Warehouse Selection in Supply Chain Network Design Based on Multi COG Clustering and MCDM

Smit Parekh and Dr. D. N. Raut
Masters of Technology Research Scholar, Production Engineering
Veermata Jijabai Technological Institute, Mumbai, India
smitparekh502@gmail.com, dnraut@pe.vjti.ac.in

Abstract

When it comes to improving logistic operations, the placement of a warehouse is generally one of the most important and strategic considerations. A single warehouse site will not be able to match the requirements of rapid delivery for customers distributed across a large area. As a result, the multi-COG model (multiple warehouse locations) is necessary. A warehouse's location is a long-term choice based on a number of quantitative and qualitative factors. To account for a variety of considerations in warehouse selection, a two-stage strategy is provided. The first part of this work involves doing macro analysis for various warehouse locations using clustering methods, with the number of warehouses determined based on the Within Cluster Sum of Squares (WCSS). In the second stage for phase-by-phase dissertation planning of warehouse are selected based on two Multiple Criteria Decision Making (MCDM) methods, Weighted Aggregated Sum Product Assessment (WASPAS) and Analytical Hierarchical Process (AHP), has been compared on five criteria: Warehouse cost per m^2, Infrastructure access, Number of customers, Delivery time, Quantity of Sales. With over 650 delivery sites, four warehouses were chosen and a two-phase warehouse implementation strategy outlined.

Keywords
Warehouse Location, Multi–COG, WCSS, WASPAS and AHP

1. Introduction

For the financial liquidity and the organization's expenses may not want to have a lot of inventory on hand. However, there would be some gap in terms of logistic management, transportation range and management of travel. Transportation will take longer time and cost more if the distance is greater. To save more, significant inventories are required and a warehouse is required to store inventory. The location of a warehouse is an important consideration when developing a supply chain network for a business. The warehouse placement problem is significant in logistics since it pertains to determining the best location for logistical facilities. To service a large number of dealers, distributors, and customers, a warehouse is required to meet all of their demands. The fundamental purpose of warehouse placement is to be as close to the consumers as feasible, given the customers' needs for quick delivery and the increased competition to attract customers for Same Day delivery. A single warehouse will not be able to provide the major demand for express delivery in a very vast geographical region, hence many warehouses will be necessary to service clients at various places. When a single client is supplied by two warehouses located at very large distance, transportation costs rise because one warehouse is closer to the consumer and the other is farther away. As a result, a suitable network should be built to determine which customers will be supplied from which warehouse, lowering transportation costs and allowing for faster delivery.

Zak and Weglinski (2014) proposed a two-stage technique for determining the most attractive warehouse site. The first step is to conduct a macro-analysis of the geographic regions in order to determine their overall potential and appropriateness for warehouse location. In the second step after macro analysis, micro-analysis is performed for individual warehouse locations in the specified areas. The selection and installation of a warehouse was a lengthy procedure that also need significant amount of capital expenditure. As a result, several criteria should be examined for phase-by-phase deployment of selected warehouses, ensuring that all criteria are taken into consideration.

© IEOM Society International

1.1 Objectives

The primary goal of this study is to demonstrate how to utilize clustering and the MCDM approach to build a warehouse network and to strategically plan for the deployment of numerous warehouse locations while taking into account decision makers' preferences. As data-driven decision-making becomes more common, it may also assist decision-makers with warehouse selection and implementation, as it is based on preliminary regional sales data of their clients.

The first goal is to use data-driven decision making to implement macroscopic analysis. The clustering method determines which customers will be served from which warehouse, and the Within Cluster Sum of Square (WCSS) method determines the number of warehouses required when there are multiple warehouses.

The second goal is to conduct microscopic analysis to establish and strategically plan for the implementation of the warehouse chosen in clustering and WCSS techniques using Multiple Criteria Decision Making (MCDM), which considers the preferences of decision makers.

Both MCDM method - Analytic Hierarchy Process (AHP) - and the hybrid MCDM method - Weighted Aggregated Sum Product Assessment (WASPAS) - are utilized in microscopic analysis. To establish the criterion weights and, more crucially, to evaluate the alternatives, the AHP technique was used. The second technique used was the WASPAS approach, which was used just to assess options while using the AHP method's criteria weights.

2. Literature Review

According to Zeferino et al. (2020), location strategy is described as the process of selecting a region and a specific place for the purpose to create factory, and it is also a crucial aspect in determining business growth. The supply chain's performance is determined by the synchronization of primary drivers such as inventory, transportation connections, and warehouse location. The organizational goal is impacted by long-term agreements with business partners, capital expenditure costs, including the cost of renting warehouse, and the equipment required to keep the warehouse operational.

According to Ravindran (2016), the supply chain is made up of two parts: (1) physical entities that operate the supply chain and (2) coordination between those organizations, such as planning and acquiring items, converting raw materials into finished goods, and distributing to clients. The supply chain may be divided into three categories based on the numerous decisions made. i.e., strategic decisions include developing a whole supply chain network, which typically takes a long time and significantly impacts a company's resources when it comes to locating new warehouses, relocating existing warehouses, or even closing present warehouses. The second category is tactical decision making, which include includes plans for the acquisition of raw materials, production planning, and distribution choices over a one or two-year period. On the other hand, operational decisions are short-term decisions such as dispatch planning, inventory replenishment, and weekly production planning.

According to Petrovi et al. (2018), various optimization algorithms are used in practice to address decision-making difficulties. When making decisions based on comparable options, however, it's critical to evaluate a variety of factors, such as alternatives in the same category and a set of distinct and opposing criteria. Yang and Nguyen (2016) revealed that the proposed clustering method gave better compact solutions of items within clusters, and simulated outcome with a 33 percent increase in retrieval efficiency.

Chen et al. (2014) introduced the TOPSIS-MCGP method to assist airline industry decision-makers logistics center. The weighted centered calculated using TOPSIS and applied to each MCGP aim, taking into considering quantitative and quantitative factors. Khaengkhan (2019) used three MCDM methods to select warehouse locations for grass flowers in Chiang Rai provinces, taking into account factors such as warehouse capacity, warehouse cost, labor cost, transportation availability, and distance from suppliers in the Simple Additive Weighting (SAW) Method, Order Preference by Similarity to Ideal Solution (TOPSIS) Method, and The Analytic Hierarchical Process (AHP) Method. Çetinkaya et al. (2021) assigned the importance level of several criteria using a Geographical Information System (GIS) and subsequently the AHP approach. To gather the data, decision makers and studies from literature are used to develop the selection criteria. The data was plotted using GIS software to evaluate the availability of suitable locations. AHP is then used to priorities the criterion.

Based on a numerous literature, Amin et al. (2019literature a case study utilizing two MCDM approaches, AHP and TOPSIS, to pick one of the five warehouses from the five for the most significant criterion out of twelve. This covers Unit Price, which is the cost per unit to store in the warehouse, Movement Flexibility, which is the ease with which products can be moved in and out of the warehouse, Warehouse Layout, distance from the plant, and Warehouse Capacity.

Dobrota et al. (2015) proposed the fuzzy AHP to the solitary Warehouse placement problem in their model of distribution center location selection model a retail enterprise in Serbia, all criteria and sub criteria are studied and used. sub-criteria include capital expenses, transportation costs, and operating costs, while qualitative criteria measure strategic elements, supply chain, and logistical issues. Experts gave quantitative variables 2/3 of the weight and qualitative considerations 1/3 of the weight, with total outbound cost receiving greater weights in sub-criteria of beneficial effects on firm image receiving the lowest weightage.

3. Methods

In this study, a data-driven approach to k-means clustering, Within Cluster Sum of Square (WCSS) approaches, and Analytic Hierarchy Process (AHP) method and hybrid MCDM method Weighted Aggregated Sum Product Assessment (WASPAS) are used to assess numerous criteria.

Stage 1: Macro Analysis, which will identify the precise number of warehouses necessary based on the sales data provided.

Stage 2: Micro Analysis: Two MCDM techniques will be evaluated in order to prepare for phase-by-phase warehouse deployment.

3.1 Macro Analysis

This methodology consists of 3 steps,
Step 1: Data preprocessing, it is required the data preparation for further analysis. Based on the location of sales data, where in Latitude and Longitude is required.

Step 2: The WCSS approach is used to determine the number of warehouses that are necessary.

The best way to figure out how many k clusters are best is to look at the total distance variation inside each cluster. The WCSS value populated for each number of k clusters and plotted on a line graph with the cluster number on the x axis and the WCSS value on the y axis using the elbow method, also known as the knee of a curve, i.e., the point at which diminishing returns from the warehouse counts is no longer worth the cost spent to set up the warehouse.

WCSS = $\sum_{i=0}^{n} \min_{\mu_j \in C}\left(||x_i - \mu_j||^2 \right)$ where n is the number of points or data to be clustered into k separate clusters C, μ_j is the centroid location of the cluster.

Step 3: K Means Clustering is a partitioning type of clustering approach in which items are divided into k clusters with the goal of placing related items in a group that is distinct from others.

Determine the number of clusters K in the first phase of the clustering process shown below (Figure 1). After finding the value of K, a random selection of cluster centroid commences where the cluster centroid may be the data point. The distance between each data point and the chosen cluster centroid is then measured, and the point with the shortest distance is given its own cluster. Up to the allotted maximum repetition, the procedure is repeated. The cluster's final centroid, also known as the sum of variance centroid, is chosen because it has the lowest WCSS value. Hence the centroid of cluster is the warehouse site will be determined by the centroid of the clusters. The plotted data is then used to identify which data belongs to the cluster. This is informed by looping the processes of first randomly assigning a centroid place in the data, then measuring a Euclidian distance from the all-clusters centroid and assigning that point to the cluster with the shortest distance from the same cluster's centroid. After mapping all points to the cluster, the total of variance of the cluster is calculated and stored until a point with the lowest sum of variation of cluster is selected.

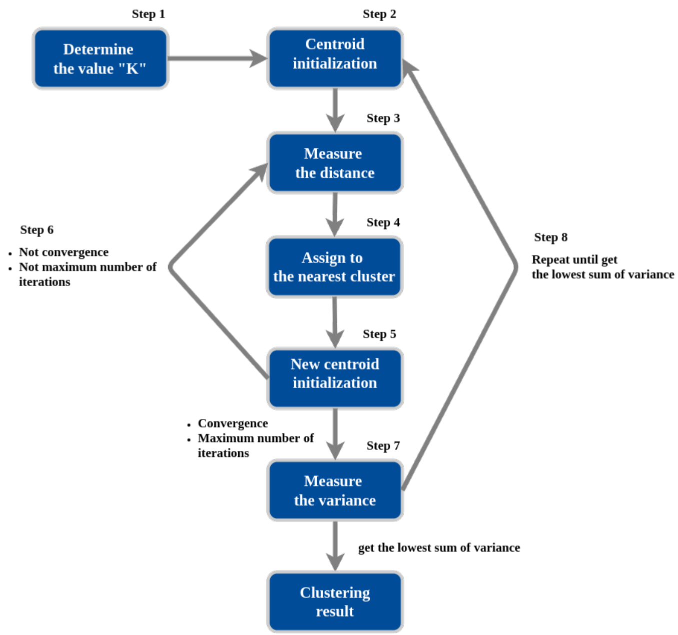

Figure 1. Steps of K-Means Clustering (Step by Step to Understanding K-means Clustering and Implementation with sklearn 2020)

3.2 Micro Analysis

For the stage 2 analysis Analytic Hierarchy Process (AHP) and Weighted Aggregated Sum-Product Assessment (WASPAS), where the weights were calculated using locations which were selected using AHP and WASPAS method.

In Analytic Hierarchy Process (AHP) can be implemented in the following steps:
Step 1: The criteria weights can be calculated by creating a pairwise comparison matrix A where each element is represented as $\boldsymbol{a_{ij}}$ i.e., importance of i^{th} criteria relative to j^{th} criteria. The values of these elements are based on values from 1 to 9 of fundamental Saaty scale. The final determination of weights w_j is based on Geometric mean method as given in the below equation.

$$\boldsymbol{GM_i = \left(\prod_{i=1}^{n} a_{ij}\right)^{1/n}}, \boldsymbol{w_i = \frac{GM_i}{\sum_{i=1}^{n} GM_i}},$$

Where n in the number of criteria and GM_i is is Geometric mean of each criterion.

Step 2: Consistency Index is Calculated to measure the consistency of the subjective comparison.

$$\boldsymbol{CI = \frac{\lambda_{max} - n}{n-1}}$$

Where λ_{max} is the maximum eigenvalue of Matrix A.

Consistency ratio (CR) is calculated using

$$CR = \frac{CI}{RI}$$

Where *R.I* is Random Index (Table 1), and the value of CR should be less than 0.1, If the CR value is more than 0.10, it might be regarded as inconsistency in pairwise comparisons or a computation error.

Table 1. Values of RI depending on number of criteria

Number of Criteria	3	4	5	6	7	8	9	10
RI	0.52	0.89	1.11	1.25	1.35	1.4	1.45	1.49

Step 3: Alternatives are compared in terms of each criterion. This stage entails constructing a pairwise alternative comparison matrix B_j, whose members b_{kl} indicate the preference of the k^{th} alternative relative to the l^{th} alternative based on criteria j. The comparisons must be made using the same values from 1 to 9 on the Saaty scale as indicated in Step 1.

Step 4: Calculate the values by multiplying weights assigned to each criterion and summing up for each location.

For WASPAS Method methodology followed is as follows:
WASPAS method was proposed by Zathe vadskas et al. (2012). Which is a hybrid of two methods i.e., Weighted sum model (WSM) and Weighted Product Model (WPM).

Step 1: Determining performance matrix for each alternative on each criterion in the *X* matrix where x_{ij} is the performance of i^{th} alternative on j^{th} criteria.

Step 2: Normalizing *X* matrix using following formulas, the where $\overline{\boldsymbol{x}}_{\boldsymbol{ij}}$ is normalized value of $\boldsymbol{x_{ij}}$.

$$\overline{\boldsymbol{x}}_{\boldsymbol{ij}} = \frac{x_{ij}}{max_i\, x_{ij}}$$, for maximizing criteria

$$\overline{\boldsymbol{x}}_{\boldsymbol{ij}} = \frac{min_i\ x_{ij}}{x_{ij}}$$, for minimizing criteria

Step 3: Similar to WPM method the total relative importance of i^{th} alternative is is determined by:

$$\boldsymbol{Q_i}^{\boldsymbol{1}} = \prod_{j=1}^{n} \overline{\boldsymbol{x}}_{\boldsymbol{ij}}{}^{w_j}$$

Step 4: Similar to WSM method total relative importance of i^{th} alternative is is determined by:

$$\boldsymbol{Q_i}^{\boldsymbol{2}} = \sum_{j=1}^{n} \overline{\boldsymbol{x}}_{\boldsymbol{ij}}\ .\boldsymbol{w_j}$$

Step 5: The importance of WPM and WSM is given by:

$$\boldsymbol{Q_i} = \frac{Q_i{}^1 + Q_i{}^2}{2}$$

Now the locations can be ranked using values of *Q* i.e., highest rank can be given to largest *Q* value.

4. Data Collection

The data for this article came from an Indian company that needed a supply chain network design and was working on a greenfield project with no existing warehouses. It presently sells in over 650 locations, and buyers will be connected to local warehouses quantity of sales product, and location of sales product were all available based on sales data. The location's latitude and longitude were discovered during data preprocessing and mapped as shown on the map (Figure 2). The cities purple circle markers indicate the cities in sales data and size of the marker indicates the quantity of goods delivered. To reduce map congestion, the size of the marker is only 20% title true size; nevertheless, this has no bearing on the final study; it is just for illustration purposes.

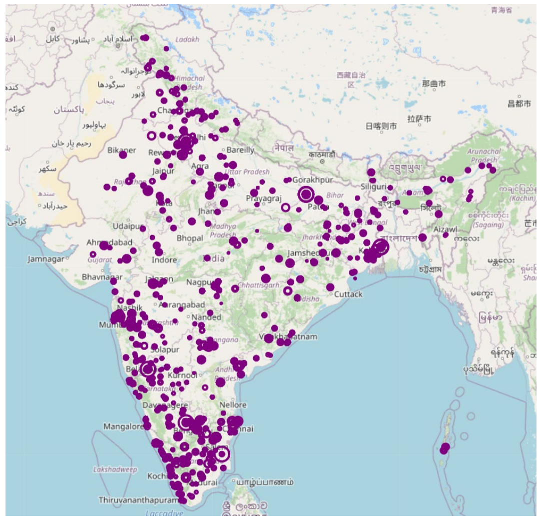

Figure 2. Mapping Sales Data on Map

Five criteria were chosen for the Multi Criteria Decision Making Method based on previous literature surveys and the needs of the firms, including warehouse cost per m^2, infrastructure access, number of customers to be supplied, delivery time, and sales quantity. Following the stage 1 procedure outlined below, four warehouses were chosen, and a Decision Matrix (Table 2) was created. Infrastructure access and delivery time criteria were evaluated by company stakeholders and a group of experts using Saaty's scale of nine numerical values. In addition, actual values for warehouse cost per square foot, number of clients to be served, and quantity of sales were taken. When the criteria were analyzed, the mean values were taken into account.

Infrastructure Access refers to the availability of transportation services, and because transportation is at the heart of logistics distribution, it should have access to a wide range of trucks and other services. For example, depending on variable demands, the truck size required may be reduced, and instead of sending small loads on larger vehicles, a smaller vehicle can be used with complete utilization and lower cost. Because customers anticipate rapid delivery, the shorter the delivery time, the better the service quality, as these factors are dependent on infrastructure access and delivery location. Warehouse cost per square foot is one of the most essential variables as it directly influences sites future developments, and extra available space also aids in warehouse expansions for better customer service and if additional inventory storage is necessary. The number of customers serviced is determines the number of customers serviced from any one warehouse site closer to it. The entire amount to be delivered to customers from the warehouses is known as the quantity of sales.

Table 2. Decision Matrix

Criteria	Warehouse Cost Per m^2	Infrastructure Access	Number of Customers	Delivery Time	Quantity of Sales
Warehouse Location	Min	Max	Max	Min	Max
Coimbatore	161	7.2	92	3.4	1010783
Bidar	140	3.6	113	5.8	1372600
Durgapur	161	4.6	75	5.6	1102966
Gurgaon	129	8.8	79	4	943647

© IEOM Society International

A pairwise comparison matrix was created by the company's stakeholders and a panel of specialists to assess the importance of the given criteria (Table 3).

Table 3. Pairwise Comparison Matrix

Criteria	Warehouse Cost Per m^2	Infrastructure Access	Number of Customers	Delivery Time	Quantity of Sales
Warehouse Cost per m^2	1.000	0.294	3.000	0.250	0.200
Infrastructure Access	3.400	1.000	7.000	0.417	1.500
Number of Customers	0.333	0.143	1.000	0.119	0.333
Delivery Time	4.000	2.400	8.400	1.000	5.200
Quantity of Sales	5.000	0.667	3.000	0.192	1.000

5. Results and Discussion

5.1 Graphical Results

On the x axis (Figure 3), at cluster number 4 is the point known as the knee of the curve, i.e., the point after which the value of the returns diminishes, and it is no longer worthwhile to invest in further warehouses. As a result, number of warehouses necessary for the whole supply chain is four.

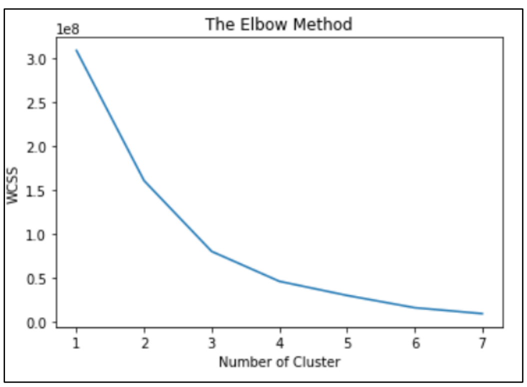

Figure 3. Plotting WCSS vs Number of Cluster

In figure 4 all cluster are mapped with the respective centroids of the cluster. The centroid of the clusters denoted by blue circular marker i.e., warehouse locations, are located in, Bidar, which completely services the Central and Western parts of India, Coimbatore, which completely services the Southern part of India, Durgapur, which completely services the Eastern parts of India, and Gurgaon, which completely services the Northern part of India. Because setting up all four warehouses is a lengthy process, it is best to deploy warehouses in stages so that it does not disrupt the firm's vast resources. In addition, express deliveries to consumers may be observed, resulting in increased sales and overall company competitiveness.

© IEOM Society International

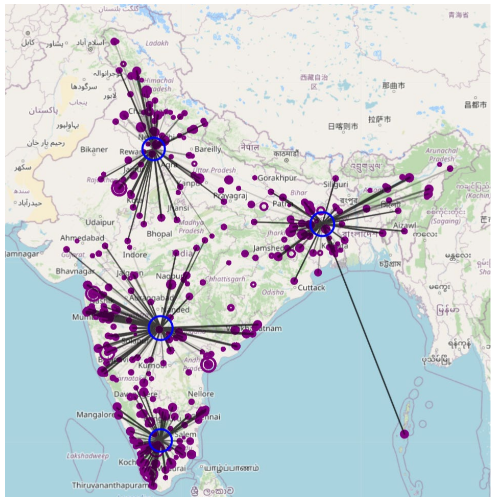

Figure 4. Supply chain Network Map with 4 Warehouse Locations

5.2 Numerical Results

After selected the warehouse location, the selection of warehouse is based on criterions selected for MCDM method i.e., First using AHP method.

Table 4. Criterion weights obtained using AHP MCDM method.

Criterion Weights	Warehouse Cost Per m^2	Infrastructure Access	Number of Customers	Delivery Time	Quantity of Sales
AHP	0.076	0.244	0.041	0.476	0.162

The Table 4, shows us weightages obtained where Delivery Time has obtained higher weightage while selecting the warehouses and Number of Customers to be Served has obtained lowest weightage. Which perfectly corresponds with the core goal of constructing a whole supply chain network, namely, to deliver goods as rapidly as feasible. The Consistency Ratio is calculated and its value in 0.081 which is less than 0.1. For ranking the locations based on AHP method and Hybrid MCDM method (WASPAS) while using weights of criteria derived from AHP method, Table 5 shows the result.

Table 5. Ranking Results of Both MCDM techniques.

Locations	Bidar	Coimbatore	Durgapur	Gurgaon
AHP	0.653	0.890	0.635	0.866
(Rankings)	3	1	4	2
AHP + WASPAS	2.287	2.439	2.286	2.426
(Rankings)	3	1	4	2

Table 5 shows that rankings are the same for both methods, with Coimbatore Warehouse Location obtaining the highest rank using both techniques and thus being the first location to implement. Durgapur, on the other hand, obtained the lowest rankings based on both MCDM techniques and were consequently the last locations to be implemented. Coimbatore and Gurgaon warehouse sites should be established as part of the implementation phase's first phase. Bidar and Durgapur warehouse sites should be implemented in phase 2 implementation. As a consequence, employing the Multi COG model and implementing it in both phases might lead to significant changes in the business

© IEOM Society International

that consumers could observe. This is because the Multi COG model helps to speed up delivery, which could lead to customer growth and increased market competition.

6. Conclusion

In this study, we suggested a two-stage macro and micro analysis technique that incorporates both data-driven judgments and many criteria to evaluate; the core goal of quick delivery may also be reached using this methodology. K Means clustering and WCSS methods aid in mapping which customers should be serviced from which warehouse, resulting in faster delivery and lower transportation costs. In this process, not only sales locations are selected for analysis, but the weights of the quantity of product to be delivered are also attached, allowing warehouse locations to be set closer to the higher density of quantity sold and a shorter distance to be travelled.

The applicability of the MCDM methodology was also demonstrated in the second stage of microanalysis, AHP and WASPAS methods were used for phase-wise implementation of warehouse site selection based on the warehouses identified in stage 1 of macro analysis. This stage has aided decision-makers to take an account for customer preferences in order to accomplish express delivery as soon as feasible when implementing the chosen warehouses.

The further scope of research is possible in selecting the number of warehouses in terms of cost as well i.e., as supply chain consist of Primary transportation cost which is the cost needed to transport goods to warehouses, warehousing cost which is the, cost needed to run warehouse, a storing cost and capital expenditure needed to start warehouses and secondary transportation cost which is the cost to deliver to the customer.

References

Al Amin, M., Das, A., Roy, S. and Shikdar, M.I., Warehouse selection problem solution by using proper mcdm process, *International Journal of Science and Qualitative Analysis*, vol. 5, no. 2, pp.43-51, 2019.

Çetinkaya, C., Özceylan, E. and Keser, İ., A GIS-Based AHP Approach for Emergency Warehouse Site Selection: A Case Close to Turkey-Syria Border. *Journal of Engineering Research*, 2021.

Chen, K.H., Liao, C.N. and Wu, L.C., A selection model to logistic centers based on TOPSIS and MCGP methods: The case of airline industry, *Journal of Applied Mathematics*, 2014.

Dobrota, M., Macura, D. and Šelmi, M., May. Multi criteria decision making for distribution center location selection-Serbia case study, *2nd Logistics International Conference*, pp. 32-37, Belgrade, Serbia, 21-23, 2015.

Khaengkhan, M., 2019, July. Using Multiple Criteria Decision Making (Mcdm) To Location Selection For Grass Flower Warehouse In Chiang Rai Province, Thailand, *International Academic Multidisciplinary Research Conference*, pp. 133-133, 2019.

Petrović G., Sekulić V., Madić M., Mihajlović J., A study of multi-criteria decision-making for selecting suppliers of linear motion guide, Facta Universitatis, vol. 14, no. 1, pp. 1–14, 2018.

Ravindran, A., *Multiple Criteria Decision Making in Supply Chain Management*, CRC Press, 2016.

Step by Step to Understanding K-means Clustering and Implementation with sklearn, Available: https://medium.com/data-folks-indonesia/step-by-step-to-understanding-k-means-clustering-and-implementation-with-sklearn-b55803f519d6, 2020.

Yang, C.L. and Nguyen, T.P.Q., Constrained clustering method for class-based storage location assignment in warehouse, *Industrial Management & Data Systems*, vol. 116, no. 4, pp. 667-689, 2016.

Żak, J. and Węgliński, S., The selection of the logistics center location based on MCDM/A methodology. *Transportation Research Procedia*, vol 3, pp.555-564, 2014.

Zeferino, E.F.S., Mpofu, K., Makinde, O.A., Ramatsetse, B.I. and Daniyan, I.A., FR/CoG multi-attribute-based comparison methods for selection of the location of a research institute. *Journal of Facilities Management*, vol. 18, no. 1, pp. 20-35, 2020.

Biographies

Smit Parekh is presently Masters of Technology Research Scholar in Production Engineering from Veermata Jijabai Technological Institute, Mumbai, India.

Dr. D. N. Raut a is Professor and Head of Production Engineering Department in Veermata Jijabai Technological Institute, Mumbai, India. He has more than 24 years' experience in Teaching, R & D and has published 32 papers in international journals and 22 papers in National journals.

© IEOM Society International