

## Penerapan Data Mining Untuk Pengelompokan Wahana Terfavorit Pada CV. Hairos Indah Menggunakan Metode K-Means

Zulham<sup>#1</sup>, Asyahri Hadi Nasyuha<sup>#2</sup>

<sup>#1</sup> Program Studi Manajemen Informatika, Universitas Dharmawangsa

<sup>#2</sup> Program Studi Sistem Informasi, STMIK Triguna Dharma

---

### Info Artikel

#### Article history:

Received Dec 12, 2017

Revised Dec 28, 2017

Accepted Jan 20, 2018

---

#### Keyword:

Algoritma K-Means  
Clustering

---

### ABSTRAK

Saat ini sepengetahuan costumers CV. Hairos Indah hanya terdapat Wahana waterpark saja, ternyata di CV. Hairos Indah terdapat banyak wahana-wahana lain untuk dinikmati, Maka dari itu informasi yang didapat mengenai wahana-wahana tidak akurat. Penelitian ini bertujuan untuk melakukan pengelompokan wahana-wahana- apa saja yang terfavorit bagi para pengunjung dan untuk memperoleh berbagai informasi tentang wahana-wahana apa saja yang terdapat pada CV. Hairos Indah, maka untuk mengetahui informasi tersebut, agar lebih mudah kita harus melakukan satu proses yang dinamakan data mining dengan metode K-Means yang sangat terkenal karena kemudahan dan kemampuannya untuk mengelompokkan data besar dan outlier dengan sangat cepat. Implementasi menggunakan bahasa pemrograman Web untuk membantu pengelompokan Wahana Terfavorit agar mempermudah informasi. Hasil dari penelitian ini dapat membantu CV. Hairos Indah dalam pengelompokan wahana berdasarkan cluster yang terbentuk.

Copyright © 2018 STMIK Triguna Dharma.  
All rights reserved.

---

#### First Author

Nama : Zulham, S.Kom, M.Kom  
Kantor : STMIK Triguna Dharma  
Program Studi : Sistem Informasi  
E-Mail : -

---

### 1. PENDAHULUAN

Seiring dengan perkembangan teknologi, semakin berkembang pula kemampuan dalam mengumpulkan dan mengolah data. Banyaknya ketersediaan data yang dihasilkan dari pengguna teknologi informasi di hampir semua bidang kehidupan, menimbulkan kebutuhan untuk dapat memanfaatkan informasi dan pengetahuan yang terkandung di dalam limbah data tersebut. Jika data dibiarkan saja, maka tersebut hanya akan menjadi sampah yang tidak berarti dan tidak mempunyai nilai guna lebih untuk keperluan di masa yang akan datang. Oleh karena itu diperlukan sebuah aplikasi yang mampu memilah data dan memilih data yang besar, sehingga dapat diperoleh informasi yang berguna bagi penggunanya. Aplikasi data mining pada pengelolaan bisnis, analisa pasar misalnya, memungkinkan diperolehnya hubungan yang dapat dimanfaatkan untuk peningkatan layanan, atau pengolaan sumber daya dengan lebih baik.

Banyaknya persaingan di dunia bisnis khususnya dalam dunia objek wisata, menuntut CV. Hairos Indah untuk menemukan strategi yang dapat meningkatkan usaha di bidang objek wisata terutama harus meningkatkan pelayanan terbaik dan pembaharuan wahana-wahana permainan yang lebih menarik. CV. Hairos Indah merupakan perusahaan yang bergerak dibidang objek wisata, yang terdiri dari wahana air (Waterpark), timezone, dan taman (Theme Park). Untuk meningkatkan pelayanan, pihak manajemen dari CV. Hairos Indah harus melakukan berbagai kebijakan – kebijakan dengan tujuan menarik para klien (Customer) untuk berkunjung ke CV. Hairos Indah, untuk meningkatkan pendapatan (Income) perusahaan. Kebijakan – kebijakan yang dapat dilakukan antara lain adalah membuat promosi, meningkatkan mutu pelayanan dan selalu

memperhatikan perkembangan serta memperbaharui wahana yang lebih menarik agar tetap menjadi objek wisata terfavorit dan tidak membosankan.

Namun untuk melakukan kebijakan-kebijakan tersebut, pihak manajemen harus mengetahui terlebih dahulu tentang hal-hal yang membuat para pengunjung senang berkunjung ke CV. Hairos Indah. Baik dari segi pelayanan, fasilitas, harga tiket, serta wahana- wahana pendukung lainnya. Untuk mengelompokkan wahana apa saja yang paling disenangi dan menjadi favorit dapat diketahui dengan menggunakan pendataan berdasarkan dari Nama - nama wahana dan beberapa kriteria penunjang yang ada di CV. Hairos Indah. Setiap pengunjung yang berkunjung ke CV. Hairos Indah harus melakukan transaksi pembelian tiket ke bagian kasir yang telah disediakan. Dengan bantuan komputer, kasir dan staff lebih mudah untuk melakukan penjualan tiket, pendataan total penjualan tiket, dan memantau laporan penjualan tiket. Dikarenakan di CV. Hairos Indah dalam proses penjualan tiket memakai teknologi yang canggih, maka dalam transaksi pembelian tiket sangatlah mudah dan praktis. Untuk memperoleh berbagai informasi tentang wahana- wahana apa saja yang jadi favorit bagi para pengunjung, maka untuk mengetahui itu semua agar lebih mudah kita harus melakukan satu proses yang dinamakan data mining.

Data Mining adalah serangkaian proses untuk menggali nilai tambah berupa informasi yang selama ini tidak diketahui proses manual dari suatu basis data dengan melakukan penggalian pola-pola dari data dengan tujuan untuk memanipulasi data menjadi informasi yang lebih berharga yang diperoleh dengan cara mengekstraksi dan menggali pola yang penting atau menarik dari data yang terdapat dalam basis data. Untuk membantu dalam proses pencarian maka metode K-Means yang dipergunakan untuk pengelompokan data non hierarki (sekatan) yang berusaha mempartisi yang ada kedalam bentuk dua kelompok atau lebih. Adapun tujuan pengelompokan data ini adalah untuk meminimalkan fungsi objektif yang di set dalam proses pengelompokan, yang pada umumnya berusaha meminimalkan variasi didalam suatu kelompok dan memaksimalkan variasi antar kelompok.

## 2. METODE PENELITIAN

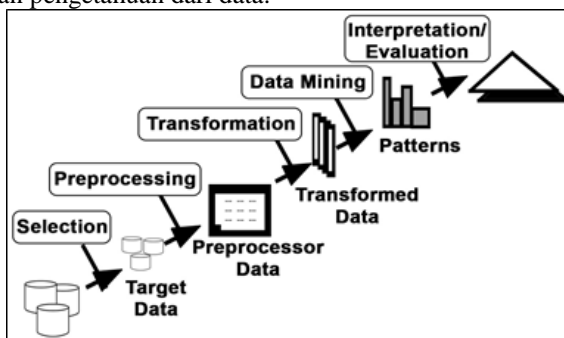
### 2.1 Data Mining

Menurut (Astuti, 2013 : 3) “Data mining adalah proses yang mempekerjakan suatu atau lebih teknik pembelajaran computer (machine learning) untuk menganalisis dan mengekstraksi pengetahuan (knowledge) secara otomatis”.

Data mining adalah kegiatan menemukan pola yang menarik dari data dalam jumlah besar, data dapat disimpan dalam database, data warehouse, atau penyimpanan informasi lainnya. Data mining berkaitan dengan bidang ilmu-ilmu lain, seperti database sistem, data warehousing, statistik, machine learning, information retrieval, dan komputasi tingkat tinggi. Selain itu, data mining didukung oleh ilmu lain seperti neural network, pengenalan pola, spatial data analysis, image database, signal processing. Data mining didefinisikan sebagai proses menemukan pola-pola dalam data. Pola yang ditemukan harus penuh arti dan pola tersebut memberikan keuntungan, biasanya keuntungan secara ekonomi. Adapun Karakteristik dalam data mining adalah sebagai berikut:

1. Data mining berhubungan dengan penemuan sesuatu yang tersembunyi dan pola data tertentu yang tidak diketahui sebelumnya.
2. Data mining biasa menggunakan data yang sangat besar. Biasanya data yang besar digunakan untuk membuat hasil lebih dipercaya.
3. Data mining berguna untuk membuat keputusan yang kritis, terutama dalam strategi.

Data Mining merupakan proses ekstraksi data menjadi informasi yang sebelumnya belum tersampaikan, dengan teknik yang tepat proses data mining akan memberikan hasil yang optimal. Data mining lebih tepat disebut sebagai penambangan pengetahuan dari data.



Gambar 1. Proses Knowledge Discovery in Database (KDD)

## 2.2 Algoritma Clustering

Clustering termasuk ke dalam descriptive methods, dan juga termasuk unsupervised learning dimana tidak ada pendefinisian kelas objek sebelumnya. Sehingga clustering dapat digunakan untuk menentukan label kelas bagi data-data yang belum diketahui kelasnya. Konsep dasar dari clustering adalah mengelompokkan sejumlah objek ke dalam cluster dimana cluster yang baik adalah cluster yang memiliki tingkat kesamaan yang tinggi antar objek di dalam suatu cluster dan tingkat ketidaksamaan yang tinggi dengan objek cluster yang lainnya. Terdapat banyak algoritma clustering yang dalam penggunaannya tergantung pada tipe data yang akan dikelompokkan dan apa tujuan dari pembuatan aplikasinya. Algoritma tersebut dapat digunakan untuk mengelompokkan objek ke dalam cluster-cluster, kemudian dari hasil clustering akan dideteksi keberadaan outlier dalam data tersebut. Sedangkan data yang digunakan bertipe data numerik (Sentika, 2016 : 499).

Algoritma *clustering* diklasifikasikan ke dalam 5 (lima) kategori, yaitu :

1. *Partitioning methods* adalah Pengelompokkan objek dimana tiap objek dimiliki oleh 1 (satu) *cluster*. Algoritma partisi dapat meminimalkan *mean* kuadrat jarak dari setiap titik data ke pusat terdekatnya. Yang termasuk ke dalam metode ini adalah algoritma *k-means*, *k-medoid* atau PAM, CLARA, dan CLARANS, PAM.
2. *Hierarchical methods* adalah Pengelompokkan objek dapat dilakukan dengan 2 (dua) cara, *agglomerative* yang dimulai dengan menggabungkan beberapa *cluster* hingga menjadi satu, atau *divisive* yang dimulai dengan *cluster* yang sama kemudian dipecah menjadi beberapa *cluster* yang lebih kecil. Yang termasuk ke dalam metode ini adalah algoritma *CURE*, *BIRCH*, dan *Chameleon*.
3. *Density-based methods* adalah Pengelompokkan objek berdasarkan tingkat kerapatan objek atau densitas. Yang termasuk dalam metode ini adalah algoritma *DBSCAN*, *DENCLUE*, dan *OPTICS*.
4. *Grid-based methods* adalah Pengelompokkan objek dengan menggunakan struktur data *grid* multiresolusi yang mampu menangani data berdimensi tinggi. Yang termasuk dalam metode ini adalah algoritma *CLIQUE*, *Wave Cluster*, dan *STING*.
5. *Model-based methods* adalah Pengelompokkan objek dengan memodelkan tiap *cluster*, dan mencoba mengoptimalkan kesesuaian data dengan model matematika. Yang termasuk ke dalam metode ini adalah algoritma *COBWEB*.

## 2.3 Algoritma K-Means

Metode K-Means adalah Metode clustering berbasis jarak yang membagi data kedalam cluster dan algoritma ini bekerja pada atribut numerik. Metode K-Means termasuk dalam partitioning clustering yang memisahkan data ke daerah bagian yang terpisah. Metode K-Means sangat terkenal karena kemudahan dan kemampuannya untuk mengelompokkan data besar dan outlier dengan sangat cepat. Dalam metode K-Means setiap data harus termasuk ke cluster tertentu pada suatu tahapan proses, pada tahapan proses berikutnya dapat berpindah ke cluster yang lain (Sentika, 2016 : 500).

Berikut ini adalah langkah-langkah dalam algoritma *K-means* :

1. Tentukan jumlah *cluster* (*k*) pada *data set*
2. Tentukan nilai pusat (*centroid*) Penentuan nilai *centroid* pada tahap awal dilakukan secara random dengan rumus menentukan target awal *k-means*, rumus tersebut digunakan untuk mendapatkan target data atau jarak antara kelompok, yaitu titik pusat awal untuk menghitung algoritma *k-means* iterasi 0 seperti pada persamaan 2.1 berikut ini:

$$\frac{\text{Jumlah Data}}{\text{Jumlah Class} + 1} \dots\dots\dots [2.1]$$

Keterangan :

Jumlah data : Jumlah data yang akan digunakan.

Jumlah *class* : Jumlah kelompok yang telah ditentukan sebelumnya seperti sangat tinggi, tinggi, normal, rendah, sangat rendah. sedangkan pada tahap *iterasi* digunakan rumus rata-rata, pada rumus tersebut dilakukan perhitungan untuk mencari nilai rata-rata seperti pada persamaan 2.2 berikut ini:

$$V_{ij} = \frac{1}{N_i} = \sum_{k=0}^{n_i} x_{kj} \dots\dots\dots [2.2]$$

Keterangan:

$V_{ij}$  : *centroid* rata-rata *cluster* ke-*i* untuk variabel ke-*j*

$N_i$  : jumlah anggota *cluster* ke-*i*

$i, k$  : indeks dari *cluster*

$j$  : indeks dari variabel

$X_{kj}$  : nilai data ke-*k* variabel ke-*j* dalam *cluster* tersebut.

1. Pada masing-masing *record*, hitung jarak terdekat dengan *centroid* Jarak *centroid* yang digunakan adalah *Euclidean Distance*, dengan rumus seperti pada persamaan 2.3:

$$De = \sqrt{(x_i - s_i)^2 + y_i - t_i)^2} \dots\dots\dots [2.3]$$

Keterangan:

$De$  : *Euclidean Distance*

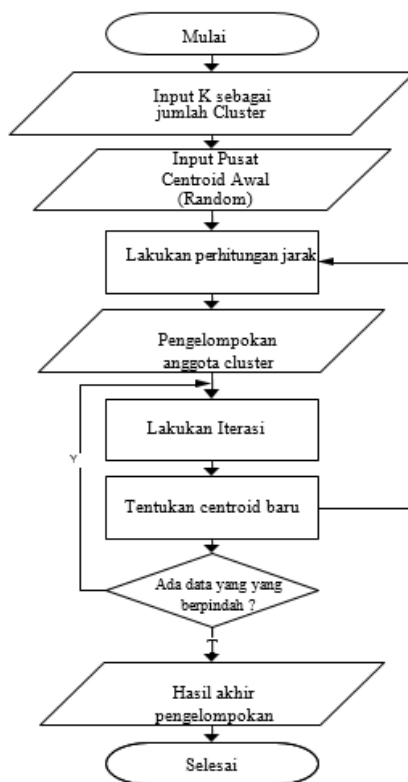
$I$  : banyaknya objek

$(x, y)$  : koordinat objek

$(s, t)$  : koordinat *centroid*

2. Kelompokkan objek berdasarkan jarak ke *centroid* terdekat.

3. Ulangi langkah ke-2, lakukan *iterasi* hingga *centroid* bernilai optimal.



Gambar 2. Flowchart Algoritma K-Means

### 3. ANALISIS DAN HASIL

#### 3.1 Analisis

Banyaknya persaingan didunia bisnis khususnya dalam dunia objek wisata, menuntut CV. Hairos Indah untuk menemukan strategi yang dapat meningkatkan usaha di bidang objek wisata terutama harus meningkatkan pelayanan terbaik dan pembaharuan wahana-wahana permainan yang lebih menarik. CV. Hairos Indah merupakan perusahaan yang bergerak dibidang objek wisata, yang terdiri dari wahana air (Waterpark), timezone, dan taman (Theme Park). Untuk meningkatkan pelayanan, pihak manajemen dari CV. Hairos Indah harus melakukan berbagai kebijakan-kebijakan dengan tujuan menarik para klien (Customer) untuk berkunjung ke CV. Hairos Indah, untuk meningkatkan pendapatan (In come) perusahaan. Kebijakan-kebijakan yang dapat dilakukan antara lain adalah membuat promosi, meningkatkan mutu pelayanan dan selalu memperhatikan perkembangan serta memperbaharui wahana yang lebih menarik agar tetap menjadi objek wisata terfavorit dan tidak membosankan.

Namun untuk melakukan kebijakan-kebijakan tersebut, pihak manajemen harus mengetahui terlebih dahulu tentang hal-hal yang membuat para pengunjung senang berkunjung ke CV. Hairos Indah. Baik dari segi pelayanan, fasilitas, harga tiket, serta wahana- wahana pendukung lainnya. Namun dengan banyaknya variasi wahana yang ada di CV. Hairos Indah, membuat pengunjung bingung memilih wahana untuk menikmati liburan. Untuk mengelompokkan wahana apa saja yang paling disenangi dan menjadi favorit dapat diketahui dengan menggunakan pendataan berdasarkan dari nama-nama wahana dan beberapa kriteria penunjang yang ada di CV. Hairos Indah. Untuk memperoleh berbagai informasi tentang wahana- wahana apa saja yang jadi favorit bagi para pengunjung, maka untuk mengetahui itu semua agar lebih mudah kita harus melakukan satu proses yang dinamakan data mining. Hal ini dapat dilakukan dengan cara mengetahui keluhan pelanggan untuk meningkatkan pelayanan tersebut dengan menggunakan algoritma K-Means sebagai solusi permasalahan tersebut.

Algoritma dalam penyelesaian masalah ini menggunakan algoritma *K-Means* untuk meningkatkan pelayanan pelanggan berdasarkan hasil dari keluhan pelanggan. Langkah-langkah pada algoritma *K-Means* adalah sebagai berikut:

1. Tentukan  $k$  sebagai jumlah *cluster* yang dibentuk.
2. Tentukan pusat (*centroid*) *cluster* awal. Pusat awal *cluster* didapatkan dari data sendiri bukan dengan menentukan titik baru, yaitu dengan mengacak (*random*) pusat awal dari data.
3. Menghitung jarak hasil perhitungan dengan melakukan perbandingan dan memilih jarak terdekat antara data dengan pusat *cluster*.
4. Lakukan pengelompokan data dengan menemukan kedekatan dengan salah satu *centroid* yang ada. Secara otomatis populasi data tersebut masuk ke dalam kelas yang memiliki *centroid* yang bersangkutan.
5. Lakukan iterasi, kemudian tentukan posisi *centroid* baru.
6. Ulangi langkah 3 (tiga) apabila masih ada data yang berpindah kelompok, atau ada perubahan nilai *centroid* diatas nilai ambang yang ditentukan.

Berikut adalah data Wahana dan Kriteria di CV. Hairos indah, dimana data nilai tersebut akan diolah menggunakan algoritma *K-Means*.

Tabel 1. Tabel Inisialisasi Kriteria

No.	Nama Kriteria	Inisialisasi Kriteria
1	Harga	K1
2	Fasilitas	K2
3	Tingkat Keamanan	K3
4	Usia	K4
5	Tingkat Kepuasan	K5

Tabel 2. Tabel Inisialisasi Bobot Kriteria (Harga)

No	Bobot / Nilai	Keterangan
1	Rp. 60.000,-	1
2	Rp. 20.000,-	2
3	Rp. 15.000,-	3
4	Rp. 10.000,-	4
5	Rp. 5.000,-	5

Tabel 3. Tabel Inisialisasi Bobot Kriteria (Fasilitas)

No	Bobot / Nilai	Keterangan
1	Sangat Lengkap Dan Terbaru	5
2	Lengkap	4
3	Cukup Lengkap	3
4	Terbaru	2
5	Kurang Lengkap	1

Tabel 4. Tabel Inisialisasi Bobot Kriteria (Tingkat Keamanan)

No	Bobot / Nilai	Keterangan
1	Sangat Aman	5
2	Aman	4
3	Cukup Aman	3
4	Kurang Aman	2
5	Tidak Aman	1

Tabel 5. Tabel Inisialisasi Bobot Kriteria (Usia)

No	Bobot / Nilai	Keterangan
1	$\geq 1$ Tahun	5
2	$\geq 17$ Tahun	4
3	$\geq 12$ Tahun	3
4	6-12 Tahun	2
5	1-5 Tahun	1

Tabel 6. Tabel Inisialisasi Bobot Kriteria (Tingkat Kepuasan)

No	Bobot / Nilai	Keterangan
1	Sangat Puas	5
2	Puas	4
3	Cukup Puas	3
4	Kurang Puas	2
5	Tidak Puas	1

Tabel 7. Data Wahana Di CV. Hairos Indah

NO	NAMA WAHANA	ALTERNATIF	K1	K2	K3	K4	K5
1	Water Park	P1	1	5	3	5	5
2	Kolam Renang Mini	P2	2	3	4	5	3
3	Mini Zoo	P3	5	2	2	5	2
4	Sepeda Air	P4	4	1	1	3	4
5	Kapal Mesin Full AC	P5	3	3	3	5	5
6	Kuda Pusing	P6	4	1	5	2	3
7	Mandi Bola	P7	4	2	4	1	4
8	Kolam Pancing	P8	4	2	1	4	3
9	Sizzler	P9	4	1	2	3	3
10	Happy Dragon	P10	4	1	3	2	4
11	Bum-Bum Car	P11	4	1	2	3	3
12	Baling-baling	P12	4	1	2	3	2
13	Odong- Odong	P13	5	2	2	5	3
14	Mini Train	P14	5	1	4	1	1
15	Banteng Gila	P15	4	1	3	3	3
16	Sado	P16	4	1	2	5	2
17	Mini Flight	P17	4	1	3	1	3
18	Flying Fox	P18	4	2	2	3	3
19	Kapal Mesin Naga	P19	4	2	2	5	4
20	Taman	P20	5	4	5	5	2

(Sumber : CV. Hairos Indah)

### 3.2.1 Perhitungan Manual Menggunakan Algoritma K-Means

Setelah data terkumpul, dilakukan perhitungan manual sebagai berikut:

#### 1. Iterasi ke-1

##### a. Penentuan pusat (*centroid*) awal *cluster*

Untuk menentukan pusat (*centroid*) awal ditentukan dengan mengacak (*random*) dari data nilai yang sudah ada. Pada kasus ini pusat (*centroid*) awalnya adalah :

Tabel 8. Pusat (*Centroid*) Awal *Cluster*

Centroid 1	5	2	2	5	2
Centroid 2	4	1	3	3	3

#### b. Perhitungan jarak dengan pusat *cluster* dengan rumus :

$$d(x,y) = \|x - y\| = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad ; i=1,2,3,\dots,n$$

Perhitungan jarak dari data ke-1 terhadap pusat *cluster* adalah :

$$D(1,1) = \sqrt{\begin{matrix} (1-5)^2 + (5-2)^2 + (3-2)^2 + \\ (5-5)^2 + (5-2)^2 \end{matrix}}$$

$$= 5.92$$

$$D(1,2) = \sqrt{\begin{matrix} (1-4)^2 + (5-1)^2 + (3-3)^2 + \\ (5-3)^2 + (5-3)^2 \end{matrix}}$$

$$= 5.74$$

Keterangan :

- C1 memiliki 6 anggota, yaitu data ke-3, data ke-8, data ke-13, data ke-16, data ke-19, dan data ke-20.
  - C2 memiliki 14 anggota, yaitu data ke-1, data ke-2, data ke-4, data ke-5, data ke-6, data ke-7, data ke-9, data ke-10, data ke-11, data ke-12, data ke-14, data ke-15, data ke-17, dan data ke-18.
  - Status Terdekat menjelaskan hasil perhitungan pengelompokkan *cluster* yang sudah terbentuk.
- c. Perhitungan pusat *cluster* baru.

Setelah di ketahui anggota tiap-tiap cluster kemudian pusat cluster baru di hitung berdasarkan data anggota tiap-tiap cluster sesuai dengan rumus pusat anggota *cluster* yaitu :  $v = \frac{\sum_{i=1}^n x_i}{n}$  ;  $i=1,2,3,\dots,n$ .

Sehingga di dapatkan perhitungan sebagai berikut :

- Karena C1 memiliki 6 anggota, yaitu (data ke-3, data ke-8, data ke-13, data ke-16, data ke-19, dan data ke-20) maka perhitungan pusat *cluster* baru menjadi :

$$C_{11} = \frac{5 + 4 + 5 + 4 + 4 + 5}{6} = 4.50$$

$$C_{12} = \frac{1 + 2 + 4 + 3 + 4 + 4 + 4 + 4 + 4 + 4 + 5 + 4 + 4 + 4}{14} = 3.64$$

- Dan selanjutnya hitung data pada C2 sampai berdasarkan jumlah anggota yang dimiliki. Sehingga di dapat hasil pembentukan pusat (*centroid*) baru sebagai berikut :

Tabel 9. Hasil Pembentukan Pusat *Cluster* Baru

Centroid 1	4.50	2.17	2.33	4.83	2.67
Centroid 2	3.64	1.71	2.93	2.86	3.29

## 2. Iterasi ke-2

### a. Penentuan pusat (*centroid*) awal *cluster*

Untuk menentukan pusat (*centroid*) awal ditentukan dengan mengacak (*random*) dari data nilai yang sudah ada. Pada kasus ini pusat (*centroid*) awalnya adalah :

- Perhitungan jarak dari data ke-1 terhadap pusat *cluster* baru adalah:

$$D(1,1) = \sqrt{(1 - 4.50)^2 + (5 - 2.17)^2 + (3 - 2.33)^2 + (5 - 4.83)^2 + (5 - 2.67)^2}$$

$$= 5.12$$

$$D(1,2) = \sqrt{(1 - 3.64)^2 + (5 - 1.71)^2 + (3 - 2.93)^2 + (5 - 2.86)^2 + (5 - 3.29)^2}$$

$$= 5.03$$

Keterangan :

- C1 memiliki 8 anggota, yaitu data ke-2, data ke-3, data ke-5, data ke-8, data ke-13, data ke-16, data ke-19, dan data ke-20.
  - C2 memiliki 12 anggota, yaitu data ke-1, data ke-4, data ke-6, data ke-7, data ke-9, data ke-10, data ke-11, data ke-12, data ke-14, data ke-15, data ke-17, dan data ke-18.
- b. Karena pada iterasi ke-2 data mengalami perubahan kelompok, maka di bentuk pusat (*centroid*) *cluster* baru yang di hitung berdasarkan data anggota tiap-tiap *cluster* pada iterasi ke-2 sesuai dengan rumus pusat anggota *cluster*. Sehingga di dapatkan perhitungan sebagai berikut :
- Karena C1 memiliki 8 anggota, yaitu data ke-2, data ke-3, data ke-5, data ke-8, data ke-13, data ke-16, data ke-19, dan data ke-20.
  - Maka perhitungan pusat *cluster* baru menjadi :

$$C_{11} = \frac{2 + 5 + 3 + 4 + 5 + 4 + 4 + 5}{8} = 4.00$$

$$C_{12} = \frac{1 + 4 + 4 + 4 + 4 + 4 + 4 + 4 + 5 + 4 + 4 + 4}{12} = 3.83$$

- Dan selanjutnya hitung data pada C2 berdasarkan jumlah anggota yang dimiliki. Sehingga di dapat hasil pembentukan pusat (*centroid*) baru sebagai berikut :

Tabel 10. Hasil Pembentukan Pusat *Cluster* Baru

Centroid 1	4.00	2.38	2.63	4.85	3.00
Centroid 2	3.83	1.50	2.63	2.50	3.17

## 3. Iterasi ke-3

- .hingga posisi data tidak mengalami perubahan.
- Perhitungan jarak dari data ke-1 terhadap pusat *cluster* baru adalah:

$$D(1,1) = \sqrt{\frac{(1 - 4.00)^2 + (5 - 2.38)^2 + (3 - 2.63)^2 + (5 - 4.85)^2 + (5 - 3.00)^2}{(5 - 4.85)^2 + (5 - 3.00)^2}}$$

$$= 4.47$$

$$D(1,2) = \sqrt{\frac{(1 - 3.83)^2 + (5 - 1.50)^2 + (3 - 2.63)^2 + (5 - 2.50)^2 + (5 - 3.17)^2}{(5 - 2.50)^2 + (5 - 3.17)^2}}$$

$$= 5.47$$

## Keterangan :

- C1 memiliki 9 anggota, yaitu data ke-1, data ke-2, data ke-3, data ke-5, data ke-8, data ke-13, data ke-16, data ke-19, dan data ke-20.
- C2 memiliki 11 anggota, yaitu data ke-4, data ke-6, data ke-7, data ke-9, data ke-10, data ke-11, data ke-12 data ke-14, data ke-15, data ke-17, dan data ke-18.
- Karena pada iterasi ke-3 data mengalami perubahan kelompok, maka di bentuk pusat (*centroid*) *cluster* baru yang di hitung berdasarkan data anggota tiap-tiap *cluster* pada iterasi ke-3 sesuai dengan rumus pusat anggota *cluster*. Sehingga Karena C1 memiliki 9 anggota, yaitu data ke-1, data ke-2, data ke-3, data ke-5, data ke-8, data ke-13, data ke-16, data ke-19, dan data ke-20. Maka perhitungan pusat *cluster* baru menjadi :

$$C_{11} = \frac{1 + 2 + 5 + 3 + 4 + 5 + 4 + 4 + 5}{9}$$

$$= 3.67$$

$$C_{12} = \frac{4 + 4 + 4 + 4 + 4 + 4 + 4 + 4 + 5 + 4 + 4 + 4}{11}$$

$$= 4.09$$

- Dan selanjutnya hitung data pada C2 berdasarkan jumlah anggota yang dimiliki. Sehingga di dapat hasil pembentukan pusat (*centroid*) baru sebagai berikut :

Tabel 11. Hasil Pembentukan Pusat *Cluster* Baru

Centroid 1	3.67	2.67	2.67	4.89	3.22
Centroid 2	4.09	1.18	2.82	2.27	3.00

## 4. Iterasi ke-4

- Hitung jarak dengan pusat (*centroid*) *cluster* baru yang telah di bentuk hingga posisi data tidak mengalami perubahan.

Perhitungan jarak dari data ke-1 terhadap pusat *cluster* baru adalah:

$$D(1,1) = \sqrt{\frac{(1 - 3.67)^2 + (5 - 2.67)^2 + (3 - 2.67)^2 + (5 - 4.89)^2 + (5 - 3.22)^2}{(5 - 4.89)^2 + (5 - 3.22)^2}}$$

$$= 3.98$$

$$D(1,2) = \sqrt{\frac{(1 - 4.09)^2 + (5 - 1.18)^2 + (3 - 2.82)^2 + (5 - 2.27)^2 + (5 - 3.00)^2}{(5 - 2.27)^2 + (5 - 3.00)^2}}$$

$$= 5.97$$

- Dan seterusnya di lakukan perhitungan jarak untuk data ke-2 sampai data ke-20. Kemudian akan didapatkan hasil perhitungan jarak setiap data terhadap pusat *cluster* baru sebagai berikut :



Tabel 12. Hasil Perhitungan Jarak Setiap Data Terhadap Pusat *Cluster* Baru Pada Iterasi Ke-4

No	C1	C2
1	3.98	5.97
2	2.17	4.07
3	2.04	3.26
4	3.14	2.21
5	1.96	4.00
6	4.09	2.21
7	4.25	2.16
8	2.05	2.64
9	2.64	1.12
10	3.46	1.07
11	2.64	1.12
12	2.90	1.50
13	1.65	3.10
14	5.14	2.80
15	2.57	0.78
16	2.20	3.03
17	4.26	1.30
18	2.15	1.37
19	1.28	3.13
20	3.24	4.69

## b. Pengelompokan data

Jarak hasil perhitungan akan dilakukan perbandingan dan di pilih jarak terdekat antara data dengan pusat *cluster*, jarak ini menunjukkan bahwa data tersebut berada dalam satu kelompok dengan pusat *cluster*. Berikut ini adalah hasil pengelompokan data berdasarkan jarak terdekat dengan pusat *cluster*.

Tabel 13. Hasil Pengelompokan Data Berdasarkan Jarak Terdekat Terhadap Pusat *Cluster* Baru Pada Iterasi Ke-4

NO	C1	C2
1	Terdekat	
2	Terdekat	
3	Terdekat	
4		Terdekat
5	Terdekat	
6		Terdekat
7		Terdekat
8	Terdekat	
9		Terdekat
10		Terdekat
11		Terdekat
12		Terdekat

13	Terdekat	
14		Terdekat
15		Terdekat
16	Terdekat	
17		Terdekat
18		Terdekat
19	Terdekat	
20	Terdekat	

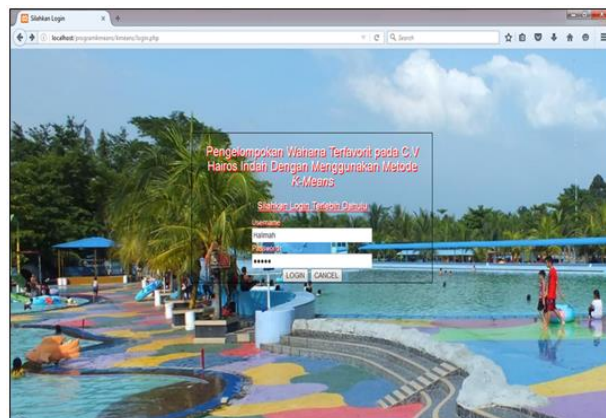
Keterangan :

- C1 memiliki 8 anggota, yaitu data ke-1, data ke-2, data ke-3, data ke-5, data ke-13, data ke-16, data ke-19, dan data ke-20.
- C2 memiliki 12 anggota, yaitu data ke-4, data ke-6, data ke-7, data ke-8, data ke-9, data ke-10, data ke-11, data ke-12, data ke-14, data ke-15, data ke-17, dan data ke-18.
- c. Karena pada iterasi ke-4 data sudah tidak berubah kelompok maka iterasi di hentikan atau selesai. Dan ditemukan hasil sebagai berikut:

Tabel 14. Hasil Proses *K-Means*

No	Nama Wahana	C1	C2
1	Water Park	Terfavorit	
2	Kolam Renang Mini	Terfavorit	
3	Mini Zoo	Terfavorit	
4	Sepeda Air		Tidak Terfavorit
5	Kapal Mesin Full AC	Terpavorit	
6	Kuda Pusing		Tidak Terpavorit
7	Mandi Bola		Tidak Terpavorit
8	Kolam Pancing	Terpavorit	
9	Sizzler		Tidak Terpavorit
10	Happy Dragon		Tidak Terpavorit
11	Bum-Bum Car		Tidak Terpavorit
12	Baling-baling		Tidak Terpavorit
13	Odong- Odong	Terpavorit	
14	Mini Train		Tidak Terpavorit
15	Banteng Gila		Tidak Terpavorit
16	Sado	Terpavorit	
17	Mini Flight		Tidak Terpavorit
18	Flying Fox		Tidak Terpavorit
19	Kapal Mesin Naga	Terpavorit	
20	Taman	Terpavorit	

### 3.2 Hasil



Gambar 3. Tampilan Halaman Login

Halaman Wahana merupakan tampilan yang berisi informasi tentang data nama-nama wahana yang terdapat pada CV.Hairos Indah Bentuk tampilan halaman wahana dapat dilihat pada gambar di bawah ini

No.	ID Wahana	Nama Wahana	Action
1.	P1	Water Park	[X] [✓]
2.	P2	Kolam Renang Mini	[X] [✓]
3.	P3	Mini Zoo	[X] [✓]
4.	P4	Sepeda Air	[X] [✓]
5.	P5	Kapal Mesin Full AC	[X] [✓]

Gambar 4. Tampilan Halaman Wahana

Halaman Kriteria merupakan tampilan yang berisi Kriteria-kriteria yang terdapat pada CV.Hairos Indah Bentuk tampilan halaman Kriteria dapat dilihat pada gambar di bawah ini

No.	ID Kriteria	Nama Kriteria	Action
1.	K1	Harga	[X] [✓]
2.	K2	Fasilitas	[X] [✓]
3.	K3	Tingkat Keamanan	[X] [✓]
4.	K4	Usia	[X] [✓]
5.	K5	Tingkat Kepuasan	[X] [✓]

Gambar 5. Tampilan Halaman Kriteria

Halaman Penilaian merupakan tampilan yang berisi menu-menu yang berfungsi menampilkan form inputan penilaian. Bentuk tampilan halaman penilaian dapat dilihat pada gambar di bawah ini

No.	ID Wahana	K1	K2	K3	K4	K5
1.	P1 - Water Park	1	5	3	5	5
2.	P2 - Kolam Renang Mini	2	3	4	5	3
3.	P3 - Mini Zoo	5	2	2	5	2
4.	P4 - Sepeda Air	4	1	1	3	4
5.	P5 - Kapal Mesin Full AC	3	3	3	5	5
6.	P6 - Kuda Pusing	4	1	5	3	2
7.	P7 - Mandi Bola	4	2	4	1	4
8.	P8 - Kolam Pancing	4	2	1	4	3
9.	P9 - Sizzler	4	1	2	3	3

Gambar 6. Tampilan Halaman Penilaian

Tampilan halaman perhitungan merupakan halaman untuk melakukan perhitungan K\_Means. Bentuk Tampilan halaman perhitungan dapat dilihat pada gambar di bawah ini

Centroid 1	5	2	2	5	2
Centroid 2	4	1	3	3	3

Iterasi Ke 1		
No.	C1	C2
1.	5.92	5.74
2.	3.87	3.61
3.	0	2.83
4.	3.32	2.24
5.	3.87	3.61
6.	4.58	2.24
7.	5	2.65
8.	2	2.45
9.	2.65	1
10.	4	1.41

Pengelompokan Data Ke 1		
No.	C1	C2
1.		✓
2.		✓
3.	✓	
4.		✓
5.		✓
6.		✓
7.		✓
8.	✓	
9.		✓
10.		✓

Gambar 7. Tampilan Halaman Perhitungan

Halaman Laporan merupakan tampilan yang berisi Laporan yang sudah diselesaikan dengan menggunakan metode K-Means. Bentuk tampilan laporan seperti gambar dibawah ini :

No.	ID Wahana	Nama Wahana	C1	C2	Cluster
1.	P1	Water Park	3.98	5.95	Terfavorit
2.	P2	Kolam Renang Mini	2.17	4.01	Terfavorit
3.	P3	Mini Zoo	2.04	3.16	Terfavorit
4.	P4	Sepeda Air	3.14	2.22	Tidak Terfavorit
5.	P5	Kapal Mesin Full AC	1.95	3.98	Terfavorit
6.	P6	Kuda Pusing	3.66	2.45	Tidak Terfavorit
7.	P7	Mandi Bola	4.25	2.25	Tidak Terfavorit
8.	P8	Kolam Pancing	2.05	2.59	Terfavorit
9.	P9	Sizzer	2.64	1.05	Tidak Terfavorit
10.	P10	Happy Dragon	3.45	1.15	Tidak Terfavorit

Gambar 8. Tampilan Halaman Laporan

#### 4. KESIMPULAN



Berdasarkan penelitian yang dilakukan mengenai perancangan dan implementasi penerapan *Metode K-Means* dalam Pengelompokan Wahana Terfavorit pada CV. Hairos Indah, maka didapat kesimpulan sebagai berikut:

1. Menggunakan bidang keilmuan Data Mining dengan metode *K-Means* dengan cara melakukan perhitungan berdasarkan nama alternatif dan kriteria-kriteria berdasarkan data dari perusahaan serta melakukan angket dari pengunjung. Sehingga dapat mempermudah pihak instansi dalam menganalisis permasalahan khususnya dapat mempromosikan variasi wahana-wahana yang terdapat pada CV. Hairos Indah serta bertujuan untuk Pengelompokan Wahana Terfavorit dan mempermudah pengunjung untuk memilih wahana terfavorit serta memilih wahana-wahana apa saja yang tepat untuk berlibur.
2. Perancangan program pada skripsi ini menggunakan aplikasi berbasis *Web Programming* dengan mengkoneksikan *Database MySQL*. Dengan perancangan tersebut dapat mempermudah dalam melakukan pengelompokan wahana terfavorit.
3. Dengan menggunakan aplikasi berbasis *Web Programming* maka dapat mempermudah instansi dalam menampilkan laporan data dan pengunjung juga lebih dipermudah mengakses untuk mengetahui wahana-wahana apa saja yang terdapat pada CV. Hairos Indah secara *online*.
4. Dengan menggunakan sistem terkomputerisasi pada proses Pengelompokan Wahana Terfavorit di CV. Hairos Indah dengan menggunakan metode *K-Means*, maka proses perolehan data akan lebih cepat dan akurat serta meminimalisir adanya penilaian yang tidak jelas karena seluruh data diproses oleh sistem.

#### DAFTAR PUSTAKA

- [1] Hermawati, Fazar Astuti. 2013. Data Mining. Yogyakarta : Andi
- [2] <https://dickynofriansyah.wordpress.com>
- [3] Rosa, A.S dan M. 2015. Salahuddin Rekayasa Perangkat Lunak. Bandung:Informatika
- [4] <https://fti.uajy.ac.id>
- [5] R.H. Sianipar, Eng. Pemrograman Database Menggunakan MySQL. Yogyakarta: Andi
- [5] Priyanto Hidayatullah dan Jauhari Khairul Kawistara. 2014. Pemrograman Web. Bandung : Informatika

**BIOGRAFI PENULIS**

	<p><b>Asyahri Hadi Nasyuha, S.Kom, M.Kom</b>, Pria kelahiran Medan 29 April 1986 merupakan peneliti yang menyenangi bidang keilmuan sistem pendukung keputusan. Pengampu beberapa mata kuliah diantaranya : Pemograman Visual Basic I,II, III, Sistem Pendukung Keputusan, Aplikasi Terapan (Excel), Animasi (Macro Media Flash 8), Paket Program Niaga (Offices Word), Pemrograman Web I, Desain Grafis (Corel Draw), dan beberapa matakuliah pendukung Sistem Informasi. Lulusan D3 Manajemen Informatika STMIK Triguna Dharma pada Tahun 2011, lulusan S1 Sistem Informasi pada Tahun 2012, lulusan S2 Sistem Informasi Universitas Putra Indonesia YPTK Padang dan sedang menjalani pendidikan S3 Fakultas Teknologi Kejuruan Universitas Negeri Padang. Menjadi Narasumber Untuk Tema Ilmu Adalah Investasi &amp; Ethical Hacking.</p>
	<p><b>Zulham, S.Kom, M.Kom</b> Pria kelahiran, Medan, 30 Juli 1982. S1 Tamat di STMIK Logika Medan tahun 2007. S2 Tamat 2015 di Univ. Putra Indonesia YPTK Padang. Saat ini sedang ambil Doktorat di Universitas Negeri Padang. Matakuliah yg diambil: Pengolahan Citra, Pengenalan Pengolahan Dasar Elektronik, Pemograman Database Client Server, Struktur Data. Anak ke-4 dari 4 bersaudara.</p>