

**REKOMENDASI PEMASARAN KENDARAAN BERMOTOR RODA DUA
MENGGUNAKAN ALGORITMA K-MEANS CLUSTERING BERBASIS GEOSPASIAL**
Usep Tatang Suryadi^{*1}, Lina Meilawati.^{#2}

Program Studi Teknik Informatika, STMIK Subang
Jl. Marsinu No. 5 - Subang, Tlp. 0206-417853 Fax. 0206-411873
E-mail: usep_is_ugie@yahoo.co.id^{*1}, lina_meilawati@yahoo.com^{#2}

ABSTRAKSI

Pada tahun 2013, perbandingan jumlah alat transportasi kendaraan angkutan jenis umum dan luas daerah di Kabupaten Subang adalah 661:2052. Dengan jumlah kendaraan yang tidak layak guna sebanyak 396 kendaraan atau 60% dari total. Data KOPJAS pada 2018, jumlah trayek/rute angkutan umum sebanyak 420 trayek/rute, sedangkan jumlah trayek/rute elf sebanyak 227 trayek/rute. Jadi, total seluruh trayek/rute di Kabupaten Subang sebanyak 647 trayek/rute. Dengan kondisi tersebut jelas, masyarakat perlu kendaraan untuk bertransportasi. Dimana persebarannya sendiri tidak merata sehingga beberapa daerah hanya bisa diakses dengan kendaraan pribadi dan sebagian besar hanya bisa diakses oleh kendaraan roda dua.

Berdasarkan pemaparan di atas, prospek bisnis/pemasaran kendaraan roda dua di Kabupaten Subang sangat menjanjikan. Tentu harus didukung dengan pemasaran yang baik, diantaranya memilih daerah-daerah prioritas dengan melakukan analisis klasterisasi daerah persebaran kendaraan umum. Untuk kemudian dijadikan acuan prioritas pemasaran kendaraan roda dua. Penerapan teknik Data Mining metode Clustering menggunakan Algoritma K-Means Clustering pada data KOPJAS Kabupaten Subang dari 26 Kecamatan untuk melakukan klasterisasi persebaran Trayek di Kabupaten Subang sebagai bahan pertimbangan pemasaran kendaraan roda dua.

Berikutnya data ditabulasi, dilakukan pemilihan 2 centroid secara acak, untuk mendapatkan dua klaster. Selanjutnya menghitung besaran rasionalnya. Kemudian dilakukan pencarian nilai jarak baru ke centroid baru, penghitungan jarak berhenti ketika nilai rasio tidak berubah lagi, ditentukan minimal jarak dan kuadrat minimal jarak, sehingga didapatkan 4 kali iterasi. Hasilnya K-means Clustering mampu melakukan klasterisasi Kecamatan di Kabupaten Subang. Aplikasi yang digunakan dalam mengolah data tersebut yaitu Xampp, Php, MySql dan Notepad++.

Kata Kunci: *Clustering, Data Mining, K-Means, MySQL, PHP, Transportasi.*

1. Pendahuluan

1.1. Latar Belakang

Alat transportasi yang banyak digunakan oleh manusia salah satunya adalah sepeda motor. Secara utilitas, sepeda motor merupakan alat transportasi yang fleksibel, irit, dan harganya terjangkau. Oleh karena itu banyak masyarakat di Indonesia utamanya di Pagaden yang menggunakan alat transportasi sepeda motor.

Secara kuantitas maupun kualitas, kondisi angkutan umum dikota Subang belum mampu mengakomodir mobilitas masyarakat Subang, hal ini disebabkan keterbatasan trayek/rute dari angkutan kota yang belum menjangkau kawasan padat penduduk secara keseluruhan yang mendorong masyarakat lebih memilih untuk menggunakan sarana transportasi pribadi seperti motor, dibandingkan angkutan umum. Dari data yang diperoleh jumlah trayek/rute angkutan umum sebanyak 420 trayek/rute, sedangkan jumlah trayek/rute elf sebanyak 227 trayek/rute. Jadi, total seluruh trayek/rute

di Kabupaten Subang sebanyak 647 trayek/rute. Dilihat dari fungsional sepeda motor tersebut menyebabkan terjadinya peningkatan pembelian sepeda motor di dealer secara kredit.

Kredit ialah, sebuah kepercayaan seseorang atau lembaga yang memiliki barang, uang, atau jasa untuk dikreditkan (kreditur) kepada seseorang (debitur) yang harus dibayar dalam waktu yang telah di sepakati atau ditentukan (Kurniawan, 2010).

Perkembangan Teknologi Informasi sudah sedemikian pesat. Perkembangan yang pesat tidak hanya teknologi perangkat keras dan perangkat lunak saja, tetapi metode komputasi juga ikut berkembang. Salah satu metode komputasi yang cukup berkembang saat ini adalah metode Sistem Pengambilan Keputusan (*Decisions Support System*). Konsep Sistem Pendukung Keputusan (SPK) / Decision Support Sistem (DSS) pertama kali diungkapkan Morton (1970) dengan istilah *Management Decision System*. Istilah SPK (Sistem Pengambilan Keputusan) mengacu pada suatu sistem yang memanfaatkan dukungan komputer dalam proses pengambilan keputusan.

Menurut (Turban,2009), mendefinisikan sistem pendukung keputusan merupakan suatu sistem informasi yang ditujukan untuk membantu manajemen dalam memecahkan masalah yang dihadapinya. Dalam pengambilan keputusan guna memasarkan produk, tentu terlebih dulu memilih daerah atau pasar yang berpotensi untuk menaikkan penjualan, salah satu algoritma yang mampu untuk klasterisasi daerah pemasaran adalah algoritma K-Means.

Algoritma K-Means sangat terkenal karena kemudahan dan kemampuannya untuk mengklasifikasi data besar dan *outlier* dengan sangat cepat (Kusumadewi, 2009).

Pada penelitian yang dilakukan oleh Apriadi Bahar (2016), yang berjudul Penentuan Strategi Penjualan Alat-Alat Tattoo Di Studio Sonyx Tattoo Menggunakan Metode K-Means Clustering., Maka dapat ditarik kesimpulan yaitu sistem penentuan strategi ini dapat menerapkan metode *data mining* algoritma *K-Means Clustering* yaitu dengan mempartisi data yang ada ke dalam bentuk satu atau lebih *Cluster* atau dapat dikatakan memiliki tujuan untuk membagi data menjadi beberapa kelompok dalam menentukan pemasaran kendaraan bermotor roda dua.

12 Identifikasi Masalah

Masalah yang dirumuskan dalam penelitian ini adalah sebagai berikut :

1. Diperlukan sistem yang dapat membantu menentukan prioritas daerah pemasaran kendaraan bermotor.
2. Diperlukan metode clustering untuk pengelompokan daerah pemasaran kendaraan bermotor yang efisien.
3. Belum diketahui kemampuan algoritma K-Means Clustering dalam mengelompokkan daerah berdasarkan data KOPJAS per-Kecamatan di Kabupaten Subang.

13 Tujuan

Tujuan yang diperoleh dari penelitian ini adalah membangun sistem yang mampu merekomendasikan pemasaran kendaraan bermotor roda dua menggunakan algoritma *K-Means*, kemudian mengetahui kemampuan algoritma *K-Means Clustering* dalam mengelompokkan daerah pemasaran berdasarkan data KOPJAS per-Kecamatan di Kabupaten Subang

14 Manfaat

Manfaat yang diperoleh pada penelitian ini adalah membantu pihak dealer dengan aplikasi tersebut untuk memudahkan proses menentukan strategi pemasaran yang tepat sasaran sesuai dengan *cluster* yang telah ditentukan. Serta dapat menambah kepustakaan penelitian komputerisasi pada bidang pemasaran untuk dijadikan sebagai referensi pada penelitian berikutnya di kemudian hari.

15. Metodologi Penelitian

Dalam penelitian ini menggunakan beberapa metode yaitu :

1. Kepustakaan.

Metode pengumpulan data yaitu mengumpulkan data baik dari buku, artikel, jurnal, situs dan lain sebagainya. Yaitu tentang referensi-referensi metode K-Means, sistem pendukung keputusan (SPK), data-data yang dapat mendukung dalam penelitian ini.

2. Observasi.

Metode observasi/survei yaitu mencari informasi tentang proses dalam penjualan sepeda motor kepada pelanggan dan syarat-syarat yang digunakan dalam pengajuan kredit sepeda motor pada PT Trijaya Motor.

3. Wawancara

Wawancara (Interview) dengan melakukan percakapan secara langsung kepada nara sumber dan pihak-pihak yang bisa memberikan informasi serta kendala-kendala yang sering muncul saat pengambilan keputusan. Adapun yang akan di tanyakan adalah data apa saja yang harus di serahkan konsumen untuk mendapatkan kredit sepeda motor dan bagaimana cara kerja sistem yang sedang berjalan.

2 Tinjauan Pustaka

21. Data Mining

Data mining adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan didalam database. *Data mining* adalah proses yang menggunakan teknik statistika, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar (Turban, dkk. 2005).

Data mining dibagi menjadi beberapa kelompok berdasarkan tugas yang dapat dilakukan, yaitu, Deskripsi, Estimasi, Prediksi, Klasifikasi, Pengklusteran, dan Asosiasi (Larose, 2005)

22. Algoritma K-Means

K-Means merupakan algoritma *clustering* yang berulang-ulang. Algoritma Kmeans dimulai dengan pemilihan secara acak K, K disini merupakan banyaknya *cluster* yang ingin dibentuk. Kemudian tetapkan nilai-nilai K secara *random*, untuk sementara nilai tersebut menjadi pusat dari *cluster* atau biasa disebut dengan *centroid*, *mean* atau “*means*”. Hitung jarak setiap data yang ada terhadap masingmasing *centroid* menggunakan rumus *Euclidian distance* hingga ditemukan jarak yang paling dekat dari setiap data dengan *centroid*. *K-Means* merupakan salah satu metode *Clustering* non hirarki yang berusaha mempartisi data yang ada ke dalam bentuk satu atau lebih *Cluster* atau dapat dikatakan memiliki tujuan untuk membagi data menjadi beberapa kelompok. Algoritma *K-Means* ditemukan oleh beberapa orang yaitu Lloyd (1957, 1982), Forgey (1965), Friedman and Rubin (1967), and McQueen (1967). Ide dari *Clustering* pertama kali ditemukan oleh Lloyd pada tahun 1957, namun hal tersebut baru dipublikasi pada tahun 1982. Pada tahun 1965, Forgey juga mempublikasi teknik yang sama sehingga terkadang dikenal sebagai Lloyd-Forgey pada beberapa sumber.

Berikut ini algoritma *K-Means* mempartisi suatu *dataset* ke dalam *Cluster-Cluster* beserta contoh kasusnya :

1. Algoritma menerima jumlah *Cluster* untuk mengelompokkan data ke dalamnya, dan *dataset* yang akan di *Cluster* sebagai nilai input.
2. Algoritma kemudian membuat sebanyak *K Cluster* awal (*K* = jumlah *Cluster* yang diperlukan) dari *dataset*, sekaligus memilih *K record* data secara acak (*random*) dari *dataset*.
3. Algoritma *K-Means* menghitung rata-rata aritmatika dari setiap *Cluster* yang dibentuk dalam *dataset*. Rata-rata dari suatu *Cluster* adalah rata-rata dari semua *record* yang terdapat di dalam *Cluster* tersebut. Karena di dalam semua *K Cluster* pertama hanya ada satu *record*

- maka rata-ratanya adalah rata-rata *record* tersebut. Rata-rata dari suatu *record* adalah kumpulan nilai-nilai yang membangun *record* tersebut.
4. Selanjutnya, *K-Means* mengirim *K record* di dalam *dataset* ke hanya salah satu dari *Cluster* awal.
 5. *K-Means* mengkalkulasi ulang rata-rata aritmatika dari semua *Cluster*. Rata-rata dari suatu *Cluster* adalah rata-rata dari semua *record* di dalam *Cluster* tersebut.
 6. *K-Means* mengirimkan lagi setiap *record* di dalam dataset ke hanya salah satu dari *Cluster*-*Cluster* baru yang terbentuk. Suatu *record* atau titik-titik data dilewatkan ke *Cluster* terdekat, seperti sebelumnya.
 7. Langkah-langkah sebelumnya diulang sampai terbentuk *Cluster*-*Cluster* stabil dan prosedur *K-Means* selesai. *Cluster* stabil terbentuk ketika iterasi atau perulangan dari *K-Means* tidak membuat *Cluster* baru sebagai pusat *Cluster* atau nilai rata-rata aritmatika dari semua *Cluster* baru sama dengan *Cluster* lama. Terdapat beberapa teknik untuk menentukan kapan suatu *Cluster* stabil terbentuk atau kapan algoritma *K-Means* berakhir.

Proses pengelompokan data ke dalam suatu *Cluster* dapat dilakukan dengan cara menghitung jarak terdekat dari suatu data ke sebuah titik pusat atau *centroid*. Pembaharuan suatu titik *centroid* dapat dilakukan dengan rumus berikut :

$$\mu_k = \frac{1}{N_k} \sum_{q=1}^{N_k} x_q$$

Dimana:

μ_k = titik *centroid* dari *Cluster* ke-K

N_k = banyaknya data pada *Cluster* ke-K

x_q = data ke-q pada *Cluster* ke-K

3. Analisa

31 Deskripsi Sistem

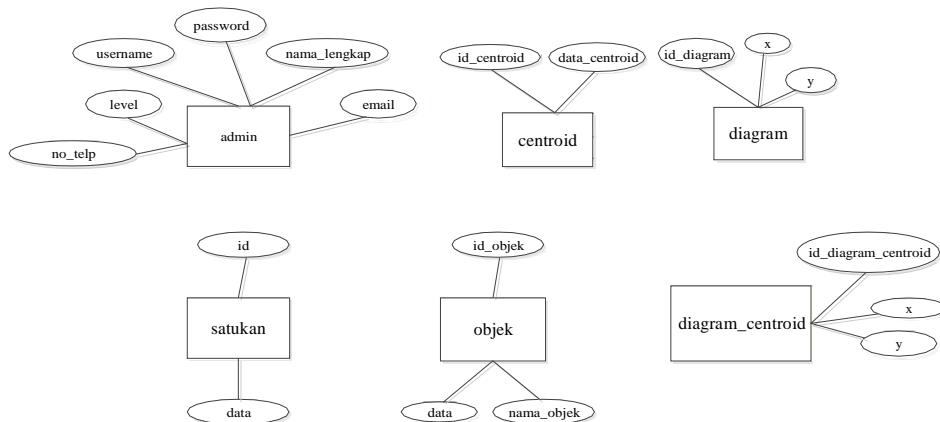
Pembuatan program aplikasi pada penelitian ini mengikuti langkah-langkah pengembangan perangkat lunak model *waterfall*. Model ini dipilih dengan alasan untuk membangun sistem ini dibutuhkan beberapa tahap yang berbeda, yang diawali dengan analisa kebutuhan, perancangan, implementasi dan pengujian sistem.

Pada penelitian ini sistem yang akan dibangun membutuhkan data input yang berisi data Trayek/rute angkutan umum terdiri dari beberapa atribut seperti No, Nama Kecamatan, Angkot/Minibus, Elp/Mikrobus. Proses yang digunakan untuk mengolah data input adalah teknik *clustering* dengan algoritma *K-means*. Namun dalam penelitian ini hanya 26 data kecamatan yang memiliki trayek/rute angkutan umum yang diambil sebagai sample dalam proses perhitungan k-means clustering dengan berdasarkan data yang didapat dari KOPJAS.

32 Model Proses

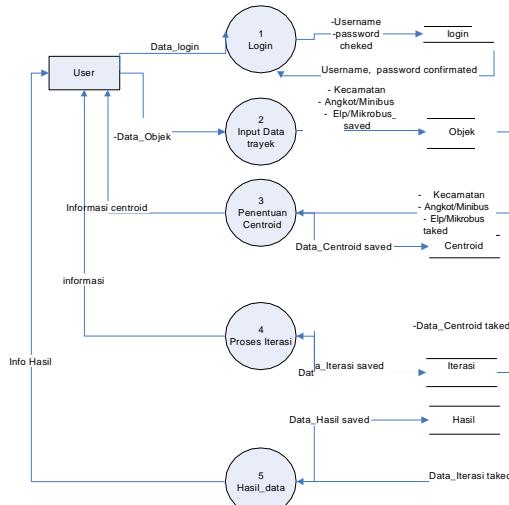
Menurut Brady dan Loonam (2010), Entity Relationship diagram (ERD) merupakan teknik yang digunakan untuk memodelkan kebutuhan data dari suatu organisasi, biasanya oleh System Analyst dalam tahap analisis persyaratan proyek pengembangan system. Sementara seolah-olah teknik diagram atau alat peraga memberikan dasar untuk desain database relasional yang mendasari sistem informasi yang dikembangkan. ERD bersama-sama dengan detail pendukung merupakan model data yang pada gilirannya digunakan sebagai spesifikasi untuk database.

Berikut adalah diagram entitas pada Rekomendasi Pemasaran Kendaraan Bermotor Berdasarkan Sebaran Angkutan Umum Menggunakan Algoritma K-Means Clustering Di Kabupaten Subang,



Gambar 1. Entitas diagram

Data Flow Diagram adalah diagram yang digunakan untuk memodelkan sistem secara logik. penulis menggambarkan sistem dari keseluruhan mulai dari aktifitas user menginputkan apapun informasi yang diingat tentang sistem rekomendasi pemasaran kendaraan bermotor sampai admin melakukan olah data yang ada pada sistem. Untuk user hanya bisa mencari dan melihat hasil perhitungan k-means hingga mendapatkan iterasi yang diproses oleh sistem. Dan admin dapat melakukan olah data baik input data sebaran angkutan umum serta dapat melihat data-data didalamnya.



Gambar 2. DFD sistem rekomendasi pemasaran

4. Hasil dan Pembahasan

Pada penelitian ini, data input yang berisi data Trayek/rute angkutan umum terdiri dari beberapa atribut seperti No, Nama Kecamatan, Angkot/Minibus, Elp/Mikrobus.

Tabel 1. Data Trayek Kabupaten Subang

No	Kecamatan	angkot/minibus	elp/mikrobus
1	Binong	5	11
2	Ciasem	1	15
3	Cilater	18	4
4	Cibogo	41	3
5	Cijambe	26	18
6	Cipeundeuy	11	5
7	Cipunagara	5	0
8	Cisalak	8	15
9	Compreng	14	0
10	Dawuan	29	3
11	Jalan Cagak	32	9
12	Kalijati	29	13
13	Kasomalang	1	5
14	Pabuaran	2	26
15	Pagaden Barat	17	0
16	Pagaden	25	23
17	Pamanukan	1	19
18	Patokbeusi	1	12
19	Purwadadi	21	3
20	Pusakajaya	6	5
21	Pusakanegara	0	5
22	Sadangserang	1	0
23	Sagalaherang	6	3
24	Subang	112	18
25	Tambakdahan	0	2
26	Tanjungsiang	9	10

Kemudian tahap selanjutnya dilakukan proses perubahan data, tujuannya adalah agar data di atas dapat diolah dengan menggunakan metode *K-Means Clustering*, maka data yang berjenis data kategorikal seperti Kecamatan harus diinisialisasikan terlebih dahulu dalam bentuk angka/numerikal.

Tabel 2. Inisialisasi Kecamatan di Kabupaten Subang

Kecamatan	Inisial
Binong	1
Ciasem	2
Ciater	3
Cibogo	4
Cijambe	5
Cipeundeuy	6
Cipunagara	7
Cisalak	8
Compreng	9
Dawuan	10
Jalan Cagak	11
Kalijati	12
Kasomalang	13
Pabuaran	14
Pagaden Barat	15
Pagaden	16
Pamanukan	17
Patokbeusi	18
Purwadadi	19
Pusakajaya	20
Pusakanegara	21
Sadangserang	22
Sagalaherang	23
Subang	24
Tambakdahan	25
Tanjungsiang	26

Selain kecamatan, cluster juga termasuk ke dalam jenis data nominal sehingga perlu diinisialisasikan ke dalam bentuk angka. Setelah semua data Anggota KOPJAS per-Kecamatan ditransformasikan ke dalam bentuk angka, maka data-data tersebut telah dapat dikelompokan dengan menggunakan algoritma *K-Means Clustering*.

$$d_{ij} = \sqrt{\sum_{k=1}^p \{x_{ik} - x_{jk}\}^2}$$

Dimana :

d_{ij} = Jarak objek antara objek i dan j

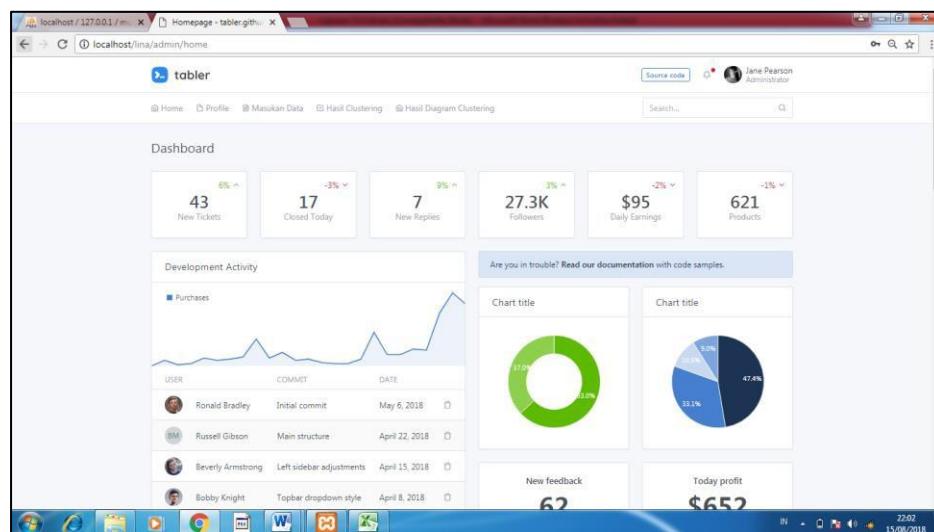
P = Dimensi data

X_{ik} = Koordinat dari objek i pada dimensi k

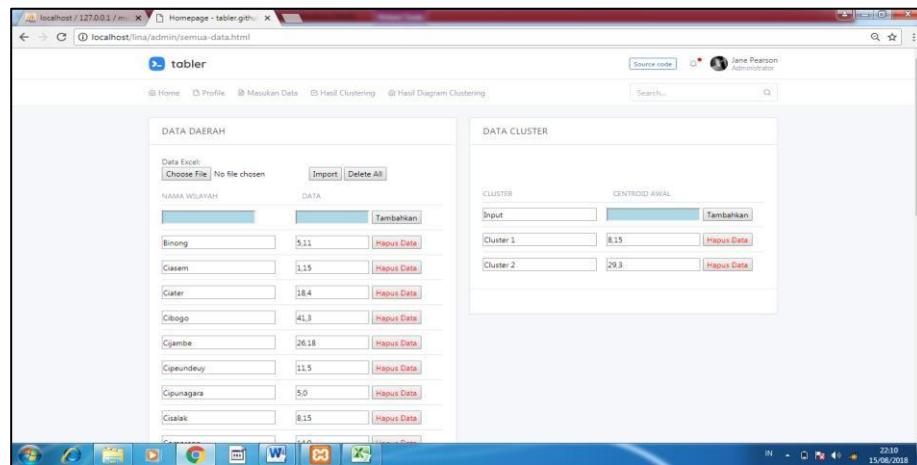
X_{jk} = Koordinat dari objek j pada dimensi k

4.1 Implementasi

Hasil tampilan pembuatan aplikasi Sistem rekomendasi pemasaran kendaraan bermotor roda dua berdasarkan sebaran angkutan umum adalah sebagai berikut:



Gambar 2. Halaman Utama Admin



Gambar 3. Halaman input masukkan data

Jika admin telah input masukkan data maka selanjutnya anda masuk ke menu hasil clustering untuk mengetahui hasil dan iterasinya.

OBJSK	Proses Clustering "				
	ITERASI 1		ITERASI 2		ITERASI 3
	DATA_1	DATA_2	CLUSTER_1	CLUSTER_2	
Binong	5	11	OK	Null	
Ciasem	1	15	OK	Null	
Clater	18	4	Null	OK	
Cibogo	41	3	Null	OK	
Cijambe	26	18	Null	OK	
Cipeundeuy	11	5	OK	Null	
Cipunagara	5	0	OK	Null	
Cisalak	8	15	OK	Null	
Compreng	14	0	Null	OK	
Dawuan	29	3	Null	OK	
Total	27	9	Null	OK	

Gambar 4. Halaman data iterasi ke-1

CLUSTER_1	
6.6842105263158	7.3684210526316
CLUSTER_2	
42	12.428571428571

Gambar 5. Halaman hasil centroid

5. Simpulan

Hasil dari penelitian ini dapat disimpulkan:

1. Sistem yang dibangun dapat dijadikan pendukung keputusan pengambil kebijakan dalam menyusun strategi memasarkan kendaraan bermotor, berdasarkan pengelompokkan/klasterisasi banyaknya trayek perkecamatan di Kabupaten Subang.
2. Metode *K-Means* menghasilkan klaster yang ideal. Dimana jumlah anggota klasternya tidak ada yang tunggal.

Pustaka

Akses kendaraan umum di kota subang : https://id.m.wikipedia.org/wiki/kabupaten_subang. Diakses pada 16 juli 2018 pukul 19.29.

Ali, Utsman. *Pengertian Diagram Konteks dan Data Flow Diagram (DFD)*. [online] : <http://www.pengertianpakar.com/2017/04/pengertian-diagram-konteks-dan-data-flow-diagram-dfd.html#> . Diakses pada 10 Agustus 2018 pukul 09.40.

Ardi, Ardiansyah. *Keterkaitan MySQL dan SQL*. [online] : <https://beritatenologi.blogspot.com/2017/10/keterkaitan-mysql-dan-sql.html> . Diakses pada 02 Agustus 2018 pukul 08.10.

Bahar, Apriadi. *Penentuan Strategi Penjualan Alat-Alat Tattoo Di Studio Sonyx Tattoo Menggunakan Metode K-Means Clustering*. Jurusan Teknik Informatika Fakultas Teknik Universitas Halu Oleo Kendari 2016. Diakses pada 27 juli 2018 pukul 22.27.

Darmi, Yulia dkk. *Penerapan Metode Clustering K-Means dalam Pengelompokan Penjualan Produk*. Program Studi Teknik Informatika Universitas Muhammadiyah Bengkulu Jl. Bali, Kel. Kampung Bali, Kec. Teluk Segara, Kota Bengkulu, Bengkulu 38119. Diakses pada 03 Agustus 2018 pukul 19.49.

Firman, Astria dkk. *Sistem Informasi Perpustakaan online berbasis web*. Jurusan Teknik Elektro Fakultas Teknik UNSRAT. Diakses pada 02 Agustus 2018 pukul 07.11.

Laila, Nur dkk. *Sistem Informasi Pengolahan Data Inventory Pada Toko Buku Studi Cv. Aneka Ilmu Semarang*. [online] : <https://media.neliti.com/media/publications/140910-ID-none.pdf> . Diakses pada 02 Agustus 2018 pukul 20.08.

Purwanto, Yusuf Eko dkk. *Pembangunan Sistem Informasi Pengolahan Perpustakaan Berbasis Website Pada Sekolah Dasar Negeri Jetis Lor 1 Kabupaten Pacitan*. [online] : <http://download.portalgaruda.org/article.php?article=132281&val=4926> . Diakses pada 01 Agustus 2018 pukul 21.01.

Ratna, Adis Lena Kusuma. *Pengertian PHP dan MySQL* [online] : <http://ilmuti.org/wp-content/uploads/2014/05/Adis-Lena-Kusuma-ratna-Pengertian-PHP-dan-MySQL.pdf> . Diakses pada 01 Agustus 2018 pukul 20.50.

Rabani, Aden. *Pengertian dan macam-macam software program database server*. Universitas Gunadarma. [online] : <https://adenrabani.wordpress.com/2014/04/10/pengertian-macam-macam-software-program-database-server/>. Diakses pada 02 Agustus 2018 pukul 07.41.

Supriatna, Encep. *Pengertian ERD menurut Bredy dan Loonam.* [online] : <https://encepsupriatna.files.wordpress.com/2009/02/entity-relationship-diagram.pdf> . Diakses pada 10 Agustus 2018 pukul 09.25.

Syahminan, Surya Adiatmaja Permana. *Perancangan Sistem Pendukung Keputusan Calon Pembeli Kredit Sepeda Motor Menggunakan K-Means Clustering.* Jurnal. Teknik Informatika Universitas Kanjuruhan Malang. [online].<https://media.neliti.com/media/publications/184634-ID-pembuatan-aplikasi-monitoring-dan-tracki.pdf>. Diakses pada 4 juni 2018 pukul 20.22.

Ufhie,Rofi. *Pengertian dan jenis database.* [online] : rofhiah.blogspot.com/2014/09/pengertian-dan-jenis-database-server.html. Diakses pada 02 Agustus 2018 pukul 08.20.

Usep Tatang Suryadi dan Siti Nurazizah. *Sistem Pendukung Keputusan Untuk Menentukan Strategi Marketing Kampus Menggunakan Metode K-Means Clustering.* Sekolah Tinggi Manajemen Informatika dan Komputer Subang. Diakses pada 10 Agustus 2018 pukul 09.18.