



## PENERAPAN DATA MINING PADA TINGKAT PENGHUNIAN KAMAR PADA HOTEL BERBINTANG BERDASARKAN PROVINSI MENGGUNAKAN K-MEANS CLUSTERING

Lambas Perdana Naibaho<sup>1</sup>, Agus Perdana Windarto<sup>2</sup>

<sup>1</sup>Program Studi Teknik Informatika, STIKOM Tunas Bangsa, Pematangsiantar, Indonesia

<sup>2</sup>Program Studi Sistem Informasi, STIKOM Tunas Bangsa, Pematangsiantar, Indonesia

E-Mail : <sup>1</sup>lambasperdanaibaho@gmail.com, <sup>2</sup>agus.perdana@amiktunasbangsa.ac.id

### Article Info

#### Article history:

Received Aug 31, 2023

Revised Sept 05, 2023

Accepted Sept 8, 2023

#### Kata Kunci:

Clustering  
Data Mining  
Hotel  
Kamar  
K-Means

#### Keywords:

Clustering  
Data Mining  
Hotels  
Rooms  
K-Means

### ABSTRAK

Hotel adalah sebuah tempat akomodasi yang menyediakan fasilitas penginapan bagi tamu yang datang dari luar kota ataupun luar negeri. Tujuan utama dari hotel adalah memberikan pengalaman menginap yang nyaman dan memuaskan bagi tamu. Seiring dengan perkembangan jumlah Hotel Berbintang di Indonesia maka persaingan akan layanan dan fasilitas juga ikut bersaing, oleh karena itu perlu adanya hotel-hotel berbintang yang terjamin kualitas dan kuantitasnya secara statistik. Dengan kualitas pelayanan yang baik akan memberikan kesan yang baik bagi pengunjung dan akan mengakibatkan pelanggan loyal serta mempercayai jasa yang diberikan. Penelitian ini membahas tentang penerapan data mining pada tingkat penghunian kamar pada hotel berbintang di Indonesia berdasarkan provinsi menggunakan K-Means Clustering. Sumber data penelitian ini dikumpulkan berdasarkan dokumen-dokumen keterangan tingkat penghunian kamar pada hotel berbintang yang dihasilkan oleh Badan Pusat Statistik Nasional. Data yang digunakan dalam penelitian ini adalah data dari bulan januari-desember tahun 2022. Data akan diolah dengan melakukan clustering dengan 3 cluster yaitu cluster tingkat penghunian kamar tinggi, cluster tingkat penghunian kamar rendah, cluster tingkat penghunian kamar sedang. Centroid data untuk cluster tingkat penghunian kamar tinggi 741.97, Centroid data untuk cluster tingkat penghunian kamar rendah 343.76, dan Centroid data untuk cluster tingkat penghunian kamar sedang 538.98. Sehingga diperoleh penilaian berdasarkan indeks tingkat penghunian kamar pada hotel berbintang dengan 12 provinsi cluster tingkat penghunian kamar tinggi yakni Sumatera Selatan, Lampung, DKI Jakarta, Jawa Barat, DI Yogyakarta, Jawa Timur, Banten, Kalimantan Barat, Kalimantan Tengah, Kalimantan Selatan, Kalimantan Timur, Sulawesi Tengah. 7 provinsi cluster tingkat penghunian kamar rendah yakni Aceh, Kep. Bangka Belitung, Bali, Nusa Tenggara Barat, Sulawesi Barat, Maluku, Papua. 15 provinsi cluster tingkat penghunian kamar sedang yakni Sumatera Utara, Sumatera Barat, Riau, Jambi, Bengkulu, Kep. Riau, Jawa Tengah, Nusa Tenggara Timur, Kalimantan Utara, Sulawesi Utara, Sulawesi Selatan, Sulawesi Tenggara, Gorontalo, Maluku Utara, Papua Barat. Hal ini dapat menjadi masukan kepada pemerintah dan pihak perhotelan, provinsi yang menjadi perhatian lebih pada tingkat penghunian kamar pada hotel berbintang berdasarkan cluster yang telah dilakukan.

### ABSTRACT

Hotel is an accommodation place that provides lodging facilities for guests who come from out of town or abroad. The main goal of the hotel is to provide a comfortable and satisfying stay for guests. Along with the development of the number of star hotels in Indonesia, competition for services and facilities also competes, therefore it is necessary to have star hotels that are guaranteed for quality and quantity statistically. With good service quality will give a good impression to visitors and will result in loyal customers and trust the services provided. This study discusses the application of data mining at the level of room occupancy in star hotels in Indonesia by province using K-Means Clustering. The data source for this

---

research was collected based on documents describing room occupancy rates in star hotels produced by the Central Bureau of National Statistics. The data used in this study is data from January-December 2022. The data will be processed by clustering with 3 clusters, namely high room occupancy rate clusters, low room occupancy rate clusters, medium room occupancy rate clusters. The data centroid for the high occupancy rate cluster is 741.97, the data centroid for the low room occupancy rate cluster is 343.76, and the data centroid for the moderate room occupancy rate cluster is 538.98. So that an assessment is obtained based on the index of room occupancy rates in star hotels with 12 province clusters with high room occupancy rates, namely South Sumatra, Lampung, DKI Jakarta, West Java, DI Yogyakarta, East Java, Banten, West Kalimantan, Central Kalimantan, South Kalimantan, East Kalimantan. , Central Sulawesi. 7 cluster provinces with low room occupancy rates namely Aceh, Kep. Bangka Belitung, Bali, West Nusa Tenggara, West Sulawesi, Maluku, Papua. The 15 province clusters have moderate room occupancy levels, namely North Sumatra, West Sumatra, Riau, Jambi, Bengkulu, Kep. Riau, Central Java, East Nusa Tenggara, North Kalimantan, North Sulawesi, South Sulawesi, Southeast Sulawesi, Gorontalo, North Maluku, West Papua. This can be input to the government and hoteliers, provinces that are of more concern to the level of room occupancy in star-rated hotels based on the clusters that have been carried out.

This is an open access article under the [CC BY-NC](#) license.



---

*Corresponding Author:*

Lambas Perdana Naibaho,

Teknik Informatika, STIKOM Tunas Bangsa, Pematangsiantar, Indonesia,

Jl. Sudirman No. 1,2 & 3, Kota Pematang Siantar, Sumatera Utara, 21142, Indonesia.

Email: lambasperdananaibaho@gmail.com

---

## 1. PENDAHULUAN

Indonesia merupakan negara yang berada diantara Benua Asia dan Benua Australia serta dua Samudera, Samudera Pasifik dan Samudera Hindia. Posisi strategis tersebut menjadikan Indonesia memiliki peran penting dalam dunia ekonomi dan iklim. Sebagai negara dengan dua musim yang berbeda, Indonesia memiliki musim hujan dan musim kemarau yang keduanya akan selalu berganti setiap enam bulan sekali. Dari segi ekonomi, Indonesia berperan sebagai tempat persimpangan lalu lintas dunia. Hal ini menjadikan Indonesia sebagai negara yang sering dikunjungi oleh negara lain seperti Jepang, Korea, dan China yang berada di kawasan Asia, Afrika, atau bahkan Eropa. Banyak Wisatawan mancanegara yang masuk ke Indonesia sehingga perlu adanya tempat yang aman sebagai tempat persinggahan untuk beristirahat saat memasuki Indonesia. Tempat yang paling sering digunakan Wisatawan mancanegara untuk beristirahat adalah hotel (Sari & Hartama, 2018). Hotel merupakan sejenis akomodasi yang menyediakan fasilitas dan pelayanan penginapan, makan dan minum, serta jasa-jasa lainnya untuk umum yang dapat tinggal untuk sementara waktu, dan dikelola secara komersial. Hotel bukan hanya dapat digunakan bagi orang-orang yang sedang melakukan perjalanan atau menjalankan liburan saja, namun juga dapat dikunjungi oleh orang-orang yang sedang melakukan perjalanan bisnis (Siregar & Siagian, 2021). Hotel merupakan salah satu badan usaha yang mempunyai potensi untuk dapat tumbuh dan berkembang yang membutuhkan investasi yang tidak sedikit yang berdiri baik ditengah kota maupun di daerah-daerah tujuan wisata (Annur, 2022).

Penilaian untuk tingkat penentuan hotel berbintang dapat dilakukan dengan menggunakan analisis multivariat karena dapat langsung melakukan analisis data terhadap beberapa variabel secara bersamaan. Dalam analisis multivariat, terdapat teknik yang disebut analisis cluster yang dapat digunakan pada kasus pengelompokan hotel berbintang yang ada di Indonesia berdasarkan Provinsi. Analisis klaster adalah teknik analisis multivariat yang bertujuan untuk mengelompokkan kumpulan data sesuai dengan karakteristik yang membedakannya, dengan mengklasifikasikan setiap kumpulan data terhadap kumpulan data lainnya yang memiliki karakteristik yang sama (Awalluddin & Taufik, 2017). K-Means Cluster adalah teknik analisis cluster yang digunakan untuk menempatkan objek dalam satu atau lebih cluster berdasarkan karakteristiknya, sehingga objek dengan karakteristik tertentu ditempatkan dalam cluster yang sama dengan objek lain yang memiliki karakteristik tersebut (Sari & Sukestiyarno, 2021).

Pada penelitian sebelumnya, menurut Johan Oscar Ong data mining digunakan untuk menentukan strategi marketing President University (Ong, 2013), pada penelitian lainnya menurut Sukai Minsas, dkk data

mining digunakan untuk pengelompokan hotel berbintang di kota Pontianak (Minsas et al., 2023), pada penelitian lainnya menurut Romiwita Irene Klarita, dkk data mining digunakan untuk analisis target pemasaran (Klarita & Kaunang, 2020), pada penelitian lainnya menurut Achmad Bahauddin, dkk data mining digunakan untuk clustering provinsi di Indonesia berdasarkan tingkat kemiskinan (Bahauddin et al., 2021), pada penelitian lainnya menurut Corie Mei Hellyana data mining digunakan terhadap kunjungan wisatawan asing di hotel berbintang di Indonesia (Mei Hellyana, 2023), pada penelitian lainnya menurut Taofik Hidayat data mining digunakan untuk analisis karakteristik konsumen hotel (Hidayat & Putro, 2020), pada penelitian lainnya menurut Fatimah Isyarah, dkk data mining digunakan untuk clustering daerah miskin di provinsi Riau, pada penelitian lainnya menurut Rahmawati data mining digunakan untuk pengelompokan penjualan produk pada hotel newton (Rahmawati & Nopriadi, 2020). Pada penelitian ini, penulis berfokus pada clustering hotel berbintang dengan metode K-Means Cluster yang pada akhirnya akan menentukan level atau peringkat klaster yang ada di Indonesia berdasarkan Provinsi. Data yang digunakan dalam penelitian ini berdasarkan dokumen-dokumen yang dihasilkan oleh Badan Pusat Statistik Nasional melalui situs <https://www.bps.go.id>.

Dalam penelitian ini, penulis mengangkat topik tingkat penghunian kamar pada hotel berbintang berdasarkan provinsi dimana metode yang digunakan adalah clustering. Hasil dari cluster dapat dijadikan masukan bagi pemerintah dan pihak perhotelan agar provinsi yang masuk ke dalam cluster rendah mendapat perhatian lebih. Proses cluster dibagi kedalam 3 (tiga) cluster yakni tingkat penghunian kamar tinggi, tingkat penghunian kamar rendah dan tingkat penghunian kamar sedang. Penelitian ini menghasilkan penilaian berdasarkan indeks tingkat penghunian kamar pada hotel berbintang dengan 12 provinsi cluster tingkat penghunian kamar tinggi yakni Sumatera Selatan, Lampung, DKI Jakarta, Jawa Barat, DI Yogyakarta, Jawa Timur, Banten, Kalimantan Barat, Kalimantan Tengah, Kalimantan Selatan, Kalimantan Timur, Sulawesi Tengah. 7 provinsi cluster tingkat penghunian kamar rendah yakni Aceh, Kep. Bangka Belitung, Bali, Nusa Tenggara Barat, Sulawesi Barat, Maluku, Papua. 15 provinsi cluster tingkat penghunian kamar sedang yakni Sumatera Utara, Sumatera Barat, Riau, Jambi, Bengkulu, Kep. Riau, Jawa Tengah, Nusa Tenggara Timur, Kalimantan Utara, Sulawesi Utara, Sulawesi Selatan, Sulawesi Tenggara, Gorontalo, Maluku Utara, Papua Barat. Hasil yang didapat dari penelitian dapat menjadi masukan kepada pemerintah maupun pihak perhotelan, provinsi yang menjadi perhatian lebih pada tingkat penghunian kamar pada hotel berbintang berdasarkan cluster yang telah dilakukan.

## **2. METODE PENELITIAN**

### **2.1. Tahap Penelitian**

Tahap dalam penelitian ini yang digunakan yaitu metode Data Mining sebagai berikut:

- a. Tahap pengumpulan data,
- b. Tahap pengolahan data,
- c. Tahap clustering dan
- d. Tahap analisis

### **2.2. Dataset**

Data yang digunakan dalam penelitian ini adalah data Tingkat Penghunian Kamar pada Hotel Bintang di tahun 2022 dari bulan januari - desember yang terdiri dari 34 provinsi yang berasal dari Badan Pusat Statistik (BPS) melalui situs <https://www.bps.go.id>. Variabel yang digunakan jumlah Tingkat Penghunian Kamar pada Hotel Bintang yang sudah diakumulasikan. Data akan diolah dengan melakukan clustering Tingkat Penghunian Kamar pada Hotel Bintang berdasarkan provinsi dalam 3 cluster yakni cluster tingkat penghunian kamar tinggi, cluster tingkat penghunian kamar rendah dan cluster tingkat penghunian sedang.

### **2.3. Hotel**

Hotel merupakan sejenis akomodasi yang menyediakan fasilitas dan pelayanan penginapan, makan dan minum, serta jasa lainnya untuk umum yang dapat tinggal untuk sementara waktu, dan dikelola secara komersial. Hotel bukan hanya dapat digunakan bagi orang-orang yang sedang melakukan perjalanan atau menjalankan liburan saja, namun juga dapat dikunjungi oleh orang-orang yang sedang melakukan perjalanan bisnis (Siregar & Siagian, 2021).

## 2.4. Data Mining

Data mining adalah proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Teknik-teknik dan metode atau algoritma dalam data mining sangat bervariasi. Pemilihan metode atau algoritma yang tepat akan bergantung pada tujuan dan proses Knowledge Discovery in Database (KDD) secara keseluruhan (Mardi, 2017).

## 2.5. Clustering

Clustering merupakan salah satu metode Data Mining yang bersifat tanpa arahan (unsupervised). Ada dua jenis data clustering yang sering digunakan dalam proses pengelompokan data yaitu hierarchical (hirarki) data clustering dan non-hierarchical (non hirarki) data clustering (Rosmini et al., 2018). Clustering dapat dikatakan juga sebagai identifikasi kelas objek yang memiliki kemiripan. Dengan menggunakan teknik clustering kita bisa lebih lanjut mengidentifikasi kepadatan dan jarak daerah dalam objek ruang dan dapat menemukan secara menyeluruh pola distribusi dan korelasi antara atribut.

## 3. HASIL AND PEMBAHASAN

Dalam melakukan clustering, data yang diperoleh akan dihitung terlebih dahulu berdasarkan jumlah tingkat penghunian kamar pada hotel bintang di tahun 2022 dari bulan januari - desember berdasarkan provinsi. Hasil penjumlahan berdasarkan 1 kriteria penilaian yakni jumlah penghunian kamar contoh dapat dilihat pada tabel 1.

Tabel 1. Data Tingkat Penghunian Kamar Pada Hotel Bintang Menurut Provinsi

Provinsi	Januari	...	November	Desember	Total
ACEH	19.09	...	35.36	36.82	386.92
SUMATERA UTARA	42.62	...	50.14	51.77	533.12
SUMATERA BARAT	45.19	...	55.82	58.62	579.87
RIAU	40.84	...	40.65	43.36	499.39
JAMBI	37.88	...	58.31	58.24	582.65
SUMATERA SELATAN	49.94	...	60.32	58.62	642.51
BENGKULU	35.43	...	45.11	51.43	481.81
LAMPUNG	51.19	...	56.98	58.74	666.26
KEP. BANGKA BELITUNG	26.11	...	35.24	36.99	402.83
KEP. RIAU	26.39	...	45.67	62.53	483.32
DKI JAKARTA	52.26	...	60.81	59.63	643.21
JAWA BARAT	46.99	...	56.64	60.18	611.44
JAWA TENGAH	39.38	...	51.33	54.74	538.98
DI YOGYAKARTA	59.90	...	64.95	72.87	684.94
JAWA TIMUR	44.71	...	57.72	61.86	630.69
BANTEN	41.64	...	55.40	51.21	605.5
BALI	20.71	...	48.91	53.75	423.85
NUSA TENGGARA BARAT	31.35	...	42.42	40.67	411.05
NUSA TENGGARA TIMUR	29.85	...	39.33	42.03	478.45
KALIMANTAN BARAT	45.27	...	57.09	52.05	598.05
KALIMANTAN TENGAH	43.20	...	58.46	56.76	621.29
KALIMANTAN SELATAN	41.08	...	56.77	58.38	602.23
KALIMANTAN TIMUR	60.78	...	66.51	67.52	741.97
KALIMANTAN UTARA	50.43	...	47.62	47.61	568.89
SULAWESI UTARA	40.11	...	49.83	50.08	510.79
SULAWESI TENGAH	44.40	...	63.09	60.66	600.52
SULAWESI SELATAN	38.90	...	55.46	54.26	579.74
SULAWESI TENGGARA	38.29	...	46.93	49.16	480.89
GORONTALO	29.60	...	59.83	46.89	525.1
SULAWESI BARAT	11.86	...	33.82	29.92	343.76
MALUKU	40.36	...	36.71	39.47	404.34
MALUKU UTARA	38.87	...	44.50	42.88	556.43
PAPUA BARAT	44.07	...	56.55	49.97	578.07
PAPUA	39.54	...	33.72	39.06	447.55

### 3.1. Pengolahan Data

Untuk mendapatkan hasil dari penelitian yang dilakukan, berikut uraian perhitungan proses clustering tingkat penghunian kamar pada hotel bintang menggunakan k-means. Data pada tabel 1 kemudian diakumulasikan berdasarkan 1 kriteria yaitu jumlah tingkat penghunian kamar contoh dapat dilihat pada tabel 2.

Tabel 2. Data Akumulasi Tingkat Penghunian Kamar Pada Hotel Bintang

Provinsi	Total
ACEH	386.92
SUMATERA UTARA	533.12
SUMATERA BARAT	579.87
RIAU	499.39
JAMBI	582.65
SUMATERA SELATAN	642.51
BENGKULU	481.81
LAMPUNG	666.26
KEP. BANGKA BELITUNG	402.83
KEP. RIAU	483.32
DKI JAKARTA	643.21
JAWA BARAT	611.44
JAWA TENGAH	538.98
DI YOGYAKARTA	684.94
JAWA TIMUR	630.69
BANTEN	605.5
BALI	423.85
NUSA TENGGARA BARAT	411.05
NUSA TENGGARA TIMUR	478.45
KALIMANTAN BARAT	598.05
KALIMANTAN TENGAH	621.29
KALIMANTAN SELATAN	602.23
KALIMANTAN TIMUR	741.97
KALIMANTAN UTARA	568.89
SULAWESI UTARA	510.79
SULAWESI TENGAH	600.52
SULAWESI SELATAN	579.74
SULAWESI TENGGARA	480.89
GORONTALO	525.1
SULAWESI BARAT	343.76
MALUKU	404.34
MALUKU UTARA	556.43
PAPUA BARAT	578.07
PAPUA	447.55

Setelah data diakumulasikan maka akan didapatkan nilai dari seluruh tingkat penghunian kamar pada hotel berbintang menurut provinsi. Kemudian data tersebut akan masuk ke tahapan clustering dengan menerapkan algoritma K-Means untuk mengcluster data menjadi tiga cluster.

#### A. Menentukan Jumlah Cluster

Jumlah Cluster yang digunakan pada data tingkat penghunian kamar pada hotel berbintang sebanyak 3 cluster yakni cluster tingkat tinggi (C1), cluster tingkat rendah (2), cluster tingkat sedang (C3).

#### B. Menentukan Centroid

Penentuan titik cluster ini dilakukan dengan mengambil nilai terbesar (maksimum) untuk cluster tingkat tinggi (C1), nilai terkecil (minimum) untuk cluster tingkat rendah (C2) dan nilai tengah untuk cluster tingkat sedang (C3). Data centroid tersebut contoh dapat dilihat pada tabel 3.

Tabel 3. Data Centroid

Cluster	Nilai
Centroid 1	741.97
Centroid 2	343.76
Centroid 3	538.98

### C. Menghitung Jarak dari Centroid

Untuk menghitung jarak antara titik centroid dengan titik tiap objek menggunakan Euclidian Distance. Hasil dari keseluruhan perhitungan untuk menghitung jarak dari centroid contoh dapat dilihat pada tabel 4.

Tabel 4. Perhitungan Jarak Data Iterasi ke-1

Provinsi	Tingkat Penghunian Kamar	C1	C2	C3	Jarak Terpendek
ACEH	386.92	355.05	43.16	152.06	43.16
SUMATERA UTARA	533.12	208.85	189.36	5.86	5.86
SUMATERA BARAT	579.87	162.1	236.11	40.89	40.89
RIAU	499.39	242.58	155.63	39.59	39.59
JAMBI	582.65	159.32	238.89	43.67	43.67
SUMATERA SELATAN	642.51	99.46	298.75	103.53	99.46
BENGKULU	481.81	260.16	138.05	57.17	57.17
LAMPUNG	666.26	75.71	322.5	127.28	75.71
KEP. BANGKA BELITUNG	402.83	339.14	59.07	136.15	59.07
KEP. RIAU	483.32	258.65	139.56	55.66	55.66
DKI JAKARTA	643.21	98.76	299.45	104.23	98.76
JAWA BARAT	611.44	130.53	267.68	72.46	72.46
JAWA TENGAH	538.98	202.99	195.22	0	0
DI YOGYAKARTA	684.94	57.03	341.18	145.96	57.03
JAWA TIMUR	630.69	111.28	286.93	91.71	91.71
BANTEN	605.5	136.47	261.74	66.52	66.52
BALI	423.85	318.12	80.09	115.13	80.09
NUSA TENGGARA BARAT	411.05	330.92	67.29	127.93	67.29
NUSA TENGGARA TIMUR	478.45	263.52	134.69	60.53	60.53
KALIMANTAN BARAT	598.05	143.92	254.29	59.07	59.07
KALIMANTAN TENGAH	621.29	120.68	277.53	82.31	82.31
KALIMANTAN SELATAN	602.23	139.74	258.47	63.25	63.25
KALIMANTAN TIMUR	741.97	0	398.21	202.99	0
KALIMANTAN UTARA	568.89	173.08	225.13	29.91	29.91
SULAWESI UTARA	510.79	231.18	167.03	28.19	28.19
SULAWESI TENGAH	600.52	141.45	256.76	61.54	61.54
SULAWESI SELATAN	579.74	162.23	235.98	40.76	40.76
SULAWESI TENGGARA	480.89	261.08	137.13	58.09	58.09
GORONTALO	525.1	216.87	181.34	13.88	13.88
SULAWESI BARAT	343.76	398.21	0	195.22	0
MALUKU	404.34	337.63	60.58	134.64	60.58
MALUKU UTARA	556.43	185.54	212.67	17.45	17.45
PAPUA BARAT	578.07	163.9	234.31	39.09	39.09
PAPUA	447.55	294.42	103.79	91.43	91.43

### D. Menentukan Pengelompokan Data

Dalam menentukan pengelompokan dengan mencari nilai cluster berdasarkan nilai minimal dari nilai cluster dan diletakkan pada cluster yang sesuai dengan nilai minimal pada iterasi 1. Contoh dapat dilihat pada tabel 5.

Tabel 5. Perhitungan Jarak Data Iterasi ke-1

No	Provinsi	Kelompok Data		
		C1	C2	C3
1	ACEH		1	
2	SUMATERA UTARA			1
3	SUMATERA BARAT			1
4	RIAU			1
5	JAMBI			1
6	SUMATERA SELATAN	1		

No	Provinsi	Kelompok Data		
		C1	C2	C3
7	BENGKULU			1
8	LAMPUNG	1		
9	KEP. BANGKA BELITUNG		1	
10	KEP. RIAU			1
11	DKI JAKARTA	1		
12	JAWA BARAT			1
13	JAWA TENGAH			1
14	DI YOGYAKARTA	1		
15	JAWA TIMUR			1
16	BANTEN			1
17	BALI		1	
18	NUSA TENGGARA BARAT		1	
19	NUSA TENGGARA TIMUR			1
20	KALIMANTAN BARAT			1
21	KALIMANTAN TENGAH			1
22	KALIMANTAN SELATAN			1
23	KALIMANTAN TIMUR	1		
24	KALIMANTAN UTARA			1
25	SULAWESI UTARA			1
26	SULAWESI TENGAH			1
27	SULAWESI SELATAN			1
28	SULAWESI TENGGARA			1
29	GORONTALO			1
30	SULAWESI BARAT		1	
31	MALUKU		1	
32	MALUKU UTARA			1
33	PAPUA BARAT			1
34	PAPUA			1

Dalam metode K-Means, perhitungan akan terus beriterasi sampai pengelompokan data sama dengan pengelompokan data iterasi sebelumnya. Maka, proses akan terus melakukan iterasi sampai data pada iterasi terakhir sama dengan iterasi sebelumnya. Pada iterasi 1 diperoleh cluster data tingkat penghunian kamar pada hotel berbintang berdasarkan provinsi. Contoh dapat dilihat pada tabel 4. Proses iterasi tersebut berhenti pada iterasi ke 6, pada iterasi 6 akan dilakukan proses mencari nilai centroid. Contoh dapat dilihat pada tabel 6.

Tabel 6. Data Centroid Iterasi 6

Cluster	Nilai
Centroid 1	6.373.842
Centroid 2	402.9
Centroid 3	5.318.333

Setelah mendapatkan nilai centroid, proses sama dilakukan dengan mencari jarak terdekat. Proses pencarian jarak terdekat, pengelompokan data pada iterasi 6 dan Clustering data contoh dapat dilihat pada tabel 7.

Tabel 7. Perhitungan Jarak Data Iterasi ke-6

Provinsi	Tingkat Penghunian Kamar	C1	C2	C3	Jarak Terpendek
ACEH	386.92	250.4641667	15.98	144.9133333	15.98
SUMATERA UTARA	533.12	104.2641667	130.22	1.286666667	1.286666667
SUMATERA BARAT	579.87	57.51416667	176.97	48.03666667	48.03666667
RIAU	499.39	137.9941667	96.49	32.44333333	32.44333333
JAMBI	582.65	54.73416667	179.75	50.81666667	50.81666667
SUMATERA SELATAN	642.51	5.125833333	239.61	110.6766667	5.125833333
BENGKULU	481.81	155.5741667	78.91	50.02333333	50.02333333
LAMPUNG	666.26	28.87583333	263.36	134.4266667	28.87583333
KEP. BANGKA BELITUNG	402.83	234.5541667	0.07	129.0033333	0.07
KEP. RIAU	483.32	154.0641667	80.42	48.51333333	48.51333333
DKI JAKARTA	643.21	5.825833333	240.31	111.3766667	5.825833333
JAWA BARAT	611.44	25.94416667	208.54	79.60666667	25.94416667

Provinsi	Tingkat Penghunian Kamar	C1	C2	C3	Jarak Terpendek
JAWA TENGAH	538.98	98.40416667	136.08	7.146666667	7.146666667
DI YOGYAKARTA	684.94	47.55583333	282.04	153.1066667	47.55583333
JAWA TIMUR	630.69	6.694166667	227.79	98.85666667	6.694166667
BANTEN	605.5	31.88416667	202.6	73.66666667	31.88416667
BALI	423.85	213.5341667	20.95	107.9833333	20.95
NUSA TENGGARA BARAT	411.05	226.3341667	8.15	120.7833333	8.15
NUSA TENGGARA TIMUR	478.45	158.9341667	75.55	53.38333333	53.38333333
KALIMANTAN BARAT	598.05	39.33416667	195.15	66.21666667	39.33416667
KALIMANTAN TENGAH	621.29	16.09416667	218.39	89.45666667	16.09416667
KALIMANTAN SELATAN	602.23	35.15416667	199.33	70.39666667	35.15416667
KALIMANTAN TIMUR	741.97	104.5858333	339.07	210.1366667	104.5858333
KALIMANTAN UTARA	568.89	68.49416667	165.99	37.05666667	37.05666667
SULAWESI UTARA	510.79	126.5941667	107.89	21.04333333	21.04333333
SULAWESI TENGAH	600.52	36.86416667	197.62	68.68666667	36.86416667
SULAWESI SELATAN	579.74	57.64416667	176.84	47.90666667	47.90666667
SULAWESI TENGGARA	480.89	156.4941667	77.99	50.94333333	50.94333333
GORONTALO	525.1	112.2841667	122.2	6.733333333	6.733333333
SULAWESI BARAT	343.76	293.6241667	59.14	188.0733333	59.14
MALUKU	404.34	233.0441667	1.44	127.4933333	1.44
MALUKU UTARA	556.43	80.95416667	153.53	24.59666667	24.59666667
PAPUA BARAT	578.07	59.31416667	175.17	46.23666667	46.23666667
PAPUA	447.55	189.8341667	44.65	84.28333333	44.65

Dari tabel jarak centroid diatas, maka pengelompokan pada iterasi ke-6 dapat dilihat pada tabel 8.

Tabel 8. Perhitungan Jarak Data Iterasi ke-6

No	Provinsi	Kelompok Data		
		C1	C2	C3
1	ACEH		1	
2	SUMATERA UTARA			1
3	SUMATERA BARAT			1
4	RIAU			1
5	JAMBI			1
6	SUMATERA SELATAN	1		
7	BENGKULU			1
8	LAMPUNG	1		
9	KEP. BANGKA BELITUNG		1	
10	KEP. RIAU			1
11	DKI JAKARTA	1		
12	JAWA BARAT	1		
13	JAWA TENGAH			1
14	DI YOGYAKARTA	1		
15	JAWA TIMUR	1		
16	BANTEN	1		
17	BALI		1	
18	NUSA TENGGARA BARAT		1	
19	NUSA TENGGARA TIMUR			1
20	KALIMANTAN BARAT	1		
21	KALIMANTAN TENGAH	1		
22	KALIMANTAN SELATAN	1		
23	KALIMANTAN TIMUR	1		
24	KALIMANTAN UTARA			1
25	SULAWESI UTARA			1
26	SULAWESI TENGAH	1		
27	SULAWESI SELATAN			1
28	SULAWESI TENGGARA			1
29	GORONTALO			1
30	SULAWESI BARAT		1	
31	MALUKU		1	
32	MALUKU UTARA			1
33	PAPUA BARAT			1
34	PAPUA		1	

## E. Analisis Data

Pada iterasi 6, pengelompokan data yang dilakukan terhadap 3 cluster dengan iterasi 5 didapatkan hasil yang sama. Dari 34 data tingkat penghunian kamar pada hotel berbintang berdasarkan provinsi dapat diketahui, 12 provinsi cluster tingkat penghunian kamar tinggi yakni Sumatera Selatan, Lampung, DKI Jakarta, Jawa Barat, DI Yogyakarta, Jawa Timur, Banten, Kalimantan Barat, Kalimantan Tengah, Kalimantan Selatan, Kalimantan Timur, Sulawesi Tengah. 7 provinsi cluster tingkat penghunian kamar rendah yakni Aceh, Kep. Bangka Belitung, Bali, Nusa Tenggara Barat, Sulawesi Barat, Maluku, Papua. 15 provinsi cluster tingkat penghunian kamar sedang yakni Sumatera Utara, Sumatera Barat, Riau, Jambi, Bengkulu, Kep. Riau, Jawa Tengah, Nusa Tenggara Timur, Kalimantan Utara, Sulawesi Utara, Sulawesi Selatan, Sulawesi Tenggara, Gorontalo, Maluku Utara, Papua Barat.

## 4. KESIMPULAN

Hasil pada penelitian ini terhadap tingkat penghunian kamar pada hotel berbintang menurut provinsi dapat diterapkan metode clustering K-Means. Data diolah untuk memperoleh nilai dari tingkat penghunian kamar pada hotel berbintang menurut provinsi. Data tersebut diolah menggunakan Ms. Excel untuk ditentukan nilai centroid dalam 3 cluster yaitu cluster tingkat penghunian kamar tinggi, cluster tingkat penghunian kamar rendah dan cluster tingkat penghunian kamar sedang. Centroid data untuk cluster tingkat penghunian kamar tinggi 637.3842, Centroid data untuk cluster tingkat penghunian kamar rendah 402.9 dan Centroid data untuk cluster tingkat penghunian kamar sedang 531.8333. Sehingga diperoleh penilaian berdasarkan indeks indeks tingkat penghunian kamar pada hotel berbintang dengan 12 provinsi cluster tingkat penghunian kamar tinggi yakni Sumatera Selatan, Lampung, DKI Jakarta, Jawa Barat, DI Yogyakarta, Jawa Timur, Banten, Kalimantan Barat, Kalimantan Tengah, Kalimantan Selatan, Kalimantan Timur, Sulawesi Tengah. 7 provinsi cluster tingkat penghunian kamar rendah yakni Aceh, Kep. Bangka Belitung, Bali, Nusa Tenggara Barat, Sulawesi Barat, Maluku, Papua. 15 provinsi cluster tingkat penghunian kamar sedang yakni Sumatera Utara, Sumatera Barat, Riau, Jambi, Bengkulu, Kep. Riau, Jawa Tengah, Nusa Tenggara Timur, Kalimantan Utara, Sulawesi Utara, Sulawesi Selatan, Sulawesi Tenggara, Gorontalo, Maluku Utara, Papua Barat. Hasil yang didapat dari penelitian dapat menjadi masukan kepada pemerintah maupun pihak perhotelan, provinsi yang menjadi perhatian lebih pada tingkat penghunian kamar pada hotel berbintang berdasarkan cluster yang telah dilakukan.

## ACKNOWLEDGEMENTS

Terima kasih disampaikan kepada pihak-pihak yang telah mendukung terlaksananya penelitian ini.

## REFERENCES

- Annur, H. (2022). Penerapan Algoritma Naïve Bayes Berbasis Backward Elimination Untuk Prediksi Pemesanan Kamar Hotel. *Jurnal Ilmiah Ilmu Komputer Banthayo Lo Komputer*, 1(1), 1–5. <https://doi.org/10.37195/balok.v1i1.99>
- Awalluddin, A. ., & Taufik, I. (2017). Analisis Cluster Data Longitudinal pada Pengelompokan Daerah Berdasarkan Indikator IPM di Jawa Barat. *Prosiding Seminar Nasional Metode Kuantitatif*, 978, 187–194.
- Bahauddin, A., Fatmawati, A., & Permata Sari, F. (2021). Analisis Clustering Provinsi Di Indonesia Berdasarkan Tingkat Kemiskinan Menggunakan Algoritma K-Means. *Jurnal Manajemen Informatika Dan Sistem Informasi*, 4(1), 1–8. <https://doi.org/10.36595/misi.v4i1.216>
- Di, B., & Pontianak, K. (2023). PENDEKATAN K - MEANS CLUSTER PADA PENGELOMPOKAN HOTEL BERBINTANG DI KOTA PONTIANAK. 8(1), 412–419.
- Hidayat, T., & Putro, B. E. (2020). Analisis Karakteristik Konsumen Hotel “X” dengan Menggunakan Metode K-Means Clustering. *Jurnal Media Teknik Dan Sistem Industri*, 4(2), 53. <https://doi.org/10.35194/jmtsi.v4i2.995>
- Isyarah, F., Hasan, A., & Wiza, F. (2020). Clustering Daerah Miskin Di Provinsi Riau Menggunakan Metode K-Means. *Prosiding-Seminar Nasional Teknologi Informasi & Ilmu Komputer (SEMESTER)*, 1(1), 1–12.
- Klarita, R. I., & Kaunang, F. J. (2020). Analisis Target Pemasaran Menggunakan Algoritma K-Means Clustering (Studi Kasus: Hotel Sutan Raja Soreang). *TeIka*, 10(2), 107–115. <https://doi.org/10.36342/teika.v10i2.2380>
- Mardi, Y. (2017). Data Mining : Klasifikasi Menggunakan Algoritma C4.5. *Edik Informatika*, 2(2), 213–219. <https://doi.org/10.22202/ei.2016.v2i2.1465>
- Mei Hellyana, C. (2023). Penerapan Algoritma K-Means Terhadap Kunjungan Wisatawan Asing Di Hotel Berbintang di Indonesia. *Jurnal Sains Dan Manajemen*, 11(1), 67–77.
- Ong, J. O. (2013). Implementasi Algoritma K-means clustering untuk menentukan strategi marketing president university. *Jurnal Ilmiah Teknik Industri*, vol.12, no(juni), 10–20.
- Rahmawati, M., & Nopriadi. (2020). *Jurnal Comasie*. Comasie, 3(3), 21–30.
- Rosmini, R., Fadlil, A., & Sunardi, S. (2018). Implementasi Metode K-Means Dalam Pemetaan Kelompok Mahasiswa Melalui Data Aktivitas Kuliah. *It Journal Research and Development*, 3(1), 22–31. [https://doi.org/10.25299/itjrd.2018.vol3\(1\).1773](https://doi.org/10.25299/itjrd.2018.vol3(1).1773)
- Sari, D. N. P., & Sukestiyarno, Y. L. (2021). Analisis Cluster dengan Metode K-Means pada Persebaran Kasus Covid-19

- Berdasarkan Provinsi di Indonesia. PRISMA, Prosiding Seminar Nasional Matematika, 4, 602–610. <https://journal.unnes.ac.id/sju/index.php/prisma/>
- Sari, & Hartama, D. (2018). Data Mining : Algoritma K-Means Pada Pengelompokan Wisata Asing ke Indonesia Menurut Provinsi. Seminar Nasional Sains & Teknologi Informasi (SENSASI), 322–326.
- Siregar, V. M. M., & Siagian, N. F. (2021). Sistem Informasi Front Office Untuk Peningkatan Pelayanan Pelanggan Dalam Reservasi Kamar Hotel. Jurnal Teknik Informasi Dan Komputer (Tekinkom), 4(1), 77–82. <https://doi.org/10.37600/tekinkom.v4i1.279>