

## ABSTRAK

Nama : Muhamad Irfan Fadhilah  
NIM : 41518010067  
Program Studi : Teknik Informatika  
Judul Laporan Skripsi : Clusterisasi Kepadatan Penduduk di DKI Jakarta  
dengan Algoritma K-Means dan Agglomerative  
Cluster  
Pembimbing : Muhammad Rifqi, S.Kom, M.Kom

Indonesia menjadi salah satu negara dengan jumlah penduduk terbesar di dunia. Salah satu kota terbesar di Indonesia adalah DKI Jakarta yang merupakan ibu kota Indonesia, hal itu dibuktikan dengan meningkatnya jumlah penduduk di DKI Jakarta yang mengakibatkan peningkatan jumlah penduduk di DKI Jakarta dari tahun ke tahun. Sehingga populasi di DKI Jakarta semakin padat dan lahan-lahan pemukiman semakin berkurang. Untuk menghindari penumpukan penduduk di satu kelurahan maka diperlukan clustering kepadatan penduduk di setiap kelurahan di DKI Jakarta. Oleh sebab itu dilakukan clustering dengan menggunakan 2 algoritma K-Means dan Agglomerative. Tahap penelitian ini dilakukan dengan menggunakan data dari tahun 2017–2020 di DKI Jakarta kemudian dilakukan pencarian K terbaik dari kedua algoritma dan didapatkan hasil  $K = 3$ . Langkah selanjutnya dilakukan clustering menggunakan algoritma K-Means dan Agglomerative yang menghasilkan nilai RMSE dan MAE, untuk algoritma K-Means menghasilkan nilai RMSE 0,996 dan nilai MAE 0,735 sedangkan algoritma Agglomerative menghasilkan nilai RMSE 1.734 dan nilai MAE 1,536, jadi dapat disimpulkan algoritma K-Means lebih baik dari algoritma Agglomerative dengan jumlah RMSE yang lebih rendah dibanding algoritma Agglomerative.

**Kata Kunci** : Clustering, Data Mining, K-Means, Agglomerative, Kepadatan Penduduk

## ABSTRACT

Name : Muhamad Irfan Fadhilah  
NIM : 41518010067  
Study Program : Informatics Engineering  
Title Thesis : Clusterisasi Kepadatan Penduduk di DKI Jakarta  
dengan Algoritma K-Means dan Agglomerative  
Cluster  
Counsellor : Muhammad Rifqi, S.Kom, M.Kom

*Indonesia is one of the countries with the largest population in the world. one of the largest cities in Indonesia is DKI Jakarta which is the capital city of Indonesia, this is evidenced by the increasing population in DKI Jakarta which has resulted in an increase in the number of residents in DKI Jakarta from year to year. So that the population in DKI Jakarta is getting denser and residential areas are decreasing. To avoid population accumulation in one sub-district, population density clustering is needed in each sub-district in DKI Jakarta. Therefore, clustering is carried out using 2 K-Means and Agglomerative algorithms. This research phase was carried out using data from 2017 – 2020 in DKI Jakarta and then a search for the best K of the two algorithms was carried out and the results obtained  $K = 3$ . The next step was clustering using the K-Means and agglomerative algorithms which produced RMSE and MAE values, for the algorithm K-Means produces an RMSE value of 0.996 and an MAE value of 0.735 while the Agglomerative algorithm produces a RMSE value of 1,734 and an MAE value of 1.536, so it can be concluded that the K-Means algorithm is better than the Agglomerative algorithm with a lower RMSE number than the Agglomerative algorithm.*

**Keyword** : Clustering, Data Mining, K-Means, Agglomerative, Population Density