Proceeding
International Conference
on Social, Sciences and Information Technology
Kisaran, August 19th, 2020, page. 273 - 282
DOI: https://doi.org/10.33330/icossit.v1i1.799
Available online at https://jurnal.stmikroyal.ac.id/index.php/ICdoSSIT

ISSN 2723-4509 (Online)

E-COMMERCE CLUSTERING ANALYSIS BASED ON LARGEST VISITORS

Nurwati
Computer Engineering, Sekolah Tinggi Manajemen Informatika dan Komputer Royal, Indonesia

Corresponding author:
nurwati763@gmail.com

Keywords:
Clastering
E-Commerce
Biggest Visitor

ABSTRACT

E-Commerce is an Electric Commerce which is very much and popularly used by all people. The increase in buying and selling online, making the number of e-commerce in Indonesia and the great benefits of e-commerce has made researchers conduct e-commerce cluster analysis so that online shop owners can join the most clusters so that they can increase their business. E-commerce clustering analysis based on the number of visitors visiting the e-commerce site or website. Clustering is used to produce the Most, Most, and Enough E-commerce groups. From the results of the clustering, it was found that the e-commerce names Tokopedia, Bukalapak and Shopee were the clusters with the most visitors. Lazada and Blibli are clusters of Many Visitors and JD Id, Orami, Sociolla, zalora, bhinneka, elevenia, blanja, laku6, jakarta Notebook and Ralali are included in the Quite Many Visitors cluster. Clustering data is based on the number of visitors starting in Quartiles 1-2019, Quartiles 2-2019, Quartiles 3-2019, Quartiles 4-2019 and Quartiles 1-2020.

INTRODUCTION

Analyzing a data object is important in producing correct information. Cluster analysis is part of grouping data (objects) which is based on the existence of information about data objects and relations [1]. Cluster analysis is one of the multiple variable analysis because it is useful in classifying objects into several groups based on the similarity of the variables, resulting in several groups of similarities of objects [2].

E-Commerce is a business activity that utilizes the internet network [3]. E-Commerce is an electronic marketing technique that does not require high costs but has a large or broad marketing reach [4]. E-Commerce is the basis of an m-commerce which includes Business to Busines (B2B) and business to Customer (B2C). business to business is part of inter-business, while B2C is part of business between consumers [5]. The development of computerized technology and the development of e-commerce in plain view has increased very high. This increase also affects all aspects, especially the Indonesian economy. The magnitude of the influence of the results of the analysis requires good quantitative skills, as well as the ability to think structured in accordance

Proceeding
International Conference
on Social, Sciences and Information Technology
Kisaran, August 19th, 2020, page. 273 - 282
DOI: https://doi.org/10.33330/icossit.v1i1.799
Available online at https://jurnal.stmikroyal.ac.id/index.php/ICdoSSIT
ISSN 2723-4509 (Online)

with the expected results.The clustering of e-commerce that the researchers conducted to obtain the largest group of visitors based on the number of visitors used a non-hierarchical method, in the form of K-Means Meaning Data [7], [8].

E-commerce clustering is based on research data that researchers obtained from databoks.katadata.co.id, which are as follows:

Table 1. Data on the number of visitors

NO	Shop Name	Year				
		Q1-2019	**Q2-2019**	**Q3-2019**	**Q4-2019**	**Q1-2020**
1	TOKOPEDIA	137200900	140414500	65953400	67900000	69800000
2	BUKALAPAK	115256600	89765800	42874100	39263300	37633300
3	SHOPEE	74995300	90705300	55964700	72973300	71533300
4	LAZADA	52044500	49620200	27995900	28383300	24400000
5	BLIBLI	32597200	38453000	21395600	26863300	17600000
6	JD ID	10656900	7102300	5524000	13539300	6066700
7	ORAMI	8380600	9813100	3906400	3708300	5642500
8	SOCIOLLA	4838300	5101800	3988300	2704300	3050000
9	ZALORA	4343000	5218300	2804100	2926300	2416700
10	BHINNEKA	3446500	7678900	5037700	5145700	4450000
11	ELEVENIA	3394400	3001400	856700	1084300	1166700
12	BLANJA	3225200	5511600	3571500	2709300	403400
13	LAKU6	2512900	2976700	1150500	1177000	940000
14	JAKARTA NOTE-BOOK	2450100	2554500	1769500	1597300	1350000
15	RALALI	2164800	2680000	3583400	2139300	1366700

Based on the data above, the researcher will then carry out a cluster analysis with the aim of making objects that are combined in related clusters and not related to objects in other clusters[9].

Data mining is part of a scientific discipline that produces patterns from data. The K-Means method is part of data mining. K is the number of clusters, Means is the way to generate cluster analysis[10].

METHOD

Data meaning K-Means clustering is a method that researchers apply in classifying the best e-commerce based on the number of visitors, whose results are obtained by appropriate clusters or groups [11]. Data mining has several stages, namely as follows:

1. Data Cleaning
   Data cleaning in data mining functions to remove noise and irrelevant data.

Proceeding **ISSN 2723-4509 (Online)**
International Conference
on Social, Sciences and Information Technology
Kisaran, August 19th, 2020, page. 273 - 282
DOI: https://doi.org/10.33330/icossit.v1i1.799
Available online at https://jurnal.stmikroyal.ac.id/index.php/ICdoSSIT

2. Data Integration
   Data integration in data mining functions as a merger of data from various databases into one new database.
3. Data Selection
   Data selection in data mining is only data that is suitable for analysis which will be taken from the database.
4. Data transformation
   Data transformation in data mining is converting data into a format suitable for the datamining process
5. The mining process
   The mining process is the main process when applying the methods used to discover new knowledge.
6. Evaluate patterns
   Pattern evaluation is the process of identifying patterns that attract new knowledge.
7. Presentation of knowledge
   Presentation of knowledge is a useful visualization to obtain knowledge that users gain.

Data mining is divided into six groups based on the tasks and work to be carried out, one of which is clustering or grouping. Clustering is widely used in identifying a limited set within a category to describe the data being studied.

The k-means method processes cases as a whole where k is the number of groups (most, many and quite a lot). The stages involved in the k-means cluster analysis include:

1. Specifies multiple clusters
2. Determine the initial cluster center (centroid).

$$v = \frac{\sum_{i=1}^{n}(xi)}{n} \; ; i = 1,2,3,\ldots..n$$

3. Get the value from the nearest centroid
   Namely by calculating the distance of each data to each centroid using the Euclidean Distance formula

$$d(x,y) = |x - y| = \sqrt{\sum_{i=1}^{n}(xi - yi)2} \; ; i = 1,2,3,\ldots.,n$$

4. Group data based on its minimum distance
5. Looking for the new centroid from the previous results based on member-ship, namely by calculating the average of the data for each cluster.
6. Back to stage three
7. Looping stops when no more data is moved.

Proceeding
International Conference
on Social, Sciences and Information Technology
Kisaran, August 19th, 2020, page. 273 - 282
DOI: https://doi.org/10.33330/icossit.v1i1.799
Available online at https://jurnal.stmikroyal.ac.id/index.php/ICdoSSIT
ISSN 2723-4509 (Online)

Required data

The data needed in this study is the number of visitors visiting e-commerce sites in Indonesia, namely, Tokopedia, Bukalapak, Shopee, Lazada, Blibli, Jd Id, Orami, Sociolla, Zalora, Bhinneka, Elevenia, Blanja, Laku6, Jakarta Notebook. , Ralali. The variable data used is the number of visitors in 2019 and 2020, starting from Quarter 1 2019, Quarter 2 2019, Quarter 3 2019, Quarter 4 2019 and Quarter 1 2020.

RESULT AND DISCUSSION

As for the results of the analysis of e-commerce clustering based on the most visitors, the researchers started from collecting data on the number of visitors starting from Q1 2019, Q2 2019, Q3 2019, Q4 2019 to Q1 2020 and 15 e-commerce. in Indonesia, as in table 1. The following is an example of data calculated using K-Means Clustering, namely many clusters = 5, many data = 15 and many attributes = 3.

After evaluating the initial cluster center, the next step is to calculate the cluster center distance using the equation, which aims to be able to get C1, C2 and C3, as follows:

Table 2. The results of the cluster center calculation

Calculation of Cluster Center Distance										
NO	Shop Name	Year					C1	C2	C3	Shortest distance
		Q1-2019	Q2-2019	Q3-2019	Q4-2019	Q1-2020				
1	Tokopedia	137200900	140414500	65953400	67900000	69800000	0	223837316	221597603	0
2	Bukalapak	115256600	89765800	42874100	39263300	37633300	73717395	156878643	154650443	73717395
3	Shopee	74995300	90705300	55964700	72973300	71533300	80430455	160428623	158588945	80430455
4	Lazada	52044500	49620200	27995900	28383300	24400000	143383119	80663607	78488439	78488439
5	Blibli	32597200	38453000	21395600	26863300	17600000	166530267	58291587	56077514	56077514
6	Jd Id	10656900	7102300	5524000	13539300	6066700	210841785	15731765	14523588	14523588
7	Orami	8380600	9813100	3906400	3708300	5642500	213864407	10505515	8581339	8581339
8	Sociolla	4838300	5101800	3988300	2704300	3050000	219943402	4040963	3154087	3154087
9	Zalora	4343000	5218300	2804100	2926300	2416700	220633540	3676506	2454561	2454561
10	Bhinneka	3446500	7678900	5037700	5145700	4450000	217780684	6876912	5404343	5404343
11	Elevenia	3394400	3001400	856700	1084300	1166700	224051305	3194234	4113772	3194234
12	Blanja	3225200	5511600	3571500	2709300	403400	221597603	3224189	0	0
13	Laku6	2512900	2976700	1150500	1177000	940000	224550841	2690037	3928131	2690037
14	Jakarta Notebook	2450100	2554500	1769500	1597300	1350000	224419654	1918703	3837319	1918703
15	Ralali	2164800	2680000	3583400	2139300	1366700	223837316	0	3224189	0.00

Proceeding
International Conference
on Social, Sciences and Information Technology
Kisaran, August 19th, 2020, page. 273 - 282
DOI: https://doi.org/10.33330/icossit.v1i1.799
Available online at https://jurnal.stmikroyal.ac.id/index.php/ICdoSSIT
ISSN 2723-4509 (Online)

The results of the calculations above, then are grouping. Matrix grouping is marked with 1. The number 1 in this matrix means that the data is included in the grouping data group.

Tabel 3. Grouping data 1

No.	C1	C2	C3
1	1	0	0
2	1	0	0
3	1	0	0
4	0	0	1
5	0	0	1
6	0	0	1
7	0	0	1
8	0	0	1
9	0	0	1
10	0	0	1
11	0	1	0
12	0	0	1
13	0	1	0
14	0	1	0
15	0	1	0

Table 4. New cluster calculation results

No	Shop Name	Year					New Claster		
		Q1-2019	**Q2-2019**	**Q3-2019**	**Q4-2019**	**Q1-2020**	C1	C2	C3
1	Tokopedia	137200900	140414500	65953400	67900000	69800000	109150933	2630550	14941525
2	Bukalapak	115256600	89765800	42874100	39263300	37633300	106961867	2803150	16062400
3	Shopee	74995300	90705300	55964700	72973300	71533300	54930733	1840025	9277938
4	Lazada	52044500	49620200	27995900	28383300	24400000	60045533	1499475	10747475
5	Blibli	32597200	38453000	21395600	26863300	17600000	59655533	1205850	8003663
6	Jd Id	10656900	7102300	5524000	13539300	6066700			
7	Orami	8380600	9813100	3906400	3708300	5642500			
8	Sociolla	4838300	5101800	3988300	2704300	3050000			
9	Zalora	4343000	5218300	2804100	2926300	2416700			
10	Bhinneka	3446500	7678900	5037700	5145700	4450000			
11	Elevenia	3394400	3001400	856700	1084300	1166700			
12	Blanja	3225200	5511600	3571500	2709300	403400			
13	Laku6	2512900	2976700	1150500	1177000	940000			
14	Jakarta Notebook	2450100	2554500	1769500	1597300	1350000			
15	Ralali	2164800	2680000	3583400	2139300	1366700			

Proceeding
International Conference
on Social, Sciences and Information Technology
Kisaran, August 19th, 2020, page. 273 - 282
DOI: https://doi.org/10.33330/icossit.v1i1.799
Available online at https://jurnal.stmikroyal.ac.id/index.php/ICdoSSIT

ISSN 2723-4509 (Online)

After evaluating the center of the second cluster, the next step is to calculate the cluster center distance using the equation, which aims to get C1, C2 and C3, as follows:

Table 5. Results of the calculation of the second New cluster

Cluster Center Distance Calculation										
No	Shop Name	Year					C1	C2	C3	Shortest distance
		Q1-2019	**Q2-2019**	**Q3-2019**	**Q4-2019**	**Q1-2020**				
1	Tokopedia	137200900	140414500	65953400	67900000	69800000	46818600	224211451	201762499	46818600
2	Bukalapak	115256600	89765800	42874100	39263300	37633300	37352728	157111395	135333201	37352728
3	Shopee	74995300	90705300	55964700	72973300	71533300	41715235	161081915	138801082	41715235
4	Lazada	52044500	49620200	27995900	28383300	24400000	97571237	81107417	58591646	58591646
5	Blibli	32597200	38453000	21395600	26863300	17600000	120616220	58850864	36217525	36217525
6	Jd Id	10656900	7102300	5524000	13539300	6066700	164768400	16280722	11148074	11148074
7	Orami	8380600	9813100	3906400	3708300	5642500	168188338	10537322	12887124	10537322
8	Sociolla	4838300	5101800	3988300	2704300	3050000	174195102	4430071	18423398	4430071
9	Zalora	4343000	5218300	2804100	2926300	2416700	174907213	3632749	19084516	3632749
10	Bhinneka	3446500	7678900	5037700	5145700	4450000	171992709	7647438	16260587	7647438
11	Elevenia	3394400	3001400	856700	1084300	1166700	178363659	1328006	22692820	1328006
12	Blanja	3225200	5511600	3571500	2709300	403400	175921708	3576999	20088089	3576999
13	Laku6	2512900	2976700	1150500	1177000	940000	178857252	833110	23092224	833110
14	Jakarta Notebook	2450100	2554500	1769500	1597300	1350000	178683558	360155	22866408	360155
15	Ralali	2164800	2680000	3583400	2139300	1366700	178067428	1925278	22201600	1925278

The results of the calculation of the second cluster center distance above are grouped. Matrix grouping is marked with 1. The number 1 in this matrix means that the data is included in the grouping data group.

Table 6. Grouping data 2

No.	**C1**	**C2**	**C3**
1	1	0	0
2	1	0	0
3	1	0	0
4	0	0	1
5	0	0	1
6	0	0	1
7	0	1	0
8	0	1	0
9	0	1	0
10	0	1	0

Proceeding
International Conference
on Social, Sciences and Information Technology
Kisaran, August 19th, 2020, page. 273 - 282
DOI: https://doi.org/10.33330/icossit.v1i1.799
Available online at https://jurnal.stmikroyal.ac.id/index.php/ICdoSSIT
ISSN 2723-4509 (Online)

Table 6. Grouping data 2

No.	C1	C2	C3
11	0	1	0
12	0	1	0
13	0	1	0
14	0	1	0
15	0	1	0

Tabel 7.New cluster calculation results

NO	Shop Name	Year					New Cluster		
		Q1-2019	**Q2-2019**	**Q3-2019**	**Q4-2019**	**Q1-2020**	C1	C2	C3
1	Tokopedia	137200900	140414500	65953400	67900000	69800000	109150933.3	3861755.56	31766200
2	Bukalapak	115256600	89765800	42874100	39263300	37633300	106961866.7	4948477.78	31725166.7
3	Shopee	74995300	90705300	55964700	72973300	71533300	54930733.33	2963122.22	18305166.7
4	Lazada	52044500	49620200	27995900	28383300	24400000	60045533.33	2576866.67	22928633.3
5	Blibli	32597200	38453000	21395600	26863300	17600000	59655533.33	2309555.56	16022233.3
6	Jd Id	10656900	7102300	5524000	13539300	6066700			
7	Orami	8380600	9813100	3906400	3708300	5642500			
8	Sociolla	4838300	5101800	3988300	2704300	3050000			
9	Zalora	4343000	5218300	2804100	2926300	2416700			
10	Bhinneka	3446500	7678900	5037700	5145700	4450000			
11	Elevenia	3394400	3001400	856700	1084300	1166700			
12	Blanja	3225200	5511600	3571500	2709300	403400			
13	Laku6	2512900	2976700	1150500	1177000	940000			
14	Jakarta Notebook	2450100	2554500	1769500	1597300	1350000			
15	Ralali	2164800	2680000	3583400	2139300	1366700			

After evaluating the center of the third cluster above the same as the previous steps, namely calculating the cluster center distance using the equation, which aims to get C1, C2 and C3, as follows:

Table 8. The results of the calculation of the Third New cluster

Cluster Center Distance Calculation										
No	Shop Name	Year					C1	C2	C3	Shortest distance
		Q1-2019	**Q2-2019**	**Q3-2019**	**Q4-2019**	**Q1-2020**				
1	Tokopedia	137200900	140414500	65953400	67900000	69800000	46818600	221179153	173535798	46818600
2	Bukalapak	115256600	89765800	42874100	39263300	37633300	37352728	154237923	108059493	37352728
3	Shopee	74995300	90705300	55964700	72973300	71533300	41715235	158022041	111137747	41715235
4	Lazada	52044500	49620200	27995900	28383300	24400000	97571237	78088023	30418618	30418618

Proceeding
International Conference
on Social, Sciences and Information Technology
Kisaran, August 19th, 2020, page. 273 - 282
DOI: https://doi.org/10.33330/icossit.v1i1.799
Available online at https://jurnal.stmikroyal.ac.id/index.php/ICdoSSIT
ISSN 2723-4509 (Online)

Table 8. The results of the calculation of the Third New cluster

Cluster Center Distance Calculation										
No	Shop Name	Year					C1	C2	C3	Shortest distance
		Q1-2019	Q2-2019	Q3-2019	Q4-2019	Q1-2020				
5	Blibli	32597200	38453000	21395600	26863300	17600000	120616220	55782287	8571816	8571816
6	Jd Id	10656900	7102300	5524000	13539300	6066700	164768400	13844209	37450214	13844209
7	Orami	8380600	9813100	3906400	3708300	5642500	168188338	7573832	41370394	7573832
8	Sociolla	4838300	5101800	3988300	2704300	3050000	174195102	1610166	47076434	1610166
9	Zalora	4343000	5218300	2804100	2926300	2416700	174907213	680641	47750736	680641
10	Bhinneka	3446500	7678900	5037700	5145700	4450000	171992709	4807486	44792894	4807486
11	Elevenia	3394400	3001400	856700	1084300	1166700	178363659	3461275	51306184	3461275
12	Blanja	3225200	5511600	3571500	2709300	403400	175921708	2177935	48698974	2177935
13	Laku6	2512900	2976700	1150500	1177000	940000	178857252	3581651	51740891	3581651
14	Jakarta Notebook	2450100	2554500	1769500	1597300	1350000	178683558	3320984	51517861	3320984
15	Ralali	2164800	2680000	3583400	2139300	1366700	178067428	3080720	50828890	3080720

The results of the calculation of the third cluster center distance above are then grouped. The steps are continued as the previous steps until the data is no longer changing or there are no changes in the group data. The following table shows the results of each step.

Tabel 9. Grouping data 3

No.	C1	C2	C3
1	1	0	0
2	1	0	0
3	1	0	0
4	0	0	1
5	0	0	1
6	0	1	0
7	0	1	0
8	0	1	0
9	0	1	0
10	0	1	0
11	0	1	0
12	0	1	0
13	0	1	0
14	0	1	0
15	0	1	0

Proceeding
International Conference
on Social, Sciences and Information Technology
Kisaran, August 19th, 2020, page. 273 - 282
DOI: https://doi.org/10.33330/icossit.v1i1.799
Available online at https://jurnal.stmikroyal.ac.id/index.php/ICdoSSIT
ISSN 2723-4509 (Online)

Table 10.New Cluster Calculation Results

NO	Shop Name	Year					New Cluster		
		Q1-2019	**Q2-2019**	**Q3-2019**	**Q4-2019**	**Q1-2020**	C1	C2	C3
1	Tokopedia	137200900	140414500	65953400	67900000	69800000	109150933	4541270	42320850
2	Bukalapak	115256600	89765800	42874100	39263300	37633300	106961867	5163860	44036600
3	Shopee	74995300	90705300	55964700	72973300	71533300	54930733	3219210	24695750
4	Lazada	52044500	49620200	27995900	28383300	24400000	60045533	3673110	27623300
5	Blibli	32597200	38453000	21395600	26863300	17600000	59655533	2685270	21000000
6	Jd Id	10656900	7102300	5524000	13539300	6066700			
7	Orami	8380600	9813100	3906400	3708300	5642500			
8	Sociolla	4838300	5101800	3988300	2704300	3050000			
9	Zalora	4343000	5218300	2804100	2926300	2416700			
10	Bhinneka	3446500	7678900	5037700	5145700	4450000			
11	Elevenia	3394400	3001400	856700	1084300	1166700			
12	Blanja	3225200	5511600	3571500	2709300	403400			
13	Laku6	2512900	2976700	1150500	1177000	940000			
14	Jakarta Notebook	2450100	2554500	1769500	1597300	1350000			
15	Ralali	2164800	2680000	3583400	2139300	1366700			

Table 11. Calculation results for the Fourth New Cluster

Cluster Center Distance Calculation										
NO	SHOP NAME	Year					C1	C2	C3	Shortest distance
		Q1-2019	**Q2-2019**	**Q3-2019**	**Q4-2019**	**Q1-2020**				
1	Tokopedia	137200900	140414500	65953400	67900000	69800000	46818600	220128081	154908908	46818600
2	Bukalapak	115256600	89765800	42874100	39263300	37633300	37352728	153218502	90296135	37352728
3	Shopee	74995300	90705300	55964700	72973300	71533300	41715235	156862241	93987081	41715235
4	Lazada	52044500	49620200	27995900	28383300	24400000	97571237	76999548	12196497	12196497
5	Blibli	32597200	38453000	21395600	26863300	17600000	120616220	54643548	12196497	12196497
6	Jd Id	10656900	7102300	5524000	13539300	6066700	164768400	12459788	56175322	12459788
7	Orami	8380600	9813100	3906400	3708300	5642500	168188338	6750894	59692258	6750894
8	Sociolla	4838300	5101800	3988300	2704300	3050000	174195102	1324839	65519598	1324839
9	Zalora	4343000	5218300	2804100	2926300	2416700	174907213	918937	66209536	918937
10	Bhinneka	3446500	7678900	5037700	5145700	4450000	171992709	4014181	63235128	4014181
11	Elevenia	3394400	3001400	856700	1084300	1166700	178363659	4536628	69750518	4536628
12	Blanja	3225200	5511600	3571500	2709300	403400	175921708	2848321	67110407	2848321
13	Laku6	2512900	2976700	1150500	1177000	940000	178857252	4738570	70191007	4738570
14	Jakarta Notebook	2450100	2554500	1769500	1597300	1350000	178683558	4401743	69996639	4401743
15	Ralali	2164800	2680000	3583400	2139300	1366700	178067428	4005122	69307350	4005122

Proceeding
International Conference
on Social, Sciences and Information Technology
Kisaran, August 19th, 2020, page. 273 - 282
DOI: https://doi.org/10.33330/icossit.v1i1.799
Available online at https://jurnal.stmikroyal.ac.id/index.php/ICdoSSIT
ISSN 2723-4509 (Online)

CONCLUSION

Based on the analysis and results above, a conclusion can be drawn as follows: The use of the K-Means clustering data mining method, the process of classifying E-Commerce based on the number of visitors, is easier than the manual method. Clustering results can be used as recommendations for e-commerce users in determining which e-commerce is the most visited. The variables and attributes used in the clustering greatly affect the results.

BIBLIOGRAPHY

[1] D. R. Ningrat, D. A. I. Maruddani, and T. Wuryandari, "Analisis Cluster Dengan Algoritma K-Means Dan Fuzzy C-Means Clustering Untuk Pengelompokan Data Obligasi Korporasi," *None*, vol. 5, no. 4, pp. 641–650, 2016.

[2] J. Matematika, F. Matematika, D. A. N. Ilmu, P. Alam, and U. N. Semarang, *Perbandingan Kinerja Metode Single Linkage , Metode Complete Linkage Dan Metode K-Means Dalam Analisis Cluster*. 2011.

[3] F. A. Dewa and M. T. Jatipaningrum, "SEGMENTASI E-COMMERCE DENGAN CLUSTER K-MEANS DAN FUZZY C-MEANS (Studi Kasus : Media Sosial di Indonesia yang diunduh di Play Store)," vol. 4, no. 1, pp. 53–67, 2019.

[4] R. Hastanti, Puji, B. eka purnama, and I. ully wardati, "Sistem Penjualan Berbasis Web (E-Commerce) Pada Tata Distro Kabupaten Pacitan," *J. Bianglala Inform.*, vol. 3, no. 2, pp. 1–9, 2015, doi: 10.1007/s13226-018-0284-5.

[5] W. Fitri and P. Nugroho, "Peningkatan Daya Saing Bisnis ' Batik Kahuripan ,'" *Semin. Ris. Unggulan Nas. Inform. dan Komput. FTI UNSA*, vol. 2, no. 1, pp. 25–32, 2013.

[6] K. Fatmawati and A. P. Windarto, "Data Mining: Penerapan Rapidminer Dengan K-Means Cluster Pada Daerah Terjangkit Demam Berdarah Dengue (Dbd) Berdasarkan Provinsi," *Comput. Eng. Sci. Syst. J.*, vol. 3, no. 2, p. 173, 2018, doi: 10.24114/cess.v3i2.9661.

[7] V. Novita Sari, Y. Yupianti, and D. Maharani, "Penerapan Metode K-Means Clustering Dalam Menentukan Predikat Kelulusan Mahasiswa Untuk Menganalisa Kualitas Lulusan," *Jurteksi*, vol. 4, no. 2, pp. 133–140, 2018, doi: 10.33330/jurteksi.v4i2.53.

[8] B. Poerwanto and R. Y. Fa'rifah, "Analisis Cluster K-Means dalam Pengelompokan Kemampuan Mahasiswa," *J. Sci. Pinisi*, vol. 2, no. 2, pp. 92–96, 2016.

[9] E. Irfiani and S. S. Rani, "Algoritma K-Means Clustering untuk Menentukan Nilai Gizi Balita," *J. Sist. dan Teknol. Inf.*, vol. 6, no. 4, p. 161, 2018, doi: 10.26418/justin.v6i4.29024.

[10] J. Xu and K. Lange, "Power k-means clustering," *36th Int. Conf. Mach. Learn. ICML 2019*, vol. 2019-June, pp. 11977–11991, 2019.