Multiple Vehicle Routing Problem

Date: 18/05/2025	**Version: 2.0**
Course: Managerial Decision Making & Modeling	

Principal investigator(s)	
Name	**email**
Samuel Aidoo	-
Elisa Massano	-

Document review				
Version	**Date**	**Modified pages / sections**	**Description / Modifications**	**Transmitted on date**
V1.0	10/04/2025	All	First draft	10/04/2025
V2.0	18/05/2025	All	Final version	

Abstract

This project addresses a **Closed Multi-vehicle Routing Problem (CVRP)**, where a fleet of trucks departs from a central depot to deliver goods to multiple customer locations and returns to the depot upon completion. The primary objective is to minimize the total travel distance, thereby reducing overall transportation costs and improving logistical efficiency. To solve this optimization problem, we utilized Google OR-Tools, leveraging a range of built-in solvers to generate cost-effective routing plans under vehicle capacity and route feasibility constraints.

Real geographic coordinates were used to compute actual road distances via OpenStreetMap, and demands were split into virtual nodes to handle cases where a customer's demand exceeded vehicle capacity. We initially applied a Nearest Neighbor heuristic, followed by improving approaches such as *Guided Local Search*, *Tabu Search*, *Simulated Annealing*, *Automatic*, and *Greedy Descent* to refine the solution. Additionally, we explored the Cluster-First, Route-Second method to assess alternative routing strategies. These methods allowed us to test and compare solution quality across different heuristic strategies.

For visual analysis and validation, QGIS was used to map the optimized routes onto the real-world road network. The visual outputs facilitated clearer interpretation and decision-making, and supported practical application in logistics planning. The results demonstrated significant improvements in route efficiency and cost reduction, validating the effectiveness of combining advanced routing algorithms with geospatial visualization.

Acronyms and definitions

Nodes	Customers to whom glass bottles are delivered.
Depot	Facility where glass bottles are stored, organized, and loaded onto trucks for distribution to various destinations, serving as the departure point for transportation operations.
Vehicle	Trucks used for deliveries.

Summary

Abstract 2
Acronyms and definitions 3
Problem statement 5
System 5
  Elements 6
    Agents/DMs 6
    Entities 6
  Relationships among elements 6
  Other constituents of the system 7
Mathematical model(s) 8
  Sets 8
  Variables 8
    Decision variables 8
    Service variables 8
  Parameters 8
  Constraints 8
  Objectives 9
  Model 9
Data collection 10
Scenario analysis 10
  Scenario 1 10
  Scenario 2 11
  Scenario 3 11
Attachments 12

Problem statement

O-I is a global leader in the manufacturing and distribution of glass containers and bottles for various industries, including food, beverages, and personal care products. Among its many international operations, the company has a significant presence in Italy, where the San Polo di Piave facility in Treviso plays a key role in producing high-quality glass bottles for a diverse range of customers, such as major beverage brands, wineries, and consumer goods companies. At this facility, O-I faces logistical challenges related to efficiently managing deliveries. With multiple vehicles departing daily to serve various customer locations, the plant must plan routes that ensure timely and accurate shipments while optimizing the use of available resources. The company aims to minimize transportation costs, reduce delivery times, and maximize vehicle utilization, all while meeting specific customer demands. This project focuses on developing a robust multi-vehicle routing optimization system tailored to the operations of the San Polo di Piave facility. The goal is to streamline its distribution network, improve operational efficiency, reduce logistical costs, and enhance customer satisfaction. The motivation behind this optimization effort lies in a strong commitment to advancing supply chain performance and reducing operational costs through effective and efficient logistics planning, ensuring the reliable, timely delivery of O-I's high-quality glass products across the region.

System

The system consists of a central depot operated by O-I at its San Polo di Piave facility in Treviso, Italy, a manufacturer and distributor of high-quality glass bottles. Orders are received from customers located across the Veneto region. A fleet of delivery trucks is available to transport glass bottles from the depot to customer locations. Each truck has a specific loading capacity, and each customer has a known demand in terms of pallets, each containing between 26 to 34 pallets of glass bottles. The objective of the system is to determine the optimal set of delivery routes for the trucks so that all customers are served at the minimum total travel cost, while ensuring that no truck exceeds its capacity.

Assumptions:

1. Each truck has a fixed capacity measured in pallets, which depends on its size. The possible capacities are either 26 or 34 pallets.
2. A maximum of 60 trucks departs daily from the depot.
3. All customers must be served.
4. Each vehicle must start its route at the depot, serve one or more customers, and then return to the depot without the possibility of reloading during the route.

Elements

Agents/DMs

General manager (decision maker): All routing decisions are made centrally at the depot, which manages the assignment of trucks to customer locations. This characterizes the problem as a centralized vehicle routing problem.

Entities

Depot: The depot is the central hub where all the supplies are coordinated and distributed to all customers. It is the supply source for fulfilling customer demands.
Customers: Recipients of the glass bottle deliveries, each with a specific demand and location within the Veneto region.
Vehicles: Delivery trucks used to transport glass bottles and containers from the depot to customer locations, each with a fixed loading capacity.

Relationships among elements

1. **Depot ↔ Vehicles**

The depot acts as the supply source for fulfilling customer demands.

- **Relationship:** The depot is the starting and ending point for all vehicles. Vehicles depart from the depot fully loaded and must return to the depot after completing their routes.
- **Attributes:** Number of available vehicles (fleet size), and capacity of each vehicle (26 or 34 pallets).
- **Constraint Imposed:**
    - ➢ The total number of vehicles dispatched cannot exceed the fleet size.
    - ➢ Each vehicle must start and end its route at the depot without reloading.
    - ➢ Vehicle capacities limit the total demand served per route.

2. **Vehicles ↔ Customers (Virtual Nodes)**

Vehicles deliver goods to assigned customers along routes. In the optimization problem, vehicles follow the shortest feasible routes to minimize costs and delivery time.

- **Relationship:** Vehicles travel along routes visiting virtual customer nodes to deliver pallets, fulfilling demand. Virtual nodes allow modeling of multiple visits or splitting demand.
- **Attributes:** Customer demand (expressed in number of pallets), and geographical coordinates (used in distance matrix).
- **Constraints Imposed:**
    - ➢ The total demand must not exceed the combined capacity of the vehicles.
    - ➢ Each virtual customer node must be visited exactly once to satisfy demand fully.
    - ➢ Customers may be served multiple times through virtual nodes, but overall demand must be met.
    - ➢ Routes must be feasible within loading/unloading constraints.

3. **Customers (Virtual Nodes) ↔ Customers (Original Nodes)**

Since certain customers have demands exceeding the limited capacity of a single truck, their demand is split into multiple virtual nodes.

- **Relationship**: Virtual nodes correspond to physical customers, possibly splitting one customer's demand into multiple virtual nodes.
- **Attributes**: Demand splitting scheme, demand per virtual node, geographical coordinates, net profit per virtual node.
- **Constraints Imposed**:
  - ➢ The sum of demands of virtual nodes associated with a customer must equal the customer's total demand.
  - ➢ The demand of each virtual node must not exceed the vehicle capacity limit.

4. **General Manager (DM) ↔ All Elements**

The General Manager (DM) makes centralized routing and vehicle assignments.

- **Relationship:** The DM oversees fleet deployment, routing decisions, and ensures service quality and cost efficiency.
- **Attributes:** Objective priorities (minimize total distance/cost, ensure timely delivery), fleet size, vehicle capacities.
- **Constraints Imposed:**
  - ➢ The DM ensures full coverage of all customer demands without any shortfalls.
  - ➢ Routing decisions must respect capacity and operational constraints.
  - ➢ Optimization must balance cost minimization and service level requirements.

Other constituents of the system

Various external and internal factors may influence the decision maker's actions, many of which lie beyond their direct control. These are disruptions that originate outside the core operational scope but can significantly impact planned delivery routes and schedules.

- **Traffic congestion**: Unexpected traffic jams due to accidents, roadwork, or peak travel hours can increase travel times and interfere with delivery schedules. This may result in delayed deliveries, higher fuel consumption, driver overtime, and customer dissatisfaction.
- **Road closures**: Both planned and unplanned closures caused by construction, events, or emergencies can force detours, increasing travel distances and delivery times.
- **Vehicle breakdowns**: Delivery trucks are subject to mechanical failures, which can occur without warning. Such breakdowns can halt deliveries, require vehicle replacements or repairs, and cause delays in reaching customers.

Mathematical model(s)

In this section, we formulate the mathematical model for the considered problem. The model is designed to capture the essential characteristics and constraints of the system in a structured and quantitative manner.

Sets

- N: Set of all virtual nodes (depot and customers), $N = \{0,1,2,\ldots,n\}$, where:
  - 0 represents the depot, and
  - $1,2,\ldots,n$ represents the customer nodes.
- C: Set of virtual customers (excluding depot), $C \subset N$.
- V: Set of vehicles $V = \{1,2,\ldots,m\}$, where each $v \in V$ represents a vehicle.

Variables

Decision variables

- x_{ij}: Binary variable used to model the decision of traveling from node i to node j, $\forall i,j \in N$, where:
  - $x_{ij} = 1$ if a vehicle travels directly from node i to node j, or
  - $x_{ij} = 0$ otherwise.
- u_i: Integer variable that represents the load on the vehicle just before arriving at node i, where $u_i \in [0,Q]$.
  - When travelling from node i to node j, the load changes as $u_j \le u_i - d_i$, $\forall i,j \in N$.

Parameters

- $c_{i\,j}$: Distance (cost) from node i to node j; $c_{i\,j} = 0$ when $i = j$ (i.e., the distance from a node to itself).
- d_i : Demand of customer $i \in C$, expressed in number of pallets.
- Q: Capacity of each vehicle $v \in V$.

Constraints

- **Each virtual customer is visited exactly once.** This ensures that every virtual node is entered and exited exactly once, satisfying both customer service and flow conservation requirements.

$$\sum_{i \in N} x_{ij} = 1 \quad \forall j \in C$$

$$\sum_{j \in N} x_{ij} = 1 \quad \forall i \in C$$

- **Each vehicle starts and ends at depot.** Assuming that the depot is node 0, we ensure that at most V vehicles depart from and return to the depot.

$$\sum_{j \in C} x_{0j} \leq V$$

$$\sum_{i \in C} x_{i0} \leq V$$

- **Capacity constraint.** This is the Miller-Tucker-Zemlin (MTZ) formulation, used to eliminate subtours and ensure that vehicle capacity is not exceeded.

$$u_j - u_i + d_i \leq Q\left(1 - x_{ij}\right) \quad \forall i, j \in C, i \neq j$$
$$d_i \leq u_i \leq Q \quad \forall i \in C$$

Objective(s)

The objective of our project is to find the optimal routes that serve all customers while minimizing the total travel distance (and thus the cost). Formally, the objective function is:

$$min \sum_{i \in N} \sum_{j \in N} c_{ij} x_{ij}$$

Model

The full model can be displayed as follows:

$$min \sum_{i \in N} \sum_{j \in N} c_{ij} x_{ij}$$
$$\textstyle\sum_{i \in N} x_{ij} = 1 \quad \forall j \in C, \sum_{j \in N} x_{ij} = 1 \quad \forall i \in C$$
$$\textstyle\sum_{j \in C} x_{0j} \leq V \text{ , } \sum_{i \in C} x_{i0}\, x_{i0} \leq V$$
$$u_j - u_i + d_i \leq Q(1 - x_{ij}) \quad \forall i, j \in C; i \neq j$$
$$d_i \leq u_i \leq Q \quad \forall i \in C$$
$$x_{ij} \in \{0,1\} \quad \forall i, j \in N$$
$$u_i \in [0, Q] \quad \forall i \in N$$

Data collection

The data used for this project was kindly provided by a relative of one of the investigators, who works for the company under consideration. They provided general insights into the types and relative sizes of some of the clients served by the San Polo di Piave facility of O-I. Based on this, we identified the clients and retrieved the geographical coordinates for each of them. We then approximately estimated both their demand and their net profit — the latter being a feature used not to build the model, but to evaluate the quality of the resulting solutions. The collected data was initially stored in a txt file and later converted into a JSON file for use in the analysis.

Assumption: The clients of the San Polo di Piave facility, located in Treviso, are all wineries of various sizes. Based on domain knowledge, larger wineries are capable of bottling around 150,000 bottles of wine per day, whereas smaller ones produce approximately 4 million bottles per year. From this, we arbitrarily estimated the daily pallet demand in the following way.

- Large wineries: 100 – 150 pallets per day
- Medium wineries: 50 – 100 pallets per day
- Small wineries: fewer than 50 pallets per day

Assumption: To estimate the net profit for each customer, we assumed that profits typically exceed costs by approximately 25 – 30%, accounting for production and fixed operational expenses, as suggested by our source. Specifically, we used the following information: the standard selling price for a pallet containing glass bottles is approximately €150. We assumed a production cost of €80 per pallet, which is a reasonable estimate for standard glass manufacturing.
Using these values, the approximate daily net profit for each customer (in terms of their daily demand) was calculated as:

$$daily\ demand\ \cdot (selling\ price\ per\ pallet - production\ cost)$$

Assumption: Based on our source's guidance, we estimated a transportation cost of €1.75 per kilometer. To calculate the true overall net profit, we subtracted the total transportation cost from the overall net profit before travel cost, that is €90,115. The formula used is:

$$net\ profit\ of\ all\ nodes\ before\ travel\ cost - (total\ travel\ distance * cost\ per\ km)$$

Scenario analysis

Scenario 1

We initially approached the problem using a heuristic method, specifically the Nearest Neighbor algorithm. To model real-world conditions more accurately, we computed actual road distances between customers using OpenStreetMap, rather than relying on simple Euclidean distances. These were organized into a distance matrix representing the cost of travel between virtual nodes, which we created by splitting original nodes whose demand exceeded the vehicle's capacity. We formulated the problem in OR-Tools, defining constraints such as vehicle capacity and optimizing for the shortest total travel distance. The model output includes optimized vehicle routes, specifying both the customer visit sequence and the total distance traveled.

We first generated a feasible solution using a constructive heuristic — *Path Cheapest Arc* — and then applied *Guided Local Search* to iteratively improve the solution and escape local optima.

The first solution yielded a total travel distance of 9,184 km, successfully delivering 1,517 pallets and serving all customers. While it respects vehicle capacity and ensures full service coverage, this approach tends to treat deliveries in isolation. As a result, many vehicles complete routes that serve only one or two customers before returning to the depot. This leads to excessive travel distances and underutilization of vehicle capacity, inflating variable costs such as fuel consumption, driver wages, and vehicle maintenance. Although the solution meets logistical requirements, it does so inefficiently from an economic perspective.

In contrast, the improved solution produced by *Guided Local Search* retains the same service level, i.e. all customers are served, and the full demand of 1,517 pallets is delivered but reduces the total travel distance to 8,971 km. This improvement does not stem from reducing service coverage, but rather from smarter route optimization and more efficient use of vehicle capacity. Economically, this enhancement is significant: reduced travel distances directly lower operational costs and increase fleet productivity. More efficient routing also decreases vehicle idle time and, if scaled, could reduce fleet size without compromising service. Thus, the *GLS*-based solution represents a substantial improvement in cost-efficiency, logistical effectiveness and overall profitability, with the first solution yielding a net profit of €74,043, compared to €74,415.75 achieved by the *GLS*-based approach.

Scenario 2

After the Nearest Neighbor approach, we explored several local search strategies to assess potential improvements in the solution quality. In combination with the *Path Cheapest Arc* construction, we tested *Tabu Search*, *Automatic*, *Simulated Annealing*, and *Greedy Descent*.

Interestingly, *Tabu Search* yielded the same optimized solution as *Guided Local Search*, demonstrating that different metaheuristic strategies can converge on a similarly high-quality outcome. This reinforces the validity of the improved solution and highlights its robustness across optimization methods. *Tabu Search*'s success in this context may stem from its ability to systematically avoid cycling back to recently visited solutions, enabling it to explore the solution space more effectively and escape local optima — an advantage that is especially useful in combinatorial problems like vehicle routing.

On the other hand, *Simulated Annealing*, *Greedy Descent*, and the *Automatic* strategy all reproduced the initial constructive solution, with a total distance of 9,184 km. This suggests that while these methods may offer moderate improvements over naive heuristics, they can struggle to escape local optima or sufficiently explore the solution space under certain configurations. As a result, their performance aligns with the basic constructive heuristic, making them less suitable when deeper optimization is required.

Scenario 3

In the final stage of our analysis, we implemented the Cluster-First, Route-Second approach to vehicle routing, evaluating its performance under both geographic and demand-based clustering schemes. This two-step method — first grouping deliveries into coherent clusters, then optimizing routes within each cluster — is widely used in logistics for its scalability and intuitive appeal. For route optimization within each cluster, we used Google OR-Tools, applying a two-phase process:

first, a constructive approach using the *Path Cheapest Arc* heuristic to generate initial routes, followed by the *Guided Local Search* metaheuristic to refine and improve them. This setup allowed us to assess the effectiveness of clustering when coupled with practical optimization techniques.

When applying geographic clustering, the method successfully delivered all 1,517 pallets while respecting the truck capacity limit per vehicle. The total distance traveled was 9,180 km, nearly identical to the 9,184 km observed using our initial Nearest Neighbor heuristic. This outcome demonstrates that proximity-based clustering can produce feasible and reasonably efficient routing solutions. However, inefficiencies emerged: many vehicles were assigned to serve only a single customer before returning to the depot, despite the presence of nearby stops within the same cluster. This pattern led to underutilized vehicle capacity and elevated operational costs, as each route incurred fixed expenses (fuel, driver time, vehicle depreciation) regardless of the number of deliveries made.

To assess the robustness of the approach, we repeated the process using customer demand as the clustering criterion. The results were consistent: all 1,517 pallets were delivered, and the total distance remained at 9,180 km. This consistency suggests that while the clustering basis changed, the route configuration and underlying inefficiencies were largely unaffected.

In conclusion, the Cluster-First, Route-Second method offers a practical strategy for decomposing complex vehicle routing problems into manageable subproblems. However, to achieve true cost-efficiency and route optimization, more advanced intra-cluster techniques, such as the integration of local search metaheuristics like Guided Local Search, as implemented here, are essential.

Attachments

The project documentation includes the following files:

- **data.json**: JSON file containing node information, used to build the models.
- **Project.ipynb**: Google Colab notebook where the analysis and computation were performed.
- **distance_matrix_km.csv**: Output file containing the distance matrix computed using OpenStreetMap for the original nodes.
- **first_solution_routes.geojson**: GeoJSON output file containing the initial solution based on the constructive Nearest Neighbor approach.
- **second_solution_routes.geojson**: GeoJSON output file containing the complete Nearest Neighbor solution, including both the constructive phase and the local search improvement.
- **clustering_by_closeness.geojson**: GeoJSON output file representing the solution from the Cluster-First, Route-Second approach using geographical proximity for clustering.
- **clustering_by_demand.geojson**: GeoJSON output file representing the solution from the Cluster-First, Route-Second approach using daily demand for clustering.
- **Project_visualization.qgz**: QGIS project file used to visualize the routing solutions on a map of the Veneto region.