Decision Science Letters 14 (2025) 327–338

Contents lists available at GrowingScience

Decision Science Letters

homepage: www.GrowingScience.com/dsl

Designing of a dynamic logistics platform for optimization of truck assignment and its route for KINZA company

Osamah Abdulhameed[a*], Sali Ghanem[a], Rafal Sadeq[a], Reemas Al Ghamdi[a], Dalal Al Mazyad[a] and Naveed Ahmed[a]

[a]*Industrial Engineering Department, College of Engineering, Al-Yamamah University, Riyadh-11512, Saudi Arabia*

C H R O N I C L E

Article history:
Received: October 16, 2024
Received in revised format: November 30, 2024
Accepted: January 22 2025
Available online: January 22, 2025

Keywords:
Supply Chain Management
Vehicle Routing Problem (VRP)
Travel salesman problem (TSP)
K-mean clustering

A B S T R A C T

A consideration of the integral variables of customer location, traffic flow, and road conditions to determine the best feasible delivery routes is a big challenge in Logistical operations. A poor routing strategy that delivers products places an ineffective gloss and eventually converts into high operating expenses, over-consumption of fuel, and shipment delays. The paper's goal is to build a model for the logistics management of the company which aims for effective management of the truck allocation and vehicle routing using K-means clustering and TSP. K-means clustering is often used to classify the sites of delivery based on their closeness in space, hence simplifying the problem by reducing its dimensionality. The proposed algorithm considered customer location prioritization in deliveries, delivery task allocation, and truck allocation to enable timely delivery. Therefore, this paper presented a solution to enhance the logistics operations of beverage brand "KINZA" by optimizing its truck loading and delivery route. The model would ensure that each truck is able to travel optimally, with vehicle-routing algorithms applied in a way to avoid all unnecessary waste of time and distance. Finally, the main scope of this paper is to develop and design a dynamic logistics platform for the KINZA Company distribution network.

1. Introduction

AL-JAMEEL Company is one of the leading companies in the Saudi Arabian beverage industry, with several products under its well-liked brand called KINZA has grown rapidly in the marketplace because of its commitment to quality and customer satisfaction. Since the brand keeps on expanding, it faces some serious logistical hindrances related to delivering its products effectively to the entire Saudi Arabian landscape.

The size of the country, coupled with increasing demand for timely delivery, has brought inefficiencies to the supply chain operations at KINZA Company. The inefficiencies are manifested in the form of poor truck assignments and suboptimal routing that result in excessive fuel consumption, increased operating costs, and delayed customers' demands. The company feels the need to automate data-driven optimization in logistics operations if it has to remain competitive. The paper will design and develop an automated supply chain management platform to support KINZA Company in improving its vehicle assignments and delivery routes. Advanced algorithms combined with real-time data analysis create one of the most efficient logistic operations to best service customers of the company.

* Corresponding author.
E-mail address o_abdulhameed@yu.edu.sa (O. Abdulhameed)

ISSN 1929-5812 (Online) - ISSN 1929-5804 (Print)

doi: 10.5267/j.dsl.2025.1.005

2. Problem Definition

Saudi Arabia's vast geographical area poses transportation challenges that result in inefficiencies in product delivery which leads to increased fuel consumption, operational costs, and long delivery times. KINZA Company-the beverage brand of AL JAMEEL GLOBAL Co.-has been reeling under some logistics inefficiency related to its truck routing problem and route optimization of deliveries. This situation is further exacerbated by the fact that Saudi Arabia, where Al-Jameel operates, comprises a land area of about 2.15 million square kilometers. Therefore, the most significant problem becomes that the current logistics structure of KINZA Company is not well optimized enough to enable proper truck loading and routing. This present system does not recognize the difference in distances of various places of delivery and the capability of the trucks. It therefore results in unwanted journeys, excess usage of petrol. This one aspect of inefficiency trickles down through longer time taken for delivery, more consumption of petrol, and increase in operational costs. In this direction, the proposed research has focused on the conceptualization and development of a customizable KINZA Company Automated Supply Chain Management Platform. It shall be developed to ease processes related to truck allocation and route planning for delivery using a data-driven model and algorithm. The paper objectives hence targeted upgrading a number of aspects of the KINZA company processes.

1. To minimize fuel consumption and reduce the number of trips needed for delivery,
2. To improve the allocation of trucks and ensure better utilization of delivery resources,
3. To develop an optimized delivery route model for KINZA distribution operations,
4. To enhance delivery times and overall operational efficiency.

The outcomes of the first objective will relate to developing a sturdy and data-driven model capable of realizing delivery with proper route planning. The existing structure laid down by KINZA Company has not considered distances between consumers and their geographical spread, which has meant lower efficiency in delivery. The best model will consider the integral variables of customer location, traffic flow, and road conditions to determine the best feasible delivery routes. The model would ensure that each truck is able to travel optimally, with vehicle-routing algorithms applied in a way to avoid all unnecessary waste of time and distance. The process described here will contribute more toward the improvement of operating efficiency in the logistics wing at KINZA Company, thus helping the development plan cope with larger volumes.

Coming to the second objective, logistics require immense fuel consumption, especially when the countries are large in area, like Saudi Arabia, and the routes of delivery can be hundreds of kilometers. One of the key objectives of the present research is to reduce the number of trips to reduce the overall fuel expenditure. This would include load capacity, vehicle performance, and criticality of delivery factors in the advocated methodology for loading of trucks for certain orders to ensure that the vehicle is utilized to full potential in every delivery trip. In this way, with a minimum possible non-value-added drive time and distance, KINZA Company can reduce the substantial fuel consumption and, correspondingly, be more ecologically friendly with regard to its distribution operation.

An equally important objective is to improve the allocation of trucks within KINZA Company's distribution operations. The current system has an inefficient truck assignment by either underutilizing or overloading truck's capacities. The paper thus aims to dynamically assign the correct number of trucks for each Customer cluster with the efficient utilization for each truck. This would eliminate the nuisance of trucks playing around with partial or overloaded conditions beyond their capacity and thus deploying delivery resources in a much better way. Similarly, proper allocation will entail that each fleet of trucks would achieve higher delivery efficiency and prolong operational lives by reducing wear and tear on vehicles. In a competitive industry like beverages, Delivery time plays an important role in customer satisfaction and loyalty. Therefore, by considering the truck capacities and fuel consumption the data driven algorithm will make a huge difference in improving KINZA Company's overall operations. Therefore, helping KINZA Company meet increasing demand, eliminate distribution bottlenecks, and stay competitive in the market.

Finally, the main scope of this paper is to develop and design a dynamic logistics platform for the KINZA Company distribution network. This system will utilize client location information, Order quantities, and vehicle capacity data to optimize delivery routes. This will assist the logistics and distribution teams at KINZA Company in creating a more automated and efficient method for vehicle allocation and delivery routing. Although KINZA Company's delivery efficiency is the primary goal, the solution could be easily applied to other beverage logistics or any business facing similar distribution challenges. The long-term scalability of this solution makes it a valuable tool for both KINZA Company's current operational needs and its future market expansion.

3. Literature Review

The consideration of logistical processes becomes a common study field that largely covers vehicle routing, with the scheduling of delivery times included within the scope of operations research and supply chain management studies. The core of this field is the so-called Vehicle Routing Problem (VRP), which is focused on how to find an optimal route to distribute goods or render services by using a fleet while minimizing the operational expenses (Johnson, 1990). Common Vehicle Routing Problem variations include the Capacitated Vehicle Routing Problem, Vehicle Routing Problem with Time Windows, and Multi-Depot Vehicle Routing Problem. Each of these variants adds additional constraints: vehicle capacity is limited, time windows for deliveries, multiple depots, respectively (Applegate, Bixby, Chvátal, & Cook, 1998). Considering these challenges, investigators have devised a range of heuristic and metaheuristic methodologies for approximating optimal solutions. Probably the most popular techniques in this domain are K-means clustering and TSP (Hoffman, Padberg, & Rinaldi, 2001). K-means clustering is often used to classify the sites of delivery based on their closeness in space, hence simplifying the problem by reducing its dimensionality (Lloyd, 1982). On the other hand, TSP is the problem of finding the shortest possible tour that passes through a set of locations and returns to the initial location; it constitutes the core of most vehicle routing solutions (Rosenkrantz, Stearns, & Lewis, 1977).

In the field of logistics optimization, mainly on route planning for deliveries, several studies have focused their attention on the use of clustering algorithms and heuristic methods in solving complex routing problems, such as the Traveling Salesman Problem (TSP) which is a computational problem that seeks to find the most appropriate path which links a set of predetermined points or locations to be visited. In this case, it also refers to the sets of cities which a salesman should travel through. The role of the salesman in general is to reduce the cost of traveling and the total traveled distance (Johnson, 1990). The use of such methods becomes imperative in industries that strive to reduce costs of delivery, optimize resource allocation, and minimize waste in operations (Applegate et al., 1998; Lee et al., 2022). The literature on the Traveling Salesman Problem (TSP) demonstrates significant advances in both theoretical approaches and practical applications. The TSP is a widely recognized NP-hard combinatorial optimization problem. It searches the most effective route through a predetermined number of cities, visiting each one once before returning to the beginning. Despite the simplicity of its description, efficiently solving large instances of the TSP remains a formidable challenge, and various strategies have been investigated to address this complexity. As stated by Hoffman, Padberg, and Rinaldi, TSP has become representative since, although its small-sized instances are solvable with enumerations using brute force, for larger-sized instances-for example, 16 cities, yielding over 653 billion possible tours-more sophisticated mathematical formulations have to be invoked to bypass explicit enumerations (Hoffman et al., 2001). Their work aims at the improvement of algorithms by avoiding enumerations without round trips and focuses on cut-plane methods combined with relaxation techniques in solving large-sized TSP problems, including those occurring in practical applications, such as VLSI systems that can allow thousands of cities - Applegate et al., 1998. Johnson, 1990 presents a detailed survey of heuristic algorithms related to TSP and lays considerable stress on the local optimization methods. He first mentions the considerable role of approximation algorithms since it is unlikely that there might be a polynomial-time method to identify the optimal solution. Then he focuses on the latest progress made by the local optimization techniques of 2-Opt, 3-Opt, and the Lin-Kernighan algorithms. The most important argument herein is that these methods have been found to work in practice, especially within the paradigm of metaheuristic search techniques such as simulated annealing and genetic algorithms (Johnson, 1990). He further makes a very worthy contribution to the study of the PLS-completeness theory, with significant contribution to an understanding of the use of computational time for various methods of local optimization (Rosenkrantz et al., 1977). Applegate, Bixby, Chvátal, and Cook (1998) give full information about the latest progress in solving big-sized instances of the Traveling Salesman Problem (TSP) using the cutting-plane method. In fact, such cutting plane algorithms have, since the seminal work of Dantzig, Fulkerson, and Johnson in 1954, formed the methods of choice for solving, in practice, TSPs containing several hundred cities (Applegate et al., 1998). The general idea is to iteratively tighten the linear programming relaxations of TSP, thereby bringing the feasible region closer to the true solution. These developments mean that a TSP instance as large as 13,509 cities could be solved and showed scalability for such a method when combined with modern computational methods (Hoffman et al., 2001). In this context, the work of Rosenkrantz, Stearns, and Lewis (1977) has reviewed intensively a number of heuristic approaches in order to solve the TSP.

The researchers seek polynomial time approximation algorithms like Nearest Neighbor and Insertion heuristics, which produce "good" tours if not necessarily optimum ones. They extend their work of analyzing the performance bounds of the techniques. To date, the authors have been able to show that whereas the methods cannot guarantee optimum solutions, they provide near-optimum solutions for big problems. They show that the nearest neighbor and insertion technique logarithmic bounds; hence, in cases where the exact algorithms are computationally infeasible, such methods are helpful (Rosenkrantz et al., 1977). In other words, TSP remained a test bed for enormous amounts of different novel optimization techniques, in which the exact algorithms, heuristics, and metaheuristics had been applied to the solution of a problem.

The most important approaches to dealing with growing sizes of TSP instances include cutting-plane methods, algorithms of local optimization, and various kinds of heuristics (Applegate et al., 1998). None of these methods proved to be universally better for all possible purposes, but in real life, this eternal problem is well solved by combining theoretical knowledge and practical heuristics (Hoffman et al., 2001).

The K-means algorithm, particularly its application in customer clustering for truck assignment, highlights its usefulness in enhancing logistical efficiency and resource allocation. The K-means algorithm, a popular clustering method, works by minimizing within-cluster variance to divide datasets into K different groups. It iteratively processes data points through initialization, assignment, update, and convergence to form clusters, making it suitable for logistics tasks like customer clustering in truck assignments (Lloyd, 1982).

Several studies have explored K-means in this context, particularly for optimizing truck assignments. Yadav et al. (2023) demonstrated how the algorithm can segment customer bases by factors such as delivery frequency and geographic proximity, leading to more effective route planning and reduced delivery times. Similarly, Patel and Desai (2023) investigated the dynamic assignment of trucks to clustered customer groups by integrating K-means with real-time data analytics, allowing for adjustments based on fluctuating demands. The literature also discusses the combination of K-means with other methods to enhance truck assignment efficiency. For example, Kumar and Singh (2023) proffered a hybrid model that combines K-means with genetic algorithms, which improved fleet efficiency and the preciseness of the assignment. However, the power of K-means suffers from certain limitations in performance, especially regarding sensitivity to initializations. Initial centroid choices may also more greatly affect the outcome, and Huang et al. (2022) suggested the use of methods like K-means++ to help improve cluster results and initializations for K-means in logistics applications. Other works, which are on truck assignment related to the application of K-means for customer segmentation, sound very promising, especially with regard to the real-time integration of IoT technologies with big data analytics. According to Patel and Desai (2023), this may enlarge the algorithm's potential to instantly deliver insights into the consumers' behavior and, hence, further optimize the logistical industry.

In summary, K-means is quite powerful in the mechanism of enhancing truck assignments by clustering customers. K-means is very important in the contemporary management of transportation, whereby its potentiality of enhancement of customers' service towards higher logistical efficiency is highly regarded. Further development in data analytics will, therefore, help to provide proficient applications in this area for further assurance of improvement in transport logistics.

4. Design Methodology

This paper is unique in that it integrates real-time data analytics and route optimization algorithms specific to the logistics requirements of KINZA Company. Much of the literature, up to this point, has modeled theoretical models or applied general applications of API Distance Matrix. The paper is practical in nature, based on these APIs supplying real-time updates for traffic conditions, such as the Google Maps Distance Matrix API. Consequently, it allows the proposed platform to dynamically re-choose delivery routes based on current road conditions-a capability that is not often used in the traditional VRP study. The methodology used for the duration of the study is presented in the flowchart below shown in Fig. 1.

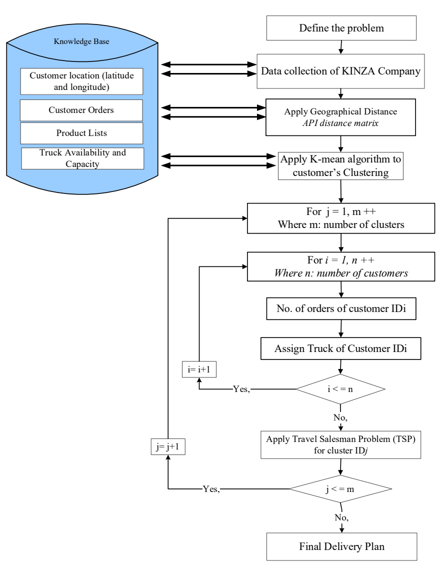


Fig. 1. Design Methodology

4.1 Mathematical formulation

The optimization aimed at minimizing fuel consumption where this constraint ensures that the load on the vehicle decreases by the demand of the customer served which are:

1. Real-Time Data:

The integration of real-time data such as traffic, weather, and road conditions add complexity to the route optimization process, necessitating a robust and adaptable system.

2. Vehicle Capacity:

The optimization model must ensure trucks are used efficiently without being overloaded or underutilized, which impacts both operational costs and delivery timelines.

3. Fuel Consumption:

The optimization model must reduce fuel consumption, by minimize the total distance traveled and optimize the number of trips.

4. Time Constraints:

Deliveries must be made within specified time frames to ensure customer satisfaction and maintain operational efficiency. Late deliveries could negatively impact customer relationships and overall service levels.

Below is the key formulation used where decision variables are:

X_{kij}: Binary variable that is 1 if vehicle k travels from customer i to customer j, and 0 otherwise.

y_{ki}: Load of vehicle k after serving customer i.

d_i: Demand of customer i.

Q_k: Capacity of vehicle k. $\forall k = 1, \ldots .8$

F_k: Fuel consumption rate of vehicle k per unit distance.

c_{ij}: Distance between customer i and customer j.

T_k: is the maximum number of trips allowed for truck kkk.

Sets:

C: Set of customers, indexed by i,j. (1,2,3….n) n=56

V: Set of vehicles, indexed by k. (1,2,….p) p=8

D: Depot, represented as node 0.

The objective function:

$$\min z = \sum_{k=1}^{n} \sum_{i=0}^{n} \sum_{j=0, i \neq j}^{n} F_k \cdot C_{ij} \cdot x_{k_{ij}} \cdot \left(\frac{y_i^k + Y_j^k}{2} \right)$$

The goal is to minimize the total fuel consumption, which is influenced by the load of the vehicle (since fuel consumption increases with load). We assume fuel consumption between two customers *i* and j is proportional to both the distance traveled and the load carried. At the start of the trip between customer *i* and customer *j*, the truck has a certain load y_{ik} after visiting customer *i*. At the end of the trip (after visiting customer *j*), the truck has a reduced load y_{jk}, since it has delivered goods to customer *j*. Instead of choosing one of these two loads (either y_{ik} or y_{jk}) to represent the truck's load during the trip from *i* to *j*, we take the average load between these two points.

subject to

$$\sum_{k=1}^{p} \sum_{i=0, i \neq j}^{n} X_{kij} = 1, \qquad \forall j \in \{1, \dots, n\}, \tag{1}$$

This constrains Ensures that each customer must be visited exactly once by one vehicle.

$$\sum_{j=1}^{n} x_{k0j} = 1, \sum_{i=1}^{n} x_{kj0} = 1, \qquad \forall k \in \{1, \dots, p\}, \tag{2}$$

Each vehicle must start and end its route at the depot.

$$\sum_{i=0, i \neq j}^{n} x_{kij} = \sum_{i=0}^{n} x_{rji}, \qquad Vk \in V, i \in C \tag{3}$$

If a vehicle arrives at a customer i, it must leave from that customer.

$$\sum_{i=0}^{n} \sum_{j=1, i \neq j}^{n} d_i k_{kij} \leq Q \tag{4}$$

The capacity constraints are stated, making sure that the sum of the demands of the customers visited in a route is less than or equal to the capacity of the vehicle performing the service.

$$\sum_{\mathrm{i}=0}^{n} \sum_{\mathrm{j}=0}^{n} x_{k\mathrm{ij}} \leq T_k \tag{5}$$

For every truck k, the sum of all trips (edges) from customer i to customer j cannot exceed the number of trips Tk allowed for that truck.

$$Y_{kj} \geqslant Y_{ki} - d_r x_{jk} \qquad \forall\, i, j \in c, k \in v \tag{6}$$

5. Result and discussion

5.1 Data Collection and Preparation

The initial phase involves gathering data needed to build the logistics platform. This data includes customer ID with their location (latitude and longitude), Product list and dimensions, Order quantities, and available truck capacities. The collected data is then cleaned and structured in a database, preparing it for further processing. Incomplete or erroneous entries are filtered out to avoid skewing the optimization process.

5.2 Application of Geographical Distance Using API Distance Matrix

After preparing the customer location data, the Google Maps Distance Matrix API computes the geographical distances between pairs of delivery points, also between KINZA Company warehouse location and every customer location. This API provides real-time traffic congestion, road closures, and alternative route information to compute the most up-to-date statistics on travel time and distances. The output from this step is a From-To matrix that is integrated into the optimization model, ensuring the routes reflect real-world conditions. This dynamic approach continuously updates travel times as traffic conditions change, maintaining the accuracy of the optimization process. Below is an explanation of the key parts of the code used to interact with the API Fig. 2. By using this real-time distance information, the system can dynamically adjust routes to account for current traffic conditions, ensuring optimal delivery efficiency.

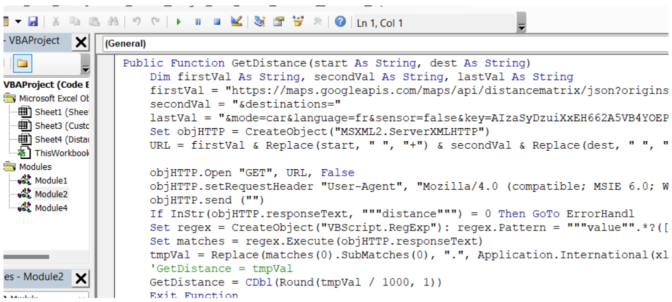


Fig. 1. Google Maps Distance Matrix API

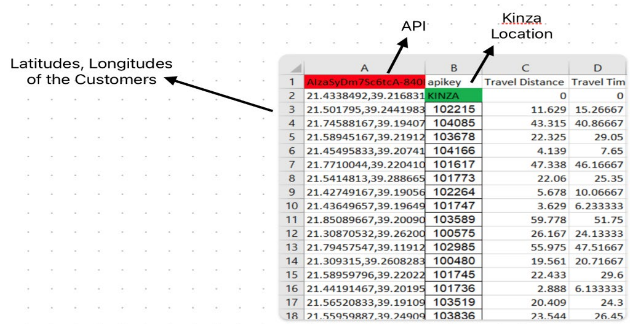


Fig. 2. Travel Distance and Time between KINZA Company Warehouse and Customers

The figure below shows how the locations of customers were fed into the system with their latitude and longitude, respectively. From the use of the API key, one can observe how it has been able to trace the calculation of travel distance and travel time from the main depot-that is, the KINZA Company warehouse-to each of the customer locations. Distances in kilometers and travel time in minutes between the warehouse and various delivery locations are as shown Fig. 3. As a result, the From-to matrix is now generated after finding Customer-to-Customer and KINZA Company-to-Customer Distance Fig. 4.

From/To	102215	104085	103678	104166	101617	101773	102264	101747	103589	100575	102985	100480	101745	101736	103519	103836
102215	0	31.3	15.2	9.8	35.3	10	14.4	11.5	47.7	30.8	44.4	24.2	15.3	10.7	13.3	12.5
104085	32.1	0	23	41	6.2	32.9	48	42.6	19	60.1	15.7	53.4	22.4	41.9	26.5	26.6
103678	12.6	20.5	0	21.1	24.6	13.4	28.4	22.7	37	40.5	33.7	33.9	0.1	22	9	7.1
104166	8.8	40.5	19.5	0	44.5	19.2	5.6	2.9	57	28.1	53.2	21.5	19.6	2.2	17.6	20.7
101617	38.3	6.7	29.2	47.2	0	39.2	54.2	48.9	14.3	66.3	17.5	59.7	28.6	48.1	32.8	32.8
101773	10.1	30.8	13.8	20.5	34.8	0	26.8	22.1	47.2	38.6	43.9	32	13.9	21.4	15.8	8.6
102264	15.2	47.9	25.9	8.2	51.9	25.7	0	4.9	61.9	27.2	55.9	17.5	26	5.8	23.2	29.6
101747	10.3	42	21	2.8	46.1	20.8	2.4	0	58.5	29.6	52.6	18.1	21.1	0.9	19.1	22.3
103589	47.3	18.4	38.2	56.2	15.8	48.1	57.1	57.8	0	75.3	18.2	68.7	37.6	57.1	41.7	41.8
100575	27.4	52.3	35.3	24.2	56.3	31.3	27.2	20.9	68.7	0	65.4	15.8	35.4	19.3	37.3	33.5
102985	40.1	16.6	30.9	48.5	15.1	40.9	48.5	50.1	16.8	68	0	61.4	30.3	49.4	34	34.5
100480	34.1	58.9	41.9	27.1	62.9	37.9	23.6	23.8	75.4	6.7	72	0	42.1	24.6	43.9	40.1
101745	12.5	20.7	0.1	21	24.7	13.3	28.3	22.6	37.1	40.4	33.8	33.8	0	21.9	8.9	6.9
101736	9.5	41.2	20.2	2	45.2	19.9	3.2	1.2	57.6	28.8	53.8	18.8	20.3	0	18.3	21.4
103519	13.2	27.8	6.1	17.3	31.8	15.3	19.4	19.6	44.2	42.5	37.7	35.9	6.2	16	0	8.5
103836	11.7	25.1	7.8	21.8	29.1	14.6	25.3	23.5	41.6	39.7	38.2	33.1	7.9	22.8	9.8	0
100956	17.4	24	8.8	19.1	29.8	20.1	23.4	23.7	37.6	47.2	31.6	40.6	8.9	20.1	5.5	12.7
103514	43.9	20.4	34.8	52.3	19	44.8	52.4	54	20.7	71.9	4.7	65.3	34.2	53.2	37.9	38.4
100192	33.5	0.8	24.4	42.4	5.9	34.3	49.4	44	14.9	61.5	15.6	54.9	23.8	43.3	28	28
102100	7.5	33	12.4	12.9	37	14.4	14.5	14.7	49.4	40.2	43.1	33.6	12.5	13.9	8.2	14.2

Fig. 3. Distance matrix

5.3 K-Means Clustering

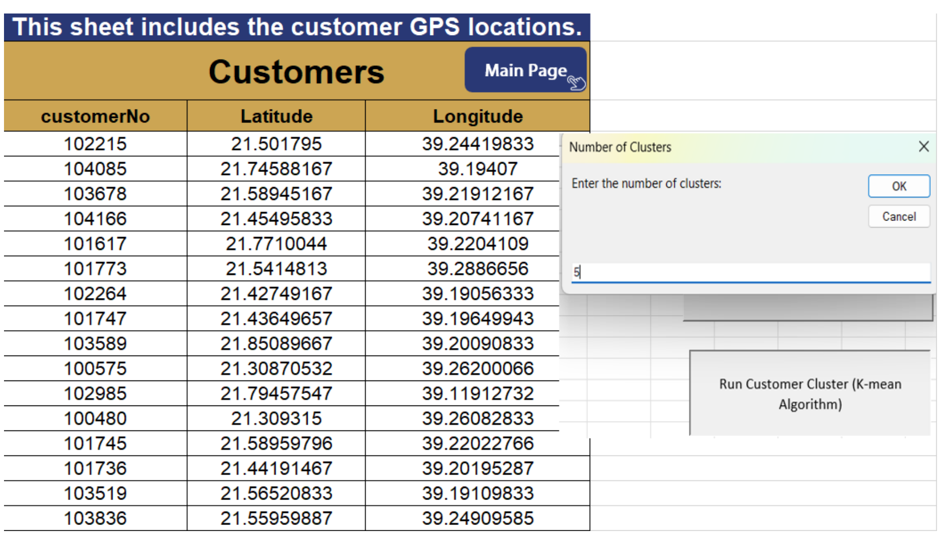
Next step, the K-means clustering algorithm is used to group customer locations based on geographical proximity. Initial cluster centroids are assigned, and delivery points are clustered based on distances obtained from the API. The centroids are iteratively refined until the clusters stabilize. This clustering process helps balance the delivery workload across trucks and simplifies the overall problem complexity. Here's a breakdown of the K-means clustering code Fig. 5.

```
[26]: n_clusters = 5
      kmeans = KMeans(n_clusters=n_clusters, random_state=42)
      sheet_data['KMeans_Cluster'] = kmeans.fit_predict(coords)

[28]: output_file_path_kmeans = 'kmeans_clustered_customers.xlsx'
      sheet_data.to_excel(output_file_path_kmeans, index=False)

[30]: kmeans_cluster_distribution = sheet_data['KMeans_Cluster'].value_counts()
      print("K-Means Clustering completed. Cluster distribution:")
      print(kmeans_cluster_distribution)

      K-Means Clustering completed. Cluster distribution:
      KMeans_Cluster
      2    20
      3    14
      1    14
      4     6
      0     3
      Name: count, dtype: int64

[32]: results = pd.DataFrame()

[34]: for cluster in range(n_clusters):
          cluster_data = sheet_data[sheet_data['KMeans_Cluster'] == cluster]

          if not cluster_data.empty:
              cluster_info = cluster_data[['customerNo', 'Latitude', 'Longitude']].copy()
              cluster_info.columns = [f'Cluster_{cluster}_customerNo', f'Cluster_{cluster}_Latitude', f'Cluster_{cluster}_Longitude']
              results = pd.concat([results, cluster_info.reset_index(drop=True)], axis=1)

[36]: results = results.fillna('')

[38]: output_file_path_results = 'clustered_results.xlsx'
      results.to_excel(output_file_path_results, index=False)
      print("\nCluster results saved to:", output_file_path_results)

      Cluster results saved to: clustered_results.xlsx
```

Fig. 4. application of K-mean algorithm.

- **n_clusters = 5**: This variable sets the number of clusters, which inserted by customer as shown in Figure 6.
- **kmeans.fit_predict(coords)**: The function fit_predict performs the clustering based on the delivery points' coordinates. This step assigns each delivery point to one of the clusters, which is stored in the column KMeans_Cluster.
- **Cluster distribution**: After clustering, the distribution of points across clusters is evaluated to ensure a balanced workload across trucks. The results are saved in an Excel file, making it easy to analyze and adjust the clusters if necessary.

Below, we outline the clustering process as executed within our customized Excel platform. The clustering process is initiated with a click on the button named “Run Customer Cluster (K-mean Algorithm)” as indicated in Fig. 6. The user is guided through the clustering setup by a sequence of prompts that are triggered by this button such as number of clusters as shown in Fig. 6.



customerNo	Latitude	Longitude
102215	21.501795	39.24419833
104085	21.74588167	39.19407
103678	21.58945167	39.21912167
104166	21.45495833	39.20741167
101617	21.7710044	39.2204109
101773	21.5414813	39.2886656
102264	21.42749167	39.19056333
101747	21.43649657	39.19649943
103589	21.85089667	39.20090833
100575	21.30870532	39.26200066
102985	21.79457547	39.11912732
100480	21.309315	39.26082833
101745	21.58959796	39.22022766
101736	21.44191467	39.20195287
103519	21.56520833	39.19109833
103836	21.55959887	39.24909585

Fig. 6. clusters No

Once all inputs are given, the platform runs the Python script, processing the customer's data and running the K-means algorithm. Figures 7 and 8 show the result of such clustering where customers were divided into five clusters. Each cluster defines a geographic group that can be served with a more efficient route plan.

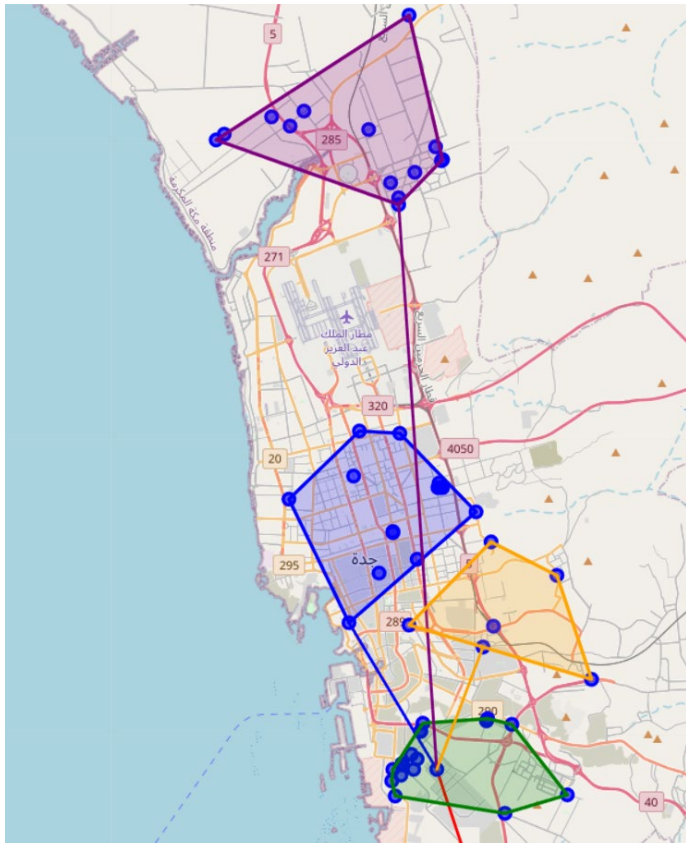


Fig. 7. Cluster Map

Cluster # 1			Cluster # 2			Cluster # 3			Cluster # 4			Cluster # 5		
customer ID	**Latitude**	**Longitude**	**customer ID**	**Latitude**	**Longitude**	**customer ID**	**Latitude**	**Longitude**	**customer ID**	**Latitude**	**Longitude**	**customer ID**	**Latitude**	**Longitude**
100575	21.30870532	39.26200066	103678	21.58945167	39.21912167	KINZA	21.43385	39.21683	104085	21.74588167	39.19407	102215	21.501795	39.24419833
100480	21.309315	39.26082833	101745	21.58959796	39.22022766	104166	21.45495833	39.20741167	101617	21.7710044	39.2204109	101773	21.5414813	39.2886656
101894	21.27643333	39.276445	103519	21.56520833	39.19109833	102264	21.42749167	39.19056333	103589	21.85089667	39.20090833	103836	21.55959887	39.24909585
			100956	21.5959682	39.1674466	101747	21.43649657	39.19649943	102985	21.79457547	39.11912732	102100	21.51347202	39.20027886
			100388	21.576593	39.24061	101736	21.44191467	39.20195287	103514	21.78516347	39.09091693	101906	21.51312667	39.25092833
			101725	21.56484246	39.19117289	100096	21.4607257	39.2468377	100192	21.74981428	39.19414742	102854	21.48383112	39.3092635
			101718	21.59208767	39.21878148	101720	21.41888093	39.19221498	103947	21.78748459	39.1762178			
			103345	21.62124689	39.17101239	104174	21.455005	39.20734667	100364	21.75840166	39.189875			
			100297	21.58334166	39.12929166	103932	21.45966	39.20858167	100997	21.78201066	39.08563463			
			100491	21.61951666	39.19491833	102309	21.462075	39.24696333	101605	21.76398492	39.20398858			
			100258	21.51479333	39.16489833	101743	21.4337653	39.19091618	100022	21.78962485	39.12966513			
			101936	21.55039	39.205445	100099	21.40930802	39.25782021	100574	21.79755704	39.13822987			
			101726	21.58954062	39.21799285	102231	21.43965667	39.20573667	100495	21.77768574	39.21662264			
			104054	21.54267	39.18264667	101709	21.43381696	39.20308027	101765	21.770426	39.219738			
						101734	21.45884577	39.26184909						
						100452	21.41975666	39.29488833						
						101732	21.430755	39.195945						
						101737	21.43699533	39.19666294						
						101730	21.43623019	39.19652984						
						101728	21.43699533	39.19666294						

Fig. 8. Clustering Results

5.4 Route Optimization

With the delivery points now clustered, the TSP algorithm is applied to optimize routes within each cluster. The TSP seeks to discover minimum total distance through the path of the shortest length that will visit each delivery site in a cluster only once and returns to the starting point. By developing a custom code that applies TSP in every cluster and hence get even better route efficiency, such that the delivery trucks can travel the minimum distance in each assigned cluster as shown in Fig. 9.

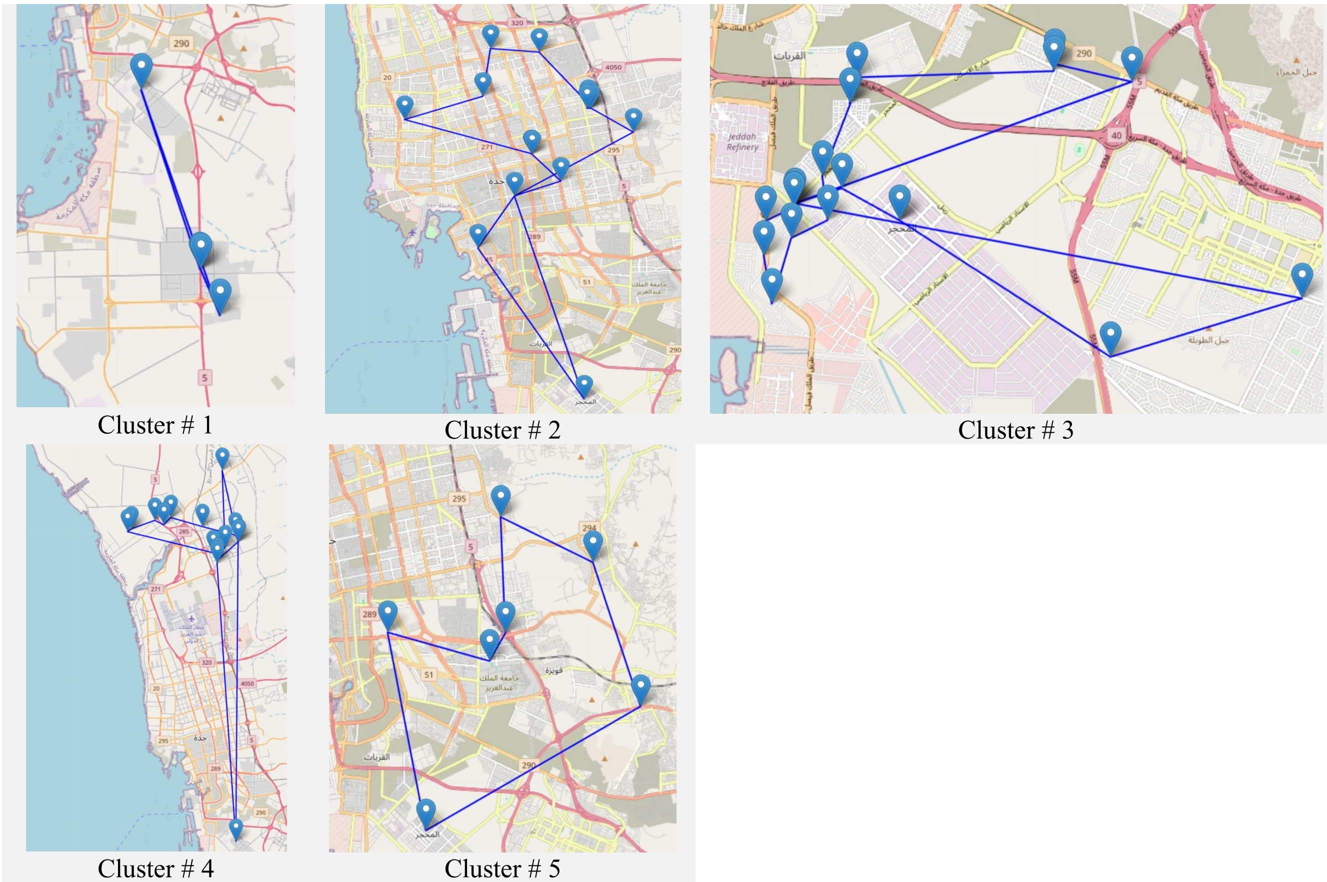


Fig. 9. TSP result for each cluster

5.5 Truck assignment

After that, an assignment of trucks for each cluster are performed based on the orders and truck availability and priority for highest utilization. This step ensures optimum usage of trucks because the delivery loads match truck capacity without overloading or leaving too much free space as shown in Table 1 and Fig. 10 which presented a comparison between different numbers of clusters and selection of the optimum cluster equal 5 to minimize the number of runs to keep operational costs and fuel consumption as low as possible.

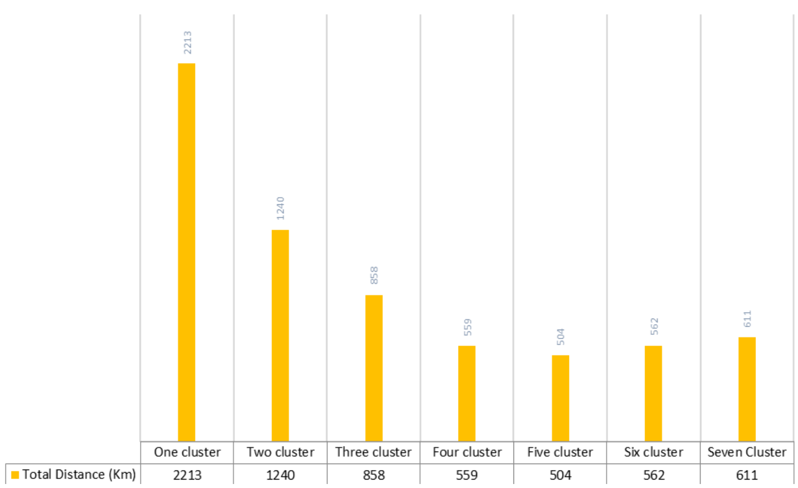


Fig. 10. A comparison between total distances in Km

Table 1
A comparison between different numbers of clusters and total distances in Km

No. of Cluster	Route (Customer IDs for each cluster)	Cluster's Distance (km)	Assigned Trucks	No. of trips	Total Distance
One Cluster	0→25→10→52→6→54→27→46→20→3→45→19→12→51→50→38→48→42→55→9→56→53→44→47→4→17→2→40→32→24→1→11→22→13→28→39→30→5→31→21→29→16→18→37→41→35→14→26→7→23→57→43→34→36→15→49→33→8→0	276.59	8	1	2212.735
Two Cluster	0→33→29→14→30→26→35→32→2→17→15→3→10→27→37→12→21→38→34→36→19→13→5→43→24→18→31→39→20→6→25→41→7→40→42→11→22→23→4→28→1→9→8→16→0	179.09	5	1	1239.967
	0→3→15→8→13→12→10→6→5→4→14→11→9→7→2→1→0	114.82	3	1	
Three Cluster	0→4→2→3→1→0	37.16	1	1	857.87
	0→4→14→3→15→11→8→13→5→6→10→12→9→7→2→1→0	119.12	3	1	
	0→40→14→2→21→15→20→19→4→9→25→1→28→37→39→7→38→36→17→6→22→13→31→35→33→18→10→27→12→26→30→24→34→3→8→16→11→5→29→23→32→0	115.83	4	1	
Four Cluster	0→1→3→2→4→0	37.16	1	1	559.4601
	0→1→8→15→13→5→10→18→16→3→6→9→4→2→17→11→14→12→7→0	79.38	2	1	
	0→19→16→15→3→10→11→8→12→17→2→7→23→18→14→1→9→4→13→21→5→22→20→6→0	49.39	3	1	
	0→1→2→7→9→11→15→3→14→4→8→13→12→5→6→10→0	107.67	2	1	
Five Cluster	0→1→3→2→4→0	37.16	1	1	504.266
	0→10→5→9→11→8→14→2→3→6→13→4→7→15→12→1→0	59.06	2	1	
	0→16→12→15→10→6→9→8→2→5→18→20→4→19→11→3→7→17→14→13→1→0	32.941	2	1	
	0→4→14→3→15→11→9→8→13→5→6→10→12→7→2→1→0	120.18	2	1	
	0→1→5→2→6→4→3→7→0	42.73	1	1	
Six Cluster	0→1→3→2→4→0	37.16	1	1	562.3982
	0→9→11→4→7→13→6→3→2→14→8→5→10→15→12→1→0	70.59	2	1	
	0→15→10→6→9→8→2→20→18→4→19→11→3→7→17→14→5→13→1→12→16→0	38.51	2	2	
	0→6→5→2→3→4→1→0	88.48	1	1	
	0→7→3→4→6→2→5→1→0	42.73	1	1	
	0→1→2→5→7→8→10→3→9→6→4→0	98.77	1	1	
Seven Cluster	0→4→2→3→1→0	37.16	1	1	610.9888
	0→10→5→9→11→8→14→2→3→6→13→4→7→15→12→1→0	59.05	1	2	
	0→9→8→2→5→13→14→19→4→18→20→11→3→7→17→1→6→10→15→12→16→0	36.81	2	2	
	0→6→5→2→3→4→1→0	88.48	1	1	
	0→7→3→4→6→2→5→1→0	42.73	1	1	
	0→1→2→0	92.8	1	1	
	0→5→8→3→9→7→6→4→2→1→0	84.42	1	1	

In conclusion, the Clustering with Real-Time Route Optimization Platform is an important tool in the logistic framework of KINZA Company, which transforms a formerly reactive methodology into a proactive and data-driven one. Real-time distance assessment coupled with dynamic clustering empowers the KINZA Company to make fast and agile routing decisions. Such a platform not only increases delivery efficiency in a wider circle while reducing fuel costs, accommodating more places as the business progresses, but is also scalable to the logistic needs of the KINZA Company.

6. Conclusion

The basic aim of this research is to optimize the logistics of KINZA Company to make an easy routing plan, save fuel, and deliver products in time. In the context of this research, an actually flexible and appropriate solution to problems in logistics that KINZA Company faces will be introduced, modernly driven by data approaches which are going to be based on the use of the Google Maps API Distance Matrix in conjunction with the K-means clustering algorithm. Integrating real-time traffic information into the route-optimization mechanism allowed it to react dynamically to prevailing roadway conditions, hence ensuring routes that get progressively better and increasingly fine-tuned along the course of delivery. Secondly, an implementation of the K-means algorithm was created in order to classify customers according to geographic proximity to a point and thus reduce the TSP to as many smaller, hence more manageable, sub-problems as one wants. The sum of distance traveled by the fleet of trucks would decrease while there would be the assurance that none of the trucks is travelling under or over its full capacity. The incorporation of clustering helps fairly distribute the delivery destinations across trucks, hence the further optimization regarding truck capacity and fuel efficiency. Moreover, the mathematical framework developed in this work enhances the minimization of fuel consumption concerning cargo that each vehicle carries and the distance to be covered. To that end, the distance information obtained from the API gave this model a full-fledged structure, considering both operational costs and the environmental impact of the delivery process in great detail. In other words, it is expected that the methodologies and algorithms that will be put into consideration within this paper will increase efficiency by reducing costs while maintaining customer satisfaction with timely delivery concerning the last-mile operations of KINZA Company. Moreover, it should be underlined that further testing and refining will be carried out in the future for better performance.

Acknowledgement

Authors would like to acknowledge Al Yamamah University in supporting this research.

References

Applegate, D., Bixby, R., Chvátal, V., & Cook, W. (1998). On the Solution of Traveling Salesman Problems. Documenta Mathematica (Extra Volume ICM 1998), 645-656.

Hoffman, K. L., Padberg, M., & Rinaldi, G. (2001). Traveling Salesman Problem: Combinatorial Optimization Approaches. In Kluwer Academic Publishers (pp. 1-40). DOI: 10.1007/1-4020-0611-X_1068.

Huang, Y., Li, J., & Wang, T. (2022). Enhancing K-means clustering performance with a K-means++ initialization approach for logistics applications. *Journal of Business Research, 142*, 429-439.

Johnson, D. S. (1990). Local Optimization and the Traveling Salesman Problem. AT&T Bell Laboratories, Murray Hill, NJ. *SIAM Journal on Computing, 6*(3), 446-478.

Kumar, A., & Singh, R. (2023). A hybrid approach for optimizing truck routes using K-means clustering and genetic algorithms. *Journal of Transportation Management, 35*(2), 29-45.

Lee, J., Kim, S., & Park, C. (2022). Evaluating clustering techniques for logistics optimization: A comparative study. *Journal of Retailing and Consumer Services, 67*, 102899.

Patel, S., & Desai, R. (2023). Dynamic truck assignment using K-means clustering and real-time data analysis. *Transportation Research Part E: Logistics and Transportation Review, 169,* 102892.

Rosenkrantz, D. J., Stearns, R. E., & Lewis, P. M. (1977). An Analysis of Several Heuristics for the Traveling Salesman Problem. *SIAM Journal on Computing,* 6(3), 563-581.

Yadav, V., Srivastava, R., & Sharma, A. (2023). Market segmentation for logistics optimization using K-means clustering. *Journal of Business Research, 150*, 121-135.