

Perancangan dan Implementasi Metode K-Means Clustering pada Aplikasi Pengiriman Barang Jadi (Studi Kasus pada Salah Satu Perusahaan Manufaktur Otomotif di Jakarta)

Design and Implementation of the K-Means Clustering Method in the Application of Shipment of Finished Goods

Intan Rinjani¹, Ulil Hamida²

^{1,2}Sistem Informasi Industri Otomotif, Politeknik STMI, Jakarta
Email: rinjani.intan19@gmail.com, ulil.hamida@gmail.com

ABSTRAK

Pengiriman barang jadi ke beberapa pelanggan menjadi tanggung jawab Bagian Delivery. Jalur pengiriman tersebut dibagi oleh Bagian Delivery menjadi beberapa jalur. Pembagian jalur yang ditetapkan sering mengakibatkan jarak antar sopir satu dengan sopir yang lainnya tidak merata. Untuk mengatasi masalah tersebut diperlukan suatu metode pembagian jalur pengiriman barang jadi. Metode yang digunakan yaitu Metode K-Means Clustering. Metode K-Means Clustering merupakan salah satu metode pengklasteran yang mempartisi data ke dalam kluster-kluster sehingga data yang memiliki kesamaan berada pada satu kluster yang sama. Metode tersebut kemudian digunakan dalam aplikasi pengiriman jadi. Metodologi pengembangan aplikasi yang digunakan adalah waterfall. Perancangan navigasi yang ada pada sistem menggunakan analisis perancangan Windows Navigation Diagram, pemodelan data menggunakan Entity Relationship Diagram dan menggunakan tools pemodelan Unified Modelling Language (UML) yaitu usecase diagram, activity diagram, dan deployment diagram sebagai pemodelan sistem. Sistem informasi ini dibangun menggunakan CodeIgniter 3.1.4 sebagai framework dan MariaDB 10.3.16 sebagai perangkat lunak manajemen basis data. Dalam penelitian ini metode K-Means Clustering terbukti dapat meminimalisir perbedaan jarak antar sopir dan PIC satu dengan sopir dan PIC yang lain.

Kata kunci: *Metode K-Means Clustering; Codeigniter; Mariadb; Jalur Pengiriman; Otomotif; Sistem Informasi*

ABSTRACT

Shipping process of finished goods to several customers is the responsibility of the delivery division. This shipping lane is divided by the delivery division into several lanes. The allocation of the specified path often results in the uneven distance between one driver and another. To overcome this problem, a method of dividing the shipping path of finished goods is needed. The method used was the K-Means Clustering Method. It is a clustering method that partitioned data into clusters so that data that has similarities are in the same cluster. This method is then used in a finished delivery application. The application development methodology used is the waterfall. The navigation design in the system uses Windows Navigation Diagram design analysis, data modeling uses Entity Relationship Diagrams and uses Unified Modeling Language (UML) modeling tools, namely use-case diagrams, activity diagrams, and deployment diagrams as system modeling. This information system was built using CodeIgniter 3.1.4 as a framework and MariaDB 10.3.16 as a database management software. In this study, the K-Means method Clustering is proven to be able to minimize the difference in distance between one driver and PIC and another driver and PIC.

Keywords : *K-Means Clustering Method; Codeigniter; MariaDB; Delivery Line; Automotive; Information Systems*

1. PENDAHULUAN

Pada proses pengiriman barang terdapat jalur-jalur pengiriman yang sudah dibuatkan oleh bagian *Delivery*. Jalur pengiriman ditulis setiap harinya pada papan jalur pengiriman dan akan disalin ke dalam laporan jalur pengiriman dalam bentuk file kertas, hal ini membuat proses pembuatan laporan pengiriman yang membutuhkan waktu yang lama dan proses pengerjaan yang berulang. Pemilihan jalur pengiriman dibuat sesuai dengan posisi perusahaan yang ditujukan. Apabila dalam satu jalur pengiriman terdapat beberapa perusahaan yang dilewati maka dapat digabungkan dalam satu pengiriman.

Pengiriman dilakukan ke beberapa perusahaan yang tersebar di daerah Jabodetabek, Banten, dan Karawang sehingga tidak dapat jika hanya ditangani oleh satu sopir dan satu PIC saja, maka harus dilakukan pembagian jalur pengiriman yang dilakukan oleh sopir dan PIC sesuai dengan daerah yang dituju. Saat ini pembagian jalur pengiriman masih dilakukan secara manual, yaitu dibuatkan jadwal sesuai dengan pemikiran pimpinan sub bagian *Delivery*. Kelemahan dari pembagian secara manual di antaranya pembagian jarak yang tidak sama, beban sopir yang tidak merata, dan informasi jalur pengiriman masih ditulis dalam papan tulis dan laporan pengiriman dalam file tertulis.

Metode *clustering* yang cukup dikenal adalah *K-Means Clustering*. *K-Means* merupakan metode pengklasteran yang memisahkan data ke dalam kelompok yang berbeda. *K-Means* mampu meminimalkan rata-rata jarak setiap data ke klasternya referensi [1]. *K-Means Clustering* merupakan salah satu metode pembagian data ke dalam *cluster-cluster* sehingga data yang memiliki kesamaan berada pada satu *cluster* yang sama dan data yang memiliki ketidaksamaan berada pada *cluster* yang lain seperti [2]. Metode ini telah banyak digunakan untuk berbagai permasalahan seperti yang telah dilakukan oleh [1],[2],[3],[4].

2. METODE DAN DATA

2.1 Data Mining

Referensi [5] menyatakan bahwa terdapat beragam pengertian mengenai data mining, yaitu:

- Penguraian (yang tidak sederhana) dari sekumpulan data menjadi informasi yang memiliki potensi secara implisit (tidak nyata/jelas) yang sebelumnya tidak diketahui.
- Penggalan dan analisis, dengan menggunakan peranti otomatis atau semi otomatis, dari sejumlah besar data yang bertujuan untuk menemukan pola yang memiliki arti
- Data mining merupakan bagian dari *knowledge discovery* dalam *database* (KDD).

Data mining berasal dari irisan berbagai disiplin ilmu pengetahuan yang meliputi: *machine learning/pattern recognition*, kecerdasan buatan, dan sistem basis data, peranan data mining meliputi:

- *Classification*
- *Clustering*
- *Association Rule Discovery*
- *Sequential Pattern Discovery*
- *Regression*
- *Deviation Detection*

2.2 Metode K-Means Clustering

Algoritma *k-means* adalah algoritma yang mempartisi data ke dalam *cluster-cluster* sehingga data yang memiliki kesamaan berada pada satu *cluster* yang sama dan data yang memiliki ketidaksamaan berada pada *cluster* yang lain [2]. Referensi [2] menyatakan, secara lebih jelas algoritma *k-means* adalah sebagai berikut:

- a. Menentukan *k* sebagai jumlah *cluster* yang ingin dibentuk.
- b. Membangkitkan nilai *random* untuk pusat *cluster* awal (*centroid*) sebanyak *k*
- c. Menghitung jarak setiap data *input* terhadap masing-masing *centroid* menggunakan rumus jarak *Euclidean* (*Euclidean Distance*) hingga ditemukan jarak yang paling dekat dari setiap data dengan *centroid* seperti pada persamaan (1). Berikut adalah persamaan *Euclidean Distance*:

$$d(x_i, \mu_j) = \sqrt{\sum (x_i - \mu_j)^2} \quad (1)$$

Keterangan:

x_i : data kriteria,

μ_j : *centroid* pada *cluster* ke-*j*

- a. Mengklasifikasikan setiap data berdasarkan kedekatannya dengan *centroid* (jarak terkecil).
- b. Memperbaharui nilai *centroid*. Referensi [4] menyatakan, nilai *centroid* baru di peroleh dari rata-rata *cluster* yang bersangkutan dengan menggunakan rumus (2):

$$C_k = \frac{1}{n_k} \sum d_i \quad (2)$$

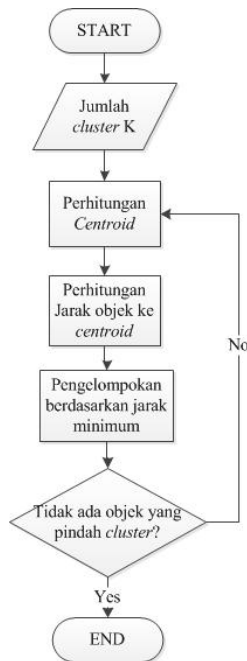
Keterangan:

n_k = jumlah data dalam *cluster* *k*

d_i = jumlah dari nilai jarak yang masuk dalam masing-masing *cluster*.

1. Melakukan perulangan dari langkah 2 hingga 5 sampai anggota tiap *cluster* tidak ada yang berubah.

Jika langkah terakhir telah terpenuhi, maka nilai pusat *cluster* (μ_j) pada iterasi terakhir akan digunakan sebagai parameter untuk menentukan klasifikasi data. Dapat dilihat pada Gambar 1.



Gambar 1 Flowchart Metode K-Means Clustering

2.3 Standar Deviasi

Rumus untuk standar deviasi digunakan untuk mengetahui penyimpangan antara data sampel dengan rata-ratanya. Secara definisi, standar deviasi adalah salah satu atau bagian dari analisis statistik yang digunakan untuk mengetahui sebaran data yang ada pada data sampel. Semakin besar nilai standar deviasi yang didapatkan maka menunjukkan semakin besar juga penyimpangannya, sebaliknya semakin kecil nilai deviasinya maka semakin kecil juga penyimpangannya [6].

Standar Deviasi dan Varians Salah satu teknik statistik yang digunakan untuk menjelaskan homogenitas kelompok. Varians merupakan jumlah kuadrat semua deviasi nilai-nilai individual terhadap rata-rata kelompok. Sedangkan akar dari varians disebut dengan standar deviasi atau simpangan baku, seperti [7]. Rumus:

$$s = \sqrt{\sum (x_i - \bar{x})^2} \quad (3)$$

Keterangan :

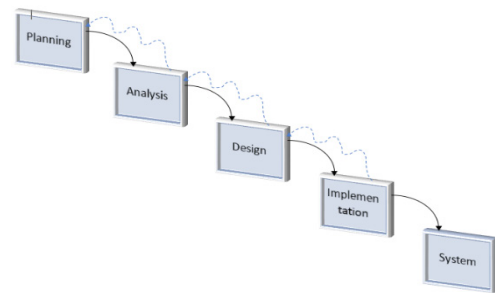
S = Deviasi Data

x_i = data ke i

\bar{x} = nilai rata-rata

2.4 Metode Pengembangan Sistem Waterfall

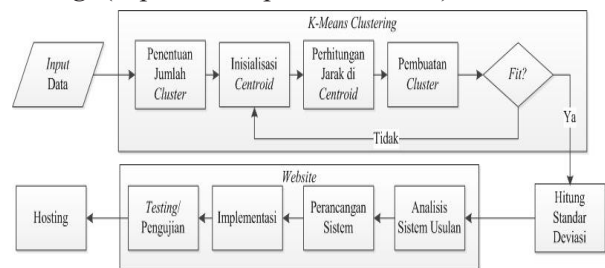
Referensi [8] menyatakan, terdapat 4 langkah dalam metode *waterfall* (dapat dilihat pada Gambar 2), adalah



Gambar 2. Model Waterfall

- Planning** (Perencanaan): pada tahap ini ditentukan secara detail rencana kerja yang harus dikerjakan, durasi yang diperlukan masing-masing tahap, sumber daya manusia, perangkat lunak, dokumentasi, perangkat keras, maupun *financial* diestimasi.
- Analysis** (Analisis): Tahap kedua, adalah tahap analisis, yaitu tahap dimana kita berusaha mengenali segenap permasalahan yang muncul pada pengguna dengan mendekomposisi dan merealisasikan komponen-komponen sistem.
- Design** (Perancangan): Tahapan mengubah kebutuhan yang masih berupa konsep menjadi spesifikasi sistem yang nyata untuk diimplementasikan.
- Implementation** (Implementasi): Tahap implementasi, dimana kita mengimplementasikan perancangan sistem ke situasi nyata atau desain harus diterjemahkan ke dalam bentuk mesin yang bisa dibaca.

Secara umum, metode penelitian yang digunakan untuk perancangan dan implementasi metode *K-Means Clustering* pada pengiriman barang jadi dapat dilihat pada Gambar 3. Tahapan proses pada penelitian ini dimulai dari data *input*, penentuan jumlah *cluster*, inisialisasi *centroid*, perhitungan jarak setiap data terhadap masing-masing *centroid*, mengelompokkan data berdasarkan jarak terdekat antara data dengan *centroid*, memperbaharui nilai *centroid*, hitung Standar Deviasi, analisis sistem usulan, perancangan sistem, implementasi sistem, uji coba sistem pengklasteran, dan terakhir adalah *hosting*. (dapat dilihat pada Gambar 3)



Gambar 3. Metode Penelitian

2.5 Data Input

Penelitian ini akan menggunakan data pengiriman barang, berupa daftar pelanggan dan koordinat lokasi yang didapatkan dari *Google Maps*. Data diperoleh dari salah satu perusahaan manufaktur otomotif selama proses penelitian sejumlah 12 perusahaan pelanggan, dan berdasarkan perhitungan koordinat relatif perusahaan yang didapat dari hasil perhitungan antara titik koordinat perusahaan pelanggan dikurangi dengan titik koordinat perusahaan manufaktur otomotif seperti pada rumus (4):

$$\text{Jarak relatif} = x_i - x_y$$

Keterangan: x_i terdiri dari koordinat lintang dan koordinat bujur, merupakan perusahaan pelanggan dan x_y perusahaan manufaktur otomotif.

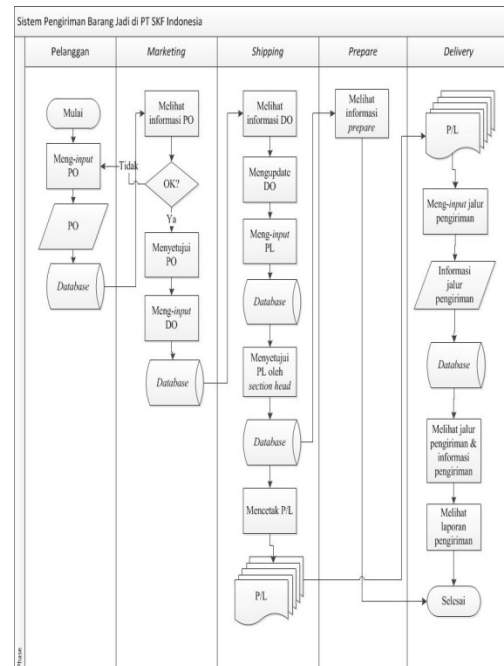
3. HASIL DAN PEMBAHASAN

3.1 Prosedur Sistem Informasi Pengiriman Barang Jadi Usulan

Prosedur sistem informasi pengiriman barang jadi yang diusulkan pada sistem informasi yang akan dibangun dapat diuraikan melalui tahapan berikut:

- Pelanggan akan mengirimkan *Purchase Order* (PO) melalui sistem pengiriman barang jadi PT SKF Indonesia yang akan tersimpan ke dalam *database*.
- Bagian *Marketing* akan menerima informasi PO yang dikirimkan oleh pelanggan.
- Apabila PO disetujui oleh Bagian *Marketing*, Bagian *Marketing* akan membuat *Delivery Order* (DO).
- Bagian *Shipping* akan menerima informasi DO.
- Bagian *Shipping* akan meng-*update* DO dan membuat *packing list* (PL) berdasarkan informasi DO yang diterima.
- Bagian *Shipping*, khususnya *section head* yang akan melakukan *approved packing list*.
- Sub bagian *Prepare* akan melihat informasi persiapan sesuai dengan tanggal persiapan.
- Packing list* yang sudah di-*approved* akan dicetak sebanyak 5 rangkap.
- Sub bagian *Delivery* akan menerima dokumen PL yang sudah tercetak dari Bagian *Shipping*.
- Sub bagian *Delivery* akan melakukan pembagian jalur pengiriman dengan meng-*input* data pada *form* pembagian jalur pengiriman.
- Output* yang dihasilkan yaitu informasi jalur pengiriman.

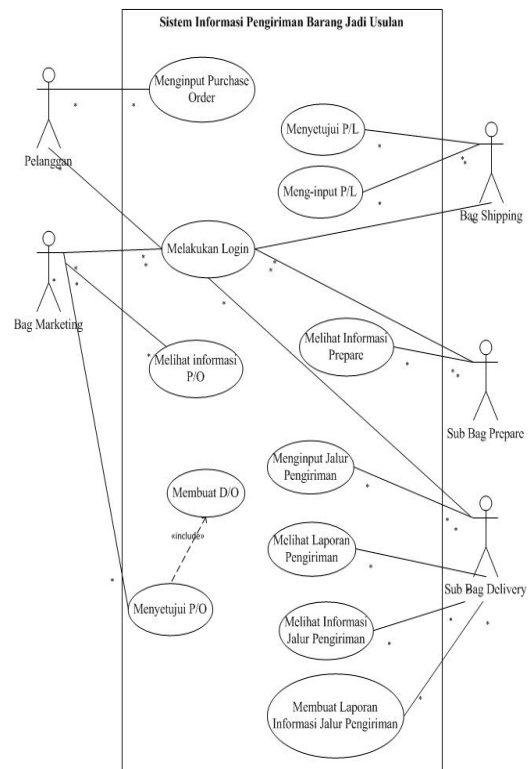
- Sub bagian *Delivery* dapat melihat informasi jalur pengiriman dan informasi pengiriman sesuai dengan tanggal/bulan yang terpilih.



Gambar 4. Flowmap Sistem Informasi Pengiriman Barang Jadi

3.2 Use Case Diagram

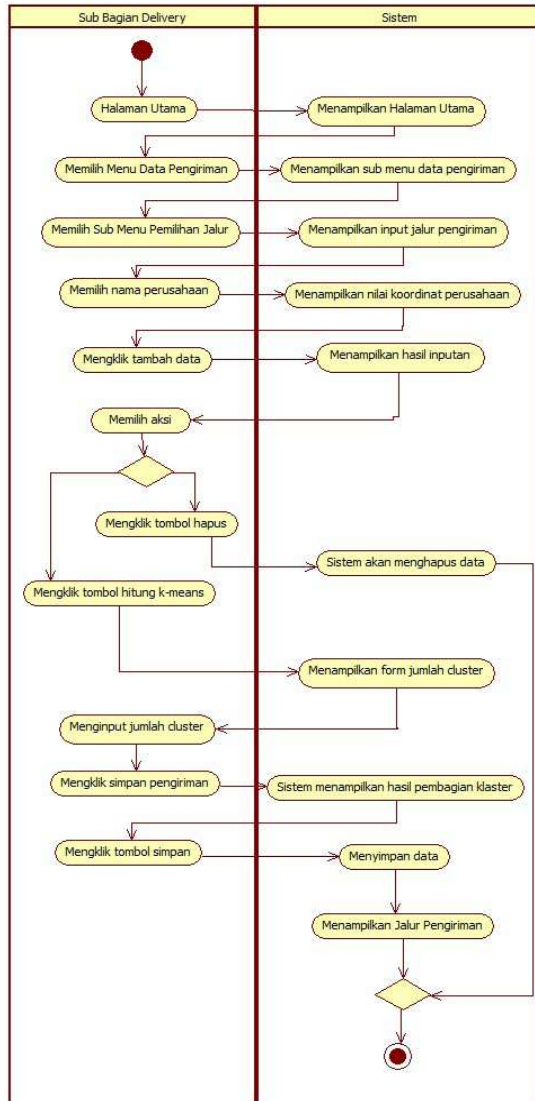
Use case diagram mendeskripsikan sebuah interaksi antara satu atau lebih aktor dengan sistem informasi yang akan dibuat. Rancangan *use case diagram* sistem informasi pengiriman barang jadi dapat dilihat pada Gambar 5.



Gambar 5. Use Case Diagram Sistem Informasi Pengiriman Barang Jadi Usulan

3.3 Activity Diagram

Activity diagram merupakan sebuah diagram yang menggambarkan urutan aktivitas proses bisnis pada suatu sistem. Berikut merupakan *activity diagram* meng-*input* jalur pengiriman dari sistem informasi pengiriman barang jadi usulan dapat dilihat pada Gambar 6.



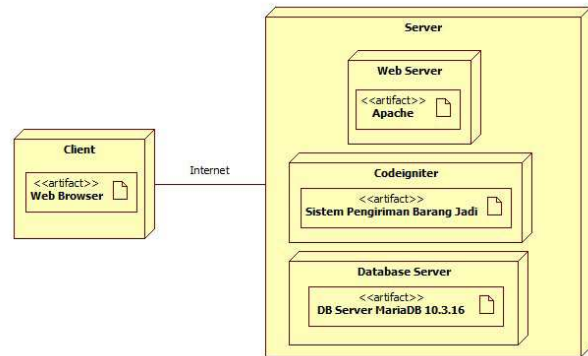
Gambar 6. Activity Diagram Meng-input Jalur Pengiriman

Activity diagram ini menjelaskan alur aktivitas yang dilakukan Sub bagian *Delivery* untuk membuat jalur pengiriman berdasarkan *packing list* yang sudah tercetak dengan menerapkan implementasi metode *k means clustering* pada sistem informasi pengiriman barang jadi.

3.4 Deployment Diagram

Deployment diagram pada usulan sistem informasi pengiriman barang jadi digunakan untuk menggambarkan komponen *software* dan bagaimana *software* ditempatkan di atas arsitektur fisik atau

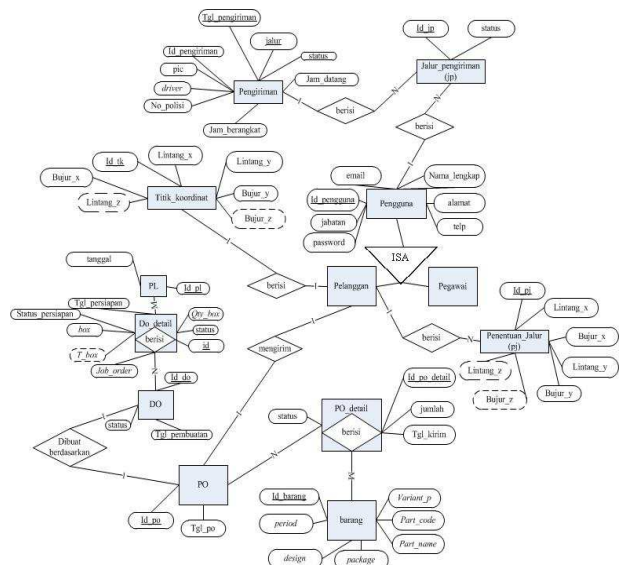
infrastruktur dari suatu informasi, berikut Gambar 7 merupakan *deployment diagram* sistem informasi pengiriman barang jadi.



Gambar 7. Deployment Diagram Sistem Informasi Pengiriman Barang Jadi Usulan

3.5 Pemodelan Data Sistem Usulan

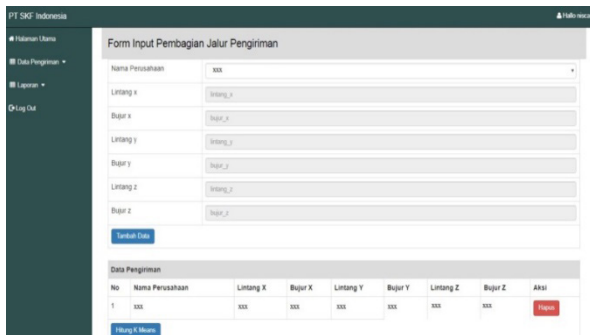
Pemodelan data pada sistem informasi pengiriman barang jadi usulan menggunakan *Entity Relationship Diagram* (ERD) yang digunakan untuk menjelaskan hubungan antar data dalam basis data. Gambar 8 menggambarkan penggambaran ERD Sistem Informasi Pengiriman Barang Usulan.



Gambar 8. ERD Usulan Sistem Pengiriman Barang Jadi

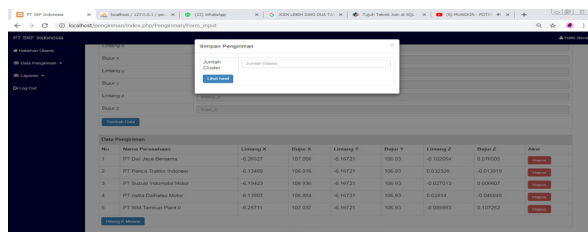
3.6 Perancangan Sistem Antarmuka

Rancangan *interface* (antarmuka) dari program sistem informasi pengiriman barang jadi ini bertujuan untuk menggambarkan antarmuka sistem yang telah dibuat. Rancangan antarmuka ini merupakan bagian yang berhubungan langsung antara pengguna dengan sistem. Halaman *form input* pembagian jalur pengiriman merupakan halaman yang digunakan sebagai *form input* perusahaan tujuan pengiriman yang nantinya akan tersimpan ke dalam *database* dan dilakukan peng-*cluster*-an nya. (Dapat dilihat pada Gambar 9)



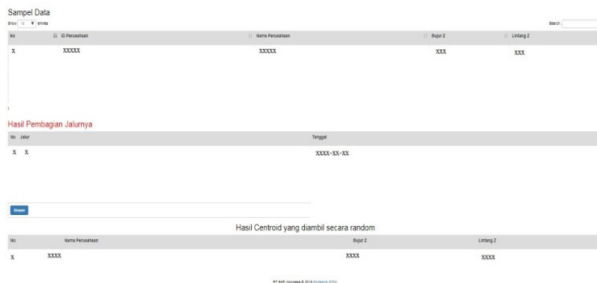
Gambar 9. Halaman Form Input Pembagian Jalur Pengiriman

Halaman *input* jumlah kluster merupakan halaman yang digunakan sebagai *form input* jumlah kluster. (Dapat dilihat pada Gambar 10)



Gambar 10. Halaman Input Jumlah Kluster

Halaman hasil pembagian jalur pembagian jalur pengiriman merupakan halaman yang berisi sampel data, hasil centroid secara random dan hasil pembagian jalurnya. (Dapat dilihat pada Gambar 11)



Gambar 11. Halaman Hasil Pembagian Jalur

Halaman jalur pengiriman digunakan untuk mengetahui jalur pengiriman yang telah ditetapkan. (Dapat dilihat pada Gambar 12)



Gambar 12. Halaman Jalur Pengiriman

3.7 Hasil Analisis Metode

Metode *K-Means Clustering* yang telah diimplementasikan ke dalam sistem informasi barang jadi kemudian dilakukan analisis perbandingan

dengan pembagian jalur sebelum penggunaan metode. Hasil realisasi pembagian jarak antar supir sebelum penggunaan metode *K-Means Clustering* terlihat pada Tabel 1 dan sesudah penggunaan metode 2.

Tabel 1. Perhitungan Jarak Tempuh Sebelum Menggunakan Metode

No	Tujuan	Jarak
1	SKF – Suzuki Indomobil Motor Penggilingan	3,8 km
	Suzuki Indomobil Motor Penggilingan – SKF	4,3 km
Jumlah Jarak Sopir 1		8,1 km
2	SKF - Astra Daihatsu Motor Sunter	10 km
	Astra Daihatsu Motor Sunter - Panca Traktor Indonesia	8,8 km
	Panca Traktor Indonesia – SKF	5,1 km
Jumlah Jarak Sopir 2		23,9 km
3	SKF –Mesin Isuzu	9,7 km
	Mesin Isuzu - SIM Tambun	12 km
	SIM Tambun - <i>Musashi Autoparts Indonesia</i>	20 km
	<i>Musashi Autoparts Indonesia</i> - Dwi Jaya Bersama	18 km
	Dwi Jaya Bersama – ASKI	42 km
Jumlah Jarak Sopir 3		152,7 km
Rata-rata jarak per sopir		61,57 km

Tabel 2. Perhitungan Jarak Tempuh Setelah Menggunakan Metode

No	Tujuan	Jarak
1	SKF – Astra Daihatsu Motor Sunter	10 km
	Astra Daihatsu Motor Sunter – Panca Traktor	8,8 km
	Panca Traktor – SKF	5,1 km
Jumlah Jarak Sopir 1		23,9 km
2	SKF – Suzuki Indomobil Motor Penggilingan	3,8 km
	Suzuki Indomobil Motor Penggilingan – Mesin Isuzu	9,4 km
	Mesin Isuzu – SKF	8,9 km
Jumlah Jarak Sopir 2		22,1 km
3	SKF – Suzuki Indomobil Motor Tambun	22 km
	Suzuki Indomobil Motor Tambun – <i>Musashi Autoparts</i>	20 km
	<i>Musashi Autoparts</i> – Dwi Jaya Bersama	5,2 km
	Dwi Jaya Bersama – ASKI	42 km
	ASKI – SKF	5,1 km
Jumlah Jarak Sopir 3		94,3 km
Rata-rata jarak per sopir		46,767 km

Berdasarkan realisasi pembagian pada Tabel 1 dan Tabel 2, dilakukan perbandingan dengan

menggunakan metode standar deviasi. Tabel 3 merupakan perhitungan standar deviasi data sebelum diterapkannya metode *K-Means Clustering*. Tabel 4 menunjukkan standar deviasi setelah diterapkan metode *K-Means Clustering*.

Tabel 3. Perhitungan Standar Deviasi Manual

Sopir ke	Jarak	Rata-rata jarak persopir	$s = \sum \sqrt{(x_1 - \bar{x})^2}$
1	8,1	61,57	53,47
2	23,9		37,67
3	152,7		91,13
Jumlah Standar Deviasi Manual			182,27

Tabel 4. Perhitungan Standar Deviasi Setelah Menggunakan Metode *K-Means Clustering*

Sopir ke	Jarak	Rata-rata jarak persopir	$s = \sum \sqrt{(x_1 - \bar{x})^2}$
1	23,9	46,767	22,867
2	22,1		24,667
3	94,3		47,533
Jumlah Standar Deviasi Setelah menggunakan metode <i>K Means Clustering</i>			95,067

Berdasarkan perbandingan antara Tabel 3 dan Tabel 4, standar deviasi jarak supir dengan rata-rata jarak yang ditempuh adalah 182,27 km untuk sebelum penggunaan metode, dan 95,067 km setelah penggunaan metode *K-Means Clustering*. Hasil tersebut menunjukkan bahwa metode *K-Means* mampu mengurangi standar deviasi jarak tempuh setiap supir dengan rata-rata jarak seluruh supir sebesar 87,203 km. Hal ini dapat disimpulkan bahwa metode *K-Means Clustering* membuat pembagian jarak tempuh supir lebih merata.

4. KESIMPULAN

Aplikasi pengiriman barang jadi telah dibangun dengan menggunakan metode pengembangan sistem menggunakan metodologi *waterfall* melalui tahapan perencanaan, analisis, desain dan implementasi. Pengembangan dilakukan dengan pendekatan berbasis objek dan menggunakan UML. Aplikasi yang telah dibangun yang mengimplementasikan metode *K-Means Clustering* mampu meminimalisir perbedaan jarak antar sopir. Selisih jarak antara metode manual dengan metode *K-Means Clustering* sebesar 87,203 km.

DAFTAR PUSTAKA

- [1] Agustin, F. E., Fitria, A., & S, A. H. (2015). Implementasi Algoritma K-Means untuk Menentukan Kelompok Pengayaan Materi Mata Pelajaran Ujian Nasional (Studi Kasus: SMAN 101 Jakarta). *Teknik Informatika*, 73-78.
- [2] Rohmawati, N., Defiyanti, S., & Jajuli, M. (2015). Implementasi Algoritma K-MEANS dalam Pengklasteran Mahasiswa Pelamar Beasiswa. *Jurnal Ilmiah Teknologi Informasi Terapan*, 62-67.
- [3] Dhanachandra, N., & Chanu, K. M. (2015). Image Segmentation using K-means Clustering Algorithm and Subtractive Clustering Algorithm. *Procedia Computer Science*, 764-771.
- [4] Rahman, A. T., Wiranto, & Anggrainingsih, R. (2017). Coal Trade Data Clustering Using K-Means (Case Study PT Global Bangkit Utama. *Jurnal Ilmiah Teknologi dan Informasi*, 24-30.
- [5] Muflikhah, L., Ratnawati, D. E., & Putri, R. R. (2018). *Data Mining*. Malang: Tim UB Press
- [6] spsstatistik, a. (2018, May 14). *Rumus Standar Deviasi*. Dipetik Juli 30, 2019, dari spsstatistik.com: <https://www.spsstatistik.com/rumus-standar-deviasi/>
- [7] Adiarsa, A. (2012, Maret 14). *Mean, Median, Modus dan Standar Deviasi*. Dipetik Juli 30, 2019, dari blog.ub.ac.id: <http://blog.ub.ac.id/adiarsa/2012/03/14/mean-median-modus-dan-standar-deviasi/>
- [7] Dennis, A., Wixom, B. H., & Tegarden, D. (2015). *System Analysis & Design An Object-Oriented Approach with UML*. Danvers: John Wiley & Sons, Inc