International Journal Of Health, Engineering And Technology (IJHET) *E-ISSN 2829 - 8683*
Volume 1, Number 2, July 2022, *Page. 124 - 134*
Email : editorijhess@gmail.com

Data Mining Motor Vehicle Testing Based On Vehicle Type Using The K-Means Method Case Study Binjai City Transportation Service

Aldo Kristian[1)], Budi Serasi Ginting[2)], Suci Ramadani[3)]
[1,2,3)] Information Systems Study Program, STMIK Kaputama Binjai, Binjai, Indonesia

*Corresponden Author :
Email: aldokristian3@gmail.com

Abstract

Motor vehicle testing is a series of activities to check the components in the vehicle. Motor vehicle testing is very important, because vehicles operated on the road have the potential to cause accidents. So if periodic inspections are not carried out, it cannot provide technical safety to vehicle users, because it is not known what components are lacking and what components must be repaired. In this study, motor vehicle test data will be calculated using the K-Means. The K-means algorithm is aclusteralgorithm non-hierarchical. Cluster analysis is a tool for grouping data based on variables or features. The purpose of k-means clustering, like other clustering methods, is to obtain clusters of data by maximizing the similarity of characteristics within the cluster and maximizing the differences between clusters.groups K-means clustering algorithm data based on the distance between the data and the centroid cluster Cluster with the number of motorized vehicle data based on the type of vehicle as many as 4 vehicles, namely, freight cars, MPU, buses, and betor. Cluster 1 there are 7 groups with 7 types of vehicles: 2 BUS and 5 betor where there is one type of vehicle (betor) that does not pass the test due to the type of damage at the time of testing motor vehicle. Cluster 1 is the type of vehicle that passes the motor vehicle test the most with the lowest level of damage; Cluster 2 there are 4 groups by type of vehicle: 4 Cars of Freight where 2 of them did not pass the test because of the type of damage during the motor vehicle test; Cluster 3 has 9 groups with the types of vehicles: 2 freight cars, 4 MPUs, and 3 BUS. 3 BUS and 1 MPU did not pass the test due to damage during the motor vehicle test. cluster is the cluster of vehicle types that do not pass the test with the most types of damage.

Keywords: Motor Vehicle Testing; Data Mining; K-Means

INTRODUCTION

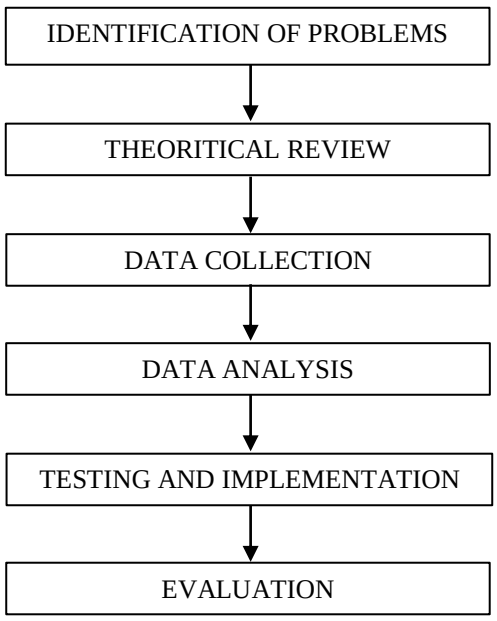
Motor vehicle testing is a series of activities for the examination of the components in the vehicle. Motor vehicle testing is very important, because vehicles operated on the road have the potential to cause accidents. So if periodic inspections are not carried out, it cannot provide technical safety to vehicle users, because it is not known what components are lacking and what components must be repaired. With the testing of motorized vehicles can provide technical safety to motorized vehicle users (Mara, 2013). This motor vehicle testing must be carried out periodically in accordance with Law No. 22 of 2009 concerning road traffic and transportation, namely regarding the transportation of Freight and public passenger cars. Testing of motorized vehicles is carried out visually, while for roadworthiness using tools, all of which have a threshold that must be met. After testing the motorized vehicle, the relevant party shall provide a certificate for repairing the motor vehicle (Nabila, 2021).

In this study, the author wants to apply data mining to motorized vehicle testing data at the Binjai City Transportation Service using the K-Means algorithm. Data mining is a data collection technique with the aim of analyzing large amounts of data. The K-Means algorithm is one of the clustering algorithms that is included in the unsupervised learning group which is used to group data into several groups with a partition system (Nofitri, 2019). Vehicles will be grouped according to the characteristics of each vehicle. To perform the grouping process, there is an algorithm that is often used because it is relatively fast and easy to adapt, namely the K-Means algorithm. K-Means is a non-hierarchical data clustering method that groups one or more data clusters. Data that has the same characteristics will be grouped in one cluster and data that has different characteristics are grouped with other clusters, so that data in one cluster has a small level of variation (Pulungan, 2019).

RESEARCH METHODS

Methodology Research

Methodology is carried out to search for something systematically by using scientific methods and applicable sources. Methodology is also an analysis of a method or method. In conducting this research, the writer follows the methodological stages in this study, namely as follows:



There are several stages of the research methodology carried out in solving problems. These stages are as follows:

1. Identifications of Problems
   stage is the initial stage used to identify problems with the aim of observing and looking for problems that are being faced in the object of research.
2. Theoretical Review
   collection of theories related to the subject matter such as the theory of the definition of analysis, the theory of motor vehicle testing and the methods used and the design application of the required system. In this stage, theory is collected from several sources such as books, journals, articles and other references.
3. Data Collection
   At this stage the researcher collects data as a basic material in testing motorized vehicles found in the data. The data used comes from the Department of Transportation of the City of Binjai.
4. Data Analysis
   At this stage the researcher analyzes the data used in the motor vehicle testing process, with existing guidelines on supporting theories from books and journals related to the subject matter.
5. Testing and Implementation
   At the stage of testing and implementing in motor vehicle testing, the method used is the grouping method or K-Means. After determining the method and testing the design system that has been made and coding according to the programming language used to create the system.
6. Evaluation
   At this stage, it contains the results of experiments carried out to improve the system to achieve a good system. in the evaluation to determine whether the system is feasible or not.

Algorithm K-Means

The K-means algorithm is a non-hierarchical cluster analysis algorithm. Cluster analysis is a tool for grouping data based on variables or features.uses *K-Means Cluster Analysis algorithm* the distance calculation method (distance) to measure the level of proximity between objects and the center point (centroid).is *K-Means algorithm* not affected by the order of objects used, this is proven when the author tries to randomly determine the starting point of the *cluster* of one of the objects at the beginning of the calculation. *K-Means* groups data based on the similarity/closeness of a data to the centroid or center of each data group. This algorithm has a level of complexity that is quite efficient and easy to understand so that this algorithm is often used in various cases.

.

RESULTS AND DISCUSSION

Data

Analysis In the analysis of testing the data mining method used in determining the grouping of motorized vehicle tests based on the type of vehicle using the K-Means method, the following data is used as an analysis of testing the method:

Tabel 1. Research Data

No.	Type of Vehicle	Test Results	Type of Damage
1	Betor	Passed	Stable Steering Performance
2	Freight Cars	Do Not Pass	Oil Leakage
3	BUS	Does Not Pass	Hard Brake When Stepped On
4	Freight Cars	Pass	Stable Steering Performance
5	MPU	Pass	Stable Steering Performance
6	Freight Cars	Pass	Smooth Engine Sound
7	MPU	Pass	Stable Steering Performance
8	BUS	Does Not Pass	Brakes Loud when stepped on
9	BUS	Pass	Oil Leakage
10	Betor	Passed	Smooth Engine Sound
11	BUS	Passed	Stable Steering Performance
12	Freight Carss	Passed	Stable Steering Performance
13	BUS	Passed	Smooth Engine Sound
14	Betor	Passed	Hard Brake When stepped on
15	MPU	Passed	Steering Stable Performance
16	Freight Cars	Does Not Pass	Leakage Of
17	MPU	Does Not Pass	Hard Brake When Stepped on
18	Betor	Passed	Smooth Engine Sound
19	Betor	Passed	Smooth Engine Sound
20	Freight Cars	Passed	Smooth Engine Sound

Application of K-Means Method

Grouping motorized vehicle testing based on vehicle type using *the K-Means algorithm* :

1. From the sample research data above, data transformation can be carried out to make it easier to calculate clustering results. The following is a data table after the transformation.

Type of Vehicle		Test Result		Type of Damage	
Freight Cars	: 1	Passed	: 1	Smooth engine sound	: 1
MPU	: 2	Did not pass	: 2	Steering performance is stable	: 2

BUS	: 3	Brakes hard when stepped on	: 3
Betor	: 4	Leakage of oil compartment	: 4

Tabel 2. Data Transformation Table

Attribute 1 (X) Type of Vehicle	Attribute 2 (Y) Test Result	Attribute 3 (Z) Type of Damage
4	1	2
1	2	4
3	2	3
1	1	2
2	1	2
1	1	1
2	1	2
3	2	3
3	2	4
4	1	1
3	1	2
1	1	2
3	1	1
4	2	3
2	1	2
1	2	4
2	2	3
4	1	1
4	1	1
1	1	1

2. After the data transformation is done, proceed to the process of determining the number of groups where the number of groups determined is 3 groups .
3. To calculate the first iteration, determine the centroid at random. The following centroid values are determined randomly.

K = 3 → Input from user → 3 Centroid
Centroid 1 (m1) = (4,1,2)
Centroid 2 (m2) = (1,2,4)
Centroid 3 (m3) = (3,2,3)

4. Counting distance of object to centroid with Euclidean Distance formula:

$$d_{Euclidean}(X,Y,Z) = \sqrt{\sum_i \quad (X_i - Y_i - Z_i)^2}$$

Calculation of object distance to centroid with 1st data.

D(4,1,2) distance from m1 = $\sqrt{(4-4)^2 + (1-1)^2 + (2-2)^2 = 0}$
D(4,1,2) distance from m2 = $\sqrt{(4-1)^2 + (1-2)^2 + (2-4)^2 = 3.74}$
D(4,1,2) distance from m3 = $\sqrt{(4-3)^2 + (1-2)^2 + (2-3)^2 = 1.73}$
Calculation of object distance to centroid with 2nd data.
D(1,2,4) distance from m1 = $\sqrt{(1-4)^2 + (2-1)^2 + (4-2)^2 = 3.74}$
D(1,2,4) distance from m2 = $\sqrt{(1-1)^2 + (2-2)^2 + (4-4)^2 = 0}$
D(1,2,4) distance from m3 = $\sqrt{(1-3)^2 + (2-2)^2 + (4-3)^2 = 2,23}$

Calculation of object distance to centroid with 3rd data.
D(3,2,3) distance from m1 = $\sqrt{(3-4)^2+(2-1)^2+(3-2)^2} = 1{,}73$
D(3 ,2,3) distance from m2 = $\sqrt{(3-1)^2+(2-2)^2+(3-4)^2} =$
2,23 D(3,2,3) distance from m3 = $\sqrt{(3-3)^2+(2-2)^2+(3-3)^2} = 0$
Calculation of object distance to centroid with 4th data.
D(1,1,2) distance from m1 = $\sqrt{(1-4)^2+(1-1)^2+(2-2)^2} = 3$
D(1,1,2) distance from m2 = $\sqrt{(1-1)^2+(1-2)^2+(2-4)^2} = 2{,}23$
D(1,1,2) distance from m3 = $\sqrt{(1-3)^2+(1-2)^2+(2-3)^2} = 2{,}44$
Calculation of object distance to centroid with 5th data.
D(2,1,2) distance from m1 = $\sqrt{(2-4)^2+(1-1)^2+(2-2)^2} = 2$
D(2,1,2) distance from m2 = $\sqrt{(2-1)^2+(1-2)^2+(2-4)^2} = 2.44$
D(2,1,2) distance from m3 = $\sqrt{(2-3)^2+(1-2)^2+(2-3)^2} = 1.73$
Calculation of object distance to centroid with 6th data.
D(1,1,1) distance from m1 = $\sqrt{(1-4)^2+(1-1)^2+(1-2)^2} = 3.16$
D(1,1,1) distance from m2 = $\sqrt{(1-1)^2+(1-2)^2+(1-4)^2} = 3.16$
D(1,1,1,) distance from m3 = $\sqrt{(1-3)^2+(1-2)^2+(1-3)^2} = 3$
Calculation of object distance to centroid with 7th data.
D(2,1,2) distance from m1 = $\sqrt{(2-4)^2+(1-1)^2+(2-2)^2} = 2$
D(2,1,2) distance from m2 = $\sqrt{(2-1)^2+(1-2)^2+(2-4)^2} = 2{,}44$
D(2,1,2) distance from m3 = $\sqrt{(2-3)^2+(1-2)^2+(2-3)^2} = 1{,}73$
Calculation of object distance to centroid with 8th data.
D(3,2,3) distance from m1 = $\sqrt{(3-4)^2+(2-1)^2+(3-2)^2} = 1.73$
D(3,2,3) distance from m2 = $\sqrt{(3-1)^2+(2-2)^2+(3-4)^2} = 2.23$
D(3,2,3) distance from m3 = $\sqrt{(3-3)^2+(2-2)^2+(3-3)^2} = 0$
Calculation of object distance to centroid with 9th data.
D(3,2,4) distance from m1 = $\sqrt{(3-4)^2+(2-1)^2+(4-2)^2} = 2.44$
D(3,2,4) distance from m2 = $\sqrt{(3-1)^2+(2-2)^2+(4-4)^2} = 2$
D(3,2,4) distance from m3 = $\sqrt{(3-3)^2+(2-2)^2+(4-3)^2} = 1$
Calculation of object distance to centroid with 10th data.
D(4 ,1,1) distance from m1 = $\sqrt{(4-4)^2+(1-1)^2+(1-2)^2} = 1$
D(4,1,1) distance from m2 = $\sqrt{(4-1)^2+(1-2)^2+(1-4)^2} = 4.35$
D(4,1,1,) distance from m3 = $\sqrt{(4-3)^2+(1-2)^2+(1-3)^2} = 2.44$
Calculation of object distance to centroid with 11th data.
D(3,1,2) distance from m1 = $\sqrt{(3-4)^2+(1-1)^2+(2-2)^2} = 1$
D(3,1,2) distance from m2 = $\sqrt{(3-1)^2+(1-2)^2+(2-4)^2} = 3$
D(3,1,2) distance from m3 = $\sqrt{(3-3)^2+(1-2)^2+(2-3)^2} = 1.41$
Calculation of object distance to centroid with 12th data.
D(1,1,2) distance from m1 = $\sqrt{(1-4)^2+(1-1)^2+(2-2)^2} = 3$
D(1,1,2) distance from m2 = $\sqrt{(1-1)^2+(1-2)^2+(2-4)^2} = 2.23$
D(1,1,2) distance from m3 = $\sqrt{(1-3)^2+(1-2)^2+(2-3)^2} = 2{,}44$
Calculation of object distance to centroid with 13th data.
D(3,1, 1) distance from m1 = $\sqrt{(3-4)^2+(1-1)^2+(1-2)^2} = 1{,}41$

D(3,1,1) distance from m2 = $\sqrt{(3-1)^2 + (1-2)^2 + (1-4)^2} = 3.74$
D(3,1,1,) distance from m3 = $\sqrt{(3-3)^2 + (1-2)^2 + (1-3)^2} = 2.23$
Calculation of object distance to centroid with 14th data.
D(4,2,3) distance from m1 = $\sqrt{(4-4)^2 + (2-1)^2 + (3-2)^2} = 1.41$
D (4,2,3) distance from m2 = $\sqrt{(4-1)^2 + (2-2)^2 + (3-4)^2} = 3{,}16$
D(4,2,3) distance from m3 = $\sqrt{(4-3)^2 + (2-2)^2 + (3-3)^2} = 1$
Calculation of object distance to centroid with 15th data.
D(2,1,2) distance from m1 = $\sqrt{(2-4)^2 + (1-1)^2 + (2-2)^2} = 2$
D(2,1,2) distance from m2 = $\sqrt{(2-1)^2 + (1-2)^2 + (2-4)^2} = 2.44$
D (2,1,2) distance from m3 = $\sqrt{(2-3)^2 + (1-2)^2 + (2-3)^2} = 1{,}73$
Calculation of object distance to centroid with 16th data.
D(1,2 ,4) distance from m1 = $\sqrt{(1-4)^2 + (2-1)^2 + (4-2)^2} = 3.74$
D(1,2,4) distance from m2 = $\sqrt{(1-1)^2 + (2-2)^2 + (4-4)^2} = 0$
D(1,2,4) distance from m3 = $\sqrt{(1-3)^2 + (2-2)^2 + (4-3)^2} = 2,23$
Calculation of object distance to centroid with 17th data.
D(2,2,3) distance from m1 = $\sqrt{(2-4)^2 + (2-1)^2 + (3-2)^2} = 2{,}44$
D(2 ,2,3) distance from m2 = $\sqrt{(2-1)^2 + (2-2)^2 + (3-4)^2} = 1.41$
D(2,2,3) distance from m3 = $\sqrt{(2-3)^2 + (2-2)^2 + (3-3)^2} = 1$
Calculation of object distance to centroid with 18th data.
D(4,1,1) distance from m1 = $\sqrt{(4-4)^2 + (1-1)^2 + (1-2)^2} = 1$
D(4 ,1,1) distance from m2 = $\sqrt{(4-1)^2 + (1-2)^2 + (1-4)^2} = 4{,}35$
D(4,1,1,) distance from m3 = $\sqrt{(4-3)^2 + (1-2)^2 + (1-3)^2} = 2.44$
Calculation of object distance to centroid with 19th data
D(4,1,1) distance from m1 = $\sqrt{(4-4)^2 + (1-1)^2 + (1-2)^2} = 1$
D(4,1,1) distance from m2 = $\sqrt{(4-1)^2 + (1-2)^2 + (1-4)^2} = 4{,}35$
D(4,1,1,) distance from m3 = $\sqrt{(4-3)^2 + (1-2)^2 + (1-3)^2} = 2{,}44$
Calculation of object distance to centroid with 20th data.
D(1 ,1,1) distance from m1 = $\sqrt{(1-4)^2 + (1-1)^2 + (1-2)^2} = 3.16$
D(1,1,1) distance from m2 = $\sqrt{(1-1)^2 + (1-2)^2 + (1-4)^2} = 3.16$
D(1,1,1,) distance from m3 = $\sqrt{(1-3)^2 + (1-2)^2 + (1-3)^2} = 3$

5. After getting the result of the object's distance to the centroid, then group objects based on the closest distance. The following table is the distance of the object to the centroid in iteration-1.

Tabel 3. Distance from Object to Centroid in Iteration-1

X	Y	Z	Distance from m1	Distance from m2	Distance from m3
4	1	2	0	3.74	1.73
1	2	4	3.74	0	2.23
3	2	3	1.73	2.23	0
1	1	2	3	2,23	2,44
2	1	2	2	2,44	1,73
1	1	1	3,16	3,16	3
2	1	2	2	2,44	1,73
3	2	3	1,73	2,23	0
3	2	4	2,44	2	1

4	1	1	1	4,35	2,44
3	1	2	1	3	1,41
1	1	2	3	2,23	2,44
3	1	1	1,41	3,74	2,23
4	2	3	1.41	3.16	1
2	1	2	2	2.44	1.73
1	2	4	3.74	0	2.23
2	2	3	2.44	1.41	1
4	1	1	1	4.35	2.44
4	1	1	1	4,35	2,44
1	1	1	3,16	3,16	3

6. Groups based on the distance *centroid* nearestK = 3, *centroid* 1 (m1) = (4,1,2), *centroid* 2 (m2) = (1,2,4), *centroid* 3 (m3) = (3,2,3). If m1 < m2 < m3 = group 1, m2 < m1 < m3 = group 2, m3 < m1 < m2 = group 3. Following are the results of grouping in iteration-1.

Tabel 4. Results of Grouping in the 1st Iteration

X	Y	Z	Distance from m1	Distance from m2	Distance from m3	Group
4	1	2	0	3.74	1.73	1
1	2	4	3.74	0	2.23	2
3	2	3	1.73	2, 23	0	3
1	1	2	3	2,23	2,44	2
2	1	2	2	2,44	1,73	3
1	1	1	3,16	3,16	3	3
2	1	2	2	2,44	1,73	3
3	2	3	1.73	2.23	0	3
3	2	4	2.44	2	1	3
4	1	1	1	4.35	2.44	1
3	1	2	1	3	1.41	1
1	1	2	3	2.23	2.44	2
3	1	1	1.41	3.74	2.23	1
4	2	3	1.41	3.16	1	1
2	1	2	2	2.44	1.73	3
1	2	4	3.74	0	2.23	2
2	2	3	2.44	1.41	1	3
4	1	1	1	4,35	2,44	1
4	1	1	1	4,35	2,44	1
1	1	1	3,16	3,16	3	3

Groups based on the minimum distance to the *centroid* nearest
Old group = { 0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0}
New group = {1,2,3, 2,3,3,3,3,3,1,1,2,1,1,3,2,3,1,1,3}
A group change occurs, then proceed to the next iteration.

7. Calculates centroid using the result of the group in each clusterr.calculation centroid is:

$$\text{Centroid 1 (m1)} = \frac{((4+4+3+3+4+4+4)}{7}, \frac{(1+1+1+1+2+1+1)}{7}, \frac{(2+1+2+1+3+1+1))}{7}$$

$$= (3,71), (1,14), (1,57)$$

$$\text{Centroid 2 (m2)} = \frac{((1+1+1+1)}{4}, \frac{(2+1+1+2)}{4}, \frac{(4+2+2+4))}{4} = (1), (1,50), (3)$$

$$\text{Centroid 3 (m3)} = \frac{((3+2+1+2+3+3+2+2+1)}{9}, \frac{(2+1+1+1+2+2+1+2+1)}{9}, \frac{(3+2+1+2+3+4+2+3+1))}{9}$$

= (2,11), (1,44), (2,33)

8. Continue iteration-2 by doing steps 4 to 6. The following is the calculation of the object distance to the *centroid* using the formula.

Calculation of object distance to *centroid* with 1st data.

D(4,1,2) distance from m1 = $\sqrt{(4-3.71)^2+(1-1.14)^2+(2-1.57)^2} = 0.53$

D(4,1,2) distance from m2 = $\sqrt{(4-1)^2+(1-1.50)^2+(2-3)^2} = 3.20$

D(4,1,2) distance from m3 = $\sqrt{(4-2.11)^2+(1-1.44)^2+(2-2,33)^2} = 1.96$

Calculation of object distance to centroid with 2nd data.

D(1,2,4) distance from m1 = $\sqrt{(1-3.71)^2+(2-1.14)^2+(4-1.57)^2} = 3.74$

D(1,2,4) distance from m2 = $\sqrt{(1-1)^2+(2-1.50)^2+(4-3)^2} = 1.11$

D(1,2,4) distance from m3 = $\sqrt{(1-2.11)^2+(2-1,44)^2+(4-2,33)^2} = 2.08$

Calculation of object distance to centroid with 3rd data.

D(3,2,3) distance from m1 = $\sqrt{(3-3,71)^2+(2-1,14)^2+(3-1.57)^2} = 1.81$

D(3,2,3) distance from m2 = $\sqrt{(3-1)^2+(2-1.50)^2+(3-3)^2} = 2,06$

D(3,2,3) distance from m3 = $\sqrt{(3-2,11)^2+(2-1,44)^2+(3-2,33)^2} =$

Calculation of object distance to centroid with 4th data.

D(1,1,2) distance from m1 = $\sqrt{(1-3.71)^2+(1-1.14)^2+(2-1.57)^2} = 2.87$

D(1,1,2) distance from m2 = $\sqrt{(1-1)^2+(1-1.50)^2+(2-3)^2} = 1.11$

D(1,1,2) distance from m3 = $\sqrt{(1-2.11)^2+(1-1.44)^2+(2-2,33)^2} = 1,28$

Calculation of object distance to centroid with 5th data.

D(2,1,2) distance from m1 = $\sqrt{(2-3,71)^2+(1-1,14)^2+(2-1,57)^2} = 1.76$

D(2,1,2) distance from m2 = $\sqrt{(2-1)^2+(1-1.50)^2+(2-3)^2} = 1.50$

D(2,1,2) distance from m3 = $\sqrt{(2-2.11)^2+(1-1.44)^2+(2-2.33)^2} = 0.56$

Calculation of object distance to centroid with 6th data.

D(1,1,1) distance from m1 = $\sqrt{(1-3.71)^2+(1-1.14)^2+(1-1.57)^2} = 2.77$

D(1,1,1) distance from m2 = $\sqrt{(1-1)^2+(1-1.50)^2+(1-3)^2} = 2.06$

D(1.1,1) distance from m3 = $\sqrt{(1-2.11)^2+(1-1.44)^2+(1-2,33)^2} = 1.78$

Calculation of object distance to centroid with 7th data.

D(2,1,2) distance from m1 = $\sqrt{(2-3.71)^2+(1-1.14)^2+(2-1.57)^2} = 1.76$

D(2,1,2) distance from m2 = $\sqrt{(2-1)^2+(1-1.50)^2+(2-3)^2} = 1.50$

D(2, 1,2) distance from m3 = $\sqrt{(2-2,11)^2+(1-1,44)^2+(2-2,33)^2} = 0,56$

Calculation of object distance to centroid with 8th data.

D(3,2, 3) distance from m1 = $\sqrt{(3-3.71)^2+(2-1.14)^2+(3-1.57)^2} = 1.81$

D(3,2,3) distance from m2 = $\sqrt{(3-1)^2+(2-1.50)^2+(3-3)^2} = 2.06$

D(3,2,3) distance from m3 = $\sqrt{(3-2.11)^2+(2-1.44)^2+(3-2.33)^2} = 1.24$

Calculation of object distance to centroid with 9th data.

D(3,2,4) distance from m1 = $\sqrt{(3-3.71)^2+(2-1.14)^2+(4-1.57)^2} = 2,67$

D(3,2,4) distance from m2 = $\sqrt{(3-1)^2+(2-1.50)^2+(4-3)^2} = 2.29$

D(3,2,4) distance from m3 = $\sqrt{(3-2,11)^2+(2-1,44)^2+(4-2,33)^2} = 1,97$

Calculation of object distance to centroid with 10th data.

D(4,1,1) distance from m1 = $\sqrt{(4-3,71)^2+(1-1.14)^2+(1-1.57)^2} = 2.93$

D(4,1.1) distance from m2 = $\sqrt{(4-1)^2+(1-1.50)^2+(1-3)^2} = 4.38$

D(4,1,1) distance from m3 = $\sqrt{(4-2.11)^2+(1-1.44)^2+(1-2{,}33)^2} = 3.44$
Calculation of object distance to centroid with 11th data.
D(3,1,2) distance from m1 = $\sqrt{(3-3.71)^2+(1-1.14)^2+(2-1.57)^2} = 0.84$
D(3,1,2) distance from m2 = $\sqrt{(3-1)^2+(1-1.50)^2+(2-3)^2} = 2.29$
D(3,1,2) distance from m3 = $\sqrt{(3-2.11)^2+(1-1.44)^2+(2-2{,}33)^2} = 1.04$
Calculation of object distance to centroid with 12th data.
D(1,1,2) distance from m1 = $\sqrt{(1-3.71)^2+(1-1.14)^2+(2-1.57)^2} = 2.87$
D(1,1,2) distance from m2 = $\sqrt{(1-1)^2+(1-1.50)^2+(2-3)^2} = 1.11$
D(1,1, 2) distance from m3 = $\sqrt{(1-2.11)^2+(1-1.44)^2+(2-2.33)^2} = 1.28$
Calculation of object distance to centroid with 13th data.
D(3,1,1) distance from m1 = $\sqrt{(3-3.71)^2+(1-1.14)^2+(1-1.57)^2} = 0.92$
D(3,1.1) distance from m2 = $\sqrt{(3-1)^2+(1-1{,}50)^2+(1-3)^2} = 2.87$
D(3,1,1) distance from m3 = $\sqrt{(3-2.11)^2+(1-1.44)^2+(1-2{,}33)^2} = 1{,}65$
Calculation of object distance to centroid with 14th data.
D(4,2,3) distance from m1 = $\sqrt{(4-3.71)^2+(2-1.14)^2+(3-1.57)^2} = 1.69$
D(4,2,3) j wine from m2 = $\sqrt{(4-1)^2+(2-1.50)^2+(3-3)^2} = 3.04$
D(4,2,3) distance from m3 = $\sqrt{(4-2.11)^2+(2-1.44)^2+(3-2{,}33)^2} = 2.08$
Calculation of object distance to centroid with 15th data.
D(2,1,2) distance from m1 = $\sqrt{(2-3{,}71)^2+(1-1{,}14)^2+(2-1.57)^2} = 1.76$
D(2,1,2) distance from m2 = $\sqrt{(2-1)^2+(1-1.50)^2+(2-3)^2} = 1.50$
D(2, 1,2) distance from m3 = $\sqrt{(2-2{,}11)^2+(1-1{,}44)^2+(2-2{,}33)^2} = 0{,}56$
Calculation of object distance to centroid with 16th data.
D(1,2, 4) distance from m1 = $\sqrt{(1-3.71)^2+(2-1.14)^2+(4-1.57)^2} = 3.74$
D(1,2,4) distance from m2 = $\sqrt{(1-1)^2+(2-1.50)^2+(4-3)^2} = 1.11$
D(1,2,4) distance from m3 = $\sqrt{(1-2.11)^2+(2-1.44)^2+(4-2{,}33)^2} = 2.08$
Calculation of object distance to centroid with 17th data.
D(2,2,3) distance from m1 = $\sqrt{(2-3.71)^2+(2-1.14)^2+(3-1.57)^2} = 2{,}38$
D(2,2,3) distance from m2 = $\sqrt{(2-1)^2+(2-1.50)^2+(3-3)^2} = 1.11$
D(2,2,3) distance from m3 = $\sqrt{(2-2{,}11)^2+(2-1{,}44)^2+(3-2{,}33)^2} = 0.88$
Calculation of object distance to centroid with 18th data.
D(4,1,1) distance from m1 = $\sqrt{(4-3{,}71)^2+(1-1.14)^2+(1-1.57)^2} = 2.93$
D(4,1.1) distance from m2 = $\sqrt{(4-1)^2+(1-1.50)^2+(1-3)^2} =$
4,38 D(4,1,1) distance from m3 = $\sqrt{(4-2{,}11)^2+(1-1{,}44)^2+(1-2{,}33)^2} = 3.44$
Calculation of object distance to centroid with 19th data.
D(4,1,1) distance from m1 = $\sqrt{(4-3{,}71)^2+(1-1{,}14)^2+(1-1.57)^2} =$
2,93 D(4,1,1) distance from m2 = $\sqrt{(4-1)^2+(1-1.50)^2+(1-3)^2} = 4.38$
D(4,1.1) distance from m3 = $\sqrt{(4-2.11)^2+(1-1.44)^2+(1-2{,}33)^2} = 3.44$
Calculation of object distance to centroid with 20th data.
D(1,1,1) distance from m1 = $\sqrt{(1-3{,}71)^2+(1-1{,}14)^2+(1-1{,}57)^2} = 2.77$
D(1,1,1) distance from m2 = $\sqrt{(1-1)^2+(1-1.50)^2+(1-3)^2} = 2.06$
D(1,1 ,1) distance from m3 = $\sqrt{(1-2.11)^2+(1-1.44)^2+(1-2.33)^2} = 1.78$

9. After getting the result of the *object* to the *centroid*, then group *objects* based on the closest distance. The following table is the distance of the object to the *centroid* in iteration-2.

Tabel 5. Distance of Object to Centroid in 2nd Iteration

X	Y	Z	Distance from m1	Distance from m2	Distance from m3
4	1	2	0,53	3,20	1,96
1	2	4	3,74	1,11	2,08
3	2	3	1, 81	2.06	1.24
1	1	2	2.87	1.11	1.28
2	1	2	1.76	1.50	0.56
1	1	1	2.77	2.06	1.78
2	1	2	1.76	1 .50	0.56
3	2	3	1.81	2.06	1.24
3	2	4	2.67	2.29	1.97
4	1	1	2.93	4.38	3.44
3	1	2	0.84	2.29	1.04
1	1	2	2.87	1.11	1.28
3	1	1	0.92	2.87	1.65
4	2	3	1.69	3.04	2.08
2	1	2	1.76	1.50	0, 56
1	2	4	3.74	1.11	2.08
2	2	3	2.38	1.11	0.88
4	1	1	2.93	4.38	3.44
4	1	1	2.93	4.38	3.44
1	1	1	2.77	2.06	1.78

10. Group based on shortest distance. K = 3, *centroid* 1 (m1) = ((3.71), (1.14), (1.57)), *centroid* 2 (m2) = ((1), (1.50), (3)), *centroid* 3 (m3) = ((2,11), (1,44), (2,33)). If m1 < m2 < m3 = group 1, m2 < m1 < m3 = group 2, m3 < m1 < m2 = group 3. Following are the results of grouping in iteration-2.

Tabel 6. Results of Grouping in the 2nd Iteration

X	Y	Z	Distance from m1	Distance from m2	Distance from m3	Group
4	1	2	0,53	3,20	1,96	1
1	2	4	3,74	1,11	2,08	2
3	2	3	1 ,81	2.06	1.24	3
1	1	2	2.87	1.11	1.28	2
2	1	2	1.76	1.50	0.56	3
1	1	1	2.77	2.06	1.78	3
2	1	2	1.76	1.50	0.56	3
3	2	3	1.81	2.06	1.24	3
3	2	4	2.67	2.29	1.97	3
4	1	1	2.93	4.38	3.44	1
3	1	2	0.84	2.29	1.04	1
1	1	2	2.87	1.11	1.28	2
3	1	1	0.92	2.87	1.65	1
4	2	3	1.69	3.04	2, 08	1
2	1	2	1.76	1.50	0.56	3
1	2	4	3.74	1.11	2.08	2
2	2	3	2.38	1.11	0.88	3
4	1	1	2.93	4.38	3.44	1

4	1	1	2.93	4.38	3.44	1
1	1	1	2.77	2.06	1.78	3

Groups based on the minimum distance to the *centroid* nearest
Old group = {1,2,3,2, 3,3,3,3,3,1,1,2,1,1,3,2,3,1,1,3}
New group = {1,2,3,2,3,3,3, 3,3,1,1,2,1,1,3,2,3,1,1,3}
No group changes, operation complete.

CONCLUSION

The calculation results from the grouping of motorized vehicle tests based on the type of vehicle are as follows:

1. *Cluster* 1 there are 7 groups with 7 types of vehicles: 2 BUS and 5 betor where there is one type of vehicle (betor) that does not pass the test due to the type of damage at the time of testing motor vehicle. *Cluster* 1 is the type of vehicle that passes the motor vehicle test the most with the lowest level of damage.
2. *Cluster* 2 there are 4 groups by type of vehicle: 4 freight cars where 2 of them did not pass the test because of the type of damage during the motor vehicle test.
3. *Cluster* 3 has 9 groups with the types of vehicles: 2 freight cars, 4 MPUs, and 3 BUS. 3 BUS and 1 MPU did not pass the test due to damage during the motor vehicle test. *cluster* is the *cluster* of vehicle types that do not pass the test with the most types of damage.

REFERENCES

Nabila, Z., Rahman Isnain, A., & Abidin, Z. (2021). DATA MINING ANALYSIS FOR CLUSTERING OF COVID-19 CASES IN LAMPUNG PROVINCE WITH K-MEANS ALGORITHM. Journal of Technology and Information Systems (JTSI), 2(2), 100. http://jim.teknokrat.ac.id/index.php/JTSI

Nofitri, R., & Irawati, N. (2019). DATA ANALYSIS OF PROFIT RESULTS USING RAPIDMINER SOFTWARE. JURTEKSI (Journal of Technology and Information Systems), 5(2), 199–204. https://doi.org/10.33330/jurteksi.v5i2.365

Mara, N., & Intisari, NS (2013). CLASSIFICATION OF CHARACTERISTICS WITH K-MEANS CLUSTER ANALYSIS METHOD. In Scientific Bulletin Matt. stats. and Its Application (Bimaster) (Vol. 02, Issue 2).

Pulungan, W., Poningsih, P., & Satria, H. (2019). GROUPING ON MOTOR VEHICLES BY USING THE K-MEANS DATA MINING METHOD. KOMIK (National Conference on Information and Computer Technology), 3(1). https://doi.org/10.30865/komik.v3i1.1687

Syarif, R., Furqon, MT, & Adinugroho, S. (2018). Comparison of K-Means Algorithm and Fuzzy C Means (FCM) Algorithm in GPS-Based Transport Mode Clustering (Vol. 2, Issue 10). http://j-ptiik.ub.ac.id