TESIS – EE185401

CLUSTERING DISTRIBUSI SPASIAL DAN TEMPORAL KAPAL PERIKANAN BERDASARKAN DATA *VESSEL MONITORING SYSTEM (VMS)* DENGAN ALGORITMA K-MEANS (STUDI KASUS DI WPPNRI-711)

SUNARMO
NRP 07111850067015

DOSEN PEMBIMBING
Dr. Ir. Achmad Affandi, DEA
Dr. Surya Sumpeno, ST., M.Sc.

PROGRAM MAGISTER
BIDANG KEAHLIAN TELEMATIKA - PETIK
DEPARTEMEN TEKNIK ELEKTRO
FAKULTAS TEKNOLOGI ELEKTRO DAN INFORMATIKA CERDAS
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA
2020

TESIS – EE185401

CLUSTERING DISTRIBUSI SPASIAL DAN TEMPORAL KAPAL PERIKANAN BERDASARKAN DATA VESSEL MONITORING SYSTEM (VMS) DENGAN ALGORITMA K-MEANS

(STUDI KASUS DI WPPNRI-711)

SUNARMO
NRP 07111850067015

DOSEN PEMBIMBING
Dr. Ir. Achmad Affandi, DEA
Dr. Surya Sumpeno, ST., M.Sc.

PROGRAM MAGISTER
BIDANG KEAHLIAN TELAMATIKA - PETIK
DEPARTEMEN TEKNIK ELEKTRO
FAKULTAS TEKNOLOGI ELEKTRO DAN INFORMATIKA CERDAS
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA
2020

LEMBAR PENGESAHAN TESIS

Tesis disusun untuk memenuhi salah satu syarat memperoleh gelar
Magister Teknik (MT)
di
Institut Teknologi Sepuluh Nopember
Oleh:

SUNARMO
NRP 07111850067015

Tanggal Ujian : 02 Juli 2020
Periode Wisuda : September 2020

Disetujui oleh:
Pembimbing:

1. Dr. Ir. Achmad Affandi, DEA
   NIP: 196510141990021001 .....

2. Dr. Surya Sumpeno, ST., M.Sc.
   NIP: 196906131997021003 .....

Penguji:

1. Dr. Adhi Dharma Wibawa, ST., MT.
   NIP: 197605052008121003 .....

2. Eko Setijadi, ST.,MT.,Ph.D.
   NIP: 197210012003121002 .....

3. Dr. Diah Puspito Wulandari, S.T., M.Sc.
   NIP: 198012192005012001 .....

Kepala Departemen Teknik Elektro
Fakultas Teknologi Elektro dan Informatika Cerdas

Dedet Candra Riawan, ST., M.Eng., Ph.D.
NIP: 197311192000031001

Halaman ini sengaja dikosongkan

PERNYATAAN KEASLIAN TESIS

Dengan ini saya menyatakan bahwa isi keseluruhan Tesis saya dengan judul "**CLUSTERING DISTRIBUSI SPASIAL DAN TEMPORAL KAPAL PERIKANAN BERDASARKAN DATA VESSEL MONITORING SYSTEM (VMS) DENGAN ALGORITMA K-MEANS (STUDI KASUS DI WPPNRI-711)**" 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 pustaka. Apabila ternyata pernyataan ini tidak benar, saya bersedia menerima sanksi sesuai peraturan yang berlaku.

Surabaya, Juli 2020

Sunarmo

NRP. 07111850067015

Halaman ini sengaja dikosongkan

CLUSTERING DISTRIBUSI SPASIAL DAN TEMPORAL KAPAL PERIKANAN BERDASARKAN DATA *VESSEL MONITORING SYSTEM (VMS)* DENGAN ALGORITMA K-MEANS (STUDI KASUS DI WPPNRI-711)

Nama mahasiswa : SUNARMO
NRP : 07111850067015
Pembimbing : 1. Dr. Ir. Achmad Affandi, DEA
2. Dr. Surya Sumpeno, ST., M.Sc.

ABSTRAK

Manajemen pengelolaan sumber daya laut yang berkelanjutan merupakan permasalah nasional dan juga global, pengelolaan perikanan mempunyai masalah yang kompleks sehingga diperlukan lebih banyak penelitian dengan berbagai pendekatan yang lebih komprehensif. Pemerintah Indonesia melalui Kementerian Kelautan dan Perikanan membuat Sistem Pemantauan Kapal Perikanan (SPKP) /*Vessel Monitoring System (VMS)*. Data VMS yang berisi tentang posisi, pergerakan dan aktivitas armada kapal penangkap ikan dimanfaatkan dalam penelitian menggunakan teknik data mining dengan metode unsupervised dari machine learning. Penelitian ini terdiri tiga langkah: i) mencari jumlah cluster uptimum dengan metode elbow, ii) melakukan clustering dengan alogaritma K-Means dengan nilai k optimum yang telah ditetapkan, iii) menganalisis distribusi sebaran data VMS secara spasial dan temporal. Secara umum diperoleh hasil jumlah cluster optimum adalah 7 dengan hasil kekompakan anggota cluster persentasenya 90,7%, dan validasi clustering dengan uji *Silhoutte Coefficient* diperoleh nilai 0.6677. Secara spasial distribusi sebaran data VMS di WPPNRI-711 tidak merata dan secara temporal sangat fluktuatif. Hasil dari penelitian ini dapat dipergunakan untuk mengevaluasi dan mengetahui intensitas serta lokasi penangkapan ikan guna mencegah terjadinya overfishing.

Kata kunci: *Data Mining, Unsupervised Learning,* Alogaritma K-Means, Pola Spasial dan Temporal, *Vessel Monitoring System.*

Halaman ini sengaja dikosongkan

CLUSTERING SPATIAL TEMPORAL DISTRIBUTION OF FISHING VESSEL BASED ON VESSEL MONITORING SYSTEM (VMS) DATA USING K-MEANS ALGORITHM (CASE STUDY IN WPPNRI-711)

By : SUNARMO
Student Identity Number : 07111850067015
Supervisor(s) : 1. Dr. Ir. Achmad Affandi, DEA
2. Dr. Surya Sumpeno, ST., M.Sc.

ABSTRACT

Management of sustainable marine resources is a national and global problem, and fisheries management has a complex problem, more research is need with a more comprehensive approach. The Government of Indonesia, through the Ministry of Marine Affairs and Fisheries, has made the Vessel Monitoring System (VMS). VMS data contains the position, movement, and activity of the fishing vessels utilized in this research. Data mining techniques and machine learning are using, and this study consists of three steps: i) Finding the number of optimum clusters by the Elbow Method, ii) Conducting clustering with the K-Means algorithm with the optimum k-value that has set, iii) Analyze the distribution of VMS data spatially and temporally. Overall, the optimum number of clusters obtained is 7 with the results of the compactness of the cluster members the percentage is 90.7%, and clustering validation using the Silhoutte Coefficient test, obtained a value of 0.6677. Spatially the distribution of VMS data in the Fisheries Management Area WPPNRI-711 is uneven and temporally very volatile. The results of this study can provide information about the intensity and location of fishing activity and for preventing overfishing.

Keywords: Data Mining, Unsupervised Learning, K-Means Algorithm, Spasial and Temporal Pattern, Vessel Monitoring System.

Halaman ini sengaja dikosongkan

KATA PENGANTAR

Alhamdulillah segala puji bagi Allah SWT, sehingga penulis bisa menyelesaikan tesis dengan judul **"CLUSTERING DISTRIBUSI SPASIAL DAN TEMPORAL KAPAL PERIKANAN BERDASARKAN DATA VESSEL MONITORING SYSTEM (VMS) DENGAN ALGORITMA K-MEANS (STUDI KASUS DI WPPNRI-711)"**. Tesis ini disusun untuk memenuhi salah satu syarat kelulusan pada Program Magister Bidang Keahlian Telematika, Departemen Teknik Elektro, Fakultas Teknologi Elektro dan Sistem Cerdas, Institut Teknologi Sepuluh Nopember, Surabaya.

Penyusunan dan penyelesaian tesis ini tidak terlepas dari bantuan berbagai pihak. Pada kesempatan ini penulis mengucapkan terima kasih dan penghormatan kepada:

1. Bapak Dr. Ir. Achmad Affandi, DEA dan Bapak Dr. Surya Sumpeno, ST., M.Sc. selaku dosen pembimbing, atas waktu, motivasi dan bimbingannya dalam penyusunan tesis.
2. Para Dosen Program Magister, Bidang Keahlian Telematika, Departemen Teknik Elektro, Fakultas Teknologi Elektro dan Sistem Cerdas, Institut Teknologi Sepuluh Nopember Surabaya atas ilmu, inspirasi, motivasi, dan pengalaman yang diberikan selama penulis menempuh studi.
3. Bapak Dr. Adhi Dharma Wibawa, S.T., M.T. selaku koordinator bidang keahlian telematika atas arahan dan motivasinya sejak awal perkuliahan sampai akhir penulisan tesis.
4. Bapak Dedet Candra Riawan, ST., M.Eng., Ph.D selaku Kepala Departemen Teknik Elektro Fakultas Teknologi Elektro dan Sistem Cerdas, Institut Teknologi Sepuluh Nopember.
5. Bapak Dr. I Ketut Eddy Purnama,ST.,MT. selaku Dekan Fakultas Teknologi Elektro dan Sistem Cerdas, Institut Teknologi Sepuluh Nopember.
6. Direktorat Pemantauan dan Operasi Armada, Direktorat Jenderal Pengawasan Sumber Daya Kelautan dan Perikanan, Kementerian Kelautan

dan Perikanan Republik Indonesia, yang telah memberikan support data SPKP/ VMS untuk penelitian ini.

7. Ayahanda Samanto, Ibunda Kasinah, dan Mertuaku Bapak Daryanto dan Ibu Menik atas semua dukungan dan doa yang tiada terputus selama penulis belajar S-2 di ITS.
8. Istriku tercinta, Gilang Hardini Putri, S.Sos dan anakku tersayang, Naura Auni Syakira. Terimakasih atas semua dukungan, pengertian, motivasi, pengorbanan, ketabahan serta kesabaran yang luar biasa selama mendampingi penulis menyelesaikan studi S-2 di ITS.
9. Kepala Badan Litbang SDM Kementerian Komunikasi dan Informatika yang telah memberikan beasiswa untuk mengikuti pendidikan Program Magister Bidang Keahlian Telematika, Departemen Teknik Elektro, Fakultas Teknologi Elektro dan Sistem Cerdas, Institut Teknologi Sepuluh Nopember Surabaya.
10. Keluarga Besar Lembaga Pemasyarakatan Kelas IIA Tanjungpinang atas dukungan selama penyelesaian studi.
11. Rekan-rekan S2 Telematika angkatan 2018 maupun beda angkatan yang selalu memberikan keceriaan, dukungan, motivasi maupun bantuan lainnya.
12. Seluruh Dosen dan seluruh staf sekretariat Program Magister Departemen Teknik Elektro, Fakultas Teknologi Elektro dan Sistem Cerdas, Institut Teknologi Sepuluh Nopember Surabaya atas dukungan dan kerjasamanya selama penyelesaian studi.
13. Semua pihak yang telah membantu dalam proses penyelesaian studi dan penyelesaian tesis yang tidak dapat penulis sebutkan satu persatu.

Penulis menyadari bahwa tesis ini masih sangat jauh dari sempurna. Oleh karena itu, penulis sangat mengharapkan kritik dan saran yang bersifat membangun agar tesis ini menjadi lebih baik. Akhir kata, penulis berharap tesis ini memberikan manfaat untuk pengembangan ilmu pengetahuan dan dapat menjadi amal kebaikan.

Surabaya, Juni 2020
Sunarmo

DAFTAR ISI

LEMBAR PENGESAHAN TESIS ..... **iii**
PERNYATAAN KEASLIAN TESIS ..... **v**
ABSTRAK ..... **vii**
ABSTRACT ..... **ix**
KATA PENGANTAR ..... **xi**
DAFTAR ISI ..... **xiii**
DAFTAR GAMBAR ..... **xvii**
DAFTAR TABEL ..... **xix**
BAB I PENDAHULUAN **1**
1.1 Latar Belakang 1
1.2 Rumusan Masalah 5
1.3 Tujuan 6
1.4 Batasan Masalah 6
1.5 Kontribusi 6
BAB II KAJIAN PUSTAKA **7**
2.1 Analisis Spasial dan Temporal 7
2.2 Definisi Kapal Perikanan 9
2.3 *Vessel Monitoring System (VMS)* 10
2.4 Wilayah Pengelolaan Perikanan Negara Republik Indonesia (WPPNRI) - 711 13
2.5 *Data Mining* 14
2.6 *Machine Learning* 16
2.7 *Clustering* 17
2.8 *Algoritma K-Means* 18
2.9 Metode *Elbow* 20
2.10 Pengujian *Silhouette Coefficient* 22
BAB III METODOLOGI PENELITIAN **25**
3.1 Studi Literatur 26
3.2 Mengumpulkan Data Set 26
3.3 Akusisi Data 26

3.4 Pembersihan Data (*Data Cleansing*). 27
3.5 Seleksi Fitur 27
3.6 Penentuan Nilai k dengan Metode *Elbow* 27
3.7 Implementasi *Clustering* 28
3.7.1 Algoritma *K-Means* 29
3.8 Validasi Hasil *Clustering* Dengan *Silhouette Coefficient* 31
3.9 Analisis 31
3.10 Kesimpulan 31
BAB IV HASIL PENELITIAN DAN PEMBAHASAN 33
4.1 Tahap *Preprocessing Data* 33
4.1.1 Akusisi Data 33
4.1.2 Pembersihan Data (*Data Cleansing*) 34
4.1.3 Konversi Data 36
4.1.4 Seleksi Fitur 38
4.2 Implementasi *Clustering* 40
4.2.1 Mencari Jumlah *Cluster* dengan *Elbow Method* 40
4.2.1.1 Berdasarkan Data Tahun 2016 41
4.2.1.2 Berdasarkan Data Tahun 2017 42
4.2.1.3 Berdasarkan Data Tahun 2018 44
4.2.1.4 Berdasarkan Data Tahun 2016 sd 2018 45
4.2.2 *Clustering* dengan *Alogaritma K-Means* 47
4.2.2.1 Visualisasi 48
4.2.3 Uji *Validitas Clustering* dengan *Silhouette Coefficient.* 49
4.3 Analisa Hasil Clustering 50
4.3.1 Analisis Cluster 1 59
4.3.2 Analisis Cluster 2 63
4.3.3 Analisis Cluster 3 67
4.3.4 Analisis Cluster 4 71
4.3.5 Analisis Cluster 5 75
4.3.6 Analisis Cluster 6 79
4.3.7 Analisis Cluster 7 83
BAB V KESIMPULAN DAN SARAN 87

5.1 Kesimpulan 87
5.2 Saran 89
DAFTAR PUSTAKA 91
BIOGRAFI PENULIS 97

Halaman ini sengaja dikosongkan

DAFTAR GAMBAR

Gambar 2. 1 Contoh Pola *Clustered, Dispersed* dan *Random* [27] 8
Gambar 2. 2 Cara Kerja *Vessel Monitoring System (VMS)* 11
Gambar 2. 3 Peta WPP-NRI 711..... 14
Gambar 2. 4 Proses *Data Mining* [18] 15
Gambar 2. 5 *Elbow Method* [53]..... 21
Gambar 3. 1 Diagram Penelitian..... 25
Gambar 3. 2 *Flowchart Algoritma K-Means* [53]..... 30
Gambar 4. 1 Tampilan Data Awal 33
Gambar 4. 2 Fitur Data di Tahap Awal..... 34
Gambar 4. 3 Data Setelah Dibersihkan dari Duplikasi 35
Gambar 4. 4 *Missing Data* 35
Gambar 4. 5 Posisi Sebaran Data VMS Berdasarkan Titik Koordinat Longitude dan Latitude..... 37
Gambar 4. 6 Frequensi Distribusi Data Tahun 2016 s/d 2018..... 38
Gambar 4. 7 Fitur-fitur yang Dipergunakan Pada Proses Selanjutnya 39
Gambar 4. 8 *Summary Data*..... 39
Gambar 4. 9 Data Longitude dan Latitude..... 40
Gambar 4. 10 Grafik *Elbow Method* Berdasarkan Data Tahun 2016 41
Gambar 4. 11 Grafik *Elbow Method* Berdasarkan Data Tahun 2017 43
Gambar 4. 12 Grafik *Elbow Method* Berdasarkan Data Tahun 2018 44
Gambar 4. 13 Grafik *Elbow Method* Berdasarkan Data Tahun 2016 s/d 2018 45
Gambar 4. 14 Hasil *Clustering* dengan *K-Means* 47
Gambar 4. 15 Plot hasil *K-Means Clustering* 48
Gambar 4. 16 Visualisasi Hasil *K-Means Clustering* 49
Gambar 4. 17 Hasil Uji *Silhouette Coefficient*..... 50
Gambar 4. 18 Visualisasi Hasil K-Means *Clustering* dengan Split Map..... 51
Gambar 4. 19 Distribusi Data per Cluster Periode Tahunan..... 52
Gambar 4. 20 Distribusi Data Bulanan Pada Tahun 2016, 2017 dan 2018 54
Gambar 4. 21 Distribusi Data per Bulan..... 55
Gambar 4. 22 Distibusi Data Tahun 2016 per Cluster Periode Bulanan 56

Gambar 4. 23 Distibusi Data Tahun 2017 per Cluster Periode Bulanan 56
Gambar 4. 24 Distibusi Data Tahun 2018 per Cluster Periode Bulanan 57
Gambar 4. 25 Distibusi Data Tahun 2016 s/d 2018 per Cluster Periode Bulanan 57
Gambar 4. 26 Visualisasi Cluster 1 59
Gambar 4. 27 Distribusi Data Berdasarkan Nama Alat Tangkap Cluster 1 61
Gambar 4. 28 Grafik Distribusi Data *Cluster* 1 61
Gambar 4. 29 Diagram Batang Aktivitas Kapal Perikanan *Cluster* 1 62
Gambar 4. 30 Visualisasi *Cluster* 2 63
Gambar 4. 31 Distribusi Data Berdasarkan Nama Alat tangkap *Cluster* 2 65
Gambar 4. 32 Grafik Distribusi Data *Cluster* 2 65
Gambar 4. 33 Diagram Batang Aktivitas Kapal Perikanan Cluster 2 66
Gambar 4. 34 Visualisasi *Cluster* 3 67
Gambar 4. 35 Distribusi Data Berdasarkan Nama Alat Tangkap Cluster 3 69
Gambar 4. 36 Grafik Distribusi Data Cluster 3 69
Gambar 4. 37 Diagram Batang Aktivitas Kapal Perikanan *Cluster* 3 70
Gambar 4. 38 Visualisasi *Cluster* 4 71
Gambar 4. 39 Distribusi Data Berdasarkan Nama Alat Tangkap *Cluster* 4 73
Gambar 4. 40 Grafik Distribusi Data *Cluster* 4 73
Gambar 4. 41 Diagram Batang Aktivitas Kapal Perikanan *Cluster* 4 74
Gambar 4. 42 Visualisasi *Cluster* 5 75
Gambar 4. 43 Distribusi Data Berdasarkan Nama Alat Tangkap *Cluster* 5 77
Gambar 4. 44 Grafik Distribusi Data Cluster 5 77
Gambar 4. 45 Diagram Batang Aktivitas Kapal Perikanan *Cluster* 5 78
Gambar 4. 46 Visualisasi Cluster 6 79
Gambar 4. 47 Distribusi Data berdasarkan Nama Alat Tangkap *Cluster* 6 81
Gambar 4. 48 Grafik Distribusi Data *Cluster* 6 81
Gambar 4. 49 Diagram Batang Aktivitas Kapal Perikanan *Cluster* 6 82
Gambar 4. 50 Visualisasi Cluster 7 83
Gambar 4. 51 Distribusi Data Berdasarkan Nama Alat Tangkap *Cluster* 7 85
Gambar 4. 52 Grafik Distribusi Data Cluster 7 85
Gambar 4. 53 Diagram Batang Aktivitas Kapal Perikanan *Cluster* 7 86

DAFTAR TABEL

Tabel 2. 1 Kategori Nilai Silhouette Coeffisien dan Interpretasinya 23
Tabel 3. 1 Perbandingan waktu komputasi *Elbow Method, Gab Statistic, Sihouette Coefficient* dan *Canopy* [58]. 28
Tabel 4. 1 Jumlah Data Duplikasi 35
Tabel 4. 2 Hasil *Distance Cluster* Berdasarkan Data Tahun 2016 42
Tabel 4. 3 Hasil *Distance Cluster* Berdasarkan Data Tahun 2017 43
Tabel 4. 4 Hasil *Distance Cluster* Berdasarkan Data Tahun 2018 44
Tabel 4. 5 Hasil *Distance Cluster* Berdasarkan Data Tahun 2016 s/d 2018..... 46
Tabel 4. 6 Data Hasil K-Means *Clustering* 52
Tabel 4. 7 Distribusi Data per Cluster Periode Tahunan 53
Tabel 4. 8 Distribusi Data Bulanan Pada Tahun 2016, 2017 dan 2018 54
Tabel 4. 9 Analisis Cluster 1 60
Tabel 4. 10 Analisis *Cluster* 2 64
Tabel 4. 11 Analisa *Cluster* 3 68
Tabel 4. 12 Analisis *Cluster* 4 72
Tabel 4. 13 Analisis *Cluster* 5 76
Tabel 4. 14 Analisis *Cluster* 6 80
Tabel 4. 15 Analisis Cluster 7 84

Halaman ini sengaja dikosongkan

BAB I
PENDAHULUAN

1.1 Latar Belakang

Indonesia merupakan Negara Kepulauan terbesar didunia yang memiliki sumber daya laut yang sangat kaya dengan aneka ragam hayatinya. Indonesia dikenal juga sebagai Negara Maritim karena memiliki perairan yang lebih luas daripada daratan, dengan luas wilayah 7,81 juta km² [1], 2/3 nya merupakan lautan dan memiliki 17.504 pulau [1]. Indonesia menempati urutan kedua negara dengan garis pantai terpanjang setelah Kanada dengan panjang garis pantai 99.093 km [2]. Dengan total luas wilayah perairan Indonesia mencapai 6,32 juta km² yang terdiri luas wilayah kedaulatan 3,37 juta km² dan luas wilayah perairan berdaulat 2,94 juta km² [1] dalam rangka optimalisasi pengelolaan perikanan maka dibagi menjadi 11 Wilayah Pengelolaan Perikanan (*fisheries management areas*) yang dikenal dengan Wilayah Pengelolaan Perikanan Negara Republik Indonesia (WPPNRI) [3], pembagian ini didasarkan pada ekologi, karakteristik wilayah, dan sumber daya ikan yang digunakan sebagai dasar pengelolaan perikanan secara lestari dan berkelanjutan.

Salah satu dari 11 wilayah tersebut adalah Wilayah Pengelolaan Perikanan Negara Republik Indonesia (WPPNRI)-711 yang memiliki potensi jumlah tangkapan sumber daya ikan yang diizinkan mencapai 1.143.341ton pertahun. Ada 5 (lima) kelompok sumber daya ikan di WPPNRI-711 diantaranya adalah ikan demersal sebesar 400,517 ton/ tahun, ikan pelagis kecil sebesar 395,451 ton/ tahun, ikan pelagis besar sebesar 198,994 ton/ tahun, udang penaeid 78,005 ton/ tahun, dan cumi cumi sebesar 35,155 ton/ tahun [4]. Secara administratif WPPNRI-711 masuk kedalam 7 (tujuh) pemerintah Provinsi yang meliputi Provinsi Kepulauan Riau, Provinsi Riau, Provinsi Jambi, Provinsi Sumatera Selatan, Provinsi Kepulauan Bangka Belitung, Provinsi Kalimantan Barat, dan Provinsi Kalimantan Tengah. Sebagian besar wilayahnya berada di Provinsi Kepulauan Riau yang secara geografis terletak pada posisi yang sangat strategis dan berbatasan langsung dengan beberapa negara di Asia Tenggara, seperti Malaysia, Singapura, Vietnam, dan

Kamboja. Didalam WPPNRI-711 juga terdapat kawasan konservasi perairan nasional yang berada di Kabupaten Kepulauan Anambas, Provinsi Kerpulauan Riau, dengan luas 1.262.686 ha [1].

Laut Indonesia berperan penting bagi berbagai kegiatan ekonomi seperti bisnis perikanan, pelayaran, maupun pariwisata. *United Nations Development Programme* (UNDP) juga menyebut perairan Indonesia sebagai habitat bagi 76 persen terumbu karang dan 37 persen ikan karang dunia. Hasil laut berupa ikan menjadi sumber protein penting bagi masyarakat Indonesia, menurut UNDP sebanyak 54 persen kebutuhan protein nasional dipenuhi dari ikan, karena ikan merupakan sumber protein yang paling ekonomis dibandingkan sumber protein lainnya. Potensi perikanan yang sangat besar akan memberikan manfaat yang maksimal secara berkelanjutan bagi negara dan masyarakat Indonesia bila dikelola dengan baik dan bertanggungjawab. Upaya untuk mewujudkan pembangunan kelautan dan perikanan yang menitik beratkan pada kedaulatan (*sovereignty*), keberlanjutan (*sustainability*), dan kesejahteraan (*prosperity*) harus dilakukan melalui proses terencana, terpadu, dan berkesinambungan [5], hasilnya akan ditandai dengan meningkatnya kualitas hidup dan kesejahteraan masyarakat serta terjaganya kelestarian sumber daya ikan dan ekosistemnya.

Secara umum, aktivitas perikanan di Indonesia belum menunjukkan kinerja yang berkelanjutan karena lemahnya sistem pengelolaan perikanan (*fisheries management system*). Hal ini, dapat dilihat dengan belum banyak usaha perikanan di Indonesia yang bisa bertahan lama [6], sebanyak 115 perusahaan pengelolaan ikan nasional gulung tikar akibat tidak mendapatkan pasokan ikan, karena kapal-kapal pencuri ikan membawa hasil tangkapannya langsung ke negara asalnya [2]. Kompleksnya permasalahan sektor perikanan nasional mengakibatkan ketidakteraturan dan menyebabkan aktivitas perikanan nasional menjadi tidak berkelanjutan. Berdasarkan survei BPS periode 2003-2013, jumlah rumah tangga nelayan turun dari 1,6 juta menjadi hanya sekitar 800.000, karena sumberdaya laut tidak bisa lagi menjadi sumber mata pencaharian nelayan tradisional, pendapatan mereka semakin menurun karena tidak dapat bersaing dengan kapal-kapal besar. Sedangkan di Indonesia 81,72% merupakan nelayan kecil yaitu nelayan tanpa perahu dan yang beroperasi dengan perahu motor tempel [7]. Jumlah sumberdaya

manusia yang bergantung pada kegiatan usaha kelautan dan perikanan sangat besar, namun pengetahuan, keterampilan, penguasaan teknologi dan akses terhadap infrastruktur dan informasi belum memadai dan belum merata di seluruh wilayah Indonesia, terutama di wilayah kepulauan [8].

Untuk mengatasi masalah tersebut pemerintah Indonesia melalui Kementerian Kelautan dan Perikanan membuat Sistem Pemantauan Kapal Perikanan (SPKP) / *Vessel Monitoring System (VMS)* yang merupakan salah satu sistem pengawasan kapal perikanan dengan menggunakan peralatan yang telah ditentukan untuk mengetahui pergerakan dan aktifitas kapal perikanan dengan berbasis satelit. Implementasi VMS merupakan bentuk komitmen Indonesia memenuhi ketentuan Internasional, Regional, maupun Nasional untuk melaksanakan ketentuan-ketentuan konservasi dan pengelolaan perikanan yang berkelanjutan. Sejak tahun 2003, Sistem Pemantauan Kapal Perikanan (SPKP) telah diterapkan dengan memasang alat pemancar atau transmitter pada kapal-kapal perikanan berukuran di atas 30 GT. Berdasarkan Pasal 12 ayat (1) Peraturan Menteri Kelautan dan Perikanan Nomor 42/PERMEN-KP/2015 tentang Sistem Pemantauan Kapal Perikanan disebutkan bahwa setiap kapal perikanan berukuran lebih dari 30 GT yang beroperasi di Wilayah Pengelolaan Perikanan Negara Republik Indonesia (WPPNRI) dan di laut lepas wajib memasang transmiter SPKP.

Manfaat SPKP/VMS antara lain meningkatkan efektifitas pengelolaan perikanan, meningkatkan ketaatan kapal perikanan yang melakukan kegiatan penangkapan ikan dan pengangkutan ikan terhadap ketentuan peraturan perundang-undangan, memperoleh data dan informasi tentang kegiatan kapal perikanan dalam rangka pengelolaan perikanan secara bertanggungjawab dan berkelanjutan serta meningkatkan pelaksanaan penegakan hukum dibidang perikanan. Kegiatan ilegal di laut sangat sulit diamati karena terlalu luasnya wilayah yang harus diawasi dan minimnya sarana dan prasarana untuk pengawasannya. Dengan memanfaatkan data yang diperoleh dari rekaman VMS berupa informasi tentang data lokasi kapal, lintasan kapal dan alat penangkap ikan yang digunakan. Teknologi VMS ini dapat memonitor melalui satelit dari luar angkasa posisi kapal penangkap ikan mendekati waktu yang sebenarnya [9].

Data VMS dapat memberikan informasi yang berharga tentang pola spasial dan temporal aktifitas kapal penangkap ikan dengan beberapa jenis sekala [10], data VMS menyediakan data serial waktu dan posisi kapal perikanan yang dapat digunakan untuk memperhitungkan dimensi spasial dan temporal operasi kapal perikanan, sehingga dapat membuka wawasan baru untuk menganalisis masa depan pengelolaan perikanan yang berkelanjutan. Hal ini merupakan salah satu perkembangan yang penting dalam penelitian dibidang perikanan pada dekade terakhir, secara keseluruhan estimasi tawur data VMS yang telah divalidasi dengan aktivitas tawur kapal contoh diperoleh tingkat kesalahan deteksi (*misdetection*) dan penyimpangan posisi estimasi tawur dengan aktivitas tawur kurang dari 10%, sehingga memungkinkan penggunaan data VMS untuk menyimpulkan informasi tentang aktivitas kapal perikanan dan upaya penangkapan. Selain itu, penggunaan data VMS setidaknya memberi keuntungan untuk menetapkan waktu yang dihabiskan di laut [11] dan saat ini semakin banyak sistem untuk mengeksploitasi data VMS yang bermanfaat dalam pengelolaan dan ekologi perikanan [12]. Dengan demikian data VMS dapat dipergunakan sebagai dasar penelitian ini untuk memperoleh gambaran tentang kondisi usaha penangkapan ikan di WPPNRI-711.

Saat ini semakin banyak sistem untuk mengeksploitasi data VMS yang bermanfaat dalam manajemen pengelolaan dan ekologi perikanan [12][13]. Beberapa penelitian dengan menggunakan data VMS diantaranya untuk mengidentifikasi alat tangkap yang digunakan kapal penangkap ikan [14][15]. Selain itu data VMS juga digunakan untuk mengidentifikasi perilaku kapal penangkap ikan [16], dan meneliti hubungan kerjasama antar kapal penangkap ikan [17].

Sedangkan data mining merupakan sebuah proses dan langkah analisis untuk menemukan informasi ataupun pengetahuan didalam sebuah basis data (*knowledge discovery in data bases*) [18]. Beberapa penelitian telah menggunakan metode data mining untuk mencari pola temporal [19][20][21], dan juga penelitian di bidang spasial [22][23][24], dengan melakukan klasterisasi untuk mencari pola spasial [25][26]. Pola spasial adalah sesuatu yang menunjukkan penempatan atau susunan benda-benda di permukaan bumi [27]. Selain itu juga ada penelitian dengan domain distribusi data spasial dan temporal [28], berdasarkan hubungan

topologi antara komponen spasial dan temporal [29] juga penelitian untuk mencari deviasi dan pola hubungan antara sub kelompok data temporal, spasial dan tekstual [30]. Dan masih banyak penelitian lainnya tentang data mining di bidang temporal, spasial dan spasial-temporal [31].

Penelitian tentang pola spatial dan temporal dengan Algoritma K-Means sebelumnya telah dilakukan oleh para peneliti. Seperti penelitian untuk mengetahui pola permintaan layanan transportasi online Didi services di Chengdu, China. Penelitian ini menunjukkan secara temporal ada pola permintaan yang berulang secara harian dan mingguan sehingga dapat diprediksi dengan karakteristik yang berbeda, sedangkan secara spatial masing-masing zona cluster memiliki luas dan jumlah pengemudi yang berbeda juga [32]. Penelitian lain juga mengadaptasi K-Means clustering untuk menemukan pola berbagai jenis badai per jam berdasarkan bentuk dan ukurannya. Hasilnya menunjukkan bentuk dan ukuran centroid cluster mewakili bentuk dan ukuran badai yang termasuk dalam cluster itu dengan tingkat kemiripan mencapai 60% [33].

Pendekatan yang diambil dalam penelitian ini adalah untuk *clustering* distribusi spasial dan temporal aktivitas kapal perikanan berdasarkan data VMS di WPPNRI-711. Informasi yang diperoleh dapat digunakan untuk mendapatkan gambaran aktifitas penangkapan ikan disetiap cluster berdasarkan lokasi juga waktu. Informasi yang akurat tentang pola kegiatan penangkapan ikan secara spasial sangat penting bagi para peneliti, pembuat kebijakan dan nelayan [34], sehingga penelitian ini dapat berkontribusi dalam mangement pengelolaan sumberdaya laut dan perikanan. Penelitian sebelumnya telah menunjukkan bahwa data VMS dapat dimanfaatkan untuk mengetahui lokasi dan jalur penangkapan ikan [35], dan untuk memperkirakan intensitas penangkapan ikan [36], sehingga dapat digunakan untuk mencegah *over fishing* demi menjaga kelestarian sumberdaya perikanan [37].

1.2 Rumusan Masalah

Untuk mengatasi kompleksnya permasalahan sistem pengelolaan perikanan (*fisheries management system*), diperlukan banyak informasi serta kajian dari berbagai sudut pandang dan bidang keilmuan tentang pengelolaan perikanan

yang selama ini berlangsung. Diantaranya adalah informasi tentang sebaran spasial dan temporal aktifitas kapal perikanan yang beroperasi di Wilayah Pengelolaan Perikanan Negara Republik Indonesia 711(WPPNRI-711).

1.3 Tujuan

Tujuan penelitian ini adalah untuk melakukan *clustering* sebaran spasial dan temporal aktivitas kapal perikanan berdasarkan fitur data *Vessel Monitoring System (VMS)* yang beroperasi di Wilayah Pengelolaan Perikanan Negara Republik Indonesia 711 (WPPNRI-711) dengan memanfaatkan alogaritma *K-Means*. Informasi dari hasil penelitian ini dapat dipergunakan untuk mengetahui intensitas dan lokasi penangkapan ikan yang selama ini berlangsung guna mencegah terjadinya *overfishing*.

1.4 Batasan Masalah

Penelitian ini dibatasi pada hal-hal sebagai berikut:

1. Data yang diteliti adalah data *vessel monitoring system* (*VMS)* Tahun 2016 s/d 2018 yang diperoleh dari Direktorat Pemantauan dan Operasi Armada, Direktorat Jenderal Pengawasan Sumber Daya Kelautan dan Perikanan, Kementerian Kelautan dan Perikanan Republik Indonesia.
2. Data VMS kapal perikanan yang beroperasi di Wilayah Pengelolaan Perikanan Negara Republik Indonesia 711 (WPPNRI-711).

1.5 Kontribusi

Informasi yang diperoleh dari penelitian ini dapat digunakan untuk mengevaluasi kegiatan penangkapan ikan yang selama ini berlangsung dan sebagai salah satu dasar dalam pengembangan manajemen pengelolaan perikanan yang berkeadilan dan berkelanjutan di masa yang akan datang.

BAB II
KAJIAN PUSTAKA

Pada bab ini akan dibahas rujukan yang digunakan sebagai kajian pustaka meliputi Pola Spasial dan Temporal, Kapal Perikanan, *Vessel Monitoring System (VMS)*, *Data Mining*, *Machine Learning* dan *K-Means Clustering*.

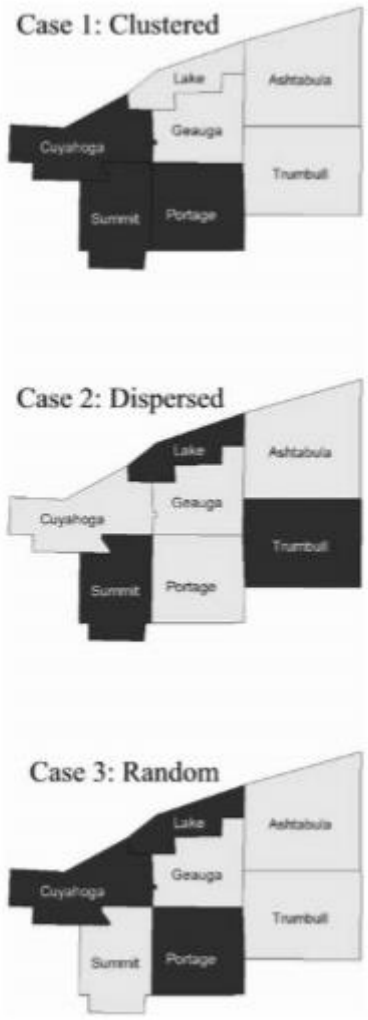
2.1 Analisis Spasial dan Temporal

Data spasial merupakan sebuah data yang berorientasi geografis, memiliki sistem koordinat tertentu sebagai dasar referensinya seperti koordinat Longitude dan Latitude. Sedangkan Pola spasial adalah sesuatu yang menunjukkan penempatan atau susunan benda-benda di permukaan bumi [27]. Pola spasial menjadi suatu konsep statistika, ketika pola tersebut menunjukkan bagaimana objek geografis terdistribusi pada suatu waktu tertentu atau memiliki dimensi temporal. Pola spasial menjelaskan tentang bagaimana fenomena geografis terdistribusi dan bagaimana perbandingannya dengan fenomena lainnya. Dalam hal ini, analisa spasial merupakan alat yang banyak digunakan untuk mendiskripsikan dan menganalisis pola spasial tersebut, yaitu bagaimana objek-objek geografis terjadi dan berubah di suatu lokasi. Selain itu juga dapat membandingkan pola objek disuatu lokasi dengan pola objek yang ditemukan di lokasi lain.

Data spasial bersifat sangat kompleks dan dinamis, untuk lebih memahaminya perlu menguranginya menjadi beberapa model representatif sederhana. Tujuan membangun model adalah untuk mengurangi kompleksitas masalah yang dipelajari ke tingkat yang bisa dikelola sehingga dapat lebih fokus pada masalah yang ada dan solusi untuk menyelesaikannya. Dengan memodelkan distribusi benda atau objek tertentu yang terletak di berbagai tempat di permukaan bumi menjadi lebih sederhana dapat dirumuskan hubungan antar objek. Bentuk distribusi data pada pola spasial, diantaranya [27] :

1. *Clustered*, beberapa area membentuk suatu kelompok dan saling berdekatan.
2. *Dispersed*, setiap area berada secara merata dan berjauhan dengan area-area lainnya.

3. *Random*, beberapa area terletak secara acak di beberapa lokasi, posisi suatu area tidak dipengaruhi oleh posisi area lainnya.



Gambar 2. 1 Contoh Pola *Clustered*, *Dispersed* dan *Random* [27]

Pola temporal adalah cara tertentu di mana sesuatu dilakukan, diatur, atau sesuatu yang terjadi secara berulang. Sedangkan menurut Kamus Besar Bahasa Indonesia [38], yang dimaksud dengan temporal adalah sesuatu yang berhubungan atau berkenaan dengan waktu-waktu tertentu. Dengan demikian pola temporal atau disebut juga pola urutan waktu dapat diartikan sebagi sesuatu yang terjadi secara berulang yang dirunut berdasarkan alur waktu atau *time series*. Analisis *time series* merupakan kumpulan data-data yang dicatat dan diamati dalam serangkaian waktu. Data yang dikumpulkan bisa berbentuk data harian, data mingguan, data bulanan maupun data tahunan, penyajian data pola temporal biasanya dilakukan dalam

bentuk grafik atau diagram garis. Dengan mengetahui pola temporal akan sangat membantu dalam melakukan analisis *time series* karena data yang disajikan lebih sederhana dan mudah dipahami.

Analisis data spasial dan temporal adalah bidang penelitian yang muncul karena pengembangan dan penerapan teknik komputasi yang memungkinkan untuk menganalisis *database* yang berukuran besar. Model spasial temporal muncul ketika data yang dikumpulkan melintasi waktu dan juga ruang dan memiliki setidaknya satu posisi spasial dan satu waktu temporal. Suatu peristiwa dalam dataset spasial temporal menggambarkan fenomena yang ada pada waktu dan lokasi tertentu. Contohnya adalah epidemi demam chikungunya di kota Yogyakarta tahun 2008-2018, di mana data spasial adalah lokasi dan geometri objek kota Yogyakarta dengan informasi epidemi demam chikungunya, dan data temporal adalah *timestamp* atau *interval* waktu untuk objek spasial yang valid pada tahun 2008-2018.

Pesatnya pertumbuhan data spasial dan temporal menarik untuk diteliti dari berbagai sudut pandang disiplin ilmu, karena kumpulan besar data spasial temporal ini seringkali menyembunyikan informasi yang menarik dan juga pengetahuan yang berharga. Lingkungan pola spasial dalam penelitian ini adalah lingkup yang menekankan kepada tempat atau koordinat beroperasinya kapal perikanan, dimana pembagiannya dibatasi berdasarkan aspek geografis seperti letak dan lokasinya. Sedangkan lingkup pola temporal dalam penelitian ini adalah lingkup yang menekankan kepada waktu, yang dipilah melalui periodeisasi menjadi beberapa periode seperti bulan dan tahun.

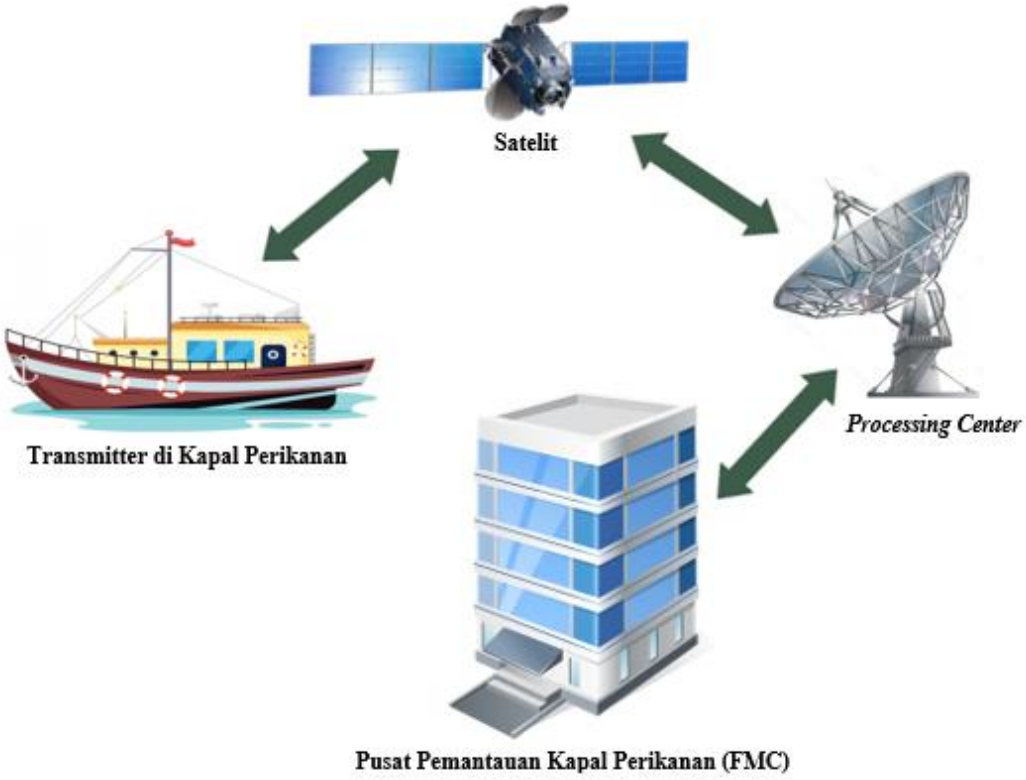
2.2 Definisi Kapal Perikanan

Undang-undang Nomor 45 Tahun 2009 tentang Perikanan [39] dan Peraturan Menteri Kelautan Dan Perikanan Republik Indonesia Nomor 42/Permen-KP/2015 Tentang Sistem Pemantauan Kapal Perikanan [40], mendefinisikan kapal perikanan adalah kapal, perahu, atau alat apung lainnya yang digunakan untuk melakukan penangkapan ikan, mendukung operasi penangkapan ikan, pelatihan perikanan, dan penelitian atau eksplorasi perikanan.

2.3 *Vessel Monitoring System (VMS)*

Vessel Monitoring System (VMS) atau Sistem Pemantauan Kapal Perikanan, yang selanjutnya disingkat SPKP, adalah salah satu sistem pengawasan kapal perikanan dengan menggunakan peralatan yang telah ditentukan untuk mengetahui pergerakan dan aktifitas kapal perikanan [40]. SPKP/ VMS merupakan salah satu bentuk sistem yang digunakan untuk pengawasan dan pengendalian dibidang penangkapan dan pengangkutan ikan, dengan menggunakan satelit dan peralatan transmitter yang ditempatkan pada kapal perikanan guna mempermudah pengawasan dan pengendalian terhadap kegiatan atau aktifitas kapal ikan berdasarkan posisi kapal yang terpantau di Pusat Pemantauan Kapal Perikanan (*Fisheries Monitoring Center*).

Cara kerja VMS adalah transmitter yang terpasang pada kapal perikanan mentransmisikan data posisi kapal berdasarkan *Global Positioning System* (GPS) setiap jam atau kurang ke satelit komunikasi (Argos, Inmarsat, Iridium, dan Garuda-1), data yang terpantau oleh satelit diolah di *Processing Center*, kemudian disampaikan ke Pusat Pemantauan Kapal Perikanan (FMC) [14], Direktorat Jenderal Pengawasan dan Pengendalian Sumberdaya Kelautan dan Perikanan. Selain data posisi kapal sebagai bahan analisa dan evaluasi juga didapatkan informasi mengenai kecepatan kapal, pola gerakan kapal dan rekaman data terdahulu maupun *near real time* (mendekati saat terjadi). Sistem pemantauan kapal perikanan (*vessel monitoring system*) selain dapat digunakan untuk tujuan penegakkan hukum perikanan, juga memberikan informasi tentang distribusi spasial dan temporal dari aktivitas penangkapan ikan yang bisa digunakan sebagai pedoman dalam pengelolaan perikanan yang berkelanjutan [12]. Ilustrasi cara kerja VMS dapat dilihat pada Gambar 2.2.



Gambar 2. 2 Cara Kerja *Vessel Monitoring System (VMS)*

Untuk menjalankan sistem pengelolaan perikanan yang bertanggung jawab, maka setiap kapal perikanan diwajibkan untuk memasang transmitter *Vessel Monitoring System* (*VMS*), sebagaimana diamanatkan dalam Undang-Undang Nomor 45 tahun 2009 [41] tentang Perubahan Atas Undang-Undang Nomor 31 tahun 2004 tentang Perikanan. Peraturan Menteri Nomor 5 tahun 2007 tentang Penyelenggaraan Sistem Pemantauan Kapal Perikanan, yang mengamanatkan kewajiban kapal-kapal perikanan untuk memasang transmitter *Vessel Monitoring System (VMS*) dan Peraturan Menteri Kelautan Dan Perikanan Republik Indonesia Nomor Per.05/Men/2008 Tentang Usaha Perikanan Tangkap [42], pada Pasal 88 ayat 1 Setiap kapal penangkap ikan dan atau kapal pengangkut ikan berbendera asing, wajib memasang dan mengaktifkan transmitter atau sistem pemantauan kapal (*Vessel Monitoring System*), ayat 2 Setiap kapal penangkap dan atau kapal pengangkut ikan berbendera Indonesia berukuran lebih dari 30 (tiga puluh) GT wajib memasang dan mengaktifkan transmitter atau sistem pemantauan kapal (*Vessel Monitoring System*), ayat 3 Pelaksanaan pemasangan atau pengaktifan transmitter atau sistem pemantauan kapal (*Vessel Monitoring System*) sebagaimana

dimaksud ayat 1 dan ayat 2 dilakukan sesuai dengan Peraturan Menteri yang mengatur mengenai penyelenggaraan sistem pemantauan kapal perikanan yaitu Nomor 10/PERMEN-KP/2019[43].

Penyelenggaraan sistem pemantauan kapal perikanan memiliki tujuan untuk meningkatkan efektifitas pengelolaan sumberdaya ikan melalui pengendalian dan pemantauan terhadap kapal perikanan, meningkatkan efisiensi dan efektifitas pengelolaan usaha perikanan yang dilakukan oleh perusahaan perikanan, meningkatkan ketaatan kapal perikanan terhadap ketentuan peraturan perundang-undangan yang berlaku serta memperoleh data dan informasi kegiatan kapal perikanan dalam rangka pengelolaan sumberdaya ikan secara berkelanjutan. Kesemuanya ini dilaksanakan untuk menjaga dan mengendalikan sumberdaya ikan agar dapat dimanfaatkan secara lestari dan bertanggung jawab. Keberadaan sistem pemantauan kapal perikanan, memiliki manfaat bagi pemerintah Indonesia, dimana dapat melindungi Zona Ekonomi Eksklusif Indonesia (ZEEI) dari kegiatan pelanggaran kapal perikanan, melacak dan mengidentifikasi tindakan *illegal fishing*.

Manfaat selanjutnya dapat menunjukkan penyebaran kapal perikanan di wilayah penangkapan ikan dan membantu penegak hukum untuk memeriksa apakah kapal-kapal tersebut beroperasi di jalur penangkapan ikan yang telah ditetapkan. Dapat segera memberikan informasi posisi kapal yang meminta bantuan saat situasi darurat, seperti perompakan, kebakaran atau kecelakaan sehingga bantuan dapat diberikan secara cepat dan efektif. Sistem ini juga memberikan manfaat yang besar bagi pengusaha atau pemilik kapal, karena mereka dapat memanfaatkan informasi dari *Vessel Monitoring System* untuk mengetahui posisi, pergerakan dan aktivitas armada kapal perikanan, meningkatkan efisiensi dalam melakukan usaha penangkapan ikan, menjamin kelangsungan usaha penangkapan ikan yang kondusif, serta penyelamatan (*save and rescue*) terhadap kapal perikanan yang menghadapi masalah di laut.

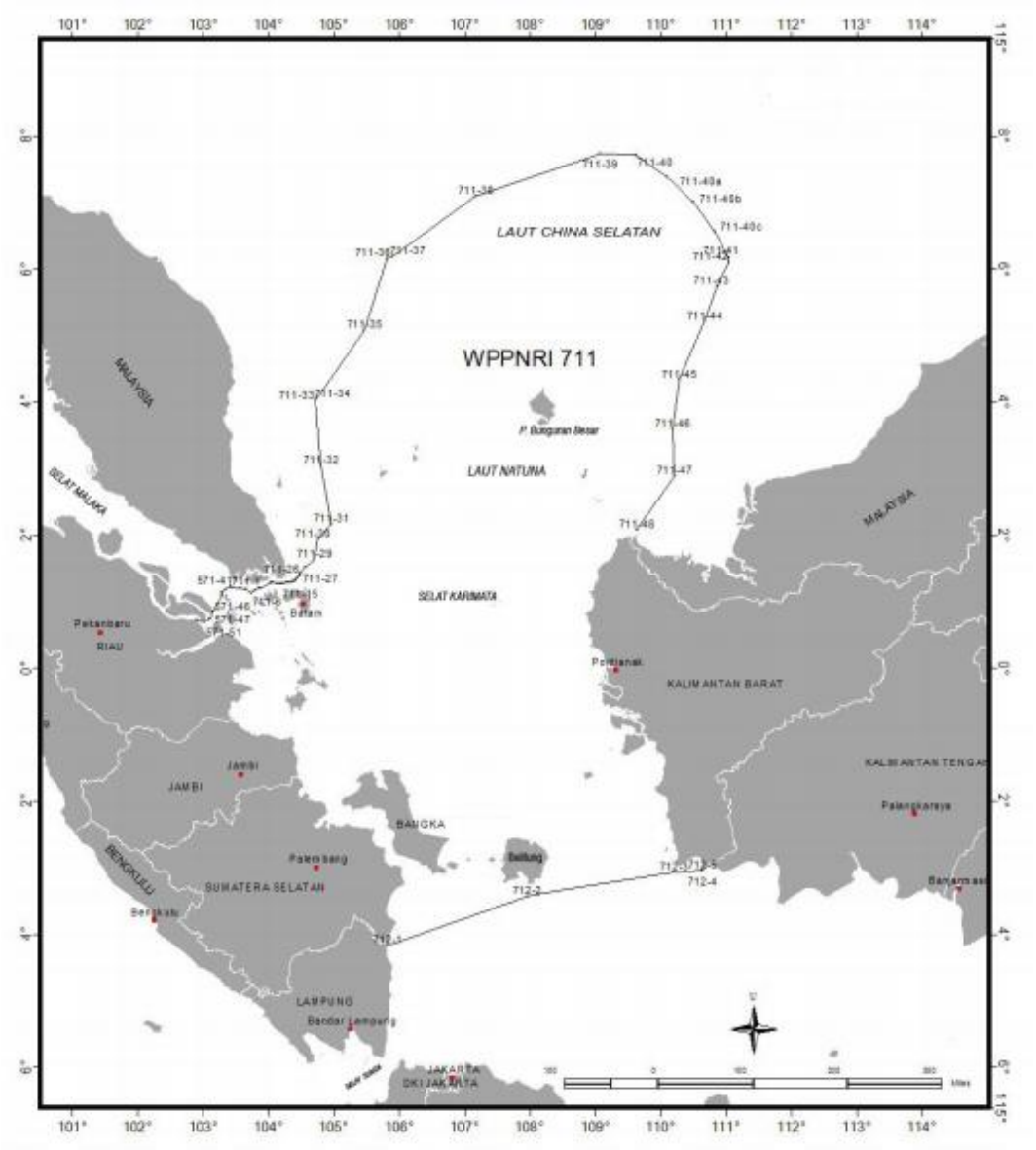
2.4 Wilayah Pengelolaan Perikanan Negara Republik Indonesia (WPPNRI) -711

Menurut Peraturan Menteri Kelautan dan Perikanan Nomor 18/PERMEN-KP/2014 [3] tentang Wilayah Pengelolaan Perikanan Negara Republik Indonesia atau disingkat dengan WPPNRI merupakan wilayah pengelolaan perikanan untuk penangkapan ikan, pembudidayaan ikan, konservasi, penelitian, dan pengembangan perikanan yang meliputi perairan pedalaman, perairan kepulauan, laut teritorial, zona tambahan, dan zona ekonomi ekslusif Indonesia (ZEEI). WPPNRI ditetapkan terbagi menjadi 11 wilayah pengelolaan yaitu WPP571, WPP572, WPP573, WPP711, WPP712, WPP713, WPP714, WPP715, WPP716, WPP717, WPP718. Wilayah pengelolaan perikanan ini menggambarkan pembagian wilayah yang didasarkan pada ekologi, karakteristik wilayah, dan stok potensi sumber daya ikan yang digunakan sebagai dasar manajemen pengelolaan perikanan secara lestari dan berkelanjutan.

Wilayah Pengelolaan Perikanan Negara Republik Indonesia (WPPNRI) 711 meliputi Selat Karimata, Laut Natuna dan Laut Cina Selatan. Secara administratif masuk kedalam tujuh (7) wilayah Provinsi yaitu Prov. Riau, Prov. Kepulauan Riau, Prov. Jambi, Prov. Sumatera Selatan, Prov. Kepulauan Bangka Belitung, Prov. Kalimantan Tengah dan Prov. Kalimantan Barat, WPPNRI-711 secara umum memiliki batasan sebagai berikut:

1. Sebelah Utara berbatasan dengan batas terluar ZEE Indonesia – Vietnam;
2. Sebelah Timur berbatasan dengan batas terluar ZEE Indonesia – Malaysia, perbatasan darat Indonesia – Malaysia di Prov. Kalimantan Barat;
3. Sebelah Selatan berbatasan dengan Kab. Tulangbawang, Prov. Lampung, Kab. Ogan Komiring Ilir, Prov. Sumatera Selatan dan Kab. Belitung, Kab. Belitung Timur, Prov. Bangka Belitung;
4. Sebelah Barat berbatasan dengan Kab. Kampar, Kab. Bengkalis, Prov. Riau batas laut Indonesia – Singapura, batas terluar ZEE Indonesia – Malaysia.

Untuk lebih jelasnya tentang wilayah WPPNRI-711 dapat dilihat pada Gambar 2.3 berikut [3] :



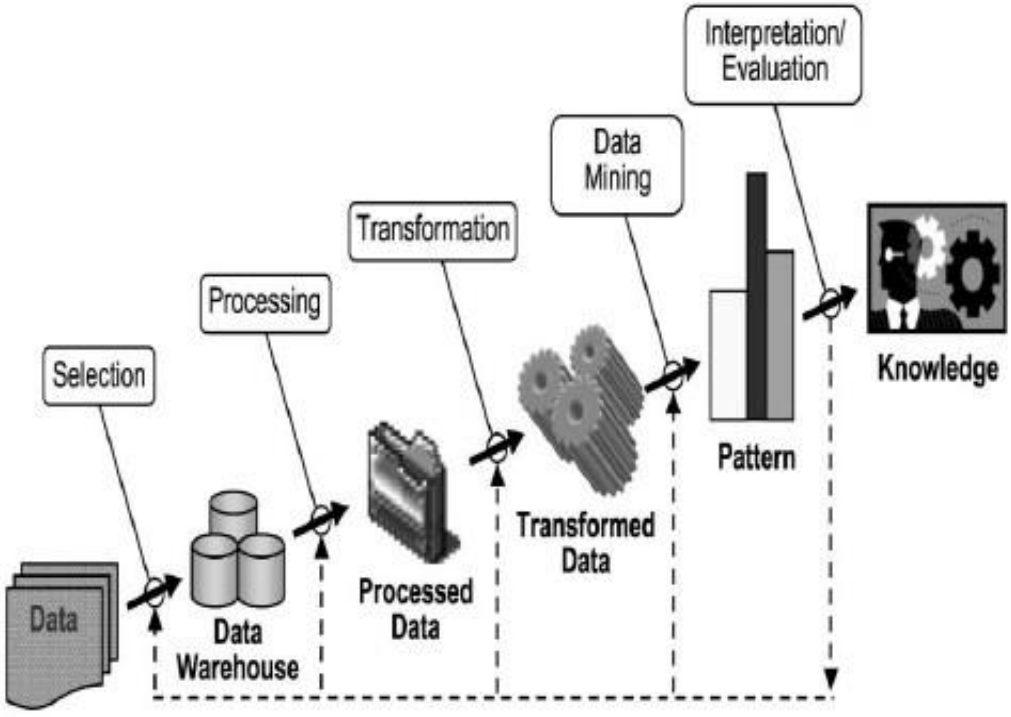
Gambar 2. 3 Peta WPPNRI -711

2.5 *Data Mining*

Data Mining adalah sebuah proses dan langkah analisis untuk menemukan informasi ataupun pengetahuan didalam sebuah basis data (*knowledge discovery in data bases*) [18]. *Data mining* merupakan proses eksplorasi dan analisis untuk menemukan pola dan aturan yang penting pada data yang berukuran besar [44], Informasi atau pengetahuan ini bisa berupa pola data atau relasi antar data yang valid [45] diperoleh melalui suatu proses yang cukup rumit seperti menggunakan artificial intelligence, teknik statistik, ilmu matematika, *machine learning*, dan lain

sebagainya. Data mining disebut juga *pattern recognition* merupakan pengolahan data untuk menemukan pola yang tersembunyi dari sebuah basis data dan hasil dari pengolahannya dapat digunakan untuk mengambil keputusan di masa depan. Data tersebut bisa berasal dari suatu basis data, data *warehouse*, atau tempat penyimpanan data lainnya [46].

Data mining digunakan untuk data yang berskala besar dan banyak diaplikasikan di berbagai bidang kehidupan baik industri, kesehatan, pendidikan, perdagangan, cuaca, ilmu pengetahuan dan teknologi serta dibidang lainnya. Secara umum penelitian *data mining* membahas metode-metode seperti *clustering*, *klasifikasi*, *regresi*, *seleksi variabel*, dan *market basket analisis.* Tahapan yang dilakukan pada proses data mining diawali dari seleksi data, tahap *preprocessing* untuk memperbaiki kualitas data, transformasi, proses *data mining* serta tahap interpretasi dan evaluasi yang menghasilkan output berupa pengetahuan atau informasi baru. Secara detail dapat dilihat pada Gambar 2.4 berikut [18]:



Gambar 2. 4 Proses *Data Mining* [18]

Proses *data mining* biasanya dilakukan dengan menggunakan teknik pengenalan pola seperti teknik statistik dan matematika. Urutan proses penemuan pengetahuan melalui langkah-langkah berikut [46]:

1. Pembersihan data, merupakan proses awal yang dilakukan dengan menghapus data yang dianggap mengganggu penelitian (*noise*) dan data yang tidak konsisten.
2. Integrasi data, proses ini dilakukan jika memiliki beberapa sumber data yang harus digabungkan menjadi satu.
3. Pemilihan data, yaitu proses untuk memperoleh data yang relevan dengan analisis penelitian dari basis data.
4. Transformasi data, yaitu proses mengubah format data sesuai keperluan dalam proses penelitian, dimana data ditransformasikan dan dikonsolidasikan ke dalam bentuk yang sesuai dengan kebutuhan.
5. *Data mining*, merupakan proses penting dan utama berupa metode cerdas yang akan diterapkan untuk mengekstrak pola data.
6. Evaluasi pola, untuk mengidentifikasi pola yang benar-benar menarik dan mewakili pengetahuan atau informasi yang diharapkan.
7. Presentasi pengetahuan, merupakan proses visualisasi dan representasi.

2.6 *Machine Learning*

Machine learning merupakan kemampuan komputer untuk melakukan pembelajaran tanpa harus menjelaskan secara eksplisit kepada komputer [47]. Dengan kata lain *machine learning* adalah bagaimana memberikan kemampuan kepada komputer untuk melakukan aktivitas belajar guna menyelesaikan masalah secara mandiri. Algoritma *machine learning* dapat dikelompokkan berdasarkan masukan dan keluaran yang diharapkan, adapun tipe algoritma *machine learning* diantaranya adalah *supervised learning, unsupervised learning, semi-supervised learning,* dan *reinforcement learning.*

Supervised learning merupakan algoritma yang mampu membangkitkan suatu fungsi yang memetakan input ke autput yang diinginkan. Kualitas hasil pembelajaran sangat bergantung kepada kesesuaian input dan output yang diberikan, sehingga user sangat berperan dalam menvalidasi input dan output

tersebut, biasanya digunakan untuk menyelesaikan masalah klasifikasi dan regresi. Sementara *unsupervised learning* mempunyai tujuan untuk memodelkan sekumpulan input secara otomatis tanpa ada panduan, artinya data yang dipelajari hanya input tanpa label kelas dan biasanya dipergunakan untuk masalah klasterisasi [48]. *Semi-supervised learning* ini menggabungkan antara supervised dengan unsupervised untuk menghasilkan suatu fungsi. Tipe lain dari *machine learning* adalah *reinforcement learning*, tipe ini mengajarkan bagaimana cara bertindak untuk menghadapi masalah, yang suatu tindakan itu mempunyai dampak.

2.7 *Clustering*

Clustering adalah sebuah proses untuk mengelompokan data ke dalam beberapa cluster atau kelompok sehingga data dalam satu cluster memiliki tingkat kemiripan yang maksimum dan data antar cluster memiliki kemiripan yang minimum[49]. *Clustering* merupakan proses partisi satu set objek data ke dalam himpunan bagian yang disebut *cluster*. Objek yang di dalam *cluster* memiliki kemiripan karakteristik antar satu sama lainnya dan berbeda dengan *cluster* yang lain. Partisi tidak dilakukan secara manual melainkan dengan suatu algoritma *clustering*. Oleh karena itu, *clustering* sangat berguna dan bisa menemukan *group* atau kelompok yang tidak dikenal dalam data.

Hasil *clustering* yang baik akan menghasilkan tingkat kesamaan yang tinggi dalam satu kelas dan tingkat kesamaan yang rendah antar kelas. Kesamaan yang dimaksud merupakan pengukuran secara numerik terhadap dua buah objek. Nilai kesamaan antar kedua objek akan semakin tinggi jika kedua objek yang dibandingkan memiliki kemiripan yang tinggi, begitu juga dengan sebaliknya. Kualitas hasil *clustering* sangat bergantung pada metode yang digunakan, metode *clustering* juga harus dapat mengukur kemampuannya sendiri dalam usaha untuk menemukan suatu pola tersembunyi pada data yang sedang diteliti. Terdapat berbagai metode yang dapat digunakan untuk mengukur nilai kesamaan antar objek-objek yang dibandingkan. Salah satunya ialah dengan *weighted Euclidean Distance*. *Euclidean distance* menghitung jarak dua buah point dengan mengetahui nilai dari masing-masing atribut pada kedua poin tersebut. Jarak adalah pendekatan yang umum dipakai untuk menentukan kesamaan atau ketidaksamaan dua vektor

fitur yang dinyatakan dengan ranking. Apabila nilai ranking yang dihasilkan semakin kecil nilainya maka semakin dekat/ tinggi kesamaan antara kedua vektor tersebut. Teknik pengukuran jarak dengan metode *Euclidean* menjadi salah satu metode yang paling umum digunakan.

Algoritma *clustering* membangun sebuah model dengan melakukan serangkaian pengulangan dan berhenti ketika model tersebut telah terpusat dan batasan dari segmentasi telah stabil. Hasil dari *clustering* yang bagus tergantung dengan ukuran kesamaan dan metode yang digunakan. Pendekatan dalam cluster berdasarkan saran dari Fraley dan Raftery, membagi metode pengelompokan menjadi dua kelompok utama yaitu metode hirarkis dan metode partisi [50].

1. Metode Hirarki yaitu metode yang membentuk cluster dengan mempartisi secara berulang-ulang dari atas ke bawah atau sebaliknya. Hasil dari metode hirarki berupa dendogram yang mewakili kelompok objek dan tingkat kesamaan di mana terdapat perubahan pengelompokkan. Sebuah pengelompokan objek data diperoleh dengan memotong dendrogram pada tingkat kemiripan yang diinginkan [51].
2. Metode Partisi yaitu metode yang membuat inisial partisi k di awal, dimana parameter k merupakan jumlah partisi untuk membentuk. Kemudian secara iterative menggunakan teknik relokasi dengan mencoba berulang-ulang memindahkan objek dari satu kelompok ke kelompok lain untuk mendapatkan partisi yang optimal. Tipe metode partisi ini seperti K-Means, K-Medoids dan CLARANS [46].

2.8 *Algoritma K-Means*

Algoritma K-means merupakan salah satu algoritma *clustering* yang bertujuan untuk membagi data menjadi beberapa kelompok. Algoritma ini menerima masukan berupa data tanpa label kelas (*unsupervised learning*). Proses *clustering* K-Means, dilakukan oleh komputer dengan mengelompokan sendiri data-data yang menjadi masukannya tanpa mengetahui terlebih dulu target kelasnya. Pada setiap *cluster* terdapat titik pusat (*centroid*) yang merepresentasikan *cluster* tersebut. Algoritma untuk melakukan K-Means *clustering* adalah sebagai berikut [52]:

1. Menentukan nilai k sebagai jumlah *cluster* yang dibentuk
2. Menentukan nilai awal centroid atau titik pusat *cluster*, pada tahapan ini ditentukan nilai centroid secara random, setelah algoritma berjalan nilai centroid dicari dengan menggunakan rumus 2.1 [53].

$$C = \sum_{i=1}^{n} \frac{x_i}{n} \quad (2.1)$$

Di mana:

C = centroid pada cluster

x_i = titik / objek ke-i

n = jumlah objek

3. Menghitung jarak antara titik centroid dengan titik setiap objek, dapat dilakukan dengan menggunakan *Euclidean Distance* dengan rumus 2.2 [53].

$$D_e = \sqrt{(x_i - s_i)^2 + (y_i - t_i)^2} \quad (2.2)$$

Di mana:

D_e = *Euclidean Distance*

i = banyaknya data

(x, y) = koordinat data

(s, t) = koordinat centroid.

4. Mengelompokan objek ke centroid *cluster* terdekat berdasarkan *Euclidean Distance*. Penentuan anggota *cluster* adalah dengan memperhitungkan jarak minumum objek.
5. Memperbarui nilai centroid setiap *cluster*, dengan menghitung kembali semua nilai titik centroid. Untuk mendapatkan pusat cluster baru bisa dihitung dari rata-rata nilai anggota cluster dan pusat cluster. Pusat cluster yang baru digunakan untuk melakukan iterasi selanjutnya, cara penentuan pusat cluster baru:
   - Cari jumlah anggota tiap cluster.
   - Hitung pusat baru dengan rumus:

$$Pusat\ cluster\ baru = \frac{x1 + x2 + x3 \ldots + xn + xp}{Jumlah + 1} \quad (2.\ 3)$$

Dimana:

$x1, x2, x3, \ldots, xn$ = anggota cluster

xp = pusat cluster lama.

6. Mengulangi langkah ke-2 s.d. 5 sampai nilai dari titik centroid konvergen dan stabil, dimana semua centroid yang dihasilkan pada literasi saat ini sama dengan centroid yang dihasilkan pada iterasi sebelumnya dan nilai SSE stabil tidak mengalami perubahan yang berarti.

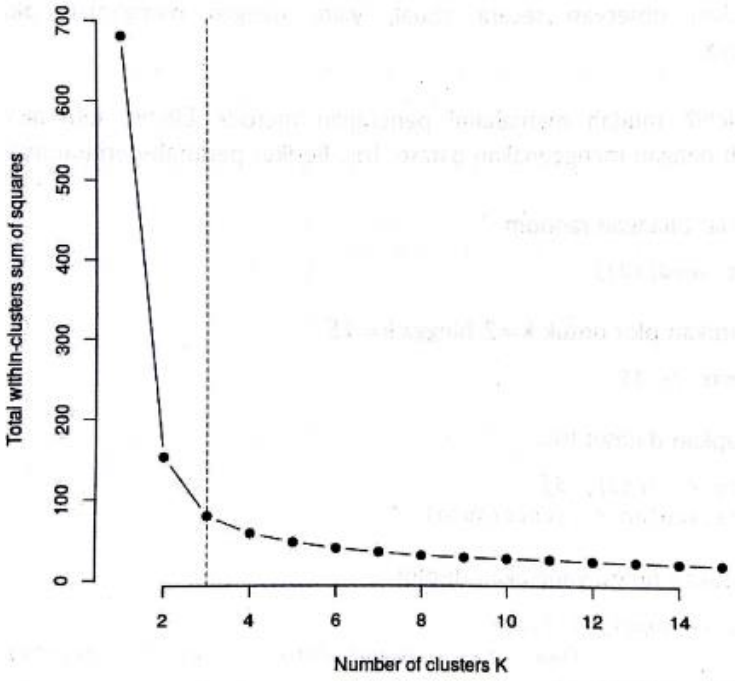
2.9 Metode *Elbow*

Metode *Elbow* memberikan ide atau gagasan cara memilih nilai cluster dengan menemukan *within cluster dissimilarity* sebagai fungsi dari jumlah cluster (k). Metode ini akan menghasilkan informasi untuk menentukan jumlah cluster terbaik dengan melihat persentase hasil perbandingan antara jumlah cluster yang membentuk siku pada suatu titik. Algoritma Metode *Elbow* dalam menentukan nilai k pada K-Means [54].

1. Mulai
2. Inisialisasi awal nilai k
3. Naikkan nilai k
4. Hitung hasil *sum of square error* dari tiap nilai k
5. Melihat hasil *sum of square error* dari nilai k yang turun secara drastis
6. Tetapkan nilai k yang beberbentuk siku
7. Selesai

Hasil persentase yang berbeda dari setiap nilai *cluster* dapat ditunjukan dengan menggunakan grafik sebagai sumber informasinya. Grafik akan menunjukkan beberapa nilai k yang mengalami penurunan paling besar dan selanjutnya hasil dari nilai k akan turun secara perlahan-lahan sampai hasil dari nilai k tersebut stabil. *Cluster* pertama akan memberikan informasi namun pada titik tertentu akan turun secara dramatis dan memberikan sebuah lekukan pada grafik yang disebut dengan “kriteria siku”. Namun metode ini juga memiliki kelemahan

dari yaitu tidak selalu dapat teridentifikasi titik sikunya [55]. Contoh nilai *sum of square error* (SSE) dihasilkan oleh cluster yang diplotkan pada sebuah grafik berdasarkan jumlah cluster, seperti ditampilkan pada Gambar 2.5.



Gambar 2. 5 *Elbow Method* [53]

Sum of Square Error (SSE) merupakan metode yang sederhana dan banyak digunakan dalam mengukur kriteria pada evaluasi pengelompokan karena SSE mengukur keseragaman antar data dalam satu cluster. Metode ini menggunakan informasi yang ada pada objek saja, sehinga termasuk ke dalam kriteria kualitas internal. Semakin kecil nilai SSE maka semakin bagus hasil clustering-nya, SSE didefinisikan sebagai *sum of the squared distance* antara setiap anggota cluster dengan *centroid*-nya. Jika diplot terhadap nilai SSE maka *error* akan berkurang karena nilai k semakin besar atau jumlah *cluster* semakin meningkat, rumus untuk mencari SSE terdapat pada persamaan 2.3.[56]

$$SSE = \sum_{i=1}^{k} \sum_{x \in Di} \left|\left|x - m_i\right|\right|^2 \qquad (2.\ 4)$$

Keterangan:

k = Jumlah *cluster*

Di = Set data D ke-i

x = Nilai set data

mi = Mean (rata-rata) vektor dari setiap *cluster*

2.10 Pengujian *Silhouette Coefficient*

Metode evaluasi yang akan digunakan pada sistem ini adalah metode *silhouette coefficient*. Metode ini berfungsi untuk menguji kualitas dari cluster yang dihasilkan. Metode ini merupakan metode validasi cluster yang menggabungkan metode *cohesion* dan *separation* [55]. Untuk menghitung nilai *silhoutte coefisient* diperlukan jarak antar objek dengan menggunakan rumus *Euclidean Distance*. Tahapan untuk menghitung nilai *silhoutte coeffisien* adalah sebagai berikut [55]:

1. Untuk setiap objek i, hitung rata-rata jarak dari objek i dengan seluruh objek yang berada dalam satu cluster. Akan didapatkan nilai rata-rata yang disebut ai.
2. Untuk setiap objek i, hitung rata-rata jarak dari objek i dengan objek yang berada di cluster lainnya. Dari semua jarak rata-rata tersebut ambil nilai yang paling kecil. Nilai ini disebut bi.
3. Setelah itu maka untuk objek i memiliki nilai *silhoutte coefisien*:

$$Si = (bi - ai) / \max(ai, bi)$$

Hasil perhitungan nilai *silhoutte coeffisien* dapat bervariasi antara -1 hingga 1. Hasil *clustering* dikatakan baik jikai nilai *silhoutte coeffisien* bernilai positif ($ai < bi$) dan ai mendekati 0, sehingga akan menghasilkan nilai *silhoutte coeffisien* yang maksimum yaitu 1 saat ai = 0. Maka dapat dikatakan, jika si = 1 berarti objek i sudah berada dalam *cluster* yang tepat. Jika nilai si = 0 maka objek i berada di antara dua *cluster* sehingga objek tersebut tidak jelas harus dimasukan ke dalam *cluster* A atau *cluster* B. Akan tetapi, jika si = -1 artinya struktur *cluster* yang dihasilkan *overlapping*, sehingga objek i lebih tepat dimasukan ke dalam *cluster* yang lain. Nilai rata-rata *silhoutte coeffisien* dari tiap objek dalam suatu *cluster* adalah suatu ukuran yang menunjukan seberapa dekat data dikelompokan

dalam cluster tersebut. Pada Tabel 2.1 berikut adalah kategori nilai *silhoutte coeffisien* berdasarkan pendapat Kaufman dan Rousseeuw [57]:

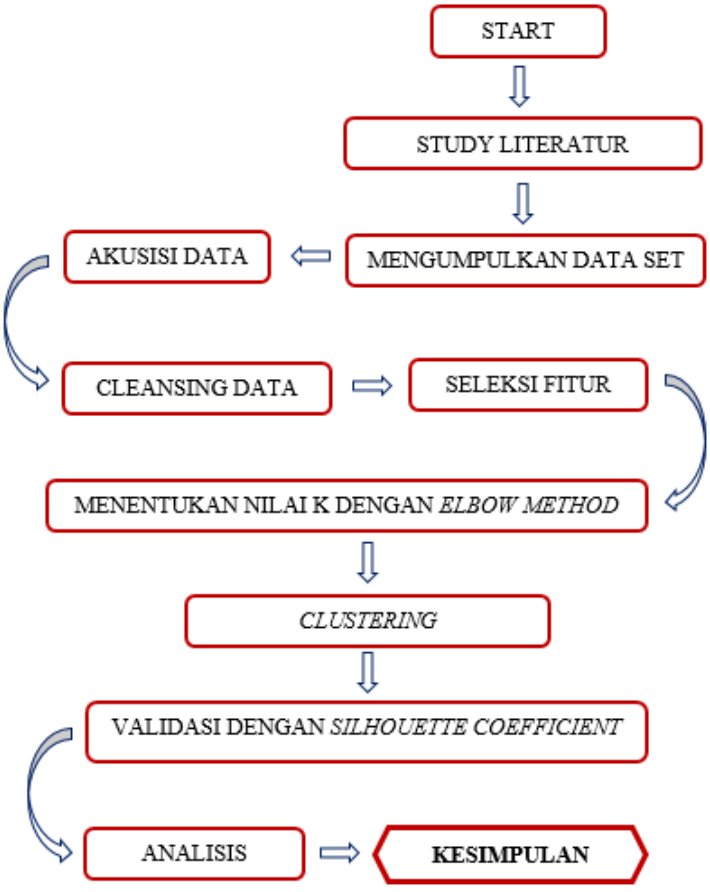
Tabel 2. 1 Kategori Nilai *Silhouette Coeffisien* dan Interpretasinya

Silhouette Coefficient (SC)	**Proposed Interpretation**
0.71 – 1.00	Strong Stucture
0.51 - 0.70	Reasonable /Medium Structure
0.26 - 0.50	Weak Structure
≤ 0.25	No structure

Halaman ini sengaja dikosongkan

BAB III
METODOLOGI PENELITIAN

Pada bab ini akan dijelaskan mengenai desain dan langkah–langkah metodologi penelitian secara sistematis yang akan dijadikan acuan dalam penelitian tentang *clustering* distribusi spasial dan temporal kapal perikanan berdasarkan data *vessel monitoring system (VMS)* di Wilayah Pengelolaan Perikanan Negara Republik Indonesia (WPPNRI)-711. Secara garis besar langkah-langkah pada penelitian ini digambarkan pada Gambar 3.1.



Gambar 3. 1 Diagram Penelitian

3.1 Studi Literatur

Studi literatur yang dilakukan oleh penulis yaitu dengan melakukan pencarian terhadap berbagai sumber tertulis, berupa buku, majalah, artikel, jurnal dan aturan perundang-undangan serta berbagai dokumen yang relevan dengan permasalahan yang dikaji. Studi literatur ini dilakukan setelah menentukan topik penelitian dan ditetapkannya rumusan permasalahan, sehingga informasi yang diperoleh dapat dijadikan rujukan dalam memperkuat argumentasi-argumentasi yang ada. Dari literatur tersebut dapat diketahui persamaan ataupun perbedaan terhadap penelitian yang dilakukan oleh penulis maupun dengan peneliti lainya. Keluaran yang didapatkan pada proses studi literatur berupa kajian pustaka sebagai acuan dalam konsep, metode dan landasan teori untuk pengerjaan penelitian ini.

3.2 Mengumpulkan Data Set

Data dipergunakan sebagai bahan untuk melihat lingkungan, objek, kejadian, ataupun suatu konsep dalam sebuah penelitian. Sesuai dengan topik dan batasan masalah pada penelitian ini maka data yang akan diambil adalah data *vessel monitoring system* (VMS) Tahun 2016 s/d 2018 yang beroperasi di WPPNRI-711. Data akan diperoleh dari Direktorat Pemantauan dan Operasi Armada, Direktorat Jenderal Pengawasan Sumber Daya Kelautan dan Perikanan, Kementerian Kelautan dan Perikanan Republik Indonesia.

3.3 Akusisi Data

Akuisisi data sebagai suatu proses yang berfungsi untuk mengambil, mengumpulkan dan menyiapkan data, hingga memprosesnya untuk menghasilkan data yang dikehendaki. Proses ini merupakan tahap yang sangat penting dalam penelitian, karena akan manentukan kualiatas data yang akan diolah pada tahap selanjutnya. Jenis serta metode yang di pilih bertujuan untuk menyederhanakan setiap langkah yang dilakukan, karena dalam sebuah penelitian data yang diperoleh tidak semuanya bisa diproses secara langsung. Hal ini disebabkan berbagai jenis format dan bentuk penyimpanan data dan juga banyaknya atribut ataupun fitur yang dimiliki sebuah data. Proses akuisisi data bertujuan untuk menyiapkan dan

menyimpan data dalam bentuk yang siap untuk dipergunakan dalam proses selanjutnya.

3.4 Pembersihan Data (*Data Cleansing*).

Data cleansing adalah proses analisa kualitas data dengan cara mengubah, membuang duplikasi data, mendeteksi data yang tidak lengkap, tidak akurat, atau memperbaiki data yang memiliki format yang salah dalam basis data guna menghasilkan data berkualitas. Pada umumnya data yang diperoleh mempunyai isi yang belum siap untuk digunakan secara langsung, seperti adanya data duplikasi dan data yang tidak lengkap, hal ini biasanya disebabkan oleh *instrument faulty*, kesalahan manusia ataupun kesalahan transmisi [56]. Pembersihan data akan mempengaruhi performa dari teknik *clustering* karena data yang ditangani akan berkurang jumlah dan kompleksitasnya, selanjutnya data diubah atau digabung ke dalam format yang sesuai dengan tujuan penelitian ini untuk proses lebih lanjut.

3.5 Seleksi Fitur

Seleksi fitur adalah bagian dari analisis yang mempertegas, memfokuskan dan membuang hal yang tidak penting, serta mengatur data agar kesimpulan akhir dapat dilakukan. Data yang ada di *data set* tidak semuanya dipergunakan untuk penelitian ini, karena hanya data yang sesuai untuk dianalisis yang akan diambil. Seleksi fitur digunakan untuk mengurangi dimensi data dan digunakan untuk meningkatkan efektifitas dan efisiensi kinerja dari algoritma *clustering*. Ada dua titik berat seleksi fitur dalam pendekatan *machine learning* yaitu memilih fitur yang akan digunakan dan mengkombinasikan fitur-fitur tersebut.

3.6 Penentuan Nilai k dengan Metode *Elbow*

Karena tidak ada satu pun metode yang benar-benar dapat dengan eksak untuk menentukan nilai k (jumlah *cluster*), sehingga telah dikembangkan berbagai metode sesuai dengan kebutuhan rill [53]. Dalam penelitian ini jumlah data yang dipergunakan mencapai 7.029.661 data VMS. Maka *Elbow Method* dipilih untuk

menentukan nilai k atau jumlah *cluster* karena memiliki kompleksitas paling sederhana dan efisien dari segi komputasinya, diketahui *Elbow Method* memiliki waktu komputasi paling cepat jika dibandingkan dengan metode *Gab Statistic, Sihouette Coefficient* dan *Canopy* [58], dengan nilai perbandingan dapat dilihat pada Tabel 3.1.

Tabel 3. 1 Perbandingan waktu komputasi *Elbow Method*, *Gab Statistic*, *Sihouette Coefficient* dan *Canopy* [58].

No.	Name	K value	Execution Time
1	Elbow Method	2	1.830 s
2	Gap Statistic	2	9.763 s
3	Silhouette Coefficient	2	8.648 s
4	Canopy	2	2.120 s

Menentukan jumlah nilai k merupan sebuah proses yang sangat penting dalam *clustering*, karena nilai k ini akan sangat mempengaruhi perolehah hasil persentase sebuah proses *clustering*, yaitu sejumlah kelompok yang akan memiliki varian, kekompakan dan kemiripan yang paling optimum. Metode yang digunakan untuk menentukan nilai k dalam penelitian ini adalah menggunakan *elbow method.* Metode *Elbow* merupakan suatu teknik yang digunakan untuk menghasilkan informasi jumlah cluster terbaik dengan cara melihat persentase hasil perbandingan jumlah cluster yang akan membentuk sebuah titik siku pada sebuah grafik.

3.7 Implementasi *Clustering*

Clustering adalah suatu metode *Data Mining* yang melakukan proses pemodelan tanpa supervisi (*unsupervised*) dan salah satu metode yang melakukan pengelompokan data dengan sistem partisi. *Clustering* digunakan untuk pengelompokkan data berdasarkan kemiripan pada objek data. Dalam penelitian ini digunakan algoritma *K-means* untuk proses *clustering* data *Vessel Monitoring System (VMS*). Algoritma *K-Means* pada dasarnya melakukan dua proses yakni proses pendeteksian lokasi pusat *cluster* dan proses pencarian anggota dari setiap *cluster*.

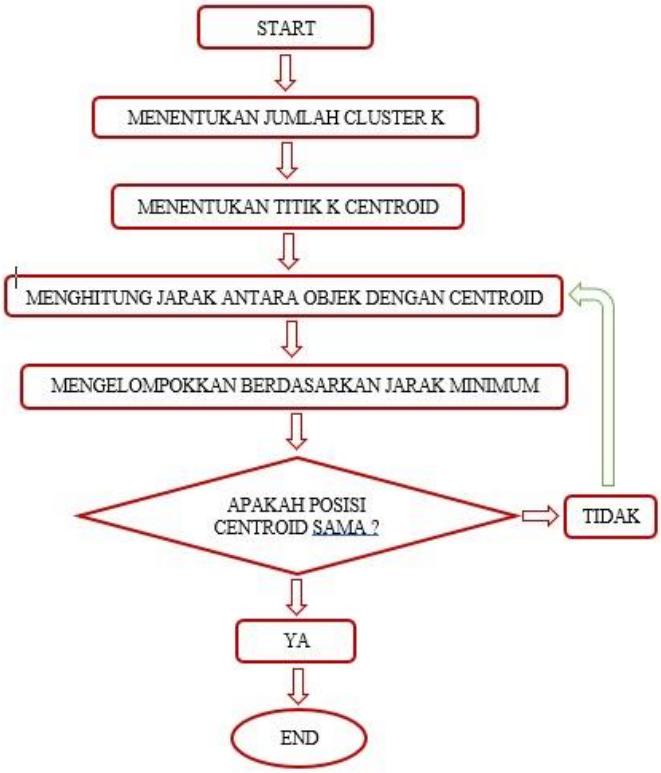
Clustering dipergunakan dalam penelitian ini untuk pemodelan secara representatif terhadap jutaan posisi spasial data VMS, yang terdistribusi secara luas di area WPPNRI-711 agar menjadi lebih sederhana dan bisa dipahami. Dengan memodelkan distribusi objek yang terletak di permukaan bumi menjadi lebih sederhana dapat dirumuskan hubungan antar objek atau membandingkan pola objek disuatu lokasi dengan pola objek di lokasi lainnya [27]. Algoritma *K-Means* dipergunakan untuk mengelompokkan data spasial berdasarkan variabel LONGITUDE dan LATITUDE dari data VMS. Dengan melakukan analisis *clustering* kita dapat memeriksa fitur apa yang biasanya muncul bersama sehingga dapat diketahui apa yang menjadi ciri atau pola dari sebuah *cluster*.

3.7.1 Algoritma *K-Means*

K-Means dipilih sebagai Algoritma yang dipergunakan dalam penelitian ini, karena *K-Means* merupakan salah satu *partitioning clustering* dengan *exclusive clusteringnya*. Artinya setiap data akan masuk menjadi angota salah satu cluster dan tidak ada data yang akan menjadi anggota dari dua cluster yang berbeda. Sehingga sesuai dengan tujuan penelitian ini untuk *clustering* secara spasial distribusi data VMS di WPPNRI-711. Selain itu karena data yang dipergunakan dalam penelitian ini cukup besar yaitu berjumlah 7.029.661 data VMS, mengutip pendapat dari para ahli ada beberapa pertimbangan lain penggunaan Algoritma *K-Means* diantaranya adalah:

1. *K-means* adalah salah satu metode yang memiliki kemudahan implementasi dan kecepatan dalam prosesnya [48].
2. Relatif efisien dan cepat, dan dapat bekerja dengan baik pada dataset yang berbentuk *spherical* (circular dua dimensi) [53].
3. Algoritma *K-Means* cukup efisien untuk pengklasteran set data yang besar, karena kompleksitas penghitungan pada algoritma ini bersifat linier terhadap jumlah titik-titik data [56].
4. *K-Means* adalah Algoritma yang simple, cepat dan efisien untuk dataset yang besar [59].
5. Sangat mudah diimplementasikan, dan kompleksitas waktunya adalah O (Nkt) sehingga cocok untuk kumpulan data yang sangat besar [60].

Proses *clustering* dalam penelitian ini dilakukan dengan bantuan aplikasi Rstudio menggunakan bahasa pemograman R. Adapun *flowchart* Alogaritma *K-Means* dapat dilihat pada Gambar 3.2.



Gambar 3. 2 *Flowchart* Algoritma *K-Means* [53]

Algoritma *K-Means* memiliki peroses alur sebegai berikut [48]:

a. Langkah pertama adalah menentukan jumlah cluster k, kemudian dari himpunan data yang akan diklasterisasi yang dalam penelitian ini adalah koordinat Longitude dan Latitude, dipilih sejumlah k objek data secara acak sebagai titik pusat (*centroid*) awal.
b. Langkah kedua adalah memasukkan setiap objek yang bukan centroid kedalam cluster terdekat berdasarkan ukuran jarak.
c. Langkah ketiga adalah setiap *centroid* akan diperbaharui berdasarkan rata-rata dari objek yang ada didalam *cluster*.
d. Langkah keempat adalah melakukan iterasi untuk langkah kedua dan ketiga sampai semua centroid konvergen dan stabil, dimana semua *centroid* yang dihasilkan pada iterasi terakhir sama dengan *centroid* yang dihasilkan pada iterasi sebelumnya.

3.8 Validasi Hasil *Clustering* Dengan *Silhouette Coefficient*

Pengujian hasil *clustering* dilakukan untuk mengetahui seberapa dekat relasi antara objek dalam sebuah *cluster* dan seberapa jauh sebuah *cluster* terpisah dengan *cluster* lain. Metode pengujian yang akan digunakan adalah *Silhouette Coefficient*. Metode *Silhouette Coefficient* merupakan gabungan dari dua metode yaitu metode *cohesion* yang berfungsi untuk mengukur seberapa dekat relasi antara objek dalam sebuah *cluster*, dan metode *separation* yang berfungsi untuk mengukur seberapa jauh sebuah *cluster* terpisah dengan *cluster* lain.

3.9 Analisis

Pada bagian ini merupakan proses intepretasi data setelah dilakukan *clustering*. Langkah berikutnya ialah mendeskripsikan pola spasial dan temporal dari setiap cluster. Hasil analisis berupa kelompok data dari setiap cluster yang disajikan dalam bentuk gambar, diagram, grafik, tabel, dan *flow chart*.

3.10 Kesimpulan

Kesimpulan merupakan jawaban dari tujuan penelitian ini, yaitu *clustering* distribusi spasial dan temporan aktivitas kapal perikanan di WPPNRI-711 berdasarkan data vessel monitoring system (VMS).

Halaman ini sengaja dikosongkan

BAB IV
HASIL PENELITIAN DAN PEMBAHASAN

4.1 Tahap *Preprocessing Data*

Data untuk penelitian ini adalah data *vessel monitoring system* (*VMS)* Tahun 2016 s/d 2018 yang diperoleh dari Direktorat Pemantauan dan Operasi Armada, Direktorat Jenderal Pengawasan Sumber Daya Kelautan dan Perikanan, Kementerian Kelautan dan Perikanan Republik Indonesia. Tahapan awal yang dilakukan pada penelitian ini adalah akusisi data, pembersihan data dan seleksi fitur. Dari tahapan preprocessing ini akan didapatkan data awal yang akan diproses pada tahapan selanjutnya.

4.1.1 Akusisi Data

Data *vessel monitoring system* (*VMS)* yang diperoleh berformat .csv dengan data awal terdiri dari 7.751.439 posisi data VMS, dengan jumlah transmitter aktif 2.130 unit, dan terdaftar menggunakan lebih dari 16 Jenis alat tangkap ikan, seperti yang terlihat pada Gambar 4.1.

```
ID_KAPAL;TRANSMITTER_NO;NAMA_ALAT_TANGKAP;NAMA_DPI;LONGITUDE;LATI
TUDE;"TO_CHAR(TR.PING_TIME,'DD-MM-YYYYHH24:MI:SS')";SPEED;HEADING
"16154";"103019";Purse Seine (Pukat Cincin) Pelagis Kecil;"L.
Cina Selatan, L. Natuna";103.374;1.199;08-02-2016 15:07:00;80;120
"16154";"103019";Purse Seine (Pukat Cincin) Pelagis Kecil;"L.
Cina Selatan, L. Natuna";103.455;1.143;08-02-2016 15:54:00;70;117
"16154";"103019";Purse Seine (Pukat Cincin) Pelagis Kecil;"L.
Cina Selatan, L. Natuna";103.465;1.138;08-02-2016 16:00:00;60;119
"16154";"103019";Purse Seine (Pukat Cincin) Pelagis Kecil;"L.
Cina Selatan, L. Natuna";103.515;1.098;08-02-2016 17:00:00;0;195
"16154";"103019";Purse Seine (Pukat Cincin) Pelagis Kecil;"L.
Cina Selatan, L. Natuna";103.515;1.098;08-02-2016 17:00:00;0;195
"16154";"103019";Purse Seine (Pukat Cincin) Pelagis Kecil;"L.
Cina Selatan, L. Natuna";103.515;1.098;08-02-2016 17:00:00;0;195
"16154";"103019";Purse Seine (Pukat Cincin) Pelagis Kecil;"L.
Cina Selatan, L. Natuna";103.515;1.098;08-02-2016 17:00:00;0;195
"16154";"103019";Purse Seine (Pukat Cincin) Pelagis Kecil;"L.
Cina Selatan, L. Natuna";103.515;1.098;08-02-2016 17:00:00;0;195
"16154";"103019";Purse Seine (Pukat Cincin) Pelagis Kecil;"L.
Cina Selatan, L. Natuna";103.579;1.031;08-02-2016 18:00:00;80;134
"16154";"103019";Purse Seine (Pukat Cincin) Pelagis Kecil;"L.
Cina Selatan, L. Natuna";103.674;0.94;08-02-2016 19:00:00;80;127
```

Gambar 4. 1 Tampilan Data Awal

Data awal yang diperoleh masih dalam format *Comma Separated Values* (.csv), untuk keperluan proses di tahap selanjutnya maka data di transformasi dalam bentuk kolom. Perangkat lunak yang digunakan dalam penelitian ini adalah Rstudio dengan menggunakan bahasa pemograman R. Data awal dipisah menjadi 9 kolom yang setiap atributnya diberi penamaan sesuai dengan data aslinya yaitu: ID_KAPAL, TRASMITER_NO, NAMA_ALAT_TANGKAP, NAMA_DPI, LONGITUDE, LATITUDE, TO_CHAR(TR.PING_TIME,'DD-MM-YYYYHH24:MI:SS'), SPEED dan HEADING. Setelah dilakukan pemecahan (*split*) data menjadi format kolom maka data terdiri dari 7.751.439 *row* dengan 9 variabel seperti yang terlihat pada Gambar 4.2.

	ID_KAPAL	TRANSMITTER_NO	NAMA_ALAT_TANGKAP	NAMA_DPI	LONGITUDE	LATITUDE	TO_CHAR(TR.PING_TIME,'DD-MM-YYYYHH24:MI:SS')	SPEED	HEADING
1	16154	103019	Purse Seine (Pukat Cincin) Pelagis Kecil	L. Cina Selatan, L. Natuna	103.374	1.199	08-02-2016 15:07:00	80	120
2	16154	103019	Purse Seine (Pukat Cincin) Pelagis Kecil	L. Cina Selatan, L. Natuna	103.455	1.143	08-02-2016 15:54:00	70	117
3	16154	103019	Purse Seine (Pukat Cincin) Pelagis Kecil	L. Cina Selatan, L. Natuna	103.465	1.138	08-02-2016 16:00:00	60	119
4	16154	103019	Purse Seine (Pukat Cincin) Pelagis Kecil	L. Cina Selatan, L. Natuna	103.515	1.098	08-02-2016 17:00:00	0	195
5	16154	103019	Purse Seine (Pukat Cincin) Pelagis Kecil	L. Cina Selatan, L. Natuna	103.515	1.098	08-02-2016 17:00:00	0	195
6	16154	103019	Purse Seine (Pukat Cincin) Pelagis Kecil	L. Cina Selatan, L. Natuna	103.515	1.098	08-02-2016 17:00:00	0	195
7	16154	103019	Purse Seine (Pukat Cincin) Pelagis Kecil	L. Cina Selatan, L. Natuna	103.515	1.098	08-02-2016 17:00:00	0	195
8	16154	103019	Purse Seine (Pukat Cincin) Pelagis Kecil	L. Cina Selatan, L. Natuna	103.515	1.098	08-02-2016 17:00:00	0	195
9	16154	103019	Purse Seine (Pukat Cincin) Pelagis Kecil	L. Cina Selatan, L. Natuna	103.579	1.031	08-02-2016 18:00:00	80	134
10	16154	103019	Purse Seine (Pukat Cincin) Pelagis Kecil	L. Cina Selatan, L. Natuna	103.674	0.940	08-02-2016 19:00:00	80	127
11	16154	103019	Purse Seine (Pukat Cincin) Pelagis Kecil	L. Cina Selatan, L. Natuna	103.705	0.932	08-02-2016 19:17:00	60	90
12	16154	103019	Purse Seine (Pukat Cincin) Pelagis Kecil	L. Cina Selatan, L. Natuna	103.708	0.933	08-02-2016 19:28:00	0	146
13	16154	103019	Purse Seine (Pukat Cincin) Pelagis Kecil	L. Cina Selatan, L. Natuna	103.708	0.933	08-02-2016 20:00:00	0	146
14	16154	103019	Purse Seine (Pukat Cincin) Pelagis Kecil	L. Cina Selatan, L. Natuna	103.708	0.933	08-02-2016 20:00:00	0	146
15	16154	103019	Purse Seine (Pukat Cincin) Pelagis Kecil	L. Cina Selatan, L. Natuna	103.708	0.933	08-02-2016 20:00:00	0	146
16	16154	103019	Purse Seine (Pukat Cincin) Pelagis Kecil	L. Cina Selatan, L. Natuna	103.708	0.933	08-02-2016 20:00:00	0	146
17	16154	103019	Purse Seine (Pukat Cincin) Pelagis Kecil	L. Cina Selatan, L. Natuna	103.708	0.933	08-02-2016 21:00:00	0	146
18	16154	103019	Purse Seine (Pukat Cincin) Pelagis Kecil	L. Cina Selatan, L. Natuna	103.706	0.934	08-02-2016 22:00:00	0	147

Gambar 4. 2 Fitur Data di Tahap Awal

4.1.2 Pembersihan Data (*Data Cleansing*)

Pada tahap ini, data yang telah diakusisi sebelumnya dibersihkan (*cleansing*). Proses pembersihan dilakukan karena adanya data duplikasi didalam data, duplikasi yang dimaksud adalah data memiliki kesamaan nilai disemua variabel data. Pada tahap pengecekan diketahui ada 721.788 data duplikasi, setelah dilakukan proses penghapusan diperoleh data baru yang berjumlah 7.029.661 baris data VMS dengan 9 variabel, jumlah data duplikasi dapat dilihat pada Tabel 4.1.

Tabel 4. 1 Jumlah Data Duplikasi

No.	Keterangan	Jumlah Row	Jumlah Variabel
1	Data Awal	7,751,439	9
2	Data Duplikasi	721,778	9
3	Data Baru	7,029,661	9

Setelah dilakukan proses pembersihan dari data duplikasi maka diperoleh data baru yang terdiri dari 7.029.331 row data dengan 9 variabel seperti yang terlihat pada Gambar 4.3.

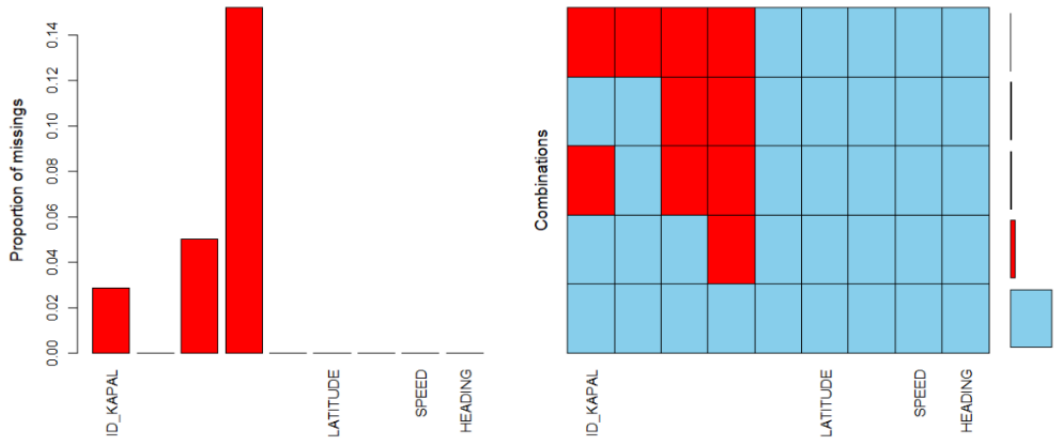
ID_KAPAL <dbl>	TRANSMITTER_NO <dbl>	NAMA_ALAT_TANGKAP <chr>	NAMA_DPI <chr>	LONGITUDE <dbl>	LATITUDE <dbl>
16154	103019	Purse Seine (Pukat Cincin) Pelagis Kecil	L. Cina Selatan, L. Natuna	103.374	1.199
16154	103019	Purse Seine (Pukat Cincin) Pelagis Kecil	L. Cina Selatan, L. Natuna	103.455	1.143
16154	103019	Purse Seine (Pukat Cincin) Pelagis Kecil	L. Cina Selatan, L. Natuna	103.465	1.138
16154	103019	Purse Seine (Pukat Cincin) Pelagis Kecil	L. Cina Selatan, L. Natuna	103.515	1.098
16154	103019	Purse Seine (Pukat Cincin) Pelagis Kecil	L. Cina Selatan, L. Natuna	103.579	1.031
16154	103019	Purse Seine (Pukat Cincin) Pelagis Kecil	L. Cina Selatan, L. Natuna	103.674	0.940
16154	103019	Purse Seine (Pukat Cincin) Pelagis Kecil	L. Cina Selatan, L. Natuna	103.705	0.932
16154	103019	Purse Seine (Pukat Cincin) Pelagis Kecil	L. Cina Selatan, L. Natuna	103.708	0.933
16154	103019	Purse Seine (Pukat Cincin) Pelagis Kecil	L. Cina Selatan, L. Natuna	103.708	0.933
16154	103019	Purse Seine (Pukat Cincin) Pelagis Kecil	L. Cina Selatan, L. Natuna	103.708	0.933
16154	103019	Purse Seine (Pukat Cincin) Pelagis Kecil	L. Cina Selatan, L. Natuna	103.706	0.934
16154	103019	Purse Seine (Pukat Cincin) Pelagis Kecil	L. Cina Selatan, L. Natuna	103.707	0.934
16154	103019	Purse Seine (Pukat Cincin) Pelagis Kecil	L. Cina Selatan, L. Natuna	103.707	0.934
16154	103019	Purse Seine (Pukat Cincin) Pelagis Kecil	L. Cina Selatan, L. Natuna	103.706	0.934
16154	103019	Purse Seine (Pukat Cincin) Pelagis Kecil	L. Cina Selatan, L. Natuna	103.752	0.922
16154	103019	Purse Seine (Pukat Cincin) Pelagis Kecil	L. Cina Selatan, L. Natuna	103.761	0.921
16154	103019	Purse Seine (Pukat Cincin) Pelagis Kecil	L. Cina Selatan, L. Natuna	103.775	0.919
16154	103019	Purse Seine (Pukat Cincin) Pelagis Kecil	L. Cina Selatan, L. Natuna	103.799	0.913
16154	103019	Purse Seine (Pukat Cincin) Pelagis Kecil	L. Cina Selatan, L. Natuna	103.857	0.894
16154	103019	Purse Seine (Pukat Cincin) Pelagis Kecil	L. Cina Selatan, L. Natuna	103.873	0.891
16154	103019	Purse Seine (Pukat Cincin) Pelagis Kecil	L. Cina Selatan, L. Natuna	103.951	0.884
16154	103019	Purse Seine (Pukat Cincin) Pelagis Kecil	L. Cina Selatan, L. Natuna	103.951	0.884

1-22 of 7,029,661 rows | 1-6 of 9 columns

Previous 1 2 3 4 5 6 ... 46 Next

Gambar 4. 3 Data Setelah Dibersihkan dari Duplikasi

Proses selanjutnya melakukan pengecekan *missing data* atau data yang kosong dari seluruh variabel. Hasil pengecekan data dapat dilihat pada Gambar 4.4.



Gambar 4. 4 *Missing Data*

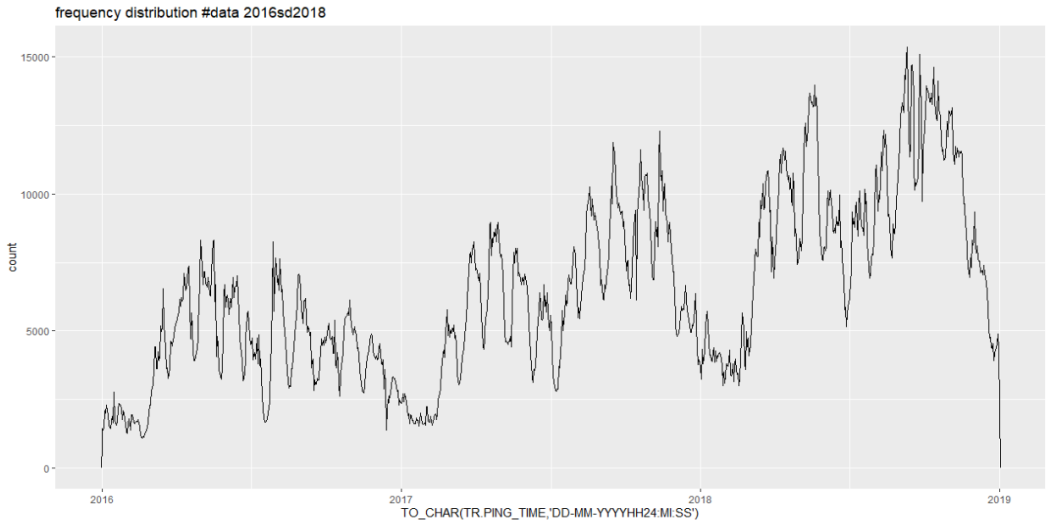
Dari Gambar 4.4 menunjukkan jumlah proporsi data yang hilang terletak pada variabel ID_KAPAL sebanyak 202.655 baris data, variabel TRASMITER_No terdapat 34 baris data yang kosong, variabel NAMA_ALAT_TANGKAP terdapat 331.003 baris data yang kosong, dan variabel NAMA_DPI memiliki paling banyak jumlah data yang kosong, sedangkan pada variabel LONGITUDE, LATITUDE, TO_CHAR(TR.PING_TIME,'DD-MM-YYYYHH24:MI:SS'), SPEED, dan HEADING tidak terdapat data yang kosong. Dalam penelitian ini variabel data yang dipergunakan sebagai dasar proses *clustering* secara spasial adalah variabel LONGITUDE dan LATITUDE, sedangkan variabel yang dipergunakan sebagai dasar untuk analisis secara temporal adalah variabel TO_CHAR(TR.PING_TIME,'DD-MM-YYYYHH24:MI:SS'), dan variabel lainnya merupakan informasi pendukung dalam penelitian ini.

4.1.3 Konversi Data

Tahap preprocessing selanjutnya adalah melakukan konversi data, data spasial dari distribusi kapal perikanan di data VMS masih berupa angka koordinat Longitude dan Latitude. Untuk mendapatkan gambaran awal tentang sebaran data VMS secara spasial maka data dikonversi ke dalam bentuk point yang ditampilkan di dalam peta statis berdasarkan titik koordinat yang sebenarnya. Hal ini dilakukan agar lebih mudah melakukan pengecekan awal secara visual, apakah ada titik koordinat yang berada diluar area WPPNRI-711. Hasilnya dapat dilihat pada Gambar 4.5 berikut ini:

Gambar 4. 5 Posisi Sebaran Data VMS Berdasarkan Titik Koordinat Longitude dan Latitude

Sedangkan untuk mendapatkan gambaran awal tentang sebaran data VMS secara temporal maka data dikonversi kedalam bentuk grafik frekuensi. Frekuensi adalah jumlah banyaknya aktifitas yang dilakukan atau terekam pada satu periode waktu tertentu. Variabel yang digunakan untuk menghasilkan grafik frekuensi temporal adalah TO_CHAR(TR.PING_TIME,'DD-MM-YYYYHH24:MI:SS'), yang berisi tanggal perekaman setiap data *vessel monitoring system (VMS)*. Sedangkan nilai *count* diperoleh dari penjumlahan seluruh posisi data VMS berdasarkan tanggal aktifitasnya, hasilnya dapat dilihat pada Gambar 4.6 berikut:



Gambar 4. 6 Frequensi Distribusi Data VMS Tahun 2016 s/d 2018

4.1.4 Seleksi Fitur

Untuk keperluan proses penelitian di tahap selanjutnya, data yang telah di bersihkan perlu dilakukan penambahan fitur baru. Penambahan fitur baru ini diperlukan untuk proses visualisasi dan analisis sebaran data VMS secara temporal. Fitur baru ini diperoleh dari pemecahan data di variabel TO_CHAR(TR.PING_TIME,'DD-MM-YYYYHH24:MI:SS'). Fitur baru yang dibuat adalah Year; Month; Day, Weekday dan Date, hasilnya seperti yang terlihat pada gambar 4.7.

ID_KAPAL	TRANSMITTER_NO	NAMA_ALAT_TANGKAP	NAMA_DPI	LONGITUDE	LATITUDE
16154	103019	Purse Seine (Pukat Cincin) Pelagis Kecil	L. Cina Selatan, L. Natuna	103.374	1.199
16154	103019	Purse Seine (Pukat Cincin) Pelagis Kecil	L. Cina Selatan, L. Natuna	103.455	1.143
16154	103019	Purse Seine (Pukat Cincin) Pelagis Kecil	L. Cina Selatan, L. Natuna	103.465	1.138
16154	103019	Purse Seine (Pukat Cincin) Pelagis Kecil	L. Cina Selatan, L. Natuna	103.515	1.098
16154	103019	Purse Seine (Pukat Cincin) Pelagis Kecil	L. Cina Selatan, L. Natuna	103.579	1.031
16154	103019	Purse Seine (Pukat Cincin) Pelagis Kecil	L. Cina Selatan, L. Natuna	103.674	0.940
16154	103019	Purse Seine (Pukat Cincin) Pelagis Kecil	L. Cina Selatan, L. Natuna	103.705	0.932
16154	103019	Purse Seine (Pukat Cincin) Pelagis Kecil	L. Cina Selatan, L. Natuna	103.708	0.933
16154	103019	Purse Seine (Pukat Cincin) Pelagis Kecil	L. Cina Selatan, L. Natuna	103.708	0.933
16154	103019	Purse Seine (Pukat Cincin) Pelagis Kecil	L. Cina Selatan, L. Natuna	103.708	0.933
16154	103019	Purse Seine (Pukat Cincin) Pelagis Kecil	L. Cina Selatan, L. Natuna	103.706	0.934
16154	103019	Purse Seine (Pukat Cincin) Pelagis Kecil	L. Cina Selatan, L. Natuna	103.707	0.934
16154	103019	Purse Seine (Pukat Cincin) Pelagis Kecil	L. Cina Selatan, L. Natuna	103.707	0.934
16154	103019	Purse Seine (Pukat Cincin) Pelagis Kecil	L. Cina Selatan, L. Natuna	103.706	0.934

1-14 of 7,029,661 rows | 1-6 of 15 columns

Previous 1 2 3 4 5 6 ... 72 Next

TO_CHAR(TR.PING_TIME,'DD-MM-YYYYHH24:MI:SS')	SPEED	HEADING	Year	Month	Day	Weekday	date
2016-02-08 15:07:00	80	120	2016	2	8	2	2016-02-08
2016-02-08 15:54:00	70	117	2016	2	8	2	2016-02-08
2016-02-08 16:00:00	60	119	2016	2	8	2	2016-02-08
2016-02-08 17:00:00	0	195	2016	2	8	2	2016-02-08
2016-02-08 18:00:00	80	134	2016	2	8	2	2016-02-08
2016-02-08 19:00:00	80	127	2016	2	8	2	2016-02-08
2016-02-08 19:17:00	60	90	2016	2	8	2	2016-02-08
2016-02-08 19:28:00	0	146	2016	2	8	2	2016-02-08
2016-02-08 20:00:00	0	146	2016	2	8	2	2016-02-08
2016-02-08 21:00:00	0	146	2016	2	8	2	2016-02-08
2016-02-08 22:00:00	0	147	2016	2	8	2	2016-02-08
2016-02-08 22:08:00	0	147	2016	2	8	2	2016-02-08
2016-02-08 22:36:00	0	147	2016	2	8	2	2016-02-08
2016-02-08 23:00:00	20	266	2016	2	8	2	2016-02-08

Previous 1 2 3 4 5 6 ...

Gambar 4. 7 Fitur-fitur yang Dipergunakan Pada Proses Selanjutnya

Untuk mendapatkan gambaran secara umum (*summary*) tentang data yang akan dipergunakan dalam penelitian ini, dapat dilihat pada Gambar 4.8.

```
summary(df2016sd2018)
```

```
   ID_KAPAL          TRANSMITTER_NO      NAMA_ALAT_TANGKAP    NAMA_DPI           LONGITUDE
Min.   :2.610e+02   Min.   :3.780e+04   Length:7029661     Length:7029661     Min.   :103.2
1st Qu.:2.168e+04   1st Qu.:4.475e+06   Class :character   Class :character   1st Qu.:106.0
Median :8.327e+04   Median :4.593e+06   Mode  :character   Mode  :character   Median :107.4
Mean   :1.155e+07   Mean   :1.641e+07                                         Mean   :107.1
3rd Qu.:8.739e+04   3rd Qu.:3.121e+07                                         3rd Qu.:108.5
Max.   :1.235e+09   Max.   :8.681e+10                                         Max.   :110.6
NA's   :202655      NA's   :34
   LATITUDE        TO_CHAR(TR.PING_TIME,'DD-MM-YYYYHH24:MI:SS')     SPEED             HEADING
Min.   :-4.0881   Min.   :2016-01-01 00:00:00                   Min.   :   0.00   Min.   :  0.0
1st Qu.:-0.6545   1st Qu.:2017-03-17 10:52:00                   1st Qu.:   0.00   1st Qu.: 60.0
Median : 1.0375   Median :2017-11-16 19:00:00                   Median :   0.50   Median :200.0
Mean   : 1.0673   Mean   :2017-10-15 22:04:06                   Mean   :  20.24   Mean   :194.4
3rd Qu.: 2.8442   3rd Qu.:2018-07-12 07:48:00                   3rd Qu.:   9.00   3rd Qu.:335.0
Max.   : 7.7380   Max.   :2018-12-31 23:59:00                   Max.   :1619.00   Max.   :511.0

      Year          Month             Day            weekday        date
Min.   :2016   Min.   : 1.000   Min.   : 1.00   Min.   :1   Min.   :2016-01-01
1st Qu.:2017   1st Qu.: 5.000   1st Qu.: 8.00   1st Qu.:2   1st Qu.:2017-03-17
Median :2017   Median : 7.000   Median :16.00   Median :4   Median :2017-11-16
Mean   :2017   Mean   : 7.111   Mean   :15.79   Mean   :4   Mean   :2017-10-15
3rd Qu.:2018   3rd Qu.:10.000   3rd Qu.:23.00   3rd Qu.:6   3rd Qu.:2018-07-12
Max.   :2018   Max.   :12.000   Max.   :31.00   Max.   :7   Max.   :2018-12-31
```

Gambar 4. 8 *Summary Data*

Tahap selanjutnya dalam proses seleksi fitur adalah memilih fitur yang akan digunakan dalam proses *clustering*. Seleksi fitur berguna untuk mengurangi dimensi data, sehingga meningkatkan efektifitas dan efisiensi kinerja dari algoritma *K-Means Clustering*. Pada tahap ini dilakukan pemilihan fitur pada variable LATITUDE dan LONGITUDE sebagai data spasial yang selanjutnya akan

dipergunakan untuk proses *clustering*, hasilnya seperti yang terlihat pada Gambar 4.9.

LONGITUDE <dbl>	LATITUDE <dbl>
103.374	1.199
103.455	1.143
103.465	1.138
103.515	1.098
103.579	1.031
103.674	0.940
103.705	0.932
103.708	0.933
103.708	0.933
103.708	0.933
103.706	0.934
103.707	0.934
103.707	0.934
103.706	0.934
103.752	0.922
103.761	0.921
103.775	0.919
103.799	0.913
103.857	0.894
103.873	0.891
103.951	0.884
103.951	0.884
103.958	0.886
104.019	0.933
104.032	0.955
104.106	0.985
104.135	1.037

1-27 of 7,029,661 rows

Gambar 4. 9 Data Longitude dan Latitude

4.2 Implementasi *Clustering*

Pada bagian ini akan menjelaskan hasil dan pembahasan dari proses pencarian nilai k optimum dengan *Elbow Method* dan *clustering* dengan algoritma *K-Means*. Bagian ini juga berisi visualisasi hasil *clustering* dan uji performa *clustering* dengan mengunakan *Silhouette Coefficient*.

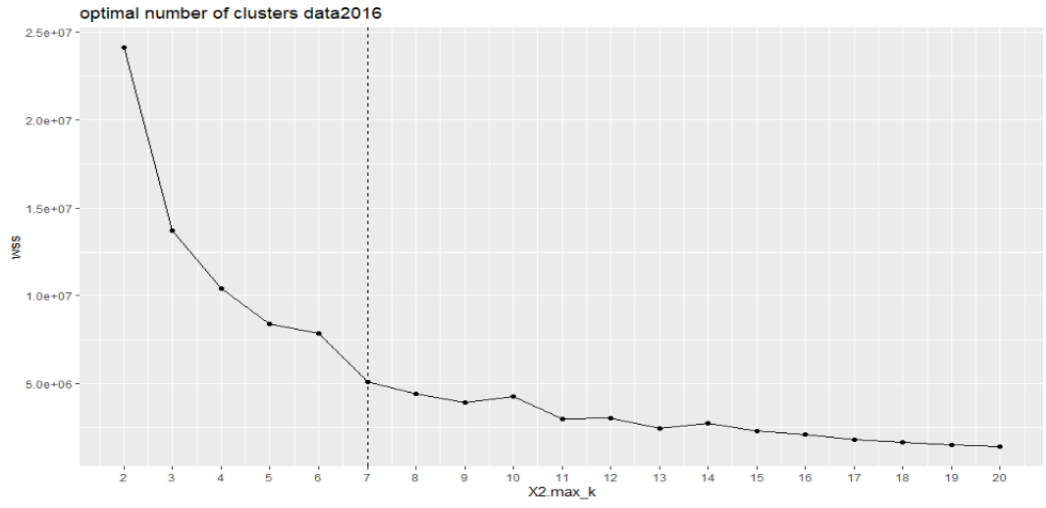
4.2.1 Mencari Jumlah *Cluster* dengan *Elbow Method*

Metode *Elbow* mengandalkan penglihatan terhadap grafik, dimulai dengan memperhatikan k = 2 dan seterusnya, pada nilai k turun secara drastis dan setelah

penurunannya grafik lebih stabil maka disinilah nilai k yang akan digunakan [55]. Tetapi metode ini juga memiliki kelemahan dan permasalah dalam penentuan nilai k optimum, yaitu metode *elbow* tidak selalu dapat mengidentifikasi titik "elbow" secara jelas, bahkan terkadang tidak ada titik siku yang terbentuk, atau justru ada beberapa pilihan titik siku di grafik [55]. Untuk mengatasi hal tersebut pada penelitian ini metode *elbow* diaplikasikan terhadap variabel LONGITUDE dan LATITUDE dari data VMS pada tahun 2016, tahun 2017, tahun 2018 dan gabungan data tahun 2016 s/d 2018. Plot dan hasil implementasi metode *Elbow* dari ke empat data set tersebut digunakan sebagai pembanding, sehingga nilai k optimum yang dihasilkan dapat mewakili keseluruhan data.

4.2.1.1 Berdasarkan Data Tahun 2016

Plot hasil implementasi metode *Elbow* untuk dataset tahun 2016 digambarkan seperti pada Gambar 4.10. Pada plot tersebut didapatkan titik siku yang terbentuk diantara titik enam dan delapan, setelah titik 7 sudah tidak lagi terjadi penurunan yang signifikan, sehingga dapat disimpulkan bahwa jumlah cluster menurut metode *Elbow* berdasarkan dataset tahun 2016 sejumlah 7 cluster.



Gambar 4. 10 Grafik *Elbow Method* Berdasarkan Data Tahun 2016

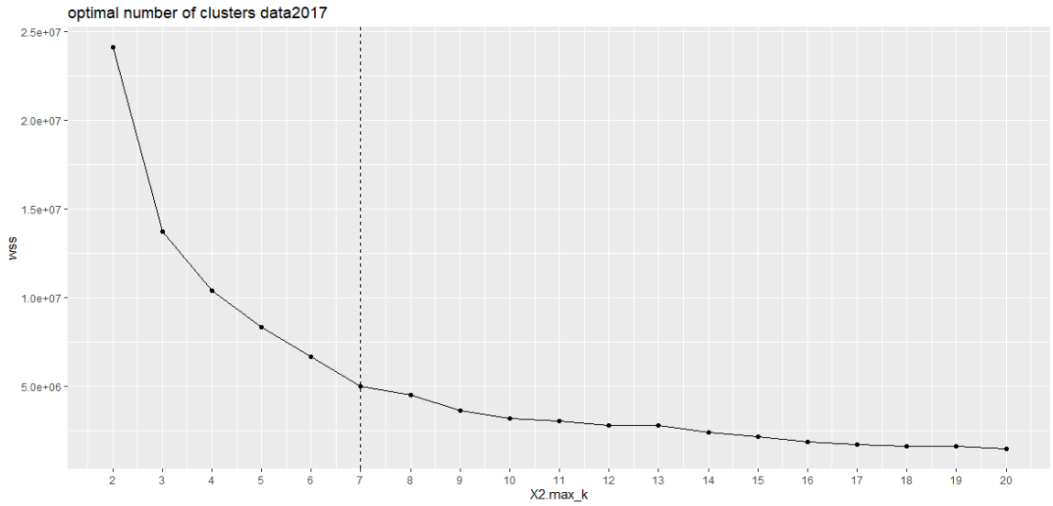
Tabel 4. 2 Hasil *Distance Cluster* Berdasarkan Data Tahun 2016

Cluster	Distance	Cluster	Distance
2	24,103,001.985	12	2,581,040.758
3	13,714,480.514	13	2,549,991.148
4	10,411,654.695	14	2,347,080.141
5	8,243,866.234	15	2,194,217.907
6	6,593,369.590	16	2,141,302.428
7	5,738,353.926	17	1,746,321.650
8	4,836,129.118	18	1,649,420.614
9	3,626,324.222	19	1,567,336.448
10	3,339,963.710	20	1,501,437.394
11	2,917,006.811		

Sebagai pembanding dapat diperhatikan nilai *distance cluster* pada Tabel 4.2, dapat di hitung bahwa selisih nilai *distance* k = 8 ke k = 9 sebesar 1,209,804.896 lebih besar dibandingkan selisih nilai *distance* k = 8 ke k = 7 sebesar 902,224.808 ataupun selisih k = 7 ke k = 6 sebesar 855,015.664. Sehingga dapat disimpulkan bahwa k = 7 atau k = 9 adalah jumlah *cluster* yang tepat untuk dipilih karena *cluster* selanjutnya menunjukkan penurunan yang stabil.

4.2.1.2 Berdasarkan Data Tahun 2017

Plot hasil implementasi metode *Elbow* untuk dataset tahun 2017 digambarkan seperti pada Gambar 4.11. Pada plot tersebut didapatkan titik siku yang terbentuk diantara titik enam dan delapan, setelah titik 7 sudah tidak lagi terjadi penurunan yang signifikan, sehingga dapat disimpulkan bahwa jumlah *cluster* menurut metode *Elbow* dengan berdasarkan data set tahun 2017 yaitu sejumlah 7 *cluster*.



Gambar 4. 11 Grafik *Elbow Method* Berdasarkan Data Tahun 2017

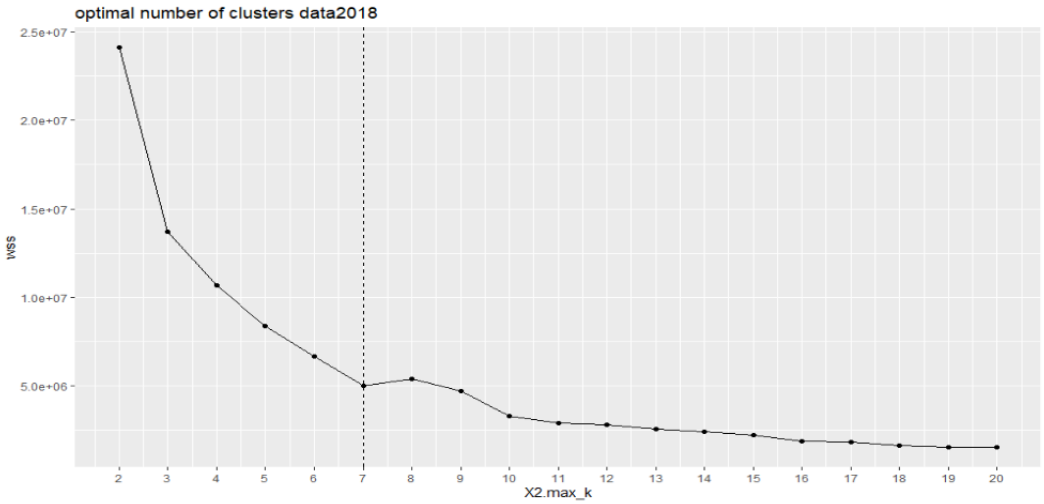
Tabel 4. 3 Hasil *Distance Cluster* Berdasarkan Data Tahun 2017

Cluster	Distance	Cluster	Distance
2	24,103,001.985	12	2,920,106.893
3	21,722,735.901	13	2,470,010.910
4	10,665,447.488	14	2,486,635.414
5	8,367,192.117	15	2,224,829.178
6	6,503,894.007	16	1,907,123.312
7	5,034,684.896	17	2,200,217.809
8	5,372,987.366	18	1,843,967.242
9	4,023,036.780	19	1,739,450.386
10	3,481,102.433	20	1,539,552.968
11	3,099,009.022		

Sebagai pembanding dapat diperhatikan nilai *distance cluster* pada Tabel 4.3, dapat di hitung bahwa selisih nilai *distance* k = 8 ke k = 9 sebesar 1,349,950.586 lebih besar dibandingkan selisih nilai *distance* k = 8 ke k = 7 sebesar -338,302.470, tetapi tidak lebih besar dari selisih k = 7 ke k = 6 sebesar 1,469,209.111. Hasil ini sesuai dengan grafik yang dihasilkan oleh metode *elbow* bahwa *cluster* 7 merupakan *cluster* optimum, karena *cluster* selanjutnya menunjukkan penurunan yang stabil.

4.2.1.3 Berdasarkan Data Tahun 2018

Plot hasil implementasi metode *Elbow* untuk dataset tahun 2018 digambarkan seperti pada Gambar 4.12. Pada plot tersebut didapatkan titik siku yang terbentuk diantara titik enam dan delapan, setelah titik 7 sudah tidak lagi terjadi penurunan yang signifikan, sehingga dapat disimpulkan bahwa jumlah *cluster* menurut metode *Elbow* dengan berdasarkan dataset tahun 2018 yaitu sejumlah 7 *cluster*.



Gambar 4. 12 Grafik *Elbow Method* Berdasarkan Data Tahun 2018

Tabel 4. 4 Hasil *Distance Cluster* Berdasarkan Data Tahun 2018

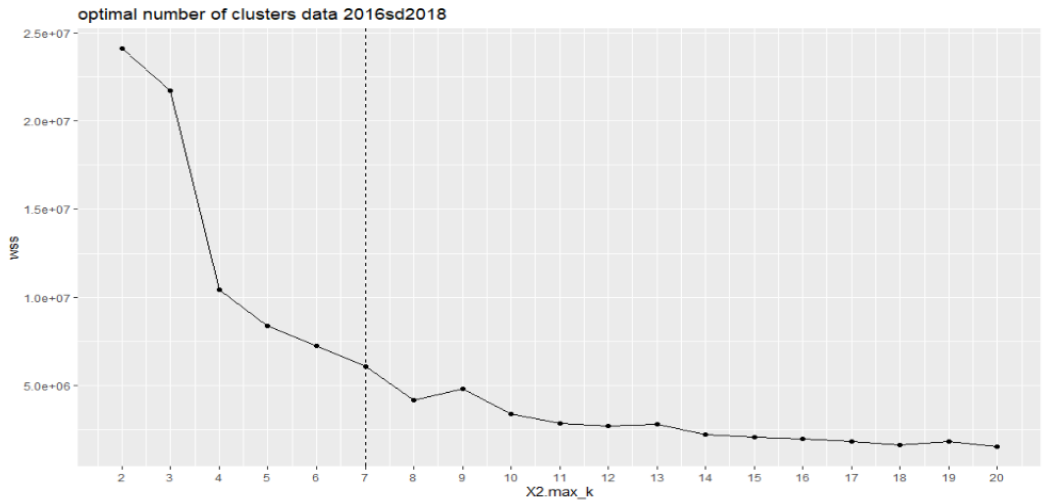
Cluster	Distance	Cluster	Distance
2	24,103,001.985	12	2,581,040.758
3	13,714,480.514	13	2,392,749.509
4	10,665,447.488	14	2,758,546.390
5	8,244,338.928	15	2,039,167.438
6	6,503,894.007	16	1,957,542.202
7	5,103,413.115	17	1,892,534.168
8	5,129,875.055	18	1,643,782.485
9	3,820,249.272	19	1,598,719.815
10	3,191,261.389	20	1,531,372.754
11	3,248,751.505		

Sebagai pembanding dapat diperhatikan nilai *distance cluster* pada Tabel 4.4, dapat di hitung bahwa selisih nilai *distance* k = 8 ke k = 9 sebesar 1,309,625.783

lebih besar dibandingkan selisih nilai *distance* k = 8 ke k = 7 sebesar -26,461.940, tetapi tidak lebih besar dari selisih k = 7 ke k = 6 sebesar 1,400,480.892. Hasil ini sesuai dengan grafik yang dihasilkan oleh metode *elbow* bahwa *cluster* 7 merupakan *cluster* optimum, karena *cluster* selanjutnya menunjukkan penurunan yang stabil.

4.2.1.4 Berdasarkan Data Tahun 2016 sd 2018

Plot hasil implementasi metode *Elbow* untuk dataset tahun 2016 s/d 2018 digambarkan seperti pada Gambar 4.13. Pada plot tersebut didapatkan titik siku yang terbentuk di titik ke-7, sehingga dapat disimpulkan bahwa jumlah *cluster* menurut metode *Elbow* dengan berdasarkan dataset tahun 2016 s/d 2018 yaitu sejumlah 7 *cluster*.



Gambar 4. 13 Grafik *Elbow Method* Berdasarkan Data Tahun 2016 s/d 2018

Tabel 4. 5 Hasil *Distance Cluster* Berdasarkan Data Tahun 2016 s/d 2018

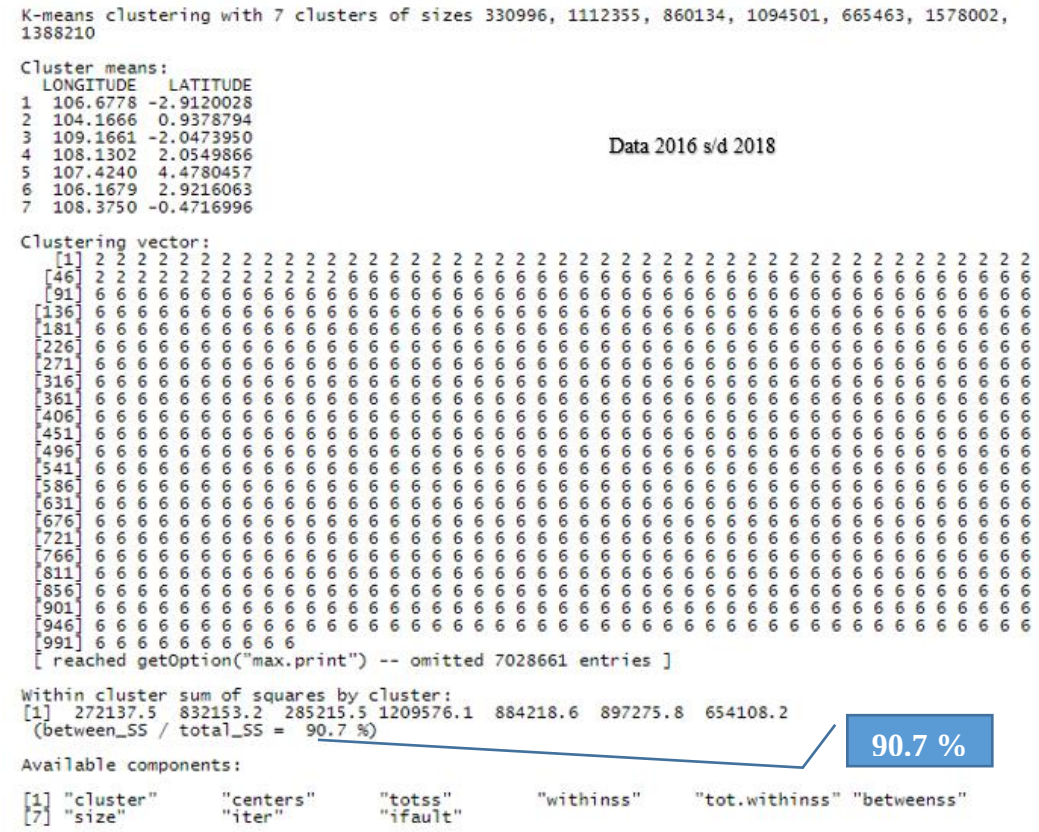
Cluster	Distance	Cluster	Distance
2	24,103,001.985	12	2,633,759.619
3	13,714,480.515	13	2,419,570.217
4	10,411,654.696	14	2,290,169.247
5	8,243,866.234	15	2,055,906.812
6	6,781,061.737	16	1,837,967.326
7	5,738,353.926	17	1,805,852.628
8	5,133,515.253	18	1,783,001.043
9	3,598,176.562	19	1,629,461.963
10	3,374,716.809	20	1,457,691.081
11	3,099,009.023		

Berdasarkan grafik pada Gambar 4.13 diatas dapat disimpulkan bahwa *cluster* 7 adalah *cluster* tepat untuk dipilih karena jumlah *cluster* selanjutnya menunjukkan penurunan yang stabil. Sebagai perbandingan perhatikan hasil penghitungan *Distance Cluster* pada Tabel 4.5, terlihat bahwa penurunan selisih nilai *distance* setelah *cluster* ke-7 ke *cluster* selanjutnya lebih stabil. Permasalah dengan metode *elbow* adalah tidak selalu dapat mengidentifikasi "*elbow"* secara jelas, bahkan terkadang tidak ada siku terbentuk, atau ada beberapa pilihan siku [55]. Seperti yang ditunjukkan pada Gambar 4.13 bahwa k = 9 justru nilainya lebih tinggi dari pada k = 8, atau pada Tabel 4.5 dapat di hitung bahwa selisih nilai *distance* k = 8 ke k = 9 sebesar 1,535,338.691 lebih besar dibandingkan selisih nilai *distance* k = 6 ke k = 7 sebesar 1,042,707.811. Sehingga dapat disimpulkan bahwa *cluster* 7 atau 9 adalah *cluster* yang tepat dipilih karena *cluster* selanjutnya menunjukkan penurunan yang stabil.

Dari keseluruhan plot dan analisis nilai *distance cluster* dari data set tahun 2016, tahun 2017, tahun 2018 dan gabungan dataset tahun 2016 s/d 2018 dapat disimpulkan bahwa nilai k optimum dipilih pada penelitian ini adalah *cluster* ke-7. Sehingga untuk proses penelitian selanjutnya jumlah yang digunakan adalah 7 *cluster*.

4.2.2 *Clustering* dengan *Alogaritma K-Means*

Setelah diperoleh jumlah *cluster* optimum atau nilai k = 7, tahap selanjutnya dilakukan proses *clustering* menggunakan algoritma *K-Means* untuk mengolah data dari variable LONGITUDE dan LATITUDE, hasilnya seperti yang terlihat pada Gambar 4.14.



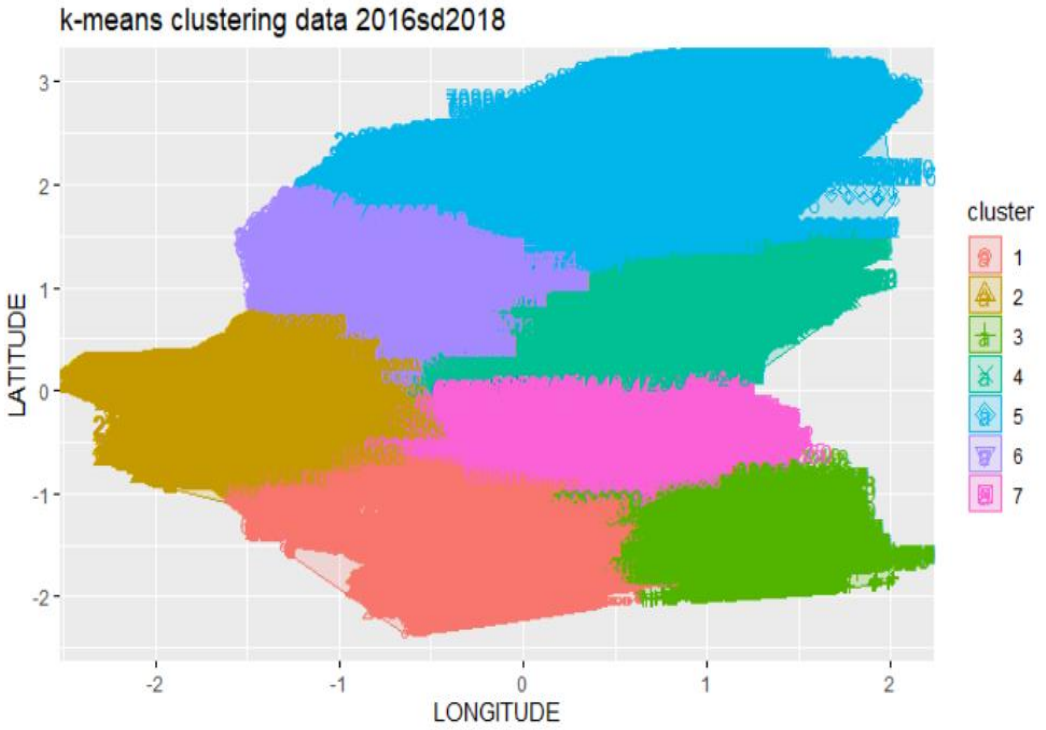
Gambar 4. 14 Hasil *Clustering* dengan *K-Means*

Fungsi *K-Means* untuk menampilkan hasil *clustering* seperti pada Gambar 4.14 digunakan untuk melihat *vektor centroid (cluster means)* serta jumlah anggota masing-masing *cluster*. *Clustering vektor* digunakan untuk melihat data 1 s/d 7.029.661 masuk kedalam anggota cluster yang mana. Selanjutnya didapatkan nilai *within cluster sum of squares* dari *cluster K-Means* dengan k = 7 yaitu, untuk *cluster* 1 sebesar 272137.5, *cluster* 2 sebesar 832153.2, *cluster* 3 sebesar 285215.5, *cluster* 4 sebesar 1209576.1, *cluster* 5 sebesar 884218.6, *cluster* 6 sebesar 897275.8 dan *cluster* 7 sebesar 654108.2. Dengan nilai *between sum of squares* dibagi nilai *total sum of squares* diperoleh hasil persentase 90.7%. Hasil ini menunjukkan

kekompakan hasil *clustering* atau tingkat kemiripan setiap anggota data dalam satu *cluster* yang sama.

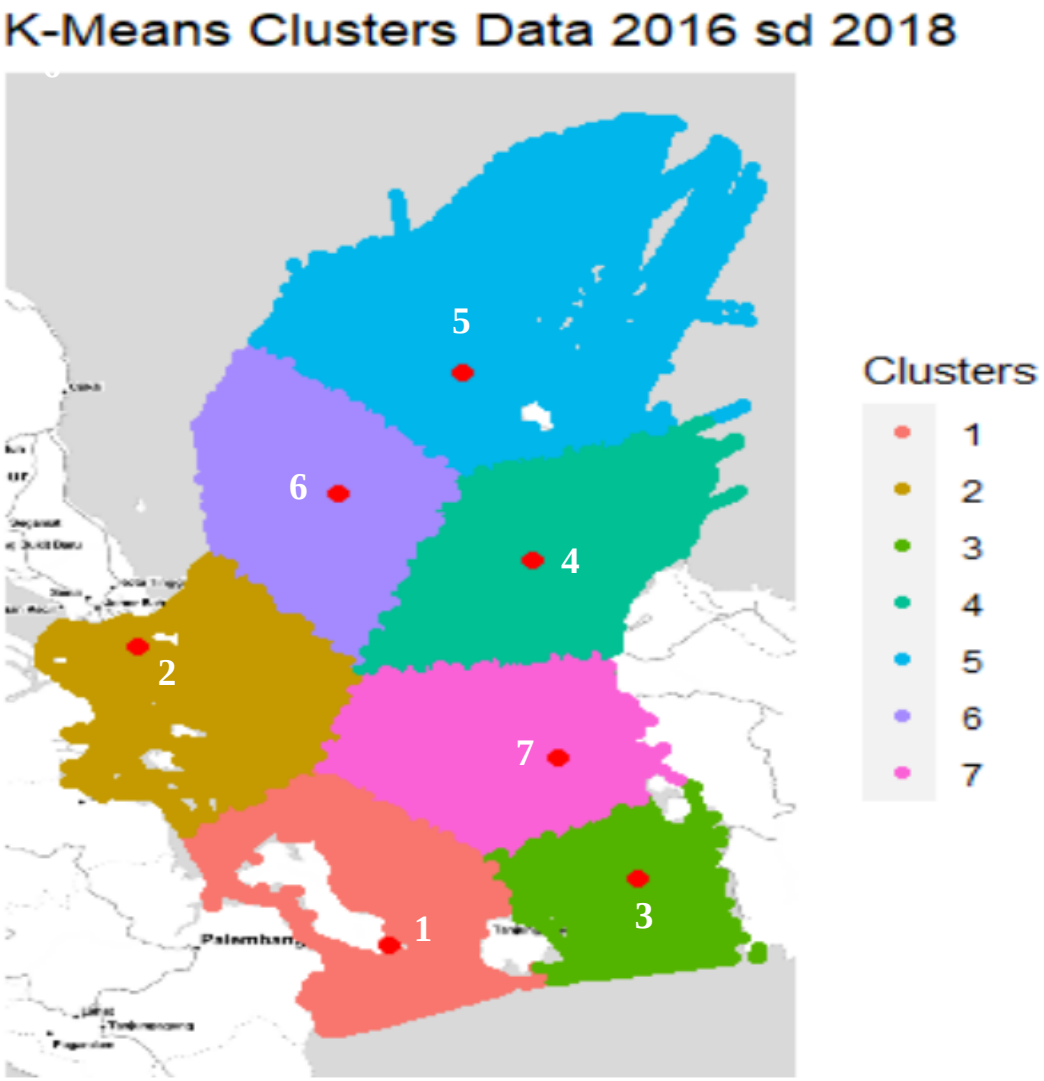
4.2.2.1 Visualisasi

Tahap ini dilakukan visualisasi hasil *clustering* dalam bentuk *cluster plot* agar lebih mudah untuk dipahami, hasilnya seperti yang terlihat pada Gambar 4.15.



Gambar 4. 15 Plot hasil *K-Means Clustering*

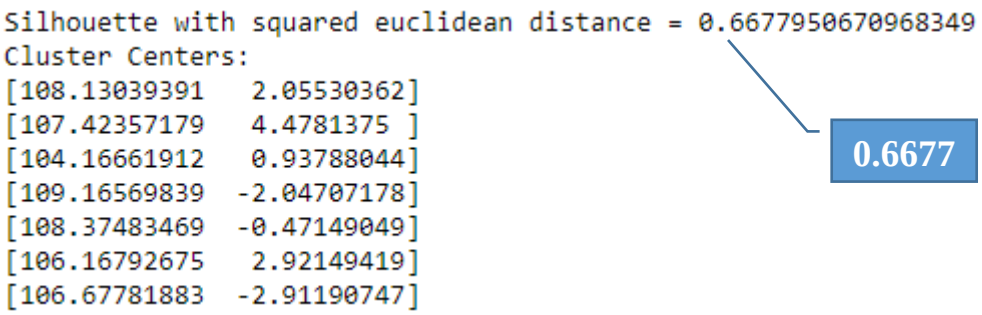
Visualisasi hasil *clustering* dengan mentranformasikan data kedalam bentuk maping, setiap data VMS akan diwakili oleh sebuah titik sesuai dengan lokasi sebenarnya berdasarkan koordinat Longitude dan Latitude. Setiap *cluster* akan dibedakan berdasarkan warna yang akan memberikan informasi abstrak dalam bentuk visual. Tujuannya untuk membantu dalam memahami, menggambarkan dan mengkomunikasikan informasi dan menganalisis data untuk mendapatkan sebuah pengetahuan baru. Visualisasi ini dapat digunakan untuk mengetahui persebaran data VMS secara spasial berdasarkan anggota masing-masing *cluster* dalam bentuk map statis, hasilnya terlihat pada Gambar 4.16.



Gambar 4. 16 Visualisasi Hasil *K-Means Clustering*

4.2.3 Uji *Validitas Clustering* dengan *Silhouette Coefficient.*

Setelah dilakukan implementasi proses *clustering*, maka selanjutnya adalah melakukan uji performa pada hasil *clustering* dengan menggunakan *Silhouette Coefficient*, dengan hasil terlihat pada Gambar 4.17.

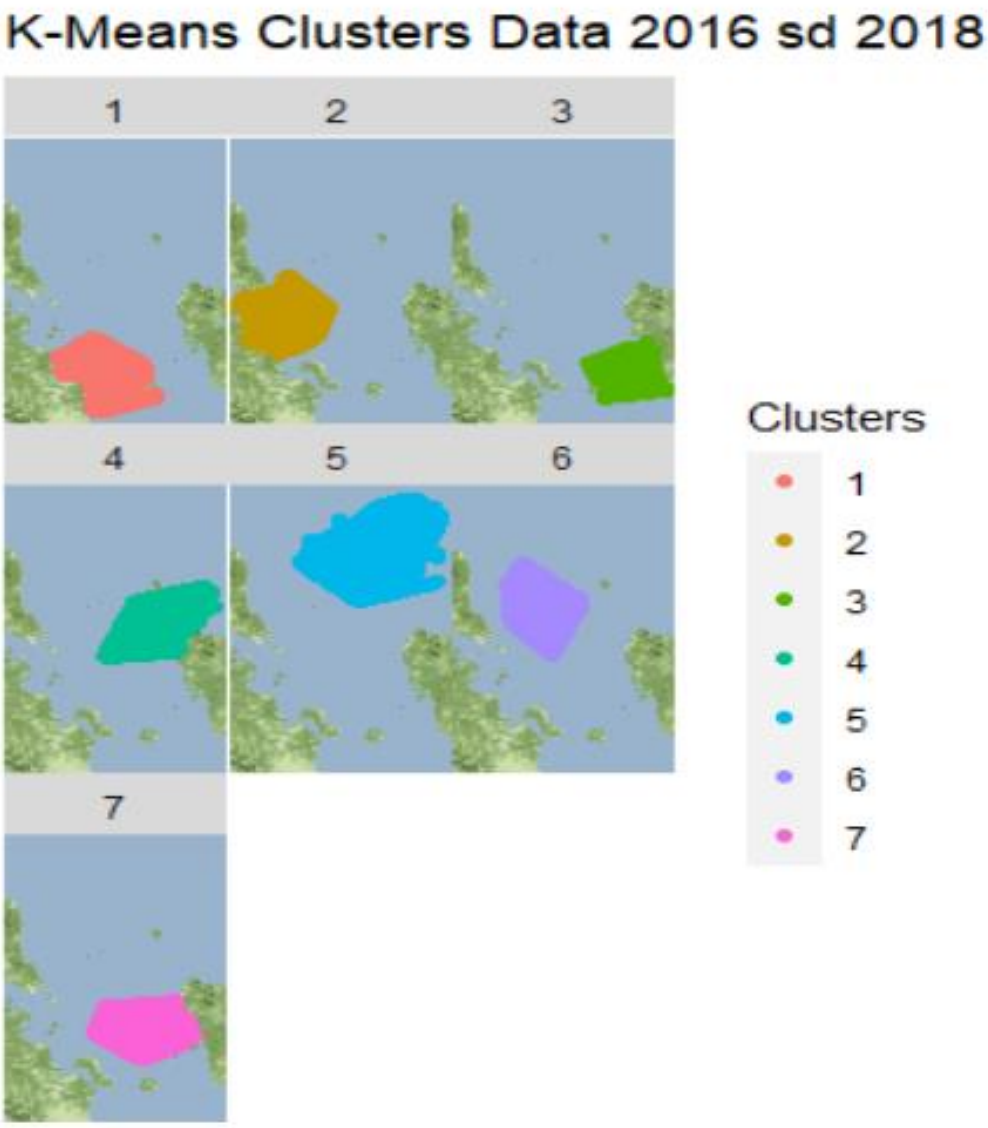


Gambar 4. 17 Hasil Uji *Silhouette Coefficient*

Hasil nilai *Silhouette Coefficient* adalah 0.6677 dan berdasarkan tabel nilai *Silhouette Coefficient* dari Kaufman dan Rousseeuw hasil ini termasuk kedalam *Reasonable /Medimum structure*. Sehingga dapat disimpulkan bahwa hasil *clustering* cukup baik dan dapat dilakukan analisa selanjutnya.

4.3 Analisa Hasil Clustering

Data VMS memberikan informasi posisi geografis dan waktu aktivitas kapal perikanan yang menjadi objek dalam penelitian ini selama periode tahun 2016 s/d 2018 tersebar secara luas di WPPNRI-711. Jutaan posisi geografis kapal perikanan tersebut menggambarkan aktivitas saat *Soak time/ idle, Hauling, Setting,* dan juga *Steaming*. Oleh karena itu, dibutuhkan pemodelan sederhana yang representatif agar data tersebut bisa dianalisis. Langkah pertama untuk memahami pola aktivitas kapal perikanan adalah dengan membagi WPPNRI-711 menjadi beberapa zona aktivitas kapal perikanan. Disini, *K-Means* membagi secara partisi area WPPNRI-711 menjadi 7 *cluster* berdasarkan data posisi geografis aktivitas kapal perikanan tersebut. Hasilnya dapat digunakan untuk menganalisis tingkat kepadatan usaha penangkapan ikan di masing-masing *cluster* dan secara temporal dapat diketahui tingkat intensitas kapal disetiap *cluster* berdasarkan periode waktu bulanan dan tahunan. Berikut gambaran visualisasi wilayah WPPNRI-711 yang dibagi menjadi 7 *cluster* dalam bentuk *Split Map* statis, hasilnya terlihat pada Gambar 4.18.



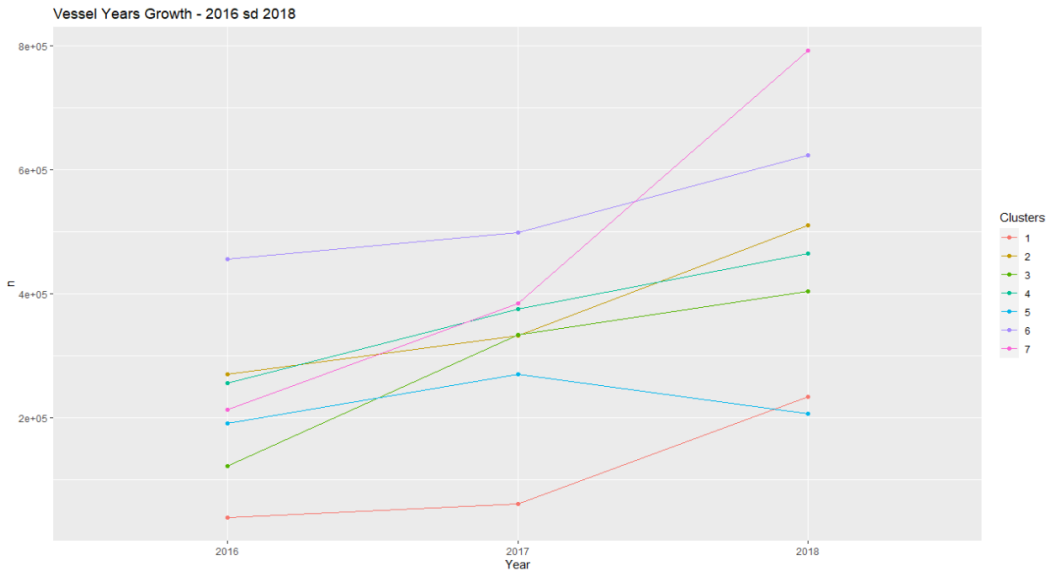
Gambar 4. 18 Visualisasi Hasil K-Means *Clustering* dengan Split Map

Setelah dilakukan pengelompokan, setiap *cluster* akan dianalisis berdasarkan data spasial dan temporalnya. Hal ini dilakukan untuk mempermudah dalam melakukan analisis hasil *cluster* tersebut. Untuk mengetahui *Cluster Means* dan jumlah anggota dari masih-masing *cluster* dapat melihat Tabel 4.6.

Tabel 4. 6 Data Hasil *K-Means Clustering*

Clusters	Clusters Means		Clusters Sizes
	Longitude	Latitude	
1	106.6778	-2.9120028	330,996
2	104.1666	0.9378794	1,112,355
3	109.1661	-2.0473950	860,134
4	108.1302	2.0549866	1,094,501
5	107.4240	4.4780457	665,463
6	106.1679	2.9216063	1,578,002
7	108.3750	-0.4716996	1,388,210

Pertumbuhan data VMS di setiap cluster berdasarkan data pertahun, akan ditampilkan dalam bentuk grafik pada Gambar 4.19 dan berbentuk data pada Tabel 4.7.



Gambar 4. 19 Distribusi Data per Cluster Periode Tahunan

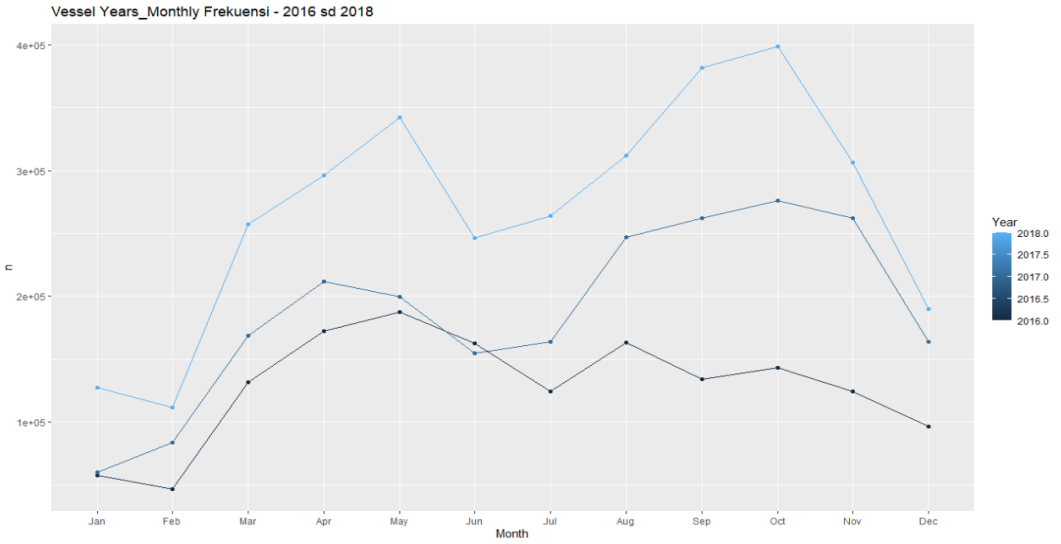
Dari Gambar 4.19 dapat diketahui bahwa distribusi data periode tahunan mengalami peningkatan di hampir semua cluster, kenaikan paling signifikan terjadi di cluster 7 dan penurunan hanya terjadi di cluster 5 pada tahun 2018 jika dibandingkan dengan tahun 2017.

Tabel 4. 7 Distribusi Data per Cluster Periode Tahunan

Clusters	Clusters Members per Years			Clusters Sizes
	2016	2017	2018	
1	38,286	60,050	232,660	330,996
2	269,260	332,452	510,643	1,112,355
3	122,084	333,871	404,179	860,134
4	255,699	374,414	464,388	1,094,501
5	190,145	269,918	205,400	665,463
6	456,134	498,634	623,234	1,578,002
7	212,613	383,474	792,123	1,388,210

Dari Tabel 4.7 dapat diambil kesimpulan bahwa *cluster* yang memiliki pertumbuhan terbanyak adalah *cluster* ke-7, karena jumlah sebaran data vms yang terekam ditahun 2016 berjumlah 212.613, tahun 2017 berjumlah 383.474 dan ditahun 2018 mengalami peningkatan yang signifikan dengan jumlah data 792.123. Jika dihitung secara presentasi maka pada cluster ke-7 dari tahun 2016 ke tahun 2018 mengalami kenaikan sebesar 272.56%. Sedangkan satu-satunya *cluster* yang mengalami penurunan jumlah data adalah *cluster* ke-5 karena jumlah data vms yang terdaftar beroperasi ditahun 2016 berjumlah 190.145, tahun 2017 berjumlah 269.918 dan ditahun 2018 mengalami penurunan dengan jumlah 205.400 Jika dihitung secara presentasi maka pada *cluster* ke-5 dari tahun 2017 ke tahun 2018 mengalami penurunan sebesar 23.90%.

Untuk kenaikan dan penurunan jumlah data periode bulanan di setiap tahunnya dapat dilihat dalam bentuk grafik pada Gambar 4.20 dan jumlah sebaran datanya dapat dilihat di Tabel 4.8. Sedangkan untuk melihat sebaran data secara keseluruhan berdasarkan distribusi bulanan dapat melihat Gambar 4.21.

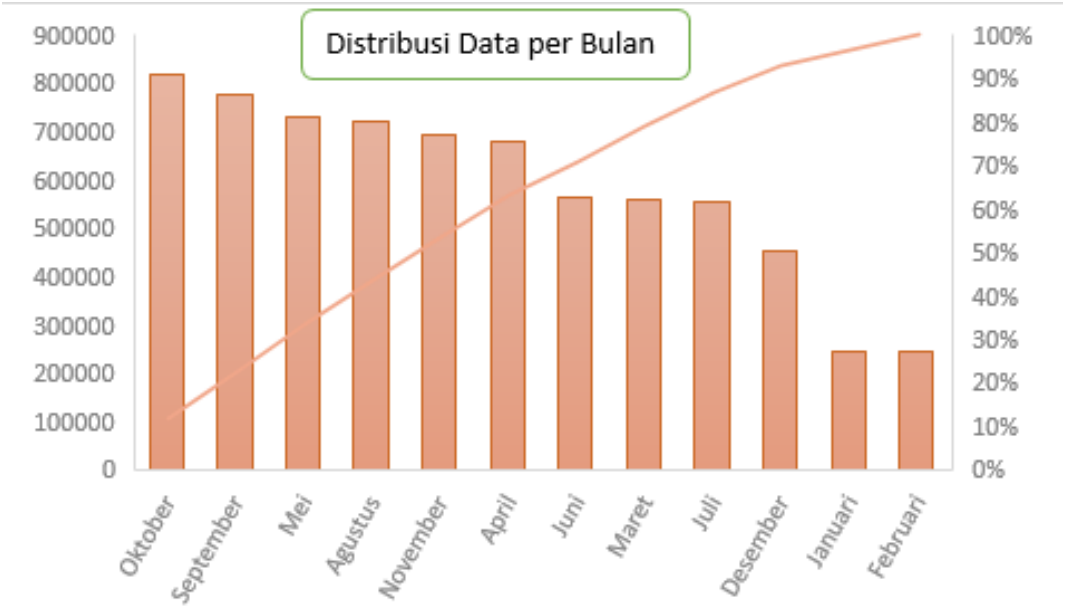


Gambar 4. 20 Distribusi Data Bulanan Pada Tahun 2016, 2017 dan 2018

Dari Gambar 4.20 dapat diketahui bahwa pola distribusi data bulanan pada tahun 2016, 2017 dan 2018 secara keseluruhan cenderung mirip, dimulai dengan bulan Januari dan Februari dengan jumlah data sedikit dan mulai meningkat dibulan Maret, April dan Mei kemudian ada sedikit penurunan di bulan Juni dan Juli berikutnya meningkat lagi dibulan Agustus, September, dan Oktober dan mulai menurun kembali di bulan November dan Desember.

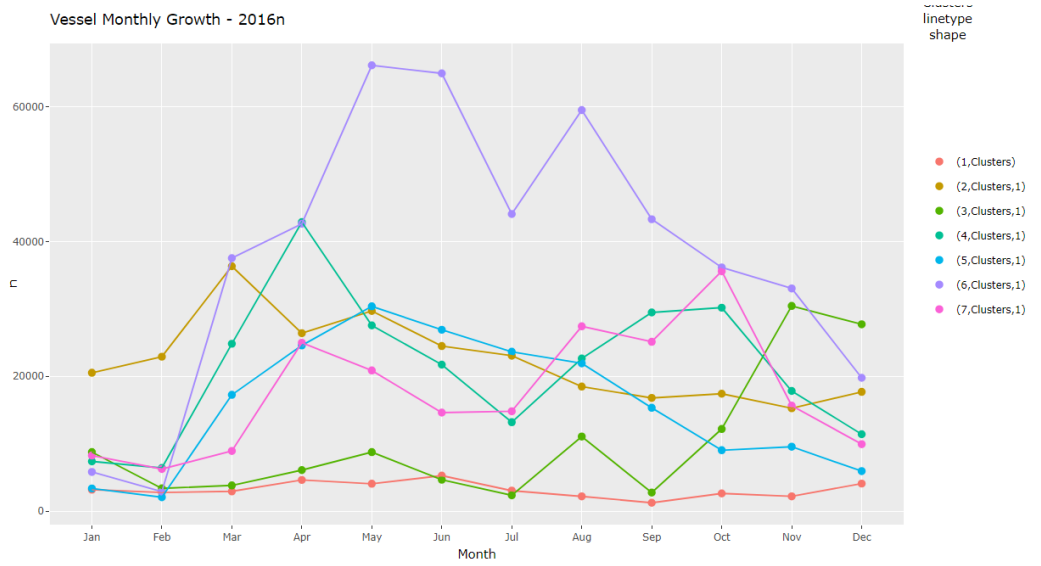
Tabel 4. 8 Distribusi Data Bulanan Pada Tahun 2016, 2017 dan 2018

Month	Distribusi data per Bulan			Jumlah Data
	2016	2017	2018	
1	57,337	60,151	127,632	245,120
2	46,716	83,655	111,376	241,747
3	131,735	168,397	257,410	557,542
4	172,318	211,820	295,814	679,952
5	187,596	199,904	342,017	729,517
6	162,750	154,640	246,640	564,030
7	124,278	163,624	263,688	551,590
8	163,351	247,090	311,679	722,120
9	134,130	262,139	381,483	777,752
10	143,297	275,842	398,586	817,725
11	124,070	262,004	306,131	692,205
12	96,643	163,547	190,171	450,361



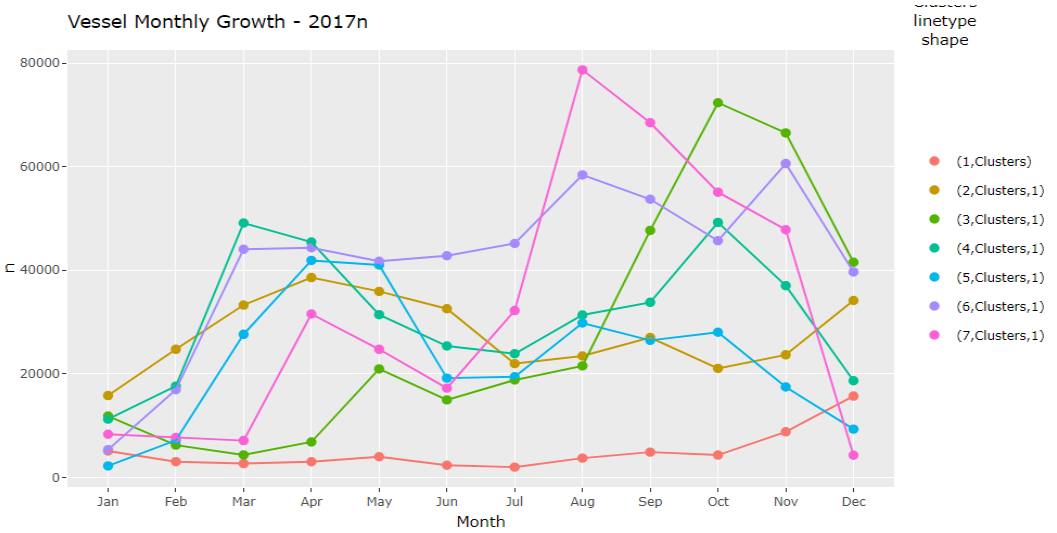
Gambar 4. 21 Distribusi Data per Bulan

Dari data Tabel 4.8 dan Gambar 4.20 diatas dapat diambil kesimpulan bahwa secara keseluruhan jumlah data VMS yang terekam mengalami peningkatan setiap tahunnya, tetapi dari frekuensi setiap bulannya cukup fluktuatif. Berdasarkan Gambar 4.21 dapat dilihat bahwa pada bulan Desember, Januari dan Februari aktifitas kapal perikanan di WPPNRI-711 terekam sedikit karena pada bulan tersebut merupakan periode Muson Asia di perairan laut Natuna dan Selat Karimata yang rawan gelombang tinggi [61], dengan tinggi gelombang diatas 30 cm dan dengan periode sekitar 5 detik [62]. Sedangkan dibulan Juni terjadi penurunan aktifitas kapal perikanan karena titik daerah potensial penangkapan ikan lebih sedikit akibat rendahnya nilai clorofil-a pada musim timur [63]. Untuk distribusi data perbulan di setiap cluster berdasarkan data tahun 2016, 2017, 2018 dan gabungan data tahun 2016 s/d 2018 dapat melihat Gambar 4.22, Gambar 4.23, Gambar 4.24 dan Gambar 4.25 dibawah ini:



Gambar 4. 22 Distibusi Data per *Cluster* Periode Bulanan di Tahun 2016

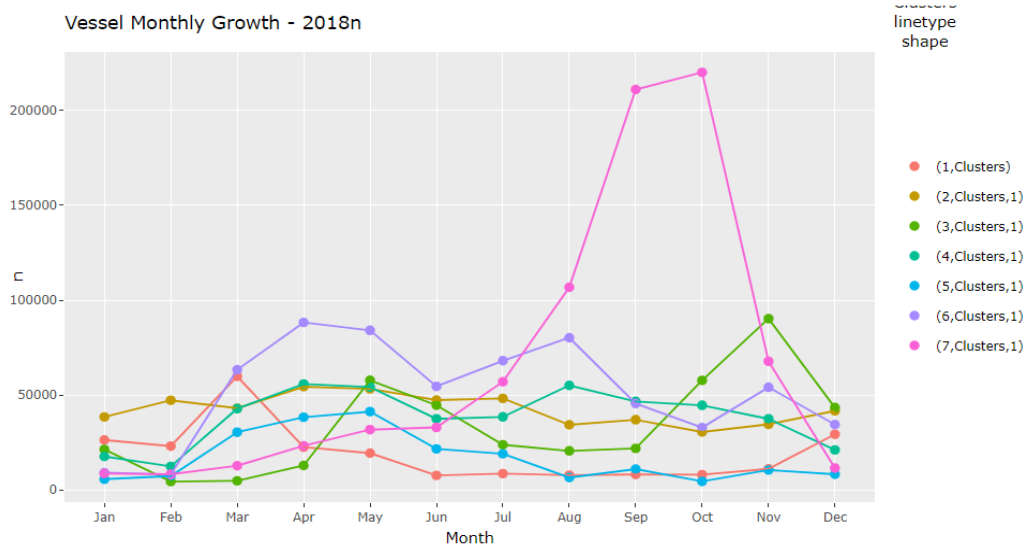
Dari Gambar 4.22 dapat diketahui bahwa pada tahun 2016 pola sebaran data bulanan disetiap cluster sangat bervariasi, cluster 6 memiliki sebaran data paling tinggi pada bulan Mei dan Juni. Sedangkan cluster 1 memiliki sebaran data paling rendah tanpa ada perubahan yang berarti mulai bulan Januari sampai dengan bulan Desember.



Gambar 4. 23 Distibusi Data per *Cluster* Periode Bulanan di Tahun 2017

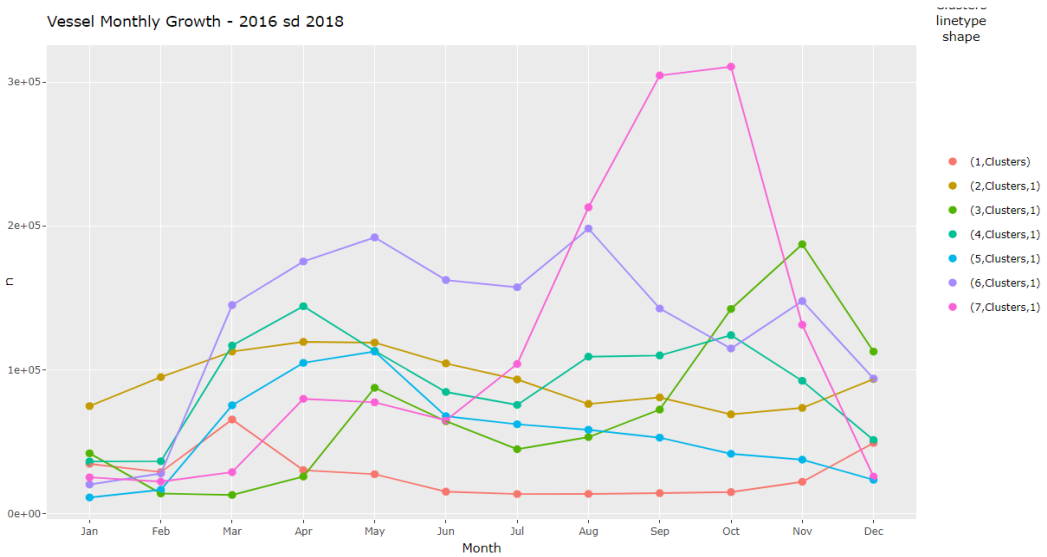
Dari Gambar 4.23 dapat diketahui bahwa pada tahun 2017 pola sebaran data bulanan disetiap cluster sangat bervariasi, cluster 7 memiliki sebaran data paling tinggi pada bulan Agustus dan cluster 3 sebaran data tertinggi pada bulan

Oktober. Sedangkan cluster 1 memiliki sebaran data paling rendah tanpa ada perubahan yang berarti mulai bulan Januari sampai dengan bulan Oktober dan hanya ada sedikit kenaikan dibulan November dan Desember.



Gambar 4. 24 Distibusi Data per *Cluster* Periode Bulanan di Tahun 2018

Dari Gambar 4.24 dapat diketahui bahwa pada tahun 2018 pola sebaran data bulanan disetiap cluster kurang bervariasi, hanya pada cluster 7 memiliki sebaran data paling tinggi pada bulan September dan Oktober. Sedangkan pada cluster lainnya sebaran data bulanan tidak mengalami perubahan yang signifikan, dengan sedikit kenaikan dan penurunan disetiap bulannya.



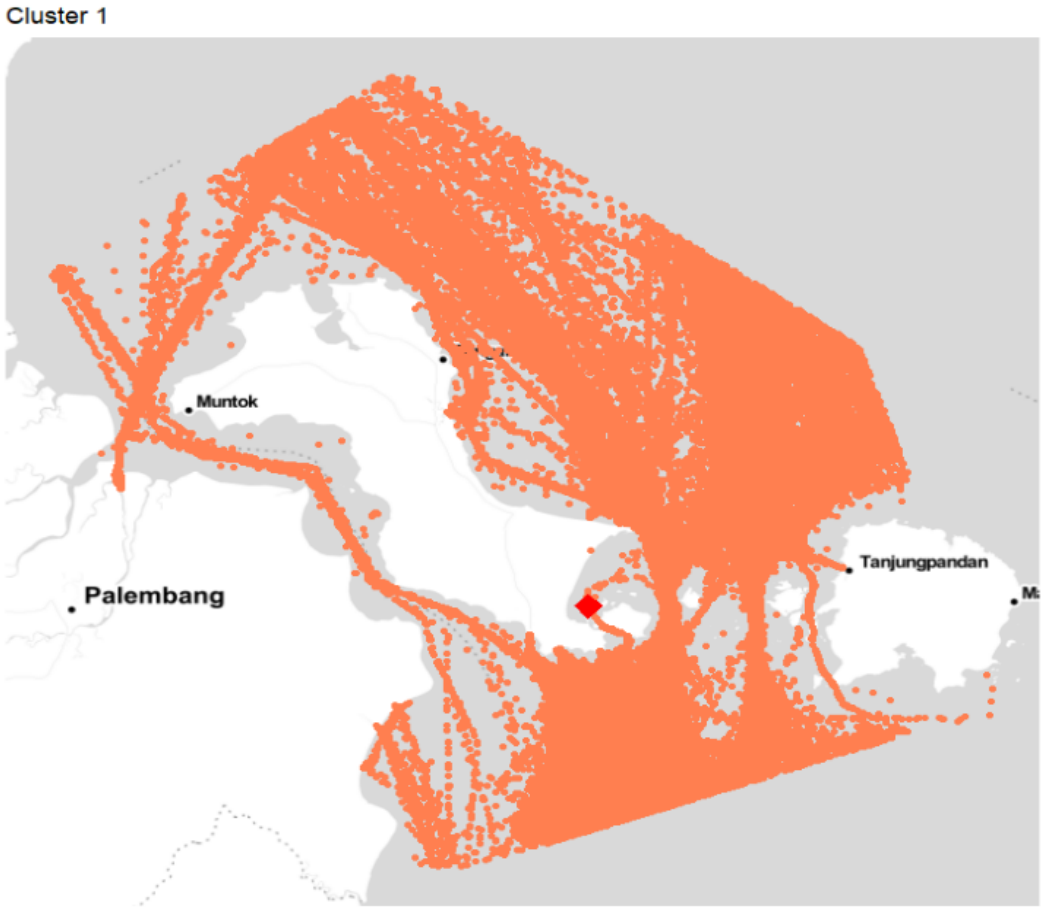
Gambar 4. 25 Distibusi Data per *Cluster* Periode Bulanan Tahun 2016 s/d 2018

Banyaknya jumlah distribusi sebaran data VMS secara spatial di setiap *cluster* tidak berbanding lurus dengan jumlah kapal yang beraktivitas di daerah tersebut. Berdasarkan data Tahun 2016 s/d 2018 di *cluster* 6 yang memiliki distribusi data VMS terbanyak dengan jumlah 1.578.002 data hanya terdapat 428 kapal aktif, sedangkan di *cluster* 1 yang memiliki jumlah sebaran data VMS paling sedikit yaitu 330.996 data terekam ada 735 kapal aktif dan jumlah kapal aktif paling banyak berada di *cluster* 7 dengan 1.144 buah kapal. Hal ini disebabkan oleh pengaruh karakteristik oseanografi di perairan Indonesia yang berhubungan dengan pola distribusi ikan [64]. Sehingga mempengaruhi intensitas dan lamanya kapal beraktifitas menangkap ikan di wilayah *cluster* tersebut, juga dipengaruhi oleh lokasi *cluster,* apakah lokasinya sebagai area penangkapan aktif, sebagai wilayah lintasan atau *home base* dari kapal perikanan.

Untuk lebih detailnya berikut dijelaskan analisis masing-masing cluster:

4.3.1 Analisis *Cluster* 1

Visualisasi distribusi data VMS di *cluster* 1 dapat dilihat pada Gambar 4.26.



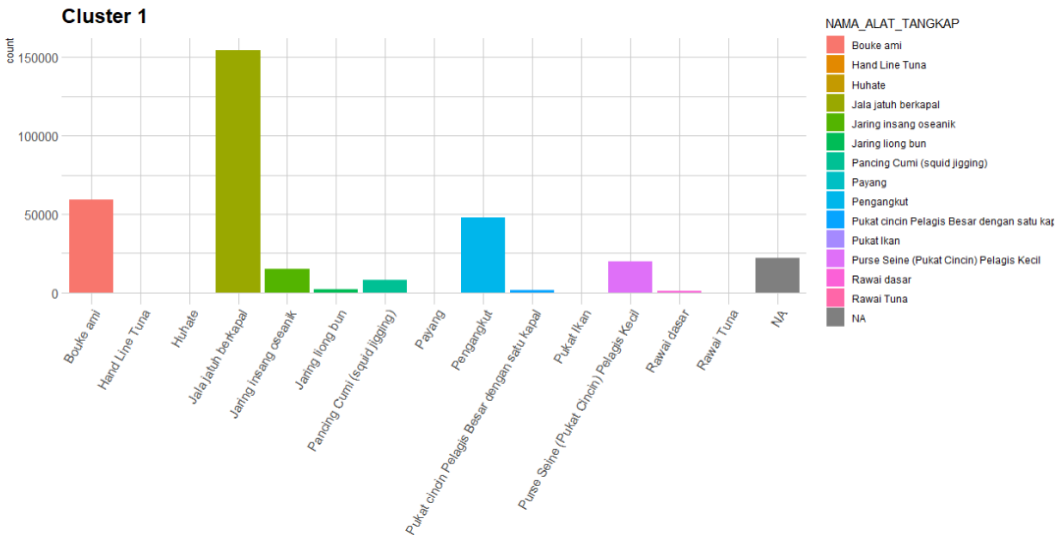
Gambar 4. 26 Visualisasi *Cluster* 1

Dari Gambar 4.26 dapat dilihat bahwa distribusi data VMS pada *cluster* 1 berada di area titik LONGITUDE Min: 104.7 s/d Max:108.2 dan titik LATITUDE Min: -4.024 s/d Max: -0.653. Berada di wilayah Provinsi Sumatera Selatan dan Provinsi Kepulauan Bangka Belitung.

Tabel 4. 9 Analisis *Cluster* 1

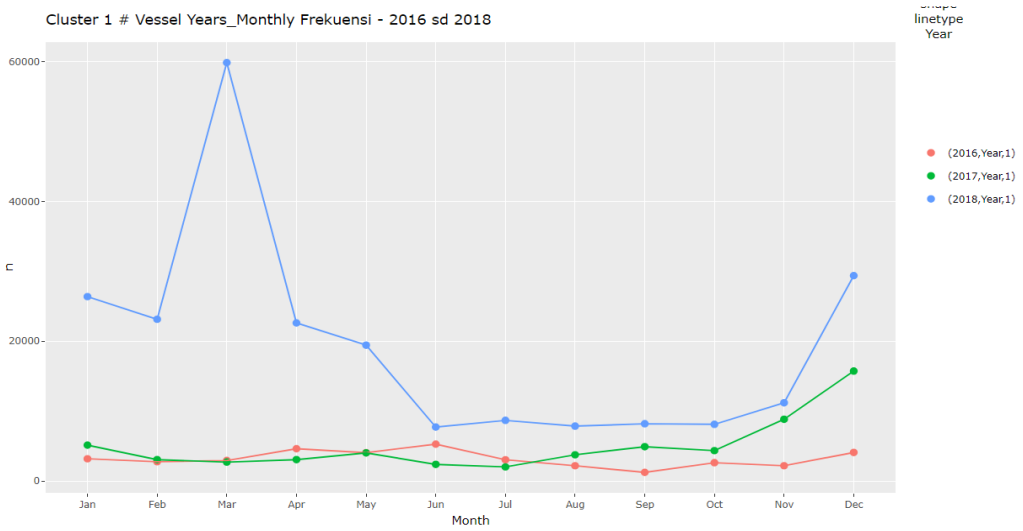
Ranking Cluster		7
Jumlah Anggota Cluster		330,996
Jumlah TRANSMITER_NO Aktif		735
NAMA_ALAT_TANGKAP		
1	Bouke ami	127
2	Hand Line Tuna	1
3	Huhate	1
4	Jala jatuh berkapal	223
5	Jaring insang oseanik	93
6	Jaring liong bun	14
7	Pancing Cumi (squid jigging)	72
8	Payang	1
9	Pengangkut	27
10	Pukat cincin Pelagis Besar dengan satu kapal	16
11	Pukat Ikan	1
12	Purse Seine (Pukat Cincin) Pelagis Kecil	77
13	Rawai dasar	5
14	Rawai Tuna	3
15	NA	74

Tabel 4.9 merupakan tabel hasil analis *cluster* 1 yang berisikan karakteristik dan distribusi data *vessel monitoring system (VMS)* berdasarkan hasil *clustering*. *Cluster* 1 memiliki jumlah anggota paling sedikit dibandingkan dengan *cluster* lainnya, beranggotakan 330.996 data VMS dengan jumlah TRANSMITER_NO aktif sebanyak 735 unit. Berdasarkan alat tangkap yang dipergunakan dapat dilihat bahwa Jala Jatuh Berkapal digunakan oleh 233 kapal perikanan, alat tangkap Bouke Ami dipergunakan oleh 127 kapal dan ada 375 kapal menggunakan alat tangkap lainnya. Sedangkan untuk jumlah distribusi data berdasarkan alat tangkap yang dipergunakan dapat melihat Gambar 4.27.

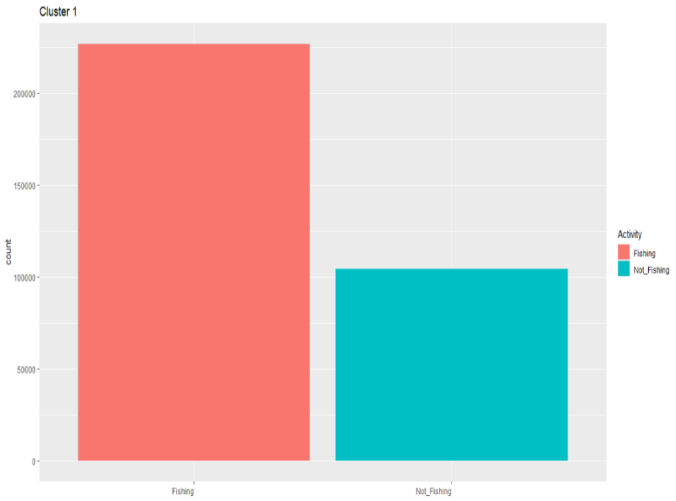


Gambar 4. 27 Distribusi Data Berdasarkan Nama Alat Tangkap *Cluster* 1

Untuk mendapatkan gambaran tentang distribusi temporal sebaran data VMS di *cluster* 1 pada Gambar 4.28 akan ditampilakan grafik distribusi data perbulan pada tahun 2016, 2017, dan 2018, serta diagram batang perbandingan aktifitas kapal penangkap ikan pada Gambar 4.29. Data VMS tidak menunjukkan secara langsung apakah posisi kapal sedang menangkap ikan atau dalam perjalanan, tetapi berdasarkan kecepatan kapal dapat mendefinisikan aktifitas kapal perikanan. Kapal dengan kecepatan <8 knot di tetapkan sedang beraktifitas menangkap ikan [65], kesimpulan ini diambil tanpa membedakan alat tangkap yang dipergunakan oleh setiap kapal.



Gambar 4. 28 Grafik Distribusi Data *Cluster* 1

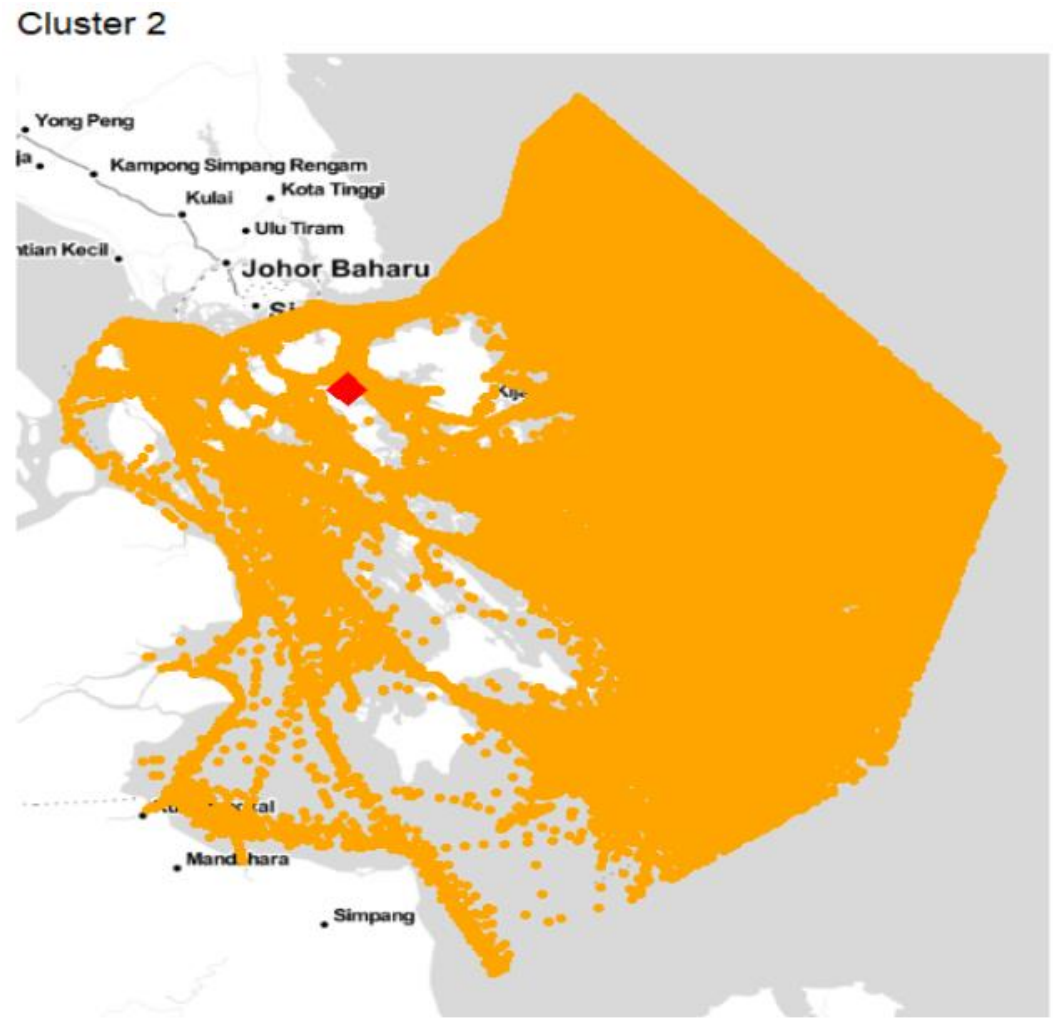


Gambar 4. 29 Diagram Batang Aktivitas Kapal Perikanan *Cluster* 1

Berdasarkan Gambar 4.28 menunjukkan grafik yang merepresentasikan sumbu x sebagai waktu time series per bulan sedangkan sumbu y merepresentasikan jumlah distribusi data VMS. Untuk pengamatan *Cluster* 1 secara temporal grafik distribusi data menunjukan bahwa pada tahun 2016 jumlah data yang terdistribusi hanya sedikit dan dari bulan Januari s/d Desember tanpa mengalami perubahan berarti, begitu juga pada tahun 2017 terjadi sedikit kenaikan hanya pada bulan November dan Desember. Pada tahun 2018 frekuensi distribusi data sangat fluktuatif, jumlah distribusi data tertinggi terjadi pada bulan Maret 2018 dan terendah rata-rata pada bulan Juni s/d November 2018. Dari Gambar 4.29 dapat dilihat perbandingan jumlah data aktifitas kapal perikanan berdasarkan variable SPEED, dengan jumlah data 226.514 kategori *fishing* dan 104.454 data kategori *not_fishing.*

4.3.2 Analisis *Cluster* 2

Visualisasi distribusi data VMS di *cluster* 2 dapat dilihat pada Gambar 4.30.



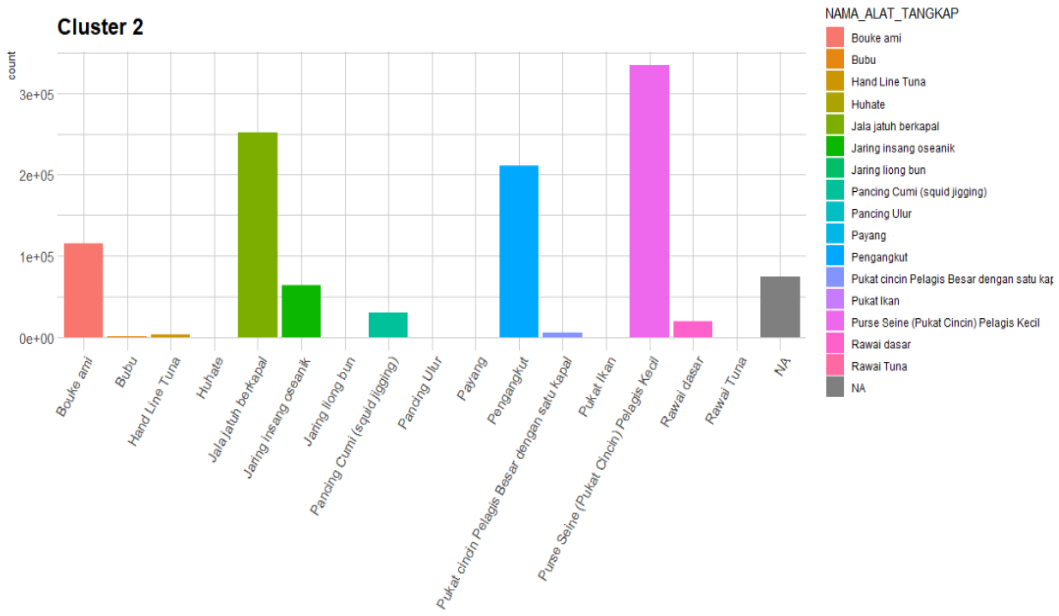
Gambar 4. 30 Visualisasi *Cluster* 2

Dari Gambar 4.30 dapat dilihat bahwa distribusi data VMS di *cluster* 2 berada di area titik LONGITUDE Min: 103.2 s/d Max:106.4 dan titik LATITUDE Min: -1.4765 s/d Max: 2.1430. Berada di wilayah Provinsi Riau, Provinsi Jambi dan Provinsi Kepulauan Riau.

Tabel 4. 10 Analisis *Cluster* 2

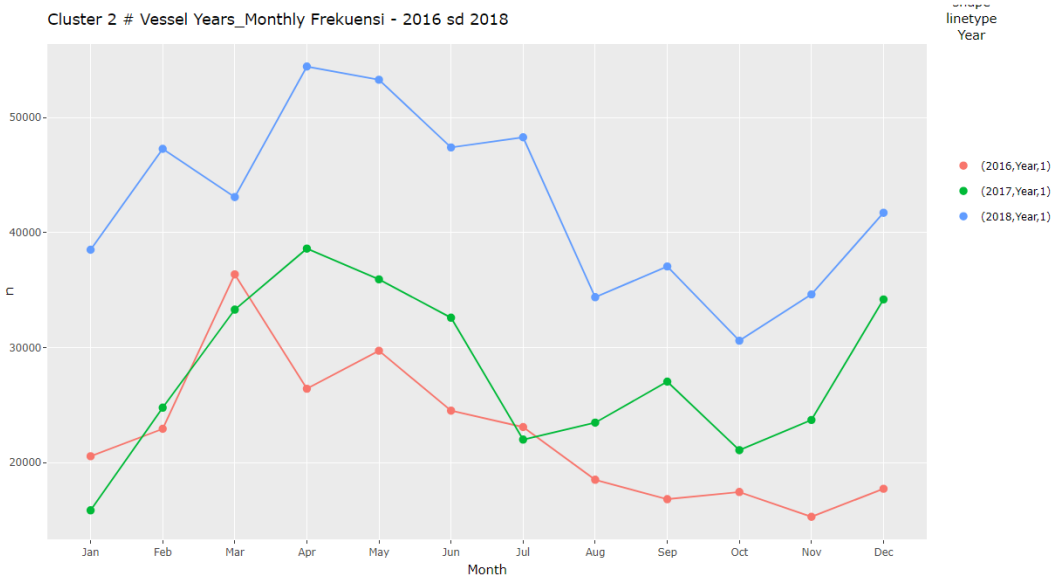
Ranking Cluster		3
Jumlah Anggota Cluster		1.112.355
Jumlah TRANSMITER_NO Aktif		848
NAMA_ALAT_TANGKAP		
1	Bouke ami	57
2	Bubu	4
3	Hand Line Tuna	3
4	Huhate	3
5	Jala jatuh berkapal	116
6	Jaring insang oseanik	90
7	Jaring liong bun	5
8	Pancing Cumi (squid jigging)	95
9	Pancing Ulur	1
10	Payang	1
11	Pengangkut	81
12	Pukat cincin Pelagis Besar dengan satu kapal	37
13	Pukat Ikan	3
14	Purse Seine (Pukat Cincin) Pelagis Kecil	195
15	Rawai dasar	32
16	Rawai Tuna	35
17	NA	90

Tabel 4.10 merupakan tabel hasil analis *cluster* 2 yang berisikan karakteristik dan distribusi data *vessel monitoring system (VMS)* berdasarkan hasil *clustering*. *Cluster* 2 memiliki jumlah anggota terbanyak ke-3 dengan beranggotakan 1.112.355 data VMS, dan jumlah TRANSMITER_NO aktif sebanyak 848 unit. Berdasarkan alat tangkap yang dipergunakan dapat dilihat bahwa Purse Seine (Pukat Cincin) Pelagis Kecil digunakan oleh 195 kapal perikanan, alat tangkap Jala Jatuh Berkapal dipergunakan oleh 116 kapal dan 487 kapal menggunakan alat tangkap lainnya. Sedangkan untuk jumlah distribusi data berdasarkan alat tangkap yang dipergunakan dapat melihat Gambar 4.31.

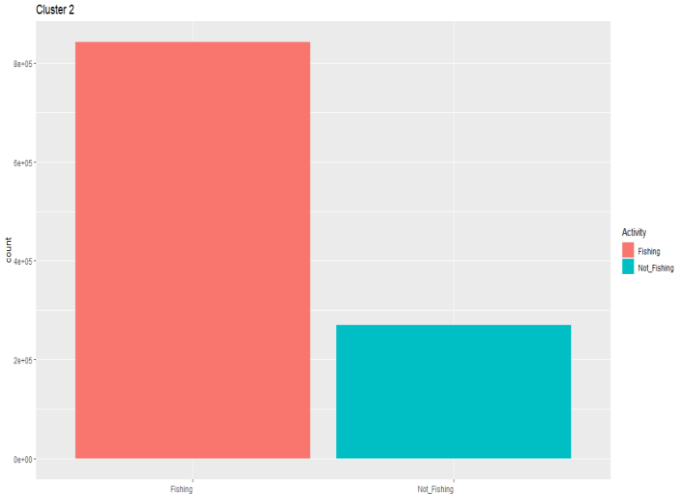


Gambar 4. 31 Distribusi Data Berdasarkan Nama Alat tangkap *Cluster* 2

Untuk mendapatkan gambaran tentang distribusi temporal sebaran data VMS di *cluster* 2 pada Gambar 4.32 akan ditampilakan grafik distribusi data perbulan pada tahun 2016, 2017, dan 2018, serta diagram batang perbandingan aktifitas kapal penangkap ikan pada Gambar 4.33.



Gambar 4. 32 Grafik Distribusi Data *Cluster* 2

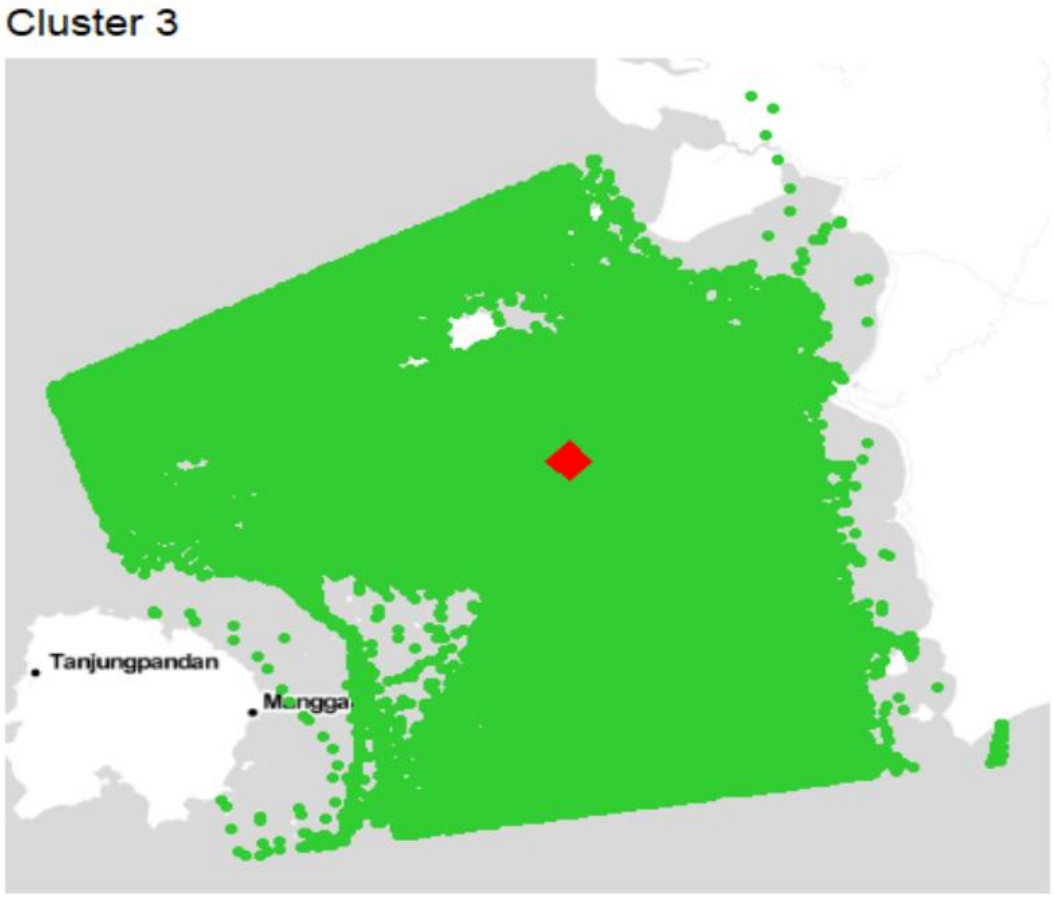


Gambar 4. 33 Diagram Batang Aktivitas Kapal Perikanan Cluster 2

Berdasarkan Gambar 4.32 menunjukkan grafik yang merepresentasikan sumbu x sebagai waktu time series per bulan sedangkan sumbu y merepresentasikan jumlah distribusi data VMS. Untuk pengamatan *Cluster* 2 secara temporal grafik distribusi data menunjukan peningkatan dari tahun ke tahun dengan fluktuasi yang cukup tinggi, data terendah terjadi pada bulan November 2016 dan data tertinggi pada bulan April 2018. Pada tahun 2016 jumlah data yang terdistribusi secara keseluruhan paling sedikit terendah di bulan November dan data tertinggi terjadi pada bulan Maret, sedangkan pada tahun 2017 data terendah pada bulan Januari dan kenaikan tertinggi pada bulan April. Pada tahun 2018 frekuensi distribusi data sangat fluktuatif, jumlah distribusi data tertinggi terjadi pada bulan April dan terendah pada bulan Oktober. Dari Gambar 4.33 dapat dilihat perbandingan jumlah data aktifitas kapal perikanan berdasarkan variable SPEED, dengan jumlah data 841.717 kategori *fishing* dan 270.610 data kategori *not_fishing.*

4.3.3 Analisis *Cluster* 3

Visualisasi distribusi data VMS di cluster 3 dapat dilihat pada Gambar 4.34.



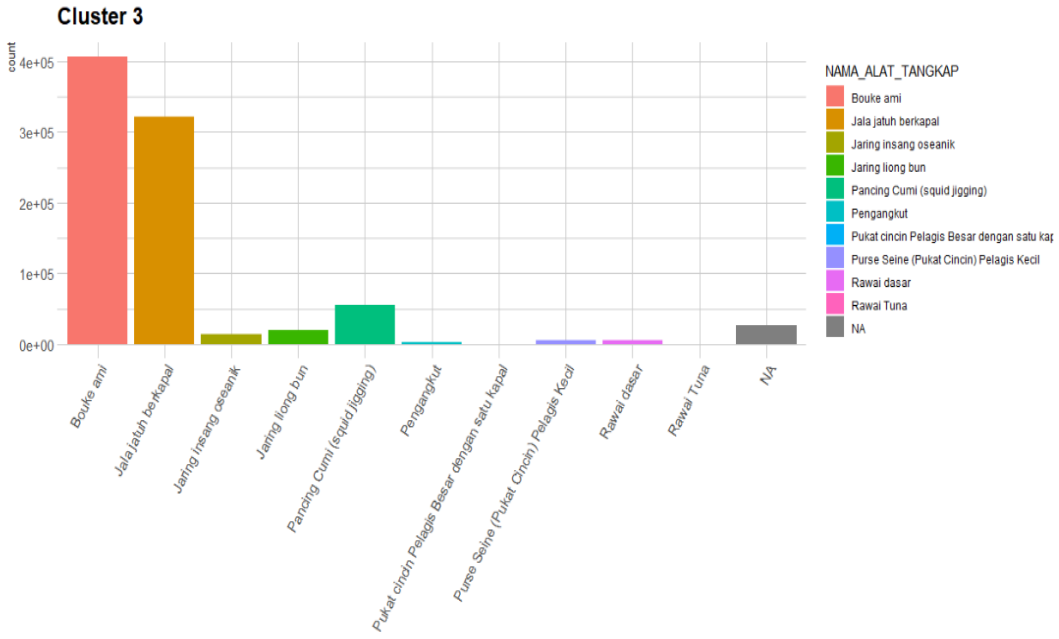
Gambar 4. 34 Visualisasi *Cluster* 3

Dari Gambar 4.34 dapat dilihat bahwa distribusi data VMS di *cluster* 3 berada di area titik LONGITUDE Min: 107.7 s/d Max:110.4 dan titik LATITUDE Min: -3.3696 s/d Max: -0.8287. Berada di wilayah Provinsi Kalimantan Barat dan Provinsi Kepulauan Bangka Belitung.

Tabel 4. 11 Analisa *Cluster* 3

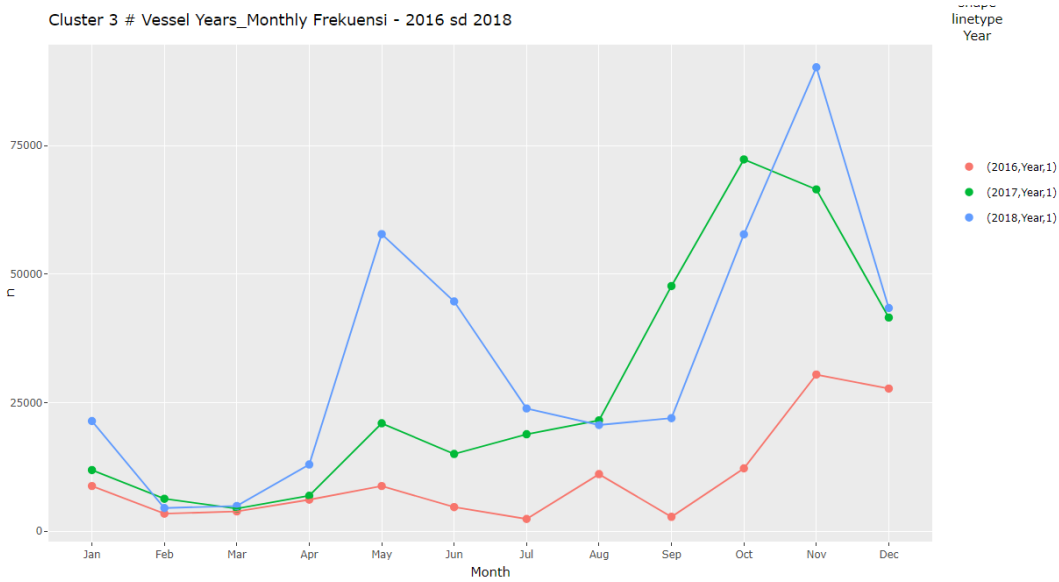
Ranking Cluster		5
Jumlah Anggota Cluster		860.134
Jumlah TRANSMITER_NO Aktif		1.205
NAMA_ALAT_TANGKAP		
1	Bouke ami	328
2	Jala jatuh berkapal	441
3	Jaring insang oseanik	120
4	Jaring liong bun	49
5	Pancing Cumi (squid jigging)	79
6	Pengangkut	14
7	Pukat cincin Pelagis Besar dengan satu kapal	3
8	Purse Seine (Pukat Cincin) Pelagis Kecil	50
9	Rawai dasar	47
10	Rawai Tuna	2
11	NA	72

Tabel 4.11 merupakan tabel hasil analis *cluster* 3 yang berisikan karakteristik dan distribusi data *vessel monitoring system (VMS)* berdasarkan hasil *clustering*. *Cluster* 3 memiliki jumlah anggota 860.134 data VMS berada di urutan ke 5 dengan jumlah TRANSMITER_NO aktif sebanyak 1.205 unit. Berdasarkan alat tangkap yang dipergunakan dapat dilihat bahwa alat tangkap Jala Jatuh Berkapal dipergunakan oleh 441 kapal penangkap ikan, alat Bouke Ami digunakan oleh 328 kapal dan 436 kapal menggunakan alat tangkap lainnya. Sedangkan untuk perbandingan jumlah distribusi data berdasarkan alat tangkap yang dipergunakan dapat melihat Gambar 4.35.

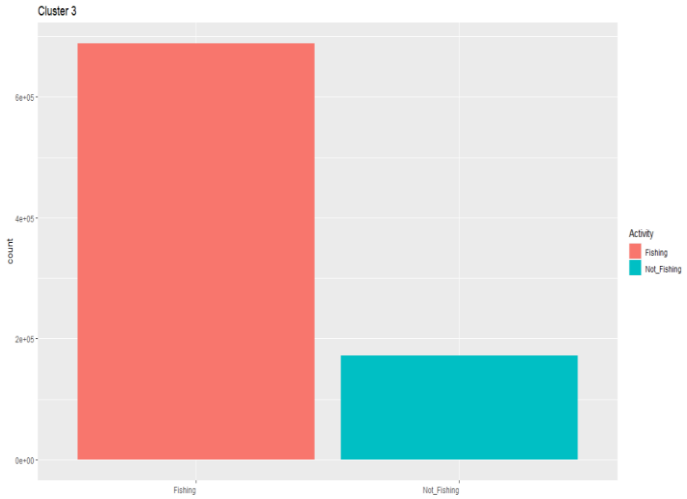


Gambar 4. 35 Distribusi Data Berdasarkan Nama Alat Tangkap *Cluster* 3

Untuk mendapatkan gambaran tentang distribusi temporal sebaran data VMS di *cluster* 3 pada Gambar 4.36 akan ditampilakan grafik distribusi data perbulan pada tahun 2016, 2017, dan 2018, serta diagram batang perbandingan aktifitas kapal penangkap ikan pada Gambar 4.37.



Gambar 4. 36 Grafik Distribusi Data *Cluster* 3

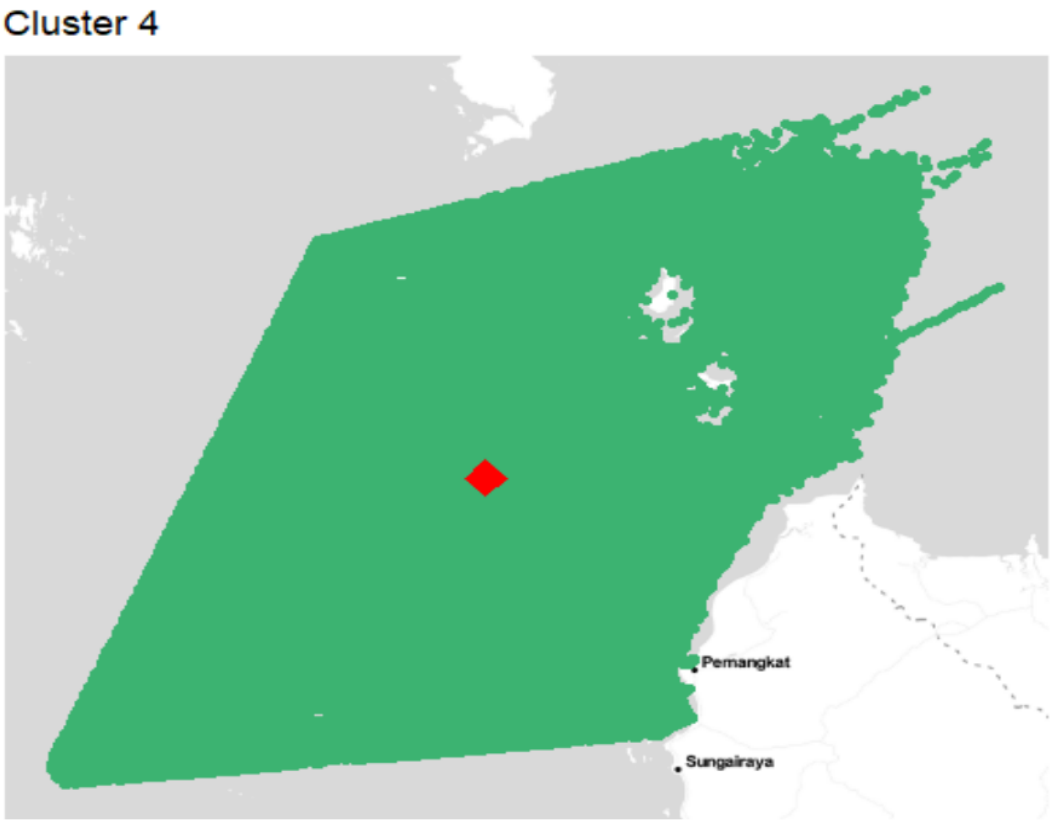


Gambar 4. 37 Diagram Batang Aktivitas Kapal Perikanan *Cluster* 3

Berdasarkan Gambar 4.36 menunjukkan grafik yang merepresentasikan sumbu x sebagai waktu time series per bulan sedangkan sumbu y merepresentasikan jumlah distribusi data VMS. Untuk pengamatan *Cluster* 3 secara temporal grafik distribusi data menunjukan peningkatan dari tahun ke tahun dengan fluktuasi yang cukup tinggi di tahun 2017 dan 2018, data terendah terjadi pada bulan Juli dan September 2016 dan data tertinggi pada bulan November 2018. Pada tahun 2016 jumlah data yang terdistribusi secara keseluruhan paling sedikit dan data tertinggi terjadi pada bulan November. Sedangkan pada tahun 2017 data terendah pada bulan Maret dan kenaikan tertinggi pada bulan Oktober, bahkan pada bulan September dan Oktober jumlah data lebih tinggi dibandingkan tahun 2018 pada bulan yang sama. Pada tahun 2018 frekuensi distribusi data sangat fluktuatif, jumlah distribusi data tertinggi terjadi pada bulan November dan terendah pada bulan Februari. Dari Gambar 4.37 dapat dilihat perbandingan jumlah data aktifitas kapal perikanan berdasarkan variable SPEED, dengan jumlah data 688.159 kategori *fishing* dan 171.791 data kategori *not_fishing*.

4.3.4 Analisis *Cluster* 4

Visualisasi distribusi data VMS di *cluster* 4 dapat dilihat pada Gambar 4.38.



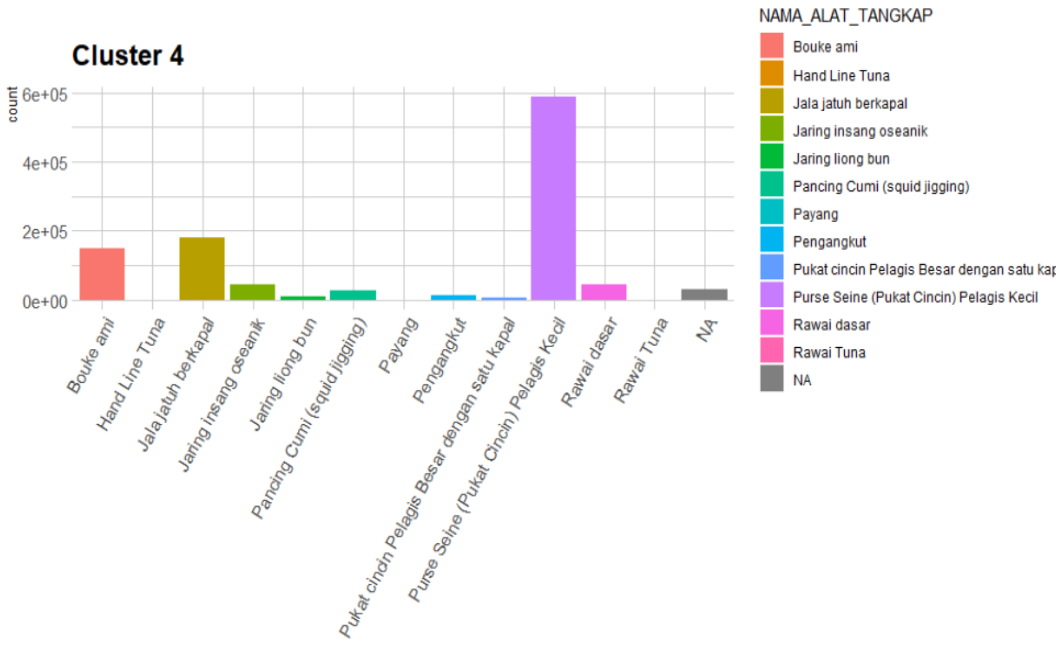
Gambar 4. 38 Visualisasi *Cluster* 4

Dari Gambar 4.38 dapat dilihat bahwa distribusi data VMS di *cluster* 4 berada di area titik LONGITUDE Min: 106.4 s/d Max:110.2 dan titik LATITUDE Min: 0.6167 s/d Max: 3.8760. Berada di wilayah Provinsi Kalimantan Barat dan Provinsi Kepulauan Riau.

Tabel 4. 12 Analisis *Cluster* 4

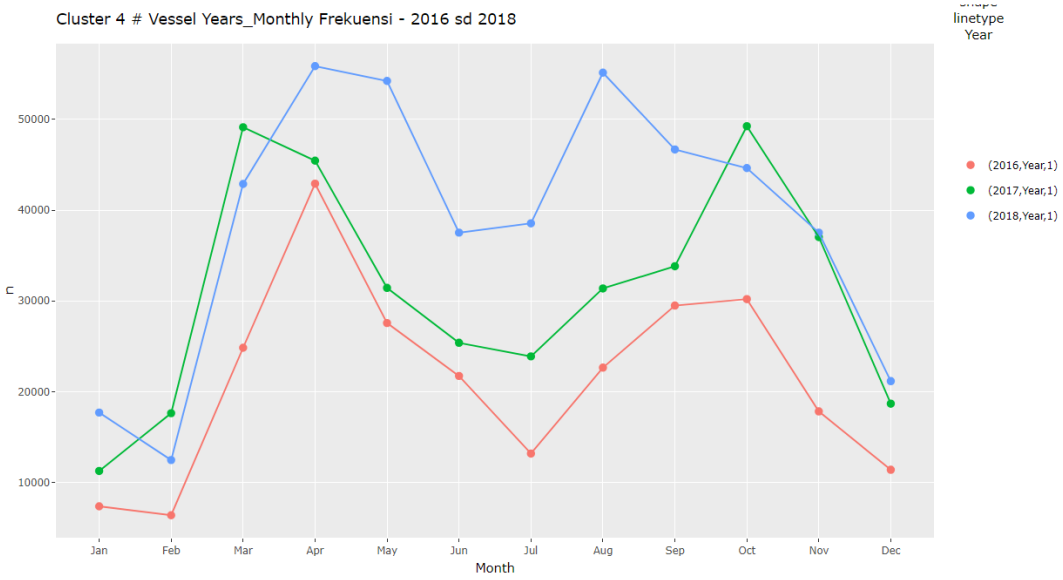
Ranking Cluster		4
Jumlah Anggota Cluster		1.094.501
Jumlah TRANSMITER_NO Aktif		659
NAMA_ALAT_TANGKAP		
1	Bouke ami	116
2	Hand Line Tuna	1
3	Jala jatuh berkapal	120
4	Jaring insang oseanik	40
5	Jaring liong bun	12
6	Pancing Cumi (squid jigging)	45
7	Payang	1
8	Pengangkut	18
9	Pukat cincin Pelagis Besar dengan satu kapal	9
10	Purse Seine (Pukat Cincin) Pelagis Kecil	229
11	Rawai dasar	26
12	Rawai Tuna	1
13	NA	41

Tabel 4.12 merupakan tabel kesimpulan hasil analis *cluster* 4 yang berisikan karakteristik dan distribusi data *vessel monitoring system (VMS)* berdasarkan hasil *clustering*. *Cluster* 4 memiliki jumlah anggota 1.094.501 data VMS berada di urutan ke 4 dengan jumlah TRANSMITER_NO aktif sebanyak 659 unit. Berdasarkan alat tangkap yang dipergunakan alat Purse Seine (Pukat Cincin) Pelagis Kecil digunakan oleh 229 kapal, alat tangkap Jala Jatuh Berkapal dipergunakan oleh 120 kapal dan alat Bouke Ami digunakan oleh 116 kapal serta ada 194 kapal menggunakan alat tangkap lainnya. Sedangkan untuk melihat jumlah distribusi data berdasarkan alat tangkap yang dipergunakan dapat melihat Gambar 4.39.

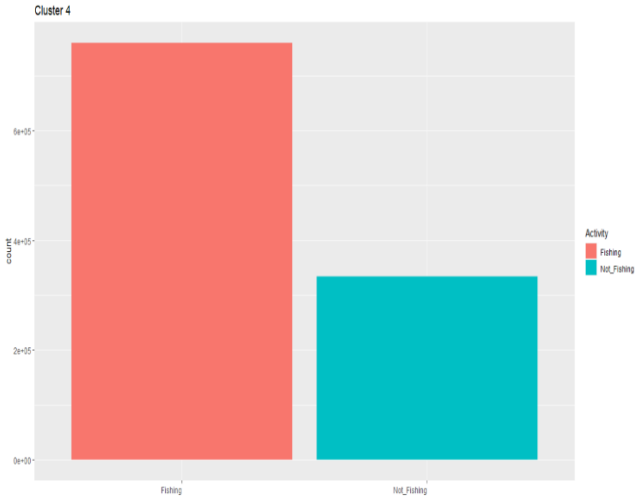


Gambar 4. 39 Distribusi Data Berdasarkan Nama Alat Tangkap *Cluster* 4

Untuk mendapatkan gambaran tentang distribusi temporal sebaran data VMS di *cluster* 4 pada Gambar 4.40 akan ditampilakan grafik distribusi data perbulan pada tahun 2016, 2017, dan 2018, serta diagram batang perbandingan aktifitas kapal penangkap ikan pada Gambar 4.41.



Gambar 4. 40 Grafik Distribusi Data *Cluster* 4

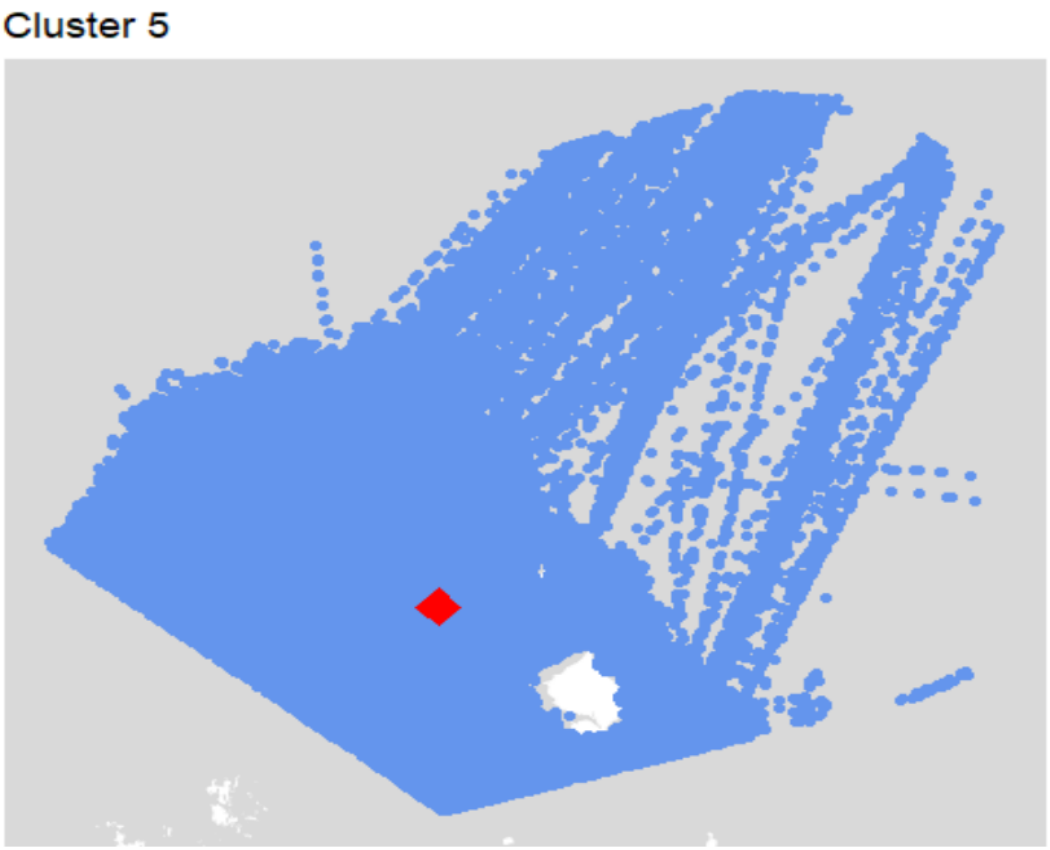


Gambar 4. 41 Diagram Batang Aktivitas Kapal Perikanan *Cluster* 4

Berdasarkan Gambar 4.40 menunjukkan grafik yang merepresentasikan sumbu x sebagai waktu time series per bulan sedangkan sumbu y merepresentasikan jumlah distribusi data VMS. Untuk pengamatan *Cluster* 4 secara temporal grafik distribusi data menunjukan peningkatan dari tahun ke tahun dengan fluktuasi yang cukup tinggi, data terendah terjadi pada bulan Februari 2016 dan data tertinggi pada bulan Agustus 2018. Pada tahun 2016 jumlah data yang terdistribusi secara keseluruhan paling sedikit dan data tertinggi terjadi pada bulan April. Sedangkan pada tahun 2017 data terendah pada bulan Januari dan kenaikan tertinggi pada bulan Maret dan Oktober, bahkan pada bulan tersebut jumlah data lebih tinggi dibandingkan tahun 2018 pada bulan yang sama. Pada tahun 2018 frekuensi distribusi data sangat fluktuatif, jumlah distribusi data tertinggi terjadi pada bulan April dan Agustus terendah pada bulan Februari. Dari Gambar 4.41 dapat dilihat perbandingan jumlah data aktifitas kapal perikanan berdasarkan variable SPEED, dengan jumlah data 759.847 kategori *fishing* dan 334654 data kategori *not_fishing*.

4.3.5 Analisis *Cluster* 5

Visualisasi distribusi data VMS di *cluster* 5 dapat dilihat pada Gambar 4.42.



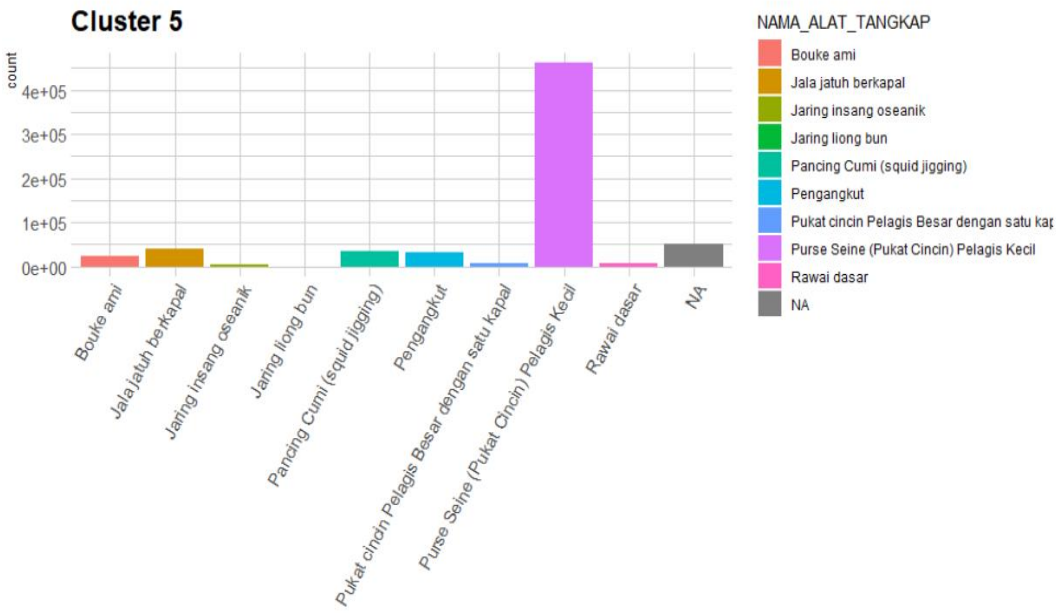
Gambar 4. 42 Visualisasi *Cluster* 5

Dari Gambar 4.42 dapat dilihat bahwa distribusi data VMS di *cluster 5* berada di area titik LONGITUDE Min: 105.4 s/d Max :110.4 dan titik LATITUDE Min: 3.174 s/d Max: 7.738. Berada di wilayah Provinsi Kepulauan Riau.

Tabel 4. 13 Analisis *Cluster* 5

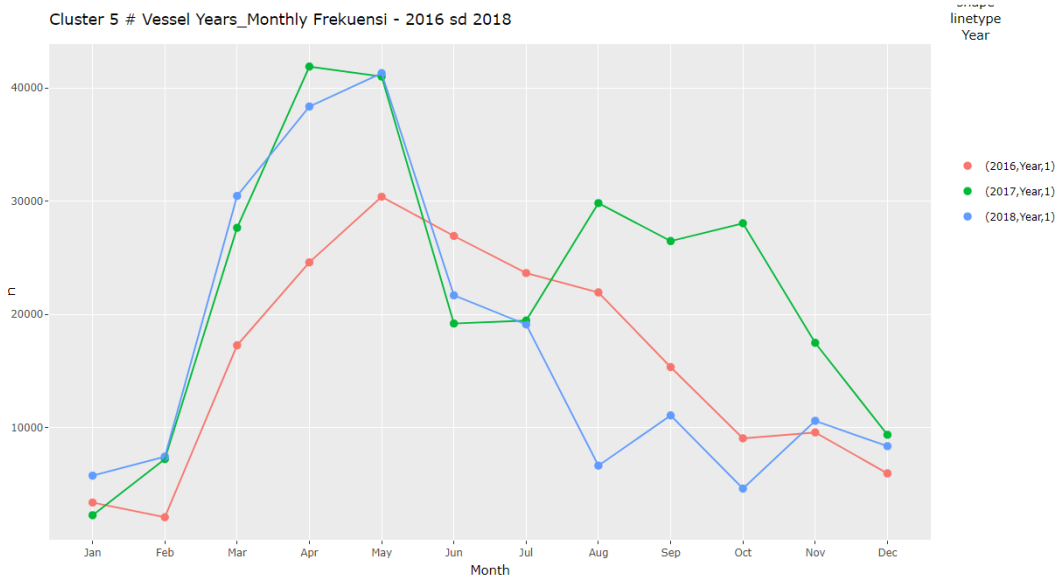
Ranking Cluster		6
Jumlah Anggota Cluster		665.463
Jumlah TRANSMITER_NO Aktif		390
NAMA_ALAT_TANGKAP		
1	Bouke ami	36
2	Jala jatuh berkapal	42
3	Jaring insang oseanik	8
4	Jaring liong bun	2
5	Pancing Cumi (squid jigging)	28
6	Pengangkut	16
7	Pukat cincin Pelagis Besar dengan satu kapal	12
8	Purse Seine (Pukat Cincin) Pelagis Kecil	195
9	Rawai dasar	14
10	NA	37

Tabel 4.13 merupakan tabel hasil analis *cluster* 5 yang berisikan karakteristik dan distribusi data *vessel monitoring system (VMS)* berdasarkan hasil *clustering*. *Cluster* 5 memiliki jumlah anggota 665.463 data VMS berada di urutan ke 6 dengan jumlah TRANSMITER_NO aktif paling sedikit dibandingkan cluster lainnya dengan jumlah 390 unit. Berdasarkan alat tangkap yang dipergunakan alat Purse Seine (Pukat Cincin) Pelagis Kecil digunakan oleh 195 kapal, alat tangkap Jala Jatuh Berkapal dipergunakan oleh 42 kapal dan alat Bouke Ami digunakan oleh 36 kapal serta ada 117 kapal menggunakan jenis alat tangkap lainnya. Sedangkan untuk melihat jumlah distribusi data berdasarkan alat tangkap yang dipergunakan dapat melihat Gambar 4.43.

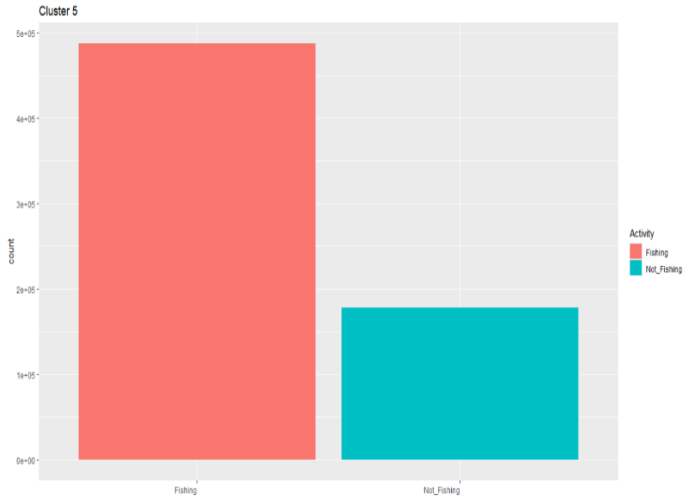


Gambar 4. 43 Distribusi Data Berdasarkan Nama Alat Tangkap *Cluster* 5

Untuk mendapatkan gambaran tentang distribusi temporal sebaran data VMS di *cluster* 5 pada Gambar 4.44 akan ditampilakan grafik distribusi data perbulan pada tahun 2016, 2017, dan 2018, serta diagram batang perbandingan aktifitas kapal penangkap ikan pada Gambar 4.45.



Gambar 4. 44 Grafik Distribusi Data Cluster 5

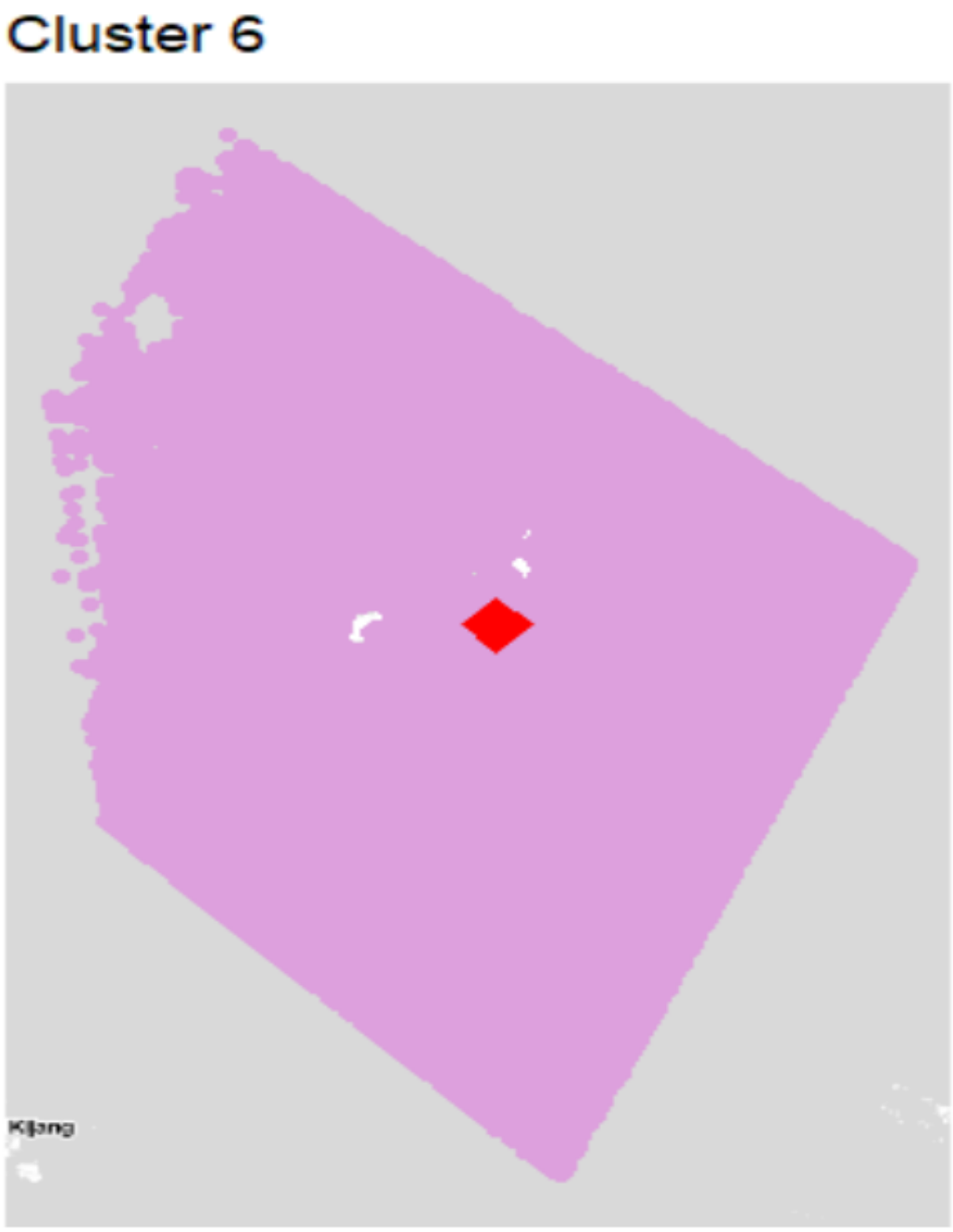


Gambar 4. 45 Diagram Batang Aktivitas Kapal Perikanan *Cluster* 5

Berdasarkan Gambar 4.44 menunjukkan grafik yang merepresentasikan sumbu x sebagai waktu time series per bulan sedangkan sumbu y merepresentasikan jumlah distribusi data VMS. Untuk pengamatan *Cluster* 5 secara temporal grafik distribusi data menunjukan fluktuasi yang cukup tinggi, data terendah terjadi pada bulan Februari 2016 dan januari 2017 sedangkan data tertinggi pada bulan April 2017. Pada tahun 2016 jumlah data yang terdistribusi paling rendah pada bulan Februari dan data tertinggi terjadi pada bulan Mei, bahkan pada bulan Juni dan Juli jumlah data lebih tinggi dibandingkan tahun 2017 dan 2018 pada bulan yang sama. Sedangkan pada tahun 2017 data terendah pada bulan Januari dan kenaikan tertinggi pada bulan April, bahkan pada bulan April, Agustus, September, Oktober, November dan Desember jumlah data lebih tinggi dibandingkan tahun 2016 dan 2018 pada bulan yang sama. Pada tahun 2018 frekuensi distribusi data sangat fluktuatif, jumlah distribusi data tertinggi terjadi pada bulan Mei dan terendah pada bulan Januari dan Oktober karena pada bulan September dan Oktober perairan Laut Natuna rawan akan gelombang tinggi [61]. Dari Gambar 4.45 dapat dilihat perbandingan jumlah data aktifitas kapal perikanan berdasarkan variable SPEED, dengan jumlah data 487.313 kategori *fishing* dan 178.115 data kategori *not_fishing*.

4.3.6 Analisis *Cluster* 6

Visualisasi distribusi data VMS di *cluster* 6 dapat dilihat pada Gambar 4.46.



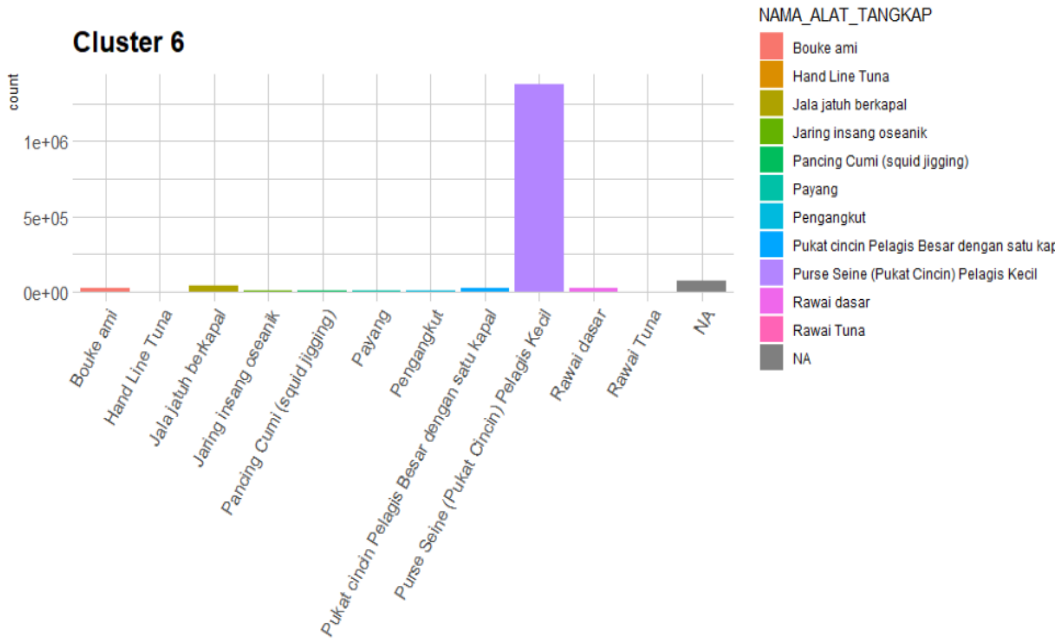
Gambar 4. 46 Visualisasi Cluster 6

Dari Gambar 4.46 dapat dilihat bahwa distribusi data VMS di *cluster* 6 berada di area titik LONGITUDE Min: 104.8 s/d Max:107.4 dan titik LATITUDE Min: 0.732 s/d Max: 4.867. Berada di wilayah Provinsi Kepulauan Riau.

Tabel 4. 14 Analisis *Cluster* 6

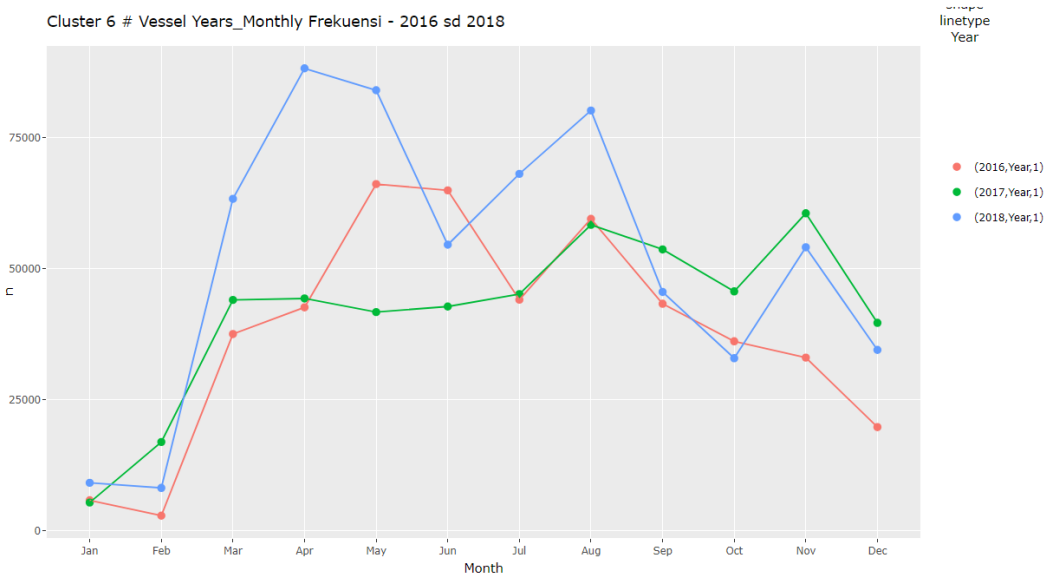
Ranking Cluster		1
Jumlah Anggota Cluster		1.578.002
Jumlah TRANSMITER_NO Aktif		428
NAMA_ALAT_TANGKAP		
1	Bouke ami	33
2	Hand Line Tuna	1
3	Jala jatuh berkapal	68
4	Jaring insang oseanik	18
5	Pancing Cumi (squid jigging)	19
6	Payang	1
7	Pengangkut	15
8	Pukat cincin Pelagis Besar dengan satu kapal	13
9	Purse Seine (Pukat Cincin) Pelagis Kecil	201
10	Rawai dasar	22
11	Rawai Tuna	1
12	NA	36

Tabel 4.14 merupakan tabel hasil analis *cluster* 6 yang berisikan karakteristik dan distribusi data *vessel monitoring system (VMS)* berdasarkan hasil *clustering*. *Cluster* 6 merupakan *cluster* yang memiliki jumlah anggota terbanyak dengan jumlah 1.578.002 data VMS dan jumlah TRANSMITER_NO aktif berjumlah 428 unit. Berdasarkan alat tangkap yang dipergunakan alat Purse Seine (Pukat Cincin) Pelagis Kecil digunakan oleh 210 kapal dan alat tangkap Jala Jatuh Berkapal dipergunakan oleh 68 kapal dan ada 150 kapal menggunakan jenis alat tangkap lainnya. Sedangkan untuk melihat jumlah distribusi data berdasarkan alat tangkap yang dipergunakan dapat melihat Gambar 4.47.

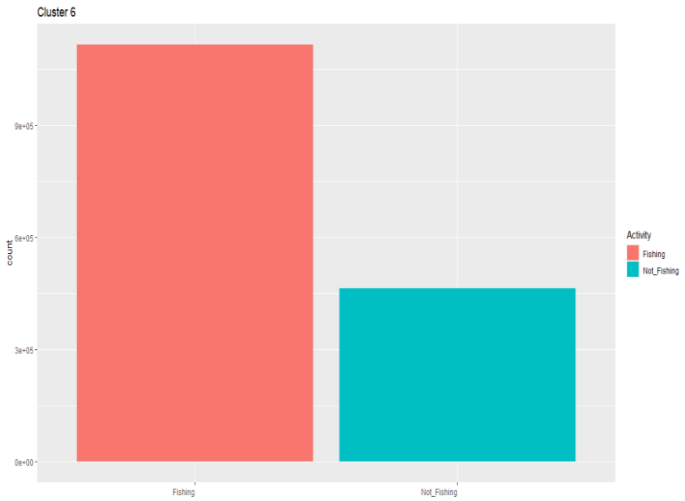


Gambar 4. 47 Distribusi Data berdasarkan Nama Alat Tangkap *Cluster* 6

Untuk mendapatkan gambaran tentang distribusi temporal sebaran data VMS di *cluster* 6 pada Gambar 4.48 akan ditampilakan grafik distribusi data perbulan pada tahun 2016, 2017, dan 2018, serta diagram batang perbandingan aktifitas kapal penangkap ikan pada Gambar 4.49.



Gambar 4. 48 Grafik Distribusi Data *Cluster* 6

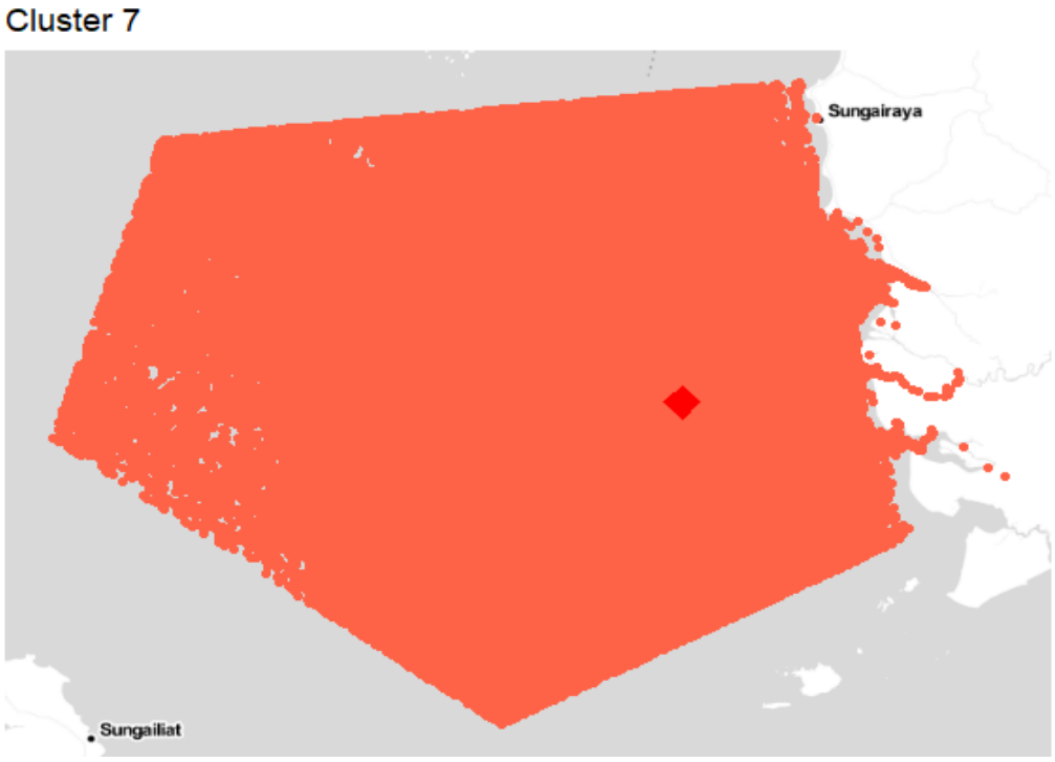


Gambar 4. 49 Diagram Batang Aktivitas Kapal Perikanan *Cluster* 6

Berdasarkan Gambar 4.48 menunjukkan grafik yang merepresentasikan sumbu x sebagai waktu time series per bulan sedangkan sumbu y merepresentasikan jumlah distribusi data VMS. Untuk pengamatan *Cluster* 6 secara temporal grafik distribusi data menunjukan fluktuasi yang cukup tinggi, data terendah terjadi pada bulan Februari 2016 sedangkan data tertinggi pada bulan April 2018. Pada tahun 2016 jumlah data yang terdistribusi paling rendah pada bulan Februari dan data tertinggi terjadi pada bulan Mei dan Juni, bahkan pada bulan Juni jumlah data lebih tinggi dibandingkan tahun 2017 dan 2018 pada bulan yang sama. Sedangkan pada tahun 2017 data terendah pada bulan Januari dan kenaikan tertinggi pada bulan November dan pada bulan tersebut jumlah data lebih tinggi dibandingkan tahun 2016 dan 2018 pada bulan yang sama. Pada tahun 2018 frekuensi distribusi data sangat fluktuatif, jumlah distribusi data tertinggi terjadi pada bulan April dan terendah pada bulan Januari dan Februari. Dari Gambar 4.49 dapat dilihat perbandingan jumlah data aktifitas kapal perikanan berdasarkan variable SPEED, dengan jumlah data 1.114.681 kategori *fishing* dan 463.321data kategori *not_fishing*.

4.3.7 Analisis *Cluster* 7

Visualisasi distribusi data VMS di *cluster* 7 dapat dilihat pada Gambar 4.50.



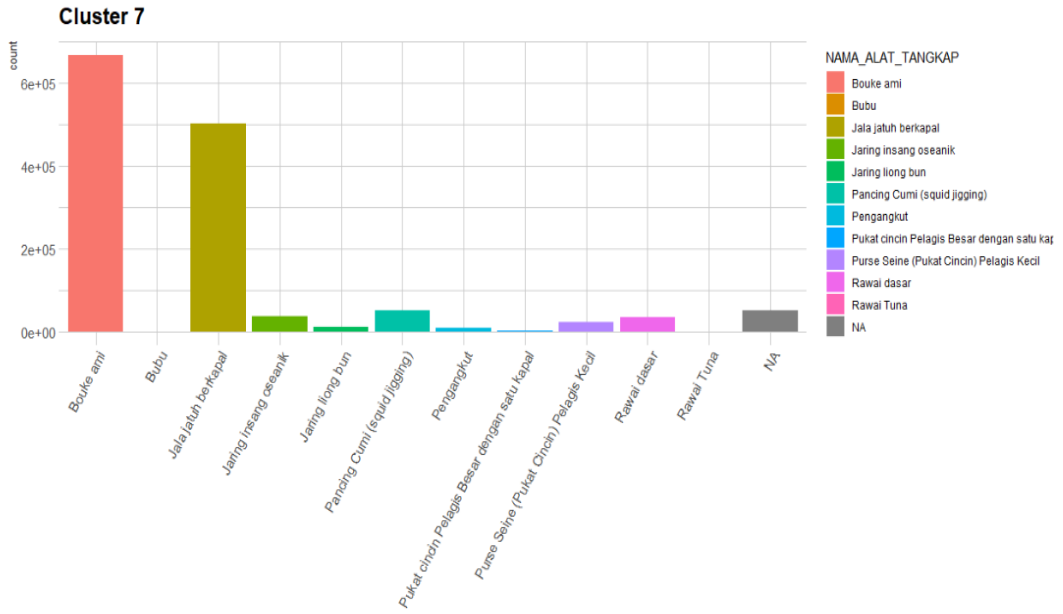
Gambar 4. 50 Visualisasi *Cluster* 7

Dari Gambar 4.50 dapat dilihat bahwa distribusi data VMS di *cluster* 7 berada di area titik LONGITUDE Min: 106.0 s/d Max:109.6 dan titik LATITUDE Min: -1.8016 s/d Max: 0.8450. Berada di wilayah Provinsi Kalimantan Barat.

Tabel 4. 15 Analisis *Cluster* 7

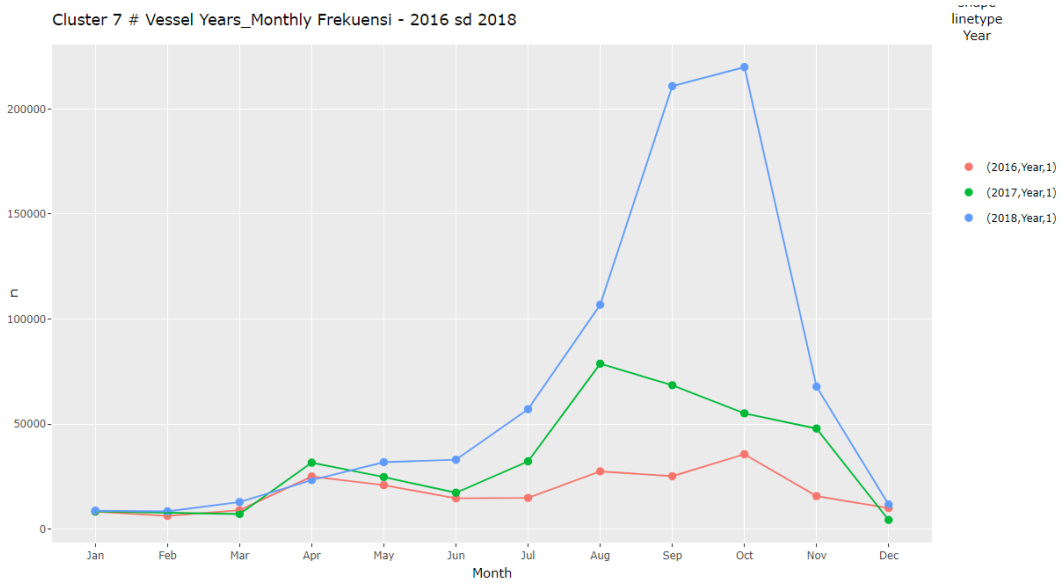
Ranking Cluster		2
Jumlah Anggota Cluster		1.388.210
Jumlah TRANSMITER_NO Aktif		1.144
NAMA_ALAT_TANGKAP		
1	Bouke ami	282
2	Bubu	2
3	Jala jatuh berkapal	398
4	Jaring insang oseanik	92
5	Jaring liong bun	33
6	Pancing Cumi (squid jigging)	58
7	Pengangkut	19
8	Pukat cincin Pelagis Besar dengan satu kapal	10
9	Purse Seine (Pukat Cincin) Pelagis Kecil	129
10	Rawai dasar	47
11	Rawai Tuna	2
12	NA	72

Tabel 4.15 merupakan tabel hasil analis *cluster* 7 yang berisikan karakteristik dan distribusi data *vessel monitoring system (VMS)* berdasarkan hasil clustering. Cluster 7 merupakan cluster yang memiliki jumlah anggota terbanyak kedua dengan jumlah 1.388.210 data VMS dan jumlah TRANSMITER_NO aktif berjumlah 1.144 unit. Berdasarkan alat tangkap yang dipergunakan alat tangkap Jala Jatuh Berkapal dipergunakan oleh 398 kapal, alat Bouke Ami digunakan oleh 287 kapal dan alat Purse Seine (Pukat Cincin) Pelagis Kecil digunakan oleh 129 kapal serta ada 330 kapal menggunakan jenis alat tangkap lainnya. Sedangkan untuk melihat jumlah distribusi data berdasarkan alat tangkap yang dipergunakan dapat melihat Gambar 4.51.

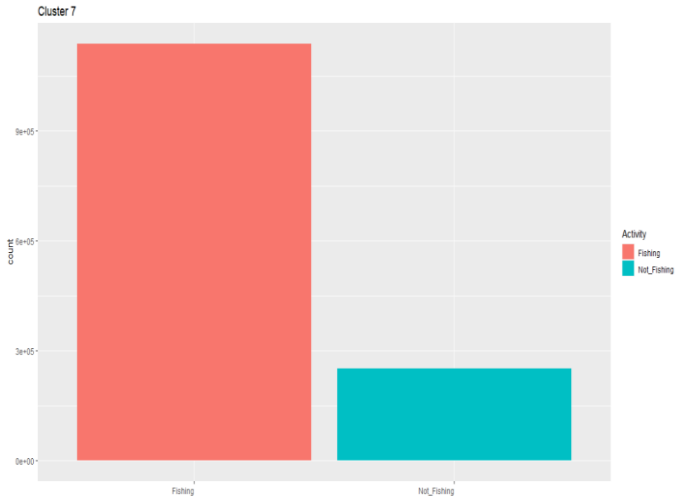


Gambar 4. 51 Distribusi Data Berdasarkan Nama Alat Tangkap *Cluster* 7

Untuk mendapatkan gambaran tentang distribusi temporal sebaran data VMS di *cluster* 7 pada Gambar 4.52 akan ditampilakan grafik distribusi data perbulan pada tahun 2016, 2017, dan 2018, serta diagram batang perbandingan aktifitas kapal penangkap ikan pada Gambar 4.53.



Gambar 4. 52 Grafik Distribusi Data *Cluster* 7



Gambar 4. 53 Diagram Batang Aktivitas Kapal Perikanan *Cluster* 7

Berdasarkan Gambar 4.52 menunjukkan grafik yang merepresentasikan sumbu x sebagai waktu time series per bulan sedangkan sumbu y merepresentasikan jumlah distribusi data VMS. Untuk pengamatan *Cluster* 7 secara temporal grafik distribusi data menunjukan bahwa pada tahun 2016 jumlah data yang terdistribusi hanya sedikit dari bulan Januari s/d Desember tanpa mengalami perubahan signifikan, sedangkan pada tahun 2017 terjadi sedikit kenaikan pada bulan Agustus, September, Oktober dan November. Pada tahun 2018 frekuensi distribusi data sangat fluktuatif, jumlah distribusi data tertinggi terjadi pada bulan September dan Oktober dan terendah rata-rata pada bulan Januari, Februari, Maret dan Desember. Dari Gambar 4.53 dapat dilihat perbandingan jumlah data aktifitas kapal perikanan berdasarkan variable SPEED, dengan jumlah data 1.137.994 kategori *fishing* dan 250.216 data kategori *not_fishing.*

BAB V
KESIMPULAN DAN SARAN

5.1 Kesimpulan

Kesimpulan yang dapat diambil dari penelitian ini adalah sebagai berikut:

1. Jumlah k optimum yang diperoleh dengan *Elbow Method* adalah 7, hasil *K-Means clustering* dengan jumlah k sama dengan 7 diperoleh hasil persentase 90.7% yaitu tingkat kemiripan setiap anggota dalam satu *cluster* atau nilai yang mewakili kekompakan hasil *clustering.* Sedangkan validasi hasil *clustering* dengan uji *Silhoutte Coefficient* diperoleh nilai 0.6677 berdasarkan tabel nilai *Silhouette Coefficient* dari Kaufman dan Rousseeuw termasuk kategori *Reasonable /Medimum Structure* atau cukup baik.
2. Hasil pengujian menunjukkan bahwa secara spatial berdasarkan data *vessel monitoring system (VMS)* kapal perikanan terbanyak beroperasi di *Cluster* 6 yang masuk ke wilayah Provinsi Kepulauan Riau, lokasi kedua berada di *Cluster* 7 yang masuk ke wilayah Provinsi Kalimantan Barat, sedangkan lokasi yang paling sedikit distribusinya berada di *Cluster* 1 yang masuk kedalam wilayah Provinsi Kepulauan Bangka Belitung, Provinsi Sumatera Selatan dan Provinsi Jambi.
3. Secara spatial, jumlah distribusi sebaran data VMS di setiap *cluster* tidak berbanding lurus dengan jumlah kapal yang aktif di daerah tersebut. Seperti di *Cluster* 6 yang memiliki distribusi data VMS terbanyak dengan jumlah 1.578.002 data hanya terdapat 428 kapal perikanan yang aktif, sedangkan di *Cluster* 1 dengan jumlah sebaran data VMS paling sedikit yaitu 330.996 data, terdapat 735 kapal perikanan aktif dan paling banyak berada di *Cluster* 7 dengan 1.144 buah kapal perikanan yang aktif. Hal ini disebabkan oleh pengaruh karakteristik oseanografi di perairan Indonesia yang berhubungan dengan pola distribusi ikan [64]. Sehingga mempengaruhi intensitas dan lamanya kapal beraktifitas menangkap ikan di wilayah *cluster* tertentu. Juga dipengaruhi oleh lokasi *cluster* tersebut, apakah lokasinya sebagai area penangkapan ikan,

sebagai wilayah lintasan kapal perikanan atau sebagai *homebase* dari kapal perikanan.

4. Secara temporal, masing-masing *cluster* memiliki pola distribusi data yang berbeda. Sedangkan setiap *cluster* berdasarkan pola distribusi bulanan yang terbentuk pada tahun 2016, 2017 dan 2018 cenderung mirip. Perubahan pola yang signifikan terjadi pada *Cluster* 1 dimana pada bulan Maret 2018 terjadi peningkatan yang tinggi dibandingkan tahun sebelumnya. Secara umum hampir semua *cluster* mengalami kenaikan distribusi jumlah data periode bulanan dari tahun 2016 s/d 2018. *Cluster* 7 mengalami kenaikan paling signifikan pada bulan September dan Oktober 2018 dan penurunan yang cukup besar hanya terjadi pada *Cluster* 5 pada bulan Agustus, September dan Oktober tahun 2018 dibandingkan tahun 2017.
5. Secara temporal, berdasarkan *data vessel monitoring system (VMS)* kapal perikanan pada tahun 2016 beroperasi terbanyak dibulan Mei dan paling sedikit di bulan Februari, di tahun 2017 terbanyak dibulan Oktober dan paling sedikit di bulan Januari, dan pada tahun 2018 terbanyak dibulan Oktober dan paling sedikit dibulan Februari, sedangkan secara keseluruhan data VMS tahun 2016 s/d 2018 terbanyak di bulan Oktober terdapat 817.725 data dan paling sedikit dibulan Februari sejumlah 241.747 data.
6. Fluktuasi kenaikan dan penurunan sebaran data VMS secara temporal dipengaruhi oleh cuaca dan musim di laut. Rendahnya distribusi data VMS pada bulan Januari, Februari dan penurunan yang signifikan pada bulan Desember di *Cluster* 3,4,5,6,7 disebabkan oleh angin Monsun Asia. Pada bulan tersebut merupakan periode Monsun Asia di perairan laut Natuna dan Selat Karimata yang rawan akan gelombang tinggi [61], dengan tinggi gelombang diatas 30cm dengan periode sekitar 5 detik dengan gelombang puncak dicapai dengan periode 10 detik yang umumnya terjadi pada bulan Desember [62]. Selanjutnya gelombang tertinggi (Hs Abs) di Laut Cina Selatan dapat mencapai ketinggian 7 meter pada bulan Desember dan Januari [61].

5.2 Saran

Adapun saran yang bisa diberikan untuk penelitian selanjutnya berdasarkan hasil yang didapat dari penelitian ini adalah sebagai berikut:

1. Penelitian ini menggunakan data *vessel monitoring system (VMS)* yang diperoleh dari Kementerian Kelautan dan Perikanan Republik Indonesia. Untuk penelitian lebih lanjut, bisa menambahkan data set lainnya sebagai pelengkap dan pembanding.
2. Penelitian selanjutnya dapat menggunakan teknik *data mining* atau a*lgoritma clustering* lainnya untuk mendapatkan informasi yang berbeda.

Halaman ini sengaja dikosongkan

DAFTAR PUSTAKA

[1] K. F. Rahmantya, A. D. Asianto, T. Wahyuni, D. Wibowo, and A. K. Zunianto, *Buku Pintar Kementerian Kelautan dan Perikanan 2018*. Jakarta: Pusat Data, Statistik dan Informasi KKP, 2018.

[2] Kementerian Kelautan dan Perikanan Republik Indonesia, *Laut Masa Depan Bangsa*. Jakarta: PT. Kompas Media Nusantara, 2018.

[3] Peraturan Menteri Kelautan Dan Perikanan Republik Indonesia, "Nomor 18/PERMEN-KP/2014 Tentang Wilayah Pengelolaan Perikanan Negara Republik Indonesia." 2014.

[4] Keputusan Menteri Kelautan dan Perikanan Republik Indonesia, "Nomor 47/KEPMEN-KP/2016 Tentang Estimasi Potensi, Jumlah Tangkapan yang Diperbolehkan, dan Tingkat Pemanfaatan Sumber Daya Ikan di Wilayah Pengelolaan Perikanan Negara Republik Indonesia." 2016.

[5] Keputusan Menteri Kelautan Dan Perikanan Republik Indonesia, "NOMOR 78/ KEPMENKP/ 2016 Tentang Rencana Pengelolaan Perikanan Wilayah Pengelolaan Perikanan Negara Republik Indonesia 711," *KKP*. 2016.

[6] Bappenas, *Kajian Strategi Pengelolaan Perikanan Berkelanjutan*. Jakarta: Kementerian Perencanaan Pembangunan Nasional, 2014.

[7] R. A. Kinseng, *Konflik Nelayan*. Jakarta: Yayasan Pustaka Obor Indonesia, 2014.

[8] Bappenas, *Kajian Strategi Industrialisasi Perikanan Untuk Mendukung Pembangunan Ekonomi Wilayah*. Jakarta: Kementerian Perencanaan Pembangunan Nasional, 2016.

[9] N. Kourti *et al.*, "Integrating Remote Sensing in Fisheries Control," pp. 295–307, 2005.

[10] C. M. Mills, S. E. Townsend, S. Jennings, P. D. Eastwood, and C. A. Houghton, "Estimating High Resolution Trawl Fishing Effort From Satellite-Based Vessel Monitoring System Data," *ICES J. Mar. Sci.*, pp. 248–255, 2007.

[11] M. Natsir and B. Atmaja, "Aktivitas Penangkapan Individu Kapal Purse Seine Di Laut Maluku : Sistem Pemantauan Kapal (Vms) Dan Observer Fishing Activity of Individual Purse Seiners in Molucca Sea : Vessel Monitoring System (Vms) and Observers," *J. Lit. Perikan. Ind.*, vol. 19, no. 1, pp. 17–24, 2013.

[12] H. Gerritsen and C. Lordan, "Integrating Vessel Monitoring Systems (VMS) Data with Daily Catch Data from Logbooks to Explore the Spatial Distribution of Catch and Effort at High Resolution," *ICES J. Mar. Sci.*, vol. 68, no. 1, pp. 245–252, 2011.

[13] T. Russo, L. D'Andrea, A. Parisi, and S. Cataudella, "VMSbase: An R-Package for VMS and logbook data management and analysis in fisheries ecology," *PLoS One*, vol. 9, no. 6, 2014, doi: 10.1371/journal.pone.0100195.

[14] M. I. Marzuki, R. Garello, R. Fablet, V. Kerbaol, and P. Gaspar, "Fishing gear recognition from VMS data to identify illegal fishing activities in Indonesia," *MTS/IEEE Ocean. 2015 - Genova Discov. Sustain. Ocean Energy a New World*, no. September 2018, pp. 10–14, 2015, doi: 10.1109/OCEANS-Genova.2015.7271551.

[15] M. I. Marzuki, P. Gaspar, R. Garello, V. Kerbaol, and R. Fablet, "Fishing Gear Identification from Vessel-Monitoring-System-Based Fishing Vessel Trajectories," *IEEE J. Ocean. Eng.*, vol. 43, no. 3, pp. 689–699, 2018, doi: 10.1109/JOE.2017.2723278.

[16] J. Zhang *et al.*, "An Automatically Learning and Discovering Human Fishing Behaviors Scheme for CPSCN," *IEEE Access*, vol. 6, no. c, pp. 19844–19858, 2018, doi: 10.1109/ACCESS.2018.2817486.

[17] H. Huang *et al.*, "FVRD: Fishing vessel relations discovery through VMS trace analysis," *2018 Ocean. - MTS/IEEE Kobe Techno-Oceans, Ocean. - Kobe 2018*, pp. 3–6, 2018, doi: 10.1109/OCEANSKOBE.2018.8559157.

[18] U. Fayyad, G. Piatetsky-Shapiro, and P. Smyth, "From Data Mining to Knowledge Discovery in Databases," *AI Mag.*, vol. 17, no. 3, pp. 37–54, 1996.

[19] Y. Li, X. Sean Wang, and S. Jajodia, "Discovering temporal patterns in multiple granularities," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 2007, no. 096314, pp. 5–19, 2001, doi: 10.1007/3-540-45244-3_2.

[20] R. J. Povinelli, "Identifying Temporal Patterns for Characterization and Prediction of Financial Time Series Events," vol. 2007, no. October 2000, 2001, doi: 10.1007/3-540-45244-3.

[21] H. J. Hamilton and D. J. Randall, "Data mining with calendar attributes," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 2007, pp. 117–132, 2001, doi: 10.1007/3-540-45244-3_10.

[22] K. Zeitouni, L. Yeh, and M. Aufaure, "Join Indices as a Tool for Spatial Data Mining," vol. 2007, no. May 2014, 2001, doi: 10.1007/3-540-45244-3.

[23] M. Santos and L. Amaral, "Knowledge Discovery in Spatial Databases : the P ADRÃO ' s qualitative approach," pp. 33–50.

[24] M. Ester, H. P. Kriegel, and J. Sander, "Spatial data mining: A database approach," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 1262 LNCS, pp. 47–66, 1997, doi: 10.1007/3-540-63238-7_24.

[25] R. Yang, K.-S. Yang, M. Kafatos, and X. S. Wang, “Value Range Queries on Earth Science Data via Histogram Clustering,” *Lect. Notes Comput. Sci. ,Lecture Notes Artif. Intell.*, vol. 2007, pp. 62–76, 2001.

[26] V. Estivill-Castro and I. Lee, “AUTOCLUSt+: Automatic clustering of point-data sets in the presence of obstacles,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 2007, pp. 133–146, 2001, doi: 10.1007/3-540-45244-3_11.

[27] L. Jay and D. W. S. Wong, *Statiscal Analysis with ArcView Gis*. Canada: John Wiley & Sons, Inc, 2001.

[28] H. J. Hamilton, L. Geng, L. Findlater, and D. J. Randall, “Spatio-temporal data mining with expected distribution domain generalization graphs,” *Proc. Int. Work. Temporal Represent. Reason.*, vol. 2003-Janua, pp. 181–191, 2003, doi: 10.1109/TIME.2003.1214895.

[29] T. Bittner, “Rough Sets in Spatio-Temporal Data Mining,” vol. 2007, no. September, 2001, doi: 10.1007/3-540-45244-3.

[30] W. Klösgen, “Deviation and association patterns for subgroup mining in temporal, spatial, and textual data bases,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 1424, pp. 1–18, 1998, doi: 10.1007/3-540-69115-4_1.

[31] J. F. Roddick and M. Spiliopoulou, “A bibliography of temporal, spatial and spatio-temporal data mining research,” *ACM SIGKDD Explor. Newsl.*, vol. 1, no. 1, pp. 34–38, 1999, doi: 10.1145/846170.846173.

[32] T. L. K. Liu, P. Krishnakumari, and O. Cats, “Exploring demand patterns of a ride-sourcing service using spatial and temporal clustering,” *MT-ITS 2019 - 6th Int. Conf. Model. Technol. Intell. Transp. Syst.*, no. June, 2019, doi: 10.1109/MTITS.2019.8883312.

[33] U. Gupta, K. Jitkajornwanich, R. Elmasri, and L. Fegaras, “Adapting K-means clustering to identify spatial patterns in storms,” *Proc. - 2016 IEEE Int. Conf. Big Data, Big Data 2016*, no. December, pp. 2646–2654, 2016, doi: 10.1109/BigData.2016.7840907.

[34] M. Vespe, M. Gibin, A. Alessandrini, F. Natale, F. Mazzarella, and G. C. Osio, “Mapping EU fishing activities using ship tracking data,” *J. Maps*, vol. 12, pp. 520–525, 2016.

[35] F. Hong *et al.*, “OceanRoads: Fishing area separation and relation discovery through VMS trace analysis,” *Ocean. 2017 - Aberdeen*, vol. 2017-Octob, pp. 1–6, 2017, doi: 10.1109/OCEANSE.2017.8084611.

[36] C. M. Mills, S. E. Townsend, S. Jennings, P. D. Eastwood, and C. A. Houghton, “Estimating high resolution trawl fishing effort from satellite-based vessel monitoring system data,” *ICES J. Mar. Sci.*, vol. 64, no. 2, pp. 248–255, 2007, doi: 10.1093/icesjms/fsl026.

[37] D. R. Mullowney and E. G. Dawe, “Development of performance indices for the Newfoundland and Labrador snow crab (Chionoecetes opilio) fishery using data from a vessel monitoring system,” *Fish. Res.*, vol. 100, no. 3, pp. 248–254, 2009, doi: 10.1016/j.fishres.2009.08.006.

[38] Pusat Bahasa Departemen Pendidikan Nasional, *Kamus Besar Bahasa Indonesia*, 4th ed. Jakarta: Pusat Bahasa, 2008.

[39] Undang Undang Republik Indonesia, “Nomor 45 Tahun 2009 Tentang Perubahan Atas Undang Undang Nomor 31 Tahun 2004 Tentang Perikanan.” 2009.

[40] Peraturan Menteri Kelautan Dan Perikanan Republik Indonesia, “Nomor 42/PERMEN-KP/2015 Tentang Sistem Pemantauan Kapal Perikanan.” 2015.

[41] Undang Undang Republik Indonesia, “Nomor 31 Tahun 2004 Tentang Perikanan.” 2004.

[42] Peraturan Menteri Kelautan Dan Perikanan Republik Indonesia, “Nomor PER.05/MEN/2008 Tentang Usaha Perikanan Tangkap.” 2008.

[43] Peraturan Menteri Kelautan Dan Perikanan Republik Indonesia, “Nomor 10/PERMEN-KP/2019 Tentang Sistem Pemantauan Kapal Perikanan.” 2019.

[44] M. J. A. Berry dan G. S. Linoff, *Mastering Data Mining The Art and Science of Customer Relationship Management*. Canada: John Wiley & Sons, Inc, 2000.

[45] Suyanto, *Data Mining Untuk Klasifikasi dan Klasterisasi Data*. Bandung: Informatika bandung, 2019.

[46] J. Han and M. Kamber, “Data Mining: Concepts and Techniques (Solution Manual),” *San Fr. CA, itd Morgan Kaufmann*, pp. 1–312, 2006, [Online]. Available: http://link.springer.com/10.1007/978-3-642-19721-5.

[47] A. L. Samuel, “Some Studies in Machine Learning Using the Game of Checkers,” *IBM J.*, pp. 210–229, 1959.

[48] Suyanto, *Machine Learning*. Bandung: Informatika bandung, 2018.

[49] P.-N. Tan, M. Steinbach, and V. Kumar, *Introduction to Data Mining*. Boston: Pearson Education, Inc., 2006.

[50] C. Fraley, “How Many Clusters? Which Clustering Method? Answers Via Model-Based Cluster Analysis,” *Comput. J.*, vol. 41, no. 8, pp. 578–588, 1998.

[51] O. Maimon and L. Rokach, *Data mining and knowledge discovery*, Second Edi. New York: Springer, 2010.

[52] P. N. Tan, M. Steinbach, and V. Kumar, *Introduction to Data Mining*. Boston: Pearson Addison-Wesley, 2006.

[53] R. Primartha, *Belajar Machine Learning*. Bandung: Informatika bandung, 2018.

[54] B. Purnima and K. Arvind, "EBK-Means: A Clustering Technique based on Elbow Method and K-Means in WSN," *Int. J. Comput. Appl.*, vol. 105, no. 9, pp. 17–24, 2014, [Online]. Available: https://www.ijcaonline.org/archives/volume105/number9/18405-9674.

[55] T. M. Kodinariya and P. R. Makwana, "Review on determining number of Cluster in K-Means Clustering," *Int. J. Adv. Res. Comput. Sci. Manag. Stud.*, vol. 1, no. 6, pp. 2321–7782, 2013.

[56] B. Santosa and A. Umam, *Data Mining dan Big Data Analytics*, 2nd ed. Yogyakarta: Penebar Media Pustaka, 2018.

[57] L. Kaufman and P. J. Rousseeuw, *Finding Groups in Data An Introduction to Cluster Analysis*. New York: John Wiley & Sons, Inc, 1990.

[58] C. Yuan and H. Yang, "Research on K-Value Selection Method of K-Means Clustering Algorithm," *J*, vol. 2, no. 2, pp. 226–235, 2019, doi: 10.3390/j2020016.

[59] A. Kassambara, *Practical Guide To Cluster Analysis in R*. sthda.com, 2017.

[60] J. Han, M. Kamber, and J. Pei, *Data Mining Concepts and Techniques*, Third Edit. 225 Wyman Street, Waltham, MA 02451, USA: Morgan Kaufmann Publishers is an imprint of Elsevier, 2012.

[61] R. Kurniawan, M. N. Habibie, and D. S. Permana, "Kajian Daerah Rawan Gelombang Tinggi Di Perairan Indonesia," *J. Meteorol. dan Geofis.*, vol. 13, no. 3, pp. 201–212, 2012, doi: 10.31172/jmg.v13i3.135.

[62] B. Pigawati, "Identifikasi Potensi dan Pemetaan Sumberdaya Pesisir Pulau - Pulau Kecil dan Laut Kabupaten Natuna - Provinsi Kepulauan Riau," *Ilmu Kelaut. - Indones. J. Mar. Sci.*, vol. 10, no. 4, pp. 229–236, 2005, doi: 10.14710/ik.ijms.10.4.229-236.

[63] B. Nababan and K. Simamora, "Variabilitas Konsentrasi Klorofil-a dan Suhu Permukaan Laut di Perairan Natuna," *J. Ilmu dan Teknol. Kelaut. Trop.*, vol. 4, no. 1, pp. 121–134, 2012.

[64] C. N. Cahya, D. Setyohadi, and D. Surinati, "Pengaruh Parameter Oseanografi terhadap Distribusi Ikan," *Oseana*, vol. 41, no. 4, pp. 1–14, 2016.

[65] H. O. Fock, "Fisheries in the context of marine spatial planning: Defining principal areas for fisheries in the German EEZ," *Mar. Policy*, vol. 32, no. 4, pp. 728–739, 2008, doi: 10.1016/j.marpol.2007.12.010.

Halaman ini sengaja dikosongkan

BIOGRAFI PENULIS

SUNARMO, lahir pada tanggal 25 oktober 1982 di Koto Agung, Kab. Dharmasraya, Sumatera Barat. Anak pertama dari dua bersaudara ini memulai pendidikan formal di SDN 16 Sungai Kalang tamat tahun 1995, masuk Madrasah Tsanawiyah Negeri Solok tamat tahun 1998, meneruskan ke SMK N 3 Solok Program Studi Mekanik Otomotif tamat tahun 2001. Kemudian melanjutkan pendidikan di Universitas Maritim Raja Ali Haji (UMRAH) Tanjungpinang pada Program Studi Ilmu Administrasi Negara, lulus tahun 2015.

Saat ini penulis bekerja sebagai Pegawai Negeri Sipil di Lembaga Pemasyarakatan Kelas IIA Tanjungpinang. Sebelumnya penulis pernah bekerja dibeberapa perusahaan swasta diantaranya PT. Tidaar Kerinci Agung-Jambi, PT. Sindo Bintan Precast-Bintan, dan Epcot Design Engineering-Tanjungpinang.

Alhamdulillah, pada tahun 2018 dengan beasiswa dari Kementerian Komunikasi dan Informatika penulis berkesempatan melanjutkan studi Pascasarjana di Institut Teknologi Sepuluh Nopember, pada Fakultas Teknologi Elektro dan Informatika Cerdas, Departemen Teknik Elektro, Bidang Keahlian TELEMATIKA dan lulus tahun 2020.

Untuk bisa berkomunikasi dengan penulis terkait dengan penelitian yang telah dikerjakan, dapat dihubungi melalui Email: sunarmo.umrah@gmail.com.