

ANALISIS DATA PENUMPANG KAPAL PROVINSI DKI JAKARTA MENGGUNAKAN METODE CLUSTERING K-MEANS

**Berdasarkan Portal Data Pemerintah Provinsi DKI Jakarta
Tahun 2018 - 2021**

SKRIPSI

**Diajukan Untuk Melengkapi Salah Satu Syarat
Memperoleh Gelar Strata 1**



Oleh :

Rani Febrianti

09011282025037

**JURUSAN SISTEM KOMPUTER
FAKULTAS ILMU KOMPUTER
UNIVERSITAS SRIWIJAYA**

2024

LEMBAR PENGESAHAN

**ANALISIS DATA PENUMPANG KAPAL PROVINSI
DKI JAKARTA MENGGUNAKAN METODE
CLUSTERING K-MEANS**

Berdasarkan Portal Data Pemerintah Provinsi DKI Jakarta
Tahun 2018 - 2021

SKRIPSI

**Diajukan Untuk Melengkapi Salah Satu Syarat
Memperoleh Gelar Strata 1**

Oleh :

RANI FEBRIANTI

09011282025037

Palembang, 30 Mei 2024

Mengetahui,

Ketua Jurusan Sistem Komputer



Dr. Ir. Sukemi, M.T.

NIP. 196612032006041001

Pembimbing Tugas Akhir

A handwritten signature in blue ink, appearing to be 'Rossi', written over the name of the supervisor.

Dr. Rossi Passarella, M.Eng.

NIP. 197806112010121004

LEMBAR PERSETUJUAN

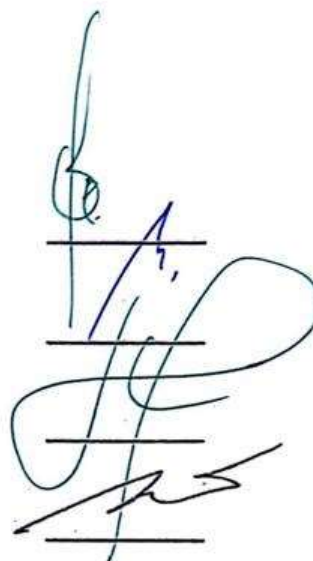
Telah diuji dan lulus pada :

Hari : Rabu

Tanggal : 15 Mei 2024

Tim Penguji :

1. Ketua : Sutarno, M.T.
2. Sekretaris : Adi Hermansyah, S.Kom., M.T.
3. Penguji : Huda Ubaya, M.T.
4. Pembimbing I : Dr. Rossi Passarella, M.Eng.



Mengetahui, 30/5/24
Ketua Jurusan Sistem Komputer



Dr. Ir. Sukemi, M.T.
NIP. 196612032006041001

HALAMAN PERNYATAAN

Yang bertanda tangan di bawah ini :

Nama : Rani Febrianti

NIM : 09011282025037

Judul : Analisis Data Penumpang Kapal Provinsi DKI Jakarta
Menggunakan Metode Clustering K-Means

Hasil pengecekan *Software Turnitin* : 4%

Menyatakan bahwa Laporan Skripsi saya merupakan hasil karya sendiri bukan hasil penjiplakan atau plagiat. Apabila ditemukan unsur penjiplakan atau plagiat dalam Laporan Skripsi ini, maka saya bersedia menerima sanksi akademik dari Universitas Sriwijaya. Demikian, pernyataan ini saya buat dalam keadaan sadar dan tidak dipaksakan.



Indralaya, 29 Mei 2024



Rani Febrianti

KATA PENGANTAR



Assalamu'alaikum Warahmatullahi Wabarakatuh

Puji dan syukur atas kehadiran Allah Swt yang telah melimpahkan Rahmat dan karunia-Nya, sehingga penulis dapat menyelesaikan penyusunan proposal tugas akhir ini dengan judul **“Analisis Data Penumpang Kapal Provinsi DKI Jakarta Menggunakan Metode Clustering K-Means”**.

Dalam laporan ini, penulis akan menjelaskan implementasi metode Clustering K-Means pada data penumpang kapal di Jakarta. Penulis berharap tulisan ini dapat bermanfaat bagi banyak orang, dan menjadi sumber pengetahuan tambahan bagi yang tertarik dalam analisis data.

Pada penyusunan laporan ini, penulis banyak mendapatkan ide, saran, dan bantuan dari berbagai pihak. Oleh karena itu, pada kesempatan ini penulis mengucapkan rasa syukur kepada Allah SWT dan mengucapkan terima kasih kepada :

1. Allah SWT., yang telah memberikan rahmat dan karunia-Nya sehingga penyusunan skripsi ini dapat berjalan dengan lancar.
2. Yang teristimewa, orang tua penulis. Ayahanda (Suwito), Ibunda (Buntari) yang menjadi saksi perjuangan penulis dalam menyelesaikan pendidikan dan penulisan skripsi ini. Ayah, terima kasih sudah menjadi ayah hebat yang walau tidak berkesempatan menempuh pendidikan di perguruan tinggi, namun ayah hebat karena mampu menuntun anaknya sampai meraih gelar sarjana. Ibu, penulis yakin semua kemudahan dalam proses meraih gelar ini bisa penulis dapat karena doa darimu. Terima kasih sudah selalu mengupayakan yang terbaik untuk penulis. Tidak banyak kata yang bisa penulis ungkapkan untuk mendeskripsikan kata terima kasih kepada orang tua penulis. Ayah, Ibu semoga kita selalu dalam lindungan Allah SWT, aamiin.
3. Almh. Nenek penulis. Nek terima kasih untuk semua Pelajaran hidup yang diberikan. Nek, sedari rani kecil nenek selalu menyampaikan pentingnya

pendidikan, selalu menjadi orang yang paling bangga dengan setiap pencapaian penulis, serta yang selalu menjadi contoh penulis dalam menjalani kehidupan. Nek, gelar ini rani hadiahkan sebagai salah satu bentuk terima kasih dan keberhasilan nenek dalam mendidik rani. Nek walau sekarang sudah tidak lagi mendampingi secara fisik, tapi rani yakin doa nenek adalah salah satu hal penting atas kemudahan yang rani peroleh selama proses mendapatkan gelar ini karena nenek langsung memintanya dari sisi Allah. Semoga nenek selalu diberikan tempat terbaik di sisi Allah SWT, aamiin.

4. Keluarga besar penulis, terutama kakak penulis Anisa Kurniasari, Eguh Susilo, dan Abim kesayangan *aunty*. Terima kasih untuk dukungan dan doa-doa baiknya hingga akhirnya penulis dapat menyelesaikan penulisan skripsi ini. Abim, terima kasih sudah menjadi penyemangat *aunty* untuk menyelesaikan pendidikan ini, biar segera jadi *rich aunty* untuk Abim.
5. Bapak Prof. Dr. Erwin, S.Si., M.Si. selaku Dekan Fakultas Ilmu Komputer Universitas Sriwijaya.
6. Bapak Dr. Ir. H. Sukemi, M.T selaku Ketua Jurusan Sistem Komputer Fakultas Ilmu Komputer Universitas Sriwijaya.
7. Bapak Dr. Rossi Passarella, M.Eng., selaku Pembimbing Tugas Akhir. Bapak, terima kasih banyak untuk semua ilmu, bantuan, serta masukan yang bapak berikan selama penulisan skripsi ini. Pak, saya tidak tau harus dengan apa membalas semua kebaikan bapak. Mungkin saya tidak mampu untuk membalas semua hal baik yang sudah bapak berikan, namun semoga Allah membalas semua hal baik yang telah bapak lakukan. Semoga bapak dan keluarga selalu dilimpahkan kesehatan dan kebaikan, serta kelancaran rezeki. Pak, terima kasih banyak. Saya sangat beruntung bisa dibimbing dan dibantu oleh bapak. Semoga ilmu yang bapak berikan menjadi amal jariyah bagi bapak, aamiin.
8. Bapak Huda Ubaya, S.T., M.T selaku Pembimbing Akademik Jurusan Sistem Komputer.
9. Bapak/Ibu Dosen, staff pengajar, dan tata usaha Fakultas Ilmu Komputer Universitas Sriwijaya.
10. Mbak Renny Virgasari dan Kak Yopi Syaputra selaku Admin Jurusan Sistem Komputer yang telah membantu penulis dalam hal administrasi.

11. Kak Dion yang selalu meluangkan waktu untuk membantu penulis dalam menyelesaikan tugas akhir.
12. Teman-teman seperjuangan Jurusan Sistem Komputer Angkatan 2020.
13. Rani Febrianti, *last but not least*, ya! diri saya sendiri. Apresiasi sebesar-besarnya karena telah bertanggung jawab untuk menyelesaikan apa yang telah dimulai. Terima kasih karena terus berusaha dan tidak menyerah, serta senantiasa menikmati setiap prosesnya yang bisa dibilang tidak mudah. Terima kasih sudah bertahan, untuk kedepannya ayo berkembang lebih jauh, lebih indah, dan lebih ceria lagi. *You've been great for getting this far, don't let irresponsible people hurt u! Someday you'll find that you're brighter than star. Just be strong, just be brave, and be sure. Yes you can!!*
14. Dan seluruh pihak yang tidak dapat penulis sebutkan satu persatu, yang telah memberikan semangat serta doa.

Penulis menyadari bahwa laporan ini masih terdapat banyak kekurangan, oleh karena itu penulis mengharapkan kritik dan saran yang membangun agar dapat menjadi lebih baik lagi di masa yang akan datang.

Akhir kata dengan segala keterbatasan, penulis berharap laporan ini dapat memberikan manfaat bagi semua pihak, terutama mahasiswa Fakultas Ilmu Komputer Universitas Sriwijaya, sebagai sumbangan pikiran dalam meningkatkan mutu pembelajaran.

Wassalamu'alaikum Warahmatullahi Wabarakatuh.

Palembang, 25 Mei 2024



Rani Febrianti

ANALISIS DATA PENUMPANG KAPAL PROVINSI DKI JAKARTA MENGGUNAKAN METODE CLUSTERING K-MEANS

Rani Febrianti (09011282025037)

Jurusan Sistem Komputer, Fakultas Ilmu Komputer, Universitas Sriwijaya

Email : ranifebrianti18@gmail.com

ABSTRAK

Pelabuhan berperan penting dalam konektivitas antarpulau dan layanan transportasi laut. Penelitian ini menggunakan metode K-Means untuk mengidentifikasi pola dan mendapatkan wawasan dari data historis. Hasil dari penelitian ini disajikan dalam bentuk tiga kelompok kapal penumpang dengan tingkat jumlah penumpang dan kapal yang berbeda, mengelompokkannya menjadi tiga kelompok berbeda: pelabuhan dengan kemampuan rendah, menengah, dan tinggi. Kelompok-kelompok ini telah diidentifikasi sebagai titik-titik fokus bagi upaya pemerintah untuk meningkatkan layanan transportasi di wilayah tersebut. Temuan ini juga menyoroti pentingnya transportasi laut dalam mendukung konektivitas dan industri pariwisata di Kabupaten Kepulauan Seribu. Analisis ini memberikan dasar bagi pengambilan keputusan berbasis data dalam manajemen pelabuhan di masa depan dan mendorong pemangku kepentingan untuk meningkatkan layanan di pelabuhan yang teridentifikasi.

Kata Kunci: Manajemen Pelabuhan; Transportasi Laut; Pengelompokan K-Means

ANALYSIS OF SHIP PASSENGER DATA IN DKI JAKARTA PROVINCE USING THE K-MEANS CLUSTERING METHOD

Rani Febrianti (09011282025037)

Dept. Of Computer Engineering, Faculty of Computer Science, Sriwijaya University

Email : ranifebrianti18@gmail.com

ABSTRACT

Port infrastructure plays an important role in supporting inter-island connectivity and marine transportation services. In this study, we used the K-means clustering method to identify patterns and gain insights from historical data. The results of this study are presented in the form of three clusters of passenger ship with different levels of passenger numbers and departing ship, grouping them into three different clusters: low capability ports, medium capability ports, and high capability ports. These clusters have been identified as focal points for government efforts to improve transportation services in the region. In addition, the findings underscore the critical role of marine transportation in facilitating connectivity and supporting the tourism industry in the Thousand Islands Regency. This analysis not only provides a comprehensive understanding of the existing situation but also provides a basis for informed decision-making in future port management strategies. The study serves as a call to action for stakeholders and policymakers to prioritize service improvements at the identified ports.

Keywords: Port Management; Marine Transportation; K-means Clustering

DAFTAR ISI

LEMBAR PENGESAHAN	ii
LEMBAR PERSETUJUAN.....	iii
HALAMAN PERNYATAAN	iv
KATA PENGANTAR	v
ABSTRAK	viii
ABSTRACT	ix
DAFTAR ISI.....	ii
DAFTAR GAMBAR	iv
DAFTAR TABEL.....	vi
BAB 1 PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	4
1.3 Batasan Masalah.....	4
1.4 Tujuan Penelitian	5
1.5 Manfaat Penelitian	5
1.6 Metodologi Penelitian	6
1.7 Sistematika Penulisan.....	7
BAB 2 TINJAUAN PUSTAKA	8
2.1 Data Analysis	8
2.1.1 Menentukan Tujuan Analisis	8
2.1.2 Pengumpulan Data	8
2.1.3 Data Preprocessing.....	9
2.1.4 Menginterpretasikan Hasil	9
2.2 Clustering	9
2.2.1 Algoritma K-Means	10
2.2.2 Metode Elbow	11
2.2.3 Metode Silhoutte	12
2.3 Penelitian Terkait	13
2.4 Landasan Teori.....	18
2.4.1 Transportasi.....	18

2.4.2 Sejarah Transportasi Laut Indonesia	19
2.4.3 Pelabuhan	23
2.4.4 Kepulauan Seribu	28
BAB 3 METODOLOGI PENELITIAN.....	36
3.1 Kerangka Kerja Penelitian	36
3.1.1 Menentukan Tujuan Analisis dan Membuat Hipotesis	37
3.1.2 Pengumpulan Data	37
3.1.3 Data Preprocessing.....	38
3.1.4 Uji Jumlah Cluster.....	39
3.1.5 Tahap Pemodelan	40
3.1.6 Perbandingan Metode.....	41
3.1.7 Hasil dan Kesimpulan	42
3.2 Spesifikasi Perangkat Keras dan Perangkat Lunak.....	43
3.2.1 Perangkat Keras	43
3.2.2 Perangkat Lunak.....	43
BAB 4 HASIL DAN PEMBAHASAN.....	45
4.1 Data Preprocessing.....	45
4.2 Uji Jumlah Cluster.....	50
4.3 Tahap Pemodelan	52
4.3.1 Analisis Data Pada Cluster 1	54
4.3.2 Analisis Data Pada Cluster 2.....	58
4.3.3 Analisis Data Pada Cluster 3.....	59
4.4 Perbandingan Metode.....	61
4.5 Hasil dan Pembahasan.....	63
BAB 5 KESIMPULAN.....	66
5.1 Kesimpulan	66
DAFTAR PUSTAKA	67

DAFTAR GAMBAR

Gambar 2.1 Algoritma K-Means.....	11
Gambar 2.2 Kapal Pertama di Dunia	20
Gambar 2.3 Relief Kapal di Candi Borobudur.....	21
Gambar 2.4 Perahu Pinisi Tahun 1947	21
Gambar 2.5 Marine Establishment Surabaya.....	22
Gambar 2.6 Kapal Baruna Jaya VIII.....	23
Gambar 2.7 Pelabuhan Tanjung Priok	27
Gambar 2.8 Pelabuhan Lembar	27
Gambar 2.9 Pelabuhan Tanjung Emas	28
Gambar 2.10 Lokasi pelabuhan kapal penumpang berdasarkan titik GPS	29
Gambar 2.11 Pelabuhan Muara Angke	30
Gambar 2.12 Pelabuhan Marina Ancol.....	30
Gambar 2.13 Pelabuhan Untung Jawa	31
Gambar 2.14 Pelabuhan Lancang	31
Gambar 2.15 Pelabuhan Pari.....	32
Gambar 2.16 Pelabuhan Tidung.....	33
Gambar 2.17 Pelabuhan Pramuka	33
Gambar 2.18 Pelabuhan Kelapa	34
Gambar 2.19 Pelabuhan Harapan.....	34
Gambar 2.20 Pelabuhan Sabira	35
Gambar 3.1 Kerangka Kerja Penelitian.....	36
Gambar 3.2 Tampilan Website Portal Data Jakarta	38
Gambar 4.1 Data Penumpang Kapal Kepulauan Seribu	45
Gambar 4.2 Hasil <i>Software Orange</i>	46
Gambar 4.3 Hasil Korelasi Antar Variabel	46
Gambar 4.4 Feature Selection	48
Gambar 4.5 Hasil MinMaxScaler	49
Gambar 4.6 <i>Data Reduction</i>	50
Gambar 4.7 Uji Jumlah Cluster Metode Elbow	51
Gambar 4.8 Uji Jumlah Cluster Metode Silhoutte	52

Gambar 4.9 Clustering K-Means	53
Gambar 4.10 Deskriptif Data Cluster 1.....	54
Gambar 4.11 Visualisasi Data Cluster 1	55
Gambar 4.12 Deskriptif Data Cluster 2.....	58
Gambar 4.13 Visualisasi Data Cluster 2	59
Gambar 4.14 Deskriptif Data Cluster 3.....	60
Gambar 4.15 Visualisasi Data Cluster 3	61
Gambar 4.16. Clustering DBSCAN.....	62

DAFTAR TABEL

Tabel 1. Jumlah Penumpang Kapal di Kepulauan Seribu.....	2
Tabel 2. Studi terkait Penelitian tentang Transportasi	13
Tabel 3. Spesifikasi Perangkat Keras	43
Tabel 4. Spesifikasi Perangkat Lunak.....	43
Tabel 5. Nilai Skewness	48
Tabel 6. Jumlah Data Setiap Cluster K-Means	53
Tabel 7. Jumlah Persebaran Data Cluster 1.....	56
Tabel 8. Rekapitulasi Distribusi Jumlah Data Cluster 1 Berdasarkan Pelabuhan. 57	
Tabel 9. Jumlah Persebaran Data Cluster 2.....	59
Tabel 10. Jumlah Persebaran Data Cluster 3.....	61
Tabel 11. Jumlah Data Setiap Cluster DBSCAN.....	62
Tabel 12. Rekapitulasi Pelabuhan, Klaster, dan Jumlah Penumpang	64

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Indonesia adalah negara kepulauan terbesar di dunia yang terdiri dari 17.499 pulau yang membentang dari Sabang sampai Merauke. Wilayah Indonesia memiliki luas total mencapai 7.81 juta km² [1].

DKI Jakarta merupakan ibukota Republik Indonesia dan juga pusat bisnis serta pemerintahan dengan jumlah penduduk yang padat. Kota Jakarta terletak di sekitar kawasan pemukiman Bogor, Depok, Tangerang, dan Bekasi (Bodetabek) yang terus berkembang. Oleh karena itu, penting untuk memiliki transportasi yang memadai guna mendukung aktivitas perekonomian [2].

Transportasi merupakan kegiatan memindahkan atau mengangkut muatan (barang atau manusia) dari tempat asal ke tempat tujuan. Transportasi memiliki peranan yang sangat penting sebagai kebutuhan dasar dalam memenuhi kebutuhan hidup manusia, serta menjadi faktor krusial dalam perekonomian dan pembangunan [3]. Transportasi di Jakarta, baik transportasi darat, transportasi udara, dan transportasi laut memiliki peranan penting dalam menunjang aktivitas warga Jakarta dan sekitarnya [2].

Transportasi laut merupakan salah satu sarana transportasi yang mempermudah pergerakan penduduk yang semakin padat, terutama di daerah transportasi untuk memfasilitasi mobilitas penduduk yang semakin tinggi, terutama di daerah yang sulit dilalui menggunakan transportasi darat dan udara. Sejalan dengan hal ini, infrastruktur pelabuhan juga berperan penting dalam mendukung konektivitas antarpulau dan pelayanan transportasi laut. Menurut Bambang Triatmodjo [4], pelabuhan memiliki peran yang sangat penting dalam mendukung mobilitas barang maupun manusia di Indonesia yang merupakan negara kepulauan dan maritim. DKI Jakarta memiliki 22 pelabuhan laut, di mana 12 pelabuhan berada di Kota Jakarta Utara dan 10 pelabuhan berada di Kabupaten Kepulauan Seribu [2].

Kabupaten Kepulauan Seribu merupakan salah satu wilayah administrasi di bawah Pemerintah Provinsi DKI Jakarta. Dari segi geografis, Kepulauan Seribu berada di antara garis Bujur Timur $106^{\circ}19'30''$ - $106^{\circ}44'50''$ dan garis Lintang Selatan $5^{\circ}10'00''$ - $5^{\circ}57'00''$ [5].

Berdasarkan data.jakarta.go.id [6], total jumlah penumpang kapal di Kabupaten Kepulauan Seribu selama periode tahun 2018-2021 dinyatakan pada Tabel 1.

Tabel 1. Jumlah Penumpang Kapal di Kepulauan Seribu

Pelabuhan	2018	2019	2020	2021
Muara Angke	339529	312188	44956	77193
Marina Ancol	181535	152381	12870	11941
Untung Jawa	133404	113795	19380	34293
Lancang	34483	29907	6289	18414
Pari	125011	111053	15131	28476
Tidung/Payung	137025	105994	16729	17590
Pramuka/Panggang	84879	69654	13270	31104
Kelapa	36427	51603	8574	15209
Harapan	69387	66762	8857	13440
Sabira	0	3979	861	2141

Manajeme transportasi laut yang baik dan efisien sangat penting bagi Indonesia sebagai negara kepulauan guna meningkatkan daya saing ekonomi dan konektivitas di seluruh wilayah Republik Indonesia [7]. Kapal merupakan moda transportasi utama yang digunakan untuk menyeberangi lautan dari satu pulau ke pulau lainnya [8].

Jumlah penumpang kapal terus meningkat seiring waktu, terutama pada akhir pekan, hari raya, dan liburan tahun baru. Setelah pandemi Covid-19 pada tahun 2022, terjadi peningkatan signifikan sebesar 50% pada jumlah wisatawan mancanegara yang mengunjungi Kepulauan Seribu [11]. Hal ini menyebabkan terjadinya kepadatan penumpang pada pelabuhan-pelabuhan yang berada di Kepulauan Seribu. Oleh karena itu, penting untuk menerapkan pendekatan pengelompokkan berdasarkan jumlah penumpang dan kapal guna mengatasi

masalah kepadatan tersebut. Dengan cara ini, kita dapat mengidentifikasi pelabuhan mana yang paling membutuhkan penyesuaian seperti peningkatan kapasitas atau penjadwalan kapal untuk mengelola aliran penumpang secara lebih efektif.

Pendekatan ini dapat dilakukan melalui analisis data lampau. Salah satu pendekatan machine learning yang cocok untuk memahami data lampau adalah Clustering. Clustering merupakan metode unsupervised learning yang tidak memerlukan label [9], sehingga menjadi pilihan yang tepat untuk data lampau yang tidak memiliki label. Dengan menerapkan Clustering, kita dapat mengelompokkan data pelabuhan berdasarkan pola-pola yang ada dalam jumlah penumpang dan kapal pada periode lampau, yang akan membantu dalam mengidentifikasi pelabuhan-pelabuhan yang memerlukan penyesuaian untuk mengelola kepadatan penumpang dengan lebih efektif di masa mendatang.

Dalam pembelajaran unsupervised learning, K-Means merupakan salah satu metode clustering yang populer dan efisien untuk tujuan ini. Dengan menggunakan K-Means, kita dapat mengelompokkan data pelabuhan ke dalam kelompok-kelompok berdasarkan pola-pola yang terdapat dalam jumlah penumpang dan kapal pada periode lampau.

Setelah kelompok-kelompok terbentuk, kita dapat menganalisis setiap kelompok untuk mengidentifikasi pelabuhan-pelabuhan yang cenderung mengalami kepadatan penumpang. Dari analisis ini, langkah-langkah penyesuaian seperti peningkatan kapasitas atau penjadwalan kapal dapat direncanakan dengan lebih tepat guna mengelola kepadatan penumpang di masa mendatang.

Metode K-Means digunakan pada penelitian yang dilakukan oleh Zeyu Hou, Ran Yan, dan Shuaian Wang pada tahun 2022 dengan judul penelitian “On the K-Means Clustering Model for Performance Enhancement of Port State Control” [10]. Hasil penelitian menunjukkan adanya tiga cluster kapal dengan tingkat risiko yang berbeda.

Selanjutnya penelitian yang dilakukan oleh Islam Ahmed Azzam, Saleh Fahed Al-Khatib dan Waed Mohammad Albataineh pada tahun 2019, dengan judul penelitian “Strategic port classification: International clustering-based approach for decision-making optimization” [11]. Penelitian ini menggunakan metode Clustering K-Means. Hasil penelitian menunjukkan adanya dua cluster utama, yaitu

pelabuhan dengan kemampuan rendah dan pelabuhan dengan kemampuan sedang hingga tinggi.

Dengan merujuk pada beberapa penelitian pembelajaran mesin yang menggunakan metode clustering, penulis berhipotesis bahwa dengan menganalisis data lampau penumpang kapal di Kepulauan Seribu, pemerintah dapat mengidentifikasi pelabuhan mana yang mengalami kepadatan penumpang sehingga tindakan penyesuaian, seperti peningkatan kapasitas atau penyesuaian jadwal kapal, dapat dilakukan secara lebih efektif. Dengan demikian, pelabuhan-pelabuhan yang membutuhkan perhatian lebih dapat diidentifikasi dengan lebih akurat, dan langkah-langkah strategis dapat diambil untuk mengatasi kepadatan penumpang dengan lebih efisien.

1.2 Rumusan Masalah

Rumusan masalah pada penelitian ini :

1. Data penumpang kapal yang terkumpul dari Portal Data Jakarta akan dianalisis untuk menentukan jumlah cluster yang dapat dibentuk dari analisis data lampau. Analisis ini juga akan membantu mengidentifikasi pelabuhan-pelabuhan yang mengalami kepadatan penumpang, sehingga tindakan penyesuaian seperti peningkatan kapasitas atau penjadwalan kapal dapat dilakukan dengan lebih efektif menggunakan metode tertentu.
2. Jumlah cluster yang dihasilkan akan dianalisis untuk menguji kebenaran hipotesis yang telah dibuat. Selanjutnya, hasil analisis tersebut akan digunakan untuk mengidentifikasi pelabuhan-pelabuhan yang mengalami kepadatan penumpang.

1.3 Batasan Masalah

Batasan masalah penelitian ini :

1. Raw data yang digunakan diperoleh dari website Portal Data Jakarta.
2. Data yang digunakan dimulai dari tanggal 1 Januari 2018 hingga 31 Desember 2021.
3. Variabel yang akan digunakan adalah nama pelabuhan, jumlah kapal berangkat, dan jumlah penumpang naik.

4. Metode yang digunakan adalah Clustering dengan algoritma K-Means.
5. Bahasa pemrograman yang dipakai adalah Python, perangkat lunak yang digunakan untuk pengolahan data adalah Jupyter Notebook, dan untuk visualisasi data digunakan perangkat lunak Tableau.

1.4 Tujuan Penelitian

Tujuan dari penelitian ini :

1. Melakukan analisis terhadap data lampau penumpang kapal yang sebelumnya hanya dikumpulkan tanpa melalui proses pengolahan data.
2. Memperoleh wawasan baru melalui pengolahan data penumpang kapal yang digunakan.
3. Menerapkan metode Clustering K-Means dalam menganalisis data lampau penumpang kapal di Kepulauan Seribu.
4. Mengidentifikasi pelabuhan-pelabuhan yang mengalami kepadatan penumpang berdasarkan hasil analisis jumlah cluster.

1.5 Manfaat Penelitian

Penelitian ini dapat memberikan manfaat sebagai berikut :

1. Mengubah data mentah menjadi data yang lebih informatif melalui penelitian yang dilakukan, sehingga dapat menghasilkan pengetahuan dan wawasan baru.
2. Mengetahui kesimpulan dari analisis yang digunakan untuk menguji kebenaran hipotesis yang telah dibuat.
3. Implementasi metode Clustering K-Means dalam menganalisis data lampau penumpang kapal di Kepulauan Seribu dapat dilakukan.
4. Membantu Pemerintah dalam mengidentifikasi serta mengatasi masalah yang terjadi di pelabuhan yang berada di Kepulauan Seribu, sehingga tindakan penyesuaian seperti peningkatan kapasitas atau penjadwalan kapal dapat dilakukan dengan lebih efektif.

1.6 Metodologi Penelitian

1. Studi Pustaka

Langkah awal yang diambil penulis adalah mengidentifikasi isu-isu yang relevan dengan penelitian ini. Penulis melakukan pencarian dan analisis terhadap berbagai sumber seperti buku, jurnal, makalah, dan internet.

2. Perencanaan

Pada tahap perencanaan ini, penulis merencanakan pendekatan untuk memecahkan masalah dengan menentukan metode atau teknik yang akan digunakan dalam menganalisis data. Penulis memilih metode atau teknik yang tepat untuk menjawab pertanyaan penelitian yang diajukan.

3. Pengujian

Dalam tahap ini, metode atau teknik diuji untuk memeriksa keefektifannya dalam menganalisis data dari penelitian ini. Penulis melakukan pengujian untuk memperoleh hasil yang dapat digunakan dalam tahap analisis selanjutnya.

4. Analisis dan Diskusi

Dalam tahap ini, penulis menganalisis hasil dari pengujian metode yang telah diterapkan sebelumnya untuk menemukan solusi dari permasalahan penelitian. Penulis kemudian mendiskusikan hasil analisis tersebut untuk memberikan pemahaman yang lebih baik tentang topik penelitian.

5. Kesimpulan

Tahap terakhir adalah menyusun kesimpulan berdasarkan hasil analisis dan pembahasan yang telah dijelaskan pada tahap sebelumnya. Penulis menarik kesimpulan yang relevan dengan pertanyaan penelitian.

1.7 Sistematika Penulisan

Penelitian ini akan disusun sesuai dengan struktur penulisan yang terdiri dari bab-bab sebagai berikut:

BAB I. PENDAHULUAN

Bab ini akan menjelaskan secara sistematis tentang topik penelitian, mencakup latar belakang, tujuan, manfaat, perumusan masalah, batasan masalah, dan metode penelitian.

BAB II. TINJAUAN PUSTAKA

Bagian ini mencakup berbagai teori yang menjadi dasar analisis dalam penelitian kasus ini.

BAB III. METODOLOGI PENELITIAN

Bagian ini akan secara detail menjelaskan kerangka kerja penelitian serta proses penelitian dari awal hingga akhir.

BAB IV. HASIL DAN ANALISA

Bagian ini memaparkan hasil dari metode yang diterapkan dalam penelitian ini dan melakukan analisis terhadap hasil tersebut.

BAB V. KESIMPULAN

Bagian ini merupakan rangkuman dari hasil penelitian yang akan menjawab tujuan yang telah diuraikan dalam Bab I Pendahuluan.

DAFTAR PUSTAKA

- [1] C. E. Wahyuningsih, “Analisis Stakeholder Dalam Kebijakan Pelayanan Pemanduan Kapal Di Pelabuhan Tanjung Perak,” pp. 1–53.
- [2] BPS Jakarta, “Statistik Transportasi DKI Jakarta 2021,” *Yolanda Wilda Artati*, vol. 2, no. 5, p. 60, 2021, [Online]. Available: <https://nasional.tempo.co/read/1566965/integrasi-jadi-kunci-perubahan-transportasi-di-jakarta>
- [3] S. A. Adisasmita, “Pengantar Perencanaan Transportasi,” *Pustaka.Ut.Ac.Id*, 2014.
- [4] B. Triatmodjo, *Perencanaan Pelabuhan*, vol. 1999, no. December. 2009.
- [5] “Kabupaten Kepulauan Seribu.” <https://pulauseribu.jakarta.go.id/> (accessed Jun. 25, 2023).
- [6] Open Data Jakarta, “Data Penumpang Kapal dari dan ke Kepulauan Seribu .” <https://data.jakarta.go.id/dataset/data-penumpang-kapal-dari-dan-ke-kepulauan-seribu-tahun-2019-kpi> (accessed Jun. 23, 2023).
- [7] “Badan Pusat Statistik.” <https://www.bps.go.id/publication/2018/11/27/ace352a6247e3e9d4856b357/statistik-transportasi-laut-2017.html> (accessed Jun. 21, 2023).
- [8] W. Utari, M. Bettiza, and A. Uperiati, “Pemodelan Regresi Non Linear Pada Prediksi Jumlah Penumpang Kapal Laut Dengan Algoritma Genetika (Studi Kasus : Rute Pelabuhan Sri Bintan Pura Jagoh),” pp. 271–282, 2021.
- [9] C. C. Peng, C. J. Tsai, T. Y. Chang, J. Y. Yeh, and P. W. Hua, “A new approach to generate diversified clusters for small data sets,” *Appl. Soft Comput. J.*, 2020, doi: 10.1016/j.asoc.2020.106564.
- [10] Z. Hou, R. Yan, and S. Wang, “On the K-Means Clustering Model for Performance Enhancement of Port State Control,” *J. Mar. Sci. Eng.*, vol. 10, no. 11, 2022, doi: 10.3390/jmse10111608.
- [11] I. A. Azzam, S. F. Al-Khatib, and W. M. Albatineh, “Strategic port classification: International clustering-based approach for decision-making optimization,” *J. Public Aff.*, vol. 21, no. 1, 2021, doi: 10.1002/pa.1963.
- [12] A. Famili, W. M. Shen, R. Weber, and E. Simoudis, “Data preprocessing and intelligent data analysis,” *Intell. Data Anal.*, 1997, doi: 10.3233/IDA-1997-1102.
- [13] Kunal Gohrani, “The 5 Steps of the Data Analysis Process.” https://medium.com/@kunal_gohrani/the-5-steps-of-the-data-analysis-process-2512ba6ac31e (accessed Jun. 27, 2023).

- [14] E. Paradis, B. O'Brien, L. Nimmon, G. Bandiera, and M. A. T. Martimianakis, "Design: Selection of Data Collection Methods," *J. Grad. Med. Educ.*, 2016, doi: 10.4300/JGME-D-16-00098.1.
- [15] V. O. Ajayi, "Primary Sources of Data and Secondary Sources of Data Benue State University.," *Disting. Between Prim. Sources Data Second. Sources Data*, 2017.
- [16] S. Lefever, M. Dal, and Á. Matthíasdóttir, "Online data collection in academic research: Advantages and limitations," *Br. J. Educ. Technol.*, 2007, doi: 10.1111/j.1467-8535.2006.00638.x.
- [17] Sana Mushtaq, "Data preprocessing in detail." <https://developer.ibm.com/articles/data-preprocessing-in-detail/> (accessed Jun. 27, 2023).
- [18] D. Barapatre and A. Vijayalakshmi, "Data preparation on large datasets for data science," *Asian J. Pharm. Clin. Res.*, 2017, doi: 10.22159/ajpcr.2017.v10s1.20526.
- [19] S. Patel, S. Sihmar, and A. Jatain, "A study of hierarchical clustering algorithms," in *2015 International Conference on Computing for Sustainable Global Development, INDIACom 2015*, 2015.
- [20] Q. Zhu, J. Pei, X. Liu, and Z. Zhou, "Analyzing commercial aircraft fuel consumption during descent: A case study using an improved K-means clustering algorithm," *J. Clean. Prod.*, 2019, doi: 10.1016/j.jclepro.2019.02.235.
- [21] M. J. Brusco, E. Shireman, and D. Steinley, "A comparison of latent class, K -means, and K -median methods for clustering dichotomous data.," *Psychol. Methods*, 2017, doi: 10.1037/met0000095.
- [22] Y. Li and H. Wu, "A Clustering Method Based on K-Means Algorithm," *Phys. Procedia*, 2012, doi: 10.1016/j.phpro.2012.03.206.
- [23] S. Ali, T. E. Ali, M. A. Khan, I. U. Khan, and M. D. Patterson, "Effective and scalable clustering of SARS-CoV-2 sequences," in *ACM International Conference Proceeding Series*, 2021. doi: 10.1145/3505745.3505752.
- [24] R. Nainggolan, R. Perangin-Angin, E. Simarmata, and A. F. Tarigan, "Improved the Performance of the K-Means Cluster Using the Sum of Squared Error (SSE) optimized by using the Elbow Method," in *Journal of Physics: Conference Series*, 2019. doi: 10.1088/1742-6596/1361/1/012015.
- [25] M. I. Firdaus, R. D. Octaviani, and I. Yusnita, "Clustering Calon Penumpang Kereta Cepat Jakarta-Bandung," *J. Manaj. Transp. Logistik*, vol. 4, no. 2, p. 193, 2017, doi: 10.54324/j.mtl.v4i2.98.
- [26] S. Sukamto, I. D. Id, and T. R. Angraini, "Penentuan Daerah Rawan Titik Api di Provinsi Riau Menggunakan Clustering Algoritma K-Means," *JUITA J. Inform.*, 2018, doi: 10.30595/juita.v6i2.3172.

- [27] K. Handoko, "Pengelompokan Data Mining Pada Jumlah Penumpang Di Bandara Hang Nadim," *Comput. Based Inf. Syst. J.*, vol. 6, no. 2, pp. 60–68, 2018, doi: 10.33884/cbis.v6i2.708.
- [28] B. S. Praja, P. D. Kusuma, and C. Setianingsih, "Penerapan Metode K-Means Clustering Dalam Pengelompokan Data Penumpang Dan Kapal Angkutan Laut Di Indonesia," in *e-Proceeding of Engineering*, 2019.
- [29] A. Hazman, F. T. Elektro, and U. Telkom, "Penerapan Metode K-Means Clustering Untuk Mengelompokkan Data Pelabuhan Dan Bongkar Muat Barang Di Indonesia," *e-Proceeding Eng.*, 2019.
- [30] E. Dianawati, P. P. Yanti, and Y. Suryandari, "Klustering Jumlah Penumpang pada Halte Bus Rapid Transit Kota Tangerang," *J. Sist. Cerdas*, 2019, doi: 10.37396/jsc.v2i3.34.
- [31] A. Supriyatna, I. Carolina, S. Janti, and A. Haidir, "Clustering Koridor Transjakarta Berdasarkan Jumlah Penumpang Dengan Algoritma K-Means," 2020.
- [32] I. C. Indah, M. N. Sari, and M. H. Dar, "Application of the K-Means Clustering Algorithm to Group Train Passengers in Labuhanbatu," *Sinkron*, vol. 8, no. 2, pp. 825–837, 2023, doi: 10.33395/sinkron.v8i2.12260.
- [33] R. Passarella, M. D. Iqbal, M. A. Buchari, and H. Veny, "Analysis of Commercial Airplane Accidents Worldwide Using K-Means Clustering," no. November, 2023, doi: 10.18280/ijssse.130505.
- [34] S. Abbas, "Jurnal Manajemen Transportasi Dan Logistik," *Jurnal Manajemen Transportasi Dan Logistik*. 2012.
- [35] J. Malisan, "Analisis Tingkat Pelayanan Terminal Penumpang Pelabuhan Balikpapan," vol. 19, pp. 76–87, 2017.
- [36] Badan Pusat Statistik, "Statistik Transportasi Laut 2017," pp. 1–141, 2019.
- [37] A. F. Minangkabau, "Tinjauan Yuridis Mutu Keselamatan Speedboat Sebagai Moda Transportasi Laut Dihubungkan Dengan Undang-Undang Nomor 17 Tahun 2008 Tentang Pelayaran".
- [38] "Kapal Pertama di Dunia." <https://velascoindonesia.com/asal-usul-kapal-pertama-di-dunia/> (accessed Sep. 08, 2023).
- [39] P. A. Samekto and R. R. Pujiastuti, "Pemanfaatan Transportasi Laut Untuk Meningkatkan Pertumbuhan Sektor Pariwisata," *J. Sains Dan Teknol. Marit.*, vol. 19, no. 2, 2019, doi: 10.33556/jstm.v19i2.204.
- [40] "Relief Kapal di Candi Borobudur." <https://kebudayaan.kemdikbud.go.id/bkborobudur/relief-kapal/> (accessed Sep. 08, 2023).
- [41] "Perahu Pinisi Tempo Dulu." <https://galerikomunitassimpulmerah.blogspot.com/2017/12/perahu-pinisi-tahun-1947.html> (accessed Sep. 08, 2023).

- [42] “Marine Establishment in Surabaya.” <https://laststandonzombieisland.com/2022/03/02/warship-wednesday-march-2-2022-burnt-java/marine-establishment-in-surabaya-photo-taken-from-victualie-warehouse-towards-the-east-3-2-42/> (accessed Sep. 08, 2023).
- [43] “Baruna Jaya VIII.” <https://www.vesselfinder.com/vessels/details/9155171> (accessed Sep. 08, 2023).
- [44] D. A. Ayuningtias and R. Purwaningsih, “Penilaian Standar Kelayakan Pelayanan Penumpang Dan Fasilitas Di Terminal Penumpang Pelabuhan Tanjung Emas Semarang,” ... *Eng. Online J.*, 2018.
- [45] “Keputusan Menteri Perhubungan,” pp. 1–61, 2004.
- [46] R. A. Takdir and I. G. N. S. Buana, “Analisis Strategi Pengembangan Pelabuhan Kendari,” *Wave J. Ilm. Teknol. Marit.*, vol. 13, no. 1, pp. 17–24, 2019, doi: 10.29122/jurnalwave.v13i1.3576.
- [47] K. Schwab, *The Global Competitiveness Report 2017-2018*. 2018.
- [48] “Pelabuhan Tanjung Priok.” <https://nasional.okezone.com/read/2021/06/11/337/2423715/si-denok-bandarwati-nama-cantik-penuh-makna-pelabuhan-tanjung-priok> (accessed Sep. 08, 2023).
- [49] “Pelabuhan Lembar Lombok.” <https://gerbanglombok.co.id/> (accessed Sep. 08, 2023).
- [50] “Pelabuhan Tanjung Emas .” <https://news.republika.co.id/berita/ptdz1t383/pelabuhan-tanjung-emas-tambah-fasilitas-shore-connection> (accessed Sep. 08, 2023).
- [51] “Pelabuhan Muara Angke.” <https://wisatahits.blog/pelabuhan-muara-angke-lebih-modern-setelah-anies-baswedan-berkantor-di-kepulauan-seribu-7085/> (accessed Sep. 08, 2023).
- [52] “Pelabuhan Marina Ancol.” <http://anekainfojakarta.blogspot.com/2012/03/pelabuhan-marina-ancol.html> (accessed Sep. 08, 2023).
- [53] “Pelabuhan Untung Jawa.” <https://www.indonesia-tourism.com/forum/showthread.php?52302-Untung-Jawa-Island-Thousand-Islands-Jakarta> (accessed Sep. 08, 2023).
- [54] “Pelabuhan Lancang.” <https://kepseribuselatan.blogspot.com/2014/07/wisatawan-diketahui-menghilang.html> (accessed Sep. 08, 2023).
- [55] “Pelabuhan Pari.” <https://jejakpiknik.com/pulau-pari/> (accessed Sep. 08, 2023).

- [56] “Pelabuhan Tidung.” <https://dontstopcrazynow.blogspot.com/2013/04/pulau-tidung-kepulauan-seribu.html> (accessed Sep. 08, 2023).
- [57] “Pelabuhan Pramuka.” <https://shipsapp.co.id/Pelabuhan/kep-seribu-pramuka.html> (accessed Sep. 08, 2023).
- [58] “Pelabuhan Kelapa.” <https://www.blj.co.id/2020/11/13/pictures-of-the-day-13-november-2020/pelabuhan-sunda-kelapa-1/> (accessed Sep. 08, 2023).
- [59] “Pelabuhan Harapan.” <https://www.tribunnews.com/images/editorial/view/1676770/pulau-harapan-di-kepulauan-seribu> (accessed Sep. 08, 2023).
- [60] “Pelabuhan Sabira.” <https://www.kompasiana.com/riri.s/54f49a2e745513932b6c8cb8/mengunjungi-sebira-pulau-bugis-di-kepulauan-seribu> (accessed Sep. 08, 2023).
- [61] S. Kelly, “What Is Python?,” in *Python, PyGame, and Raspberry Pi Game Development*, 2019. doi: 10.1007/978-1-4842-4533-0_2.
- [62] P. L. Gilabert *et al.*, “An efficient combination of digital predistortion and ofdm clipping for power amplifiers,” *Int. J. RF Microw. Comput. Eng.*, 2009, doi: 10.1002/mmce.20381.
- [63] C. R. Harris *et al.*, “Array programming with NumPy,” *Nature*. 2020. doi: 10.1038/s41586-020-2649-2.
- [64] I. Stancin and A. Jovic, “An overview and comparison of free Python libraries for data mining and big data analysis,” in *2019 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics, MIPRO 2019 - Proceedings*, 2019. doi: 10.23919/MIPRO.2019.8757088.
- [65] “Open Data Jakarta.” <https://data.jakarta.go.id/> (accessed Jun. 28, 2023).
- [66] G. Premashanti, “Penerapan Metode Klastering Dengan Algoritma K-Means Untuk Prediksi Kelulusan Mahasiswa Pada Program Studi Teknik Informatika Strata Satu,” pp. 274–282, 2020.