**Organization clustering airports using K-Means clustering algorithm**

**D L Trenggonowati*, M Ulfah, R Ekawati and V A Yusuf**  
Industrial Engineering Department, Engineering Faculty, Sultan Ageng Tirtayasa University  
*E-mail: dyahlintang@untirta.ac.id

**Abstract.** One of the right management quantity of human resources is through the establishment of an efficient organizational structure in accordance with the conditions of the airport. PT. Angkasa Pura II (Persero) is one of the State-Owned Enterprises engaged in the business of airport services in the western region which currently has 16 airports. With the growing needs of air transportation, PT. Angkasa Pura II is projected by the Ministry of Transportation to become the manager of 21 airports. With an additional projection of 5 airports, PT. Angkasa Pura II (Persero) requires projections clustering for all 21 airports to be managed in order to form the right organizational structure. Therefore, 5 clusters are formed using the k-means algorithm. This method is used because it is one of the partitional clustering methods. Partitional clustering method was chosen because it was known that the company wanted to form 5 clusters. In this study clustering was carried out based on variable aircraft movements, passenger movement, cargo, area, terminal area, runway, EBITDA and revenue. The result is obtained in cluster (1) there is 1 airport, cluster (2) there are 6 airports, cluster (3) there are 5 airports, cluster (4) there are 2 airports and cluster (5) there are 5 airports.

**1. Introduction**

Airport is an area on land and / or waters with convinced limits that are used as a place for airliner landing and taking off, boarding passengers, loading and unloading goods, and intra-and inter-mode transportation, which are armed with safety facilities and navigation security, as well as basic facilities and other aiding facilities (PM.69 of 2013 concerning the National Airport Regulations). The structure of a corporate organization is a formal framework of an organization with a framework that is divided, grouped and coordinated [4]. PT Angkasa Pura II (Persero) is one of the State-Owned Business in the Exhortation of Transportation which is committed in the field of airport benefit and airport-related benefit in the territory of West Indonesia. Until now, PT. Angkasa Pura II has managed 16 airports but over time PT. Angkasa Pura II (Persero) is projected by the Ministry of Transportation to become the manager of 21 airports by adding 5 new airports. With the additional 5 airport projections, PT. Angkasa Pura II (Persero) as the airport manager takes projection clustering all 21 airports to be managed in order to establish appropriate organizational structure at every airport in accordance with the circumstances and a load of each airport so there is no wastage of human resources. Based on the needs of clustering projections of 21 airports to be managed, this study usage the k-means algorithm to create clusters. This is because the K-Means algorithm issues forth partitional clustering arrangements. Partitional clustering method was chosen because it was known that the company wanted to form 5 airport clusters based on the projections of 21 airports to be managed.
2. Methodology

In clustering, this method attempts to place similar objects (close distance) in one cluster and make the distance between clusters as far as possible. In general the method on clustering can be classified into 2 methods that is partitional clustering and hierarchy clustering. The difference between the two based on the way the clustering is. Hierarchy clustering creates a certain level of cluster while partitional clustering divides the data set into a number of groups that do not overlap one group to another, meaning that each data only becomes one group. In the hierarchy clustering there is no known number of clusters to be made while the partitional clustering is already known to the number of clusters to be made. The hierarchical clustering method consists of complete linkage clustering, single link age clustering, average link age clustering and centroid link age clustering [2].

In partitional clustering consists of the K-Means method and DBSCAN. The K-Means approach is compared to other methods is more analyze in capturing information. K-Means is one method of non-hierarchical clustering data that attempts to partition existing data into one or more clusters or groups. This method is partitioned into clusters or groups so that the data that has the same characteristics (High intra-class similarity) are grouped into the same cluster and which have different characteristics (inter-class similarity Law) grouped in the other group [3].

The steps to Clustering with the K Means method are as follows:
1. Select the number of cluster K.
2. Initialize K cluster centers randomly.
3. Allocation of all data / objects to the nearest cluster.
   To calculate the distance of all data to each cluster center point can use the Euclidean distance theory which is formulated as follows:
   \[ D(ij) = \sqrt{\sum_{i=1}^{p} |X_{ki} - X_{kj}|^2} \]  
   Where:
   \( D(ij) \) = data distance to (i) to cluster center (j)
   \( X_{ki} \) = Data to (i) in the data attribute to (k)
   \( X_{kj} \) = Center point (j) in attribute (k)
4. Recalculate the cluster center with the current cluster membership. The cluster center is the average of all data / objects in a particular cluster.
   \[ C(jk) = \frac{\sum_{i} X_{ki}}{n} \]  
   Where:
   \( C(jk) \) = New cluster center (centroid)
   \( X_{ki} \) = Data to (i) in the data attribute to (k) at temporary centroid position
   \( n \) = the amount 1of data at temporary centroid position
5. Assign each object to use the new cluster center. If the cluster center does not change again, the clustering process is complete. Or, go back to step number 3 until the center of the cluster doesn't change again.

3. Result And Discussion

3.1. ANOVA Test

Analysis of variance (ANOVA) technique is used to determine whether there is a difference or there is an influence of a factor on a particular response. The initial hypothesis (H0) is a statement tested both statistically in the form of no relationship or no influence. Hypothesis (H1) is an opposing statement from H0 in the form of a relationship or influence.

Based on Table 1, the calculation of Analysis of Variance above, it is obtained results that the overall value of F Statistics is greater than F Critical \( F_{0.05(4,14)} = 3.112 \). This shows the influence of all these variables on the formation of airport clusters.
Table 1. ANOVA

| Cluster          | Error Mean Square | Error Df | Error Mean Square | Error Df | F     | Sig.   |
|------------------|------------------|---------|------------------|---------|-------|--------|
| Zscore(pesawat)  | 4.064            | 4       | .125             | 14      | 32.597| .000   |
| Zscore(penumpang)| 4.191            | 4       | .088             | 14      | 47.407| .000   |
| Zscore(kargo)    | 3.833            | 4       | .191             | 14      | 20.116| .000   |
| Zscore(luasterminal) | 4.196          | 4       | .087             | 14      | 48.374| .000   |
| Zscore(luaswilayah) | 2.793          | 4       | .488             | 14      | 5.724 | .006   |
| Zscore(runway)   | 2.439            | 4       | .589             | 14      | 4.143 | .020   |
| Zscore(EBITDA)   | 4.021            | 4       | .137             | 14      | 29.403| .000   |
| Zscore(Revenue)  | 4.155            | 4       | .099             | 14      | 42.168| .000   |

3.2. Two-Step Cluster

Two-step cluster analysis is performed to show the value of Silhouette Coefficient which is useful to see the quality and strength of the Cluster, how well an object is placed in a Cluster. Two step cluster analysis can be obtained using SPSS software.

Silhouette Coefficient is used to see the quality and strength of a cluster, how well an object is placed in a cluster. Here is the result of two-step cluster analysis regarding silhouette coefficient values using SPSS Software.

Based on the consequence of the two-step cluster study when adopting the number of 5 clusters, can be known that the silhouette measure of cohesion and separation value is 0.5 which is categorized in the Good silhouette coefficient category. This shows that the quality and strength of the cluster formed is good.

Table 2. Airport data that will be clustered

| Airport | Aircraft (Unit) | Passenger (Person) | Cargo (Ton) | Terminal (Meter²) | Area (Hectares) | Runaway (Meter) | EBITDA (Juta) | Revenue (Juta) |
|---------|-----------------|--------------------|-------------|-------------------|-----------------|-----------------|---------------|---------------|
| A       | 37.372          | 4.286.289          | 18.322.718  | 29.000            | 835.99         | 2.240           | 91.994        | 226.932       |
| B       | 45.118          | 4.167.777          | 18.616.951  | 32000             | 305.62         | 2.250           | 51.130        | 167.835       |
| C       | 52.571          | 5.119.286          | 17.872.318  | 34.000            | 200.36         | 3.000           | 98.007        | 234.443       |
| D       | 33.696          | 4.349.498          | 17.729.484  | 20.568            | 411             | 2.750           | 43.677        | 161.354       |
| E       | 35.982          | 3.889.802          | 14.738.128  | 17.000            | 157             | 2.220           | 95.368        | 192.960       |
| F       | 80.431          | 7.609.051          | 27.169.914  | 21.108            | 3.000           | 148.289         | 318.187       |
| G       | 21.277          | 1.942.229          | 9.746.751   | 13.000            | 169.19         | 2.220           | 30.103        | 101.341       |
| H       | 25.562          | 2.258.470          | 9.785.893   | 12.170            | 145             | 2.250           | -3.058        | 67.251        |
| I       | 6.655           | 386.746            | 2.999.426   | 8.373             | 95.62          | 2.256           | -33.205       | 14.602        |
| J       | 12.559          | 1.321.375          | 5.505.561   | 14.742            | 230             | 3.000           | -16.545       | 56.498        |
| K       | 6.468           | 310.721            | 130.363     | 1.700             | 142             | 2.250           | -18.390       | 15.570        |
| L       | 7.340           | 209.323            | 0           | 7000              | 129             | 2.250           | -27.186       | 7.440         |
| M       | 1.031           | 92.789             | 0           | 118.284           | 873             | 2.500           | 0             | 0             |
| N       | 11.990          | 953.017            | 7.326.121   | 29.144            | 388             | 2.500           | -18.390       | 15.570        |
| O       | 22.539          | 2.463.703          | 19.740.841  | 3.084             | 128             | 2.770           | -18.390       | 15.570        |
| P       | 9.532           | 1.044.084          | 10.542.871  | 2.492             | 416             | 2.400           | -18.390       | 15.570        |
| Q       | 9.719           | 1.065.479          | 11.453.800  | 2.000             | 388             | 2.250           | -18.390       | 15.570        |
| R       | 472             | 16.413             | 113.959     | 304               | 129             | 1.850           | -18.390       | 15.570        |
| S       | 3.156           | 164.508            | 1.331.337   | 1.250             | 129             | 2.260           | -18.390       | 15.570        |
3.3. K-Means Clustering Algorithm

The problem solving are determine the number of clusters K, determine the point K centroid, calculate the distance between objects and K centroid, group by minimum distance until the centroid position same.

The cluster that want to create is 5 clusters outside Soekarno Hatta Airport and Kualanamu Airport. Data collection was conducted on the aircraft movement data, the movement of passengers, cargo, area, terminal area, EBITDA (Earnings Before Interest, Taxes, Depreciation, and Amortization), and revenue. Table 2. above is the data processing has done.

The next step is how to determine the initial Centroid is random selection. Table 3. below is the initial cluster center (centroid) that used in this research.

### Table 3. Initial centroid determination

| Number | Aircraft | Passenger | Cargo | Terminal | Area | Runway | EBITDA | Revenue |
|--------|----------|-----------|-------|----------|------|--------|--------|---------|
| 1      | 80.431   | 7.609051  | 27.169914 | 22.793   | 165  | 3.000  | 148289 | 318187  |
| 2      | 52.571   | 5.119286  | 17.872318 | 36.234   | 310.2| 3.000  | 98007  | 234443  |
| 3      | 35.982   | 3.889802  | 14.738128 | 34.549   | 145  | 2.220  | 95368  | 192960  |
| 4      | 6.468    | 310.721   | 130.363 | 1.700    | 82   | 2.400  | -18390 | 15570   |
| 5      | 472      | 16.413    | 113.959 | 0        | 0    | 1.800  | 0      | 0       |

The initial Centroid is chosen randomly appropriate with the number of Clusters to be formed. In this research there were 5 initial centroids for each parameter used. The 6th data was taken as the centre of the 1st Cluster, the 3rd data was the centre of the 2nd Cluster, the 5th data was taken as the centre of the 3rd Cluster, the 11th data was taken as the 4th Cluster centre, and the taken the 18th data as the 5th Cluster centre [7].

Table 4. below is the calculation of cluster distance by determining the shortest distance using the euclidean formula in iteration 1:

### Table 4. Distance of cluster center iteration 1

| Airport | C1     | C2     | C3     | C4     | C5     | Shortest Distance |
|---------|--------|--------|--------|--------|--------|-------------------|
| A       | 9451297| 947165 | 3606617| 18623251| 18704355| 947165            |
| B       | 9221113| 1211014| 3889115| 18885477| 18963869| 1211014           |
| C       | 9625729| 2237   | 3367014| 18383794| 18478828| 2237              |
| D       | 9989201| 788588 | 3027111| 18057325| 18141439| 788588            |
| E       | 12977003| 3367069| 17549  | 15041347| 15130015| 17549             |
| F       | 1685   | 9625735| 12977009| 28009430| 28103433| 1685              |
| G       | 18323320| 8725960| 3539125| 9754322 | 9824013 | 5359125           |
| H       | 18191259| 8579861| 5216510| 9850196 | 9928667 | 5216510           |
| I       | 25229049| 15609878| 12252282| 2870116 | 2909380 | 2870116           |
| J       | 22560571| 12938611| 9584842 | 5469558 | 5547624 | 5469558           |
| K       | 28009431| 18383798| 15041376| 162    | 295814 | 162               |
| L       | 28161910| 18536386| 15192411| 165676 | 226035 | 165676            |
| M       | 28192855| 18567697| 15221179| 280520 | 181143 | 181143            |
| N       | 20933299| 11342099| 7975435 | 7224421 | 7272831 | 7224421           |
| O       | 9043718 | 3256832 | 5206388 | 19728316| 19778898| 3256832           |
| P       | 17879657| 8390024 | 5073897 | 10438302| 10479455| 5073897           |
| Q       | 17027606| 7595735 | 4337023 | 11348564| 11388292| 4337023           |
Based on the calculation of euclidean distance from each cluster, it can be seen that the shortest distance from the cluster is the temporary location of the centroid until it reaches a steady state condition or the centroid position does not change again. For example, in Iteration 0, the shortest distance obtained at Airport A is 947165 which is located in the first centroid [6].

Table 5. below is the calculation of the shortest euclidean distance, a grouping of data is made by according to the position or location of the shortest centroid. Here is the data grouping:

| No. | Airport | C1     | C2     | C3     | C4     | C5     | Shortest Distance |
|-----|---------|--------|--------|--------|--------|--------|------------------|
| 1   | A       | 1      |        |        |        |        |                  |
| 2   | B       | 1      |        |        |        |        |                  |
| 3   | C       | 1      |        |        |        |        |                  |
| 4   | D       | 1      |        |        |        |        |                  |
| 5   | E       | 1      |        |        |        |        |                  |
| 6   | F       |        |        |        |        |        |                  |
| 7   | G       | 1      |        |        |        |        |                  |
| 8   | H       | 1      |        |        |        |        |                  |
| 9   | I       | 1      |        |        |        |        |                  |
| 10  | J       | 1      |        |        |        |        |                  |
| 11  | K       | 1      |        |        |        |        |                  |
| 12  | L       |        |        |        |        |        |                  |
| 13  | M       | 1      |        |        |        |        |                  |
| 14  | N       | 1      |        |        |        |        |                  |
| 15  | O       |        |        |        |        |        |                  |
| 16  | P       | 1      |        |        |        |        |                  |
| 17  | Q       |        |        |        |        |        |                  |
| 18  | R       |        |        |        |        |        |                  |
| 19  | S       | 1      |        |        |        |        |                  |

Based on the grouping of data based on the shortest euclidean distance on iteration 1, the temporary centroid position is obtained as above. In Cluster 1 there is only one airport, namely Airport F. At Cluster 2 there are 5 airports namely A Airport, B Airport, C Airport, D Airport, and O Airport. At Cluster 3 there are 5 namely E Airport, G Airport, H Airport, P Airport, and Q Airport. In Cluster 4 there are 6 airports, namely I Airport, J Airport, K Airport, L Airport, and S Airport. In Cluster 5 there are 2 airports, namely M airport and R airport.

After obtaining a temporary Cluster position, a new centroid calculation is carried out based on Table 6. is the calculation results of the temporary cluster position data grouping [1].
Table 6. Determination the new centroids for iteration 2

| No | Aircraft | Passenger | Cargo | Terminal area | Area | Runway | EBITDA | Revenue |
|----|----------|-----------|-------|---------------|------|--------|--------|---------|
| 1  | 80431    | 7609051   | 27169914 | 22793 | 165 | 3000   | 148289 | 318187  |
| 2  | 38259    | 4077311   | 18456462 | 18082 | 189 | 2698   | 56962  | 158113  |
| 3  | 20414    | 2040013   | 11253489 | 11944 | 122 | 2282   | 24483  | 72310   |
| 4  | 10038    | 494408    | 2882135  | 4281  | 154 | 2423   | -15888 | 20908   |
| 5  | 752      | 54601     | 56980   | 41850 | 437 | 2150   | 0      | 0       |

After going through the same steps in the first iteration namely the calculation of euclidean distance from each cluster and the calculation of the shortest euclidean distance, then proceed with the grouping of data based on the shortest euclidean distance on iteration 2. The temporary centroid position are in Cluster 1 there is only one airport, namely Airport F. In Cluster 2 there are 6 airports namely A Airport, B Airport, C Airport, D Airport, E Airport and O Airport. In Cluster 3 there are 5 namely G Airport, H Airport, N Airport, P Airport, and Q Airport. In Cluster 4 there are 2 airports, namely I Airport and J Airport. In Cluster 5 there are 5 airports, namely K airport, L airport, M airport, R airport and S airport.

After obtaining a temporary Cluster position, a new centroid calculation is carried out based on Table 7.

Table 7. Determination the new centroids for iteration 3

| No | Aircraft | Passenger | Cargo | Terminal area | Area | Runway | EBITDA | Revenue |
|----|----------|-----------|-------|---------------|------|--------|--------|---------|
| 1  | 80431    | 7609051   | 27169914 | 22793 | 165 | 3000   | 148289 | 318187  |
| 2  | 38259    | 4077311   | 18456462 | 18082 | 189 | 2698   | 56962  | 158113  |
| 3  | 20414    | 2040013   | 11253489 | 11944 | 122 | 2282   | 24483  | 72310   |
| 4  | 10038    | 494408    | 2882135  | 4281  | 154 | 2423   | -15888 | 20908   |
| 5  | 752      | 54601     | 56980   | 41850 | 437 | 2150   | 0      | 0       |

Table 8. below is the calculation of cluster distance by determining the shortest distance using the euclidean formula in iteration 3.

| Airport | C1 | C2 | C3 | C4 | C5 | The shortest distance |
|---------|----|----|----|----|----|----------------------|
| A       | 9451297 | 546574 | 9035423 | 14484569 | 18476219 | 546574 |
| B       | 9221113 | 789865 | 9276784 | 14742567 | 18736624 | 789865 |
| C       | 9625729 | 1076868 | 8926618 | 14274059 | 18246333 | 1076868|
| D       | 9989201 | 322476 | 8496450 | 13923668 | 17912294 | 322476 |
| E       | 12977003 | 3102853 | 5569434 | 10918068 | 14899362 | 3102853|
| F       | 1685 | 9991790 | 18484102 | 23894623 | 28781407 | 1685  |
| G       | 18323320 | 8359389 | 569862 | 5601648 | 9599347 | 569862 |
| H       | 18191259 | 8247763 | 882310 | 5708993 | 9700959 | 882310 |
| I       | 25229049 | 15282971 | 6843821 | 1337569 | 2694101 | 1337569|
| J       | 22560571 | 12629350 | 4265989 | 1337567 | 5319312 | 1337567|
| K       | 28009431 | 18096922 | 9699557 | 4157852 | 240182 | 240182 |
| L       | 28161910 | 18245647 | 9840701 | 4301188 | 319870 | 319870 |
| M       | 28192855 | 18270725 | 9855840 | 4321636 | 337635 | 337635 |
Based on the calculation of the shortest euclidean distance, a grouping of data is made by according to the position or location of the shortest centroid. Based on Table 9, the data grouping with the shortest euclidean distance in iteration 3, the final result is obtained in the cluster as below. This is because the centroid position on the iteration 2 and the iteration 3 is fixed or does not change, this condition indicates that the steady state condition has occurred which is the final cluster position requirement [6].

| Airport | C1     | C2     | C3     | C4     | C5     | The shortest distance |
|---------|--------|--------|--------|--------|--------|-----------------------|
| N       | 20933299 | 10957621 | 2481781 | 3075341 | 7055866 | 2481781               |
| O       | 9043718  | 2481671  | 10028885 | 15571786 | 19561997 | 2481671               |
| P       | 17879657 | 7889373  | 841187  | 6293288 | 10266008 | 841187                |
| Q       | 17027606 | 7046674  | 1711633 | 7204446 | 11175534 | 1711633               |
| R       | 28103365 | 18175960 | 9752553 | 4222533 | 247450    | 247450                |
| S       | 26891989 | 16956564 | 8526456 | 3001539 | 1016447  | 1016447               |

| No. | Airport | C1 | C2 | C3 | C4 | C5 |
|-----|---------|----|----|----|----|----|
| 1   | A       | 1  |    |    |    |    |
| 2   | B       | 1  |    |    |    |    |
| 3   | C       | 1  |    |    |    |    |
| 4   | D       | 1  |    |    |    |    |
| 5   | E       | 1  |    |    |    |    |
| 6   | F       | 1  |    |    |    |    |
| 7   | G       |    | 1  |    |    |    |
| 8   | H       |    | 1  |    |    |    |
| 9   | I       |    |    | 1  |    |    |
| 10  | J       |    |    |    | 1  |    |
| 11  | K       |    |    |    | 1  |    |
| 12  | L       |    |    |    |    | 1  |
| 13  | M       |    |    |    |    |    |
| 14  | N       |    |    |    |    | 1  |
| 15  | O       |    |    |    |    | 1  |
| 16  | P       |    |    |    |    |    |
| 17  | Q       |    |    |    |    |    |
| 18  | R       |    |    |    |    |    |
| 19  | S       |    |    |    |    |    |

Table 9. Data grouping iteration 3

Table 10. Initial cluster centers

|                | Cluster      |
|----------------|--------------|
|                | 1           | 2           | 3           | 4           | 5           |
| Zscore(Aircraft)| 0.71960     | 2.77375     | -1.01407    | 0.01198     | -1.04074    |
| Zscore(Passanger)| 0.97675    | 2.52655     | -0.97918    | 0.12666     | -1.01480    |
| Zscore(Cargo)   | 0.98850     | 2.06050     | -1.23161    | 1.16033     | -1.21780    |
| Cluster Membership | Cluster | Distance |
|-------------------|---------|----------|
| Cluster 1 | 5,000 |
| Cluster 2 | 1,000 |
| Cluster 3 | 6,000 |
| Cluster 4 | 6,000 |
| Cluster 5 | 1,000 |
| Valid | 19,000 |
| Missing | 0,000 |

**Table 11. Number of Cases in each Cluster**
3.4 Recapitulation of Clustering Results

Based on Table 13, the calculation of the cluster k-means algorithm the different results of clustering between the k-means manual algorithm and SPSS are obtained, this is because the weakness of the k-means algorithm in SPSS software is dependent on the initial centroid.

In this study the results used are the manual k-means algorithm because it is more suited to the airport situation. The cluster results are in cluster 1, there is only one airport, namely Airport F. In cluster 2 there are 6 airports, namely Airport A, Airport B, Airport C, Airport D, Airport E, Airport O. In cluster 3 there are 5 namely G Airport, H Airport, N Airport, P Airport, and Q Airport. In cluster 4 there are 2 airports namely Airport I and Airport J. In cluster 5 there are 5 airports, namely K Airport, L Airport, M Airport, R Airport and S Airport. Below is the recapitulation of k-means clustering result.

| Table 13. Recapitulation of clustering results |
|-----------------------------------------------|
| Manually | SPSS |
|--------- | ---- |
| F       | A    |
|         | B    |
|         | C    |
|         | D    |
|         | O    |
| A       | F    |
| B       |      |
| C       |      |
| D       |      |
| E       |      |
| O       |      |
| G       | K    |
| H       | L    |
| N       | M    |
| P       | R    |
| Q       | S    |
|         | I    |
| I       | Q    |
| J       | P    |
|         | G    |
|         | H    |
|         | J    |
|         | N    |
| K       | E    |
| L       |      |
| M       |      |
| R       |      |
| S       |      |

4. Conclusion

Based on the ANOVA results using SPSS, it can be concluded that there is an influence between variable aircraft movements, passenger movement, cargo, area, terminal area, runway, EBITDA and revenue to the clustering process, this is indicated by the F statistic value greater than F critical.
Based on the two-step analysis, the number of 5 clusters has been optimal, this can be seen from the silhouette coefficient value of 0.6 which is classified as good from the quality and strength of the cluster.

Based on the calculation of the cluster k-means algorithm the different results of clustering between the k-means manual algorithm and SPSS are obtained, this is because the weakness of the k-means algorithm in SPSS software is dependent on the initial centroid. In this study the results used are the manual k-means algorithm because it is more suited to the airport situation.

The cluster results are in cluster 1, there is only one airport, namely Airport F. In cluster 2 there are 6 airports, namely Airport A, Airport B, Airport C, Airport D, Airport E, Airport O. In cluster 3 there are 5 namely G Airport, H Airport, N Airport, P Airport, and Q Airport. In cluster 4 there are 2 airports namely Airport I and Airport J. In cluster 5 there are 5 airports, namely K Airport, L Airport, M Airport, R Airport and S. Airport.

5. References

[1] Agus Iriaanto 2009 Statistics Basic Concepts and Applications (Jakarta: Kencana Publishing)
[2] Ali H H 2017 Applications in Data Mining Vol. 6 No. 8 (India: International Journal of Science and Research) pp. 1577 – 1584
[3] Awangga R M 2019 Operational Management Activity Vol. 7 Issue 1 (Surabaya: EMITTER International Journal of Engineering Technology) pp. 44-54
[4] Chung Tae-Won 2015 Shipping and Logistics Vol. 31 Issue 1 (Amsterdam: Elsevier) pp. 85-108
[5] Raval U R 2017 Computer Science and Information Technology Vol.5 Issue 5 (India: IJCSMC) pp. 191-203
[6] Thakare Y S 2015 Electronics and Telecommunication Vol. 110 No. 11 (India: International Journal of Computer Applications) pp. 0975 – 8887
[7] Trenggonowati D L 2018 Industrial And Engineering Applications Vol. 218 No.7 (Banten: MATEC Web Conf.) pp. 04021