



TESIS - IF185401

PENENTUAN JUMLAH CLUSTER OPTIMUM PADA SEGMENT RUTE PENERBANGAN MENGGUNAKAN DATA AUTOMATIC DEPENDENT SURVEILLANCE- BROADCAST

**Reza Prasetya Prayogo
NRP. 5116201009**

**DOSEN PEMBIMBING
Prof. Dr. Ir. Joko Lianto Buliali, M.Sc.
NIP: 19670727199203 1 002**

**PROGRAM MAGISTER
BIDANG KEAHLIAN DASAR DAN TERAPAN KOMUNIKASI
DEPARTEMEN INFORMATIKA
FAKULTAS TEKNOLOGI INFORMASI DAN KOMUNIKASI
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA
2019**

[Halaman sengaja dikosongkan]

LEMBAR PENGESAHAN

Tesis disusun untuk memenuhi salah satu syarat memperoleh gelar
Magister Komputer (M.Kom.)

di

Institut Teknologi Sepuluh Nopember Surabaya

oleh:

Reza Prasetya Prayogo
Nrp. 5116201009

Dengan judul :

Penentuan Jumlah Cluster Optimum pada Segmen Rute Penerbangan Menggunakan
Data Automatic Dependent Surveillance-Broadcast

Tanggal Ujian : 17-1-2019
Periode Wisuda : 2019 Gasal

Disetujui oleh:

Prof. Dr. Ir. Joko Lianto Buliali, M.Sc.
NIP. 196707271992031002



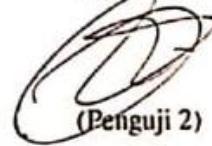
(Pembimbing 1)

Daniel Oranova Siahaan, S.Kom., M.Sc., PD.Eng.
NIP. 197411232006041001



(Penguji 1)

Tohari Ahmad, S.Kom., MIT., Ph.D.
NIP. 197505252003121002



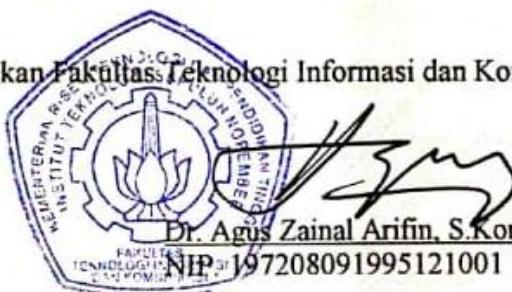
(Penguji 2)

Hadziq Fabroyir, S.Kom., Ph.D.
NIP. 1986201911089



(Penguji 3)

Dekan Fakultas Teknologi Informasi dan Komunikasi,



[Halaman ini sengaja dikosongkan]

PERNYATAAN KEASLIAN

Dengan ini saya menyatakan bahwa isi sebagian maupun keseluruhan Tesis saya dengan judul:

PENENTUAN JUMLAH CLUSTER OPTIMUM PADA SEGMENT RUTE PENERBANGAN MENGGUNAKAN DATA AUTOMATIC DEPENDENT SURVEILLANCE-BROADCAST

adalah benar-benar hasil karya intelektual mandiri, diselesaikan tanpa menggunakan bahan-bahan yang tidak diijinkan dan bukan merupakan karya pihak lain yang saya akui sebagai karya sendiri.

Semua referensi yang dikutip maupun dirujuk telah ditulis secara lengkap pada daftar pusaka.

Apabila ternyata pernyataan ini tidak benar, saya bersedia menerima sanksi sesuai dengan peraturan yang berlaku.

Surabaya, 21 Januari 2019

Reza Prasetya Prayogo

NRP: 5116201009

[Halaman ini sengaja dikosongkan]

**Penentuan Jumlah *Cluster* Optimum pada Segmen Rute Penerbangan
Menggunakan Data Automatic Dependent
*Surveillance–Broadcast***

Nama Mahasiswa : Reza Prasetya Prayogo

NRP : 5116 201 009

Pembimbing : Prof. Dr. Ir. Joko Lianto Buliali, M.Sc.

ABSTRAK

Dalam setiap rute penerbangan dibutuhkan beberapa titik lokasi tertentu agar *Air Traffic Control* dapat memantau pergerakan pesawat. Dimana dalam dunia penerbangan, titik-titik tersebut disebut *waypoint*. Pada penelitian ini penulis melakukan segmentasi untuk membagi suatu rute penerbangan menjadi beberapa area segmen berdasarkan *waypoint* yang sudah ditentukan sebelumnya oleh *Air Navigation Indonesia*. Pada tiap area segmen dilakukan analisis kebiasaan lokasi mana saja yang sudah biasa dilalui oleh pesawat menggunakan pendekatan *clustering*.

Dalam penelitian ini penulis menggunakan algoritma *clustering K-means* dengan jumlah cluster yang sudah ditentukan dan validasi pengelompokan internal *Silhouette Index*. Optimasi *centroid* juga dilakukan dengan mengimplementasikan algoritma *Antlion Optimizer* untuk kemudian diuji nilai performansinya sehingga pada akhirnya jumlah *cluster* dengan nilai validitas cluster terbaik adalah jumlah *cluster* yang optimum pada tiap segmen area penerbangan.

Pengujian dilakukan dengan membandingkan nilai *silhouette index* untuk setiap percobaan jumlah *cluster* terhadap kedua metode yaitu *K-means* dan *K-means - ALO*. Pada uji coba yang dilakukan, metode optimasi yang diusulkan menghasilkan nilai *silhouette index* yang lebih baik pada tiga segmen akan tetapi signifikan di semua segmen.

Kata Kunci: *Segmen Penerbangan, K-means, Antlion Optimizer, Kinerja dan Validitas Cluster*

[Halaman ini sengaja dikosongkan]

Determination of the Number of Optimum Clusters on Flight Route Segment using Automatic Dependent Surveillance-Broadcast Data

Nama Mahasiswa : Reza Prasetya Prayogo
NRP : 5116 201 009
Pembimbing : Prof. Dr. Ir. Joko Lianto Buliali, M.Sc.

ABSTRACT

Each flight route takes several specific location points for Air Traffic Control to monitor the movement of the aircraft that are called waypoints. In this study the authors do a segmentation to divide a flight route into several segment areas based on waypoints that have been determined previously by Air Navigation Indonesia. In each segmented area, an analysis of the location habits that have been used by the aircraft using the clustering approach is analyzed

In this study the author uses the *K-means* clustering algorithm with a predetermined number of clusters. Centroid optimization is also done by implementing the Antlion Optimizer algorithm to test the performance values. So the number of clusters with the best performance value is the optimum number of clusters in each segment of the flight area.

Testing is done by comparing the silhouette index value for each trial number of clusters against both methods, namely *K-means* and *K-means - ALO*. In the trial, the proposed optimization method produces a better silhouette index value in three segments but is significant in all segments.

Keywords: *Flight Segment, K-means, Antlion Optimizer, Cluster Performance and Validity*

[Halaman ini sengaja dikosongkan]

KATA PENGANTAR

Bismillahirrohmaanirrohiim. Puji syukur penulis panjatkan kepada Tuhan Yang Maha Esa. Atas rahmat dan kasih sayangNya, penulis dapat menyelesaikan tesis dalam bentuk buku ini yang berjudul Penentuan Jumlah Cluster Optimum pada Segmen Rute Penerbangan Menggunakan Data Automatic Dependent Surveillance–Broadcast.

Pengerjaan buku ini penulis tujuhan untuk mengeksplorasi lebih mendalam topik topik yang tidak diwadahi oleh kampus, namun banyak menarik perhatian penulis. Selain itu besar harapan penulis bahwa pengerjaan tugas akhir sekaligus pengerjaan buku ini dapat menjadi batu loncatan penulis dalam menimba ilmu yang bermanfaat.

Penulis ingin menyampaikan rasa terima kasih kepada banyak pihak yang telah membimbing, menemani dan membantu penulis selama masa pengerjaan tesis maupun masa studi.

1. Allah SWT yang selalu memberi kebahagiaan dan makna pada hidupku.
2. Ibu dan Ayah yang selalu mendukung dalam segala hal.
3. Bapak Prof. Dr. Ir. Joko Lianto Buliali, M.Sc., selaku pembimbing penulis yang telah memberikan didikan, pengajaran, dan nasihat yang telah diberikan oleh beliau semasa pengerjaan tugas akhir.
4. Rekan-rekan satu angkatan 2016 mahasiswa magister Teknik Informatika yang tidak lelah membantu penulis semasa masa studi.
5. Penulis menyadari bahwa buku ini jauh dari kata sempurna. Maka dari itu, penulis memohon maaf apabila terdapat salah kata maupun makna pada buku ini. Akhir kata, penulis mempersembahkan buku ini sebagai wujud nyata kontribusi penulis dalam ilmu pengetahuan.

Surabaya, Januari 2019

Reza Prasetya Prayogo

[Halaman ini sengaja dikosongkan]

DAFTAR ISI

LEMBAR PENGESAHAN	iii
ABSTRAK	vii
ABSTRACT	ix
KATA PENGANTAR.....	xi
DAFTAR ISI	xiii
DAFTAR GAMBAR.....	xv
DAFTAR TABEL	xvii
BAB 1 PENDAHULUAN.....	1
1.1. Latar Belakang.....	1
1.2. Perumusan Masalah	3
1.3. Tujuan Penelitian	3
1.4. Manfaat Penelitian	3
1.5. Kontribusi Penelitian	3
1.6. Batasan Masalah	3
BAB 2 KAJIAN PUSTAKA DAN DASAR TEORI	5
2.1. ADS-B	5
2.2. Metode Penentuan Cluster Centroid	8
2.3. Metode Optimasi Centroid	9
2.4. Metode Pengukuran Validitas Cluster	11
2.5. Analysis of Variance.....	11
2.6. Least Significant Difference	13
BAB 3 METODOLOGI PENELITIAN	15
3.1 Pre-Processing	16
3.2 Metode Clustering	17
3.3 Optimasi Centroid.....	19
3.4 Validasi <i>Cluster</i>	22
BAB 4 HASIL DAN PEMBAHASAN	24
4.1 Lingkungan Uji Coba	25
4.2 Hasil Pengujian Metode.....	25
4.2.1 Segmen 2	26

4.2.2	Segmen 3	26
4.2.3	Segmen 4	27
4.2.4	Segmen 5	27
4.2.5	Segmen 6	28
4.2.6	Segmen 7	28
4.3	Analisis Hasil dan Pembahasan	29
4.3.1	Segmen 2	29
4.3.2	Segmen 3	31
4.3.3	Segmen 4	33
4.3.4	Segmen 5	36
4.3.5	Segmen 6	38
4.3.6	Segmen 7	39
BAB 5 KESIMPULAN	43	
5.1	Kesimpulan	43
5.2	Saran	43
DAFTAR PUSTAKA	45	
LAMPIRAN	47	

DAFTAR GAMBAR

GAMBAR 1. 1 GRAFIK SPLINE CUBIC INTERPOLATION OPTIMAL WAYPOINT (PUSADAN ET AL., 2016)	1
GAMBAR 1. 2 SEGMENTASI RUTE PENERBANGAN MENGGUNAKAN WAYPOINT BERDASARKAN REGULASI AIR NAVIGATION	2
GAMBAR 2. 1 BLOK DIAGRAM PERPADUAN RADAR DAN ADS-B	5
GAMBAR 2. 2 KARAKTERISTIK KINERJA SENSOR PENGAWAS	7
GAMBAR 2. 3 SKEMA ADS-B.....	7
GAMBAR 2. 4 JEBAKAN BERBENTUK KERUCUT DAN CARA BERBURU ANT LION.....	9
GAMBAR 2. 5 FLOWCHART <i>K-MEANS - ANT LION OPTIMIZER</i>	10
GAMBAR 2. 6 ILUSTRASI ELEMEN-ELEMEN PADA KOMPUTASI ELEMEN <i>I</i> , OBJEK DARI CLUSTER A	11
GAMBAR 3. 1 DESAIN SISTEM PENELITIAN	16
GAMBAR 3. 2 VISUALISASI SEGMENTASI RUTE PENERBANGAN BERDASARKAN WAYPOINT PADA RUTE SURABAYA-PALU.....	16
GAMBAR 3. 3 PSEUDO CODE CLUSTERING SEGMENTASI 2 SAMPAI 7.....	18
GAMBAR 3. 4 VISUALISASI HASIL <i>K-MEANS</i> CLUSTERING.....	18
GAMBAR 3. 5 PSEUDO CODE OPTIMASI CENTROID ANT LION OPTIMIZER	21
GAMBAR 3. 6 PSEUDO CODE METODE VALIDASI CLUSTER SILHOUETTE COEFFICIENT.....	22
GAMBAR 3. 7 <i>SILHOUETTE INDEX</i> PADA METODE <i>K-MEANS</i> DENGAN 3 CLUSTER PADA SEGMENTASI 2.....	23
GAMBAR 4. 1 VISUALISASI LEAST SIGNIFICANT DIFFERENCE SEGMENTASI 2 UNTUK METODE <i>K-MEANS</i>	30
GAMBAR 4. 2 VISUALISASI LEAST SIGNIFICANT DIFFERENCE SEGMENTASI 2 UNTUK METODE <i>K-MEANS-ALO</i>	31
GAMBAR 4. 3 VISUALISASI LEAST SIGNIFICANT DIFFERENCE SEGMENTASI 3 UNTUK METODE <i>K-MEANS</i>	32
GAMBAR 4. 4 VISUALISASI LEAST SIGNIFICANT DIFFERENCE SEGMENTASI 3 UNTUK METODE <i>K-MEANS -ALO</i>	33

GAMBAR 4. 5 VISUALISASI LEAST SIGNIFICANT DIFFERENCE SEGMENT 4 UNTUK METODE K-MEANS.....	35
GAMBAR 4. 6 VISUALISASI LEAST SIGNIFICANT DIFFERENCE SEGMENT 4 UNTUK METODE K-MEANS-ALO.....	36
GAMBAR 4. 7 VISUALISASI LEAST SIGNIFICANT DIFFERENCE SEGMENT 5 UNTUK METODE K-MEANS.....	37
GAMBAR 4. 8 VISUALISASI LEAST SIGNIFICANT DIFFERENCE SEGMENT 5 UNTUK METODE K-MEANS -ALO.....	38
GAMBAR 4. 9 VISUALISASI LEAST SIGNIFICANT DIFFERENCE SEGMENT 6 UNTUK METODE K-MEANS-ALO.....	39
GAMBAR 4. 10 VISUALISASI LEAST SIGNIFICANT DIFFERENCE SEGMENT 7 UNTUK METODE K-MEANS-ALO	41

DAFTAR TABEL

TABEL 3. 1 DAFTAR WAYPOINT RUTE PENERBANGAN SURABAYA - PALU	17
TABEL 4. 1 RINCIAN LINGKUNGAN UJI COBA	25
TABEL 4. 2 PENGUKURAN VALIDITAS <i>CLUSTER SILHOUETTE COEFFICIENT SEGMENT 2</i>	26
TABEL 4. 3PENGUKURAN VALIDITAS <i>CLUSTER SILHOUETTE COEFFICIENT SEGMENT 3</i>	26
TABEL 4. 4 PENGUKURAN VALIDITAS <i>CLUSTER SILHOUETTE COEFFICIENT SEGMENT 4</i>	27
TABEL 4. 5 PENGUKURAN VALIDITAS <i>CLUSTER SILHOUETTE COEFFICIENT SEGMENT 5</i>	28
TABEL 4. 6 PENGUKURAN VALIDITAS <i>CLUSTER SILHOUETTE COEFFICIENT SEGMENT 6</i>	28
TABEL 4. 7 PENGUKURAN VALIDITAS <i>CLUSTER SILHOUETTE COEFFICIENT SEGMENT 7</i>	29
TABEL 4. 8 HASIL ANOVA SEGMENT 2 METODE <i>K-MEANS</i>	29
TABEL 4. 9 HASIL ANOVA SEGMENT 2 METODE <i>K-MEANS-ALO</i>	30

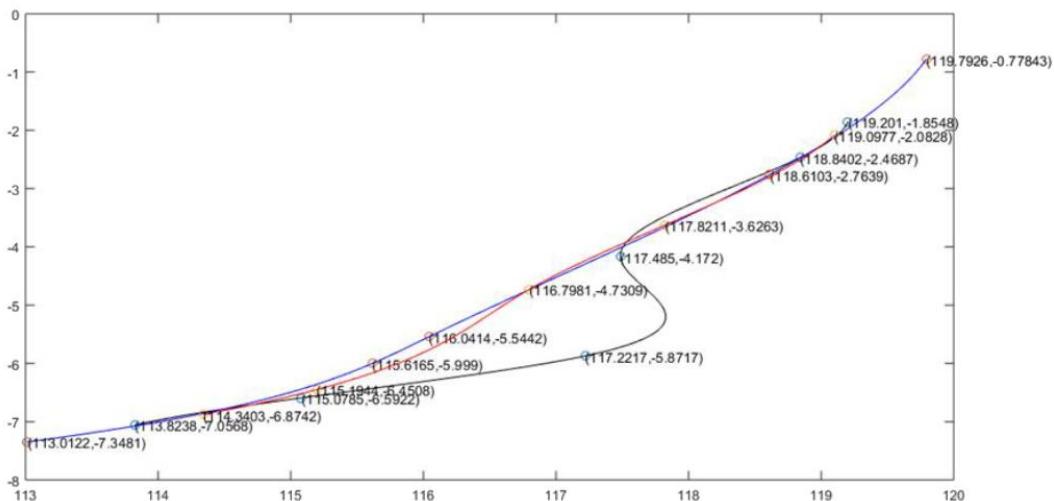
[Halaman sengaja dikosongkan]

BAB 1

PENDAHULUAN

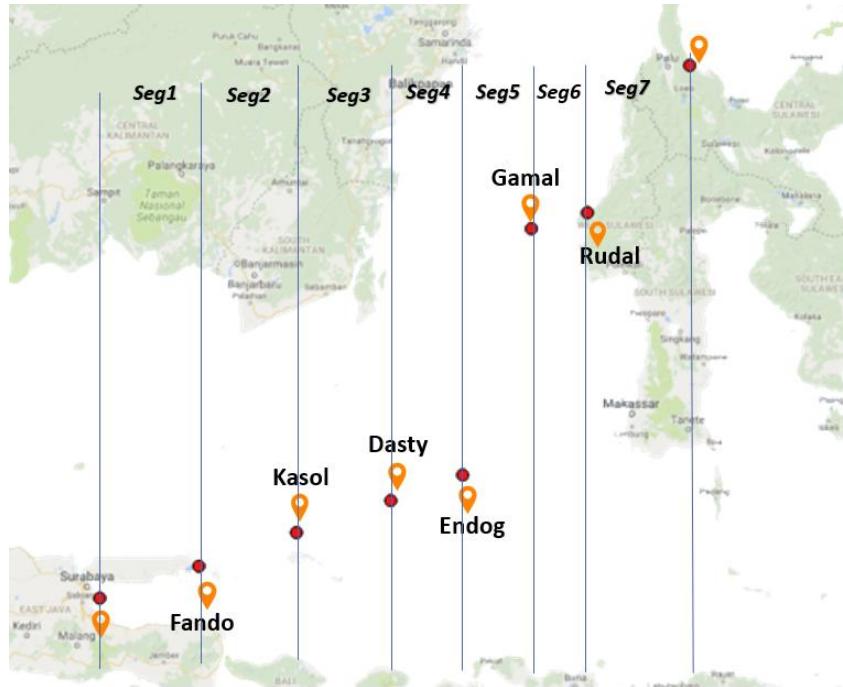
1.1. Latar Belakang

Faktor utama dalam transportasi udara adalah keselamatan dalam penerbangan. Oleh karena itu perilaku pesawat di luar kebiasaan harus bisa terdeteksi secepat mungkin untuk mengantisipasi terjadinya hal yang tidak diinginkan. Dalam setiap rute penerbangan terdapat beberapa titik lokasi tertentu agar *Air Traffic Control* dapat memantau pergerakan pesawat dalam rute penerbangannya. Pada penelitian sebelumnya tentang waypoint dan deteksi anomali penerbangan (Pusadan *et al.*, 2016), penulis meneliti tentang penentuan waypoint optimum untuk deteksi anomali pada rute penerbangan. Dari beberapa waypoint yang dihasilkan, selanjutnya ditentukan segmen-semen yang merupakan wilayah diantara suatu waypoint dengan waypoint sesudahnya.



Gambar 1. 1 Grafik Spline Cubic Interpolation Optimal Waypoint (Pusadan *et al.*, 2016)

Pada penelitian ini penulis melakukan segmentasi yang serupa dengan penelitian sebelumnya untuk membagi rute penerbangan ke dalam beberapa segmen berdasarkan waypoint yang sudah ditentukan oleh *Air Navigation Indonesia*.



Gambar 1. 2 Segmentasi rute penerbangan menggunakan waypoint berdasarkan regulasi Air Navigation

Seperti terlihat pada Gambar 1.2, tujuan utama dari penelitian ini adalah menentukan berapa jumlah area di tiap segmen pada suatu penerbangan tertentu. Oleh karena itu dibutuhkan proses pengelompokan data berdasarkan jarak sehingga penulis menggunakan algoritma *clustering*.

Algoritma clustering berbasis centroid yang digunakan dalam penelitian ini adalah *K-means* (Tang *et al.*, 2015) dengan penerapan metode optimasi centroid Ant Lion Optimizer (Mirjalili, 2015) (Majhi and Biswal, 2018). Pada algoritma tersebut jumlah k cluster harus ditentukan terlebih dahulu. Oleh karena itu diperlukan pengukuran validitas cluster (Xiong and Li, 2014) untuk menentukan jumlah cluster yang tepat, dimana dalam hal ini peneliti menyebutnya sebagai jumlah cluster yang optimum.

Dengan beberapa algoritma clustering yang digunakan seperti yang telah disebutkan di atas, penentuan jumlah cluster yang optimum menjadi faktor yang paling penting. Terdapat beberapa pendekatan sederhana yang umum dan bisa digunakan antara lain Davies-Bouldin Index (Davies and Bouldin, 1979), Calinski-Harabasz Index (Calinski and Harabasz, 1974), Dunn's Index, Silhouette Index (Rousseeuw, 1987).

1.2. Perumusan Masalah

Rumusan masalah yang diangkat dalam penelitian ini adalah sebagai berikut:

1. Bagaimana caranya menentukan jumlah *cluster* optimum pada set data penerbangan berbasis ADS-B berdasarkan parameter latitude dan longitude?
2. Bagaimana caranya mengukur kinerja dan validitas *cluster* pada tiap segmen area penerbangan?

1.3. Tujuan Penelitian

1. Tujuan penelitian ini adalah menentukan jumlah *cluster* optimum pada segmen jalur penerbangan berdasarkan pengukuran internal validitas cluster yaitu *silhouette index*.
2. Optimasi centroid pada metode *clustering K-means*.

1.4. Manfaat Penelitian

Manfaat dari penelitian ini adalah sebagai berikut :

1. Membantu *Air Traffic Management* dalam memandu dan mengamati (*surveillance*) lalu lintas penerbangan pada suatu rute berdasarkan pola kebiasaan dari data ADS-B yang diperoleh.
2. Sebagai acuan dalam penentuan deteksi dini anomali penerbangan berdasarkan analisis dari kebiasaan pesawat yang melintasi suatu rute penerbangan yang ditentukan.

1.5. Kontribusi Penelitian

Kontribusi dari penelitian ini adalah implementasi optimasi centroid pada metode pengelompokan *K-means* menggunakan metode Ant Lion Optimizer pada data ADS-B..

1.6. Batasan Masalah

Batasan masalah pada penelitian ini adalah :

1. Parameter set data ADS-B yang digunakan adalah *latitude* dan *longitude* pada posisi *cruising level* dengan jumlah data 90 hari.
2. Jumlah segmen ditentukan berdasarkan waypoint dari rute penerbangan dengan callsign LNI-860 (Surabaya - Palu).
3. Metode *cluster* yang digunakan adalah *clustering* berbasis jarak.

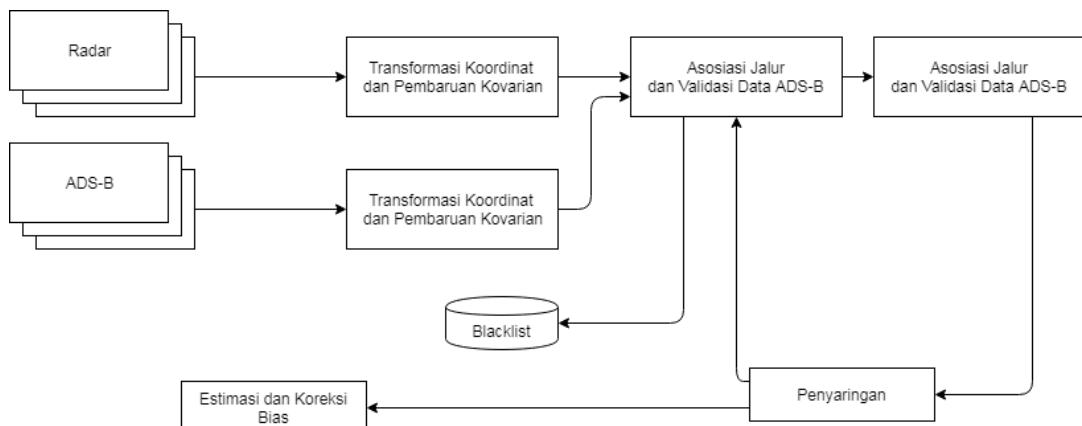
4. Metode pengukuran validitas cluster yang digunakan adalah pengukuran cluster internal, yaitu Silhouette Coefficient dan Davies-bouldin Index.

BAB 2

KAJIAN PUSTAKA DAN DASAR TEORI

2.1. ADS-B

Sebagai sumber informasi pengawasan pesawat, ADS-B diperkenalkan sebagai sensor yang dipadukan dengan radar. Sebuah penelitian tentang perpaduan radar dan ADS-B dapat dilihat pada gambar 2.1 (Jeon, Eun and Kim, 2015). Sedangkan karakteristik kinerja sensor pengawas dapat dilihat pada gambar 2.2 (Ali, 2013). Posisi pesawat pada data ADS-B dihasilkan dari pesawat itu sendiri (Pourvoyeur and Heidger, 2014). Federal Aviation Administration (FAA) memberikan perintah penggunaan sistem ADS-B kepada Air Traffic Control (ATC) karena beberapa kelebihan yaitu alat yang ringan dan biaya murah, sehingga sistem pesawat tak berawak pun tidak lagi dibatasi oleh masalah muatan dan kemampuan daya (Martel *et al.*, 2013). Pesawat yang dilengkapi dengan ADS-B dapat mengirimkan informasi dengan tingkat frekuensi lebih tinggi serta lebih komprehensif daripada radar. Selain itu ADS-B juga dapat mengirimkan informasi dengan posisi lebih tinggi dan akurasi kecepatan pesawat dengan lebih baik. Sistem ini mengirimkan data sekali per detik, lebih cepat dari radar dengan kecepatan pengiriman sekali per 4.8 detik di area terminal dan sekali per 12 detik di ruang udara. Skema sistem ADS-B dapat dilihat pada gambar 2.3.



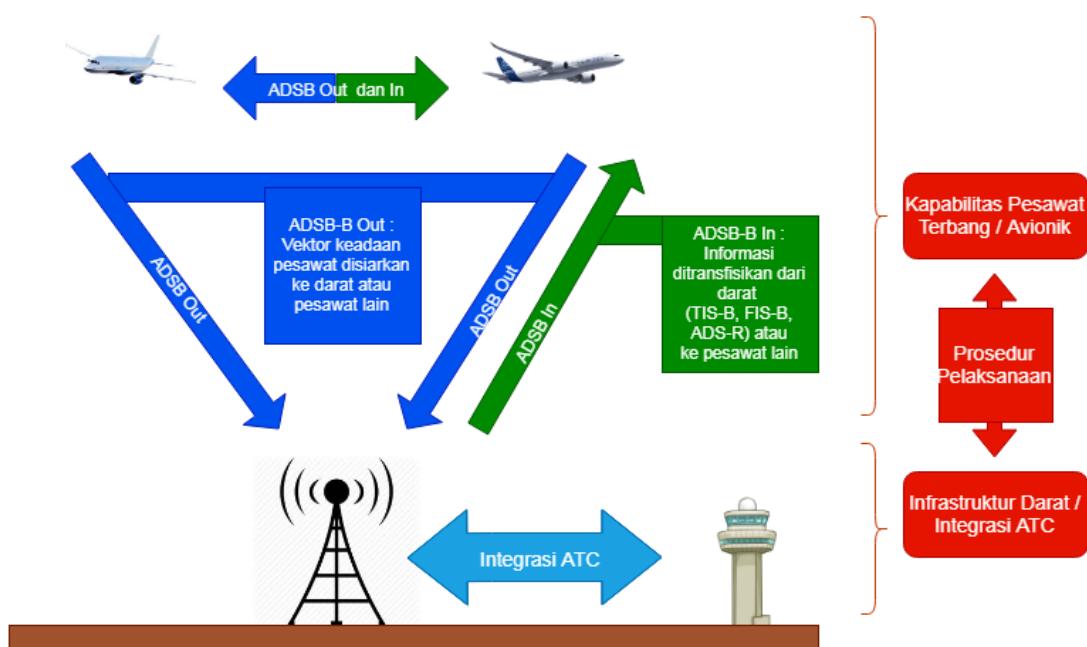
Gambar 2. 1 Blok diagram perpaduan radar dan ADS-B

Teknologi Pengawasan	Cakupan	Akurasi	Integritas	Periode Pembaruan
Primary Surveillance Radar (PSR)	S-band 60-80 NM L-band 160-220 NM	Dalam lingkup : Dalam azimuth :	Tidak ada laporan integritas yang disediakan	4 - 15 detik
Secondary Surveillance Radar (SSR) Mode (A/C)	200 NM- 250 NM	Dalam lingkup : Dalam azimuth :	Tidak ada laporan integritas yang disediakan	4 - 15 detik
Secondary Surveillance Radar (SSR) Mode (S)	200 NM- 250 NM	Sama seperti SSR(Mode A/C)	Tidak ada laporan integritas yang disediakan	4 - 12 detik
Mulilateration (MLAT)	200 NM	10-500 meter	Tidak ada laporan integritas yang disediakan	1 - 5 detik
ADS-B	200 NM- 250 NM	Ditentukan oleh avionik pesawat dan tidak tergantung dari sensor. Untuk GPS, 95% kurang dari 0,1NM	Integritas posisi dijamin ke 1 * 10-7 karena algoritma RAIM dalam avionik. Nilai integritas di-downlink-an dalam pesan ADS-B.	0.5 - 2 detik

			downlink	
--	--	--	----------	--

Gambar 2. 2 Karakteristik Kinerja Sensor Pengawas

Data yang dipancarkan dari transponder ADS-B di dalam pesawat ke arah *receiver ADS-B* di *ground station* disebut “ADS-B Out”. Data ini kemudian diteruskan untuk ditampilkan di Air Traffic Control. Pesawat lain juga dapat menerima pancaran data “ADS-B Out” dan menampilkannya pada monitor *on-board*. Kemampuan ini disebut “ADS-B In” dan dapat dilihat pada gambar 2.3 sebagai panah berwarna hijau.



Gambar 2. 3 Skema ADS-B

Terdapat sebuah penelitian tentang algoritma untuk peringatan dan kesadaran situasi lalu lintas menggunakan ADS-B dimana penelitian tersebut bertujuan untuk memberikan peringatan kepada pilot tentang adanya potensi tabrakan dan situasi berbahaya (Gariel, Kunzi and Hansman, 2011). Dalam penelitian lain, ADS-B digunakan sebagai input untuk algoritma deteksi konflik sebagai dukungan pada sistem penghindaran tabrakan di masa yang akan datang (Orefice *et al.*, 2014). Penelitian tentang algoritma deteksi konflik prediksi lintasan berbasis ADS-B juga sudah mendapatkan perhatian dari para peneliti (Baek and Bang, 2012).

2.2. Metode Penentuan Cluster Centroid

Algoritma *clustering* merupakan algoritma pengelompokan data ke dalam beberapa kelompok data tertentu, dimana data yang dikelompokkan adalah data yang tidak diketahui kebenarannya. Data yang terletak dalam suatu *cluster* harus memiliki kemiripan dan memiliki perbedaan dengan *cluster* yang lain. Pada beberapa aplikasi tertentu clustering disebut juga segmentasi data karena dapat mempartisi set data yang besar ke dalam beberapa kelompok berdasarkan kemiripan dalam satu cluster dan perbedaan dengan cluster lain (Han, Kamber and Pei, 2012).

Dalam penelitian ini algoritma *clustering* yang digunakan adalah *clustering* berdasarkan jarak kedekatan yaitu *K-means* (Jain, 2010). Clustering berbasis centroid digunakan karena tujuan dari penelitian ini adalah memperoleh area yang merupakan representasi dari cluster-cluster yang terbentuk dari sebaran data pada tiap segmen.

K-means adalah salah satu algoritma unsupervised learning yang paling sederhana yang memecahkan masalah *clustering* yang terkenal. Prosedur berikut cara sederhana dan mudah untuk mengklasifikasikan kumpulan data yang diberikan melalui sejumlah *cluster* tertentu (asumsikan k *cluster*) tetap apriori. Ide utamanya adalah mendefinisikan k center, satu untuk setiap *cluster*. Pusat-pusat ini harus ditempatkan dengan cara yang licik karena lokasi yang berbeda menyebabkan hasil yang berbeda. Jadi, pilihan yang lebih baik adalah menempatkannya sebanyak mungkin jauh dari satu sama lain. Langkah selanjutnya adalah mengambil setiap titik milik kumpulan data yang diberikan dan mengaitkannya ke pusat terdekat. Bila tidak ada titik yang tertunda, langkah pertama selesai dan usia kelompok awal selesai. Pada titik ini kita perlu menghitung ulang k centroids baru sebagai barycenter dari *cluster* yang dihasilkan dari langkah sebelumnya. Setelah kita memiliki centroid baru ini, sebuah ikatan baru harus dilakukan antara titik setel data yang sama dan pusat baru terdekat. Sebuah loop telah dihasilkan. Sebagai hasil dari loop ini kita mungkin memperhatikan bahwa pusat k mengubah lokasi mereka selangkah demi selangkah sampai tidak ada lagi perubahan yang dilakukan atau dengan kata lain pusat tidak bergerak lagi. Akhirnya, algoritma ini bertujuan untuk meminimalkan fungsi objektif yang dikenal sebagai fungsi error kuadrat. Langkah-langkah dari algoritma *K-means* adalah sebagai berikut :

$X = \{x_1, x_2, x_3, \dots, x_n\}$ adalah the set data points and $V = \{v_1, v_2, \dots, v_c\}$ adalah set titik pusat.

1. Pilih titik pusat secara acak.
2. Hitung jarak antara setiap titik data dan pusat *cluster*.
3. Tetapkan titik data ke pusat *cluster* yang jaraknya dari *cluster center* minimal dari semua pusat *cluster*.
4. Hitung ulang pusat *cluster* baru dengan menggunakan :

$$v_i = \frac{1}{C_i} \sum_{j=1}^{C_i} X_i$$

(2.1)

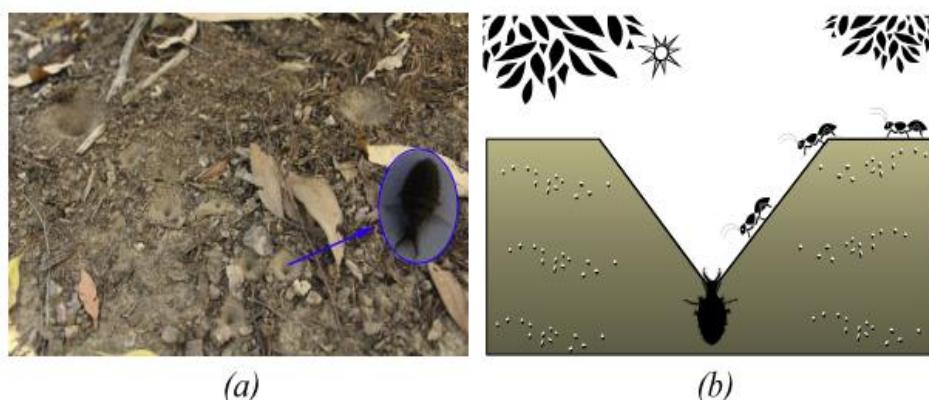
dimana C merupakan titik *centroid* dan X adalah data yang akan dihitung jaraknya.

5. Hitung ulang jarak antara setiap titik data dan pusat *cluster* yang baru didapat.

Jika tidak ada titik data yang dipindahkan maka berhenti, jika tidak ulangi dari langkah 3.

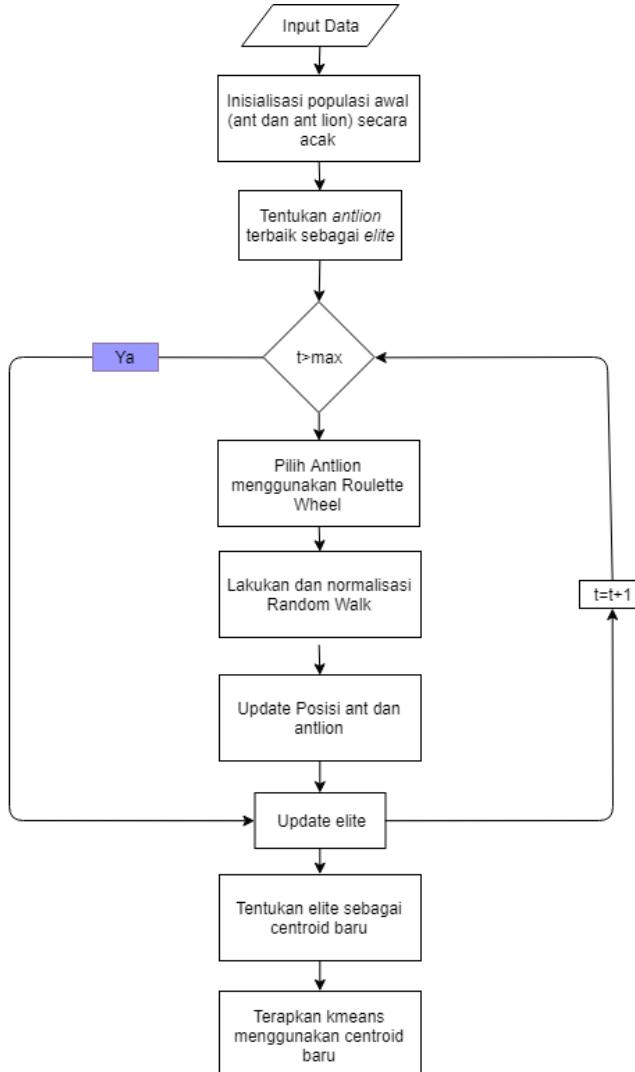
2.3. Metode Optimasi Centroid

Ant Lion Optimizer (ALO) merupakan salah satu metode yang dapat digunakan untuk mengoptimalkan penentuan centroid dalam proses clustering (Mirjalili, 2015). Seekor *Ant Lion* membuat lubang berbentuk kerucut yang merupakan jebakan untuk mangsa yaitu semut-semut yang lebih kecil. Ketika mangsa sudah memasuki lubang, *Ant Lion* akan melemparkan pasir keluar agar mangsa tergelincir dan lebih sulit untuk merangkak keluar. Terdapat lima langkah utama pada metode ini yaitu : (1) Pergerakan acak semut, (2) Membuat jebakan (lubang), (3) Menjebak mangsa di dalam lubang jebakan, (4) Menangkap mangsa dan (5) Membangun ulang jebakan.



Gambar 2. 4 Jebakan berbentuk kerucut dan cara berburu Ant Lion

Metode optimasi ini dilakukan setelah proses clustering pada dataset. Untuk setiap cluster yang terbentuk ditentukan posisi *Ant Lion*. Kemudian ukur nilai *fitness value* dari *objective function*, yaitu persamaan dari metode *K-means*. *Ant Lion* dengan nilai fitness minimum ditetapkan sebagai *Elite (Centroid)*. Flowchart dapat dilihat pada gambar 2.5.



Gambar 2. 5 Flowchart *K-means - Ant Lion Optimizer*

Random Walk untuk iterasi t menggunakan persamaan (2.1). Untuk membatasi random walk agar tetap di dalam area pencarian, dilakukan normalisasi menggunakan persamaan (2.2).

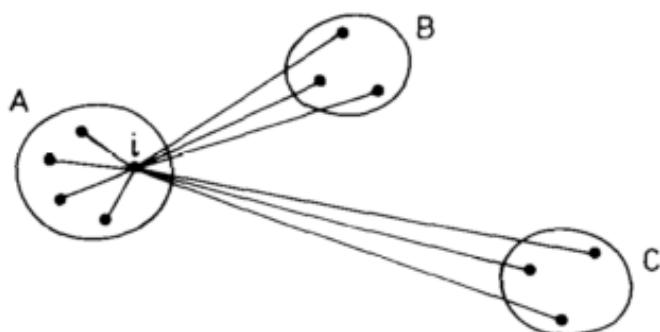
$$X[t] = [0, \text{cumsum}(2r(t_1) - 1), \text{cumsum}(2r(t_2) - 1), \dots, \text{cumsum}(2r(t_n) - 1)] \quad (2.2)$$

dimana, $r(t) = 1$ jika $\text{rand} > 0.5$ atau 0 if $\text{rand} \leq 0.5$

2.4. Metode Pengukuran Validitas Cluster

Faktor utama yang mempengaruhi validasi cluster adalah pengukuran validitas cluster internal dalam memilih jumlah cluster yang optimal. Beberapa parameter yang menjadi karakteristik adalah kerapatan pemisahan yang diukur pada jarak intra-cluster (Jegatha Deborah, Baskaran and Kannan, 2010). Dimana metode pengukuran validitas cluster dapat dikategorikan menjadi dua yaitu external dan internal (Xiong and Li, 2014), akan tetapi ada juga yang mengkategorikannya menjadi tiga yaitu validasi internal, external, dan relatif (Vij and Khandnor, 2016).

Silhouette berguna ketika mencari cluster yang rapat dan benar-benar terpisah. Ilustrasi dari perhitungan nilai silhouette index dapat dilihat pada gambar 2.3.



Gambar 2. 6 Ilustrasi elemen-elemen pada komputasi elemen i , objek dari cluster A.

Koefisien *silhouette* dari individu objek $x(j)$ didefinisikan sebagai berikut :

$$S_{x(j)} = \frac{b_{p,j} - a_{p,j}}{\max \{a_{p,j}, b_{p,j}\}}$$

(2.3)

Hitung rata-rata jarak setiap objek ke- j dengan semua objek yang ada pada cluster p , cluster yang sama dengan objek j , dilambangkan dengan $a_{p,j}$. Kemudian, hitung rata-rata jarak dari setiap objek ke- j dengan semua objek yang ada pada *cluster* q , dimana p tidak sama dengan q , disebut $d_{q,j}$. Lalu, cari $a_{p,j}$ dari minimum $d_{q,j}$, yang menunjukkan perbedaan rata-rata objek $x(j)$ untuk cluster yang terdekat dengan tetangganya. Untuk s_j yang memiliki nilai tinggi, maka semakin tepat penempatan $x(j)$ ke cluster p . Nilai koefisien silhouette biasanya di antara rentang -1 sampai dengan 1.

2.5. Analysis of Variance

Analisis of variance (ANOVA) adalah kumpulan dari model statistik yang digunakan untuk menganalisis perbedaan rata-rata antara kelompok dan prosedur terkait (seperti “variasi” antara kelompok), yang dikembangkan oleh Ronald Fisher. ANOVA

menggunakan uji f untuk menguji secara statistik. ANOVA termasuk jenis analisis statistik yang sangat umum, berguna dalam desain eksperimental yang sangat sederhana dan sangat rumit (Miller & Haden, 1988).

Dalam ANOVA hanya satu hipotesis yang digunakan, yaitu hipotesis dua arah (*two tail*) artinya hipotesis ini ingin mengetahui apakah ada perbedaan rata-rata atau tidak dan tidak spesifik yang mana yang berbeda.

ANOVA memiliki beberapa jenis salah satunya adalah *one-way* ANOVA. *one-way* ANOVA digunakan untuk menguji perbedaan tiga kelompok atau lebih berdasarkan satu variabel independen.

Berikut adalah langkah-langkah dalam perhitungan *one-way* ANOVA:

1. Tentukan k atau banyaknya perlakuan,
2. Tentukan n atau banyaknya sampel,
3. Hitung jumlah kuadrat total dengan persamaan 2.5

$$SS_T = \sum X_{ij}^2 - \frac{(\sum T_j)^2}{n} \quad (2.4)$$

4. Hitung jumlah kuadrat perlakuan dengan persamaan 2.6

$$SS_P = \sum \frac{(\sum T_j)^2}{n_j} - \frac{(\sum T_j)^2}{n} \quad (2.5)$$

5. Hitung nilai *F-value* dengan menggunakan rumus yang tertera pada Tabel 2.1.

Tabel 2. 1 Perhitungan Anova

Source	df	SS	MS	F-value
Treatment	$n - 1$	SS_P	$\frac{SS_P}{k - 1}$	$\frac{MS_P}{MS_E}$
Error	$(n - 1) - (k - 1)$	SS_E $= SS_T - SS_P$	$\frac{SS_E}{df_E}$	
Total	$n - 1$	SS_T		

6. Cari harga F tabel dengan mempertimbangkan (1) tingkat signifikansi (α), (2) df antar perlakuan, dan (3) df dalam perlakuan,
7. Bandingkan harga F Hitung dengan F table

- a. Bila $F_{\text{Hitung}} < F_{\text{tabel}}$, maka H_0 diterima, yang berarti rata-rata kedua perlakuan tidak berbeda secara signifikan,

Bila $F_{\text{Hitung}} > F_{\text{tabel}}$, maka H_0 ditolak dan H_1 diterima, yang berarti rata-rata kedua perlakuan berbeda secara signifikan.

2.6. Least Significant Difference

Tes LSD dikembangkan oleh Fisher, Ronald Aylmer (1935), yang ingin mengetahui perlakuan mana yang memiliki efek signifikan dalam analisis varian. *Least significant difference* (LSD) digunakan ketika F-rasio menunjukkan penolakan terhadap hipotesis nol H_0 , dan terjadi ketika terdapat perbedaan signifikan antara rata-rata populasi. Tes ini membantu mengidentifikasi populasi yang berbeda secara statistik.

$$LSD = \frac{t(s\sqrt{2})}{\sqrt{n}}$$

(2.6)

dimana nilai t yang ditabulasikan pada tingkat probabilitas (misalnya $P<0.01$) dan dengan derajat kebebasan yang sama dengan (jumlah varietas minus 1) x (jumlah pengukuran untuk setiap varietas minus 1)

[Halaman ini sengaja dikosongkan]

BAB 3

METODOLOGI PENELITIAN

Tahapan penelitian dibutuhkan agar tujuan dari penelitian ini dapat tercapai. Berikut tahapan dan metode yang digunakan adalah sebagai berikut :

1. Studi literatur dan pencarian dataset.

Proses ini merupakan penelusuran referensi yang terdiri dari buku, jurnal, atau artikel. Kemudian dilanjutkan dengan pencarian data ADS-B yang tersedia di Flightradar24.

2. Rancang bangun sistem untuk menghasilkan cluster optimum.

Proses ini terdiri dari penggunaan dua metode clustering berbasis jarak, yaitu *K-means* dan K-medoids. Dimana jumlah cluster diinisialisasi dalam rentang tertentu.

3. Menyusun rancangan pengujian sistem.

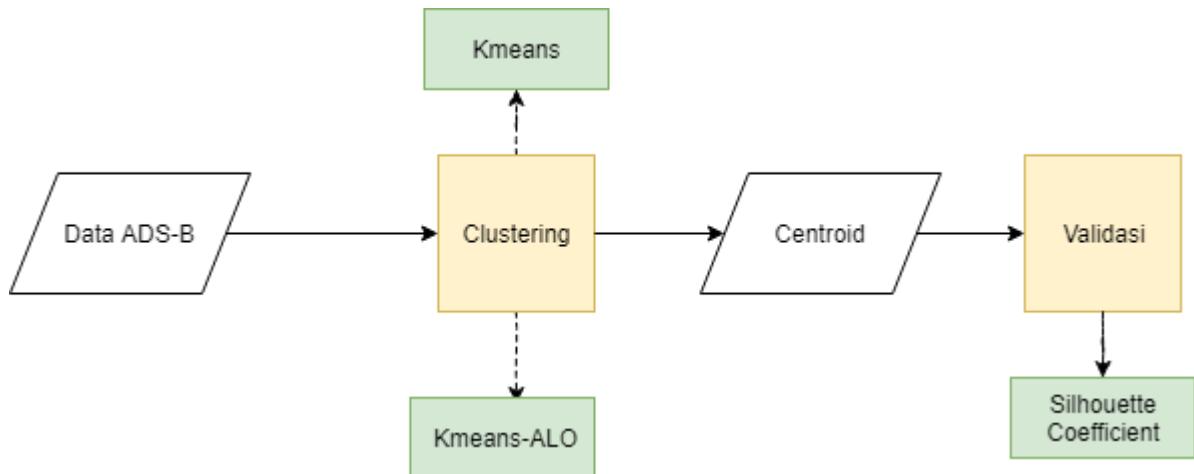
Dalam proses ini ditentukan skenario pengujian menggunakan metode Silhouette Coefficient dan Davies-Bouldin Index. Jumlah cluster dengan nilai index mendekati 1 adalah jumlah cluster yang optimum. Sedangkan untuk Davies-Bouldin, jumlah cluster optimum dihitung berdasarkan nilai index terkecil.

4. Mengimplementasikan sistem.

Sistem dibuat dalam bahasa pemrograman Python yang umum digunakan dalam kepentingan penelitian.

5. Menganalisis hasil yang diperoleh untuk menghasilkan kesimpulan.

Proses clustering dilakukan di semua segmen dengan kondisi pesawat pada kecepatan tetap (cruising level) dengan rentang inisialisasi jumlah cluster tertentu. Hasil cluster dari kedua metode tersebut kemudian diukur validitas clusternya menggunakan kedua metode Internal Validity Index.

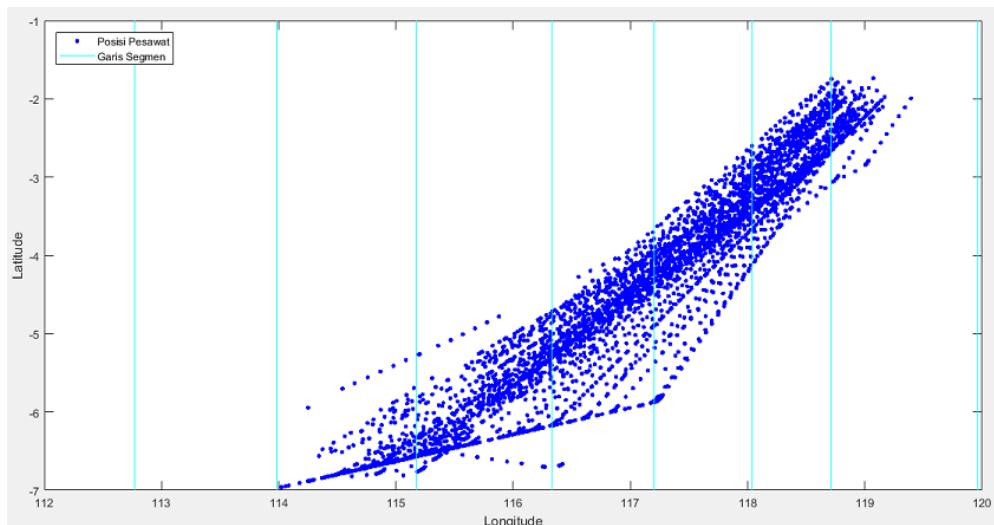


Gambar 3. 1 Desain Sistem Penelitian

Gambar 3.1 merupakan desain sistem penelitian yang diawali dengan membagi data ADS-B menjadi beberapa segmen berdasarkan waypoint. Kemudian pada tiap segmen dilakukan metode clustering yaitu K-Medoids dan *K-means*. Hasil clustering dari kedua metode tersebut akan membentuk beberapa kelompok titik koordinat yang merupakan representasi posisi pesawat pada rute penerbangannya.

3.1 Pre-Processing

Pada tahap ini dilakukan pembagian segmen berdasarkan 8 *waypoint* pada rute penerbangan Surabaya – Palu (Callsign LNI-860). Segmen yang digunakan pada penelitian ini adalah segmen dimana pesawat berada dalam posisi *cruising*, bukan *take off*, *landing*, ataupun *approaching*.



Gambar 3. 2 Visualisasi segmen rute penerbangan berdasarkan waypoint pada rute Surabaya-Palu

Pada gambar 3.2 di atas dapat dilihat representasi segmen yang dipisahkan dengan garis vertikal berwarna biru muda. Segmen terbentuk dari 2 waypoint yang ada yaitu waypoint sebelum dan sesudah. Berikut adalah daftar waypoint yang mengelompokkan data berdasarkan segmen :

Tabel 3. 1 Daftar Waypoint Rute Penerbangan Surabaya - Palu

Waypoint	Latitude	Longitude
Surabaya	-7.37333	112.77167
Fando	-6.97333	113.98500
Kasol	-6.56833	115.17333
Dasty	-6.17333	116.33000
Endog	-5.87667	117.20167
Gamal	-2.86333	118.03833
Rudal	-2.66167	118.71167
Palu	-0.88500	119.96167

3.2 Metode Clustering

Tujuan dari penelitian ini adalah mencari jumlah cluster optimum dari data ADS-B pada rute tertentu. Cluster sangat diperlukan karena data yang digunakan tidak memiliki class sebagaimana pada metode klasifikasi. Oleh karena itu dengan jumlah cluster yang optimum yang dihasilkan dapat diketahui area mana saja yang menjadi kebiasaan untuk dilewati pesawat pada rute penerbangan tertentu.

Metode yang digunakan pada pengelompokan berdasarkan jarak kedekatan adalah *K-means* dan *K-Medoids* dengan metode pengukuran jarak yaitu euclidian distance. Dataset akan dibagi berdasarkan latitude dan longitude waypoint yang dapat dilihat di tabel 3.1. Kemudian proses clustering dilakukan pada segmen 2 sampai segmen 7. Pada setiap segmen, dilakukan percobaan dengan inisialisasi cluster sejumlah 3 sampai 10 cluster.

```

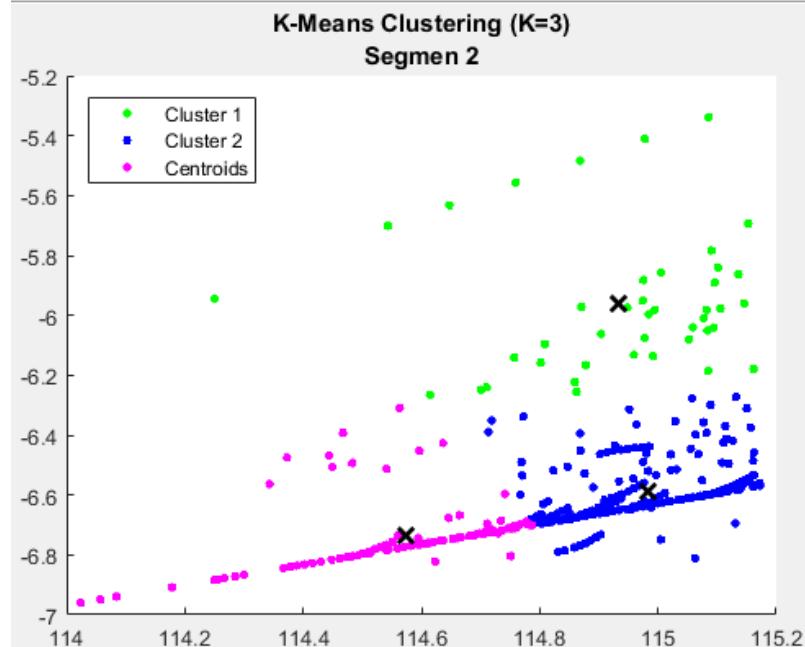
segmen = 7;
jumlah_k = 10;
z = 1;
load('waypoints.mat');
load(strcat('fdatas',int2str(segmen),'.mat'));
fdata = eval(strcat('fdatas',int2str(segmen)));
fdata = [fdata(:,2) fdata(:,1)];

for i=2:segmen
    list_segmen = [list_segmen;i];
    for j=3:jumlah_k

        [data_cluster,cluster_index_kmeans,centroid_kmeans]=index_kmeans(fdata
        ,i,j);
        [data_silhouette,    index_rerata,    index_max,    index_min]    =
        Silhouette(data_cluster);
        r_silhouette = [r_silhouette;i,j,index_rerata];
        z = z+1;
    end
end

```

Gambar 3. 3 Pseudo Code Clustering Segmen 2 Sampai 7



Gambar 3. 4 Visualisasi Hasil *K-means* Clustering

Gambar 3.4 merupakan *visualisasi* metode *K-means* dengan inisialisasi jumlah cluster sebanyak 3 cluster pada segmen 2. Tidak hanya 3, untuk setiap segmen dilakukan uji coba jumlah cluster sampai dengan 10 cluster. Data yang digunakan adalah data ADS-B pada segmen 2.

3.3 Optimasi Centroid

Pada setiap percobaan di tiap segmen dilakukan inisialisasi jumlah cluster antara 3 sampai 10. Percobaan dilakukan dengan iterasi sebanyak 100, 500, dan 1000. Centroid yang sudah terbentuk selanjutnya diperbaiki atau dioptimalisasi menggunakan metode Ant Lion Optimizer.

```
t_iterasi = 1000;
jumlah_k = 10;

for t=1:t_iterasi

[antlions_roulette_slot]=alo_roulette_wheel_slot(elites,init_antlion_n
umber);

[antlions_roulette]=alo_roulette_wheel(antlions,antlions_roulette_slot
);

disp(num2str(t));
antlions_new = [];
elites_new = [];
compare_ants = [];
for i=2:segmen
    load(strcat('Data/fdatas',int2str(i),'.mat'));
    fdata = eval(strcat('fdatas',int2str(i)));
    fdata = [fdata(:,2) fdata(:,1)];
    [edata] = random_walk(fdata);
    eval(['move' num2str(i) '= edata;']);
    move = eval(strcat('move',int2str(i)));
    for a=1:length(fdata)
        move_data(a,1) = fdata(a,1) + move(a,1);
        move_data(a,2) = fdata(a,2) + move(a,2);
        eval(['edata' num2str(i) '= move_data;']);
    end
    move_data = [0,0];
```

```

for j=3:jumlah_k

[edata_cluster,cluster_index_kmeans,get_centroid_kmeans]=kmedoids_classic(eval(strcat('edata',int2str(i))),i,j,t);

[edata_silhouette, index_rerata, index_max, index_min] =
Silhouette(edata_cluster);

edata_fitness_k = [edata_fitness_k;edata_silhouette];

eval(['ants_s' num2str(i) '_k' num2str(j) '='
edata_silhouette;']);

get_dt =
eval(strcat('ants_s',num2str(i),'_k',num2str(j)));
disp(strcat('ants_s',num2str(i),'_k',num2str(j)));
sort_ants = sortrows(get_dt,4,'descend');

for e=1:j
    for d=1:length(sort_ants)
        if sort_ants(d,3) == e
            compare_ants_tmp(1,1) = i;
            compare_ants_tmp(1,2) = j;
            compare_ants_tmp(1,3) = sort_ants(d,3);
            compare_ants_tmp(1,4) = sort_ants(d,1);
            compare_ants_tmp(1,5) = sort_ants(d,2);
            compare_ants_tmp(1,6) = sort_ants(d,4);
            compare_ants =
[compare_ants;compare_ants_tmp];
            break;
        else
            continue;
        end
    end
end
eval(['compare_ants' num2str(i) '_k' num2str(j) '='
compare_ants;']);

f_stop = size(compare_ants,1);
f_start = f_stop - j + 1;
disp(strcat(num2str(f_start),'-',num2str(f_stop)));
for f=f_start:f_stop
    if antlions_roulette(f,9) < compare_ants(f,6)

```

```

        antlions_new_tmp(1,1) = compare_ants(f,1);
        antlions_new_tmp(1,2) = compare_ants(f,2);
        antlions_new_tmp(1,3) = compare_ants(f,3);
        antlions_new_tmp(1,4) = compare_ants(f,4);
        antlions_new_tmp(1,5) = compare_ants(f,5);
        antlions_new_tmp(1,6) = compare_ants(f,6);

    else
        antlions_new_tmp(1,1) = antlions_roulette(f,3);
        antlions_new_tmp(1,2) = antlions_roulette(f,4);
        antlions_new_tmp(1,3) = antlions_roulette(f,5);
        antlions_new_tmp(1,4) = antlions_roulette(f,7);
        antlions_new_tmp(1,5) = antlions_roulette(f,8);
        antlions_new_tmp(1,6) = antlions_roulette(f,9);

    end
    antlions_new = [antlions_new;antlions_new_tmp];
    antlions_new_tmp = [];

if elites(f,8) < antlions_new(f,6)
    elites_new_tmp(1,1) = antlions_new(f,1);
    elites_new_tmp(1,2) = antlions_new(f,2);
    elites_new_tmp(1,3) = antlions_new(f,3);
    elites_new_tmp(1,4) = antlions_new(f,4);
    elites_new_tmp(1,5) = antlions_new(f,5);
    elites_new_tmp(1,6) = antlions_new(f,6);

else
    elites_new_tmp(1,1) = elites(f,3);
    elites_new_tmp(1,2) = elites(f,4);
    elites_new_tmp(1,3) = elites(f,5);
    elites_new_tmp(1,4) = elites(f,6);
    elites_new_tmp(1,5) = elites(f,7);
    elites_new_tmp(1,6) = elites(f,8);

end
elites_new = [elites_new;elites_new_tmp];
elites_new_tmp = [];

end
end
end

```

Gambar 3. 5 Pseudo Code Optimasi Centroid Ant Lion Optimizer

3.4 Validasi Cluster

Setelah ketiga metode di atas menghasilkan cluster sesuai jumlah k yang ditentukan, dibutuhkan pengukuran validasi untuk menentukan jumlah cluster yang paling sesuai untuk tiap segmen.

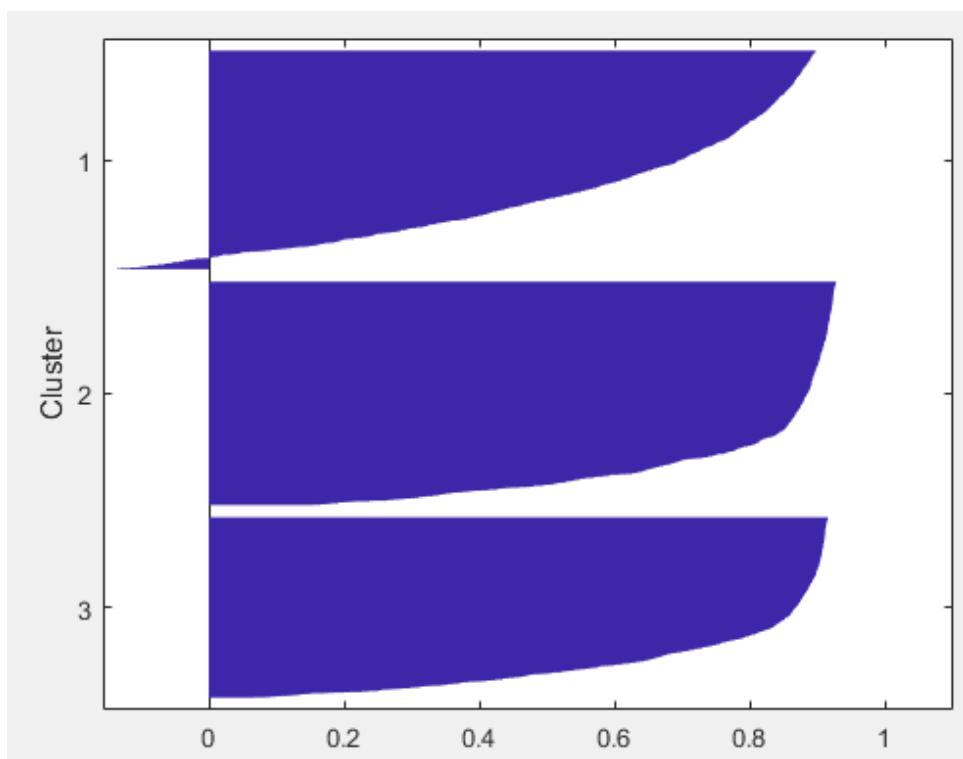
```
function[output1, output2, output3, output4]=Silhouette(dataCluster)

[s]=silhouette(dataCluster(:,1:2),           dataCluster(:,3), 'SqEuclidean');

%[s,h]
sort_s=sort(s);
max_s=max(s);
min_s=min(s);
for i=1:length(dataCluster)
    Silhouette_result(i,1) = dataCluster(i,1);
    Silhouette_result(i,2) = dataCluster(i,2);
    Silhouette_result(i,3) = dataCluster(i,3);
    Silhouette_result(i,4) = s(i,1);
    Silhouette_average(i,1) = Silhouette_result(i,4);
end
output1 = Silhouette_result;
output2 = sum(Silhouette_average)/length(Silhouette_average);
output3 = max_s;
output4 = min_s;
end
```

Gambar 3. 6 Pseudo Code Metode Validasi Cluster Silhouette Coefficient

Gambar 3.7 merupakan representasi salah satu metode pengukuran validitas cluster internal yaitu silhouette coefficient. Jumlah cluster yang optmum adalah ketika nilai indeks mendekati 1.



Gambar 3. 7 *Silhouette Index* pada metode *K-means* dengan 3 *cluster* pada segmen 2.

[Halaman ini sengaja dikosongkan]

BAB 4

HASIL DAN PEMBAHASAN

Pembahasan pertama adalah lingkungan pengujian yang terdiri dari perangkat keras dan perangkat lunak. Pembahasan kedua adalah menunjukkan hasil *clustering* berdasarkan nilai *silhouette index* pada tiap metode *clustering*. Pembahasan ketiga adalah menganalisis hasil *clustering* menggunakan uji statistik *one-way anova*.

4.1 Lingkungan Uji Coba

Sub bab ini menjelaskan ruang lingkup uji coba yang digunakan dalam penelitian ini. Lingkungan pengujian meliputi spesifikasi perangkat keras dan spesifikasi perangkat lunak. Perangkat keras yang dimaksud adalah jenis *processor* dan *memory (RAM)* yang digunakan, kapasitas *hard disk*, dan VGA yang digunakan. Untuk perangkat lunak adalah sistem operasi serta *tools* untuk pengujian. Rincian tersebut dapat dilihat pada 4.1 dibawah ini.

Tabel 4. 1 Rincian Lingkungan Uji Coba

Jenis Lingkungan	Rincian	
Perangkat Keras	<i>Processor</i>	Intel Core i5 2.80 Ghz
	<i>Memory (RAM)</i>	8192 GB
	<i>Hard disk</i>	464 GB
	<i>VGA</i>	Intel HD Graphics 4600
Perangkat Lunak	Sistem Operasi	Windows 10
	<i>Tools</i>	Matlab

4.2 Hasil Pengujian Metode

Metode yang diusulkan dibandingkan dengan *K-means* tradisional. Pengujian dilakukan pada segmen 2 sampai 7 dengan jumlah *cluster* mulai dari 3 sampai 10. *Dataset* yang digunakan adalah data penerbangan dengan posisi *cruising* (tidak mengalami perubahan ketinggian). Dalam pengujian ini didapatkan hasil evaluasi *cluster* berdasarkan nilai *silhouette index*.

4.2.1 Segmen 2

Hasil pengujian pada segmen 2 sebanyak 500 iterasi pada tabel 4.2 menghasilkan nilai rata-rata *silhouette index* tertinggi sebesar 0.6849 untuk jumlah k cluster = 3 menggunakan metode *K-means* pada percobaan pertama. Tiga titik tengah atau *centroid* yang terbentuk berdasarkan parameter *latitude* dan *longitude*, yaitu : -6.5859,114.9826; -6.7371,114.5729; dan -5.9633,114.9340. Sedangkan metode *K-means - ALO* menghasilkan nilai rata-rata *silhouette index* tertinggi sebesar 0.7732 pada jumlah k cluster = 3. *Centroid* yang terbentuk adalah sebagai berikut : -5.3384,115.0854; -5.4107,114.9777; dan -6.5637,114.8586.

Visualisasi data hasil *clustering* serta nilai pengukuran validitas *cluster* menggunakan metode *K-means* dan *K-means - ALO* dapat dilihat pada Lampiran 1 sampai dengan 4.

Tabel 4. 2 Pengukuran Validitas *Cluster Silhouette Coefficient* Segmen 2

	K = 3	K = 4	K = 5	K = 6	K = 7	K = 8	K = 9	K = 10
Classic <i>K-means</i>	0.6849	0.6823	0.6824	0.6720	0.6588	0.6452	0.6423	0.6479
<i>K-means-</i> ALO	0.7732	0.7550	0.7402	0.7284	0.5427	0.5798	0.4582	0.5332

4.2.2 Segmen 3

Hasil pengujian pada segmen 3 pada tabel 4.3 menggunakan metode *K-means* menghasilkan nilai rata-rata tertinggi sebesar 0.6849 untuk jumlah k cluster = 3 pada percobaan ke 2. *Centroid* yang terbentuk adalah sebagai berikut : -6.5859,114.9826; -6.7371,114.5730; dan -5.9633,114.9341. Sementara itu, metode *K-means - ALO* menghasilkan nilai rata-rata *silhouette index* tertinggi sebesar 0.7732 untuk jumlah k cluster = 3. *Centroid* yang terbentuk yaitu : -6.7367,114.5794; -5.9633,114.9341; dan -6.5840,114.9851. Lampiran 5 sampai 8 menggambarkan *visualisasi* data hasil *clustering* dan nilai validitas *cluster* (*silhouette index*) menggunakan metode *K-means* dan *K-means - ALO*.

Tabel 4. 3 Pengukuran Validitas *Cluster Silhouette Coefficient* Segmen 3

	K = 3	K = 4	K = 5	K = 6	K = 7	K = 8	K = 9	K = 10
Classic <i>K-means</i>	0.6849	0.6823	0.6807	0.6583	0.6592	0.6412	0.6417	0.6483

<i>K-means</i>	0.7732	0.7169	0.7402	0.7284	0.7212	0.6815	0.5427	0.4716
- <i>ALO</i>								

4.2.3 Segmen 4

Hasil pengujian untuk segmen 4 pada tabel 4.4 menunjukkan perbedaan jumlah *cluster* yang optimal antar kedua metode. Metode *K-means* menghasilkan nilai rata-rata tertinggi sebesar 0.6677 untuk jumlah k cluster = 3 pada percobaan pertama. *Centroid* yang terbentuk adalah sebagai berikut : -5.8770,116.6404; -4.3845,116.9775; -4.9750,116.6267. Sedangkan metode *K-means* - *ALO* menghasilkan nilai rata-rata *silhouette index* tertinggi sebesar 0.6579 untuk jumlah k cluster = 5. *Centroid* yang terbentuk adalah sebagai berikut : -6.861,116.3907; -6.6719,116.4119; -5.8491,116.6484; -4.3833,116.9792; dan -4.9709,116.6263. *Visualisasi* data hasil *clustering* dan nilai validitas *cluster* (*silhouette index*) menggunakan metode *K-means* dan *K-means* - *ALO* dapat dilihat pada Lampiran 9 sampai 12.

Tabel 4. 4 Pengukuran Validitas *Cluster Silhouette Coefficient* Segmen 4

	K = 3	K = 4	K = 5	K = 6	K = 7	K = 8	K = 9	K = 10
Classic Kmeans	0.6677	0.6389	0.5974	0.5908	0.5884	0.5583	0.5539	0.5564
<i>K-means</i> - <i>ALO</i>	0.6574	0.6576	0.6579	0.4960	0.4775	0.4980	0.4981	0.1853

4.2.4 Segmen 5

Hasil pengujian pada segmen 5 menunjukkan hasil yang serupa dengan segmen 4 dimana terdapat perbedaan jumlah k cluster yang optimum antar kedua metode dan dapat dilihat pada tabel 4.5. Metode *K-means* menghasilkan nilai rata-rata tertinggi sebesar 0.6620 untuk jumlah k cluster = 3 pada percobaan pertama. *Centroid* yang terbentuk adalah sebagai berikut : -3.4163,117.8145; -5.2585,117.3781; dan -4.0489,117.4892. Sedangkan metode *K-means* - *ALO* menghasilkan nilai rata-rata *silhouette index* tertinggi sebesar 0.7374 untuk jumlah k cluster = 6. *Centroid* yang terbentuk yaitu : -5.7398,117.2704; -4.8834,117.4270; -3.2380,117.8902; -4.3009,117.7199; -4.0484,117.3737; dan -3.6772,117.6983. Lampiran 13 sampai 16 menggambarkan *visualisasi* data hasil *clustering* dan nilai validitas *cluster* (*silhouette index*) menggunakan metode *K-means* dan *K-means* - *ALO*.

Tabel 4. 5 Pengukuran Validitas *Cluster Silhouette Coefficient* Segmen 5

	K = 3	K = 4	K = 5	K = 6	K = 7	K = 8	K = 9	K = 10
Classic <i>K-means</i>	0.6620	0.5727	0.5652	0.5737	0.5623	0.5657	0.5636	0.5548
<i>K-means</i> - <i>ALO</i>	0.7113	0.7324	0.7350	0.7374	0.2540	0.7029	0.2497	0.2932

4.2.5 Segmen 6

Hasil pengujian pada segmen 6 pada tabel 4.6 menghasilkan jumlah k *cluster* optimum yang sama antar kedua metode. Metode *K-means* menghasilkan nilai rata-rata tertinggi sebesar 0.9546 untuk jumlah K *cluster* = 3 pada percobaan pertama. *Centroid* yang terbentuk adalah sebagai berikut : -2.8915,61.4471; -2.6347,4.7530; dan -2.7122,88.2727. Sedangkan metode *K-means* - *ALO* menghasilkan nilai rata-rata *silhouette index* tertinggi sebesar 0.9290 untuk jumlah k *cluster* = 3. *Centroid* yang terbentuk yaitu : -2.6253,2.8874; -2.8840,63.0658; dan -2.7681,31.0385. Lampiran 17 sampai 20 menggambarkan *visualisasi* data hasil *clustering* dan nilai validitas *cluster* (*silhouette index*) menggunakan metode *K-means* dan *K-means* - *ALO*.

Tabel 4. 6 Pengukuran Validitas *Cluster Silhouette Coefficient* Segmen 6

	K = 3	K = 4	K = 5	K = 6	K = 7	K = 8	K = 9	K = 10
Classic <i>K-means</i>	0.9546	0.9488	0.9459	0.9467	0.9465	0.9464	0.9462	0.9435
<i>K-means</i> - <i>ALO</i>	0.9290	0.6259	0.8346	0.3444	0.8434	0.3516	0.3552	0.3588

4.2.6 Segmen 7

Hasil pengujian pada segmen 7 pada tabel 4.7 menghasilkan jumlah k *cluster* optimum yang berbeda antar kedua metode. Metode *K-means* menghasilkan nilai rata-rata tertinggi sebesar 0.6438 untuk jumlah K *cluster* = 4 pada percobaan ke 9. *Centroid* yang terbentuk adalah sebagai berikut : -2.1009,118.9071; -2.4900,118.7972; -2.7295,119.0248; dan -2.9888,118.7880. Sedangkan metode *K-means* - *ALO* menghasilkan nilai rata-rata *silhouette index* tertinggi sebesar 0.5932 untuk jumlah k *cluster* = 3. *Centroid* yang terbentuk yaitu : -2.1009, 118.9071; -2.8779, 118.8974; dan -2.4903, 118.8023. Lampiran 21 sampai 24 menggambarkan *visualisasi* data hasil

clustering dan nilai validitas *cluster (silhouette index)* menggunakan metode *K-means* dan *K-means - ALO*.

Tabel 4. 7 Pengukuran Validitas *Cluster Silhouette Coefficient Segmen 7*

	K = 3	K = 4	K = 5	K = 6	K = 7	K = 8	K = 9	K = 10
Classic <i>K-means</i>	0.6428	0.6438	0.6236	0.6360	0.6344	0.6296	0.6269	0.6214
<i>K-means</i> - <i>ALO</i>	0.5932	0.5642	0.5600	0.5499	0.5476	0.3264	0.1485	0.1193

4.3 Analisis Hasil dan Pembahasan

Setelah melalui tahap pengujian metode terhadap dataset, kemudian dilakukan analisis hasil *cluster* berupa nilai *silhouette index* menggunakan metode uji statistik dan uji lanjut. Uji statistik digunakan untuk mengetahui apakah ada perbedaan diantara jumlah k *cluster* dan metode lanjut digunakan untuk mendapatkan jumlah k *cluster* terbaik pada kedua metode di setiap segmen. Metode uji statistik yang digunakan adalah metode *one-way Analysis of Variance (one way anova)* dan *least significant difference* sebagai uji lanjut.

Analisis hasil pada setiap segmen dapat dilihat pada sub-sub bab berikut :

4.3.1 Segmen 2

Pengujian *anova* dilakukan pada segmen 2 untuk metode *K-means* menghasilkan nilai Prob > F atau p-value sebesar 0 dapat dilihat pada tabel 4.8 dimana nilai tersebut kurang dari 0.05 sehingga jumlah k *cluster* berpengaruh terhadap *dataset* yang digunakan. Kemudian dilakukan uji lanjut *Least Significant Difference*.

Tabel 4. 8 Hasil ANOVA Segmen 2 Metode *K-means*

<i>Source</i>	SS	df	MS	F	Prob > F
<i>Groups</i>	3,2138	7	0,4591	509,9690	0
<i>Error</i>	3,5940	3992	0,0009		
<i>Total</i>	6,8078	3999			

Pengujian *anova* juga dilakukan untuk metode *K-means - ALO* dan dapat dilihat pada tabel 4.9 dimana nilai Prob > F atau p-value sebesar 0 sehingga dapat disimpulkan bahwa jumlah k *cluster* berpengaruh terhadap *dataset* yang digunakan. Uji lanjut *Least Significant Difference* juga dilakukan terhadap metode *K-means - ALO*.

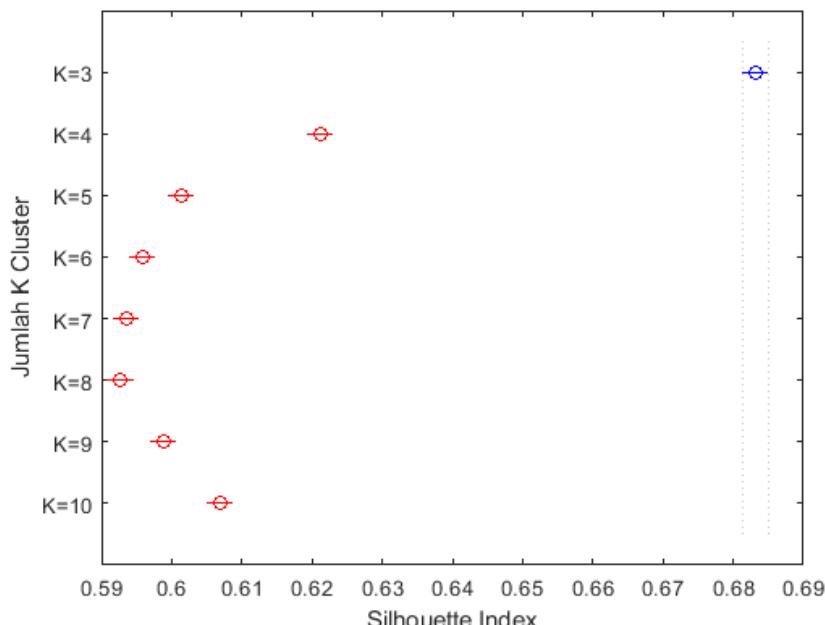
Tabel 4. 9 Hasil ANOVA Segmen 2 Metode *K-means* - *ALO*

Source	SS	df	MS	F	Prob > F
<i>Groups</i>	3.6657	7	0.5237	1556.6147	0
<i>Error</i>	1.3430	3992	0.0003		
<i>Total</i>	5.0087	3999			

Uji lanjut menggunakan *Least Significant Difference* (LSD) untuk metode *K-means* menghasilkan dua kelompok dari percobaan jumlah *k cluster*, dimana detail hasil pengujian dapat dilihat pada Lampiran 25. Kelompok pertama adalah $K = 3$ dengan nilai rata-rata *silhouette index* sebesar 0.6831 dan *standard error* sebesar 0.0013. Kelompok kedua adalah $K = 4$ sampai $K = 10$ dengan nilai rata-rata *silhouette index* sebesar 0.6015 dan *standard error* sebesar 0.0013.

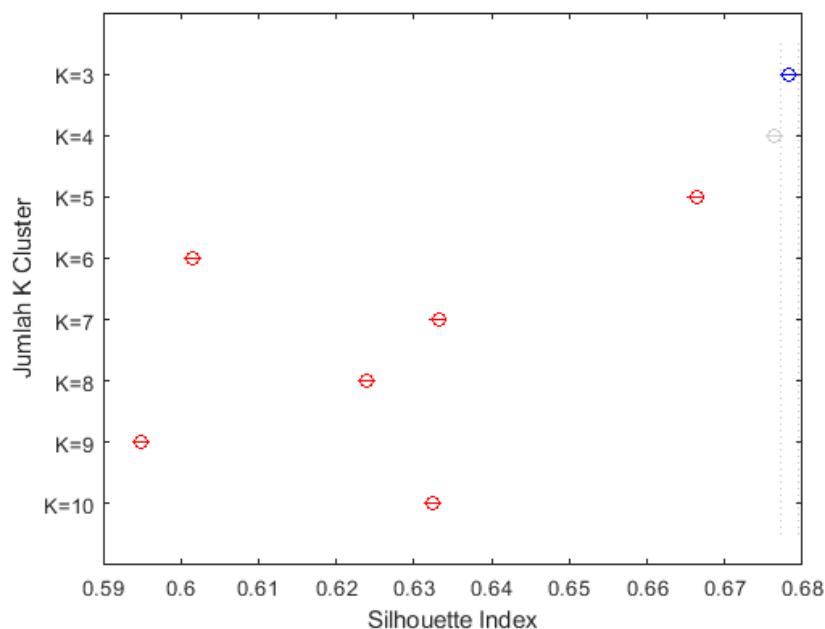
Berdasarkan kelompok yang terbentuk pada pengujian menggunakan metode LSD untuk metode *K-means* dapat ditarik kesimpulan bahwa kelompok pertama memiliki hasil *cluster* yang lebih baik dibandingkan kelompok kedua. Jumlah *k cluster* terbaik adalah $k=3$ dengan nilai rata-rata terbesar yaitu 0.6831.

Penjelasan Lampiran 25 dapat dilihat lebih mudah pada Gambar 4.1 dimana pada gambar tersebut semakin ke kanan menunjukkan semakin bagus kualitas *cluster* yang dihasilkan.



Gambar 4. 1 Visualisasi *Least Significant Difference* Segmen 2 untuk metode *K-means*

Uji lanjut menggunakan *Least Significant Difference* (LSD) untuk metode *K-means - ALO* menghasilkan enam kelompok dari percobaan jumlah k cluster dimana detail hasil pengujian dapat dilihat pada Lampiran 26. $K = 3$ menjadi jumlah *cluster* terbaik dengan nilai rata-rata *silhouette index* sebesar 0.6384 dan *standard error* sebesar 8.2026E-04. Hasil dari LSD, semua pValue yang dihasilkan bernilai dibawah 0.05 yang berarti jumlah k cluster ini merupakan jumlah *cluster* yang paling signifikan diantara yang lain. Penjelasan Lampiran 26 dapat dilihat lebih mudah pada Gambar 4.2 dimana pada gambar tersebut semakin ke kanan menunjukkan semakin bagus kualitas *cluster* yang dihasilkan.



Gambar 4. 2 Visualisasi Least Significant Difference Segmen 2 untuk metode *K-means - ALO*

4.3.2 Segmen 3

Pengujian *anova* dilakukan pada segmen 3 menghasilkan nilai Prob > F atau p-value sebesar 0 dapat dilihat pada tabel 4.10 dimana nilai tersebut kurang dari 0.05 sehingga jumlah k cluster berpengaruh terhadap *dataset* yang digunakan. Kemudian dilakukan uji lanjut LSD.

Tabel 4.10 Hasil ANOVA Segmen 3 Metode *K-means*

Source	SS	df	MS	F	Prob > F
Groups	3,3811	7	0,4830	528,8456	0
Error	3,6461	3992	0.0009		

Total	7,0272	3992			
-------	--------	------	--	--	--

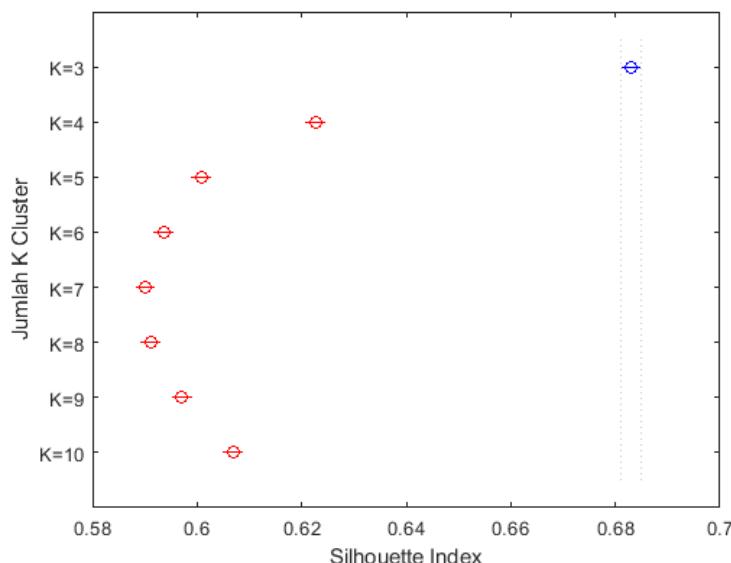
Pengujian *anova* juga dilakukan untuk metode *K-means - ALO* dan dapat dilihat pada tabel 4.11 dimana nilai Prob > F atau p-value sebesar 0 sehingga dapat disimpulkan bahwa jumlah k cluster berpengaruh terhadap *dataset* yang digunakan. Uji lanjut LSD juga dilakukan terhadap metode *K-means - ALO*.

Tabel 4.11 Hasil ANOVA Segmen 3 Metode *K-means - ALO*

Source	SS	df	MS	F	Prob > F
Groups	5.2390	7	0.7484	1841.4288	0
Error	1.6225	3992	0.0004		
Total	6.8615	3999			

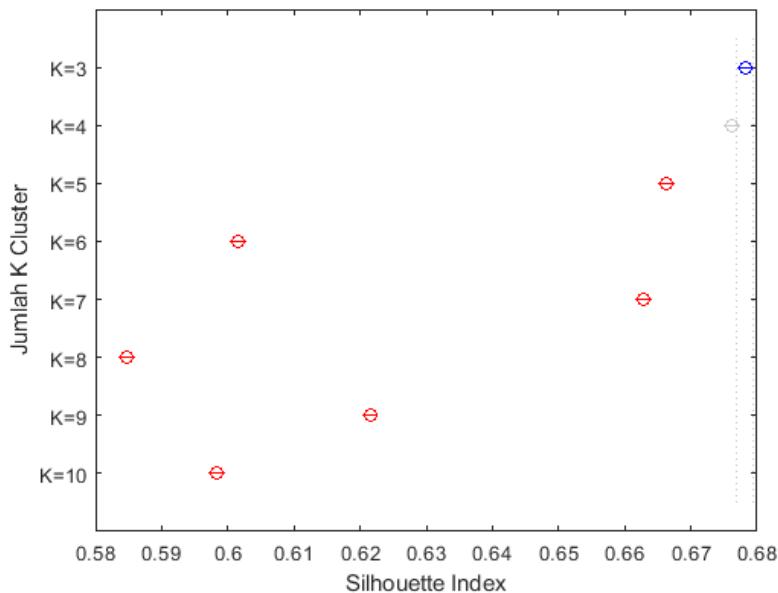
Uji lanjut menggunakan LSD untuk metode *K-means* menghasilkan tiga kelompok dari percobaan jumlah k cluster, dimana detail hasil pengujian dapat dilihat pada Lampiran 27. $K = 3$ menjadi jumlah *cluster* terbaik dengan nilai rata-rata *silhouette index* sebesar 0.6107 dan *standard error* sebesar 0.0014.

Berdasarkan kelompok yang terbentuk pada pengujian menggunakan metode LSD untuk metode *K-means* dapat ditarik kesimpulan bahwa jumlah $K = 3$ memiliki nilai dibawah 0.05 yang berarti jumlah k cluster ini merupakan jumlah *cluster* yang paling signifikan diantara yang lain. Penjelasan Lampiran 27 dapat dilihat lebih mudah pada Gambar 4.3 dimana pada gambar tersebut semakin ke kanan menunjukkan semakin bagus kualitas *cluster* yang dihasilkan.



Gambar 4. 3 Visualisasi Least Significant Difference Segmen 3 untuk metode *K-means*

Uji lanjut menggunakan LSD untuk metode *K-means - ALO* menghasilkan lima kelompok dari percobaan jumlah k cluster dimana detail hasil pengujian dapat dilihat pada Lampiran 28. $K = 3$ menjadi jumlah *cluster* terbaik dengan nilai rata-rata *silhouette index* sebesar 0.6362 dan *standard error* sebesar 0.0009. Hasil dari LSD, semua pValue yang dihasilkan bernilai dibawah 0.05 yang berarti jumlah k cluster ini merupakan jumlah *cluster* yang paling signifikan diantara yang lain. Penjelasan Lampiran 28 dapat dilihat lebih mudah pada Gambar 4.4 dimana pada gambar tersebut semakin ke kanan menunjukkan semakin bagus kualitas *cluster* yang dihasilkan.



Gambar 4. 4 Visualisasi Least Significant Difference Segmen 3 untuk metode *K-means - ALO*

4.3.3 Segmen 4

Pengujian *anova* pada segmen 4 menghasilkan nilai Prob > F atau p-value sebesar 0 dapat dilihat pada tabel 4.12 dimana nilai tersebut kurang dari 0.05 sehingga jumlah k cluster berpengaruh terhadap *dataset* yang digunakan. Kemudian dilakukan uji lanjut LSD.

Tabel 4.12 Hasil ANOVA Segmen 4 Metode *K-means*

Source	SS	df	MS	F	Prob > F
Groups	7.7631	7	1.1090	3766.0008	0
Error	1.1756	3992	0.0003		

<i>Total</i>	8.9387	3999			
--------------	--------	------	--	--	--

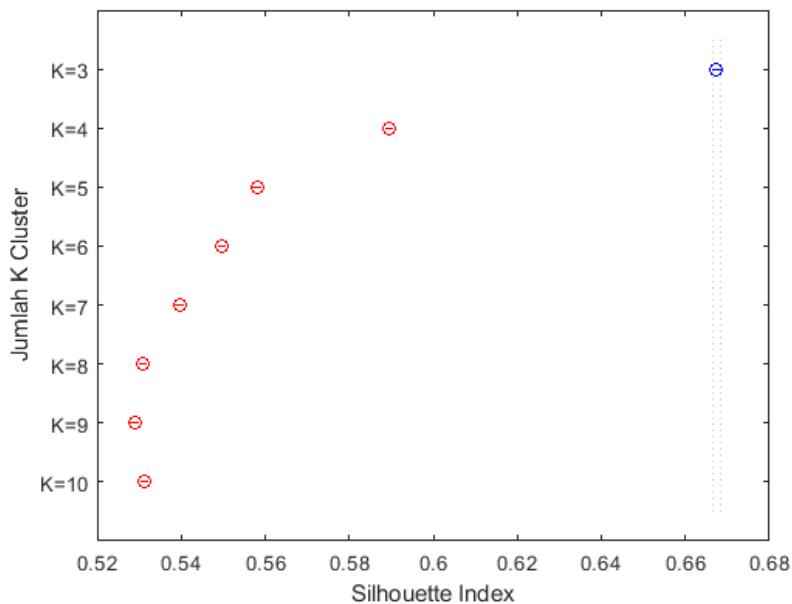
Pengujian *anova* untuk metode *K-means - ALO* dan dapat dilihat pada tabel 4.13 dimana nilai Prob > F atau p-value sebesar 0 sehingga dapat disimpulkan bahwa jumlah k cluster berpengaruh terhadap *dataset* yang digunakan. Uji lanjut LSD juga dilakukan terhadap metode *K-means - ALO*.

Tabel 4.13 Hasil ANOVA Segmen 4 Metode *K-means - ALO*

<i>Source</i>	SS	df	MS	F	Prob > F
<i>Groups</i>	0.4947	7	0.7135	7154.6510	0
<i>Error</i>	0.3981	3992	9.9730e-05		
<i>Total</i>	5.3928	3999			

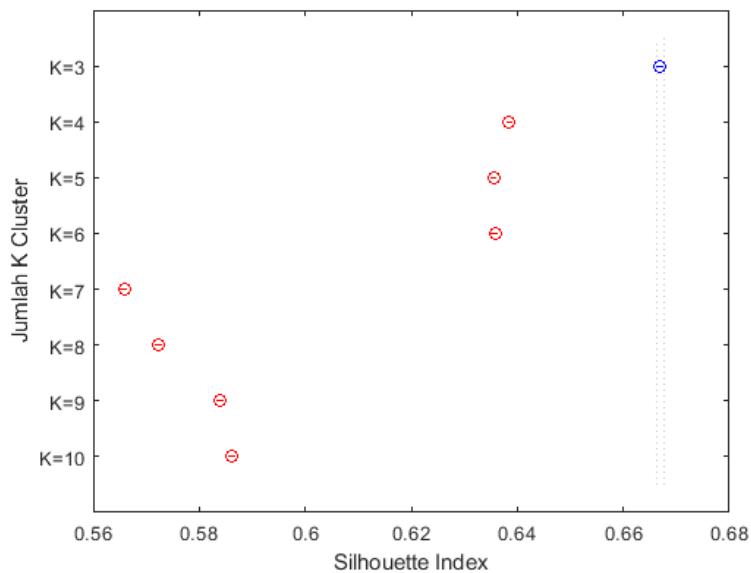
Uji lanjut menggunakan LSD untuk metode *K-means* menghasilkan tiga kelompok dari percobaan jumlah k cluster, dimana detail hasil pengujian dapat dilihat pada Lampiran 29. $K = 3$ menjadi jumlah cluster terbaik dengan nilai rata-rata *silhouette index* sebesar 0.5620 dan *standard error* sebesar 7.6743.

Berdasarkan kelompok yang terbentuk pada pengujian menggunakan metode LSD untuk metode *K-means* dapat ditarik kesimpulan bahwa jumlah $K = 3$ memiliki nilai dibawah 0.05 yang berarti jumlah k cluster ini merupakan jumlah cluster yang paling signifikan diantara yang lain. Penjelasan Lampiran 29 dapat dilihat lebih mudah pada Gambar 4.5 dimana pada gambar tersebut semakin ke kanan menunjukkan semakin bagus kualitas cluster yang dihasilkan.



Gambar 4. 5 Visualisasi Least Significant Difference Segmen 4 untuk metode *K-means*

Uji lanjut menggunakan LSD untuk metode *K-means - ALO* menghasilkan tiga kelompok dari percobaan jumlah *k cluster* dimana detail hasil pengujian dapat dilihat pada Lampiran 30. $K = 3$ menjadi jumlah *cluster* terbaik dengan nilai rata-rata *silhouette index* sebesar 0.6106 dan *standard error* sebesar 0.0004. Hasil dari LSD, semua *pValue* yang dihasilkan bernilai dibawah 0.05 yang berarti jumlah *k cluster* ini merupakan jumlah *cluster* yang paling signifikan diantara yang lain. Penjelasan Lampiran 28 dapat dilihat lebih mudah pada Gambar 4.6 dimana pada gambar tersebut semakin ke kanan menunjukkan semakin bagus kualitas *cluster* yang dihasilkan.



Gambar 4. 6 Visualisasi Least Significant Difference Segmen 4 untuk metode *K-means - ALO*

4.3.4 Segmen 5

Pengujian *anova* pada segmen 5 menghasilkan nilai Prob > F atau p-value sebesar 0 dapat dilihat pada tabel 4.14 dimana nilai tersebut kurang dari 0.05 sehingga jumlah *k* cluster berpengaruh terhadap *dataset* yang digunakan. Kemudian dilakukan uji lanjut LSD.

Tabel 4.14 Hasil ANOVA Segmen 5 Metode *K-means*

<i>Source</i>	SS	df	MS	F	Prob > F
<i>Groups</i>	6.1824	7	0.8832	6809.9628	0
<i>Error</i>	0.5177	3992	0.0001		
<i>Total</i>	6.7001	3999			

Pengujian *anova* untuk metode *K-means - ALO* dan dapat dilihat pada tabel 4.15 dimana nilai Prob > F atau p-value sebesar 0 sehingga dapat disimpulkan bahwa jumlah *k* cluster berpengaruh terhadap *dataset* yang digunakan. Uji lanjut LSD juga dilakukan terhadap metode *K-means - ALO*.

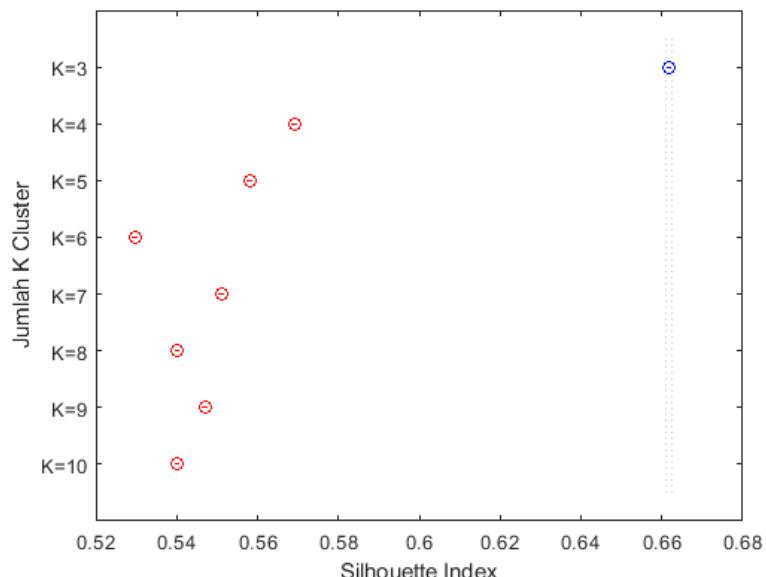
Tabel 4.15 Hasil ANOVA Segmen 5 Metode *K-means - ALO*

<i>Source</i>	SS	df	MS	F	Prob > F
<i>Groups</i>	4.8502	7	0.6929	4330.9072	0
<i>Error</i>	0.6387	3992	0.0001		

<i>Total</i>	5.4889	3999			
--------------	--------	------	--	--	--

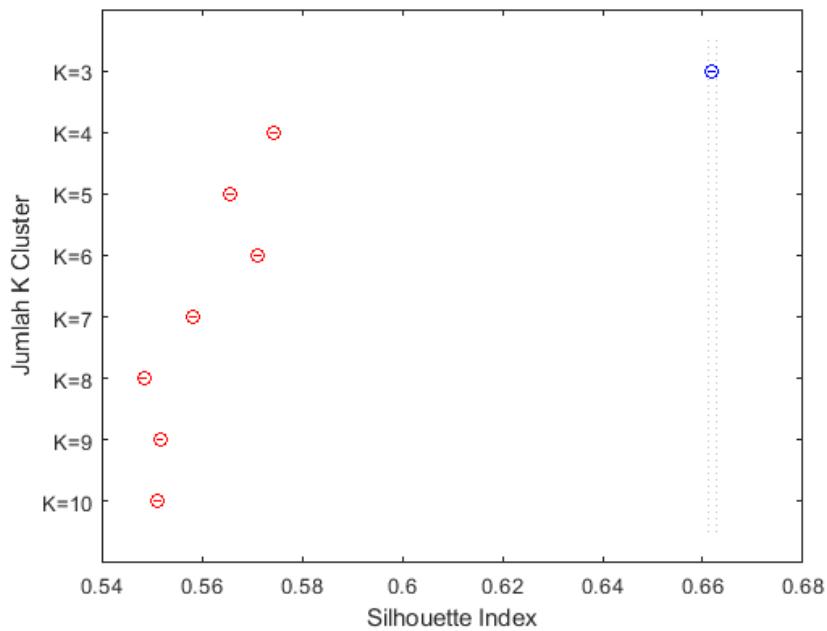
Uji lanjut menggunakan LSD untuk metode *K-means* menghasilkan tujuh kelompok dari percobaan jumlah *K cluster*, dimana detail hasil pengujian dapat dilihat pada Lampiran 31. $K = 3$ menjadi jumlah *cluster* terbaik dengan nilai rata-rata *silhouette index* sebesar 0.5622 dan *standard error* sebesar 0.0005.

Berdasarkan kelompok yang terbentuk pada pengujian menggunakan metode LSD untuk metode *K-means* dapat ditarik kesimpulan bahwa jumlah $K = 3$ memiliki nilai dibawah 0.05 yang berarti jumlah *k cluster* ini merupakan jumlah *cluster* yang paling signifikan diantara yang lain. Penjelasan Lampiran 31 dapat dilihat lebih mudah pada Gambar 4.7 dimana pada gambar tersebut semakin ke kanan menunjukkan semakin bagus kualitas *cluster* yang dihasilkan.



Gambar 4. 7 Visualisasi Least Significant Difference Segmen 5 untuk metode *K-means*

Uji lanjut menggunakan LSD untuk metode *K-means - ALO* menghasilkan tiga kelompok dari percobaan jumlah *k cluster* dimana detail hasil pengujian dapat dilihat pada Lampiran 32. $K = 3$ menjadi jumlah *cluster* terbaik dengan nilai rata-rata *silhouette index* sebesar 0.5728 dan *standard error* sebesar 0.0005. Hasil dari LSD, semua *pValue* yang dihasilkan bernilai dibawah 0.05 yang berarti jumlah *k cluster* ini merupakan jumlah *cluster* yang paling signifikan diantara yang lain. Penjelasan Lampiran 32 dapat dilihat lebih mudah pada Gambar 4.8 dimana pada gambar tersebut semakin ke kanan menunjukkan semakin bagus kualitas *cluster* yang dihasilkan.



Gambar 4. 8 Visualisasi Least Significant Difference Segmen 5 untuk metode *K-means - ALO*

4.3.5 Segmen 6

Pengujian *anova* pada segmen 6 menghasilkan nilai Prob > F atau p-value sebesar 2.3233 dapat dilihat pada tabel 4.16 dimana nilai tersebut lebih dari 0.05 sehingga jumlah *k cluster* tidak berpengaruh terhadap *dataset* yang digunakan. Maka tidak diperlukan uji lanjut menggunakan metode LSD.

Tabel 4.16 Hasil ANOVA Segmen 6 Metode *K-means*

Source	SS	df	MS	F	Prob > F
<i>Groups</i>	5.8698	7	0.8385	73.4435	2.3233
<i>Error</i>	45.5789	3992	0.0114		
<i>Total</i>	51.4487	3999			

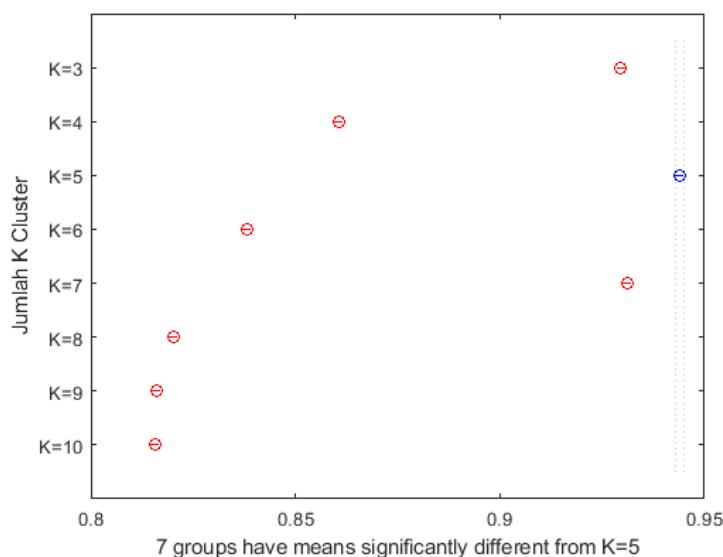
Pengujian *anova* untuk metode *K-means - ALO* dapat dilihat pada tabel 4.17 dimana nilai Prob > F atau p-value sebesar 0 sehingga dapat disimpulkan bahwa jumlah *k cluster* berpengaruh terhadap *dataset* yang digunakan. Uji lanjut LSD dilakukan terhadap metode *K-means - ALO*.

Tabel 4.17 Hasil ANOVA Segmen 6 Metode *K-means - ALO*

Source	SS	df	MS	F	Prob > F

<i>Groups</i>	11.0900	7	1.5843	5.7551	0
<i>Error</i>	1.0989	3992	2.7528		
<i>Total</i>	12.1889	3999			

Uji lanjut menggunakan LSD untuk metode *K-means - ALO* menghasilkan enam kelompok dari percobaan jumlah k cluster dimana detail hasil pengujian dapat dilihat pada Lampiran 33. $K = 5$ menjadi jumlah *cluster* terbaik dengan nilai rata-rata *silhouette index* sebesar 0.8694 dan *standard error* sebesar 7.4199. Hasil dari LSD, semua pValue yang dihasilkan bernilai dibawah 0.05 kecuali terhadap $K = 7$ yang berarti jumlah k cluster ini merupakan jumlah *cluster* yang paling signifikan diantara yang lain kecuali terhadap kedua jumlah $K = 7$. Penjelasan Lampiran 33 dapat dilihat lebih mudah pada Gambar 4.9 dimana pada gambar tersebut semakin ke kanan menunjukkan semakin bagus kualitas *cluster* yang dihasilkan.



Gambar 4. 9 *Visualisasi Least Significant Difference Segmen 6* untuk metode *K-means - ALO*

4.3.6 Segmen 7

Pengujian *anova* pada segmen 6 menghasilkan nilai Prob > F atau p-value sebesar 1.3980 dapat dilihat pada tabel 4.16 dimana nilai tersebut lebih dari 0.05 sehingga jumlah K cluster tidak berpengaruh terhadap *dataset* yang digunakan. Maka tidak diperlukan uji lanjut menggunakan metode LSD.

Tabel 4.18 Hasil ANOVA Segmen 7 Metode *K-means*

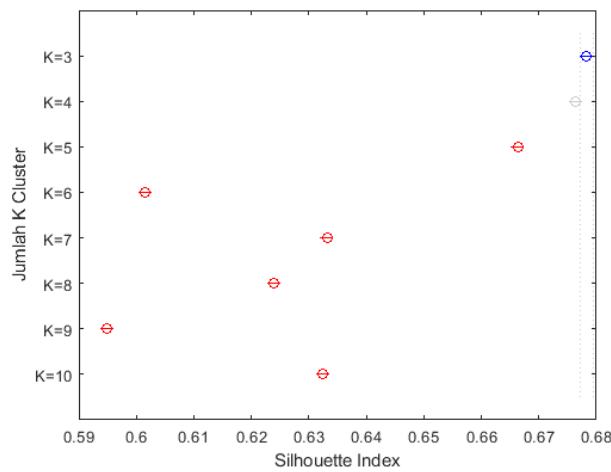
<i>Source</i>	SS	df	MS	F	Prob > F
<i>Groups</i>	0.0991	7	0.0142	12.9572	1.3980
<i>Error</i>	4.3601	3992	0.0011		
<i>Total</i>	4.4592	3999			

Pengujian *anova* untuk metode *K-means* - *ALO* dapat dilihat pada tabel 4.19 dimana nilai Prob > F atau p-value sebesar 0 sehingga dapat disimpulkan bahwa jumlah *k cluster* berpengaruh terhadap *dataset* yang digunakan. Uji lanjut LSD dilakukan terhadap metode *K-means* - *ALO*.

Tabel 4.19 Hasil ANOVA Segmen 7 Metode *K-means* - *ALO*

<i>Source</i>	SS	df	MS	F	Prob > F
<i>Groups</i>	4.8064	7	0.6866	2.0520	0
<i>Error</i>	1.3358	3992	3.3462		
<i>Total</i>	6.1422	3999			

Uji lanjut menggunakan LSD untuk metode *K-means - ALO* menghasilkan enam kelompok dari percobaan jumlah k cluster dimana detail hasil pengujian dapat dilihat pada Lampiran 33. $K = 4$ menjadi jumlah *cluster* terbaik dengan nilai rata-rata *silhouette index* sebesar 0.6424 dan *standard error* sebesar 8.1807. Hasil dari LSD, semua pValue yang dihasilkan bernilai dibawah 0.05 kecuali terhadap $K = 3$ yang berarti jumlah K cluster ini merupakan jumlah *cluster* yang paling signifikan diantara yang lain kecuali terhadap kedua jumlah $K = 3$. Penjelasan Lampiran 33 dapat dilihat lebih mudah pada Gambar 4.10 dimana pada gambar tersebut semakin ke kanan menunjukkan semakin bagus kualitas *cluster* yang dihasilkan.



Gambar 4. 10 Visualisasi Least Significant Difference Segmen 7 untuk metode *K-means - ALO*

[Halaman sengaja dikosongkan]

BAB 5

KESIMPULAN

5.1 Kesimpulan

1. Optimasi centroid *K-means* menggunakan Ant Lion Optimizer berhasil meningkatkan performa pengelompokan data, karena *K-means - ALO* selalu menghasilkan nilai validitas *cluster* yang lebih tinggi daripada *K-means classic* pada semua segmen kecuali tiga segmen saja.
2. Similaritas data pada setiap segmen berbeda-beda. Hal ini terbukti dari jumlah *cluster* optimum yang terbentuk dari pengujian statistik.

5.2 Saran

Perbaikan data sebelum di proses dan menghilangkan noise masih sangat perlu dilakukan. Di sisi lain eksplorasi metode pengelompokan data serta optimasi juga dapat membantu untuk menghasilkan nilai validitas cluster yang lebih optimal lagi.

[Halaman sengaja dikosongkan]

DAFTAR PUSTAKA

- Ali, B. S. (2013) ‘A Safety Assessment Framework for Automatic Dependent Surveillance Broadcast (ADS-B) and its Potential Impact on Aviation Safety’
Busyairah Syd Ali A thesis submitted for the degree of Doctor of Philosophy of the Imperial College London Centre for Transp’.
- Baek, K.-Y. and Bang, H.-C. (2012) ‘ADS-B based Trajectory Prediction and Conflict Detection for Air Traffic Management’, *International Journal of Aeronautical and Space Sciences*, 13(3), pp. 377–385. doi: 10.5139/IJASS.2012.13.3.377.
- Calinski, T. and Harabasz, J. (1974) ‘A dendrite method for cluster analysis’, *Communications in Statistics - Theory and Methods*, 3(1), pp. 1–27. doi: 10.1080/03610927408827101.
- Davies, D. L. and Bouldin, D. W. (1979) ‘A Cluster Separation Measure’, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-1(2), pp. 224–227. doi: 10.1109/TPAMI.1979.4766909.
- Gariel, M., Kunzi, F. and Hansman, R. J. (2011) ‘An algorithm for conflict detection in dense traffic using ADS-B’, *AIAA/IEEE Digital Avionics Systems Conference - Proceedings*, pp. 1–12. doi: 10.1109/DASC.2011.6095916.
- Han, J., Kamber, M. and Pei, J. (2012) *Data Mining: Concepts and Techniques*, San Francisco, CA, itd: Morgan Kaufmann. doi: 10.1016/B978-0-12-381479-1.00001-0.
- Jain, A. K. (2010) ‘Data clustering: 50 years beyond K-means’, *Pattern Recognition Letters*. Elsevier B.V., 31(8), pp. 651–666. doi: 10.1016/j.patrec.2009.09.011.
- Jegatha Deborah, L., Baskaran, R. and Kannan, A. (2010) ‘A Survey on Internal Validity Measure for Cluster Validation’, *International Journal of Computer Science & Engineering Survey*, 1(2), pp. 85–102. doi: 10.5121/ijcses.2010.1207.
- Jeon, D., Eun, Y. and Kim, H. (2015) ‘Estimation fusion with radar and ADS-B for air traffic surveillance’, *International Journal of Control, Automation and Systems*, 13(2), pp. 336–345. doi: 10.1007/s12555-014-0060-1.
- Kovács, F., Legány, C. and Babos, A. (2005) ‘Cluster validity measurement techniques’.

Pusadan et al. (2016) ‘Anomaly detection of flight routes through optimal waypoint’, *Journal of Physics: Conference Series*, 755, p. 011001. doi: 10.1088/1742-6596/755/1/011001.

Majhi, S. K. and Biswal, S. (2018) ‘ScienceDirect Optimal cluster analysis using hybrid K-Means and Ant Lion Optimizer’, *Karbala International Journal of Modern Science*. Elsevier Ltd. doi: 10.1016/j.kijoms.2018.09.001.

Martel, F. et al. (2013) ‘Unmanned Aircraft Systems Sense and Avoid Avionics Utilizing ADS-B Transceiver’, (April 2009), pp. 1–8.

Mirjalili, S. (2015) ‘The ant lion optimizer’, *Advances in Engineering Software*. Elsevier Ltd, 83, pp. 80–98. doi: 10.1016/j.advengsoft.2015.01.010.

Orefice, M. et al. (2014) ‘Aircraft conflict detection based on ADS-B surveillance data’, *2014 IEEE International Workshop on Metrology for Aerospace, MetroAeroSpace 2014 - Proceedings*, pp. 277–282. doi: 10.1109/MetroAeroSpace.2014.6865934.

Pourvoyeur, K. and Heidger, R. (2014) ‘Secure ADS-B usage in ATC tracking’, *2014 Tyrrhenian International Workshop on Digital Communications - Enhanced Surveillance of Aircraft and Vehicles, TIWDC/ESAV 2014*, pp. 35–40. doi: 10.1109/TIWDC-ESAV.2014.6945444.

Rousseeuw, P. J. (1987) ‘Silhouettes: A graphical aid to the interpretation and validation of cluster analysis’, *Journal of Computational and Applied Mathematics*, 20(C), pp. 53–65. doi: 10.1016/0377-0427(87)90125-7.

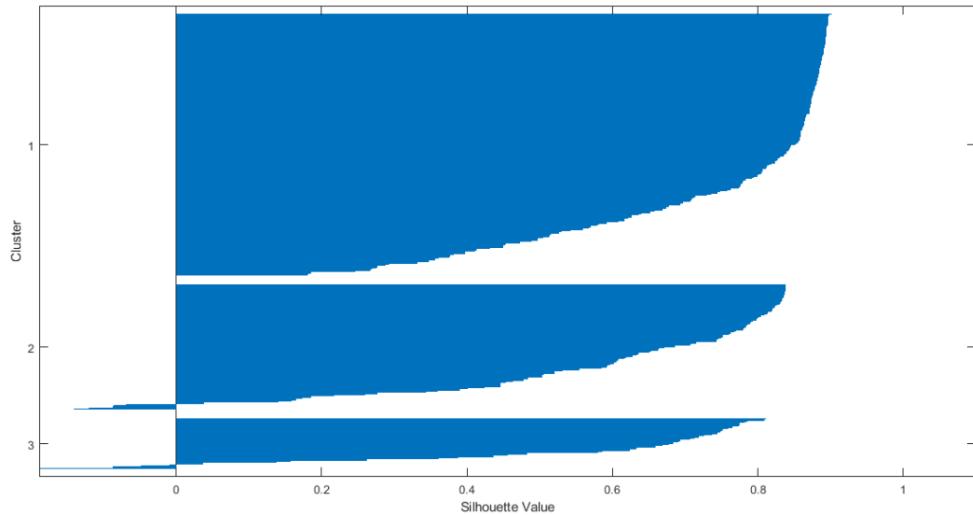
Tang, X. et al. (2015) ‘A flight profile clustering method combining twed with K-means algorithm for 4D trajectory prediction’, *2015 Integrated Communication, Navigation and Surveillance Conference (ICNS)*, pp. S3-1-S3-9. doi: 10.1109/ICNSURV.2015.7121260.

Vij, A. and Khandnor, P. (2016) ‘Validity of internal cluster indices’, *2016 International Conference on Computation System and Information Technology for Sustainable Solutions, CSITSS 2016*, pp. 388–395. doi: 10.1109/CSITSS.2016.7779392.

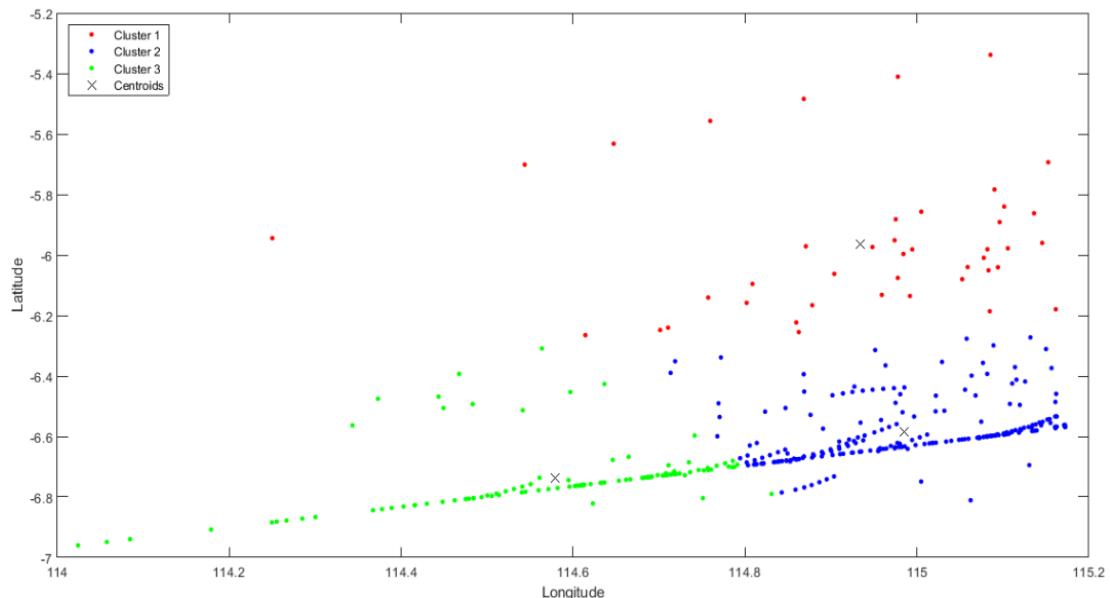
Xiong, H. and Li, Z. (2014) ‘Understanding and Enhancement of Internal Clustering Validation Measures’, Aggarwal, C. C., & Reddy, C. K. (Eds.). (2014). *Data Clustering: Algorithms and Applications*. Boca Raton, FL: CRC., 43(3), pp. 571–605. doi: doi:10.1201/b15410-24.

LAMPIRAN

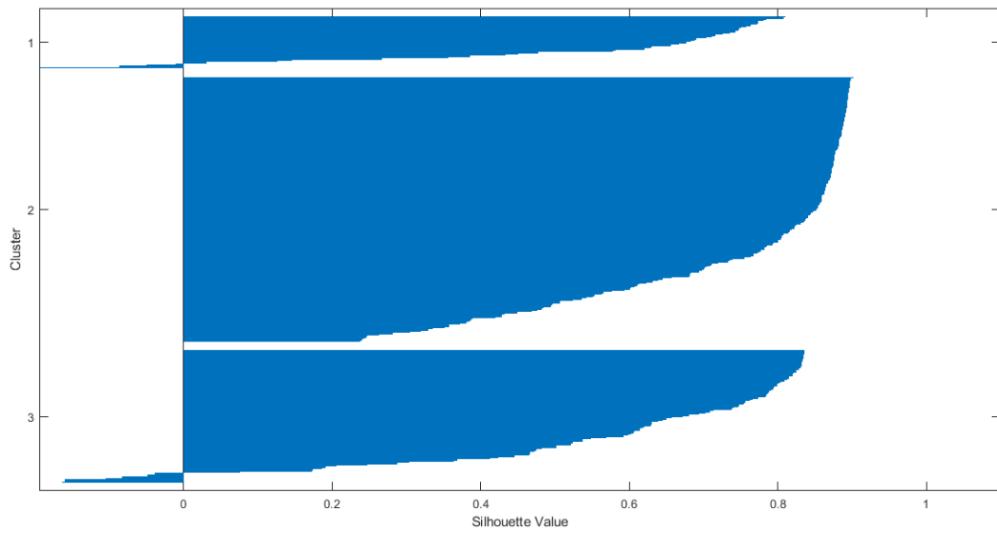
Lampiran 1. Visualisasi *Silhouette Index* pada Segmen 2 untuk jumlah k cluster = 3 dengan metode *K-means*



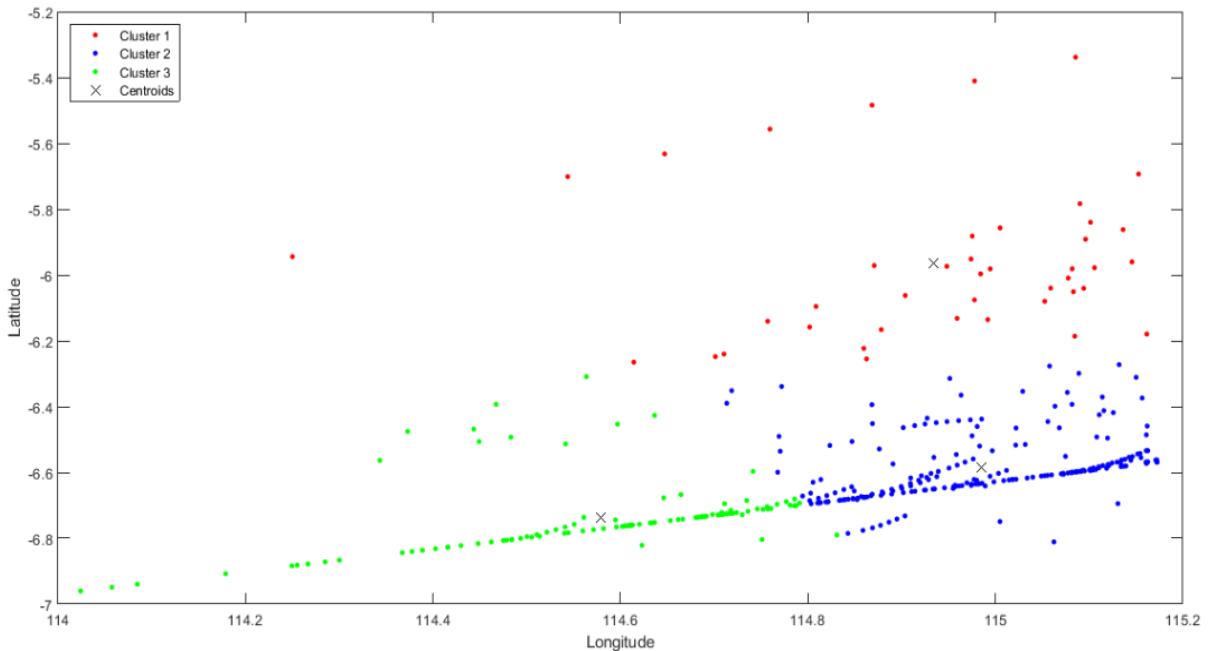
Lampiran 2. Visualisasi data pada Segmen 2 untuk jumlah k cluster = 3 dengan metode *K-means*



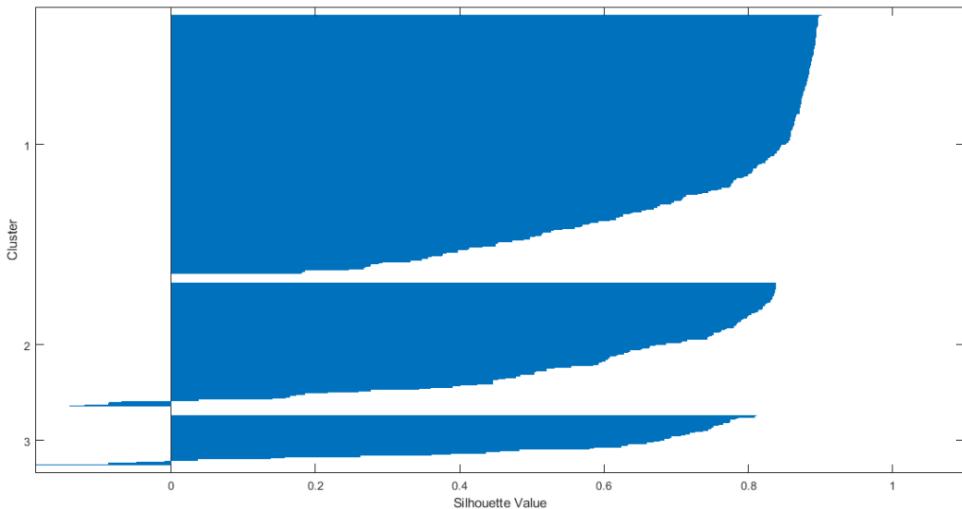
Lampiran 3. Visualisasi *Silhouette Index* pada Segmen 2 untuk jumlah k cluster = 3 dengan metode *K-means - ALO*



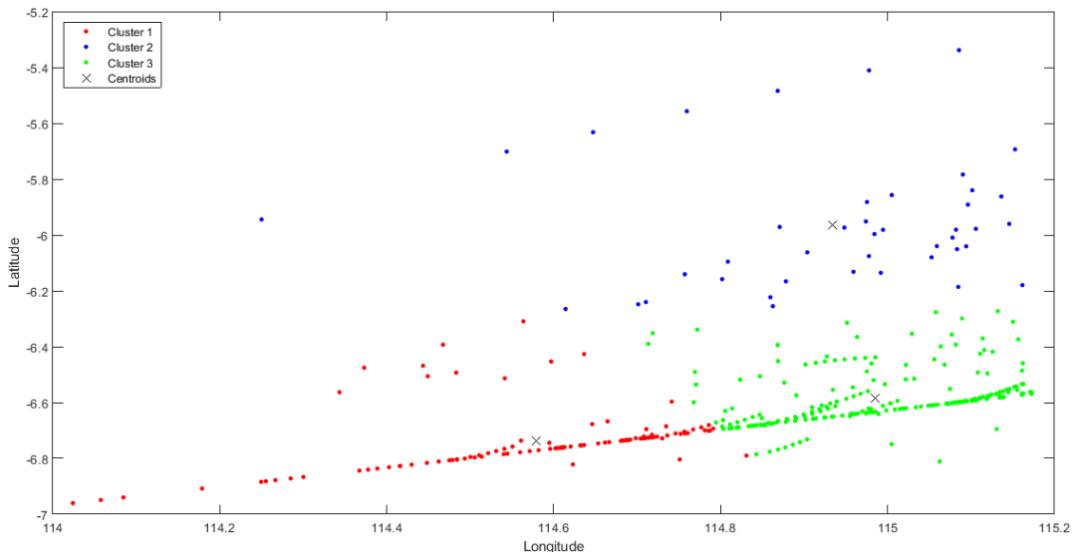
Lampiran 4. Visualisasi data pada Segmen 2 untuk jumlah k cluster = 3 dengan metode K -means - ALO



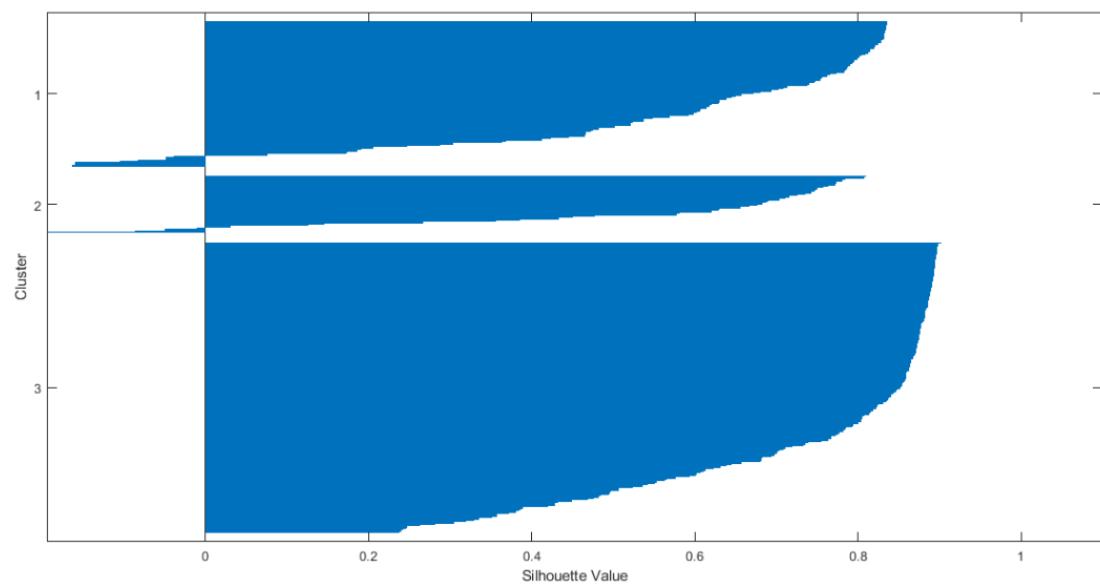
Lampiran 5. Visualisasi Silhouette Index pada Segmen 3 untuk jumlah k cluster = 3 dengan metode K -means



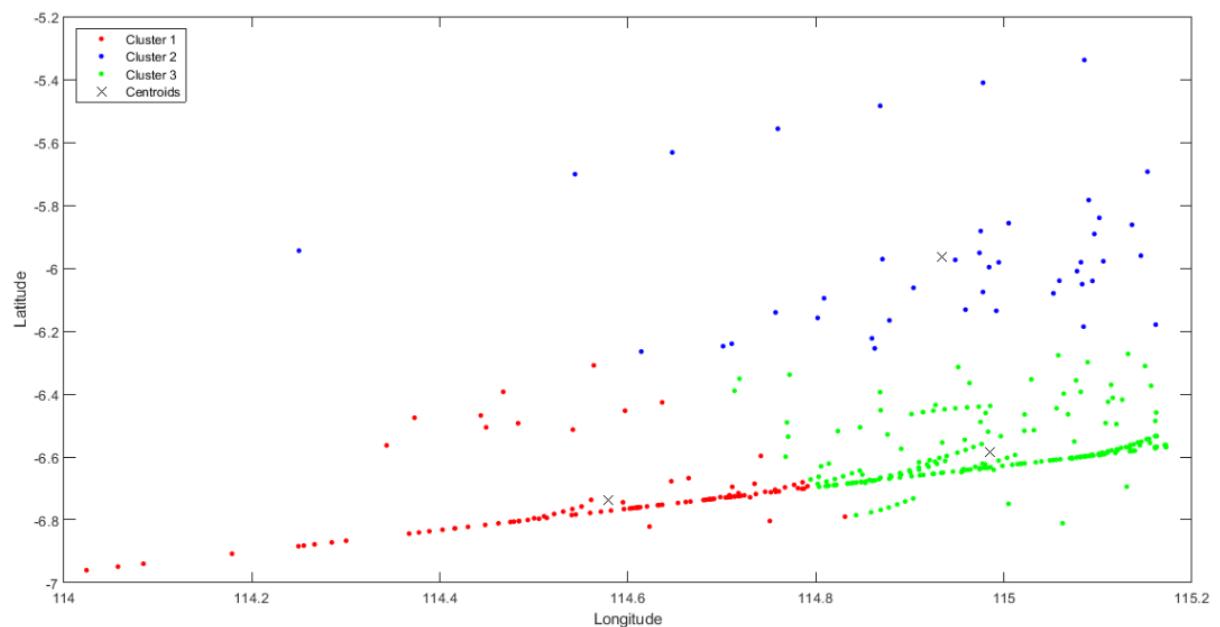
Lampiran 6. Visualisasi data pada Segmen 3 untuk jumlah k cluster = 3 dengan metode $K\text{-means}$



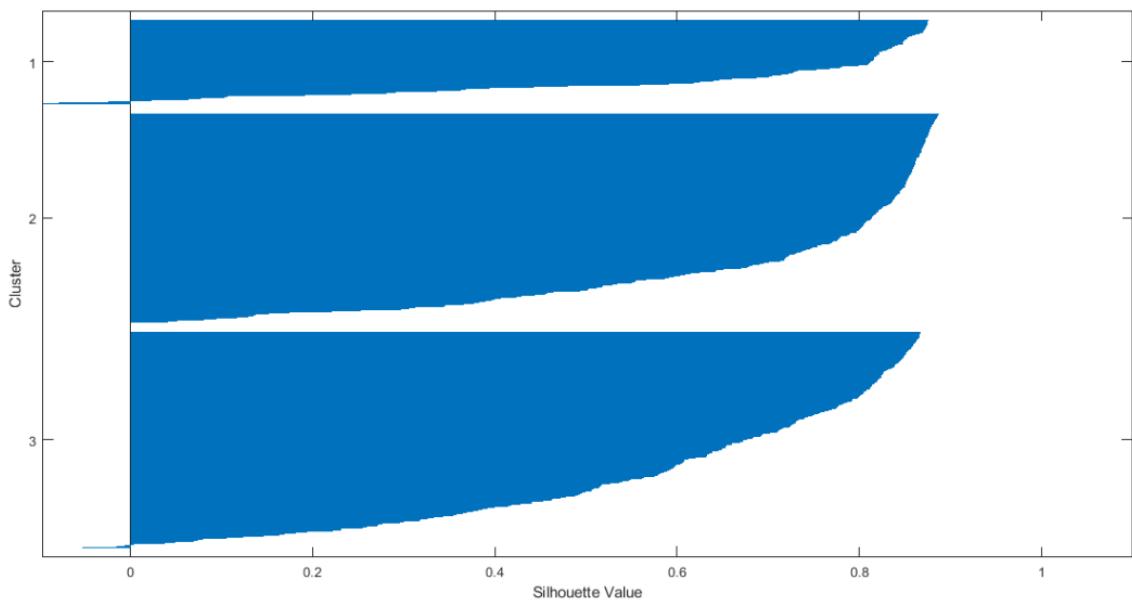
Lampiran 7. Visualisasi *Silhouette Index* pada Segmen 3 untuk jumlah k cluster = 3 dengan metode $K\text{-means}$ - ALO



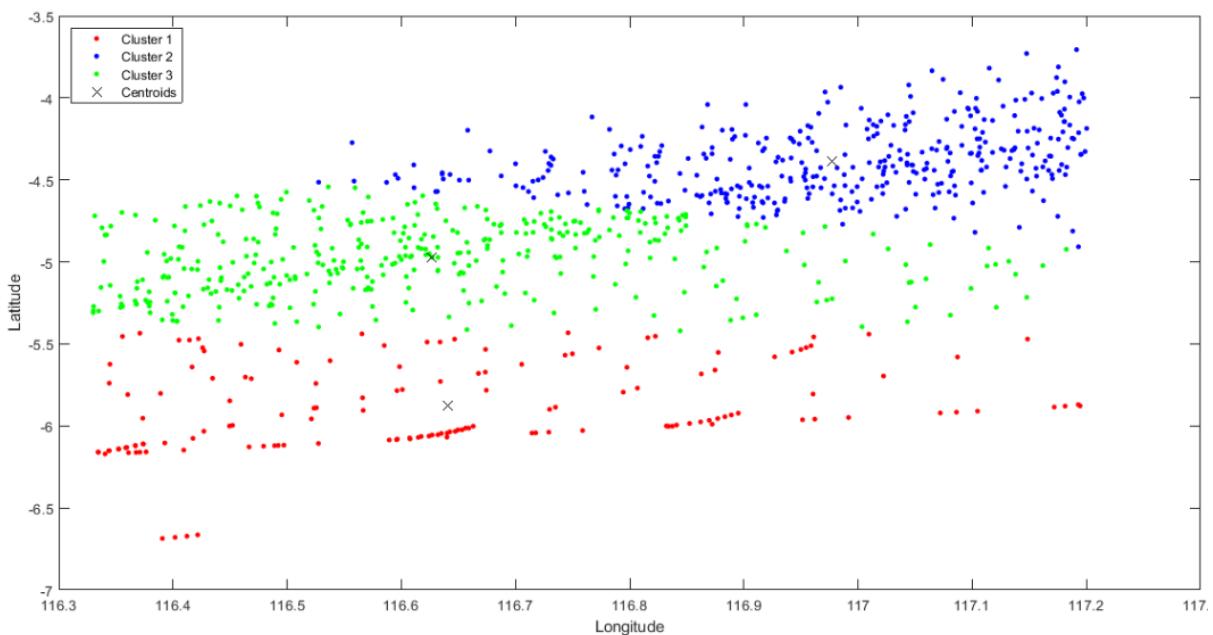
Lampiran 8. Visualisasi data pada Segmen 3 untuk jumlah k cluster = 3 dengan metode *K-means - ALO*



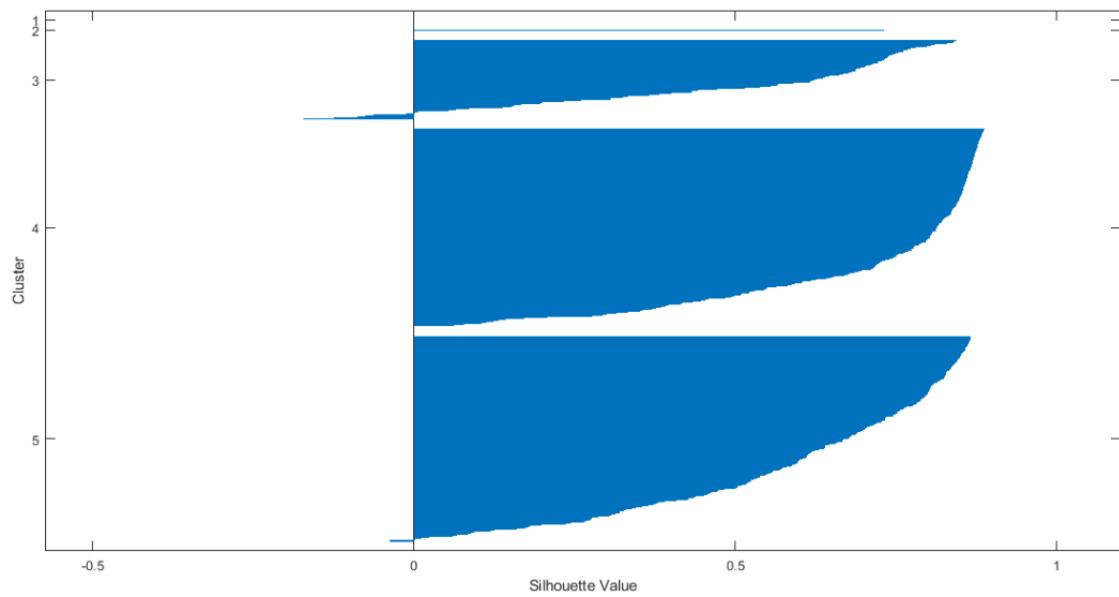
Lampiran 9. Visualisasi *Silhouette Index* pada Segmen 4 untuk jumlah k cluster = 3 dengan metode *K-means*



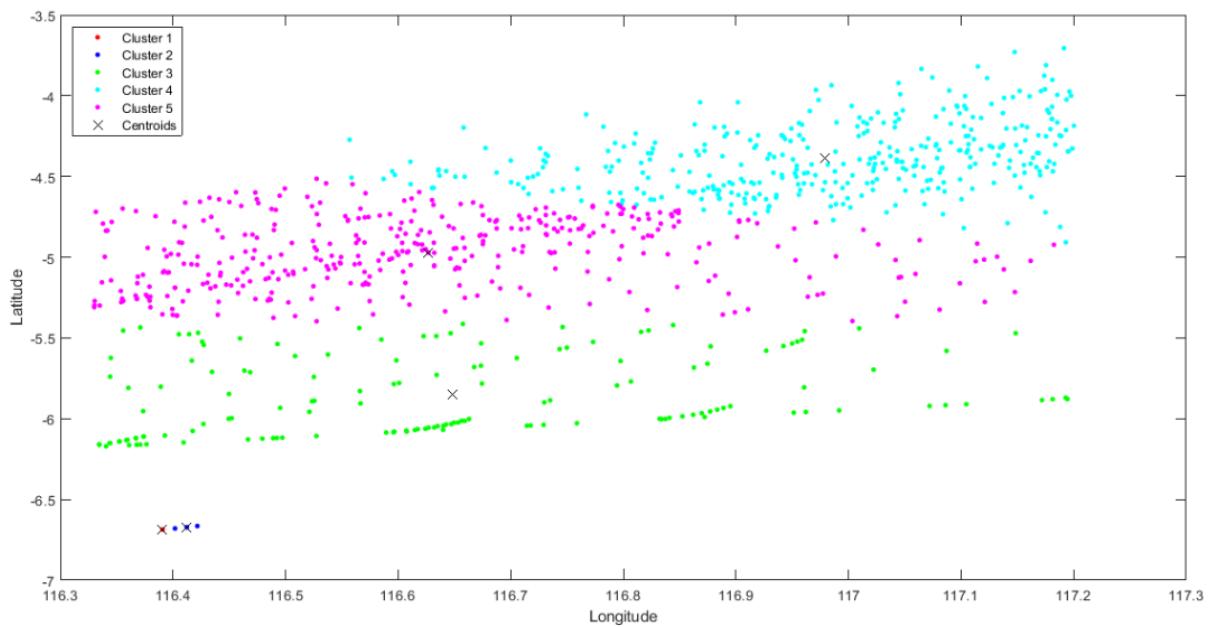
Lampiran 10. Visualisasi data pada Segmen 4 untuk jumlah k cluster = 3 dengan metode *K-means*



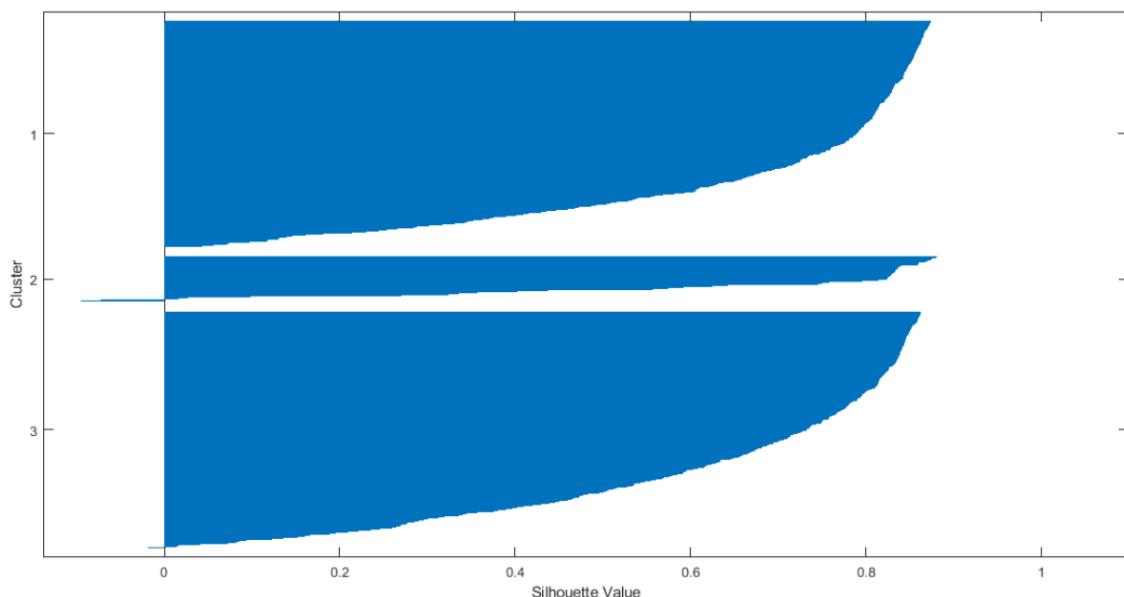
Lampiran 11. Visualisasi *Silhouette Index* pada Segmen 4 untuk jumlah k cluster = 5 dengan metode *K-means - ALO*



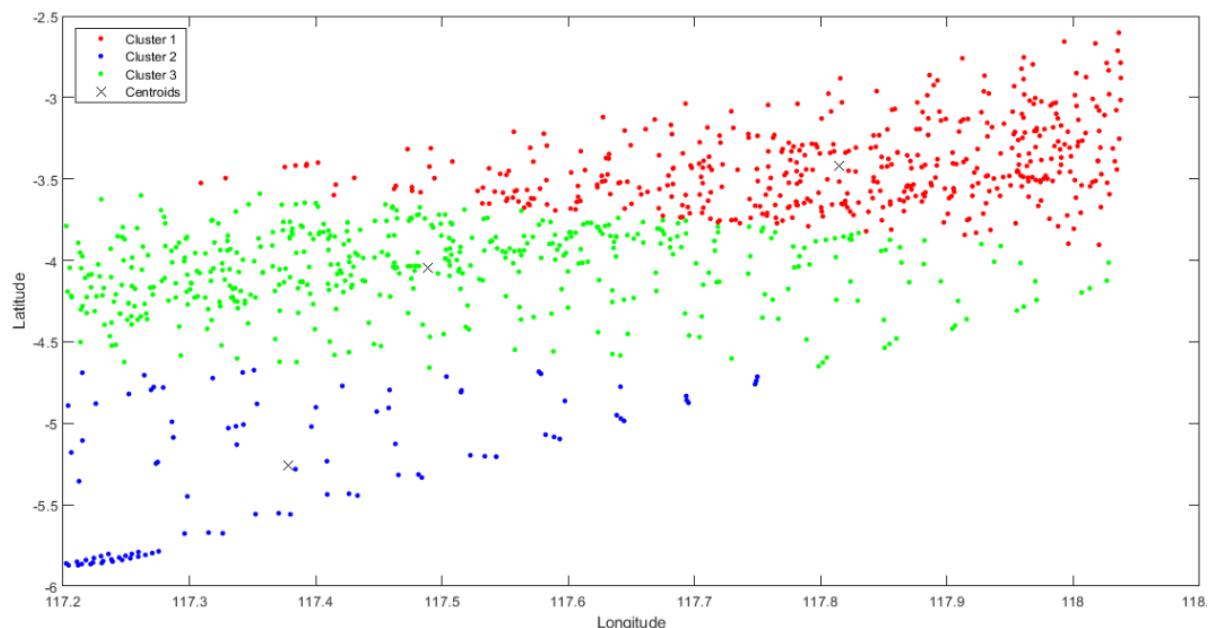
Lampiran 12. Visualisasi data pada Segmen 4 untuk jumlah k cluster = 5 dengan metode K -means - ALO



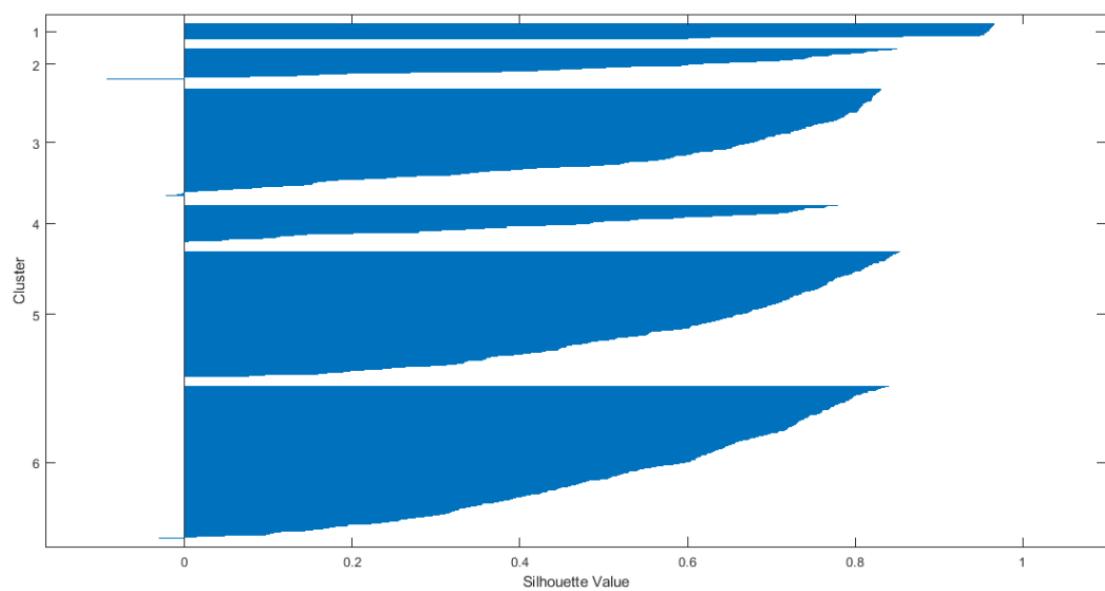
Lampiran 13. Visualisasi Silhouette Index pada Segmen 5 untuk jumlah k cluster = 3 dengan metode K -means



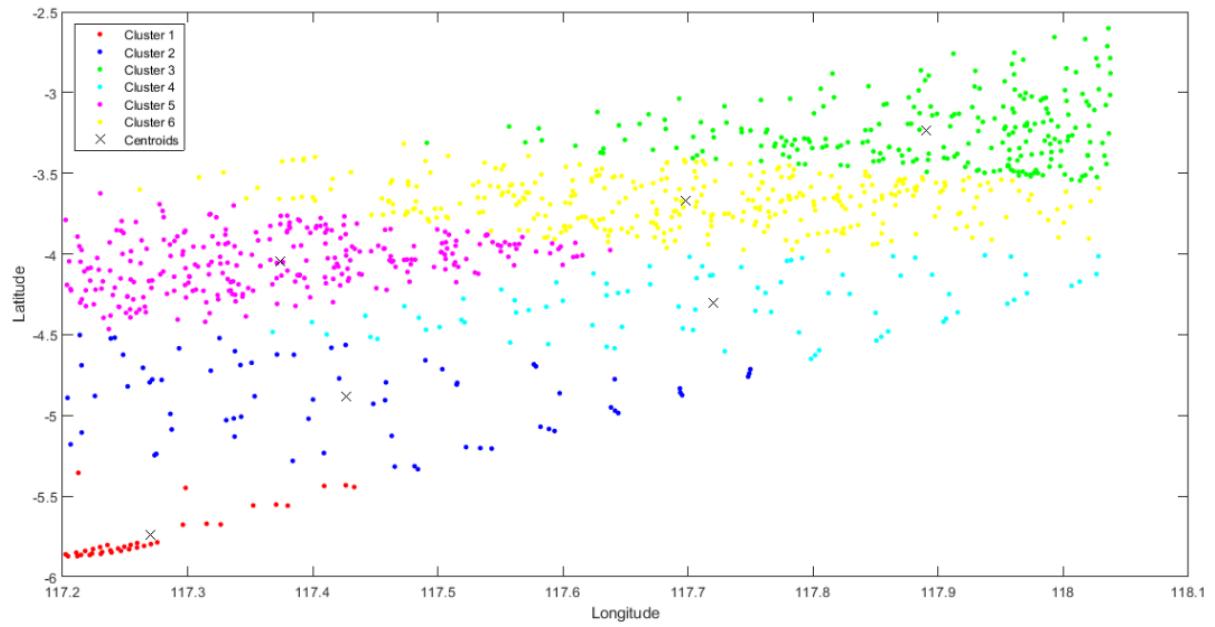
Lampiran 14. Visualisasi data pada Segmen 5 untuk jumlah k cluster = 3 dengan metode K -means



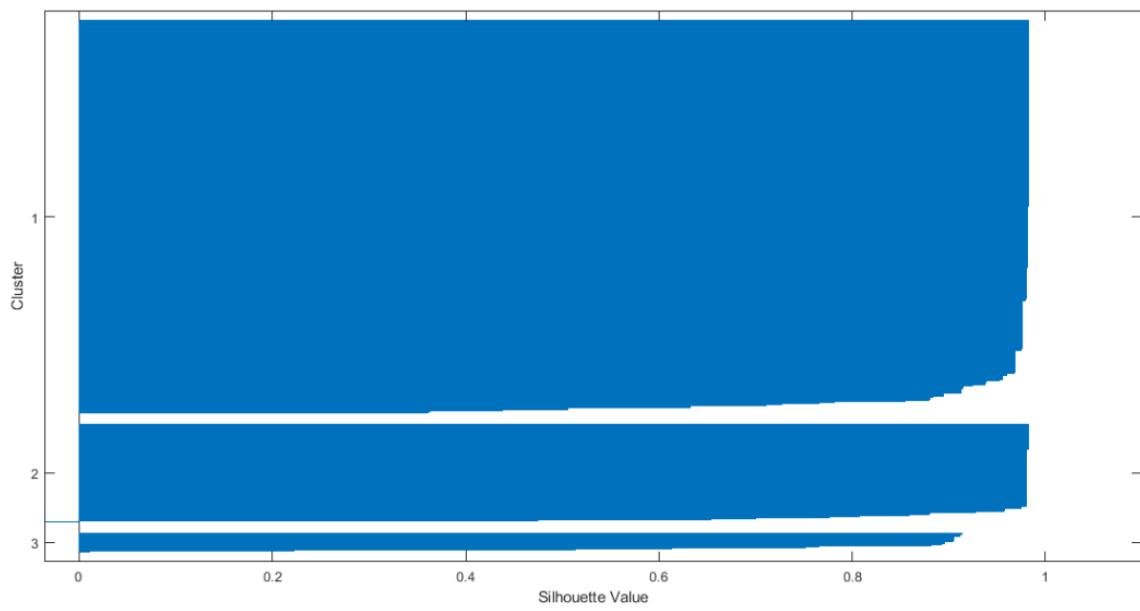
Lampiran 15. Visualisasi *Silhouette Index* pada Segmen 5 untuk jumlah k cluster = 6 dengan metode *K-means* - *ALO*



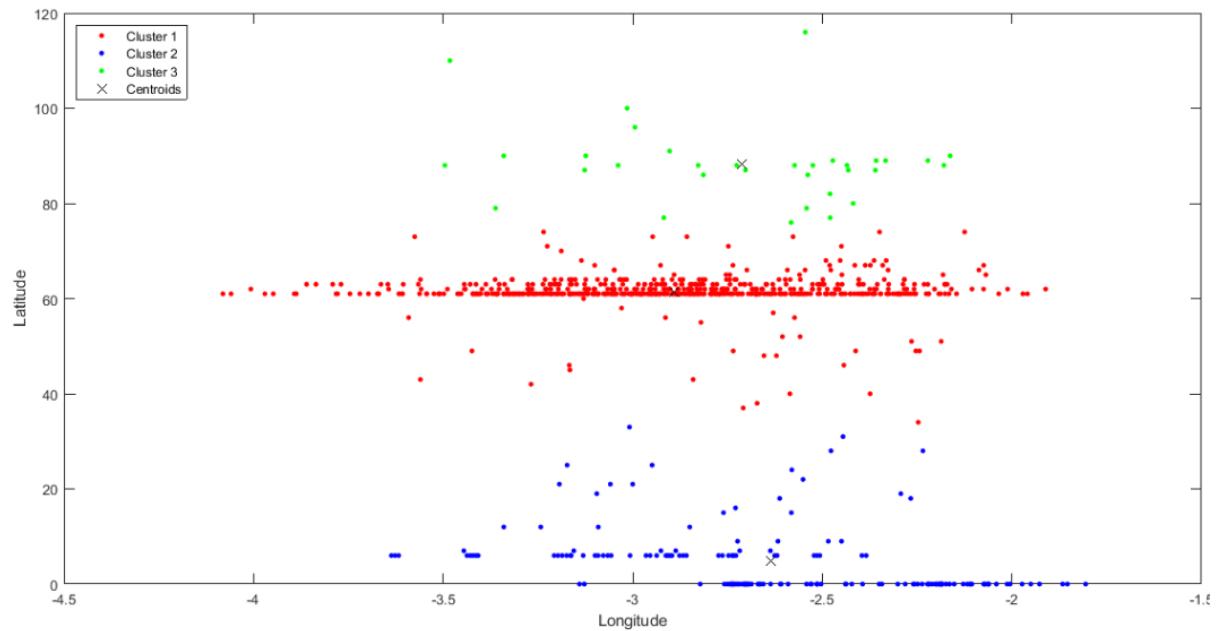
Lampiran 16. Visualisasi data pada Segmen 5 untuk jumlah k cluster = 6 dengan metode *K-means* - *ALO*



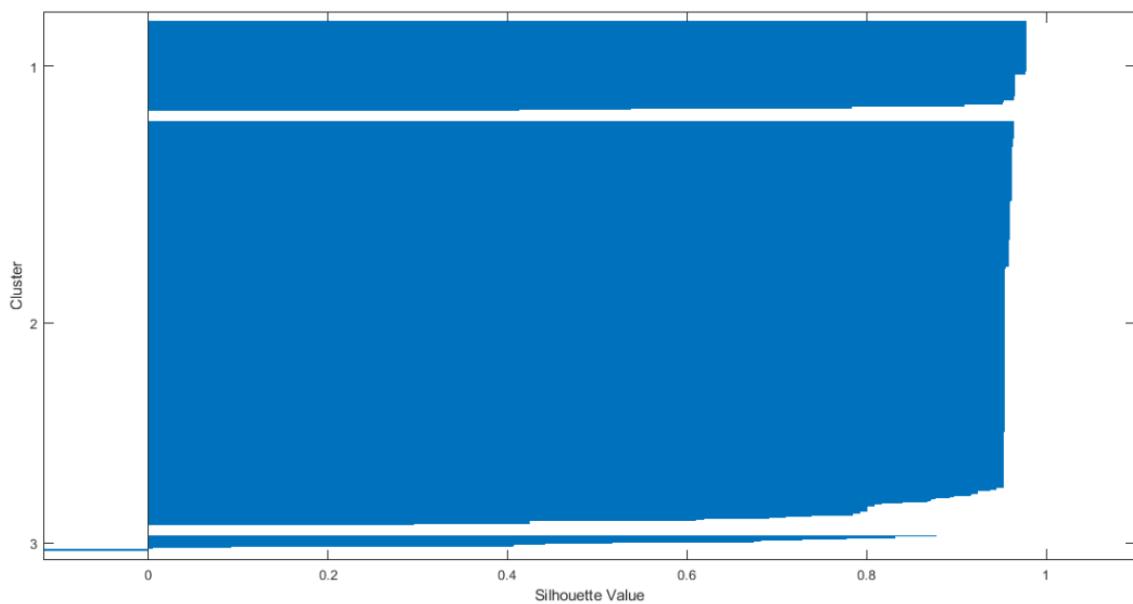
Lampiran 17. Visualisasi *Silhouette Index* pada Segmen 6 untuk jumlah k cluster = 3 dengan metode *K-means*



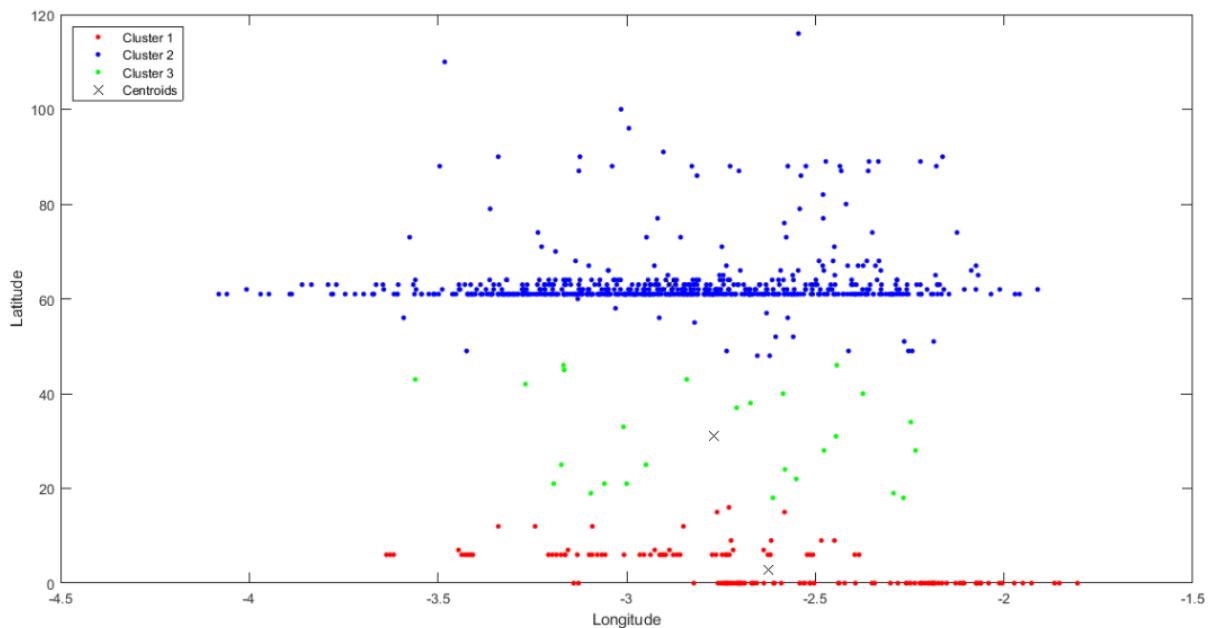
Lampiran 18. Visualisasi data pada Segmen 6 untuk jumlah k cluster = 3 dengan metode *K-means*



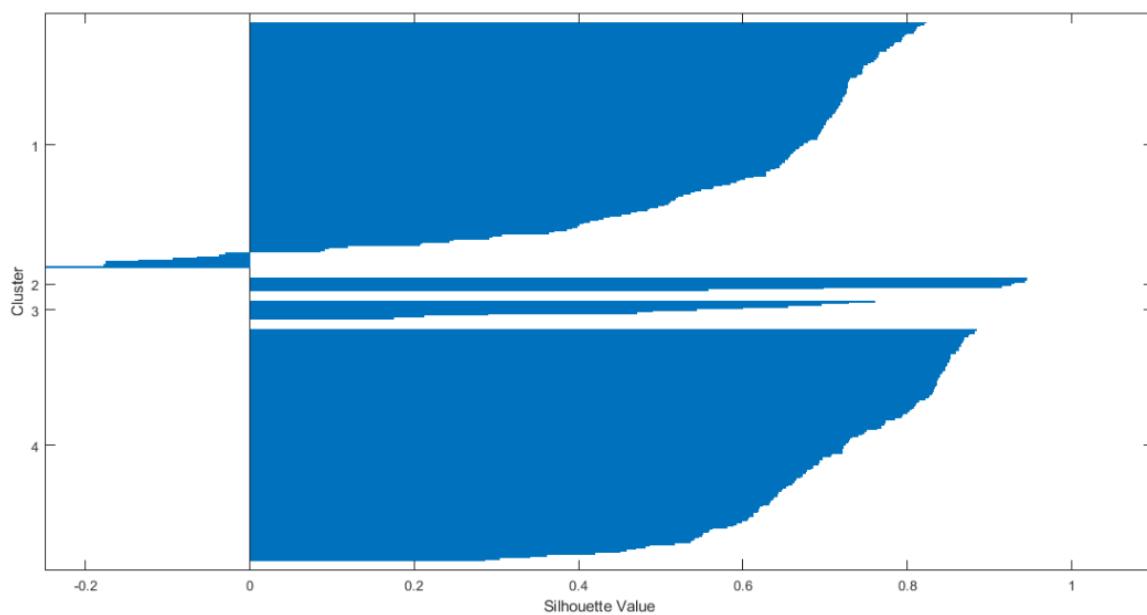
Lampiran 19. Visualisasi *Silhouette Index* pada Segmen 6 untuk jumlah k cluster = 3 dengan metode *K-means* - ALO



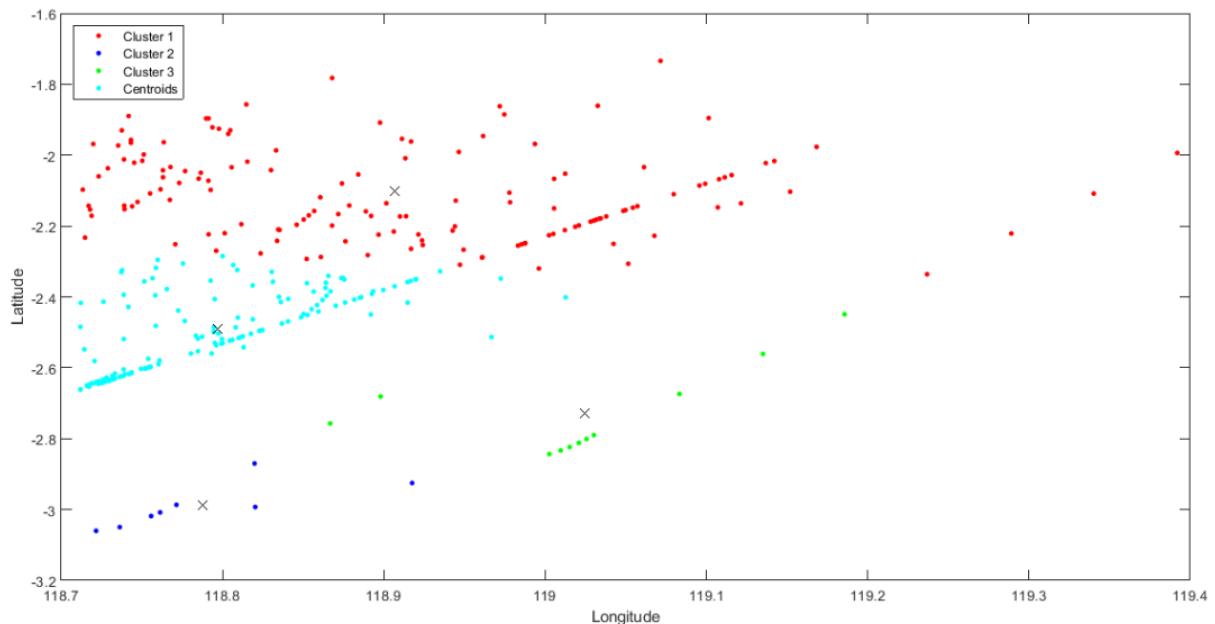
Lampiran 20. Visualisasi data pada Segmen 6 untuk jumlah k cluster = 3 dengan metode *K-means* - ALO



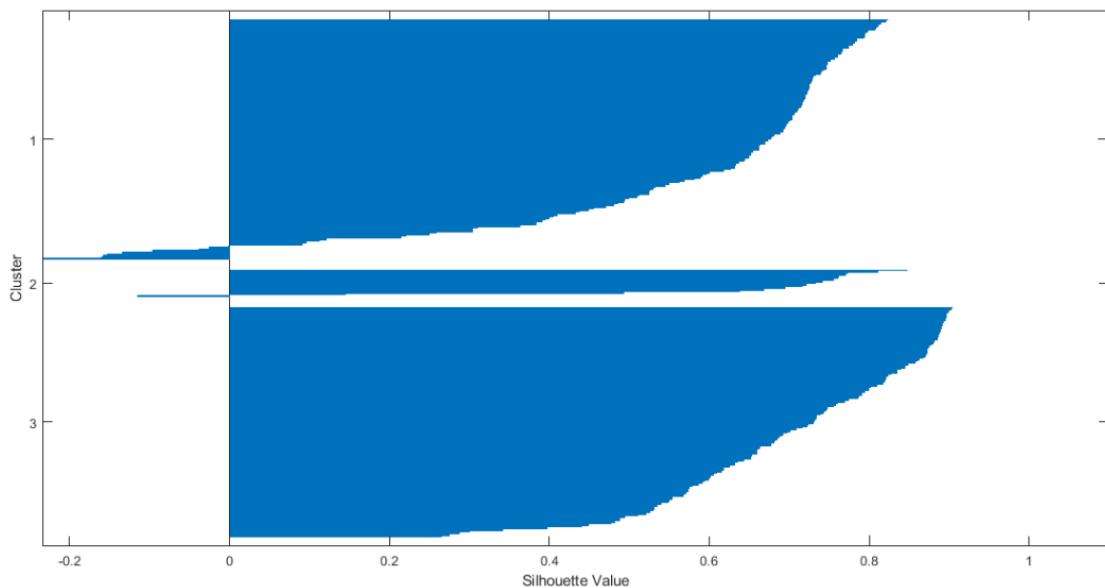
Lampiran 21. Visualisasi Silhouette Index pada Segmen 7 untuk jumlah k cluster = 4 dengan metode *K-means*



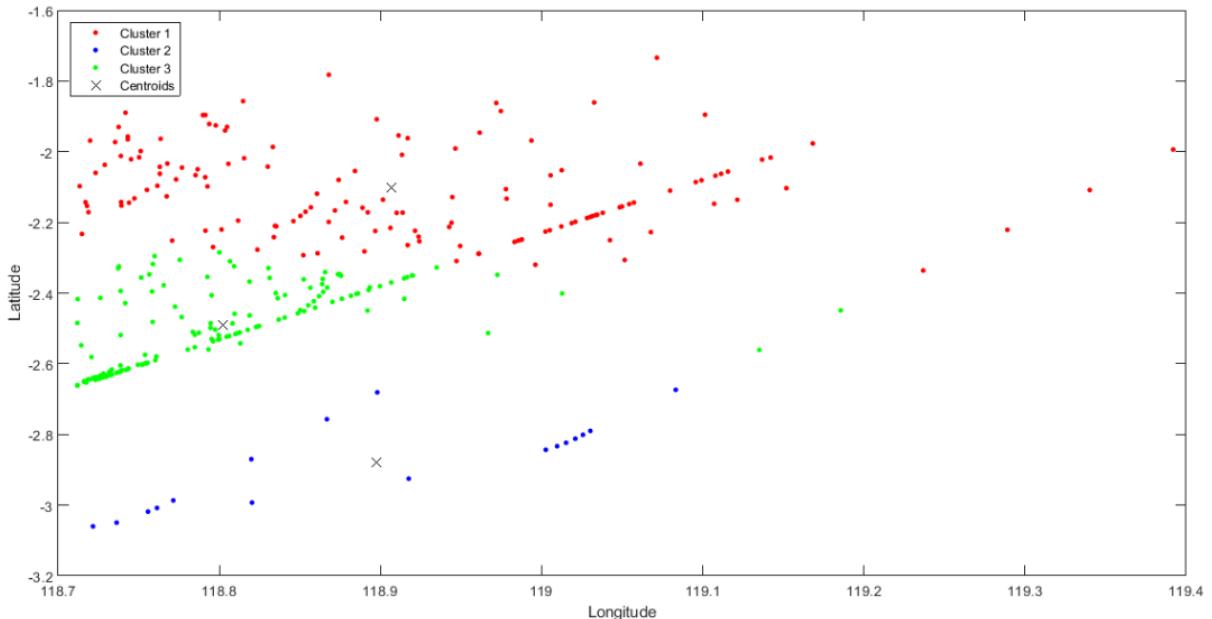
Lampiran 22. Visualisasi data pada Segmen 7 untuk jumlah k cluster = 4 dengan metode *K-means*



Lampiran 23. Visualisasi Silhouette Index pada Segmen 7 untuk jumlah k cluster = 3 dengan metode K-means - ALO



Lampiran 24. Visualisasi data pada Segmen 7 untuk jumlah k cluster = 3 dengan metode K-means - ALO



Lampiran 25. Hasil Least Significant Difference pada Segmen 2 untuk metode K-means

A	B	Lower Bound	$\mu_{AB} - \mu_{AB}$	Upper Bound	pValue

K = 3	K = 4	0.0582	0.0619	0.0656	4.37E-207
K = 3	K = 5	0.0780	0.0817	0.0854	0
K = 3	K = 6	0.0835	0.0872	0.0910	0
K = 3	K = 7	0.0859	0.0896	0.0933	0
K = 3	K = 8	0.0867	0.0904	0.0941	0
K = 3	K = 9	0.0806	0.0843	0.0881	0
K = 3	K = 10	0.0725	0.0763	0.0800	7.48E-297
K = 4	K = 5	0.0161	0.0198	0.0235	3.73E-25
K = 4	K = 6	0.0216	0.0253	0.0291	7.93E-40
K = 4	K = 7	0.0240	0.0277	0.0314	4.77E-47
K = 4	K = 8	0.0248	0.0285	0.0322	1.35E-49
K = 4	K = 9	0.0187	0.0224	0.0262	9.32E-32
K = 4	K = 10	0.0106	0.0144	0.0181	4.58E-14
K = 5	K = 6	0.0018	0.0055	0.0093	0.003494583
K = 5	K = 7	0.0042	0.0079	0.0116	3.19E-05
K = 5	K = 8	0.0050	0.0087	0.0124	4.78E-06
K = 5	K = 9	-0.0011	0.0027	0.0064	0.162444619
K = 5	K = 10	-0.0091	-0.0054	-0.0017	0.00424411
K = 6	K = 7	-0.0014	0.0024	0.0061	0.214383163
K = 6	K = 8	-0.0006	0.0031	0.0069	0.097383765
K = 6	K = 9	-0.0066	-0.0029	0.0008	0.127313495
K = 6	K = 10	-0.0147	-0.0110	-0.0073	7.88E-09
K = 7	K = 8	-0.0029	0.0008	0.0045	0.677241865
K = 7	K = 9	-0.0090	-0.0053	-0.0015	0.005685276
K = 7	K = 10	-0.0171	-0.0133	-0.0096	2.50E-12
K = 8	K = 9	-0.0098	-0.0060	-0.0023	0.00146796
K = 8	K = 10	-0.0178	-0.0141	-0.0104	1.21E-13
K = 9	K = 10	-0.0118	-0.0081	-0.0044	2.11E-05

Lampiran 26. Hasil Least Significant Difference pada Segmen 2 untuk metode K-means - ALO

A	B	Lower Bound	$\mu_{AB} - \mu_{AB}$	Upper Bound	pValue
K = 3	K = 4	-0.0004	0.0019	0.0042	0.10018

K = 3	K = 5	0.0096	0.0119	0.0142	2.64E-24
K = 3	K = 6	0.0744	0.0767	0.0789	0
K = 3	K = 7	0.0428	0.0450	0.0473	4.01E-280
K = 3	K = 8	0.0520	0.0543	0.0566	0
K = 3	K = 9	0.0811	0.0834	0.0857	0
K = 3	K = 10	0.0435	0.0458	0.0481	3.89E-288
K = 4	K = 5	0.0077	0.0100	0.0122	1.18E-17
K = 4	K = 6	0.0725	0.0747	0.0770	0
K = 4	K = 7	0.0409	0.0431	0.0454	3.52E-260
K = 4	K = 8	0.0501	0.0524	0.0547	0
K = 4	K = 9	0.0792	0.0815	0.0838	0
K = 4	K = 10	0.0416	0.0439	0.0462	4.90E-268
K = 5	K = 6	0.0625	0.0648	0.0670	0
K = 5	K = 7	0.0309	0.0332	0.0354	9.48E-164
K = 5	K = 8	0.0402	0.0424	0.0447	6.79E-253
K = 5	K = 9	0.0692	0.0715	0.0738	0
K = 5	K = 10	0.0317	0.0339	0.0362	1.59E-170
K = 6	K = 7	-0.0339	-0.0316	-0.0293	4.76E-150
K = 6	K = 8	-0.0246	-0.0223	-0.0201	4.15E-79
K = 6	K = 9	0.0045	0.0067	0.0090	6.86E-09
K = 6	K = 10	-0.0331	-0.0308	-0.0286	1.36E-143
K = 7	K = 8	0.0070	0.0093	0.0115	1.82E-15
K = 7	K = 9	0.0361	0.0383	0.0406	5.74E-212
K = 7	K = 10	-0.0015	0.0008	0.0030	0.514628
K = 8	K = 9	0.0268	0.0291	0.0314	6.72E-129
K = 8	K = 10	-0.0108	-0.0085	-0.0062	2.70E-13
K = 9	K = 10	-0.0399	-0.0376	-0.0353	1.17E-204

Lampiran 27. Hasil Least Significant Difference pada Segmen 3 untuk metode K-means

A	B	Lower Bound	$\mu_{AB} - \mu_{AB}$	Upper Bound	pValue
K = 3	K = 4	0.0566	0.0603	0.0641	1.80E-195
K = 3	K = 5	0.0784	0.0822	0.0859	0.00E+00
K = 3	K = 6	0.0857	0.0894	0.0932	0

K = 3	K = 7	0.0891	0.0928	0.0966	0.00E+00
K = 3	K = 8	0.0881	0.0918	0.0956	0
K = 3	K = 9	0.0822	0.0859	0.0896	0
K = 3	K = 10	0.0724	0.0762	0.0799	8.93E-293
K = 4	K = 5	0.0181	0.0218	0.0256	1.04E-29
K = 4	K = 6	0.0253	0.0291	0.0328	7.99E-51
K = 4	K = 7	0.0287	0.0325	0.0362	1.33E-62
K = 4	K = 8	0.0277	0.0315	0.0352	5.73E-59
K = 4	K = 9	0.0218	0.0256	0.0293	6.31E-40
K = 4	K = 10	0.0121	0.0158	0.0196	1.53E-16
K = 5	K = 6	0.0035	0.0073	0.0110	0.000149
K = 5	K = 7	0.0069	0.0107	0.0144	2.51E-08
K = 5	K = 8	0.0059	0.0097	0.0134	4.62E-07
K = 5	K = 9	0.0000	0.0037	0.0075	0.050194
K = 5	K = 10	-0.0097	-0.0060	-0.0022	1.80E-03
K = 6	K = 7	-0.0003	0.0034	0.0072	7.40E-02
K = 6	K = 8	-0.0014	0.0024	0.0061	2.10E-01
K = 6	K = 9	-0.0073	-0.0035	0.0002	6.63E-02
K = 6	K = 10	-0.0170	-0.0132	-0.0095	5.26E-12
K = 7	K = 8	-0.0048	-0.0010	0.0027	5.94E-01
K = 7	K = 9	-0.0107	-0.0069	-0.0032	2.93E-04
K = 7	K = 10	-0.0204	-0.0166	-0.0129	4.51E-18
K = 8	K = 9	-0.0097	-0.0059	-0.0022	2.01E-03
K = 8	K = 10	-0.0194	-0.0156	-0.0119	3.99E-16
K = 9	K = 10	-0.0135	-0.0097	-0.0060	3.90E-07

Lampiran 28. Hasil Least Significant Difference pada Segmen 3 untuk metode K-means - ALO

A	B	Lower Bound	$\mu_{AB} - \bar{\mu}_{AB}$	Upper Bound	pValue
K = 3	K = 4	-0.0004	0.0021	0.0046	1.04E-01
K = 3	K = 5	0.0094	0.0119	0.0144	1.96E-20
K = 3	K = 6	0.0742	0.0767	0.0792	0
K = 3	K = 7	0.0129	0.0154	0.0179	5.12E-33
K = 3	K = 8	0.0910	0.0935	0.0960	0

K = 3	K = 9	0.0542	0.0567	0.0592	0
K = 3	K = 10	0.0773	0.0798	0.0823	0.00E+00
K = 4	K = 5	0.0073	0.0098	0.0123	1.84E-14
K = 4	K = 6	0.0721	0.0746	0.0771	0.00E+00
K = 4	K = 7	0.0108	0.0133	0.0158	2.98E-25
K = 4	K = 8	0.0890	0.0915	0.0940	0.00E+00
K = 4	K = 9	0.0522	0.0547	0.0572	0.00E+00
K = 4	K = 10	0.0753	0.0778	0.0803	0.00E+00
K = 5	K = 6	0.0623	0.0648	0.0673	0
K = 5	K = 7	0.0010	0.0035	0.0060	5.75E-03
K = 5	K = 8	0.0792	0.0817	0.0842	0.00E+00
K = 5	K = 9	0.0424	0.0449	0.0474	2.17E-236
K = 5	K = 10	0.0655	0.0680	0.0705	0.00E+00
K = 6	K = 7	-0.0637	-0.0612	-0.0587	0.00E+00
K = 6	K = 8	0.0144	0.0169	0.0194	3.20E-39
K = 6	K = 9	-0.0224	-0.0199	-0.0174	1.93E-53
K = 6	K = 10	0.0007	0.0032	0.0057	1.27E-02
K = 7	K = 8	0.0756	0.0781	0.0806	0.00E+00
K = 7	K = 9	0.0388	0.0413	0.0438	7.81E-205
K = 7	K = 10	0.0619	0.0644	0.0669	0.00E+00
K = 8	K = 9	-0.0393	-0.0368	-0.0343	1.40E-166
K = 8	K = 10	-0.0162	-0.0137	-0.0112	1.32E-26
K = 9	K = 10	0.0206	0.0231	0.0256	1.55E-70

Lampiran 29. Hasil Least Significant Difference pada Segmen 4 untuk metode K-means

A	B	Lower Bound	$\bar{x}_{AB} - \bar{x}_{AB}$	Upper Bound	pValue
K = 3	K = 4	0.0757	0.0778	0.0799	0.00E+00
K = 3	K = 5	0.1072	0.1094	0.1115	0.00E+00
K = 3	K = 6	0.1155	0.1176	0.1197	0
K = 3	K = 7	0.1256	0.1277	0.1298	0.00E+00
K = 3	K = 8	0.1343	0.1364	0.1385	0
K = 3	K = 9	0.1363	0.1384	0.1405	0
K = 3	K = 10	0.1340	0.1361	0.1383	0.00E+00

K = 4	K = 5	0.0294	0.0315	0.0337	1.54E-168
K = 4	K = 6	0.0376	0.0398	0.0419	1.25E-253
K = 4	K = 7	0.0477	0.0499	0.0520	0.00E+00
K = 4	K = 8	0.0565	0.0586	0.0607	0.00E+00
K = 4	K = 9	0.0585	0.0606	0.0627	0.00E+00
K = 4	K = 10	0.0562	0.0583	0.0604	0.00E+00
K = 5	K = 6	0.0061	0.0082	0.0104	4.10E-14
K = 5	K = 7	0.0162	0.0183	0.0205	6.83E-62
K = 5	K = 8	0.0249	0.0271	0.0292	1.17E-127
K = 5	K = 9	0.0269	0.0291	0.0312	2.34E-145
K = 5	K = 10	0.0246	0.0268	0.0289	3.41E-125
K = 6	K = 7	0.0080	0.0101	0.0122	2.10E-20
K = 6	K = 8	0.0167	0.0188	0.0210	4.64E-65
K = 6	K = 9	0.0187	0.0208	0.0230	1.49E-78
K = 6	K = 10	0.0164	0.0185	0.0207	3.22E-63
K = 7	K = 8	0.0066	0.0087	0.0109	1.16E-15
K = 7	K = 9	0.0086	0.0107	0.0128	9.36E-23
K = 7	K = 10	0.0063	0.0084	0.0106	9.35E-15
K = 8	K = 9	-0.0001	0.0020	0.0041	6.62E-02
K = 8	K = 10	-0.0024	-0.0003	0.0018	7.92E-01
K = 9	K = 10	-0.0044	-0.0023	-0.0002	3.57E-02

Lampiran 30. Hasil Least Significant Difference pada Segmen 4 untuk metode K-means - ALO

A	B	Lower Bound	$\bar{q}_{AB} - \bar{q}_{AB}$	Upper Bound	pValue
K = 3	K = 4	0.0272	0.0284	0.0297	0
K = 3	K = 5	0.0301	0.0313	0.0326	0
K = 3	K = 6	0.0299	0.0312	0.0324	0
K = 3	K = 7	0.0999	0.1011	0.1023	0
K = 3	K = 8	0.0934	0.0946	0.0958	0
K = 3	K = 9	0.0817	0.0829	0.0842	0
K = 3	K = 10	0.0796	0.0808	0.0821	0
K = 4	K = 5	0.0017	0.0029	0.0041	4.5713e-06
K = 4	K = 6	0.0015	0.0027	0.0040	1.4686e-05

K = 4	K = 7	0.0714	0.0727	0.0739	0
K = 4	K = 8	0.0649	0.0662	0.0674	0
K = 4	K = 9	0.0533	0.0545	0.0557	0
K = 4	K = 10	0.0512	0.0524	0.0537	0
K = 5	K = 6	-0.0014	-0.0002	0.0011	0.8016
K = 5	K = 7	0.0685	0.0698	0.0710	0
K = 5	K = 8	0.0620	0.0633	0.0645	0
K = 5	K = 9	0.0504	0.0516	0.0528	0
K = 5	K = 10	0.0483	0.0495	0.0508	0
K = 6	K = 7	0.0687	0.0699	0.0712	0
K = 6	K = 8	0.0622	0.0634	0.0647	0
K = 6	K = 9	0.0505	0.0518	0.0530	0
K = 6	K = 10	0.0484	0.0497	0.0509	0
K = 7	K = 8	-0.0077	-0.0065	-0.0053	1.5660e-24
K = 7	K = 9	-0.0194	-0.0182	-0.0169	9.4089e-166
K = 7	K = 10	-0.0215	-0.0203	-0.0190	3.9733e-201
K = 8	K = 9	-0.0129	-0.0117	-0.0104	2.3764e-73
K = 8	K = 10	-0.0150	-0.0138	-0.0125	1.4417e-99
K = 9	K = 10	-0.0033	-0.0021	-0.0008	0.001

Lampiran 31. Hasil Least Significant Difference pada Segmen 5 untuk metode K-means

A	B	Lower Bound	$\bar{x}_{AB} - \bar{x}_{AB}$	Upper Bound	pValue
K = 3	K = 4	0.0912	0.0926	0.0940	0
K = 3	K = 5	0.1023	0.1037	0.1051	0
K = 3	K = 6	0.1307	0.1321	0.1335	0
K = 3	K = 7	0.1093	0.1107	0.1121	0
K = 3	K = 8	0.1202	0.1216	0.1230	0
K = 3	K = 9	0.1133	0.1147	0.1161	0
K = 3	K = 10	0.1203	0.1217	0.1231	0
K = 4	K = 5	0.0096	0.0111	0.0125	4.57E-06
K = 4	K = 6	0.0381	0.0395	0.0409	1.47E-05
K = 4	K = 7	0.0167	0.0181	0.0195	0
K = 4	K = 8	0.0276	0.0290	0.0304	0

K = 4	K = 9	0.0206	0.0221	0.0235	0
K = 4	K = 10	0.0277	0.0291	0.0305	0
K = 5	K = 6	0.0270	0.0284	0.0298	0.801695
K = 5	K = 7	0.0056	0.0070	0.0084	0
K = 5	K = 8	0.0165	0.0179	0.0193	0
K = 5	K = 9	0.0096	0.0110	0.0124	0
K = 5	K = 10	0.0166	0.0180	0.0194	0
K = 6	K = 7	-0.0228	-0.0214	-0.0200	0
K = 6	K = 8	-0.0119	-0.0105	-0.0091	0
K = 6	K = 9	-0.0189	-0.0174	-0.0160	0
K = 6	K = 10	-0.0118	-0.0104	-0.0090	0
K = 7	K = 8	0.0095	0.0109	0.0123	1.57E-24
K = 7	K = 9	0.0026	0.0040	0.0054	9.41E-166
K = 7	K = 10	0.0096	0.0110	0.0124	3.97E-201
K = 8	K = 9	-0.0083	-0.0069	-0.0055	2.38E-73
K = 8	K = 10	-0.0013	0.0001	0.0015	1.44E-99
K = 9	K = 10	0.0056	0.0070	0.0084	0.000992

Lampiran 32. Hasil *Least Significant Difference* pada Segmen 5 untuk metode *K-means - ALO*

A	B	Lower Bound	$\bar{x}_{AB} - \bar{x}_{AB}$	Upper Bound	pValue
K = 3	K = 4	0.0858	0.0873	0.0889	0
K = 3	K = 5	0.0945	0.0961	0.0977	0
K = 3	K = 6	0.0891	0.0906	0.0922	0
K = 3	K = 7	0.1020	0.1035	0.1051	0
K = 3	K = 8	0.1118	0.1134	0.1150	0
K = 3	K = 9	0.1086	0.1101	0.1117	0
K = 3	K = 10	0.1091	0.1107	0.1123	0
K = 4	K = 5	0.0072	0.0087	0.0103	2.32E-27
K = 4	K = 6	0.0017	0.0033	0.0049	3.80E-05
K = 4	K = 7	0.0146	0.0162	0.0177	1.23E-86
K = 4	K = 8	0.0245	0.0261	0.0276	1.54E-206
K = 4	K = 9	0.0212	0.0228	0.0243	1.53E-162
K = 4	K = 10	0.0218	0.0233	0.0249	6.31E-170

K = 5	K = 6	-0.0070	-0.0054	-0.0039	1.26E-11
K = 5	K = 7	0.0059	0.0074	0.0090	2.18E-20
K = 5	K = 8	0.0158	0.0173	0.0189	2.21E-98
K = 5	K = 9	0.0125	0.0140	0.0156	1.51E-66
K = 5	K = 10	0.0130	0.0146	0.0162	1.09E-71
K = 6	K = 7	0.0113	0.0129	0.0144	1.57E-56
K = 6	K = 8	0.0212	0.0228	0.0243	3.26E-162
K = 6	K = 9	0.0179	0.0195	0.0210	3.03E-122
K = 6	K = 10	0.0185	0.0201	0.0216	7.02E-129
K = 7	K = 8	0.0083	0.0099	0.0114	2.12E-34
K = 7	K = 9	0.0050	0.0066	0.0082	2.03E-16
K = 7	K = 10	0.0056	0.0072	0.0087	4.57E-19
K = 8	K = 9	-0.0048	-0.0033	-0.0017	4.36E-05
K = 8	K = 10	-0.0043	-0.0027	-0.0011	7.32E-04
K = 9	K = 10	-0.0010	0.0006	0.0021	0.47669

Lampiran 33. Hasil Least Significant Difference pada Segmen 6 untuk metode K-means - ALO

A	B	Lower Bound	$\bar{q}_{AB} - \bar{q}_{AB}$	Upper Bound	pValue
K = 3	K = 4	0.0668	0.0689	0.0709	0.00E+00
K = 3	K = 5	-0.0164	-0.0144	-0.0123	9.33E-42
K = 3	K = 6	0.0892	0.0912	0.0933	0.00E+00
K = 3	K = 7	-0.0037	-0.0017	0.0004	1.15E-01
K = 3	K = 8	0.1072	0.1093	0.1113	0.00E+00
K = 3	K = 9	0.1115	0.1135	0.1156	0.00E+00
K = 3	K = 10	0.1118	0.1138	0.1159	0.00E+00
K = 4	K = 5	-0.0853	-0.0832	-0.0812	0.00E+00
K = 4	K = 6	0.0203	0.0224	0.0244	1.42E-95
K = 4	K = 7	-0.0726	-0.0705	-0.0685	0.00E+00
K = 4	K = 8	0.0383	0.0404	0.0425	4.40E-276
K = 4	K = 9	0.0426	0.0447	0.0467	0.00E+00
K = 4	K = 10	0.0429	0.0450	0.0470	0.00E+00

K = 5	K = 6	0.1036	0.1056	0.1077	0.00E+00
K = 5	K = 7	0.0107	0.0127	0.0148	3.22E-33
K = 5	K = 8	0.1216	0.1236	0.1257	0.00E+00
K = 5	K = 9	0.1258	0.1279	0.1300	0.00E+00
K = 5	K = 10	0.1262	0.1282	0.1303	0.00E+00
K = 6	K = 7	-0.0950	-0.0929	-0.0908	0.00E+00
K = 6	K = 8	0.0160	0.0180	0.0201	6.43E-64
K = 6	K = 9	0.0202	0.0223	0.0243	5.75E-95
K = 6	K = 10	0.0205	0.0226	0.0247	2.18E-97
K = 7	K = 8	0.1089	0.1109	0.1130	0.00E+00
K = 7	K = 9	0.1131	0.1152	0.1172	0.00E+00
K = 7	K = 10	0.1134	0.1155	0.1176	0.00E+00
K = 8	K = 9	0.0022	0.0043	0.0063	5.06E-05
K = 8	K = 10	0.0025	0.0046	0.0066	1.41E-05
K = 9	K = 10	-0.0018	0.0003	0.0024	0.771827

Lampiran 34. Hasil Least Significant Difference pada Segmen 7 untuk metode K-means - ALO

A	B	Lower Bound	$\bar{x}_{AB} - \bar{x}_A$	Upper Bound	pValue
K = 3	K = 4	-0.0030	-0.0007	0.0015	5.18E-01
K = 3	K = 5	0.0171	0.0194	0.0217	3.80E-61
K = 3	K = 6	0.0477	0.0499	0.0522	0.00E+00
K = 3	K = 7	0.0497	0.0519	0.0542	0.00E+00
K = 3	K = 8	0.1117	0.1139	0.1162	0.00E+00
K = 3	K = 9	0.0523	0.0545	0.0568	0.00E+00
K = 3	K = 10	0.0340	0.0363	0.0386	3.27E-193
K = 4	K = 5	0.0179	0.0202	0.0224	1.22E-65
K = 4	K = 6	0.0484	0.0507	0.0529	0.00E+00
K = 4	K = 7	0.0504	0.0527	0.0549	0.00E+00
K = 4	K = 8	0.1124	0.1147	0.1170	0.00E+00
K = 4	K = 9	0.0530	0.0553	0.0576	0.00E+00
K = 4	K = 10	0.0348	0.0370	0.0393	2.46E-200

K = 5	K = 6	0.0282	0.0305	0.0328	2.12E-141
K = 5	K = 7	0.0302	0.0325	0.0348	1.19E-158
K = 5	K = 8	0.0922	0.0945	0.0968	0.00E+00
K = 5	K = 9	0.0329	0.0351	0.0374	2.57E-182
K = 5	K = 10	0.0146	0.0169	0.0191	6.01E-47
K = 6	K = 7	-0.0003	0.0020	0.0043	8.41E-02
K = 6	K = 8	0.0617	0.0640	0.0663	0.00E+00
K = 6	K = 9	0.0023	0.0046	0.0069	6.76E-05
K = 6	K = 10	-0.0159	-0.0136	-0.0114	1.47E-31
K = 7	K = 8	0.0597	0.0620	0.0643	0.00E+00
K = 7	K = 9	0.0003	0.0026	0.0049	2.38E-02
K = 7	K = 10	-0.0179	-0.0156	-0.0134	9.64E-41
K = 8	K = 9	-0.0617	-0.0594	-0.0571	0.00E+00
K = 8	K = 10	-0.0799	-0.0777	-0.0754	0.00E+00
K = 9	K = 10	-0.0205	-0.0183	-0.0160	1.88E-54

BIOGRAFI PENULIS



Reza Prasetya Prayogo lahir di Mojokerto, Jawa Timur pada tanggal 09 Februari 1990. Penulis menempuh pendidikan dasar pada tahun 1998 di MI AT TAQWA Bondowoso, pendidikan menengah pertama pada tahun 2004 di SMPN 1 Bondowoso dan menengah atas pada tahun 2006 di SMAN 2 Bondowoso di kota Bondowoso, Jawa Timur. Selanjutnya meneruskan pendidikan jenjang sarjana pada tahun 2008 di Universitas Muhammadiyah Malang di Kota Malang, Jawa Timur dengan jurusan Teknik Informatika. Setelah menyelesaikan pendidikan sarjana penulis sempat bekerja selama 3 tahun sebagai karyawan dan melanjutkan Pendidikan Magister (S2) di Institut Teknologi Sepuluh Nopember jurusan Teknik Informatika pada tahun 2016. Untuk menghubungi penulis silahkan menghubungi alamat email berikut reza.prasetya.prayogo@gmail.com.