

Perancangan Sistem Rekomendasi Paket Wisata Alam Kabupaten Rembang Jawa Tengah Dengan Metode K-Means

Design of a Recommendation System for Nature Tourism Packages in Rembang Regency, Central Java with the K-Means Method

1st Vinsa Arinaldi
Fakultas Rekayasa Industri
Universitas Telkom
Bandung, Indonesia

vinsaarinaldi@student.telkomuniversity.ac.id

2nd Rayinda Pramuditya Soesanto
Fakultas Rekayasa Industri
Universitas Telkom
Bandung, Indonesia

raysoesanto@telkomuniversity.ac.id

3rd Hilman Dwi Anggana
Fakultas Rekayasa Industri
Universitas Telkom
Bandung, Indonesia

hilmandwianggana@telkomuniversity.ac.id

Abstrak—Pariwisata merupakan sektor yang berpengaruh pada pertumbuhan ekonomi. Berdasarkan Kemenparekraf, pariwisata mengalami penurunan terhadap GDP di Indonesia sehingga dengan memanfaatkan perkembangan ekonomi digital dan kekayaan destinasi wisata di Indonesia, ekonomi Indonesia akan kembali normal. Hal tersebut dapat dilakukan dengan melakukan pemanfaatan daerah yang memiliki potensi wisata alam namun masih jarang dikunjungi. Kabupaten Rembang merupakan salah satu kabupaten di Indonesia yang memiliki banyak tempat wisata alam. Dari beberapa lokasi wisata alam tersebut masih terdapat lokasi yang minim pengunjung, dikarenakan kurangnya dukungan paket wisata yang menjadi kendala bagi wisatawan untuk mengetahui karakteristik objek wisata yang tepat. Tujuan penelitian ini yaitu mengidentifikasi karakteristik tiap objek wisata dan menyusun paket wisata berdasarkan jarak, arah lokasi wisata, biaya, jam operasional, dan ratings tempat wisata tersebut. Metode dalam penelitian ini adalah K-Means Clustering yang didukung dengan menggunakan visualisasi Interface R-Shiny di R-Studio dengan bahasa pemrograman R. Hasil dari penelitian ini adalah melakukan pencarian cluster optimal yang dilakukan sebanyak 8 kali, sehingga setelah dilakukan proses perhitungan silhouette terhadap data, hasil silhouette didapatkan dengan menghasilkan 5 pilihan paket wisata alam di Kabupaten Rembang dengan karakteristik lokasi wisata yang berbeda-beda.

Kata Kunci—sistem rekomendasi, K-Means, paket wisata, kabupaten rembang

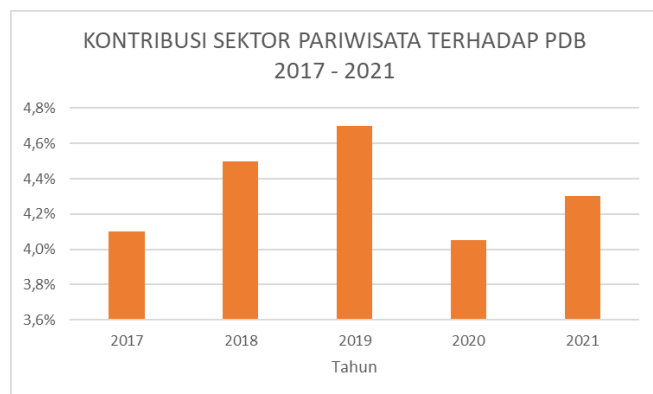
Abstract—Tourism is a sector that affects economic growth. According to the Kemenparekraf, tourism has decreased to GDP in Indonesia, so that by taking advantage of the development of the digital economy and the wealth of tourist destinations in Indonesia, the Indonesian economy will return to normal. This

can be done by utilizing areas that have natural tourism potential but are still rarely visited. Rembang Regency is one of the regencies in Indonesia that has many natural tourist. From some of natural tourist sites, there are still locations that have minimal visitors, due to the lack of support for tour packages which is an obstacle for tourists to find out the characteristics of the right tourist attraction. The purpose of this study is to identify the characteristics of each tourist attraction and compile tour packages based on distance, direction of tourist sites, costs, operating hours, and ratings of the tourist attractions. The method in this study is K-Means Clustering which is supported by using the visualization of the R-Shiny Interface in R-Studio with the R programming language. The results of this study are to search for the optimal cluster which is carried out 8 times, so that after the silhouette calculation process is carried out on the data, the silhouette results are obtained by producing 5 choices of natural tourism packages in Rembang Regency with different characteristics of tourist sites.

Keywords—system recommendation, K-Means, tour package, rembang regency

I. PENDAHULUAN

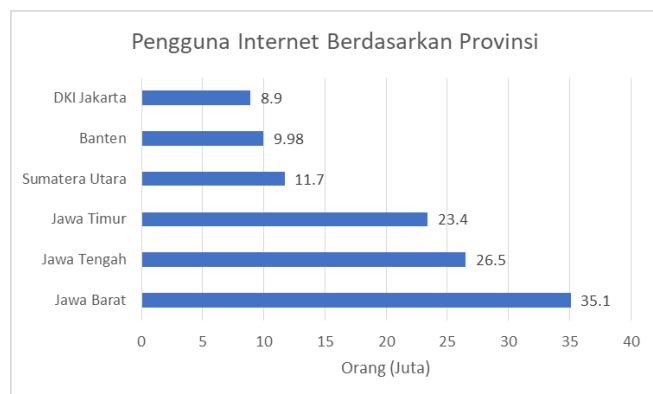
Bidang pariwisata merupakan salah satu contoh sektor ekonomi unggulan dalam pembangunan suatu negara, terutama bagi negara Indonesia. Hal ini dibuktikan dari Kementerian Pariwisata dan Ekonomi Kreatif (Kemenparekraf) yang menyatakan bahwa kontribusi sektor pariwisata di Indonesia terhadap Produk Domestik Bruto (PDB) pada tahun 2017 hingga 2019 terus meningkat hingga mencapai 4,7%. seperti yang dapat dilihat pada Gambar I.



GAMBAR I.1

(Kontribusi Sektor Pariwisata Terhadap PDB 2017-2021)

Untuk memenuhi target tersebut tentu membutuhkan suatu sistem yang dapat membantu sektor pariwisata, salah satunya dengan memanfaatkan adanya kemajuan teknologi menggunakan internet yang diharapkan dapat memberikan dampak yang sangat luar biasa, dikarenakan menurut hasil survei Asosiasi Penyelenggara Jasa Internet Indonesia (APJII) periode 2019-kuartal II/2020, Jawa Tengah menempati peringkat kedua setelah Jawa Barat sebagai provinsi sebagai pengguna internet terbanyak di Indonesia sebanyak 26,5 juta orang seperti pada Gambar I.2.



GAMBAR I.2

(Jumlah Pengguna Internet Berdasarkan Provinsi di Indonesia (Juta))

Berdasarkan fakta bahwa Jawa Tengah memiliki peringkat kedua dalam penggunaan internet di Indonesia, maka hal tersebut dapat menjadi peluang untuk pariwisata di Jawa Tengah, khususnya pariwisata di Kabupaten Rembang. Kabupaten Rembang memiliki tempat wisata alam yang cukup beragam dan menarik untuk didatangi oleh wisatawan. Tempat wisata alam tersebut dibagi menjadi 3 wilayah yang diantaranya yaitu daerah Kota Rembang bagian utara, timur, dan barat seperti yang dapat dilihat pada Tabel I.1.

TABEL I.1

(Tempat Wisata Alam di Kabupaten Rembang)

Daerah di Rembang	Nama Tempat Wisata Alam
Utara	Dampo Awang
	Dermaga Rembang
	Pantai Pandean
	Kartini

TABEL I.1

(Tempat Wisata Alam di Kabupaten Rembang (Lanjutan))

Timur	Watu Layar Lasem
	Pantai Jatisari
	Pantai KarangJahe
	Pantai Caruban
	Pasir Putih Pantai Nyamplung
	Mangrove
	Binangun
	Mangrove Desa Tireman
	Pantai Cemara kasih
	Pantai Sluke
	Pantai Dasun
	Pantai Balongan
Barat	Jembatan merah
	Pasir Putih Wates
	Dermaga Cinta wates

Berdasarkan Tabel I.1, diketahui bahwa terdapat beragam destinasi pariwisata di Kabupaten Rembang, salah satu contohnya yaitu wisata alam. Setiap wisatawan yang datang tentu memiliki kebutuhan yang berbeda, sedangkan daya tarik tempat wisata yang tersebar di Kabupaten Rembang juga memiliki keunikan yang berbeda pula, sehingga masing-masing objek wisata tersebut memiliki kapasitas yang berbeda dalam memenuhi kebutuhan wisatawan. Hal tersebut tentu akan membuat para calon wisatawan akan menghadapi beberapa masalah saat merencanakan perjalanan mereka sendiri akibat terlalu banyaknya informasi dari sumber daya online yang isinya mungkin tidak relevan dan terlalu banyak pilihan tanpa adanya pengambilan keputusan yang tepat dikarenakan proses pencarian menjadi kompleks, sehingga akan memakan waktu yang lama bagi calon wisatawan untuk memutuskan tempat apa saja yang akan mereka datangi, bahkan mungkin menyebabkan calon wisatawan tersebut tidak dapat membuat keputusan akhir dalam pemilihan wisatanya.

Pada kegiatan pariwisata, hal utama yang harus diperhatikan adalah bagaimana meningkatkan kepuasan wisatawan dan memberikan pengalaman terbaik bagi mereka. Urutan kegiatan pariwisata biasanya meliputi kegiatan pra-perjalanan, perjalanan, dan pasca-perjalanan. Pengalaman pra-perjalanan diantaranya yaitu rencana perjalanan, perencanaan, dan pencarian rute perjalanan. Proses perjalanan tersebut mencakup mengunjungi tempat wisata dan perjalanan antar lokasi. Lalu, kegiatan pasca-perjalanan mencakup evaluasi dan umpan balik tentang pengalaman pariwisatanya. Dalam keseluruhan kegiatan pariwisata, pengalaman pra-perjalanan merupakan faktor utama untuk mempengaruhi kepuasan wisatawan. Oleh karena itu, evaluasi subjektif mereka mengenai pra-perjalanan kegiatan pariwisata sangat penting dilakukan walaupun akan menghabiskan sejumlah waktu tertentu agar dapat merancang dan mengetahui rute wisata yang sesuai dengan kebutuhan dan keinginan mereka.

Maka dari itu, dengan adanya sistem rekomendasi paket wisata di Kabupaten Rembang, para calon wisatawan dapat

memilih untuk merencanakan perjalanan mereka sendiri menggunakan informasi yang tersedia secara online dengan mempertimbangkan informasi tepat yang mereka butuhkan, sehingga para calon wisatawan dapat memiliki banyak fleksibilitas dalam memilih tempat wisata alam untuk dikunjungi dengan waktu yang singkat dengan tetap memperhatikan waktu perjalanan serta pengalaman pariwisata yang maksimal.

II. KAJIAN TEORI

A. Clustering

Clustering merupakan sekumpulan data yang bersifat homogen, yaitu setiap bagian mempunyai kesamaan antar satu dan yang lain. Kualifikasi dari *cluster* harus memiliki data yang homogen dan tidak sama dengan data yang lainnya, sedangkan data yang berbeda akan berkumpul dengan kluster yang berbeda pula. [1]

B. K-Means Clustering

Pengelompokan data *K-Means clustering* merupakan teknik eksplorasi suatu kumpulan data yang memiliki objek karakteristik yang sama dimana di kelompokkan menjadi satu untuk mengklasifikasi pengolahan data lebih lanjut. Algoritma *K-Means* merupakan suatu *cluster analysis* non hirarki yang mengelompokkan suatu data berdasarkan karakteristiknya kedalam satu atau lebih *cluster*, data yang memiliki kesamaan variabel dikelompokkan dalam suatu *cluster* yang sama dan sebaliknya. [1]

C. Metode Silhouette Coefficient

Metode *silhouette coefficient* merupakan metode untuk melihat kualitas dan kekuatan dari *cluster*. Metode tersebut merupakan gabungan dari dua metode yaitu metode *cohesion* yang berfungsi untuk mengukur seberapa dekat relasi objek dalam sebuah *cluster*. [2]

D. Klasifikasi

Klasifikasi merupakan teknik dari *data mining* yang memiliki fungsi untuk menganalisis kumpulan data yang diberikan. Hal tersebut digunakan untuk mengekstrak model secara akurat untuk mendefinisikan kelas data penting dalam kumpulan data yang diberikan [3].

E. Data Mining

Data Mining adalah merupakan suatu teknik dalam pengambilan keputusan suatu informasi dengan pola yang terdapat pada sekumpulan data yang ada di database dengan cakupan jumlah yang cukup besar. Terdapat juga di suatu jurnal ilmiah, KDD adalah nama lain dari data mining. [1]

F. Perancangan Sistem

Perancangan sebuah sistem adalah sebuah kegiatan merancang dan menentukan bagaimana cara mengolah sistem informasi dari hasil analisa sistem sehingga dapat terpenuhi kebutuhan si pengguna diantaranya *user interface*, data dan aktivitas proses [4].

G. Perancangan Sistem

Basis data merupakan suatu koleksi terdistribusi dari data-data yang saling terintegrasi secara logika dan menjadi suatu penggambaran dari data yang dirancang untuk memenuhi kebutuhan informasi dari sebuah sistem. Pada umumnya data dalam basis data bersifat *integrated* dan *shared* yang memiliki arti yaitu sebagai penggabungan beberapa file yang berbeda dan merupakan

data individual yang dapat digunakan secara bersamaan dengan user yang berbeda [5].

H. Paket Wisata

Paket Wisata merupakan suatu perjalanan wisata terdiri dari satu atau beberapa tujuan kunjungan yang dapat disusun dan terdiri dari berbagai fasilitas perjalanan tertentu, serta dijual sebagai harga tunggal yang menyangkut keseluruhan komponen dari perjalanan wisata [6].

I. Pemrograman R dan R-Studio

Dalam pemrograman R terdapat berbagai *tools* statistik linier, *non linier*, *clustering*, klasifikasi, uji statistik klasik dan *tools* teknis grafis yang biasa disebut dengan *package*. Pemrograman R juga merupakan program yang bersifat *open source* dalam mengolah data Statistika berbentuk grafik yang sangat berguna untuk penelitian maupun industri. [1]

J. R-Shiny App

R-Shiny berdasarkan merupakan sebuah *package* dalam pemrograman R yang berupa aplikasi berbasis *website* yang menggunakan paket *shiny* yang berupa pengembangan dari R Studio. Secara garis besar komponen *shiny* terbagi menjadi 2 yaitu *User Interface (UI)* dan *Server* [7]. Berikut merupakan penjelasan mengenai UI dan *Server*:

1. User Interface

User Interface merupakan tampilan visual dari sebuah *website* yang dapat dilihat oleh pengguna atau *user*. UI dapat berupa bentuk, warna, dan tulisan yang di tampilkan dalam *website* tersebut.

2. Server

Server merupakan otak atau *main* dari program yang berfungsi untuk melakukan simulasi dengan berbagai data sesuai dengan pilihan pengguna atau *user* dan selanjutnya akan mengirim hasil ke bagian *output*. Pada umumnya bagian ini didukung oleh berbagai prosedur dengan menggunakan *package* yang telah tersedia di Rstudio [7].

K. Black Box Testing

Black Box Testing merupakan metode pengujian yang berfokus pada spesifikasi fungsionalitas dari perangkat lunak, yang mana pengujian ini memberikan gambaran atas sekumpulan kondisi masukan dan melakukan pengujian pada uraian fungsionalitas dari suatu program [8]. *Black Box Testing* dapat digunakan untuk mendeteksi suatu permasalahan seperti berikut [9]:

1. Fungsi yang salah atau pun hilang.
2. Kesalahan pada antarmuka (*interface*).
3. Kesalahan pada struktur dan basis data.
4. Kesalahan fungsi.
5. Kesalahan deklarasi dan terminasi.

III. METODE

Perancangan model bisnis aplikasi *rndpostudent* menggunakan metode *K-Means* melalui beberapa tahapan yang harus dilakukan, diantaranya:

A. Tahapan Pendahuluan

Tahap pendahuluan pada perancangan sistem rekomendasi tugas akhir ini dilakukan dengan melakukan beberapa proses diantaranya yaitu mengidentifikasi rumusan dari permasalahan, membuat tujuan dan batasan masalah, lalu melakukan studi pendahuluan seperti studi literatur dan juga

studi lapangan. Berikut merupakan uraian dari setiap langkah pada tahap pendahuluan di tugas akhir ini.

1. Identifikasi dan Perumusan Masalah

Untuk mengidentifikasi dan membuat rumusan masalah, dilakukan pengumpulan data primer dengan mendatangi langsung lokasi penelitian dan melakukan wawancara dengan sumber terkait di Kabupaten Rembang mengenai wisata alam. Lalu, untuk mengumpulkan data sekunder digunakan dengan mencari data menggunakan studi literatur.

2. Tujuan Masalah

Tujuan dari dilakukannya penelitian tugas akhir ini adalah untuk merancang sistem rekomendasi paket wisata alam di Kabupaten Rembang, Jawa Tengah.

3. Batasan Masalah dan Asumsi

Batasan masalah dan asumsi pada penelitian tugas akhir memiliki tujuan untuk membatasi segala hal terkait penelitian tugas akhir ini.

4. Studi Pendahuluan

Studi pendahuluan pada penelitian tugas akhir ini dilakukan dengan 2 cara, diantaranya yaitu melakukan studi literatur dengan mencari data melalui jurnal dan *google maps* dan mengumpulkan informasi dengan memperoleh data yang didapatkan dari dinas Kabupaten Rembang, Jawa Tengah dan juga observasi.

B. Tahapan Pengumpulan Data

1. Mencari informasi mengenai wisata alam di Kabupaten Rembang.
2. Mencari interval dan jam operasional tempat wisata alam yang sudah ditentukan.
3. Mencari jarak lokasi wisata alam yang telah ditentukan dengan pusat kota.
4. Mencari arah lokasi wisata alam yang telah ditentukan dengan pusat kota.
5. Mengumpulkan *ratings* pada setiap lokasi wisata alam yang telah ditentukan.

C. Tahapan Perancangan Sistem Terintegrasi

Tahapan ini dilakukan untuk melakukan pengolahan data sehingga dapat merancang sistem rekomendasi paket wisata di Kabupaten Rembang. Berikut merupakan tahapan dalam perancangan sistem rekomendasi paket wisata di Kabupaten Rembang menggunakan *K-Means*.

1. *Import* data ke *R-Studio*
2. Menentukan nilai *K-Means*
3. Menentukan nilai optimal dari hasil *cluster K-Means*
4. Menggunakan metode *Silhouette*

5. Mendapatkan hasil optimal *clustering*
6. Menentukan analisis untuk hasil *cluster*
7. Menentukan hasil kesimpulan dan analisis
8. Hasil Selesai

D. Tahapan Verifikasi dan Validasi

Tahap verifikasi dan validasi merupakan sebuah tahapan agar dapat memastikan hasil perancangan yang telah dirancang apakah sudah sesuai atau belum.

E. Kesimpulan dan Saran

Kesimpulan dan saran merupakan tahapan terakhir pada sistematika perancangan tugas akhir ini yang berisikan sebuah rancangan sistem rekomendasi paket wisata alam untuk daerah Kabupaten Rembang, Jawa Tengah.

IV. HASIL DAN PEMBAHASAN

A. Rancangan Penentuan Paket Wisata Menggunakan Metode K-Means

1. Tahap Pengumpulan Data

Pada tahap ini, dilakukan tahapan pengumpulan data dan informasi yang nantinya akan dibutuhkan dalam proses pengolahan data untuk membuat rancangan sistem rekomendasi paket wisata alam di Kabupaten Rembang. Dalam penelitian ini menggunakan berbagai sumber data yang diperlukan diantaranya yaitu:

a. Data Primer

Data Primer didapatkan dengan cara melakukan wawancara secara langsung dengan PIC (*Person in Charge*) di Kabupaten Rembang mengenai wisata alam di lokasi tersebut. Berikut merupakan data primer yang telah didapatkan yaitu nama-nama tempat wisata alam di Kabupaten Rembang dan interval jam operasional pada masing-masing tempat wisata alam.

b. Data Sekunder

Data sekunder merupakan data yang didapatkan dari sumber-sumber seperti buku, jurnal, dan internet sebagai data pendukung atau penunjang untuk penelitian tugas akhir ini. Berikut merupakan data sekunder yang akan diolah pada penelitian tugas akhir ini diantaranya seperti jarak dan arah masing-masing tempat wisata alam dari pusat kota serta *ratings* tiap tempat wisata tersebut.

Data primer dan sekunder yang telah didapatkan dapat dilihat pada Tabel IV.1. Data-data tersebut akan diolah menggunakan algoritma *K-Means clustering* dengan mengambil nilai dari setiap variabel pada data untuk menentukan paket wisata.

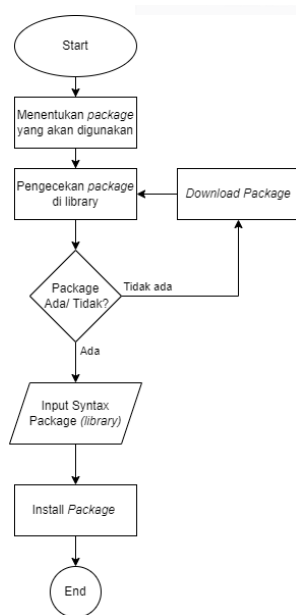
TABEL IV.1
(Data Primer dan Sekunder untuk Penentuan Paket Wisata Alam di Kabupaten Rembang)

Nama Tempat Wisata	Jam Operasional (Jam)	Jarak dari Pusat Kota (Km)	Arah dari Pusat Kota	Ratings
Dampo Awang	13	1,8	UTARA	4
Dermaga Rembang	9	2,5	UTARA	4,2
Pantai Pandean	24	0,5	UTARA	6
Pantai Kartini	10	1	UTARA	4

Pulo Siwalan / Pulau Gede	10	6,6	UTARA	4,6
Watu Layar Lasem	10	17	TIMUR	4,3
Pantai Jatisari	24	24	TIMUR	4,2
Pantai Karang Jahe	10	8,5	TIMUR	4,3
Pantai Caruban	10	11	TIMUR	4,5
Pasir Putih Pantai Nyamplung	9	6,9	TIMUR	4,1
Mangrove Desa Tireman	12	3,6	TIMUR	4,2
Pantai Cemara Kasih	24	6,6	TIMUR	4,1
Pantai Sluke	10	24	TIMUR	4,4
Pantai Dasun	12	13	TIMUR	4,3
Pantai Balongan	24	37,3	TIMUR	4,2
Jembatan Merah	24	5	SELATAN	4
Pasir Putih Wates	12	8,3	BARAT	4,2
Dermaga Cinta Wates	12	8,3	BARAT	4,4

2. Proses Perancangan

a. Memanggil Packages



GAMBAR IV.1
(Pemanggilan Packages)

Dalam melakukan pemanggilan *packages*, terdapat tahapan dalam melakukannya seperti menentukan *package* yang akan digunakan terlebih dahulu. *Packages* yang akan digunakan diantaranya yaitu:

Library(cluster) : Digunakan untuk melakukan analisis *cluster*

Library(tidyverse) : Digunakan sebagai kumpulan *tools data science*

Library(ggplot2) : Digunakan untuk menampilkan plot

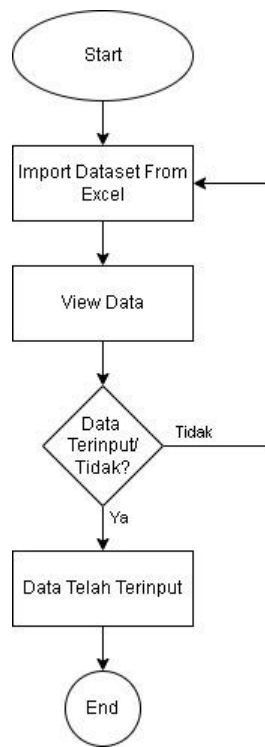
Library(factoextra) : Digunakan untuk visualisasi hasil *cluster*

Library(gridextra) : Digunakan untuk menggabungkan beberapa grafik

Library(readxl) : Digunakan untuk membaca data pada microsoft excel

Library(dplyr) : Digunakan untuk memanipulasi data.

b. Import Data



GAMBAR IV.2
(Import Data)

Dalam melakukan *import* data ke dalam aplikasi R-Studio terdapat beberapa tahapan seperti:

- Import Dataset* ke R-Studio
- View Data*

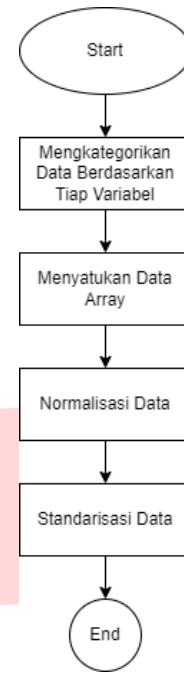
```
data1=as.data.frame(data1)
view(data1)
```

Fungsi dari *view(data1)* adalah untuk menampilkan hasil data pada aplikasi microsoft excel yang telah dilakukan *import* ke dalam aplikasi R-Studio. Tabel IV.2 merupakan hasil dari *view data* pada aplikasi R-Studio.

TABEL IV.2
(Hasil View Data)

	nama Tempat wisata	Jam Operasional	Jarak dari PK	Arah dari PK	Ratings
1	Watu Layar Lasem	10	17.0	TIMUR	4.3
2	Pantai Jatisari	24	24.0	TIMUR	4.2
3	Pantai Karanglahe	10	8.5	TIMUR	4.3
4	pantai Caruban	10	11.0	TIMUR	4.5
5	Pasir putih Wates	12	8.3	BARAT	4.2
6	Dampo awang	13	1.8	UTARA	4.0
7	Dermaga Rembang	9	2.5	UTARA	4.2
8	Pantai Pandean	24	0.5	UTARA	3.2
9	Pantai kartini	10	1.0	UTARA	4.0
10	pulo siwalan / pulau Gede	10	6.6	UTARA	4.6
11	Pasir putih pantai Nyamplung	9	6.9	TIMUR	4.1
12	Dermaga Cinta wates	12	8.3	BARAT	4.4
13	mangrove desa Tireman	12	3.6	TIMUR	4.2
14	Pantai Cemara kasih	24	6.6	TIMUR	4.1
15	Pantai sluke	10	24.0	TIMUR	4.4
16	Pantai Dasun	12	13.0	TIMUR	4.3
17	Pantai Balongan	24	37.3	TIMUR	4.2
18	Jembatan Merah	24	5.0	SELATAN	4.0

c. Mengkategorikan Data



GAMBAR IV.2
(Standarisasi Data)

Langkah pertama untuk mendapatkan standarisasi data yaitu dengan mengkategorikan data berdasarkan setiap variabel yang telah ditentukan. Kemudian melakukan penggabungan data *array* untuk menentukan *summary* data. Tabel IV.3 merupakan kategori data berdasarkan masing-masing variabel.

TABEL IV.3
(Kategori Data Berdasarkan Setiap Variabel)

Syntax	Hasil
x1=as.numeric(data1\$`Jam Operasional`) x1	[1] 10 24 10 10 12 13 9 24 10 10 9 12 12 24 10 12 24 24
x2=as.numeric(data1\$`Jarak dari PK`) x2	[1] 17.0 24.0 8.5 11.0 8.3 1.8 2.5 0.5 1.0 6.6 6.9 8.3 3.6 6.6 [15] 24.0 13.0 37.3 5.0
x3=as.factor(data1\$`Arah dari PK`) x3	[1] TIMUR TIMUR TIMUR TIMUR BARAT UTARA UTARA UTARA UTARA [10] UTARA TIMUR BARAT TIMUR TIMUR TIMUR TIMUR TIMUR SELATAN
x4=as.numeric(data1\$`Ratings`) x4	[1] 4.3 4.2 4.3 4.5 4.2 4.0 4.2 3.2 4.0 4.6 4.1 4.4 4.2 4.1 4.4 4.3 4.2 4.0

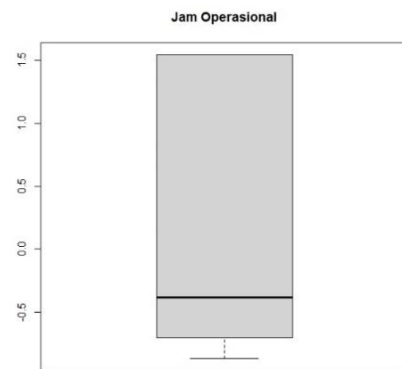
Setelah melakukan pengkategorian data sebelumnya dan mendapatkan hasil tiap data berdasarkan masing-masing variabel, dilakukan penggabungan data *array* dengan menggunakan *syntax* berikut.

```
data = cbind(x1,x2,x3,x4)
```

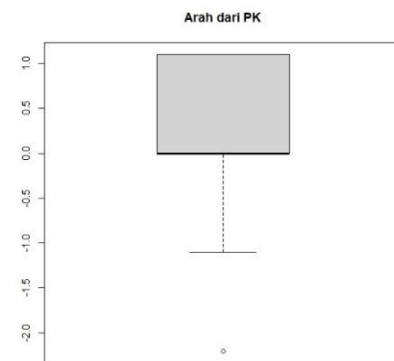
```
data <- na.omit(data)
data <- scale(data)
head(data)
summary(data)
```

TABEL IV.4
(Summary Data)

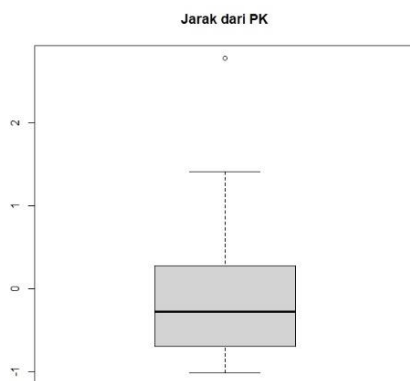
	x1	x2	x3	x4
Watu Layar Lasem	-0.7042319	0.68865791	0.000000	0.4126859
Pantai Jatisari	1.5421788	1.41114664	0.000000	0.0750338
Pantai KarangJahe	-0.7042319	-0.18864984	0.000000	0.4126859
pantai Caruban	-0.7042319	0.06938185	0.000000	1.0879901
Pasir putih Wates	-0.3833161	-0.20929237	-2.203893	0.0750338
Dampo awang	-0.2228582	-0.88017476	1.101946	-0.6002704
Dermaga Rembang	-0.8646898	-0.80792589	1.101946	0.0750338
Pantai Pandean	1.5421788	-1.01435124	1.101946	-3.3014872
Pantai kartini	-0.7042319	-0.96274491	1.101946	-0.6002704
pulo siwalan / pulau Gede	-0.7042319	-0.38475392	1.101946	1.4256422
Pasir putih pantai Nyamplung	-0.8646898	-0.35379012	0.000000	-0.2626183
Dermaga Cinta wates	-0.3833161	-0.20929237	-2.203893	0.7503380
mangrove desa Tireman	-0.3833161	-0.69439195	0.000000	0.0750338
Pantai Cemara kasih	1.5421788	-0.38475392	0.000000	-0.2626183
Pantai sluke	-0.7042319	1.41114664	0.000000	0.7503380
Pantai Dasun	-0.3833161	0.27580721	0.000000	0.4126859
Pantai Balongan	1.5421788	2.78387523	0.000000	0.0750338
Jembatan Merah	1.5421788	-0.54989420	-1.101946	-0.6002704



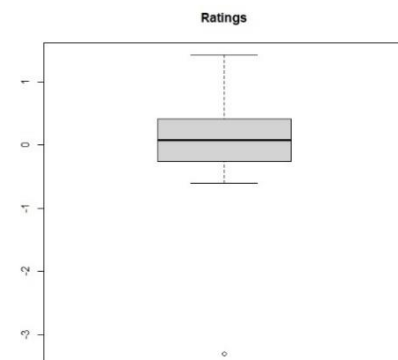
GAMBAR IV.4
(Boxplot pada Variabel Jam Operasional)



GAMBAR IV.5
(Boxplot pada Variabel Arah dari Pusat Kota)

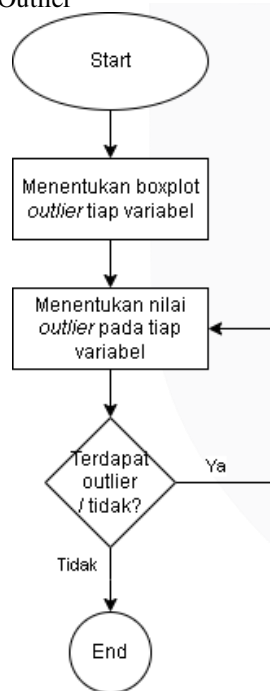


GAMBAR IV.6
(Boxplot pada Variabel Jarak dari Pusat Kota)



GAMBAR IV.7
(Boxplot pada Variabel Ratings)

d. Pengecekan Outlier

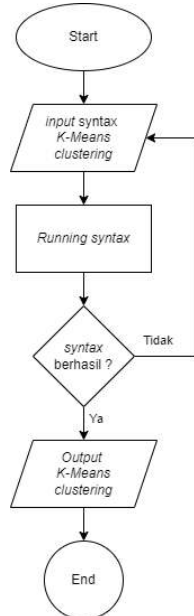


GAMBAR IV.3
(Flowchart Pengecekan Outlier)

Setelah dilakukannya pengecekan pada setiap variabel, seperti jam operasional, jarak dan arah tempat wisata dari pusat kota, dan juga *ratings* didapatkan hasil boxplot pada Gambar IV.4 hingga Gambar IV.7 sebagai berikut.

e. Pencarian Cluster Optimal

Langkah keempat yaitu menggunakan syntax K-Means yang telah dikerjakan pada langkah sebelumnya untuk mencari cluster optimal. Cluster optimal merupakan cluster yang terpilih dikarenakan data-data pada cluster tersebut merupakan data yang memiliki hasil data terbaik, sehingga hasil dari cluster optimal akan menjadi data terpilih pada penelitian ini. Gambar IV.18 menjelaskan proses dari penentuan cluster optimal dengan menggunakan syntax K-Means.



GAMBAR IV.8

(Flowchart Pencarian Cluster Optimal)

Pencarian *cluster* optimal dilakukan sebanyak 8 kali dengan *syntax* berikut ini.

```

k2 <- kmeans(data, centers = 2, nstart = 25)
k2
k3 <- kmeans(data, centers = 3, nstart = 25)
k3
k4 <- kmeans(data, centers = 4, nstart = 25)
k4
k5 <- kmeans(data, centers = 5, nstart = 25)
k5
k6 <- kmeans(data, centers = 6, nstart = 25)
k6
k7 <- kmeans(data, centers = 7, nstart = 25)
k7
k8 <- kmeans(data, centers = 8, nstart = 25)
k8
  
```

f. Metode Silhoutte



GAMBAR IV.9

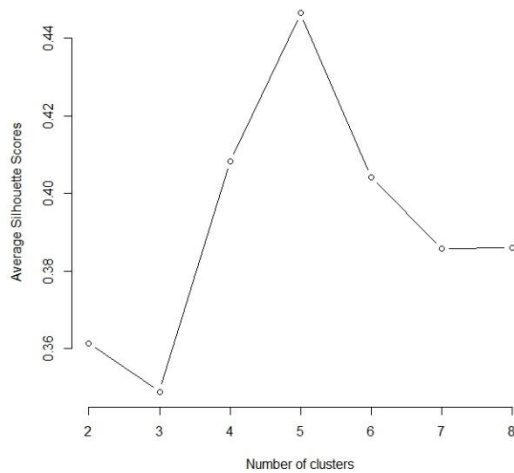
(Flowchart Metode Silhoutte)

Syntax yang dihasilkan sebelumnya akan dilanjutkan dengan menggunakan metode Silhouette untuk mencari cluster optimal diantara 8 cluster dengan *syntax* sebagai berikut:

```

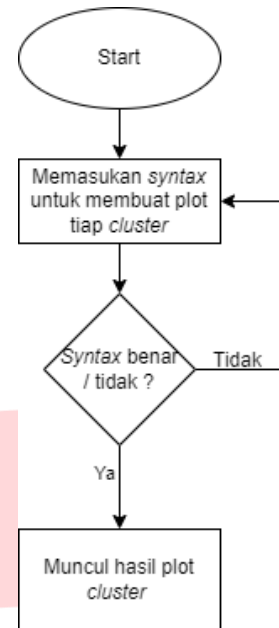
silhouette_score <- function(k){
  km <- kmeans(data, centers = k, nstart=25)
  ss <- silhouette(km$cluster, dist(data))
  mean(ss[, 3])
}
k <- 2:8
avg_sil <- sapply(k, silhouette_score)
plot(k, type='b', avg_sil, xlab='Number of clusters',
      ylab='Average Silhouette Scores', frame=FALSE)
  
```

Setelah dilakukan proses perhitungan silhouette terhadap data maka hasil silhouette didapatkan pada saat K=5. Nilai silhouette berdasarkan gambar grafik IV.10 setelah K=5 nilai cluster cenderung menurun.

g. Hasil *Clustering*

GAMBAR IV.10
(Hasil Nilai *Cluster Metode Silhouette*)

Gambar IV.10 menunjukkan bahwa hasil optimal terdapat pada *cluster* ke-5, sehingga *cluster* ke-5 merupakan *cluster* yang akan digunakan pada langkah berikutnya dalam penentuan paket wisata alam di Kabupaten Rembang.



GAMBAR IV.11
(Flowchart Plot Cluster)

TABEL IV.5
(Hasil *Centroid Data Clustering*)

	X1	X2	X3	X4
Cluster 1	-0.6312965	-0.1661307	0.4007078	0.2899033
Cluster 2	1.5421788	-0.4673241	-0.5509732	-0.4314444
Cluster 3	1.5421788	-1.0143512	1.1019463	-3.3014872
Cluster 4	-0.3833161	-0.2092924	-2.2038927	0.4126859
Cluster 5	1.5421788	2.0975109	0.0000000	0.0750338

Analisis *cluster* menunjukkan bahwa tempat wisata di *cluster* 1 masih mengalami perbedaan yang signifikan pada variabel X1 yaitu pada jam operasional buka tempat wisata tersebut yang menjadikan tempat wisata kurang akan pengunjung yang datang pada jam-jam tertentu ataupun wisatawan merasa kurang lama berada di tempat wisata tersebut.

Untuk *cluster* 2 berdasarkan data centroid untuk variabel X4 memiliki hasil ke-dua paling kecil yaitu -0.4314444, dapat disimpulkan bahwa *ratings* pada *cluster* 2 masih sangat minim dikarenakan wisatawan yang berkunjung masih belum sadar akan pentingnya *ratings* bagi pengelola tempat wisata. Dinas pariwisata Kabupaten Rembang seharusnya lebih meningkatkan penyebaran *feedback* bagi wisatawan yang sedang berkunjung ke tempat wisata yang berada di *cluster* 2 yaitu Pantai Cemara Kasih dan Jembatan Merah Mangrove.

Cluster 3 memiliki jarak terjauh diantara *cluster* lainnya berdasarkan Tabel IV.9 pada variabel X2 yaitu jarak dari

pusat kota dengan hasil -1.0143512. Hasil tertinggi pada *cluster* 5 yaitu pada hasil variabel x2 jarak dari pusat kota dengan nilai 2.0975109, pada *cluster* 5 didapati tempat wisata yang berada pada *cluster* 5 memiliki jarak yang cukup dekat dengan pusat kota. Cluster 5 memiliki kelebihan yaitu dekat dengan penginapan wisatawan yang berkunjung ke Kabupaten Rembang dikarenakan banyaknya hotel yang dibangun di area pusat kota Kabupaten Rembang.

TABEL IV.6
(Hasil *Cluster Optimal*)

5 Cluster				
1	2	3	4	5
Pasir Putih Wates	Pantai Cemara Kasih	Pantai Jatisari	Watu Laya Lasem	Pantai Pandean
Dermaga Cinta Wates	Jembatan Merah	Pantai Balongan	Pantai Karang Jahe	
			Pantai Caruban	

			Dampo Awang	
			Dermaga Rembang	
			Pantai Kartini	
			Pulo Siwalan / Pulau Gede	
			Pasir Putih Pantai Nyamplung	
			Mangrove Desa Tireman	
			Pantai Sluke	
			Pantai Dasun	

h. Visualisasi Data Statistic K-Means menggunakan R-Shiny

R-Shiny merupakan UI (*User Interface*) dari bahasa R, Shiny merupakan *User Interface* berbasis website dimana aplikasi ini akan berjalan dengan menggunakan pemrograman dan algoritma bahasa R. Berikut merupakan langkah-langkah dalam pengoperasian R-Shiny pada penelitian tugas akhir ini.

1. Input Data

2. Preview Data

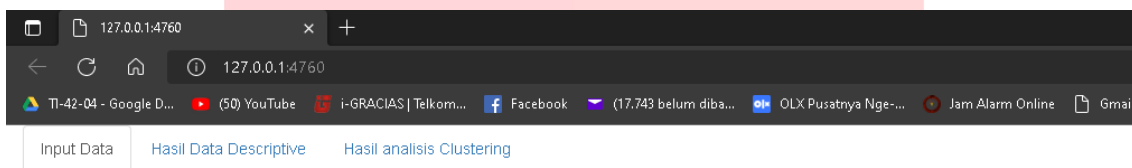
3. Summary Data Statistic

4. Outlier Data

5. Clustering Data Descriptive

6. Hasil Analisis Clustering

Berikut merupakan tampilan dari Data Statistic K-Means menggunakan R-Shiny.



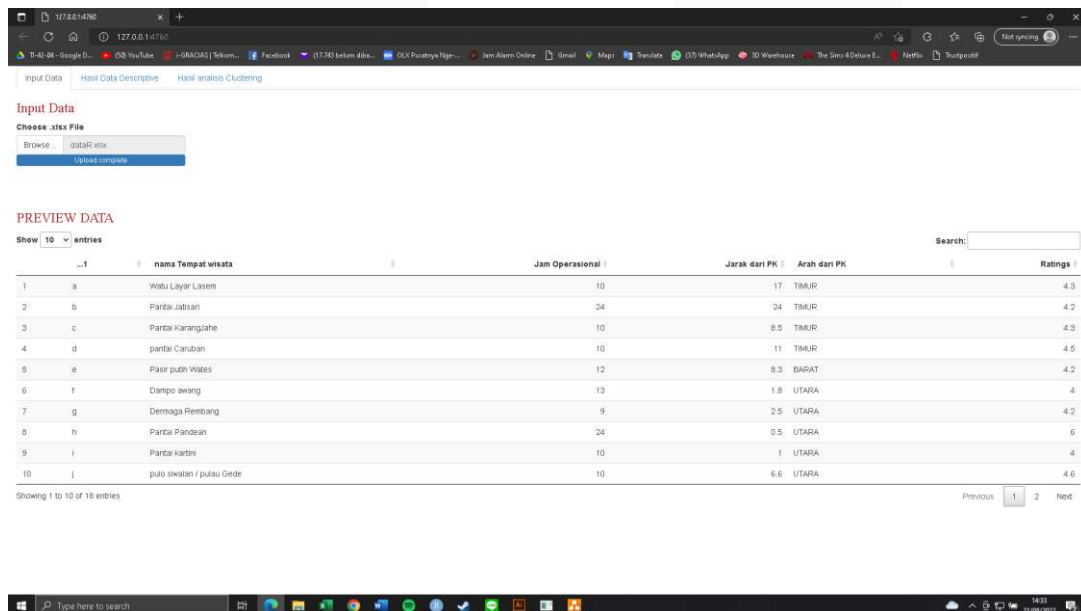
Input Data

Choose .xlsx File

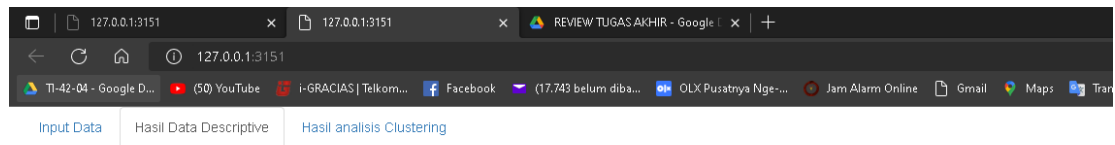
Browse...

No file selected

GAMBAR IV.12
(Input Data)



GAMBAR IV.13
(Preview Data)

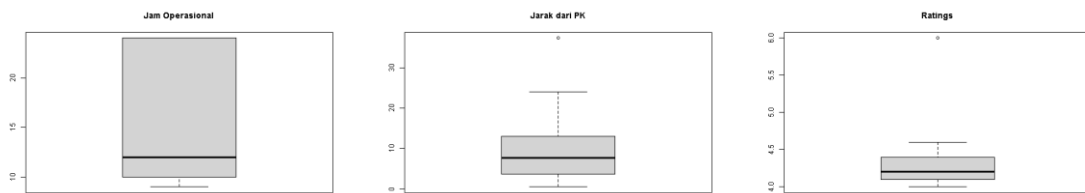


Summary statistik

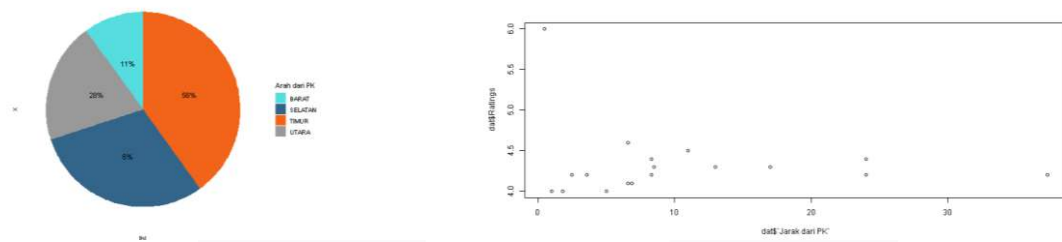
...1	nama Tempat wisata	Jam Operasional	Jarak dari PK	Arah dari PK	Ratings
Length:18	Length:18	Min. : 9.00	Min. : 0.50	Length:18	Min. :4.000
Class :character	Class :character	1st Qu.:10.00	1st Qu.: 3.95	Class :character	1st Qu.:4.125
Mode :character	Mode :character	Median :12.00	Median : 7.60	Mode :character	Median :4.200
		Mean :14.39	Mean :10.33		Mean :4.333
		3rd Qu.:21.25	3rd Qu.:12.50		3rd Qu.:4.375
		Max. :24.00	Max. :37.30		Max. :6.000

GAMBAR IV.22
(Summary Statistic)

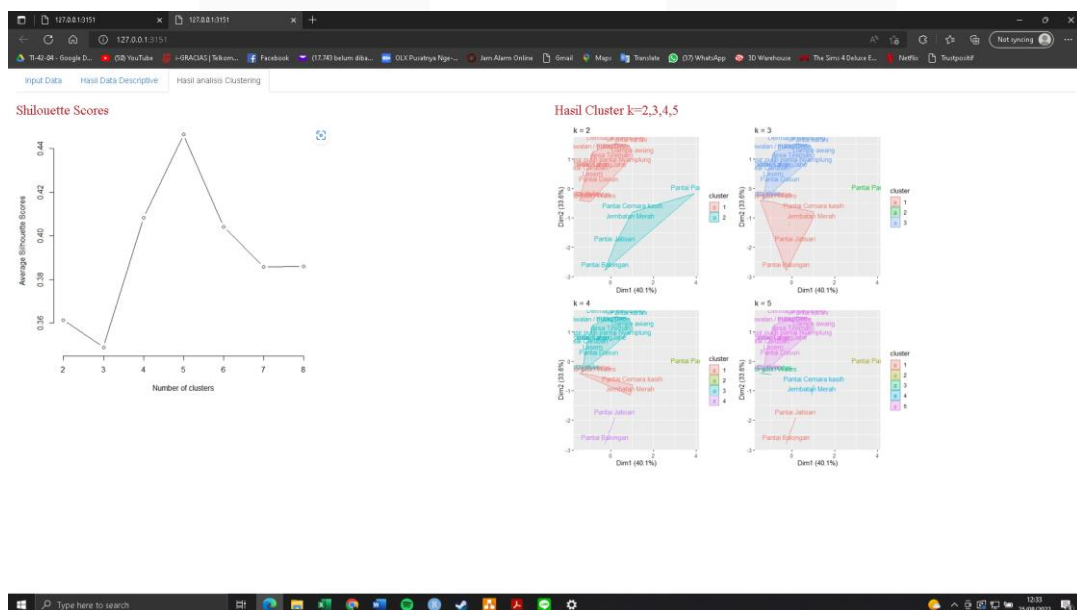
Visualisasi Outlier



GAMBAR IV.23
(Boxplot Outlier)



GAMBAR IV.24
(Hasil Clustering Data)



GAMBAR IV.25
(Page Hasil Analisis Clustering)

3. Verifikasi Hasil Rancangan

a. Penentuan Paket Wisata Alam Menggunakan K-Means

TABEL IV.7
(Verifikasi Hasil Rancangan)

No	Spesifikasi	Hasil Rancangan	Verifikasi Hasil Rancangan
1	Jam Operasional	Menjadi variabel pada penentuan paket wisata alam di Kabupaten Rembang	Terverifikasi
2	Jarak tempat wisata alam dari pusat kota		Terverifikasi

3	Arah tempat wisata alam dari pusat kota		Terverifikasi
4	<i>Ratings</i> tempat wisata alam		Terverifikasi
5	Harga masuk	Terdapat pada analisis tempat wisata alam	Terverifikasi
6	Harga parkir		Terverifikasi

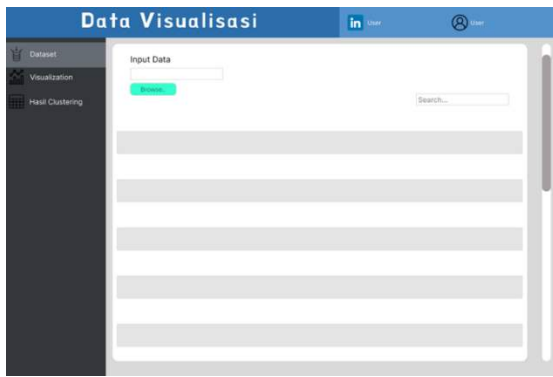
b. Black Box Testing

Berikut merupakan verifikasi menggunakan pengujian *black box* yang dapat dilihat pada Tabel IV.9.

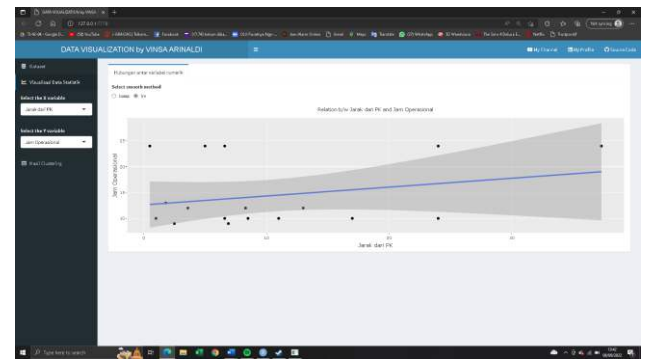
TABEL IV.8
(Pengujian Black Box)

No	Skenario	Test Case	Hasil yang Diharapkan	Hasil Pengujian	Kesimpulan
A. Tab Panel Input Data					
1	Melihat data yang akan dimasukan kedalam R-Shiny	Klik tombol “Browse”	Menampilkan <i>home file folder</i> dari perangkat	Sesuai Harapan	VALID
2	Memasukkan data kedalam R- R-Shiny	Klik tombol “Open”	Dapat menampilkan data	Sesuai Harapan	VALID
B. Tab Panel Hasil Data Descriptive					
3	Menampilkan <i>page</i> berisikan <i>summary</i> data dari variabel data pada <i>page</i> sebelumnya	Klik tab panel “data Visualisasi”	Menampilkan hasil data secara <i>summary statistic</i>	Sesuai Harapan	VALID
4.	Menampilkan <i>page</i> berisikan <i>boxplot outlier</i> data dari variabel data pada <i>page</i> sebelumnya	Klik tab panel “data <i>descriptive</i> ”	Menampilkan <i>boxplot outlier</i> data	Sesuai Harapan	VALID
C. Tab Panel Hasil Data Clustering					
5.	Menampilkan <i>page</i> berisikan <i>silhouette</i> grafik diagram dari variabel data pada <i>page</i> sebelumnya	Klik tab panel “hasil data <i>clustering</i> ”	Menampilkan hasil <i>silhouette</i> grafik diagram data	Sesuai Harapan	VALID
6.	Menampilkan <i>page</i> berisikan <i>plot</i> hasil analisis <i>cluster kmeans</i> data dari variabel data pada <i>page</i> sebelumnya	Klik tab panel “hasil data <i>clustering</i> ”	Menampilkan <i>plot</i> hasil analisis <i>cluster K-Means</i>	Sesuai Harapan	VALID

c. Tampilan Mockup Dashboard Visualisasi



GAMBAR IV.14
(Tampilan Mockup Input Data)



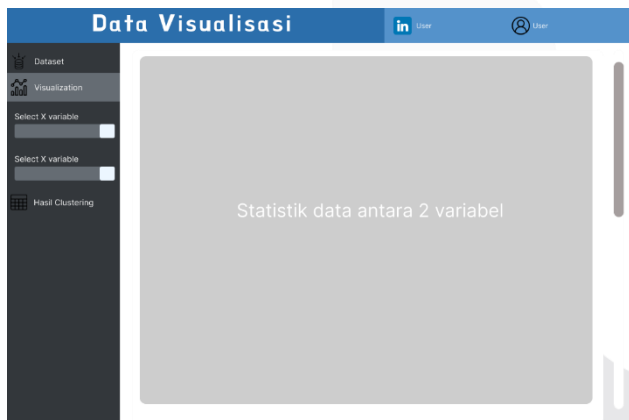
GAMBAR IV.17
(Tampilan Data Hubungan Variabel Numerik)

Index	Serial dari Pk	Yakut
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	6	6
7	7	7
8	8	8
9	9	9
10	10	10
11	11	11
12	12	12
13	13	13
14	14	14
15	15	15

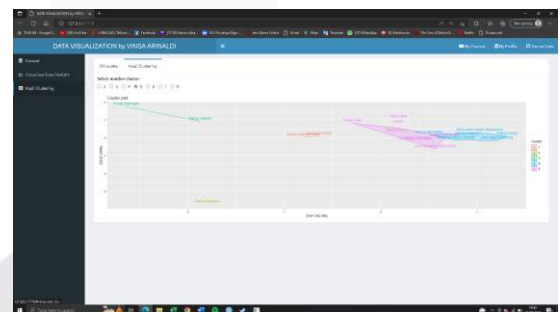
GAMBAR IV.15
(Tampilan Input Data)



GAMBAR IV.18
(Tampilan Tab Panel)



GAMBAR IV.16
(Tampilan Mockup Data Visualization)



GAMBAR IV.19
(Tampilan Hasil Clustering)

4. Validasi dan Evaluasi Hasil Rancangan

a. Validasi Hasil Rancangan

Validasi merupakan suatu pengujian yang berfungsi untuk memberikan jaminan akhir bahwa suatu sistem atau aplikasi yang telah dirancang dapat memenuhi kriteria yang diperlukan atau tidak (Nurudin, Jayanti, Saputro, Saputra, & Yulianti, 2020). Hasil rancangan dalam penelitian tugas akhir ini meliputi penentuan paket wisata alam dan tampilan sistem dengan menggunakan R-Shiny.

b. Evaluasi Hasil Rancangan

Ketika melakukan sebuah perancangan sistem pada suatu objek tentu sistem tersebut akan memiliki sebuah kelebihan dan kekurangan. Berikut merupakan kelebihan dan kekurangan pada sistem yang telah dirancang.

a) Kelebihan Sistem

- 1) Informasi yang tercantum mudah dipahami oleh pembaca.
- 2) Dapat mengubah data sesuai dengan pengguna
- 3) Dapat mengubah tampilan pada 1 tab di R studio

b) Kelemahan Sistem

- 1) Platform belum dapat di akses secara online
- 2) Kurangnya penggunaan user pada sistem
- 3) Belum terintegrasi antara tab panel 1 dan yang lainnya.

- 4) Kurang menarik tampilan untuk visualisasi data.

c. Analisis dan Rencana Implementasi Hasil Rancangan

K-Means clustering merupakan metode yang digunakan untuk melakukan perancangan pada penelitian tugas akhir ini. Metode K-Means merupakan metode yang mengelompokkan data berdasarkan jarak terdekat dengan pusatnya. Hasil cluster menggunakan aplikasi R-Studio menghasilkan 5 cluster yang terbentuk beserta anggotanya. Tabel merupakan hasil dari program R dan anggotanya.

TABEL IV.9
(Hasil Pengklasteran Metode *K-Means*)

5 Cluster
Clustering vector:
[1] 4 3 4 4 1 4 4 5 4 4 4 1 4 2 4 4 3 2

Pada Tabel IV.15 dapat diketahui hasil pengelompokan menggunakan algoritma *K-Means* dengan menggunakan yaitu pada *Cluster* 1, 2, dan 3 masing-masingnya terdiri dari 2 tempat wisata alam, *Cluster* 4 terdiri dari 11 tempat wisata alam, dan *Cluster* 5 terdiri dari 1 tempat wisata alam.

TABEL IV.10
(Hasil *Cluster* Menggunakan *K-Means*)

Cluster	Jumlah Tempat Wisata	Nama Tempat Wisata
1	2	Pasir Putih Wates dan Dermaga Cinta Wates
2	2	Pantai Cemara Kasih dan Jembatan Merah
3	2	Pantai Jatisari dan Pantai Balongan
4	11	Watu Layar Lasem, Pantai Karang Jahe, Pantai Caruban, Dampo Awang, Dermaga Rembang, Pantai Kartini, Pulo Siwalan / Pulau Gede, Pasir Putih Pantai Nyamplung, Mangrove Desa Tireman, Pantai Sluke, dan Pantai Dasun
5	1	Pantai Pandean

Tahap selanjutnya adalah pengimplementasian hasil perancangan yang diharapkan akan dapat membantu wisatawan dan Dinas Pariwisata dalam mengembangkan sistem informasi dalam sektor wisata alam. Terdapat tiga faktor utama dalam proses pengimplementasian sebagai berikut.

1) Faktor Peralatan

2) Faktor Sumber Daya Manusia

3) Faktor Sistem Informasi

d. Analisis dan Rencana Implementasi Hasil Rancangan

TABEL IV.11
(Nilai Skala Penilaian Kuisisioner UAT)

Skala	Keterangan
1	Sangat Tidak Setuju (STS)
2	Tidak Setuju (TS)
3	Kurang Setuju (KS)
4	Setuju (S)
5	Sangat Setuju (SS)

User yang akan melakukan testing Pihak *Travel Agent*. Tabel IV.12 merupakan hasil kuisisioner salah satu responden yang mewakilkan *Travel Agent*.

TABEL IV.12
(Hasil Pengolahan UAT)

Karakteristik	Pertanyaan	Nilai				
		1	2	3	4	5
Functional suitability	Apakah semua fitur yang ada pada sistem dapat berjalan dengan baik?					V
	Apakah sistem dapat memberikan informasi yang dibutuhkan?					V
Performance efficiency	Apakah sistem dapat memberikan respon yang sesuai dengan aksi yang diberikan pengguna?				V	
	Apakah sistem dapat memberikan respon yang cepat sesuai dengan aksi yang diberikan pengguna?				V	
Operability	Apakah sistem ini mudah dioperasikan?				V	
	Apakah sistem ini mudah untuk dipahami?					V
	Apakah tampilan menu dan navigasi di sistem sudah jelas?				V	
	Apakah sistem sudah sesuai dengan yang dibutuhkan?				V	
Reliability	Apakah sistem dapat memudahkan pengguna dalam menentukan paket wisata alam Kabupaten Rembang ?				V	
	Apakah sistem dapat melakukan melakukan fungsi Tambah, Update dan hapus data wisata untuk pemetaan potensi wisata di Kabupaten Rembang?				V	

Setelah melakukan wawancara dilakukannya pengolahan data berdasarkan hasil kuisioner *User Acceptance Test* (UAT) untuk mengetahui nilai persentase dari masing-masing karakteristik.

TABEL IV.13
(Hasil Pengolahan UAT)

Karakteristik	No. Pertanyaan	Frekuensi Jawaban					Skor	Total Skor	Persentase (%)
		1	2	3	4	5			
Functional suitability	1					1	15	30	100%
	2					1	15		
Performance efficiency	1				1		13	26	86,60%
	2				1		13		
Operability	1				1		13	54	90 %
	2					1	15		
	3				1		13		
	4				1		13		
Reliability	1				1		13	26	86,60%
	2				1		13		

Berdasarkan Tabel IV.13 dapat disimpulkan penilaian sistem dengan karakteristik *functional suitability*, dan *operability* berada di rentan 90%-100% masuk kedalam kualifikasi “Sangat Baik” sedangkan karakteristik *performance efficiency* dan *reliability* berada di nilai 86% dan masuk ke dalam kualifikasi “Baik” seperti yang dapat dilihat pada Tabel IV.14 mengenai interpretasi skor. Dalam hal ini keseluruhan sistem dapat disimpulkan diterima oleh pengguna berdasarkan hasil UAT yang sudah dilakukan.

TABEL IV.14
(Tabel Interpretasi Skor)

Kriteria Interpretasi Skor	
Persentase Pencapaian	Keterangan

0%-20%	Sangat Buruk
21%-40%	Buruk
41%-60%	Cukup
61%-80%	Baik
81%-100%	Sangat Baik

V. KESIMPULAN

- Proses perancangan penentuan paket wisata alam di Kabupaten Rembang menggunakan metode algoritma *K-Means Clustering* dengan menggunakan visualisasi *Interface R-Shiny* di R-Studio dengan bahasa pemrograman R.
- Penelitian dalam penentuan paket wisata mendapatkan hasil optimal dari 18 lokasi wisata alam dengan terdapat

4 variabel seperti jam operasional, jarak dan arah tempat wisata dari pusat kota, dan *ratings* tiap lokasi tempat wisata alam tersebut yaitu sebanyak 5 *cluster* yang diantaranya yaitu pada *cluster* 1, 2, dan 3 terdapat 2 lokasi wisata, lalu pada *cluster* 4 terdapat 11 lokasi wisata yang dibagi lagi menjadi 4 kelompok, dan terakhir yaitu *cluster* 5 terdapat 1 lokasi wisata alam.

A. Saran

1. Dinas Pariwisata Kabupaten Rembang, Jawa Tengah
 - a. Perlu menambahkan *platform* sistem informasi mengenai tempat wisata yang ada di Kabupaten Rembang.
 - b. Lebih memperhatikan untuk fasilitas yang terdapat di setiap tempat wisata agar lebih terjaga kebersihan, kenyamanan, dan keselamatan bagi para wisatawan.
 - c. Mengembangkan *platform* sistem informasi agar lebih mudah dan maju lagi untuk wisatawan.
2. Peneliti Selanjutnya
 - a. Penelitian selanjutnya dapat mencoba menggunakan metode *clustering* lainnya selain Algoritma *K-Means Clustering* dan metode *silhouette*, sehingga dapat membandingkan hasil data yang diperoleh ketika menggunakan metode *clustering* yang lain.
 - b. Memaksimalkan pengelompokan dengan tambahan data yang berpengaruh langsung pada penentuan paket wisata alam di Kabupaten Rembang, seperti penambahan variabel ataupun wisata lainnya.
 - c. Mengintegrasikan *interface website* menggunakan aplikasi R-Shiny dan menjalankannya secara *online*.

INTERAKTIF MENGGUNAKAN RSHINY UNTUK ANALISIS STATISTIK NONPARAMETRIK.

- [8] Vikasari, C. (2018). Pengujian Sistem Informasi Magang Industri dengan Metode Blackbox Testing Boundary Value Analysis.
- [9] Mustaqbal, M., Firdaus, R. F., & Rahmadi, H. (2015). PENGUJIAN APLIKASI MENGGUNAKAN BLACK BOX TESTING BOUNDARY VALUE ANALYSIS (Studi Kasus : Aplikasi Prediksi Kelulusan SNMPTN)

REFERENSI

- [1] Primanda, R. P., Alwi, A., & Mustikasari, D. (2021). DATA MINING SELEKSI SISWA BERPRESTASI UNTUK MENENTUKAN KELAS UNGGULAN MENGGUNAKAN METODE K-MEANS CLUSTERING (Studi Kasus di MTS Darul Fikri)
- [2] Handoyo, R., M, R., & Nasution, S. M. (2014). PERBANDINGAN METODE CLUSTERING MENGGUNAKAN METODE SINGLE LINKAGE DAN K - MEANS PADA PENGELOMPOKAN DOKUMEN.
- [3] Rajeshinigo, D., & Jebamalar, J. A. (2015). Accuracy Improvement of C4.5 using K means Clustering.
- [4] O'Brien, J. A., & Marakas, G. M. (2009). *Management Information Systems*.
- [5] Connolly, T. M., & Begg, C. E. (2004). *ataBase Systems: A Practical Approach to Design, Implementation and Management* (4th ed.).
- [6] Lumanauw, N. (2020). Perencanaan Paket Wisata Pada Biro Perjalanan Wisata Inbound (Studi Kasus Di PT. Golden Kris Tours, Bali).
- [7] Maisarah, Kusnandar, D., & Perdana, H. (2021). PENGEMBANGAN APLIKASI WEB