15.2
4	.2	.	15.3
4	.3	.	16.1
5	.1	.	16.2
5	.2	.	16.3
5	.3	.	17.1
6	.1	.	17.2
6	.2	1.0000	17.3
6	.3	.	18.1
7	.1	.	18.2
7	.2	.	18.3
8	.1	.	19.1
8	.2	.	19.2
8	.3	1.0000	19.3
9	.1	.	20.1
9	.2	.	20.2
9	.3	.	20.3
10	.1	.	21.1
10	.2	.	21.2
10	.3	.	21.3
11	.1	.	22.1
11	.2	.	22.2
11	.3	.	22.3
12	.1	.	22.3

Fig. 8. Resultado final

No se logra visualizar cómodamente el resultado, pero se debe de asignar la unidad 21 al primer cluster, la unidad 6 al segundo y la unidad 8 al tercer cluster. Con esto, se logra

minimizar el costo de transporte con las rutas previamente optimizadas. El costo total, incluyendo el costo del salario por hora, queda de \$721.30 MXN para las tres rutas establecidas.

A. Tamaño de la red

Se toman los nodos como los domicilios a visitar en cada cluster y total así como el CEDIS. En cuanto a las aristas, estas acaban teniendo el mismo número que nodos ya que se toma en consideración la ida y vuelta al CEDIS y ningún nodo es visitado más de una vez.

Ruta	Nodos	Aristas
1	34	34
2	27	27
3	38	38
Total	74	74

TABLE IV
TAMAÑO DE LAS REDES

B. Otra ciudad

Para la verificación de que la solución propuesta es de utilidad para otras ubicaciones geográficas, se buscó implementar el modelo en el municipio de García en la misma fecha (21 de enero, 2022). En esta zona geográfica, se tuvieron que entregar 65 paquetes a 45 domicilios.

Las distancias, tiempo de duración y volumen de los recorridos y rutas se establecen en las siguientes figuras:

	1	2	3	4	5	6
distancia	137063	192841	193579	262805	304905	367404
duración	14039	17916	18125	22959	26286	31113
dur total	0 days 11:23:59	0 days 12:28:36	0 days 12:32:05	0 days 13:52:39	0 days 14:48:06	0 days 15:58:33

Fig. 9. Agrupaciones de García

Aunque con $k = 1$ se obtienen las menores distancias y tiempos, se considera para la solución $k = 2$ para que se cumpla dentro del horario laboral. Con este número de agrupaciones se obtienen las siguientes rutas y sus resultados correspondientes:

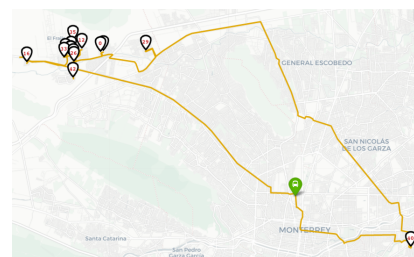


Fig. 10. Ruta 1 de García

Como se trabajó con el mismo CEDIS y fecha de entrega que con el planteamiento de San Nicolás de los Garza, se considerará a la totalidad de los camiones disponibles en Monterrey para la asignación en los 5 agrupamientos, para esto se le añadió dos columnas a la matriz de costos anterior las cuales representan los costos de los litros de gasolina gastados en cada una de las rutas de García.

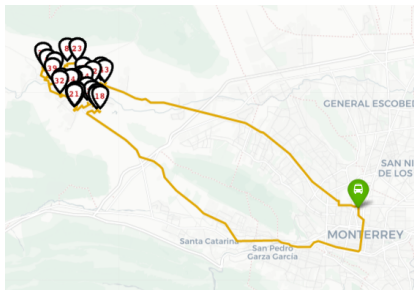


Fig. 11. Ruta 2 de García

		distance	amount	duration	tiempo_total (hrs)
ruta	vehicle				
1	0	93498	[29]	8910	0 days 05:28:30
2	0	99343	[36]	9006	0 days 07:00:06

Fig. 12. Resultados de las rutas de García

Después de realizar el modelo se asignaron los siguientes camiones a las distintas rutas, con un costo de transporte total de \$705.63 MXN, los cuales \$447 MXN corresponden a las rutas de García.

Municipio	Camión	Camión	Modelo
San Nicolás de los Garza	Ruta 1	17	716 LONG SERIE 300
	Ruta 2	2	ELF 300
	Ruta 3	8	616 LONG SERIE 300
García	Ruta 1	4	ELF 300
	Ruta 2	13	616 LONG SERIE 300

A continuación, se muestra una tabla a manera de resumen de los resultados obtenidos.

Agrupación	Distancia (km)	Tiempo (hrs)	Volumen (m^3)
1	93.50	5:28	14.72
2	99.34	7:00	18.63

TABLE V
DATOS POR AGRUPACIÓN

Como resultado final, el costo de transporte y de tiempo con las 2 rutas establecidas da un total de \$ 876.60 MXN. Este precio es mayor que el de San Nicolás de los Garza aunque sean menos puntos y paquetes a entregar, esto se debe a que como es una zona no tan poblada la demanda no es tanta como se quisiera, se podría considerar un análisis con áreas vecinas como una mejora para un futuro, siempre con el objetivo de minimizar costos totales.

VI. ANÁLISIS DE RESULTADOS

A continuación se muestra una tabla de las distintas rutas generadas por medio de *k-means* junto con sus respectivos hallazgos.

Ruta	No. de clientes	Costo MXN	Distancia (km)	Tiempo (hrs)	% tiempo muerto	% vol. utilizado	L de gasolina	Emissiones CO_2 (toneladas)
1	18	275.48	37.95	5:50	79.90	97.22	3.19	0.004086
2	27	199.18	32.98	4:06	76.91	97.27	2.47	0.004636

TABLE VI
RESULTADOS POR RUTAS

De manera general, para la ciudad de enfoque del análisis (Nuevo León) se puede observar como cada una de las rutas se completan en un promedio de cinco horas, lo que significa que cada una se completa en menos de las 8 horas laborables de cada empleado, dando tiempo extra en el caso de que existan contratiempos en el momento de la entrega. Esto es un indicador muy importante de la calidad del modelo de resolución, ya que demuestra que se puede cumplir con los objetivos de entrega utilizando muy pocos recursos físicos y humanos de Coppel.

Por otro lado, el porcentaje de volumen utilizado de cada uno de los camiones es mayor al 92% en todos los casos, lo que significa que se están aprovechando al máximo los recursos disponibles en la empresa. Así también, se puede observar un consumo de gasolina bajo en los tres casos al haber separado en rutas según la cercanía de los domicilios. Este bajo consumo no solamente tiene un beneficio para Coppel al minimizar sus costos; sino también para el planeta, al reducir las emisiones de CO_2 .

La huella de carbono es un indicador que mide la cantidad de gases de efecto invernadero derivados de distintos procesos humanos. Sin embargo, alrededor del 95% de estos lo constituyen las emisiones liberadas de dióxido de carbono CO_2 . [20] Por lo que para simplificar los cálculos de la huella de carbono, se consideraron solo las emisiones de CO_2 en toneladas. Para calcular esto, se siguió el proceso propuesto en "The Green Freight Handbook" [20]. En el mismo se indica que el primer paso es investigar la cantidad de CO_2 que se genera al transportar el volumen ocupado en cada camión. Este factor fue hallado en TEU-milla, lo cual significa la cantidad de gramos emitidos por cada unidad transportada equivalente a 20 pies (TEU) por milla recorrida. [21] Lo cual al hacer los cálculos correspondientes indica que por cada $38.55m^3$ transportados por kilómetro se generan 371.29 gramos de CO_2 . Más tarde, se hizo la conversión del volumen ocupado a TEU, lo cual fue multiplicado por 371.29 y después por los kilómetros recorridos, generando los resultados mostrados en la tabla anterior.

Para su visualización, se muestran las rutas generadas por cada cluster para el municipio de San Nicolás de los Garza.

A. Equipo y software utilizado

Por último, si se busca recrear el modelo de solución, cabe mencionar el equipo y software utilizado para la creación del mismo. Esta información se encuentra en la sección VIII.

VII. CONCLUSIONES Y RECOMENDACIONES

La optimización de rutas de entrega para una empresa es un trabajo complejo, el cual no tiene una solución única debido al cambio diario de la demanda de paquetes. Debido a esta idea, es importante plantear soluciones multidisciplinarias que permitan encontrar las mejores rutas contemplando la mayor

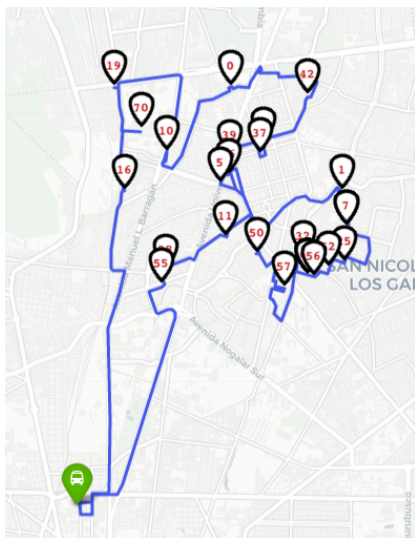


Fig. 13. Ruta del cluster 1, kmeans = 3

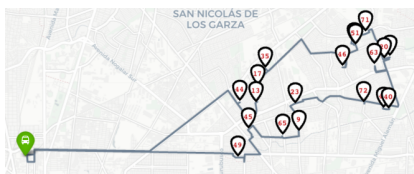


Fig. 14. Ruta del cluster 2, kmeans = 3

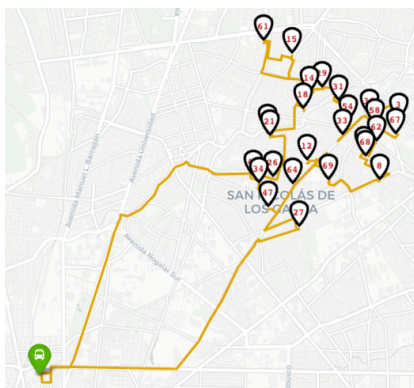


Fig. 15. Ruta del cluster 3, kmeans = 3

cantidad de condiciones y poder modelar el problema lo más parecido a la vida real. En este trabajo, se replanteó este problema de TSP como un problema de asignación posterior a haber seleccionado el número de rutas y el orden de entrega más rápido. Este replanteamiento dio la posibilidad de llegar a una respuesta óptima en términos del costo.

Por otro lado, es importante recalcar que la optimización de rutas de entrega no solamente beneficia a minimizar los costos de la empresa sino también significa el aumento de la satisfacción de los clientes. Esto debido a que, al destinar menos camiones para más entregas, se podrán entregar mayor cantidad de paquetes por día y por ende el tiempo de espera disminuirá y por ende las ventas aumentarán.

Así también, presenta una mejoría para el medio ambiente, al reducir las emisiones de dióxido de carbono, las cuales contribuyen al cambio climático y contaminación en la ciudad.

No obstante, caben mencionar las diversas limitantes del mismo reporte. Entre estas, se encuentra que, por lo pronto, únicamente se tomó en consideración una zona geográfica así como un solo día de entregas. Además, no se están utilizando todos los camiones disponibles por lo mismo de que sólo se plantea una sola zona geográfica del estado de Nuevo León. Esto podría afectar al resultado establecido como óptimo ya que otras zonas pueden requerir más tiempo de ruta y se necesitaría un camión con mejor rendimiento, por ejemplo. De igual manera, para el planteamiento de las rutas, se asume que los camiones se encuentran con el tanque lleno por lo que no se incluye posibles desviaciones a gasolineras.

En cuanto a los tiempos establecidos, únicamente se toma en consideración un tiempo promedio de espera de 10 minutos por domicilio pero no se toman en consideración otros factores como el tráfico, accidentes, o siniestros naturales; esto ya pasaría a ser un modelo estocástico pero por lo pronto está limitado a ser determinista. Otra posible mejora en cuanto a la resolución en sí sería la inclusión de la asignación de camiones al código. A pesar de que se encuentra la solución por medio de GAMS, se divide el proceso en dos y puede existir algún error si se pasan mal los datos del código al archivo .gms, además de que se necesita más tiempo para obtener la solución.

Finalmente, se recomendaría trabajar con APIs más precisas para calcular las rutas, puesto que la API utilizada es gratis y no cuenta con información importante como zonas con mayor afluencia vehicular. Sin embargo, a pesar de las limitaciones que se han tenido en este trabajo, se debe mencionar que se cuenta con resultados beneficiosos que toman en cuenta la mayor cantidad de detalles posibles dentro de un escenario determinístico.

VIII. ANEXOS

A. Mapa

<https://www.google.com/maps/d/u/0/edit?mid=1QgdEspOEnpcFyz-Faa6MpjdEQScqlhg&usp=sharing>

B. Unidades de transporte

C. Equipo y software utilizado

REFERENCES

- [1] D. P. B. Riveros and P. P. B. Silva, "Importancia de la administración logística," *Revistar UTP*, no. 1, pp. 217–222, 2008.
- [2] A. C. MARTÍNEZ, "Cooperación en los problemas del viajante (tsp) y de rutas de vehículos (vrp): una panorámica." 2011. [Online]. Available: http://eio.usc.es/pub/mte/descargas/proyectosfinmaster/proyecto_762.pdf
- [3] Coppel, "Acerca de nosotros," 2022. [Online]. Available: <https://www.coppel.com/acerca-de-nosotros>
- [4] —, "Ubica tu tienda coppel," 2022. [Online]. Available: <https://www.coppel.com/ubicacion-de-tiendas-coppel>
- [5] R. G. Cuéllar and L. A. R. Tovar, "Comercio electrónico en México: propuesta de un modelo conceptual aplicado a las pymes," *Revista Internacional de Ciencias Sociales y Humanidades, SOCIOTAM*, vol. 15, no. 1, pp. 79–116, 2005.

Unidad	Nombre	Modelo	Marca
01	CADDY MAXI	2020	Volkswagen
02	ELF 300	2013	ISUZU
03	PARTNER MAXI	2021	PEUGEOT
04	ELF 100	2022	ISUZU
05	514 SEMI LONG S 300	2019	HINO
06	MANAGER FURGON L4H2	2021	PEUGEOT
07	CHASIS CABINA	2020	NISSAN
08	616 LONG SERIE 300	2019	HINO
09	716 SUPER LONG S300	2022	HINO
10	EXPERT	2021	PEUGEOT
11	CHASIS CABINA	2019	NISSAN
12	ELF 300	2015	ISUZU
13	716 LONG SERIE 300	2022	HINO
14	ELF 100	2020	ISUZU
15	616 LONG SERIE 300	2013	HINO
16	616 LONG SERIE 300	2014	HINO
17	716 LONG SERIE 300	2019	HINO
18	616 LONG SERIE 300	2022	HINO
19	TRANSPORTER	2021	VOLKSWAGEN
20	ELF 300	2016	ISUZU
21	DUCATO VAN 11.5 M3	2021	FIAT
22	ELF 300	2014	ISUZU