

**Clustering Data Numeric dengan Algoritma Improved K- Means
Pada Inisialisasi Centroid Awal untuk Penentuan Rekomendasi
Jodoh**

SKRIPSI

Untuk memenuhi sebagian persyaratan mencapai gelar Sarjana Komputer



Disusun Oleh :

AR RAMAREGA DHERIVA YUSUF ADHITAMA

NIM. 115060807111089

**KEMENTERIAN RISET TEKNOLOGI DAN PENDIDIKAN TINGGI
UNIVERSITAS BRAWIJAYA
PROGRAM TEKNOLOGI INFORMASI DAN ILMU KOMPUTER
PROGRAM STUDI INFORMATIKA/ILMU KOMPUTER
UNIVERSITAS BRAWIJAYA**

MALANG

2015

LEMBAR PERSETUJUAN

**Clustering Data Numeric dengan Algoritma Improved K- Means
Pada Inisialisasi Centroid Awal untuk Penentuan Rekomendasi
Jodoh**

SKRIPSI

LABORATORIUM KOMPUTASI CERDAS DAN VISUALISASI

Untuk memenuhi sebagian persyaratan memperoleh gelar Sarjana Komputer



Disusun Oleh :

AR RAMAREGA DHERIVA YUSUF ADHITAMA

NIM. 115060807111089

Skripsi ini telah diperiksa dan disetujui oleh:

Dosen Pembimbing I

Dosen Pembimbing II

M. Tanzil Furqon, S.Kom, M.CompSc.
NIP. 198209302008011004

Dian Eka Ratnawati, S.Si, M.Kom.
NIP. 19730619 200212 2 001



LEMBAR PENGESAHAN

Clustering Data Numeric dengan Algoritma Improved K- Means Pada Inisialisasi Centroid Awal untuk Penentuan Rekomendasi Jodoh

SKRIPSI

LABORATORIUM KOMPUTASI CERDAS DAN VISUALISASI

Diajukan untuk memenuhi persyaratan memperoleh gelar Sarjana Komputer

Disusun Oleh :

AR RAMAREGA DHERIVA YUSUF ADHITAMA

NIM. 115060807111089

Skripsi ini telah diuji dan dinyatakan lulus pada tanggal 30 Juni 2015

Skripsi ini telah diperiksa dan disetujui oleh:

Penguji I

Penguji II

Candra Dewi, S.Kom, M.Sc
NIP. 19771114 200312 2 001

Wavan Firdaus Mahmudy, S.Si, M.T, Ph.D
NIP. 19720919 199702 1 001

Penguji III

Barlian Henryranu Prasetyo, S.T, M.T
NIK. 82102406110254

Mengetahui

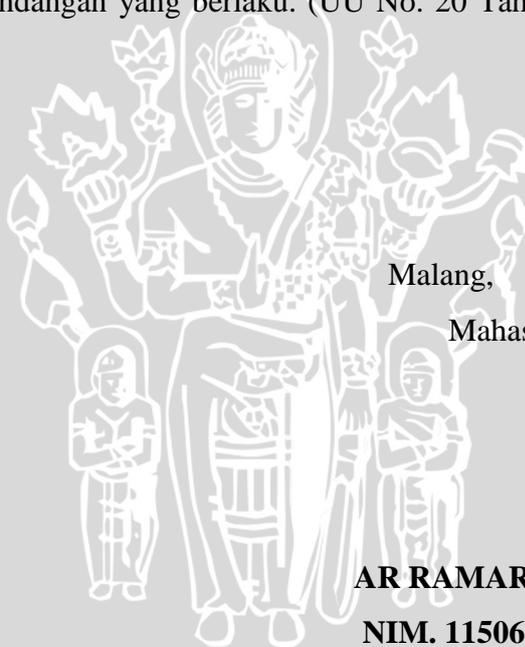
Ketua Program Studi Informatika/Ilmu Komputer

Drs. Marji, M.T.
NIP. 19670801 199203 1 001

PERNYATAAN ORISINALITAS SKRIPSI

Saya menyatakan dengan sebenar-benarnya bahwa sepanjang pengetahuan saya, di dalam naskah SKRIPSI ini tidak terdapat karya ilmiah yang pernah diajukan oleh orang lain untuk memperoleh gelar akademik di suatu perguruan tinggi dan tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis dikutip dalam naskah ini dan disebutkan dalam sumber kutipan dan daftar pustaka.

Apabila ternyata didalam naskah SKRIPSI ini dapat dibuktikan terdapat unsur-unsur PLAGIASI, saya bersedia SKRIPSI ini digugurkan dan gelar akademik yang telah saya peroleh (SARJANA) dibatalkan, serta diproses sesuai dengan peraturan perundang-undangan yang berlaku. (UU No. 20 Tahun 2003, Pasal 25 ayat 2 dan Pasal 70).



Malang, Juni 2015

Mahasiswa,

AR RAMAREGA D Y A

NIM. 115060807111089

KATA PENGANTAR

Puji syukur kehadiran Allah SWT yang selalu memberikan rahmat, hidayah dan cahaya-Nya sehingga penulis dapat menyelesaikan Skripsi dengan judul **“Clustering Data Numeric dengan Algoritma Improved K- Means Pada Inisialisasi Centroid Awal untuk Penentuan Rekomendasi Jodoh “**.

Skripsi ini diajukan untuk memenuhi sebagian persyaratan memperoleh gelar Sarjana Komputer di Fakultas Ilmu Komputer Universitas Brawijaya. Penyelesaian Skripsi ini tidak terlepas dari bimbingan dan bantuan dari berbagai pihak, untuk itu pada kesempatan ini penulis menyampaikan rasa terima kasih dan penghargaan sebesar- besarnya kepada:

1. Allah SWT yang telah memberikan kesabaran, kelancaran dan kenikmatan yang tak ternilai harganya,
2. M. Tanzil Furqon, S.Kom, M.CompSc. serta Dian Eka Ratnawati, S.Si, M.Kom. selaku dosen pembimbing atas segala bimbingan, waktu yang diluangkan serta kritik dan saran yang telah diberikan kepada penulis.
3. Ir. Sutrisno, M.T, Ir. Heru Nurwasito, M.Kom., Himawat Aryadita, S.T, M.Sc., dan Edy Santoso, S.Si., M.Kom., selaku Ketua, Wakil Ketua 1, Wakil Ketua 2 dan Wakil Ketua 3 Fakultas Ilmu Komputer, Universitas Brawijaya Malang.
4. Drs. Marji, M.T., dan Issa Arwani, S.Kom, M.Sc., selaku Ketua dan Sekretaris Program Studi Teknik Informatika, Universitas Brawijaya Malang.
5. Seluruh dosen dan Civitas Akademik Program Studi Teknik Informatika Universitas Brawijaya, atas dukungan dan kerjasamanya;
6. Kedua orangtua penulis beserta anggota keluarga lainnya yang selalu memberikan doa dan dorongan semangat hingga skripsi ini dapat terselesaikan.
7. Semua teman-teman informatika khususnya angkatan 2011 dan teman seperjuangan Program Studi Teknik Informatika atas bantuan dan semangatnya
8. Serta semua pihak yang tidak dapat disebutkan satu-persatu.

Penulis menyadari bahwa skripsi ini masih jauh dari sempurna serta banyak kekurangan disebabkan oleh keterbatasan kemampuan dan pengalaman, dengan kerendahan hati penulis mengharapkan kritik dan saran yang bersifat membangun dari pembaca.

Malang , Juli 2015

Penulis



ABSTRAK

Ar Ramarega Dheriva Yusuf Adhitama, 2015 : Clustering Data Numeric dengan Algoritma Improved K- Means Pada Inisialisasi Centroid Awal untuk Penentuan Rekomendasi Jodoh.

**Dosen Pembimbing : M. Tanzil Furqon, S.Kom, M.CompSc. ;
Dian Eka Ratnawati, S.Si, M.Kom.**

Implementasi rekomendasi ini bertujuan untuk memberikan rekomendasi jodoh kepada user terhadap data user lain yang telah mengisi sebelumnya. Sulitnya seseorang untuk mempertahankan hubungannya karena disebabkan faktor - faktor seperti ketidakcocokan dan sifat pasangannya. Sibuknya seseorang membuat peluang orang tersebut untuk bertemu dengan banyak orang menjadi lebih sedikit sehingga menjadi salah satu faktor yang menyebabkan ketidaktahuan individu tersebut terhadap lingkungannya. Metode Improve K-Means pada inisialisasi centroid awal adalah metode pengembangan pada K-Means dimana pengembangannya terletak pada inisialisasi centroidnya. Pada K-Means inisialisasi centroid ditentukan dengan random, sedangkan pada Improve K-Means ini penentuan centroid dilakukan dengan perhitungan jarak antar data. Penentuan ini memberi rekomendasi dengan menggunakan tingkat kemiripan data yang dicari dengan cluster dan data yang ada pada cluster tersebut. Pengujian dilakukan dengan menggunakan silhouette coefficient dengan range nilai antara -1 sampai 1 dimana semakin tinggi nilai dari silhouette coefficient maka semakin bagus cluster bentukan dari metode ini. Pada implementasi ini dibandingkan pula hasil pengujian pada Improv K-Means terhadap hasil dari pengujian pada K-Means sehingga dapat diketahui keefektifan dari kedua metode. Pada penentuan rekomendasi jodoh ini penggunaan 5 cluster memiliki nilai silhouette coefficient paling tinggi yaitu 0.959 sehingga penggunaan 5 cluster menjadi pilihan utama pada penentuan rekomendasi ini.

Kata kunci : *Clustering, K-means, Improve K-Means, Silhouette coefficient.*

ABSTRACT

Ar Ramarega Dheriva Yusuf Adhitama, 2015 : *numeric clustering using K-Means improved algorithm at the first initialitation of centroid for couple recomendation*

**Advisors : M. Tanzil Furqon, S.Kom, M.CompSc. ;
Dian Eka Ratnawati, S.Si, M.Kom.**

The aim of this research is to give match recommendation to user towards other users based on data which has been registered before. There are several factors that makes someone is difficult to maintain his relationship, such as mismatch and his partner's character. Too many activities make someone have limited time and it influences someone to know his partner's character well. Improving K- Means methods on early centroid initialization is a development method on K-Means where the development located on centroid initialization. On K-Means, centroid initialization was decided randomly while in Improving K-Means, the researcher decided centroid by calculating the distance of data. In giving recommendation, this implementation using rate of the similarity of the data with cluster and the data that involved in the cluster. The data examined by using silhouette coefficient with range of value between -1 until 1, higher value of silhouette coefficient made cluster in this method better. In this implementation, the researcher also compares the result of calibration on Improving K-Means towards the result of calibration on K-Means, therefore the researcher know the effectiveness of both methods. In this match engine recommender, the use of 5 clusters has highest value silhouette coefficient on 0.959. Thus, the use of 5 clusters becomes the primary option on this recommender.

Keywords: *Clustering, K-means, Improve K-Means, Silhouette coefficient.*

DAFTAR ISI

LEMBAR PERSETUJUAN.....	ii
LEMBAR PENGESAHAN	iii
PERNYATAAN ORISINALITAS SKRIPSI	iv
KATA PENGANTAR	v
ABSTRAK	vii
ABSTRACT.....	viii
DAFTAR ISI.....	ix
DAFTAR TABEL.....	xii
DAFTAR GAMBAR	xiii
DAFTAR SOURCECODE.....	xv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Tujuan Penelitian	3
1.4 Manfaat Penelitian	3
1.5 Batasan Masalah.....	3
1.6 Sistematika Penulisan	4
BAB II DASAR TEORI	5
2.1 Kajian Pustaka.....	5
2.2 Pemilihan Pasangan Hidup	6
2.3 Tipe Data.....	11
2.4 Normalisasi	13
2.5 Clustering	14
2.6 K-Means.....	15

2.7 Improve K-Means Inisialisasi Centroid Awal..... 18

2.8 *Silhouette Coefficient* 18

BAB III METODOLOGI PENELITIAN DAN PERANCANGAN..... 20

3.1 Metodologi Penelitian 20

3.1.1 Studi Literatur 21

3.1.2 Pengumpulan data 21

3.1.3 Analisis dan Perancangan Sistem..... 21

3.1.4 Implementasi Sistem 22

3.1.5 Pengujian dan Analisis Sistem 22

3.1.6 Evaluasi dan Analisis hasil..... 23

3.2 Perancangan Sistem 23

3.2.1 Deskripsi Sistem 23

3.2.2 Perancangan Perangkat Lunak 24

3.2.3 Perhitungan Manualisasi 29

3.2.4 Perancangan *User Interface* 51

3.2.5 Perancangan Pengujian dan Evaluasi 58

BAB IV IMPLEMENTASI 60

4.1 Spesifikasi Sistem 60

4.1.1 Spesifikasi Perangkat Keras..... 60

4.1.2 Spesifikasi Perangkat Lunak 61

4.2 Implementasi Algoritma..... 61

4.2.1 Normalisasi 61

4.2.2 Improve K-Means 67

4.2.3 Menghitung Ulang Centroid 70

4.3 Implementasi *Interface* 73

4.3.1 Interface *Input Data*..... 73



4.3.2	Interface Pencarian Data.....	74
4.3.3	Improve Interface.....	75
4.3.4	Interface About.....	76
BAB V PENGUJIAN DAN ANALISIS.....		77
5.1	Pengujian <i>Cluster</i>	77
5.2	Pengujian Jumlah Data.....	81
5.3	Pengujian <i>Cluster</i> K-Means.....	83
5.4	Perbandingan pengujian K-Means dan Improve K-Means.....	87
BAB VI PENUTUP.....		91
6.1	Kesimpulan.....	91
6.2	Saran.....	92
Daftar Pustaka.....		93
DAFTAR LAMPIRAN.....		96



DAFTAR TABEL

Table 3.1 Data sampel manualisasi	29
Table 3.2 Hasil proses pelabelan	31
Table 3.3 Hasil proses Normalisasi	33
Table 3.4 Perhitungan jarak Euclidean Distance	35
Table 3.5 Total jarak tiap data (v_j)	37
Table 3.6 Perhitungan jarak v_j terhadap objek j	37
Table 3.7 Jarak v_j terhadap objek j	39
Table 3.8 Hasil pengurutan objek k	40
Table 3.9 Perhitungan jarak terhadap pusat cluster	42
Table 3.10 Keanggotaan Cluster 1	44
Table 3.11 Keanggotaan Cluster 2	44
Table 3.12 Keanggotaan Cluster 3	45
Table 3.13 Perhitungan kalkulasi terhadap total jarak cluster iterasi 1	46
Table 3.14 Keanggotaan Cluster 1	48
Table 3.15 Keanggotaan Cluster 2	49
Table 3.5 Keanggotaan Cluster 3	49
Table 3.17 Perbandingan klasterisasi awal dan klasterisasi ahir	50
Table 4.1 Spesifikasi Perangkat Keras Komputer	60
Table 4.2 Spesifikasi Perangkat Lunak	61
Table 5.1 Pengujian Jumlah Cluster	77
Table 5.2 Pengujian Jumlah Data	81
Table 5.3 Pengujian Jumlah cluster K-Means	83

DAFTAR GAMBAR

Gambar 2.1 Clustering [PHC-09].....	15
Gambar 2.2 K-Means Clustering [WBS-05]	17
Gambar 3.1 Diagram Alir	20
Gambar 3.2 Diagram Alir Tahap Perancangan Sistem.....	23
Gambar 3.3 Gambaran Sistem.....	24
Gambar 3.4 Diagram Utama Sistem.....	25
Gambar 3.5 Diagram Alir Sistem.....	26
Gambar 3.6 Diagram Alir Proses Normalisasi	27
Gambar 3.7 Diagram Alir Proses Inisialisasi <i>Cluster</i> Awal.....	28
Gambar 3.8 Interface Halaman Utama.....	51
Gambar 3.9 Interface Input Data	52
Gambar 3.10 Interface pencarian data.....	53
Gambar 3.11 <i>Interface</i> rekomendasi.	54
Gambar 3.12 <i>Result Interface</i>	55
Gambar 3.13 <i>Hasil Perhitungan</i>	56
Gambar 3.14 <i>Interface About</i>	57
Gambar 4.1 Antarmuka Halaman Utama.....	73
Gambar 4.2 Interface Input Data.....	74
Gambar 4.3 Interface Pencarian Data.....	74
Gambar 4.4 <i>Interface</i> Hasil Pencarian	75
Gambar 4.5 Interface Improve.....	75
Gambar 4.6 Hasil Improve.....	76
Gambar 4.7 Interface About.....	76
Gambar 5.1 Statistik Jumlah Iterasi Cluster 3 sampai 10.....	78
Gambar 5.2 Statistik Nilai Silhouette Cluster 3 sampai 10.....	79
Gambar 5.3 Statistik Jumlah Iterasi Cluster 10 sampai 100.....	79
Gambar 5.4 Statistik Nilai Silhouette Cluster 10 sampai 100.....	80
Gambar 5.5 Statistik Jumlah Iterasi Data 10 sampai 150.....	82
Gambar 5.6 Statistik Nilai Silhouette Data 10 sampai 150.....	82
Gambar 5.7 Statistik jumlah iterasi cluster 3 sampai 10 K-Means	85

Gambar 5.8 Statistik nilai silhouette cluster 3 sampai 10 K-Means..... 86

Gambar 5.9 Statistik jumlah iterasi cluster 10 sampai 100 K-Means 86

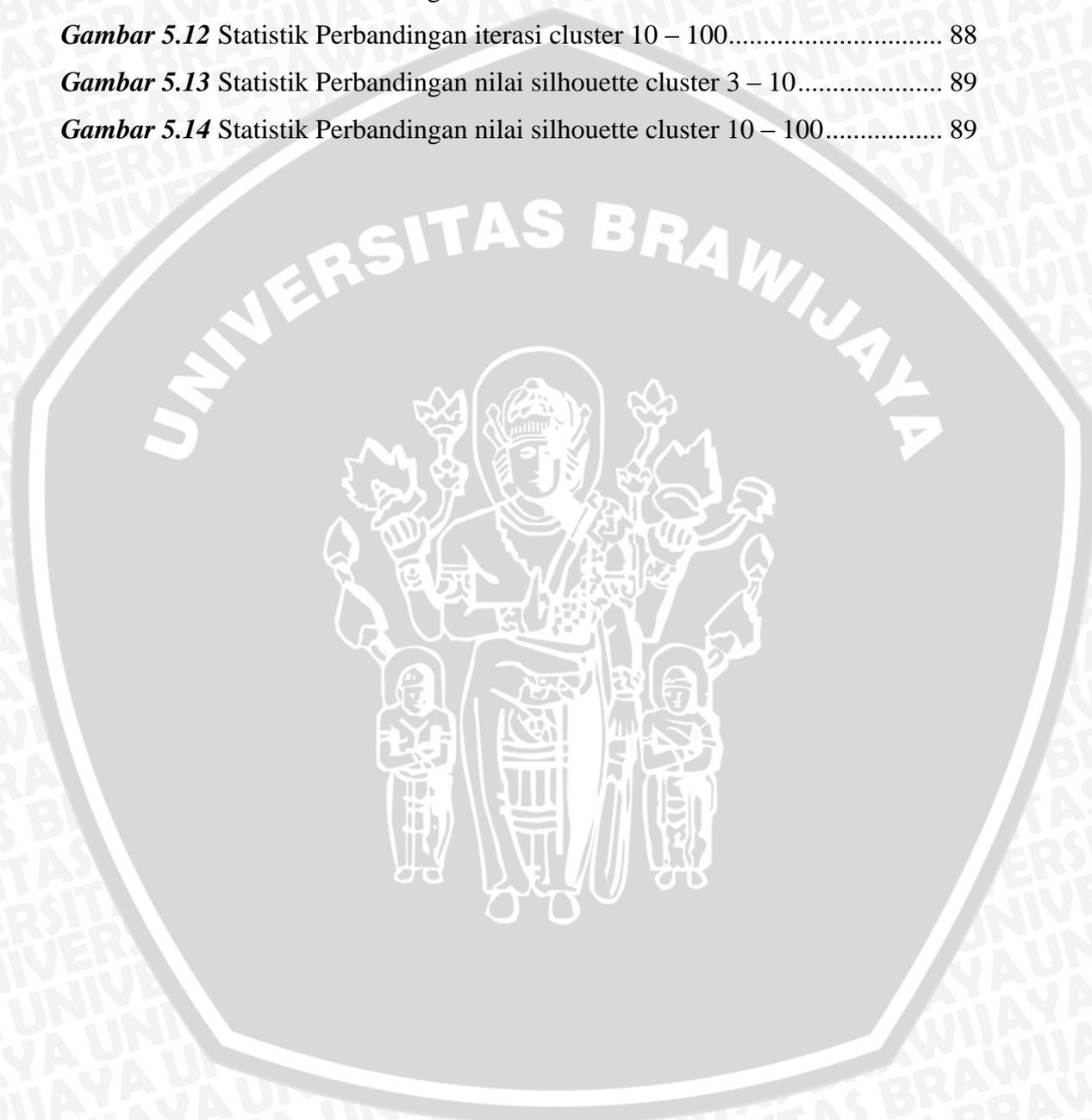
Gambar 5.10 Statistik nilai silhouette cluster 10 sampai 100 K-Means..... 87

Gambar 5.11 Statistik Perbandingan iterasi cluster 3 – 10..... 88

Gambar 5.12 Statistik Perbandingan iterasi cluster 10 – 100..... 88

Gambar 5.13 Statistik Perbandingan nilai silhouette cluster 3 – 10..... 89

Gambar 5.14 Statistik Perbandingan nilai silhouette cluster 10 – 100..... 89



DAFTAR SOURCECODE

Sourcecode 4.1 Labeling 65

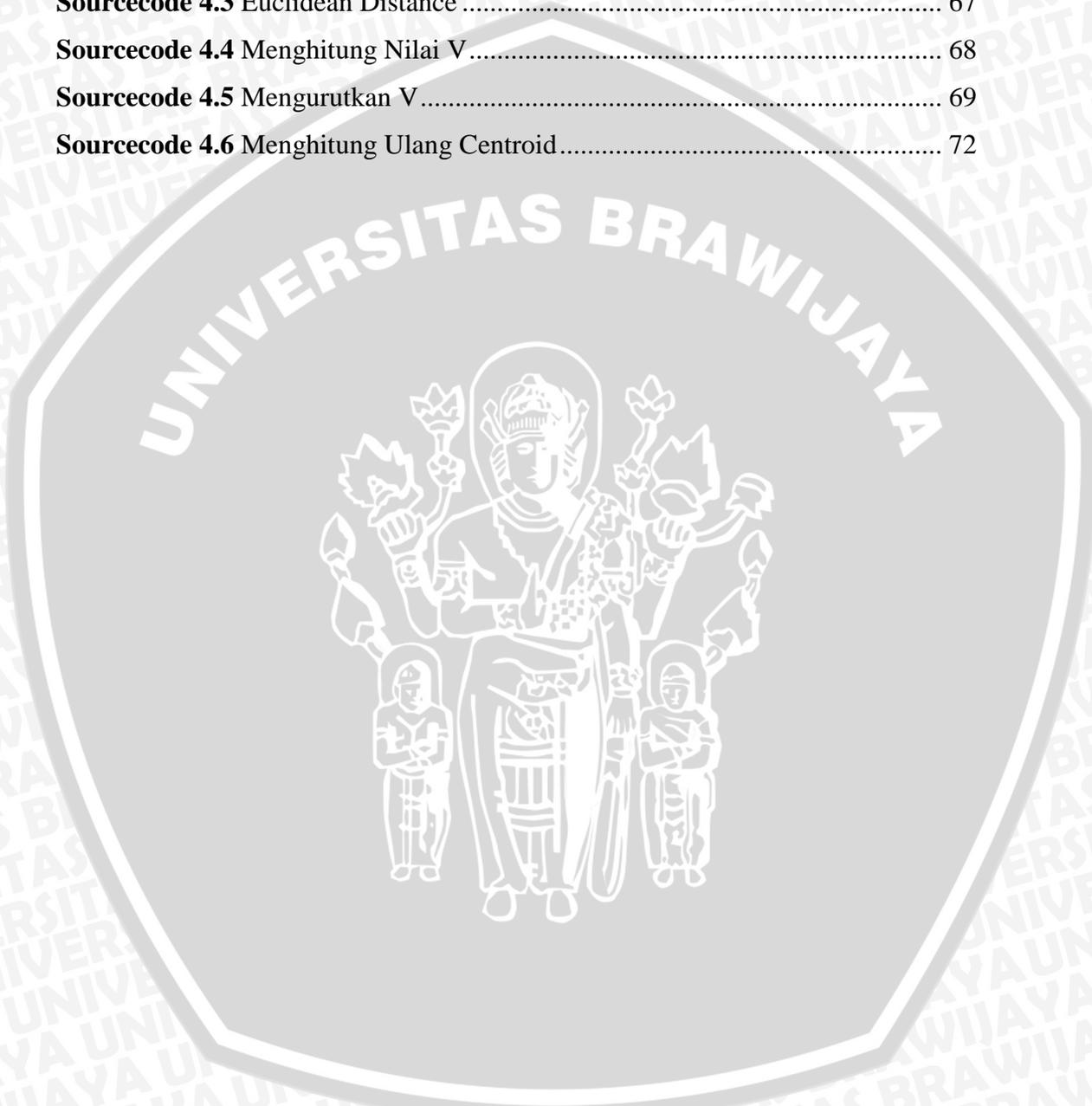
Sourcecode 4.2 Min Max 66

Sourcecode 4.3 Euclidean Distance 67

Sourcecode 4.4 Menghitung Nilai V 68

Sourcecode 4.5 Mengurutkan V 69

Sourcecode 4.6 Menghitung Ulang Centroid 72



BAB I

PENDAHULUAN

Pada pelaksanaan sebuah penelitian maka perlu dijelaskan hal hal penting yang akan menjadi dasar dan acuan dari pelaksanaan penelitian tersebut. Pada bab ini akan dijelaskan mengenai latar belakang permasalahan yang di angkat, perumusan masalah, tujuan penelitian, batasan masalah, dan manfaat yang dapat di petik dari penelitian yang dilakukan.

1.1 Latar Belakang

Semakin majunya perkembangan zaman, maka tak dipungkiri pula bahwa media media seperti media sosial menjadi salah satu ajang pencarian jodoh. Namun tidak semua orang yang dikenal dalam media tersebut tepat dan pantas untuk dijadikan pasangan hidup nantinya [HAS-14]. Ketidacocokan karakter dan sifat pada pasangan, membuat pasangan tersebut menjadi kesulitan untuk mempertahankan hubungannya.

Secara umum faktor yang menyebabkan seseorang sulit untuk menemukan calon pasangan hidup adalah karena minimnya kesempatan dari seseorang tersebut untuk bertemu dengan banyak orang, sehingga hal tersebut menjadi faktor utama yang dapat menghambat seseorang untuk bertemu dengan calon pasangan yang dia idamkan. Sulitnya seseorang yang sukses dalam karirnya dalam mencari pasangan hidup seringkali dihubungkan dengan kriteria. Semakin tinggi kedudukan seseorang maka secara umum pola pikirnya akan mencari pasangan yang tinggi pula. Namun hal tersebut tidak sepenuhnya benar, karena semakin tinggi kedudukan seseorang maka semakin sulit juga kesempatan seseorang tersebut untuk bertemu dengan orang lain [CNS-14].

Adapun sarana yang ada saat ini untuk membantu seseorang agar mendapatkan pasangan hidup adalah melalui biro jodoh. Tujuan adanya biro jodoh yaitu membantu seseorang dengan kesibukan padat dan seseorang dengan kepercayaan diri rendah agar dapat bertemu dengan calon pasangan hidupnya. Adapun keunggulan yang diberikan oleh biro jodoh kepada calon pencari pasangan yaitu memberikan rasa nyaman, menghemat biaya, praktis dan lebih mudah untuk

menyeleksinya. Mengacu pada masalah tersebut, penulis terinspirasi untuk membuat alat rekomendasi untuk pemilihan jodoh dengan tujuan agar dapat membantu banyak orang menemukan calon pasangan hidupnya terutama bagi mereka yang memiliki kesibukan tinggi dan memiliki kepercayaan diri rendah.

Dengan mengacu pada dua penelitian sebelumnya, yaitu yang pertama berjudul “Clustering Data Non-Numerik Dengan Pendekatan Algoritma *K-Means* Dan Hamming Distance Studi Kasus Biro Jodoh.” [MSN-05]. Pada penelitian ini membahas pengelompokan kriteria calon pasangan hidup dengan menggunakan algoritma *K-Means*. Kemudian pada penelitian kedua berjudul “A simple and fast algorithm for K-medoids clustering” [PHC-09]. Pada penelitian ini membahas tentang iris data *sentosa*, *versicolor*, dan *virginica*. Pada penelitian pertama dapat disimpulkan bahwa algoritma *K-Means* dapat digunakan untuk mengkluster data beratribut non-numerik atau katagorikal. Akan tetapi inisialisasi centroid pada *K-Means* dilakukan secara random sehingga tidak dapat menghasilkan nilai yang maksimal. [WBS-03]. Kemudian pada penelitian kedua didapatkan kesimpulan bahwa penggunaan inisialisasi kluster awal menghasilkan kluster yang lebih rapat.

Dengan acuan beberapa referensi di atas, penulis akan mengimplemetasikan dengan menggunakan metode *Improve K- Means pada inisialisasi centroid awal* dengan teknik *clustering* untuk *data numerik*. *Data numeric* adalah data dengan nilai-nilai dari data tersebut adalah bilangan atau ukuran angka. Pada rekomendasi jodoh akan dilakukan proses klasifikasi dari 10 kriteria yang ada kedalam sebuah kelompok dimana nilai optimalnya atau mendekati optimal akan dikelompokan menjadi satu *cluster*. Jarak antar matrik akan di hitung untuk menentukan nilai pusat cluster dengan metode *inisialisasi centroid awal*.

Dengan penelitian ini diharapkan dapat membantu para pencari jodoh untuk memilih calon pasangan hidup dengan rekomendasi jodoh. Selain itu rekomendasi jodoh ini diharapkan dapat memberikan solusi dan alternatif serta dapat memberikan gambaran mengenai calon pasangan yang ideal.

1.2 Rumusan Masalah

Mengacu pada latar belakang yang telah dikemukakan di atas, maka perlu dirumuskan agar masalah pada penelitian ini lebih terarah dan sesuai dengan tujuan.

Rumusan masalah dalam penelitian ini antara lain:

1. Bagaimana menerapkan metode *Improve K- Means pada inisialisasi centroid awal* untuk Penentuan Rekomendasi Jodoh?
2. Bagaimana kualitas dari hasil kluster pada implementasi rekomendasi Jodoh dengan metode *Improve K- Means pada inisialisasi centroid awal* dibandingkan metode *K-Means*?

1.3 Tujuan Penelitian

Adapun tujuan dari penelitian ini adalah sebagai berikut:

1. Menerapkan metode *Improve K-Means pada inisialisasi centroid awal* pada penentuan rekomendasi Jodoh.
2. Mengetahui kualitas dari hasil kluster pada penentuan rekomendasi Jodoh dengan metode *Improve K- Means pada inisialisasi centroid awal*.

1.4 Manfaat Penelitian

Adapun manfaat yang didapat pada implementasikan dari tugas ahir ini bagi peneliti adalah untuk menambah pengetahuan peneliti tentang penerapan metode *Improve K- Means pada inisialisasi centroid awal*.

Sedangkan manfaat penelitian ini bagi masyarakat adalah untuk dapat memberikan solusi dan alternatif terhadap para pencari pasangan hidup dan dapat memberikan rekomendasi pasangan yang ideal yang sesuai dengan kriteria.

1.5 Batasan Masalah

Agar penelitian ini lebih terfokus dan tidak terlalu luas pada pembahasannya, maka perlu diberikan batasan batasan masalah. Adapun batasan batasan masalah yang akan digunakan yaitu antara lain:

1. Data yang digunakan adalah data quisioner yang disebar menggunakan google form (<http://goo.gl/forms/9QBqtYUjmi>) sebanyak 150 responden.
2. Pada Penelitian ini menggunakan 10 kriteria dari calon pasangan berdasarkan pertimbangan kriteria pada penelitian sebelumnya [MSN-05] dan survei pada jajakpendapat.com

1.6 Sistematika Penulisan

Gambaran secara garis besar pembahasan dari keseluruhan isi laporan penelitian untuk setiap bab adalah sebagai berikut :

BAB I PENDAHULUAN

Pada Bab pendahuluan akan dibahas tentang Latar Belakang, Rumusan Masalah, Batasan Masalah, Tujuan, Manfaat, Serta sistematika penulisan dari penelitian ini.

BAB II DASAR TEORI

Pada Bab Dasar Teori akan dibahas mengenai kajian pustaka serta teori penunjang yang berhubungan dengan penentuan jodoh dan clustering dengan *Improve K- Means pada inisialisasi centroid awal*.

BAB III METODOLOGI PENELITIAN

Pada Bab Metodologi Penelitian akan dibahas tentang metode yang digunakan pada penelitian yang terdiri dari Studi Literatur, Perancangan Perangkat Lunak, Implementasi Perangkat Lunak, Pengujian dan Analisis, Serta Pengambilan Keputusan.

BAB IV IMPLEMENTASI

Pada Bab Implementasi akan dibahas tentang proses pembangunan penentuan rekomendasi jodoh yang sesuai dengan bahasa pemrograman yang telah ditentukan pada Bab Pendahuluan.

BAB V PENGUJIAN DAN ANALISIS

Pada Bab Pengujian dan Analisis akan menjelaskan proses dan hasil dari pengujian terhadap aplikasi yang telah dibangun dan memastikan bahwa program telah sesuai dengan perancangan beserta analisis analisisnya.

BAB VI PENUTUP

Pada Bab Penutup akan dimuat kesimpulan dari keseluruhan uraian bab-bab sebelumnya, serta saran-saran dari hasil yang diperoleh, yang diharapkan dapat bermanfaat dalam pengembangan selanjutnya.

BAB II

DASAR TEORI

Dalam penelitian diperlukan dasar teori dan argumen yang saling terhubung dengan konsep permasalahan penelitian yang akan digunakan dalam analisis. Pada bab ini akan dijelaskan tentang dasar-dasar teori dan argumen yang digunakan pada penelitian ini yaitu antara lain mengenai *Data Mining*, *Clustering*, *K-Means* dan *Improve K-Means pada inisialisasi centroid awal* untuk penentuan rekomendasi jodoh

2.1 Kajian Pustaka

Pada kajian pustaka ini akan membahas tentang penelitian sebelumnya yang terkait dengan penelitian ini. Diantaranya yaitu yang telah dilakukan oleh Darlis Heru Murti, Nanik Suciati, dan Daru Jani Nanjaya (2005) yang berjudul “Clustering Data Non-Numerik Dengan Pendekatan Algoritma K-Means Dan Hamming Distance Studi Kasus Biro Jodoh” membahas tentang pengelompokan data ke dalam beberapa *cluster* dimana data tersebut bersumber dari biro jodoh Grasco, Sakinah Surabaya, Libe, dan O’Diva dimana data telah dikonversi menjadi data non-numerik (Katagorikal). Pada penelitian ini katagori yang digunakan yaitu sebanyak 20 antara lain Jenis Kelamin, Umur, Tinggi, Berat, Bentuk Badan, Rambut, Warna Kulit, Warna Rambut, Suku, Status Pernikahan, Jumlah Anak, Agama, Pendidikan Terakhir, Pekerjaan, Merokok, Berkacamata, Aktif Berolahraga, dan Berjilbab. Katagori – katagori tersebut dijadikan parameter untuk mengelompokkan data berdasarkan tingkat kesamaan untuk dijadikan kedalam beberapa kluster. Hasil dari penelitian ini didapatkan bahwa metode *K-Means* dapat digunakan untuk pengelompokan data non-numerik yang sudah dirubah kedalam bentuk biner dan efektif untuk pencarian berdasarkan jarak [MSN-05].

Akan tetapi metode K-Means memiliki hasil pengelompokan yang tidak maksimal yang disebabkan oleh inisialisasi pusat kluster yang *random* [WBS-03]. Hal tersebut menyebabkan kualitas kluster bentukan berubah ubah atau tidak pasti. Oleh karena itu penulis mengusulkan untuk mengangkat metode pengembangan *K-Means* dimana metode ini disebut sebagai *Improve K-Means inisialisasi centroid*

awal. Algoritma *Improve K-Means* pada inisialisasi *centroid* awal memperbaiki kelemahan *K-Means* pada inisialisasi *centroid* random yang menyebabkan nilai akhir *K-Means* kurang optimal menjadi ditentukan dengan proses perhitungan. Hal tersebut meningkatkan performansi kinerja dari penggunaan metode *K-Means*. [CZS-10]

Kemudian *improve K-Means* pada penelitian ini mengadopsi dari salah satu paper penelitian dari Hae-Sang Park dan Chi Hyuck Jun (2009) yang berjudul “A simple and fast algorithm for K-medoids clustering” dimana inisialisasi *centroid* awal yang digunakan pada *K-Medoids* diterapkan pada metode *K-Means*. pada *K-Means* inisialisasi dilakukan secara *random* [MSN-05] dirubah menjadi ditentukan dengan langkah awal perhitungan yaitu dengan *Euclidean distance* kemudian dilanjutkan kalkulasi nilai v_j dan diakhiri dengan sorting nilai v_j berdasarkan nilai terkecil [PHC-09]. Dengan mengadopsi proses inisialisasi tersebut diharapkan hasil klustering menjadi lebih baik dan optimal.

2.2 Pemilihan Pasangan Hidup

1. Pengertian pemilihan pasangan hidup

Memilih pasangan hidup adalah proses memilih atau seleksi seseorang yang nantinya dapat dipercaya menjadi teman hidup dan menjadi orang tua dari anak-anaknya kelak. Pemilihan terhadap pasangan hidup yang dilakukan oleh seseorang biasanya didasari untuk sikap saling melengkapi antar kebutuhan kedua individu dan juga didasari oleh suatu pemikiran bahwa individu akan memilih pasangan yang nantinya dapat melengkapi kebutuhan yang diperlukannya [DEG-08].

2. Faktor-faktor yang Mempengaruhi Pemilihan Pasangan Hidup

Adapun faktor yang mempengaruhi suatu individu untuk memilih pasangannya menurut Degenova antara lain sebagai berikut: [DEG-08]

a) Latar Belakang Keluarga

Latar belakang keluarga adalah salah satu faktor yang sangat berpengaruh pada individu baik ketika ingin menjadi pasangan hidup atau saat proses pemilihan pasangan hidup dan setelah menjadi pasangan hidup. Dengan mempelajari latar belakang dari calon pasangan hidup akan membantu seorang individu

mengetahui sifat calon pasangannya. Adapun hal yang perlu diperhatikan dalam mempelajari latar belakang dari calon pasangan hidup antara lain:

- Kelas Sosial Ekonomi

Faktor yang mempengaruhi kepuasan dalam pernikahan adalah ekonomi. Dengan status ekonomi yang baik, maka kepuasan pernikahan dari pasangan tersebut akan tinggi. Berbeda apabila status ekonominya rendah, status kepuasan pernikahan kemungkinan akan kurang baik dibandingkan dengan individu yang memilih pasangannya dari kelas ekonomi tinggi [DEG-08]. Adapun kriteria yang digunakan pada kelas sosial ekonomi adalah diukur dari pekerjaannya.

- Pendidikan dan inteligensi

Secara umum individu akan memilih pasangan yang mempunyai perhatian terhadap dunia pendidikannya. Pernikahan kedua individu dengan latar belakang pendidikan yang sama akan lebih cocok dibandingkan dengan pernikahan kedua individu dengan latar belakang pendidikan yang berbeda [DEG-08]. Adapun kriteria yang digunakan pada kelas pendidikan dan inteligensi adalah diukur dari pendidikan terakhir.

- Agama

Agama adalah salah satu faktor yang dipertimbangkan dalam pemilihan pasangan. Pernikahan dengan latar belakang agama yang sama akan membuat pernikahan menjadi lebih stabil dan dengan agama memungkinkan anak-anak hasil pernikahan akan dapat tumbuh sesuai dengan keyakinan dan moral yang berkembang di lingkungan masyarakat [DEG-08].

- Pernikahan antar Ras atau Suku

Pernikahan berbeda rasa atau suku merupakan salah satu masalah yang sering terjadi di masyarakat. Banyak masalah terjadi apabila seorang individu memiliki pasangan dari suku atau ras yang berbeda. Permasalahan yang ada bukan berasal dari pernikahan pasangan tersebut, melainkan berasal dari keluarga, teman-teman, maupun masyarakat sekitar. Secara umum apabila tidak ada dukungan dari keluarga yang

berhubungan atau suku asal, maka pernikahanpun tidak akan terjadi [DEG-08].

b) Karakteristik Personal

Ketika individu memilih calon pasangan hidup maka kecocokan adalah hal yang perlu diperhatikan. Adapun faktor yang dapat mendukung kecocokan dari pemilihan pasangan hidup menurut Degenova antara lain:

- Sikap dan Tingkah laku

Adapun pemilihan pasangan yang didasari oleh sifat dari tiap individu cenderung berfokus pada fisik, kepribadian, dan faktor kesehatan mental. Perbedaan sifat dan kepribadian seseorang tentunya memungkinkan dapat menghambat kebahagiaan dari pernikahan pasangan tersebut. Sifat muram dalam pernikahan akan menyebabkan pernikahan tersebut menjadi negatif dan dapat menyebabkan penurunan dari kualitas hubungan antar pasangan. Sedangkan sifat ramah akan menimbulkan nilai positif dari pernikahan sehingga kualitas hubungan antar pasangan pun lebih stabil. [DEG-08]. Adapun kriteria yang digunakan yaitu tinggi badan, berat badan, jenis rambut dan keaktifan merokok.

- Usia

Perbedaan usia merupakan salah satu hal yang dipertimbangkan dalam pemilihan pasangan hidup. Secara umum rata-rata perbedaan umur tiap pasangan adalah dua tahun. Pertimbangan untuk mendapatkan hasil pernikahan yang berkualitas tentunya menuntut setiap individu untuk merenung dalam memilih pasangan yang lebih muda atau lebih tua. Sebagai contoh apabila seorang perempuan muda menikah dengan seorang pria yang umurnya jauh lebih tua maka perempuan tersebut harus siap menjadi janda di usia muda. Namun apabila kedua pasangan memiliki umur yang sama maka akan memiliki kecenderungan hidup bersama lebih lama jika menikah sejak usia muda.

- Sikap dan Nilai

Memiliki kesamaan sikap dan nilai pada sesuatu yang penting menjadi salah satu hal yang dapat meningkatkan kecocokan dalam hubungan pernikahan. Individu yang saling terbuka dalam hal sikap dan nilai maka

akan membuat pasangan pernikahannya menjadi lebih nyaman. Kecenderungan seperti stress akan dapat ditekan dengan adanya penyesuaian diri pada pasangan pernikahan.

- Peran Gender dan Kebiasaan

Kecocokan dalam pernikahan juga dipengaruhi oleh perilaku dari tiap individunya. Pasangan akan merasa lebih puas apabila pasangannya dapat berbagi harapan yang sama mengenai peran masing masing gender dan saling memberikan toleransi terhadap kebiasaan dari pasangannya. Adapun salah satu tolak ukur kecocokan dalam pernikahan adalah persamaan harapan dari kedua pasangan. Seorang pria mempunyai peran yang harus ditunjukkan kepada pasangannya sebagai sepasang suami istri. Begitupun dengan wanita yang mempunyai konsep peranan yang harus di tunjukan sebagai seorang istri dan juga harapan dari masing masing individu harus juga ditunjukkan walaupun keduanya memiliki harapan yang berbeda beda [DEG-08].

3. Proses Pemilihan Pasangan Hidup

Pemilihan pasangan hidup adalah proses seleksi dari individu untuk menentukan calon pasangan hidupnya sampai calon tersebut terpilih menjadi pasangan hidup yang sesuai menurut individu tersebut. Adapun variasi proses dalam memilih pasangan hidup menurut Degenova antara lain: [DEG-08].

- a) The Field of Eligibles (Area yang ditentukan)

Salah satu faktor yang harus dipertimbangkan dalam proses pemilihan pasangan adalah apakah calon pasangan sudah memenuhi syarat yang sesuai dengan ketentuan individu tersebut. Pada wanita, kurangnya kualitas pernikahan berasal dari pernikahan itu sendiri dan kualitas pasangannya. Kecenderungan apabila memiliki pasangan dengan status tinggi maka kualitas penikahannya akan lebih baik dibandingkan dengan pasangan dengan status rendah (diukur dari tingkat pendidikan dan pekerjaan) [DEG-08].

- b) *Propinquity* (Kedekatan)

Kedekatan adalah salah satu faktor dalam proses pemilihan pasangan hidup dimana kedekatan secara geografi akan mempengaruhi proses

pemilihan pasangan. Hal tersebut bukan berarti dikatakan dekat secara kediaman antar kedua pasangan saja, melainkan juga kedekatan secara institutional. Hal tersebut dipengaruhi oleh seberapa sering individu tersebut berjumpa dengan pasangannya seperti tempat kerja, sekolah dan lainnya [DEG-08].

c) *Attraction* (Daya Tarik)

Ketertarikan individu terhadap pasangannya dapat diartikan ketertarikan secara fisik maupun kepribadian individu secara spesifik. Pada dasarnya setiap pria dan wanita memiliki perbedaan dalam menentukan pasangannya. Setiap individu memiliki kebutuhan dan perbedaan dalam cara menentukan calon pasangannya. Banyak alasan yang menyebabkan individu dapat jatuh cinta kepada seseorang [DEG-08].

d) *Homogamy* dan *Heterogamy*

Seorang individu akan memilih pasangan hidupnya dimana pasangannya dapat berbagi kepribadian dan karakteristik sosialnya (usia, ras, etnik, pendidikan, kelas sosial dan agama). Kecenderungan memilih pasangan yang memiliki kesamaan disebut *homogamy* sedangkan kecenderungan memilih pasangan yang mempunyai perbedaan disebut dengan *heterogamy*. Pernikahan *homogamy* akan lebih stabil dibandingkan dengan pernikahan *heterogamy*. Kebanyakan individu akan lebih memilih pasangan yang mempunyai kesamaan dengannya dan kurang merasa nyaman bila berada di dekat individu yang memiliki banyak perbedaan dengannya. Faktor penting lainnya seperti tekanan sosial akan mengarah kepada *endogamy* yaitu pernikahan dimana menikah dengan pasangan yang berbeda usia atau kelompok etnik yang berbeda dimana akan menerima celaan halus dari lingkungan sekitarnya. Sedangkan lingkungan yang melarang adanya pernikahan dengan pasangan yang terlalu mirip seperti saudara kandung atau sepupu pertama disebut dengan tekanan social *exogamy* [DEG-08].

e) *Compability* (Kecocokan)

Faktor kecocokan merupakan faktor dimana kemampuan individu tersebut untuk dapat hidup bersama secara harmonis. Kecocokan akan

menilai dan menjadi evaluasi dari pemilihan pasangan berdasarkan tempramen, sikap dan niali, kebutuhan, peran dan kebiasaan. Dalam proses pencarian pasangan hidup, seorang individu akan berjuang untuk menentukan pasangan yang cocok meskipun berasal dari area yang berbeda [DEG-08].

f) *The Filtering Process* (Proses Penyaringan)

Adapun variasi dari proses pencarian pasangan hidup biasanya akan melalui tahapan yang berbeda beda pada setiap individunya. Sebagaimana contohnya yaitu mengeliminasi calon pasangan yang tidak memenuhi syarat. Sebelum membuat keputusan untuk menikah, kedua pasangan akan dibawa menuju periode terahir yaitu pertunangan. Apabila keduanya dapat bertahan pada proses ini maka kedua individu dapat dan siap untuk menikah [DEG-08].

2.3 Tipe Data

Pada bidang statistika, hasil dari sebuah proses wawancara dan pengukuran harus dicatat agar dapat dilakukan pengolahan, analisa, interpretasi, dan pembuatan kesimpulan. Data yang di catat dapat berupa huruf, kata, angka ataupun dapat berupa sebuah kalimat. Adapun data yang dicatat dapat digolongkan menjadi 2 jenis yaitu data yang bukan berupa angka di sebut dengan data *kualitatif* dimana tidak dapat dilakukan operasi matematika sedangkan data yang berupa angka disebut data *kuantitatif*. Adapun jenis-jenis dari data kualitatif menurut Santoso dan Sutanto diantaranya sebagai berikut [SSN-10] [SUT-03]:

a. Data Nominal

Data nominal adalah data yang mempunyai tingkatan paling rendah pada level pengukurannya. Data ini hanya menghasilkan satu katagori pada proses pengukurannya. Contohnya agama, suku, jenis kelamin dan lain sebagainya. Oleh karena itu data nominal tidak akan pernah memiliki dua data sekaligus.

Karena variabel yang dihasilkan mempunyai derajat nilai yang sama, maka untuk proses analisis berdasarkan data nominal harus melalui tahap konversi dari data nominal menjadi data katagorikal yaitu data yang katagorinya dirubah menjadi angka. Contohnya jenis kelamin laki-laki dikatagorikan menjadi “1” dan jenis kelamin perempuan dikatagorikan menjadi “0”. Dalam hal ini angka

yang digunakan bukanlah sebuah tingkatan, melainkan hanya sebuah lambang sehingga semua data memiliki nilai yang sama hanya saja memiliki lambang yang berbeda.

b. Data Ordinal

Data ordinal adalah data yang memiliki tingkatan yang lebih tinggi dari data nominal karena data ordinal memiliki tingkatan pada nilai datanya. Walaupun terdapat perbedaan pada setiap nilai data, namun tidak diketahui perbedaan nilai antar datanya. Contohnya adalah data mengenai pendapat terhadap produk tertentu, tingkat pendidikan, social ekonomi dan lain sebagainya.

Data ordinal tidak dapat dilakukan operasi matematika seperti penambahan, pengurangan, pembagian maupun perkalian. Adapun penerapan yang digunakan dalam pengukuran hasil yaitu melalui wawancara dimana akan diperoleh data berupa sikap atau pendapat terhadap suatu kebijakan. Adapun nilai yang sering digunakan yaitu “sangat setuju”, “setuju”, “tidak setuju”, dan “sangat tidak setuju” dimana skala tentang kesetujuan atau ketidak setujuan disebut sebagai skala *Likert*. Jarak antar nilai tersebut tidaklah jelas perbedaannya, oleh karena itu harus dilakukan konversi dari data ordinal menjadi data interval untuk dapat dinilai kuantitatifnya dimana dapat dimisalkan “sangat setuju” dikategorikan sebagai 4, “setuju” dikategorikan sebagai 3, “tidak setuju” dikategorikan sebagai 2, dan “sangat tidak setuju” dikategorikan sebagai 1.

Data kuantitatif adalah data yang berupa angka yang dapat dilakukan operasi matematika seperti penambahan, pengurangan, perkalian dan pembagian. Adapun jenis dari data kuantitatif menurut Santoso dan Sutanto diantaranya sebagai berikut [SSN-10] [SUT-03]:

a. Data Interval

Data interval adalah data yang memiliki level paling tinggi dimana data interval memiliki tingkatan interval atau jarak yang tetap, akan tetapi data interval tidak memiliki titik nol yang mutlak. Adapun contoh penggunaan data interval yaitu pada pengukuran suhu zat cair dimana suhu antara 40-59 derajat *celcius* dikatakan sebagai suhu “kurang panas”, suhu antara 60-79 derajat

celcius dikatakan sebagai suhu “cukup panas”, suhu antara 80-90 derajat *celcius* dikatakan sebagai suhu “panas”, suhu dan antara 100-119 derajat *celcius* dikatakan sebagai suhu “sangat panas”. Karena data interval tidak memiliki titik nol mutlak bukan berarti bahwa zat cair tersebut tidak memiliki suhu atau panas apabila berada di suhu nol derajat *celcius*.

b. Data Rasio

Data rasio adalah data yang memiliki tingkatan paling tinggi diantara jenis data atau skala pengukuran lain. Data rasio adalah data angka dimana dapat dilakukan operasi matematika seperti penambahan, pengurangan, pembagian, perkalian, dan tidak memiliki titik nol yang pasti. Adapun penerapan dari data rasio yaitu pada pengukuran berat, tinggi dan jumlah.

Contoh dari data rasio misalkan jumlah mahasiswa yang datang pada suatu perkuliahan dimana hari senin sebanyak 20 mahasiswa yang datang, hari selasa sebanyak 10 mahasiswa, dan hari rabu sebanyak 0 mahasiswa. Pada hari rabu dikatakan mahasiswa yang hadir dalam perkuliahan sebanyak nol mahasiswa karena memang tidak ada mahasiswa yang hadir pada hari tersebut. Selain itu sifat rasio pada contoh di atas adalah jumlah mahasiswa yang hadir pada hari senin sebanyak 2 kali lipat jumlah mahasiswa yang hadir pada hari selasa.

2.4 Normalisasi

Normalisasi adalah proses transformasi sebuah atribut numerik kedalam *range* yang kecil [JBM-11]. Adapun metode yang dapat digunakan pada proses normalisasi antara lain:

- *Min-Max*: Normalisasi *min-max* bekerja dengan melakukan pemetaan pada value v dari atribut A menjadi v' kedalam *range* $[new_min_A, new_max_A]$ [JBM-11] dimana batasan nilai baru minimum dan maksimum ditentukan oleh user itu sendiri. Berikut adalah rumus yang digunakan untuk menentukan nilai *Min-Max*:

$$v' = \frac{v - \min_A}{\max_A - \min_A} (new_max_A - new_min_A) + new_min_A \dots \dots \dots (2.1)$$

Keterangan:

- v' = Hasil normalisasi pada rentang nilai tertentu.
- v = Nilai attribut A yang akan dinormalisasi.

- \min_A = Nilai minimum atribut A.
- \max_A = Nilai maksimum atribut A.
- new_min_A = Nilai minimum baru atribut A.
- new_max_A = Nilai maksimum baru atribut A.

Nilai new_min_A dan nilai new_max_A ditentukan oleh peneliti karena digunakan untuk mengubah atribut kedalam skala atau *range* yang lebih kecil [WBS-06].

2.5 Clustering

Clustering adalah salah satu metode dalam data mining yang digunakan untuk menganalisa data dengan tujuan untuk mengelompokan data sesuai dengan karakteristiknya kedalam wilayah wilayah dimana data dengan karakteristik sama akan ditempatkan pada wilayah yang sama. Ada beberapa jenis pendekatan yang digunakan pada metode *clustering* dimana terdapat dua pendekatan utama yaitu pendekatan partisi dan pendekatan hirarki. [WBS-02].

Clustering dengan pendekatan partisi (*partition-based clustering*) bekerja dengan cara mengelompokan data dengan memilih data yang akan di analisis kedalam *cluster – cluster* yang ada. Sedangkan *Clustering* dengan menggunakan pendekatan Hirarki (*hierarchical clustering*) bekerja dengan mengelompokan data dengan membuat suatu hirarki berupa dendogram dimana data dengan kemiripan yang sama akan di tempatkan secara berdekatan sedangkan data dengan kemiripan yang jauh akan ditempatkan berjauhan. Selain kedua pendekatan di atas, terdapat juga pendekatan *automatic mapping*. Adapun metode metode yang digunakan pada setiap pendekatan antara lain sebagai berikut: [WBS-02].

Clustering dengan pendekatan Partisi:

1. *K-Means*.
2. *Mixture Modelling (Mixture Modeling)*.

Clustering dengan pendekatan Hirarki:

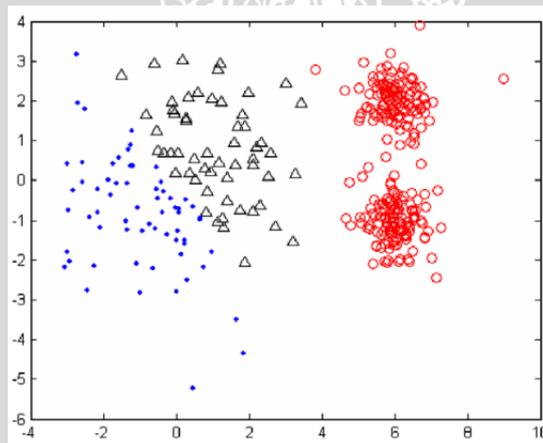
1. *Agglomerative hierarchical clustering*.
2. *Divisive hierarchical clustering*.
3. *Single Linkage*.
4. *Complete Linkage*.
5. *Average Linkage*.
6. *Average Group Linkage*.

Clustering dengan pendekatan Automatic Mapping:

1. *Self-Organising Map (SOM).*

Variasi Metode Clustering:

1. *Quality Threshold Clustering Method.*
2. *Locality Sensitive Hashing.*
3. *Algoritma Rock.*
4. *Hierarchical Frequent Term-Base Clustering.*
5. *Suffix Tree Clustering.*
6. *Single Pass Clustering.*
7. *Neighborhood Clustering.*
8. *Sequence Clustering.*
9. *Spectral Clustering.*
10. *Clustering on Frequent Tree*



Gambar 2.1 Clustering [PHC-09].

2.6 K-Means

K-Means adalah salah satu algoritma dari algoritma *clustering* dimana memiliki tujuan untuk membagi data kedalam beberapa kelompok. Algoritma ini menerima masukan data berupa data tanpa label kelas. Berbeda dengan *supervised learning* dimana menerima masukan berupa vektor $(X_1, Y_1), (X_2, Y_2), \dots, (X_i, Y_i)$ dimana X_i adalah data dari suatu data latih dan Y_i adalah label kelas dari X_i . [WBS-03].

K-Means adalah salah satu metode dari algoritma *Clustering* yang paling umum digunakan dalam berbagai aplikasi. *K-Means* adalah algoritma yang digunakan untuk mengelompokkan suatu data berdasarkan kemiripan atribut menjadi *k* kluster. Adapun alur dari algoritma *K-Means* adalah sebagai berikut: [MSN-05].

- 1) Menentukan Jumlah *k cluster*.
- 2) Mengalokasikan data sejumlah *k* untuk dijadikan pusat *cluster*.
- 3) Menentukan keanggotaan *cluster* dengan mengumpulkan data kedalam pusat *cluster* terdekat.

$$d(x_i, x_j) = (|x_{i1} - x_{j1}|^2 + |x_{i2} - x_{j2}|^2 + \dots + |x_{ip} - x_{jp}|^2)^{1/2} \dots\dots(2.4)$$

Keterangan $g = 1$ untuk menghitung *Manhattan*.
 $g = 2$ untuk menghitung *Euclidean*.
 $g = \infty$ untuk menghitung *Chebychev*.
 x_i, x_j adalah data yang dihitung jaraknya.
 p = dimensi dari data.

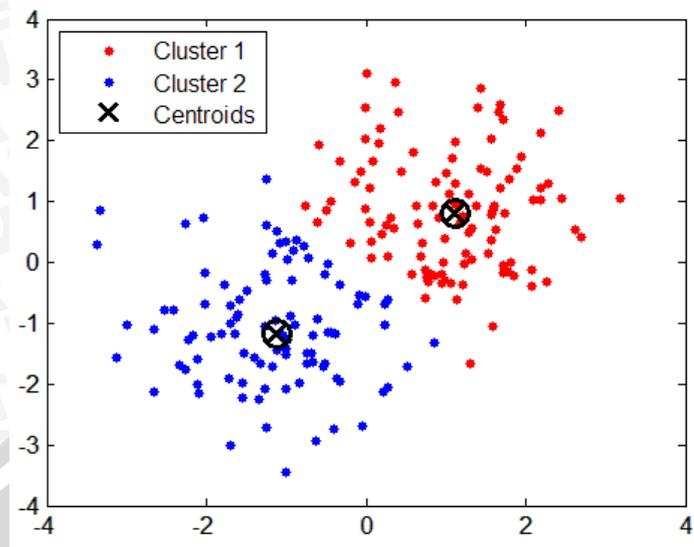
- 4) Mengkalkulasikan nilai tengah dari suatu *cluster* untuk mendapatkan pusat klaster yang baru.

$$\mu_k = \frac{1}{N_k} \sum_{q=1}^{N_k} x_q \dots\dots\dots(2.5)$$

Keterangan μ_k = Titik centroid cluster *K*
 N_k = Banyaknya data pada cluster *K*
 X_q = Data ke-*q* pada cluster *K*

- 5) Lakukan iterasi sampai tidak ada lagi perubahan pusat *cluster* [MSN-05].





Gambar 2.2 *K-Means Clustering* [WBS-05]

Adapun kelebihan dari penggunaan algoritma *K-Means* antara lain: [WBS-03].

1. Implementasi dari algoritma *K-Means* cukup mudah.
2. Algoritma *K-Means* mudah untuk dijalankan.
3. Waktu pembelajaran algoritma *K-Means* tidak membutuhkan waktu yang lama.
4. Algoritma *K-Means* mudah untuk diadaptasikan.
5. Algoritma *K-Means* umum digunakan.

Sedangkan kelemahan dari penggunaan metode *K-Means* antara lain: [WBS-03].

1. Sebelum algoritma *K-Means* dijalankan, k centroid diinisialisasikan secara acak sehingga akan menghasilkan pengelompokan data yang berbeda-beda. Penggunaan inisialisasi secara acak mengakibatkan pengelompokan menjadi kurang baik dan hasil pengelompokannya pun tidak maksimal.
2. Penggunaan algoritma *K-Means* dapat terjebak pada masalah *curse of dimensionality* dimana jika data latih memiliki dimensi yang sangat tinggi (misalnya data latih terdiri dari 30 atribut maka dimensinya pun akan menjadi 30). Cara kerja algoritma *K-Means* adalah mencari jarak antar titik pada dimensi yang ada, sehingga apabila dimensinya terlalu banyak, maka hal seperti ini akan menjadi sulit untuk dilakukan.

3. Pada penggunaan algoritma *K-Means* apabila hanya terdapat beberapa titik sample pada data, maka perhitungan dan pencarian titik terdekat kepada pusat *cluster* yang diinisialisasikan secara acak akan mudah. Akan tetapi apabila terdapat banyak sekali titik maka perhitungan dan pencarian titik akan memakan waktu yang sangat lama. Proses dapat dipercepat dengan menggunakan struktur data yang lebih rumit seperti penggunaan *kD-Tree* atau *hashing*.

2.7 Improve K-Means Inisialisasi Centroid Awal

Inisialisasi centroid awal adalah salah satu metode dari *improve K-Means* yang merupakan penggabungan dari dua perhitungan algoritma yaitu *K-Medoids* dan *K-Means*. Pada *inisialisasi centroid awal* ini penentuan pusat *cluster* menggunakan teknik penentuan pusat *cluster* pada algoritma *K-Medoids* yaitu dengan melakukan perhitungan jarak objek *i* dengan objek *j* dengan *Euclidean Distance* dilanjutkan dengan kalkulasi v_j untuk objek *j* yang kemudian diurutkan dari yang paling kecil. Setelah data terurut sesuai dengan nilai v_j yang terkecil, tentukan jumlah *k* untuk menjadi pusat klaster. Adapun rumus perhitungan kalkulasi v_j adalah sebagai berikut: [PHC-09].

$$V_j = \sum_{i=1}^n \frac{d_{ij}}{\sum_{i=1}^n d_{ij}}, \quad j = 1, \dots, n \dots \dots \dots (2.6)$$

Keterangan d_{ij} = Jarak objek *i* dengan objek *j*

Setelah proses inisialisasi pada *improve K-Means* selesai, maka dilanjutkan pada proses *K-Means* poin 4 dan poin 5.

2.8 Silhouette Coefficient

Silhouette Coefficient adalah salah satu teknik yang digunakan untuk melihat seberapa baik kualitas *cluster*. metode *Silhouette Coefficient* digunakan untuk mengukur kualitas dan kekuatan *cluster* dan seberapa baik objek pada *cluster* tersebut. Untuk tiap *cluster* nilai dari *Silhouette Coefficient* adalah nilai rata-rata SC dari tiap dokumen pada *cluster*. Nilai SC berada pada rentang -1 sampai 1. Semakin kecil nilai SC, maka semakin jelek hasil *cluster* yang diharapkan. Hal tersebut berarti objek memiliki tingkat kemiripan yang lebih mirip dengan dokumen di

cluster lain. Adapun tahapan dalam perhitungan *Silhouette Coefficient* adalah sebagai berikut: [HAR-13]

1. Mengitung rata-rata jarak data i dengan semua data lain yang ada dalam cluster yang sama.

$$a(i) = \frac{1}{|A|} - 1 \sum_{j \in A, j \neq i} d(i, j) \dots \dots \dots (2.8)$$

Keterangan:

j = Dokumen lain dalam satu cluster A.

$d(i, j)$ = Jarak antara dokumen i dan j .

$a(i)$ = rata-rata jarak.

A = jumlah anggota cluster data i

2. Hitung rata-rata dari data i dengan semua data di cluster lain kemudian ambil nilai terkecil.

$$d(i, C) = \frac{1}{|A|} \sum_{j \in C} d(i, j) \dots \dots \dots (2.9)$$

$$b(i) = \min_{C \neq A} d(i, C) \dots \dots \dots (2.10)$$

Keterangan :

$d(i, C)$ = Jarak rata – rata data i dengan semua objek pada cluster lain.

C = Jumlah anggota cluster selain A.

$d(i, j)$ = Jarak data i ke dokumen anggota cluster selain data i .

$b(i)$ = Nilai minimum dari $d(i, C)$.

3. Maka *Silhouette Coefficient* adalah:

$$s(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))} \dots \dots \dots (2.11)$$

Keterangan :

$s(i)$ = nilai *Silhouette Coefficient*



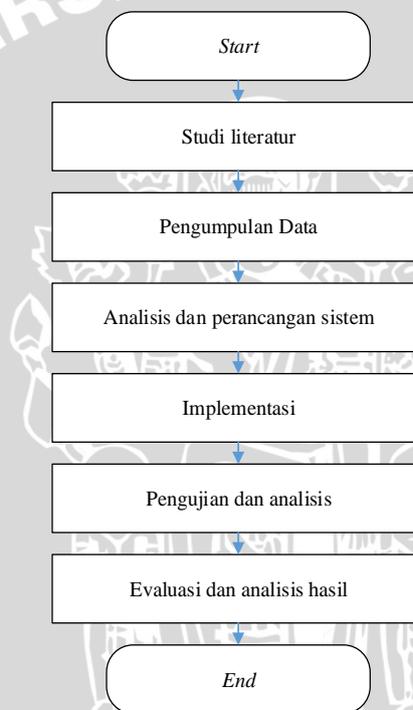
BAB III

METODOLOGI PENELITIAN DAN PERANCANGAN

Pada bab metodologi penelitian dan perancangan akan dibahas mengenai metode yang digunakan dan perancangan yang akan dibuat meliputi langkah-langkah penelitian hingga gambaran penggunaan metode pada implementasi rekomendasi jodoh menggunakan algoritma *Improve K- Means pada inisialisasi centroid*.

3.1 Metodologi Penelitian

Adapun diagram alir pada penelitian ini adalah sebagai berikut:



Gambar 3.1 Diagram Alir

Adapun penjelasan secara rinci terkait metodologi yang digunakan pada Gambar 3.1 adalah sebagai berikut:

1. Mempelajari literature-literatur yang terkait dengan metode *Improve K- Means pada inisialisasi centroid*.
2. Mengumpulkan dan menganalisis data.
3. Melakukan analisis dan perancangan sistem untuk menentukan rekomendasi pasangan hidup dengan menggunakan metode *Improve K- Means pada inisialisasi centroid*.

4. Mengimplementasikan hasil perancangan kedalam proses klustering yang terkomputerisasi.
5. Melakukan pengujian dan analisis pada aplikasi yang telah dibuat.
6. Mengevaluasi dan menganalisis tingkat kualitas kluster berdasarkan perhitungan manualisasi dan hasil dari implementasi.

3.1.1 Studi Literatur

Untuk menyelesaikan masalah pada penelitian ini agar sesuai dengan tujuan yang akan dicapai maka dibutuhkan studi literatur. Adapun pembelajarannya teori-teori yang digunakan pada penelitian contohnya pengolahan data yang belum siap digunakan agar menjadi data yang siap digunakan. Kemudian mempelajari metode *Improve K- Means pada inisialisasi centroid* sebagai dasar penelitian yang diperoleh dari jurnal penelitian, buku dan internet serta beberapa *literature* lain yang mendukung sesuai yang telah dijabarkan pada bab 2.

3.1.2 Pengumpulan data

Pengumpulan data dilakukan dengan melakukan survey menggunakan google form (<http://goo.gl/forms/9QBqtYUjmi>) kepada masyarakat terutama bagi mereka yang ingin mencari pasangan hidup. Dari hasil survey didapatkan data sebanyak 150 data yang digunakan untuk parameter klustering untuk penentuan calon pasangan hidup. Dari hasil survey didapatkan sebanyak 10 parameter yang akan digunakan yaitu jenis kelamin, Umur, tinggi, berat, jenis rambut, asal suku, warna kulit, agama, status pernikahan dan penggunaan rokok.

3.1.3 Analisis dan Perancangan Sistem

Pada analisis dan perancangan sistem membahas hal-hal yang terkait pada proses pembuatannya antaralain deskripsi sistem, perancangan sistem, manualisasi, desain *interface*, dan perancangan pengujian dan evaluasi. Secara garis besar aplikasi yang akan dibangun yaitu akan merekomendasikan pasangan hidup dengan menggunakan metode *Improve K- Means pada inisialisasi centroid* sesuai dengan kriteria yang user inginkan.

Adapun kebutuhan perangkat keras yang digunakan agar proses implementasi tidak mengalami masalah yaitu memiliki spesifikasi sebagai berikut:

1. Performa : AMD A8-4500M APU with Radeon(tm) HD Graphics (4 CPUs), ~1.90 GHz
2. Sistem Operasi : Windows 8.1 Pro 64-bit (6.3, Build 9600)
3. RAM : 8.00 GB
4. VGA card : AMD Radeon HD 7640G dan AMD Radeon HD 8600/8700M 6GB
5. Aplikasi di bangun dengan menggunakan Bahasa pemrograman JAVA
6. Database Aplikasi dibangun menggunakan aplikasi NetBeans 8.0

Tahap perancangan sistem adalah tahap pembuatan alur kerja untuk memecahkan permasalahan pengelompokan data menggunakan metode *Improve K-Means pada inisialisasi centroid* menggunakan parameter dan variabel, penggunaan data kedalam algoritma, desain aplikasi yang menampilkan proses dan hasil akurasi serta desain pengujian.

3.1.4 Implementasi Sistem

Proses implementasi adalah proses pengaplikasian sistem sesuai dengan analisis dan perancangan sistem yang telah dibuat sebelumnya. Untuk pemecahan permasalahan penentuan pasangan hidup diimplementasikan ke dalam metode *Improve K-Means pada inisialisasi centroid* yang selanjutnya akan dilakukan pengujian. Implementasi akan dilakukan dengan menggunakan Bahasa pemrograman web yang meliputi PHP, HTML, CSS, *Javascript* dengan editornya *Dreamweaver*. Implementasi basis data dilakukan dengan menggunakan *database management system (DBMS)* MySQL dengan software XAMPP.

3.1.5 Pengujian dan Analisis Sistem

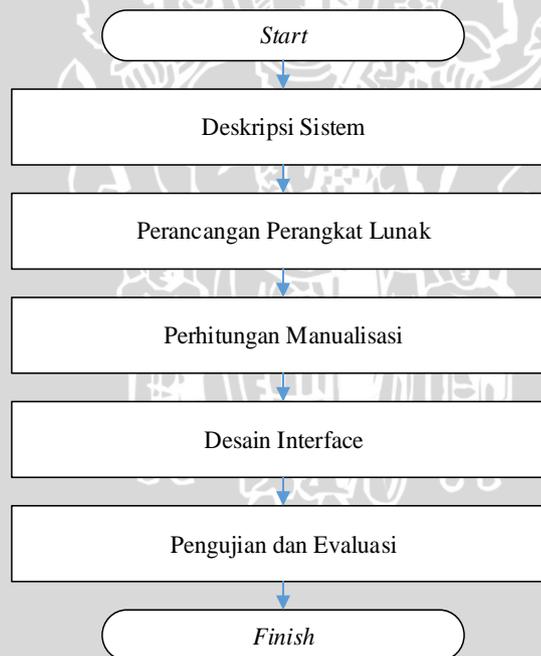
Pada tahapan ini permasalahan yang diimplementasikan ke dalam bentuk aplikasi akan diuji tingkat keberhasilannya. Hasil dari uji coba tersebut akan dijadikan bahan evaluasi agar dapat sesuai dengan tujuan yang diharapkan. Adapun pengujian yang digunakan yaitu dengan menggunakan teknik *silhouette index* untuk mengukur kesesuaian hasil pengklasteran.

3.1.6 Evaluasi dan Analisis hasil

Pada tahapan ini akan dilakukan evaluasi untuk mempertimbangkan tingkat keberhasilan implementasi metode pada aplikasi dalam menyelesaikan permasalahan apakah sudah sesuai dengan tujuan penelitian. Kemudian dilanjutkan proses penarikan kesimpulan yang merupakan tahap paling akhir dimana berisikan hasil yang diperoleh dari proses pengujian dan analisis yang telah melalui proses evaluasi. Kemudian dilanjutkan dengan kesimpulan dan saran yang dapat digunakan oleh pembaca sebagai acuan untuk memperbaiki dan mengembangkan aplikasi ke depannya.

3.2 Perancangan Sistem

Pada tahapan perancangan sistem terdiri dari deskripsi sistem secara umum, perancangan aplikasi, manualisasi, desain *interface*, dan pengujian dan evaluasi. Pada gambar 3.2 di bawah ini merupakan model perancangan sistem yang digunakan pada proses penelitian ini.



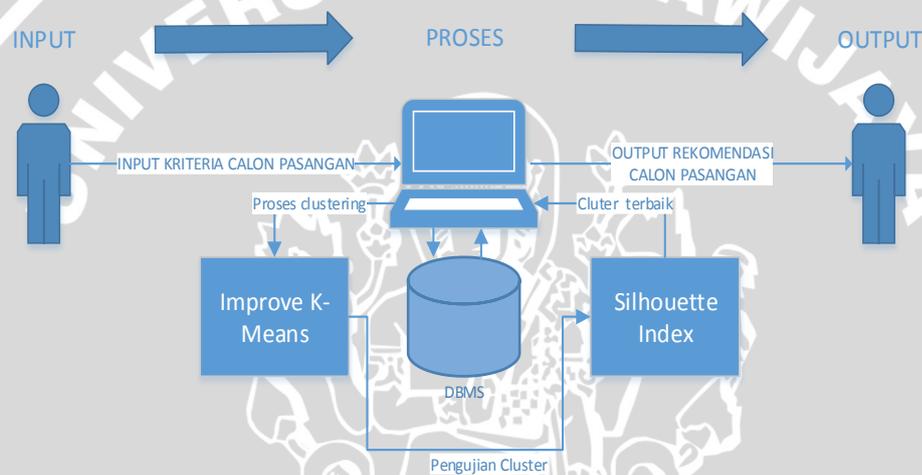
Gambar 3.2 Diagram Alir Tahap Perancangan Sistem

3.2.1 Deskripsi Sistem

Penelitian ini dilakukan dengan menerapkan metode *Improve K- Means* pada *inisialisasi centroid* dalam penentuan kluster calon pasangan hidup. Permasalahan *clustering* adalah permasalahan pengelompokan beberapa data

kedalam beberapa kelas dengan menyiapkan data training dan data testing, dimana metode *Improve K- Means pada inisialisasi centroid* digunakan untuk menghitung jarak data terhadap pusat *cluster*. Terdapat 10 parameter dari 150 responden yang akan digunakan pada penelitian ini. Terdapat input yaitu yang proses penambahan data yang akan diklasterisasi kedalam *database*.

Sedangkan *Input* pencarian yang digunakan adalah kriteria calon yang diinginkan yaitu umur, tinggi badan, berat badan, jenis rambut, suku, warna kulit, agama, aktif merokok. *Output* dari aplikasi ini yaitu kluster dimana terdapat kecocokan yang dapat dijadikan rekomendasi calon pasangan hidup. Adapun desain umum ditunjukkan pada gambar 3.3.

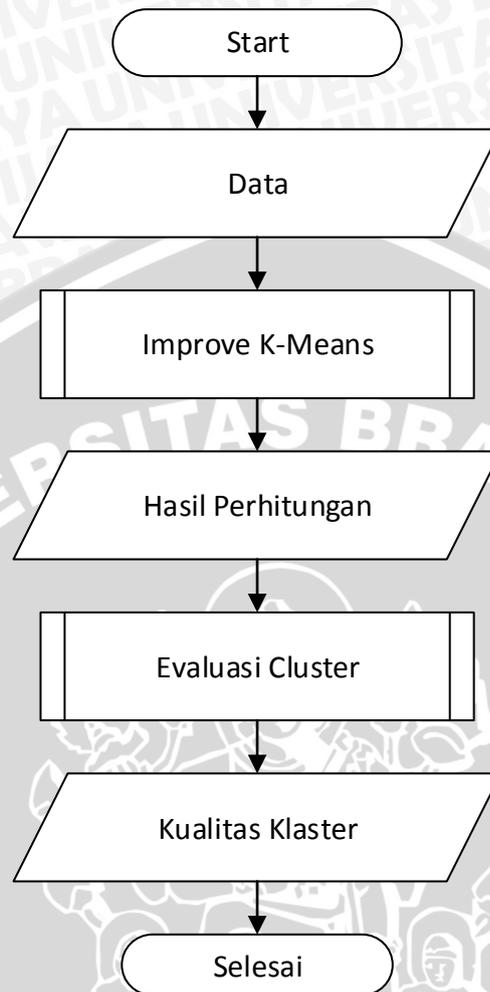


Gambar 3.3 Gambaran Sistem

3.2.2 Perancangan Perangkat Lunak

Tahap perancangan perangkat lunak pada proses penerapan metode *Improve K- Means pada inisialisasi centroid* untuk penentuan rekomendasi jodoh yang dijalankan pada program aplikasi digambarkan pada diagram alir sebelumnya.

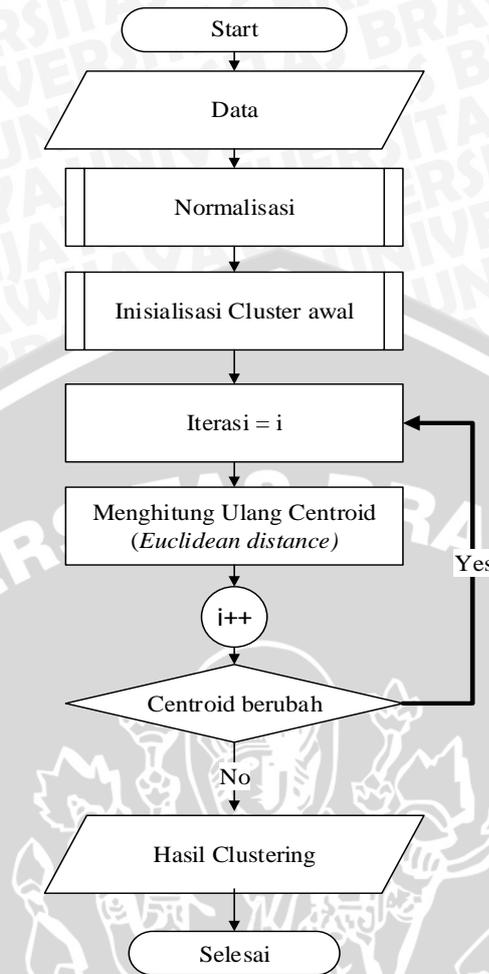
Pada prosesnya terdapat parameter yang digunakan. Adapun parameter yang digunakan yaitu jenis kelamin, umur, tinggi badan, berat badan, jenis rambut, suku, warna kulit, agama, aktif merokok. Adapun diagram utama dari jalannya proses perhitungan ditunjukkan pada gambar 3.4



Gambar 3.4 Diagram Utama Sistem

Data pada *database* sejumlah 150 data diproses menggunakan metode *Improve K-Means* pada inisialisasi *centroid* awal sampai data terklaster secara keseluruhan. Kemudian setelah didapatkan hasil klustering dilanjutkan dengan proses pengujian dengan menggunakan metode *Shilhouette Coefficient*.

Adapun diagram alir sistem yaitu tahapan proses perhitungan dengan metode *Improve K-Means* pada inisialisasi *centroid* ditunjukkan pada Gambar 3.5.



Gambar 3.5 Diagram Alir Sistem

Adapun tahapan-tahapan yang ditunjukkan pada Gambar 3.5 dapat dijabarkan sebagai berikut :

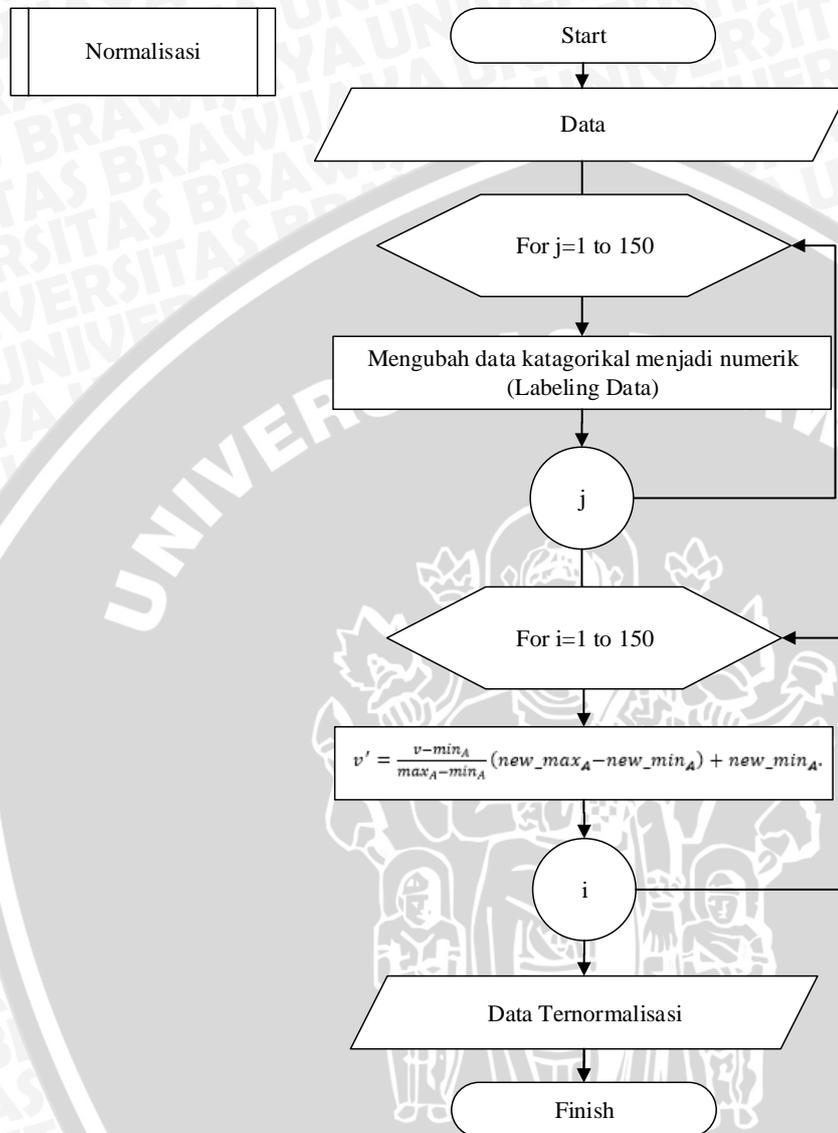
1. Masukkan data identitas diri dan kriteria calon pasangan.
2. Menjalankan proses normalisasi.
3. Mendapatkan hasil normalisasi.
4. Menjalankan proses inisialisasi cluster awal.
5. Mendapatkan hasil inisialisasi cluster awal.
6. Menghitung ulang centroid dengan menggunakan rumus:

$$d(x_i, x_j) = (|x_{i1} - x_{j1}|^2 + |x_{i2} - x_{j2}|^2 + \dots + |x_{ip} - x_{jp}|^2)^{1/2}$$

7. Apabila terjadi perubahan *cluster* pada data maka kembali ke tahap 6, apabila tidak lanjut ke tahap selanjutnya.
8. Mendapatkan hasil *cluster* terbaik.

3.2.2.1 Proses Normalisasi

Secara garis besar proses dari normalisasi ditunjukkan pada Gambar 3.6.



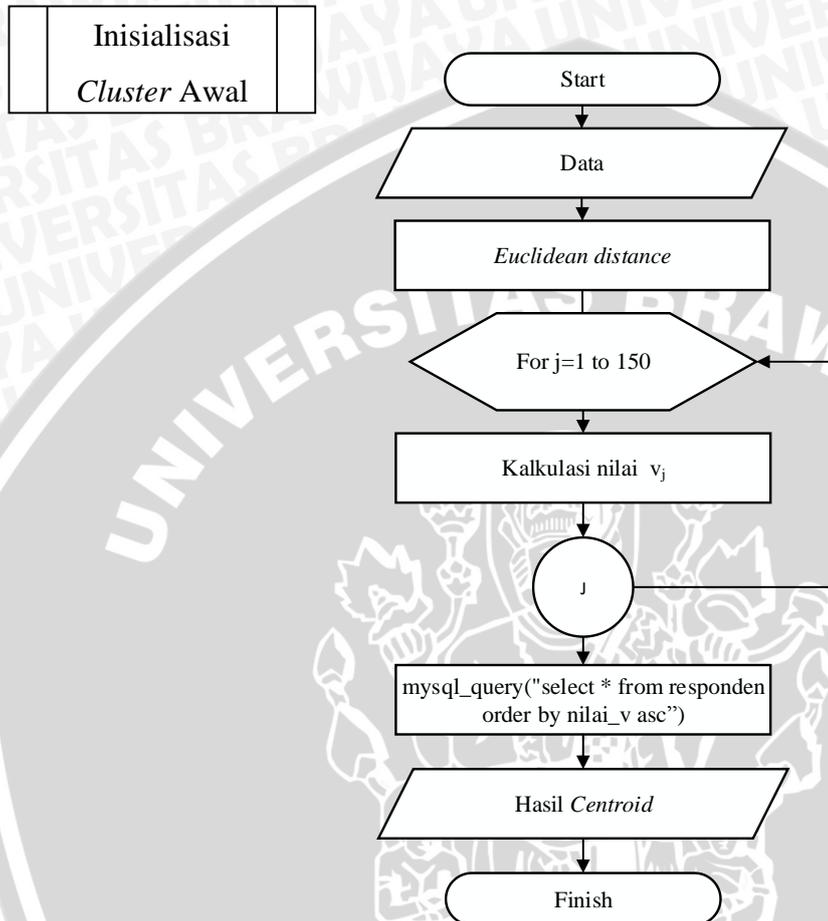
Gambar 3.6 Diagram Alir Proses Normalisasi

Adapun proses normalisasi yang ditunjukkan pada Gambar 3.6 dapat dijabarkan sebagai berikut :

1. Masukkan data.
2. Melakukan labeling untuk semua data (150 data).
3. Melakukan Normalisasi untuk semua data (150 data).
4. Proses normalisasi selesai.

3.2.2.2 Proses inisialisasi *cluster* awal

Secara garis besar proses dari Inisialisasi *Cluster* Awal ditunjukkan pada Gambar 3.7.



Gambar 3.7 Diagram Alir Proses Inisialisasi *Cluster* Awal

Adapun proses Inisialisasi *Cluster* Awal yang ditunjukkan pada Gambar 3.7 dapat dijabarkan sebagai berikut :

1. Masukkan data.
2. Menghitung jarak antar data secara keseluruhan.
3. Mengkalkulasi nilai v_j untuk semua data(150 data).
4. Mengurutkan v_j 's dari yang terkecil kemudian pilih k objek yang memiliki nilai terkecil.
5. Menentukan k objek yang memiliki nilai terkecil.
6. Proses Inisialisasi *Cluster* awal selesai.

3.2.3 Perhitungan Manualisasi

Pada sub bab ini akan membahas tentang perhitungan manualisasi dari metode *Improve K- Means pada inisialisasi centroid awal* yang akan di implementasikan. Data yang digunakan dalam perhitungan manual ini berjumlah 12 data yang telah diambil dari antara data yang akan digunakan pada proses pengujian. Data yang digunakan pada perhitungan manualisasi ditunjukkan pada tabel 3.1 data sampel manualisasi.

Table 3.1 Data sampel manualisasi

Data	JK	Umur (Tahun)	Tinggi (cm)	Berat (kg)	Rambut	Ras	Kulit	Agama	Status	Merokok
Data 1	Perempuan	20 < 25 tahun	155 < 165 cm	< 50 kg	Keriting	Jawa	Sawo Matang	Islam	Lajang	Tidak
Data 2	Perempuan	20 < 25 tahun	155 < 165 cm	< 50 kg	Ikal	Jawa	Kuning Langsung	Islam	Lajang	Tidak
Data 3	Perempuan	20 < 25 tahun	155 < 165 cm	50 < 60 kg	Lurus	Jawa	Sawo Matang	Islam	Lajang	Tidak
Data 4	Perempuan	20 < 25 tahun	< 155 cm	< 50 kg	Lurus	Jawa	Sawo Matang	Islam	Lajang	Tidak
Data 5	Laki - Laki	20 < 25 tahun	165 < 175 cm	50 < 60 kg	Ikal	Minang	Sawo Matang	Islam	Lajang	Tidak

Data	JK	Umur (Tahun)	Tinggi (cm)	Berat (kg)	Rambut	Ras	Kulit	Agama	Status	Merokok
Data 6	Perempuan	20 < 25 tahun	155 < 165 cm	60 < 70 kg	Ikal	Jawa	Sawo Matang	Islam	Lajang	Tidak
Data 7	Laki - Laki	20 < 25 tahun	165 < 175 cm	50 < 60 kg	Lurus	Jawa	Sawo Matang	Islam	Lajang	Tidak
Data 8	Perempuan	20 < 25 tahun	155 < 165 cm	50 < 60 kg	Lurus	Jawa	Kuning Langsung	Islam	Lajang	Tidak
Data 9	Laki - Laki	20 < 25 tahun	165 < 175 cm	50 < 60 kg	Ikal	Jawa	Sawo Matang	Islam	Lajang	Tidak
Data 10	Laki - Laki	20 < 25 tahun	> 175 cm	60 < 70 kg	Lurus	Jawa	Sawo Matang	Islam	Lajang	Tidak
Data 11	Laki - Laki	25 < 30 tahun	> 175 cm	60 < 70 kg	Lurus	Jawa	Sawo Matang	Islam	Lajang	Ya
Data 12	Laki - Laki	25 < 30 tahun	> 175 cm	> 70 kg	Lurus	Tionghoa	Putih	Budha	Lajang	Ya

3.2.3.1 Manualisasi Pelabelan

Manualisasi pelabelan adalah proses yang digunakan untuk melabeli kriteria dari data yang di dapat pada proses survey. Pada proses ini data akan di labeli dengan angka 0 sampai 10 dimana angka tersebut tidak akan mempengaruhi nilai dari hasil survey karena hanya sebagai simbol. Contohnya pada kriteria rambut yaitu rambut dengan tipe lurus diberi simbol “0”, rambut tipe keriting diberi simbol “1”, rambut tipe

ikal diberi simbol “2”, dan rambut tipe botak diberi simbol “3”. Adapun hasil dari proses pelabelan ditunjukkan pada tabel 3.2 hasil proses pelabelan.

Table 3.2 Hasil proses pelabelan

Data hasil transformasi	JK	Umur	Tinggi	Berat	Rambut	Ras	Kulit	Agama	Status	Merokok
Data 1	0	0	1	0	1	1	1	0	0	0
Data 2	0	0	1	0	0	1	2	0	0	0
Data 3	0	0	1	1	2	1	1	0	0	0
Data 4	0	0	0	0	2	1	1	0	0	0
Data 5	1	0	2	1	0	2	1	0	0	0
Data 6	0	0	1	2	0	1	1	0	0	0
Data 7	1	0	2	1	2	1	1	0	0	0
Data 8	0	0	1	1	2	1	2	0	0	0
Data 9	1	0	2	1	0	1	1	0	0	0
Data 10	1	0	3	2	2	1	1	0	0	0
Data 11	1	1	3	2	2	1	1	0	0	1
Data 12	1	1	3	3	2	4	0	3	0	1

3.2.3.2 Manualisasi normalisasi

Normalisasi adalah proses transformasi sebuah atribut numerik kedalam *range* yang kecil [JBM-11]. Tujuan dilakukannya normalisasi adalah untuk memperbaiki desain tabel yang kurang baik sehingga penyimpanan data yang ada lebih efisien dan bebas anomaly. Normalisasi digunakan untuk meminimalkan adanya *redundansi* data dan menjamin integritas data [WBS-06]. Adapun proses dari normalisasi adalah sebagai berikut.

Data yang digunakan

	JK	Umur	Tinggi	Berat	Rambut	Ras	Kulit	Agama	Status	Merokok
Data 1	0	0	1	0	1	1	1	0	0	0

$$v' = \frac{v - \min_A}{\max_A - \min_A} (\text{new_max}_A - \text{new_min}_A) + \text{new_min}_A$$

Data yang di normalisasi

Data 1 = 1 (Dapat dilihat pada tabel 3. 2 hasil proses pelabelan).

$$v' = (1 - 0) / (3 - 0) \times ((1 - 0) + 0) = 0,333$$

Adapun hasil secara keseluruhan ditunjukkan pada tabel 3.3 hasil proses normalisasi.

Table 3.3 Hasil proses Normalisasi

Data hasil normalisasi	JK	Umur	Tinggi	Berat	Rambut	Ras	Kulit	Agama	Status	Merokok
Data 1	0.000	0.000	0.333	0.000	0.333	0.000	0.333	0.000	0.000	0.000
Data 2	0.000	0.000	0.333	0.000	0.000	0.000	0.667	0.000	0.000	0.000
Data 3	0.000	0.000	0.333	0.333	0.667	0.000	0.333	0.000	0.000	0.000
Data 4	0.000	0.000	0.000	0.000	0.667	0.000	0.333	0.000	0.000	0.000
Data 5	1.000	0.000	0.667	0.333	0.000	0.400	0.333	0.000	0.000	0.000
Data 6	0.000	0.000	0.333	0.667	0.000	0.000	0.333	0.000	0.000	0.000
Data 7	1.000	0.000	0.667	0.333	0.667	0.000	0.333	0.000	0.000	0.000
Data 8	0.000	0.000	0.333	0.333	0.667	0.000	0.667	0.000	0.000	0.000
Data 9	1.000	0.000	0.667	0.333	0.000	0.000	0.333	0.000	0.000	0.000
Data 10	1.000	0.000	1.000	0.667	0.667	0.000	0.333	0.000	0.000	0.000
Data 11	1.000	0.333	1.000	0.667	0.667	0.000	0.333	0.000	0.000	1.000
Data 12	1.000	0.333	1.000	1.000	0.667	0.900	0.000	1.000	0.000	1.000

3.2.3.3 Manualisasi Improve K-Means inisialisasi cluster awal

Manualisasi inisialisasi *cluster* awal adalah tahapan awal pada proses perhitungan dimana digunakan untuk menentukan *centroid* awal. Adapun tahapan pada proses ini adalah sebagai berikut:

1. Menghitung jarak antara objek i dengan objek j dengan menggunakan perhitungan *Euclidean Distance*. Adapun proses dari *Euclidean Distance* adalah sebagai berikut:

Data yang digunakan:

Data hasil normalisasi	JK	Umur	Tinggi	Berat	Rambut	Ras	Kulit	Agama	Status	Merokok
Data 1	0.000	0.000	0.333	0.000	0.333	0.000	0.333	0.000	0.000	0.000
Data 2	0.000	0.000	0.333	0.000	0.000	0.000	0.667	0.000	0.000	0.000

Proses perhitungan:

$$d(x_i, x_j) = (|x_{i1} - x_{j1}|^2 + |x_{i2} - x_{j2}|^2 + \dots + |x_{ip} - x_{jp}|^2)^{1/2}$$

$$d(\text{Data 1}, \text{Data 2}) = (|0,000-0,000|^2 + |0,000-0,000|^2 + |0,333-0,333|^2 + |0,000-0,000|^2 + |0,333-0,000|^2 + |0,000-0,000|^2 + |0,333-0,667|^2 + |0,000-0,000|^2 + |0,000-0,000|^2 + |0,000-0,000|^2)^{1/2} = 0.471$$

Adapun hasil perhitungan dari *Euclidean Distance* secara keseluruhan ditunjukkan pada tabel 3.4 Perhitungan jarak Euclidean Distance

Table 3.4 Perhitungan jarak Euclidean Distance.

	Data 1	Data 2	Data 3	Data 4	Data 5	Data 6	Data 7	Data 8	Data 9	Data 10	Data 11	Data 12
Data 1	0.000	0.471	0.471	0.471	1.222	0.745	1.155	0.577	1.155	1.414	1.764	2.364
Data 2	0.471	0.000	0.816	0.816	1.222	0.745	1.333	0.745	1.155	1.563	1.886	2.501
Data 3	0.471	0.816	0.000	0.471	1.310	0.745	1.054	0.333	1.247	1.247	1.633	2.218
Data 4	0.471	0.816	0.471	0.000	1.470	1.000	1.247	0.577	1.414	1.563	1.886	2.456
Data 5	1.222	1.222	1.310	1.470	0.000	1.176	0.777	1.352	0.400	0.909	1.392	1.863
Data 6	0.745	0.745	0.745	1.000	1.176	0.000	1.291	0.816	1.106	1.374	1.732	2.243
Data 7	1.155	1.333	1.054	1.247	0.777	1.291	0.000	1.106	0.667	0.471	1.155	1.894
Data 8	0.577	0.745	0.333	0.577	1.352	0.816	1.106	0.000	1.291	1.291	1.667	2.292
Data 9	1.155	1.155	1.247	1.414	0.400	1.106	0.667	1.291	0.000	0.816	1.333	2.008
Data 10	1.414	1.563	1.247	1.563	0.909	1.374	0.471	1.291	0.816	0.000	1.054	1.773
Data 11	1.764	1.886	1.633	1.886	1.392	1.732	1.155	1.667	1.333	1.054	0.000	1.426
Data 12	2.364	2.501	2.218	2.456	1.863	2.243	1.894	2.292	2.008	1.773	1.426	0.000

Didapatkan total jarak (v_j) tiap data pada objek i terhadap objek j . Adapun contoh salah satu perhitungan adalah sebagai berikut:

Data yang digunakan:

	Data 1
Data 1	0.000
Data 2	0.471
Data 3	0.471
Data 4	0.471
Data 5	1.222
Data 6	0.745
Data 7	1.155
Data 8	0.577
Data 9	1.155
Data 10	1.414
Data 11	1.764
Data 12	2.364

Proses perhitungan

$$Total = 0.000 + 0.471 + 0.471 + 0.471 + 1.222 + 0.745 + 1.155 + 0.577 + 1.155 + 1.414 + 1.764 + 2.364 = \mathbf{11.810}$$

Adapun hasil perhitungan total secara keseluruhan ditunjukkan pada tabel 3.5 Total jarak tiap data (v_j):

Table 3.5 Total jarak tiap data (v_j).

Data 1	Data 2	Data 3	Data 4	Data 5	Data 6	Data 7	Data 8	Data 9	Data 10	Data 11	Data 12
11.810	13.255	11.548	13.373	13.093	12.974	12.150	12.048	12.592	13.478	16.927	23.039

Menghitung jarak v_j untuk objek i . Adapun contoh salah satu perhitungan adalah sebagai berikut:

Data yang digunakan:

	Data 1
Data 2	0.040
Total	11.810

Perhitungan : $\text{data1}/\text{total} = 0.040$

Adapun hasil perhitungan jarak v_j terhadap objek j secara keseluruhan ditunjukkan pada tabel 3.6:

Table 3.6 Perhitungan jarak v_j terhadap objek j

	Data 1	Data 2	Data 3	Data 4	Data 5	Data 6	Data 7	Data 8	Data 9	Data 10	Data 11	Data 12
Data 1	0.000	0.036	0.041	0.035	0.093	0.057	0.095	0.048	0.092	0.105	0.104	0.103
Data 2	0.040	0.000	0.071	0.061	0.093	0.057	0.110	0.062	0.092	0.116	0.111	0.109

	Data 1	Data 2	Data 3	Data 4	Data 5	Data 6	Data 7	Data 8	Data 9	Data 10	Data 11	Data 12
Data 3	0.040	0.062	0.000	0.035	0.100	0.057	0.087	0.028	0.099	0.093	0.096	0.096
Data 4	0.040	0.062	0.041	0.000	0.112	0.077	0.103	0.048	0.112	0.116	0.111	0.107
Data 5	0.103	0.092	0.113	0.110	0.000	0.091	0.064	0.112	0.032	0.067	0.082	0.081
Data 6	0.063	0.056	0.065	0.075	0.090	0.000	0.106	0.068	0.088	0.102	0.102	0.097
Data 7	0.098	0.101	0.091	0.093	0.059	0.100	0.000	0.092	0.053	0.035	0.068	0.082
Data 8	0.049	0.056	0.029	0.043	0.103	0.063	0.091	0.000	0.103	0.096	0.098	0.099
Data 9	0.098	0.087	0.108	0.106	0.031	0.085	0.055	0.107	0.000	0.061	0.079	0.087
Data 10	0.120	0.118	0.108	0.117	0.069	0.106	0.039	0.107	0.065	0.000	0.062	0.077
Data 11	0.149	0.142	0.141	0.141	0.106	0.133	0.095	0.138	0.106	0.078	0.000	0.062
Data 12	0.200	0.189	0.192	0.184	0.142	0.173	0.156	0.190	0.159	0.132	0.084	0.000

Didapatkan total jarak v_j terhadap objek j . Adapun contoh salah satu perhitungan adalah sebagai berikut:

Data yang digunakan:

	Data 1	Data 2	Data 3	Data 4	Data 5	Data 6	Data 7	Data 8	Data 9	Data 10	Data 11	Data 12
Data 1	0.000	0.000	0.041	0.035	0.093	0.057	0.095	0.048	0.092	0.105	0.104	0.103

Perhitungan : Total = $0.000 + 0.000 + 0.041 + 0.035 + 0.093 + 0.057 + 0.095 + 0.048 + 0.092 + 0.105 + 0.104 + 0.103 = 0.809$

Adapun hasil perhitungan jarak v_j terhadap objek j secara keseluruhan ditunjukkan pada tabel 3.7 perhitungan jarak v_j terhadap objek j :

Table 3.7 Jarak v_j terhadap objek j .

	Nilai V_j
Data 1	0.809
Data 2	0.922
Data 3	0.793
Data 4	0.929
Data 5	0.948
Data 6	0.912
Data 7	0.872
Data 8	0.831
Data 9	0.903
Data 10	0.988
Data 11	1.293
Data 12	1.801

2. Urutkan v_j 's dalam urutan kemudian pilih k objek yang memiliki nilai terkecil. Urutan dari k ditunjukkan pada tabel 3.8 hasil pengurutan objek k

Table 3.8 Hasil pengurutan objek k

	Nilai V_j
Data 3	0.793
Data 1	0.809
Data 8	0.831
Data 7	0.872
Data 9	0.903
Data 6	0.912
Data 2	0.922
Data 4	0.929
Data 5	0.948
Data 10	0.988
Data 11	1.293
Data 12	1.801

Dari hasil perhitungan didapatkan 3 *centroid* yaitu data 3 pada pusat *centroid* 1 dengan total jarak 0.793024, Data 1 pada pusat *centroid* 2 dengan total jarak 0.808815, dan Data 8 pada pusat *centroid* 3 dengan total jarak 0.830576.

3. Hitung jumlah jarak dari semua objek terhadap pusat *cluster*. Adapun contoh salah satu perhitungan adalah sebagai berikut:

Data yang digunakan. Pusat Cluster:

	JK	Umur	Tinggi	Berat	Rambut	Ras	Kulit	Agama	Status	Merokok
Data 1 (Pusat Cluster 1)	0,000	0,000	0.333	0.333	0.666	0,000	0.333	0,000	0,000	0,000
Data 2 (Pusat Cluster 2)	0,000	0,000	0.333	0,000	0.333	0,000	0.333	0,000	0,000	0,000
Data 2 (Pusat Cluster 2)	0,000	0,000	0.333	0.333	0.666	0,000	0.666	0,000	0,000	0,000

Data yang dihitung:

	JK	Umur	Tinggi	Berat	Rambut	Ras	Kulit	Agama	Status	Merokok
Data 4	0,000	0,000	0,000	0,000	0.666	0,000	0.333	0,000	0,000	0,000

Proses perhitungan:

$$d(x_i, x_j) = (|x_{i1} - x_{j1}|^2 + |x_{i2} - x_{j2}|^2 + \dots + |x_{ip} - x_{jp}|^2)^{1/2}$$

$$d(\text{Pusat Cluster 1}, \text{Data 4}) = (|0,000-0,000|^2 + |0,000-0,000|^2 + |0,000-0,333|^2 + |0.333-0,000|^2 + |0.666-0,666|^2 + |0,000-0,000|^2 + |0,333-0,333|^2 + |0,000-0,000|^2 + |0,000-0,000|^2 + |0,000-0,000|^2)^{1/2} = 0.471$$

Adapun hasil perhitungan dari *Euclidean Distance* secara keseluruhan ditunjukkan pada tabel 3.9 Perhitungan jarak terhadap pusat cluster:

Table 3.9 Perhitungan jarak terhadap pusat cluster.

	Jarak C1	Jarak C2	Jarak C3	Cluster	Cluster 1	Cluster 2	Cluster 3
Data 1	0.471	0.000	0.577	C2	0	1	0
Data 2	0.816	0.471	0.745	C2	0	1	0
Data 3	0.000	0.471	0.333	C1	1	0	0
Data 4	0.471	0.471	0.577	C1	1	0	0
Data 5	1.310	1.222	1.352	C2	0	1	0
Data 6	0.745	0.745	0.816	C1	1	0	0
Data 7	1.054	1.155	1.106	C1	1	0	0
Data 8	0.333	0.577	0.000	C3	0	0	1
Data 9	1.247	1.155	1.291	C2	0	1	0
Data 10	1.247	1.414	1.291	C1	1	0	0
Data 11	1.633	1.764	1.667	C1	1	0	0
Data 12	2.218	2.364	2.292	C1	1	0	0

Dari perhitungan didapatkan hasil 7 data berada pada *cluster 1* (Data 3, Data 4, Data 6, Data 7, Data 10, Data 11, dan Data 12), 4 data berada pada *cluster 2* (Data 1, Data 2, Data 5, dan Data 9), dan 1 data berada pada *cluster 3* (Data 8).

4. Menentukan keanggotaan *cluster* dengan mengumpulkan data kedalam pusat *cluster* terdekat. Adapun contoh salah satu perhitungan adalah sebagai berikut:

Data yang digunakan. Hasil Normalisasi Data 1

Data hasil normalisasi	JK	Umur	Tinggi	Berat	Rambut	Ras	Kulit	Agama	Status	Merokok
Data 1	0.000	0.000	0.333	0.000	0.333	0.000	0.333	0.000	0.000	0.000

Dan hasil clustering

	Jarak C1	Jarak C2	Jarak C3	Cluster	Cluster 1	Cluster 2	Cluster 3
Data 1	0.471	0.000	0.577	C2	0	1	0

Maka

	JK	Umur	Tinggi	Berat	Rambut	Ras	Kulit	Agama	Status	Merokok
Data 1	0.000 x 1	0.000 x 1	0.333 x 1	0.000 x 1	0.333 x 1	0.000 x 1	0.333 x 1	0.000 x 1	0.000 x 1	0.000 x 1

Hasil akan ditampilkan pada tabel 3.12 keanggotaan cluster 3

Adapun hasil pengumpulan data kedalam pusat *cluster* secara lengkap ditunjukkan pada table 3.10:

Table 3.10 Keanggotaan Cluster 1

	JK	Umur	Tinggi	Berat	Rambut	Ras	Kulit	Agama	Status	Merokok
Data 3	0.000	0.000	0.333	0.333	0.667	0.000	0.333	0.000	0.000	0.000
Data 4	0.000	0.000	0.000	0.000	0.667	0.000	0.333	0.000	0.000	0.000
Data 6	0.000	0.000	0.333	0.667	0.000	0.000	0.333	0.000	0.000	0.000
Data 7	1.000	0.000	0.667	0.333	0.667	0.000	0.333	0.000	0.000	0.000
Data 10	1.000	0.000	1.000	0.667	0.667	0.000	0.333	0.000	0.000	0.000
Data 11	1.000	0.333	1.000	0.667	0.667	0.000	0.333	0.000	0.000	1.000
Data 12	1.000	0.333	1.000	1.000	0.667	0.900	0.000	1.000	0.000	1.000
Total	0.571	0.095	0.619	0.524	0.571	0.129	0.286	0.143	0.000	0.286

Perhitungan total = $(0.000 + 0.000 + 0.000 + 1.000 + 1.000 + 1.000 + 1.000) / \text{jumlah data dalam cluster}$

= $(0.000 + 0.000 + 0.000 + 1.000 + 1.000 + 1.000 + 1.000) / 7 = \mathbf{0.571}$ (ditunjukkan pada tabel 3.10 keanggotaan cluster 1)

Table 3.11 Keanggotaan Cluster 2

	JK	Umur	Tinggi	Berat	Rambut	Ras	Kulit	Agama	Status	Merokok
Data 1	0.000	0.000	0.333	0.000	0.333	0.000	0.333	0.000	0.000	0.000
Data 2	0.000	0.000	0.333	0.000	0.000	0.000	0.667	0.000	0.000	0.000
Data 5	1.000	0.000	0.667	0.333	0.000	0.400	0.333	0.000	0.000	0.000

	JK	Umur	Tinggi	Berat	Rambut	Ras	Kulit	Agama	Status	Merokok
Data 9	1.000	0.000	0.667	0.333	0.000	0.000	0.333	0.000	0.000	0.000
Total	0.500	0.000	0.500	0.167	0.083	0.100	0.417	0.000	0.000	0.000

Table 3.12 Keanggotaan Cluster 3

	JK	Umur	Tinggi	Berat	Rambut	Ras	Kulit	Agama	Status	Merokok
Data 8	0.000	0.000	0.333	0.333	0.667	0.000	0.667	0.000	0.000	0.000
Total	0.000	0.000	0.333	0.333	0.667	0.000	0.667	0.000	0.000	0.000

3.2.3.4 Manualisasi Perhitungan K-Means

Manualisasi *K-Means* adalah tahapan ahir pada proses perhitungan dimana digunakan untuk menentukan *cluster* yang tepat. Pada tahapan ini adalah tahapan selanjutnya setelah proses perhitungan yang telah dilakukan pada tahapan Manualisasi *Improve K-Means inisialisasi cluster awal*. adapun tahapan pada proses ini adalah sebagai berikut:

1. Mengkalkulasikan nilai tengah dari *cluster* untuk mendapatkan pusat klaster yang baru. Pada tahapan ini akan terus berulang sampai tidak ada data yang berubah *clusternya*. Adapun contoh salah satu perhitungan adalah sebagai berikut:

Data yang digunakan: Total dari *cluster* 1

	JK	Umur	Tinggi	Berat	Rambut	Ras	Kulit	Agama	Status	Merokok
Total	0.571	0.095	0.619	0.524	0.571	0.129	0.286	0.143	0.000	0.286

Data 1 (Tabel 3. 1 Hasil proses normalisasi)

Data normalisasi	JK	Umur	Tinggi	Berat	Rambut	Ras	Kulit	Agama	Status	Merokok
Data 1	0.000	0.000	0.333	0.000	0.333	0.000	0.333	0.000	0.000	0.000

Perhitungan

$$d(x_i, x_j) = (|x_{i1} - x_{j1}|^2 + |x_{i2} - x_{j2}|^2 + \dots + |x_{ip} - x_{jp}|^2)^{1/2}$$

$$d(\text{Data 1, Total}) = (|0,000-0.571|^2 + |0,000-0.095|^2 + |0,333-0.333|^2 + |0,000-0.619|^2 + |0,333-0.524|^2 + |0,000-0.571|^2 + |0,333-0.129|^2 + |0,000-0.143|^2 + |0,000-0,000|^2 + |0,000-0.286|^2)^{1/2} = 0.932$$

Adapun hasil perhitungan pada iterasi 1 secara keseluruhan ditunjukkan pada tabel 3.13.

Table 3.13 Perhitungan kalkulasi terhadap total jarak cluster iterasi 1

	Jarak C1	Jarak C2	Jarak C3	Cluster	Cluster 1	Cluster 2	Cluster 3
Data 1	0.932	0.620	0.577	C3	0	0	1
Data 2	1.132	0.620	0.745	C2	0	1	0
Data 3	0.764	0.814	0.333	C3	0	0	1
Data 4	1.060	0.941	0.577	C3	0	0	1
Data 5	0.858	0.640	1.352	C2	0	1	0
Data 6	0.941	0.743	0.816	C2	0	1	0
Data 7	0.601	0.814	1.106	C1	1	0	0

	Jarak C1	Jarak C2	Jarak C3	Cluster	Cluster 1	Cluster 2	Cluster 3
Data 8	0.852	0.848	0.000	C3	0	0	1
Data 9	0.824	0.574	1.291	C2	0	1	0
Data 10	0.699	1.052	1.291	C1	1	0	0
Data 11	0.982	1.489	1.667	C1	1	0	0
Data 12	1.595	2.112	2.292	C1	1	0	0

Adapun contoh salah satu perhitungan dari penentuan keanggotaan klaster adalah sebagai berikut:
Data yang digunakan. Hasil Normalisasi Data 1

Data Normalisasi	JK	Umur	Tinggi	Berat	Rambut	Ras	Kulit	Agama	Status	Merokok
Data 1	0.000	0.000	0.333	0.000	0.333	0.000	0.333	0.000	0.000	0.000

Dan hasil clustering

	Jarak C1	Jarak C2	Jarak C3	Cluster	Cluster 1	Cluster 2	Cluster 3
Data 1	0.932	0.620	0.577	C3	0	0	1

Maka

	JK	Umur	Tinggi	Berat	Rambut	Ras	Kulit	Agama	Status	Merokok
Data 1	0.000 x 1	0.000 x 1	0.333 x 1	0.000 x 1	0.333 x 1	0.000 x 1	0.333 x 1	0.000 x 1	0.000 x 1	0.000 x 1

Hasil akan ditampilkan pada tabel 3.16 keanggotaan klaster 3

Adapun hasil pengumpulan data kedalam *cluster* baru ditunjukkan pada table 3.16:

Table 3.14 Keanggotaan Cluster 1

	JK	Umur	Tinggi	Berat	Rambut	Ras	Kulit	Agama	Status	Merokok
Data 7	1.000	0.000	0.667	0.333	0.667	0.000	0.333	0.000	0.000	0.000
Data 10	1.000	0.000	1.000	0.667	0.667	0.000	0.333	0.000	0.000	0.000
Data 11	1.000	0.333	1.000	0.667	0.667	0.000	0.333	0.000	0.000	1.000
Data 12	1.000	0.333	1.000	1.000	0.667	0.900	0.000	1.000	0.000	1.000
TOTAL	1.000	0.167	0.917	0.667	0.667	0.225	0.250	0.250	0.000	0.500

Perhitungan total = $(1.000 + 1.000 + 1.000 + 1.000 + 1.000)/\text{jumlah data dalam cluster}$

= $(1.000 + 1.000 + 1.000 + 1.000 + 1.000)/4 = \mathbf{1.000}$ (ditunjukkan pada table 3.14 keanggotaan klaster 1)

Table 3.15 Keanggotaan Cluster 2

	JK	Umur	Tinggi	Berat	Rambut	Ras	Kulit	Agama	Status	Merokok
Data 2	0.000	0.000	0.333	0.000	0.000	0.000	0.667	0.000	0.000	0.000
Data 5	1.000	0.000	0.667	0.333	0.000	0.400	0.333	0.000	0.000	0.000
Data 6	0.000	0.000	0.333	0.667	0.000	0.000	0.333	0.000	0.000	0.000
Data 9	1.000	0.000	0.667	0.333	0.000	0.000	0.333	0.000	0.000	0.000
TOTAL	0.500	0.000	0.500	0.333	0.000	0.100	0.417	0.000	0.000	0.000

Table 3.16 Keanggotaan Cluster 3

	JK	Umur	Tinggi	Berat	Rambut	Ras	Kulit	Agama	Status	Merokok
Data 1	0.000	0.000	0.333	0.000	0.333	0.000	0.333	0.000	0.000	0.000
Data 3	0.000	0.000	0.333	0.333	0.667	0.000	0.333	0.000	0.000	0.000
Data 4	0.000	0.000	0.000	0.000	0.667	0.000	0.333	0.000	0.000	0.000
Data 8	0.000	0.000	0.333	0.333	0.667	0.000	0.667	0.000	0.000	0.000
TOTAL	0.000	0.000	0.250	0.167	0.583	0.000	0.417	0.000	0.000	0.000

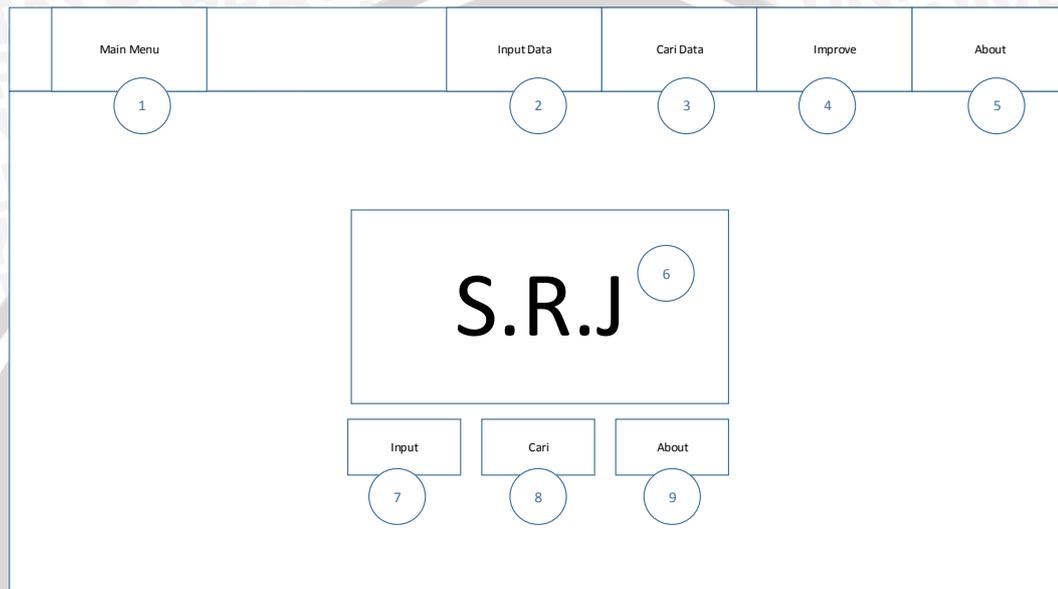
Lakukan proses iterasi sampai tidak ada perubahan kluster pada semua data. Pada perhitungan ini, iterasi berhenti pada iterasi ke 5 dikarenakan hasil klusterisasi pada iterasi 4 tidak berubah pada iterasi ke 5. Adapun perbandingan *cluster* ditunjukkan pada tabel 3.17 Perbandingan klusterisasi awal dan klusterisasi ahir.

Table 3.17 Perbandingan klusterisasi awal dan klusterisasi ahir.

Iterasi 4				Iterasi 5			
Cluster	Cluster 1	Cluster 2	Cluster 3	Cluster	Cluster 1	Cluster 2	Cluster 3
C3	0	0	1	C3	0	0	1
C3	0	0	1	C3	0	0	1
C3	0	0	1	C3	0	0	1
C3	0	0	1	C3	0	0	1
C2	0	1	0	C2	0	1	0
C3	0	0	1	C3	0	0	1
C2	0	1	0	C2	0	1	0
C3	0	0	1	C3	0	0	1
C2	0	1	0	C2	0	1	0
C2	0	1	0	C2	0	1	0
C1	1	0	0	C1	1	0	0
C1	1	0	0	C1	1	0	0

3.2.4 Perancangan *User Interface*

Pada sub bab perancangan *user interface* pada implementasi penentuan rekomendasi jodoh menggunakan algoritma *improve K-Means* yang terdiri dari input data, pencarian data, hasil klustering, dan *about*. Adapun tampilan halaman utama ditunjukkan pada gambar 3.8 interface halaman utama.



Gambar 3.8 Interface Halaman Utama

Keterangan:

1. Main Menu adalah menu untuk menuju halaman utama.
2. Input Data adalah menu untuk menambahkan data baru kedalam database.
3. Cari Data adalah menu untuk mencari kriteria pasangan yang kita inginkan.
4. Improve adalah menu perhitungan dan mengatur jumlah kluster yang akan digunakan.
5. About adalah menu untuk melihat informasi tentang aplikasi.
6. Adalah logo dari aplikasi.
7. Input adalah *shortcut* untuk menambahkan data baru kedalam database.
8. Cari adalah *shortcut* untuk mencari kriteria pasangan yang kita inginkan.
9. About adalah *shortcut* untuk melihat informasi tentang aplikasi.

3.2.4.1 Interface Input Data

Pada halaman pertama akan menginputkan data baru yang nantinya akan di proses dan dilakukan klasterisasi dengan menggunakan algoritma *Improve K-Means*. *Interface input data* ditunjukkan seperti pada Gambar 3.9 Interface Input Data.

Gambar 3.9 Interface Input Data

Keterangan:

1. Main Menu adalah menu untuk menuju halaman utama
2. Input Data adalah menu untuk menambahkan data baru kedalam database
3. Cari Data adalah menu untuk mencari kriteria pasangan yang kita inginkan
4. Improve adalah menu perhitungan dan mengatur jumlah klaster yang akan digunakan.
5. About adalah menu untuk melihat informasi tentang aplikasi.
6. Kriteria data yang wajib diisikan.
7. Tombol untuk menginputkan data yang sudah diisikan sebelumnya.

Halaman ini berfungsi sebagai tempat menginputkan data baru kedalam database dimana data tersebut nantinya akan diproses bersama data yang sudah ada dalam aplikasi rekomendasi jodoh.

3.2.4.2 Interface Pencarian Data

Adalah halaman dimana tempat menginputkan kriteria kriteria yang user inginkan. *Interface input data* ditunjukkan seperti pada gambar 3.10 Interface pencarian data.

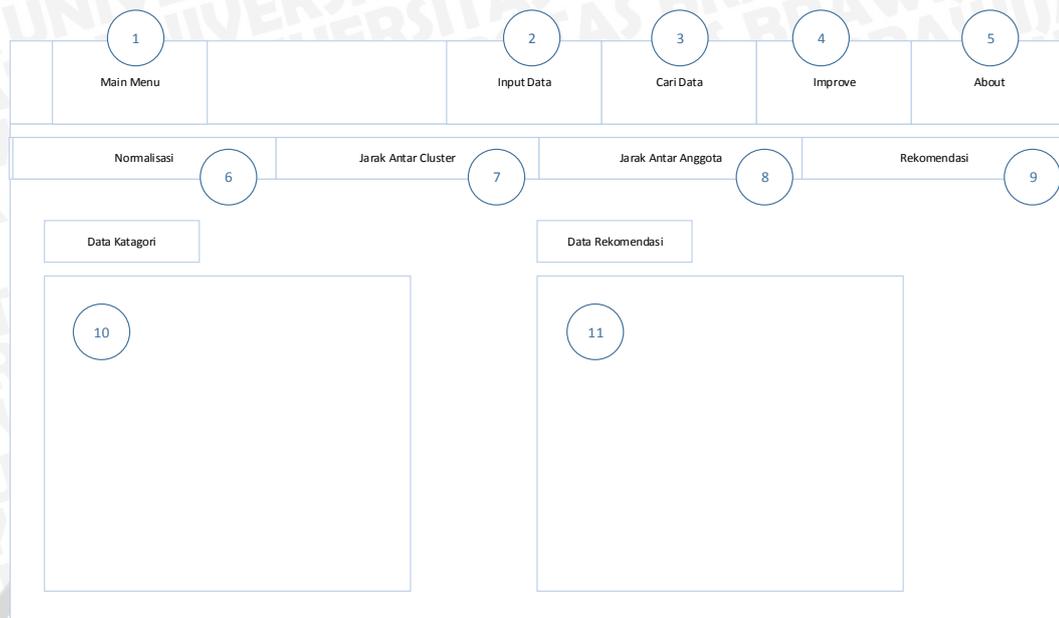
The screenshot shows a web interface titled 'CARI DATA'. At the top, there is a navigation menu with five items: 'Main Menu' (1), 'Input Data' (2), 'Cari Data' (3), 'Improve' (4), and 'About' (5). The 'Cari Data' section contains a search form with two columns of criteria, each with a dropdown menu and a checkmark icon. The criteria are: Jenis Kelamin, Agama, Umur, Tinggi Badan, Berat Badan, and Jenis Rambut on the left; and Ras, Warna Kulit, Status Pernikahan, and Merokok on the right. A large bracket on the right side of the criteria lists is labeled with the number 6. At the bottom center, there is a 'Cari' button and a small circle labeled with the number 7.

Gambar 3.10 Interface pencarian data

Keterangan:

1. Main Menu adalah menu untuk menuju halaman utama
2. Input Data adalah menu untuk menambahkan data baru kedalam database
3. Cari Data adalah menu untuk mencari kriteria pasangan yang kita inginkan
4. Improve adalah menu perhitungan dan mengatur jumlah klaster yang akan digunakan.
5. About adalah menu untuk melihat informasi tentang aplikasi.
6. Kriteria data yang wajib diisikan.
7. Tombol untuk memproses kriteria yang telah diisikan.

Halaman ini berfungsi sebagai tempat meninputkan kriteria yang akan di cari berdasarkan data yang ada dimana akan dicocokkan dengan kriteria yang ada di dalam database. Kemudian hasil yang di tampilkan dari proses pencarian ini adalah ditunjukkan pada gambar 3.11 *Interface rekomendasi*



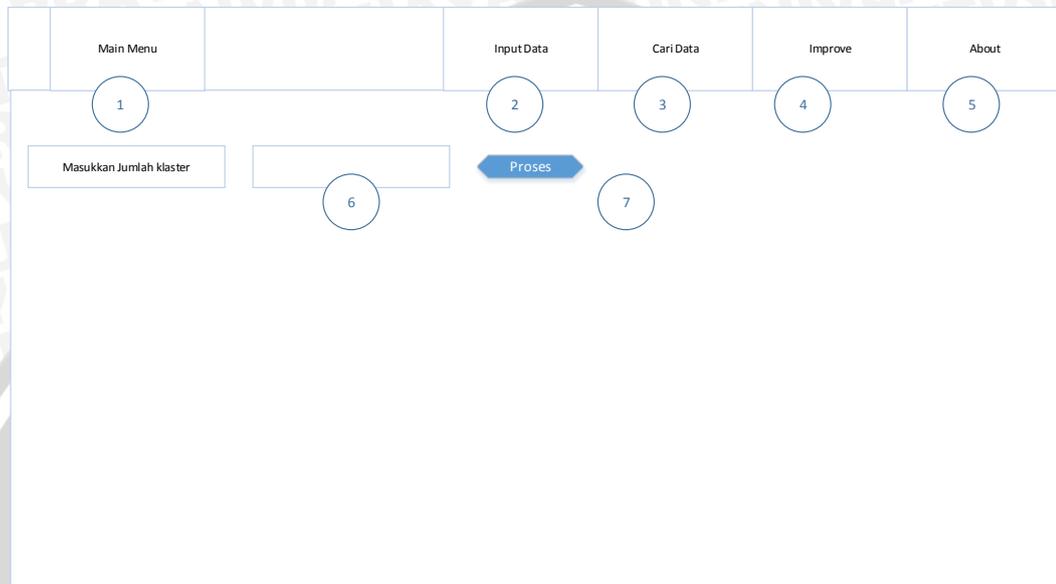
Gambar 3.11 Interface rekomendasi.

Keterangan:

1. Main Menu adalah menu untuk menuju halaman utama
2. Input Data adalah menu untuk menambahkan data baru kedalam database
3. Cari Data adalah menu untuk mencari kriteria pasangan yang kita inginkan
4. Improve adalah menu perhitungan dan mengatur jumlah klaster yang akan digunakan.
5. About adalah menu untuk melihat informasi tentang aplikasi.
6. Tab menampilkan proses normalisasi.
7. Tab Menampilkan perhitungan jarak terhadap klaster.
8. Tab menampilkan perhitungan jarak terhadap data paa klaster.
9. Tab rekomendasi yang sesuai dengan kriteria.
10. Tempat menampilkan kriteria masukan.
11. Tempat menampilkan kriteria rekomendasi.

3.2.4.4 Result Improve

Result Improve berisi berupa proses dari klasterisasi berdasarkan jumlah klaster yang diinginkan. *Result Improve* ditunjukkan seperti pada Gambar 3.12 Result Interface.

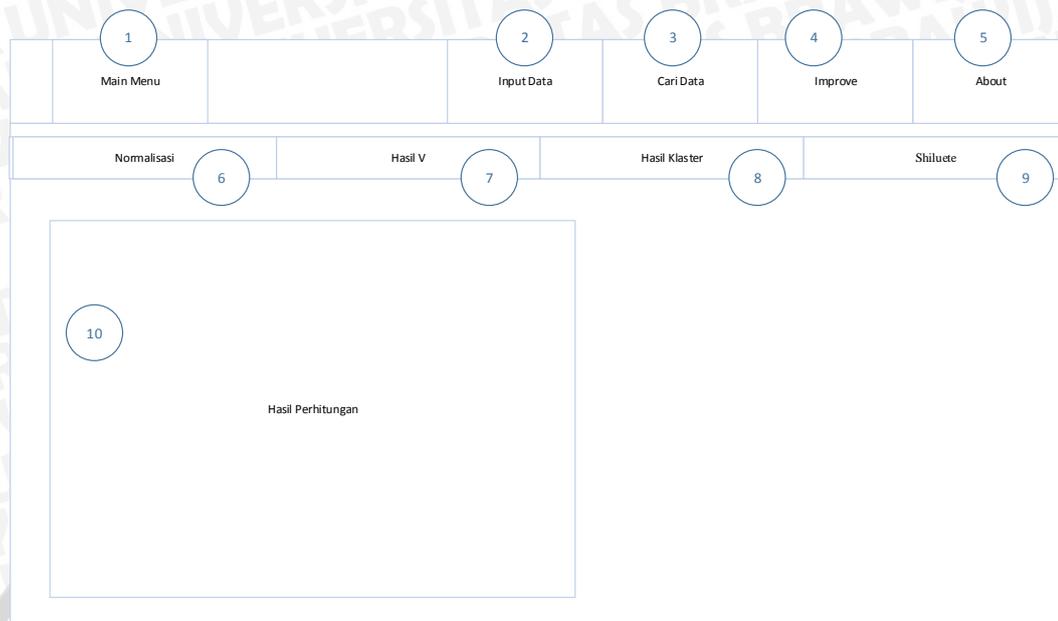


Gambar 3.12 Result Interface.

Keterangan:

1. Main Menu adalah menu untuk menuju halaman utama
2. Input Data adalah menu untuk menambahkan data baru kedalam database
3. Cari Data adalah menu untuk mencari kriteria pasangan yang kita inginkan
4. Improve adalah menu perhitungan dan mengatur jumlah klaster yang akan digunakan.
5. About adalah menu untuk melihat informasi tentang aplikasi.
6. Tempat memasukkan jumlah klaster yang akan digunakan.
7. Memproses masukan jumlah klaster serta melakukan perhitungan *Improve K-Means* dan *K-Means*.

Halaman ini berfungsi sebagai tempat menampilkan data yang sudah dicari pada form sebelumnya dimana data yang ditampilkan sudah melewati proses perhitungan dan klasterisasi. Adapun hasil dari perhitungan metode ditunjukkan pada gambar 3.13 Hasil Perhitungan



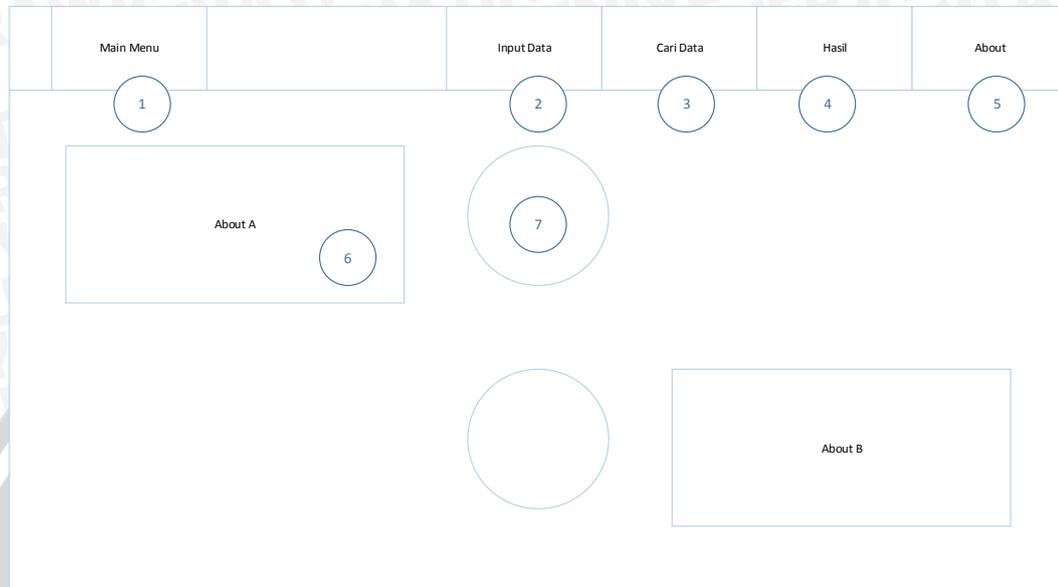
Gambar 3.13 Hasil Perhitungan

Keterangan:

1. Main Menu adalah menu untuk menuju halaman utama
2. Input Data adalah menu untuk menambahkan data baru kedalam database
3. Cari Data adalah menu untuk mencari kriteria pasangan yang kita inginkan
4. Improve adalah menu perhitungan dan mengatur jumlah klaster yang akan digunakan.
5. About adalah menu untuk melihat informasi tentang aplikasi.
6. Tab menampilkan proses normalisasi.
7. Tab menampilkan perhitungan nilai v .
8. Tab menampilkan pusat *centroid* dan jumlah data pada klaster untuk metode *Improve K-Means* dan *K-Means*.
9. Tab menampilkan hasil perhitungan kualitas klaster untuk metode *Improve K-Means* dan *K-Means*.
10. Tempat menampilkan hasil perhitungan.

3.2.4.4 Interface About

Interface about berisi berupa informasi tentang aplikasi. *Interface about* ditunjukkan seperti pada Gambar 3.14 Interface About.



Gambar 3.14 Interface About

Keterangan:

1. Main Menu adalah menu untuk menuju halaman utama
2. Input Data adalah menu untuk menambahkan data baru kedalam database
3. Cari Data adalah menu untuk mencari kriteria pasangan yang kita inginkan
4. Improve adalah menu perhitungan dan mengatur jumlah klaster yang akan digunakan.
5. About adalah menu untuk melihat informasi tentang aplikasi.
6. Tempat menampilkan informasi.
7. Tempat menampilkan gambar.

Halaman ini berfungsi sebagai tempat menampilkan informasi terkait aplikasi dan pembuat serta sebagai sarana panduan aplikasi.

3.2.5 Perancangan Pengujian dan Evaluasi

Pada tahap ini adalah tahapan dimana proses membuat rancangan uji coba dan evaluasi sistem. Pengujian dilakukan dengan tujuan untuk memenuhi tujuan dari penelitian sebagaimana yang telah dijelaskan subbab sebelumnya. Selain itu pengujian ini juga digunakan untuk mengetahui kualitas klaster yang terbentuk.

Kualitas klaster berhubungan erat dengan tingkat validasi data yaitu bagaimana data tersebut dikelompokkan dalam satu klaster. Untuk mengetahui tingkat validasi data digunakan metode *Silhouette Coefficient* seperti yang dijelaskan pada subbab sebelumnya. Adapun analisis yang akan dilakukan antara lain:

1. Pengujian nilai kualitas dari *clustering* dengan menggunakan 10 fitur pada 150 data responden. Pada algoritma *Improve K- Means* pada *inisialisasi centroid awal* yang digunakan dilakukan percobaan menggunakan jumlah 3, 4 5, 6 7, 8, 9 10, 20, 25, 50 dan 100 *cluster* untuk mengetahui berapa jumlah *cluster* terbaik untuk penelitian ini.
2. Pengujian jumlah data ini dilakukan untuk mengetahui seberapa besar efek dari banyak data terhadap hasil dari pengujian *Shiluite index*. Pengujian jumlah data dilakukan dengan menggunakan 5 *cluster* karena pada pengujian sebelumnya memiliki nilai *Shiluite index* yang paling bagus. Jumlah data yang akan di uji pada pengujian ini adalah 10, 30, 50, 80, 110, 130, dan 150 data.
3. Pengujian nilai kualitas dari *clustering* dengan menggunakan 10 fitur pada 150 data responden untuk metode *K-Means* dengan percobaan menggunakan jumlah 3, 4 5, 6 7, 8, 9 10, 20, 25, 50 dan 100 *cluster* untuk mengetahui berapa jumlah *cluster* terbaik untuk penelitian metode *K-Means* dan dibandingkan dengan metode *Improve K- Means*.

Tahapan pengujian kualitas *cluster* yang akan dilakukan yaitu:

1. Labeling dan normalisasi data dalam range yg sama dan hitung jarak tiap data.
2. Tentukan jumlah *cluster* untuk percobaan yang telah dijabarkan diatas.

3. Perhitungan jumlah *cluster* menghasilkan nilai *cluster* yang paling baik dengan perhitungan dengan menggunakan *Shiluite index*.

Tahapan pengujian jumlah data yang akan dilakukan yaitu:

1. Labeling dan normalisasi data dalam range yg sama dan hitung jarak tiap data dengan jumlah data 10, 30, 50, 80, 110, 130, dan 150 data.
2. Menentukan jumlah *cluster* yang digunakan yaitu 5 *cluster*.
3. Perhitungan nilai *Shiluite index* dengan masing masing jumlah data.

Tahapan pengujian kualitas *cluster* K-Means yang akan dilakukan yaitu:

1. Labeling dan normalisasi data dalam range yg sama dan hitung jarak tiap data.
2. Tentukan jumlah *cluster* untuk percobaan yang telah dijabarkan diatas.
3. Perhitungan jumlah *cluster* menghasilkan nilai *cluster* yang paling baik dengan perhitungan dengan menggunakan *Shiluite index*.
4. Membandingkan hasil *Shiluite index* K-Means terhadap hasil *Shiluite index Improve K- Means*.



BAB IV IMPLEMENTASI

Pada bab ini akan dijabarkan tentang implementasi perangkat lunak berdasarkan analisis kebutuhan dan perancangan perangkat lunak yang sebelumnya telah dijabarkan pada Bab III. Pembahasan dalam bab ini terdiri dari penjelasan tentang spesifikasi dan implementasi *interface*.

4.1 Spesifikasi Sistem

Analisis kebutuhan dan perancangan perangkat lunak yang telah dijabarkan pada Bab III akan dijadikan dasar untuk melakukan implementasi agar dapat digunakan dan sesuai dengan kebutuhan. Implementasi dari spesifikasi dari aplikasi yang akan di bangun terdapat pada bagian spesifikasi perangkat keras dan spesifikasi perangkat lunak.

4.1.1 Spesifikasi Perangkat Keras

Implementasi *Sublime Text* digunakan sebagai pendukung pembuatan aplikasi pada komputer. Adapun spesifikasi perangkat keras dari computer yang digunakan akan dijabarkan pada table 4.1. Spesifikasi perangkat keras computer.

Table 4.1 Spesifikasi Perangkat Keras Komputer

Nama Hardware	Spesifikasi
Prosesor	AMD A8-4500M APU with Radeon(tm) HD Graphics (4 CPUs), ~1.90 GHz
Memori (RAM)	8 GB
Hardisk	500 GB
VGA card	AMD Radeon HD 7640G dan AMD Radeon HD 8600/8700M 6GB
Monitor	14"

4.1.2 Spesifikasi Perangkat Lunak

Implementasi *Improve K- Means* pada inisialisasi centroid untuk penentuan rekomendasi jodoh akan menggunakan perangkat lunak yang akan di jabarkan pada tabel 4.2

Table 4.2 Spesifikasi Perangkat Lunak

Nama Hardware	Spesifikasi
Sistem Operasi	Microsoft Windows 8 Pro 64x
Bahasa pemrograman	Web
<i>Tools</i> Pemrograman	Sublime Text 2
<i>Server Localhost</i>	XAMPP Server Version 3.2.1
DBMS	MySQL

4.2 Implementasi Algoritma

Implementasi yang akan dibahas menggunakan bahasa pemrograman Java dan menggunakan DBMS MySQL. Bahasa pemrograman Java sendiri digunakan untuk memproses algoritma. Database MySQL digunakan untuk menyimpan data. Implementasi algoritma ini mengacu pada bab perancangan.

Implementasi yang akan dibahas menggunakan Bahasa pemrograman Web dan menggunakan DBMS MySQL. Bahasa pemrograman web sendiri digunakan untuk memproses dan mengimplementasikan algoritma pada penelitian ini. Kemudian Database MySQL digunakan untuk media penyimpanan data. Proses implementasi pada bab ini mengacu pada perancangan yang telah disampaikan pada Bab III sebelumnya.

4.2.1 Normalisasi

Proses normalisasi merupakan algoritma yang digunakan untuk menormalkan seluruh data yang ada agar sesuai dengan perancangan yang telah disusun di awal.

4.2.1.1 Labeling

Implementasi proses labeling pada penelitian ini meliputi semua nilai dari kriteria setiap data. Pada proses ini adalah proses mengubah data katagorikal

menjadi numerik agar mudah dilakukan proses perhitungan klustering. pada saat proses penambahan data baru akan langsung dikonversi kedalam nilai numerik menggunakan “*option value*”. Adapun *source code* pada proses labeling ditunjukkan pada *source code 4.1*

```

1 <td id="tbl-cell-1" align="left" valign="middle" rowspan="2">
2 <table border="1" width="296" align="left"><select name="1" id="tbl-cell-1">
3 <option value="0">Banjar</option>
4 <option value="1">Balikpapan</option>
5 <option value="2">Betawi</option>
6 <option value="3">Bugis</option>
7 <option value="4">Dayak</option>
8 <option value="5">Jawa</option>
9 <option value="6">Melayu</option>
10 <option value="7">Papua</option>
11 <option value="8">Sunda</option>
12 <option value="9">Tionghoa</option>
13 <option value="10">Lainnya</option>
14 </select></td>
15 </tr>
16 <tr>
17 <td border="1" height="36" align="center" valign="middle">Jenis
Kelamin</td>
18 <td border="1" width="144" height="36" align="left"
19 valign="middle"><select name="7" id="tbl-cell-2">
20 <option value="1">Laki - Laki</option>
21 <option value="0">Perempuan </option>
22 </select></td>
23 <td border="1" width="162" align="left" valign="middle">Warna
24 <table border="1" width="296" align="left" id="tbl-cell-3"><select name="8"
25 id="tbl-cell-3">
26 <option value="0">Biru</option>
27 <option value="1">Kuning</option>
28 <option value="2">Merah</option>
29 <option value="3">Hijau</option>
30 <option value="4">Coklat</option>
31 <option value="5">Lainnya</option>
32 </select></td>
33 </tr>
34 </table>

```




```

53 <tbody>
54 <tr>
55 <td width="144" height="30" align="left" valign="middle">Tinggi
56 Badan</td>
57 <td width="144" height="30" align="left"
58 valign="middle"><select name="tinggi" id="tinggi">
59 <option value="0">&lt;155 Cm</option>
60 <option value="1">&lt;165 Cm</option>
61 <option value="2">&lt;175 Cm</option>
62 <option value="3">&lt;185 Cm</option>
63 </select></td>
64 <td width="162" align="left" valign="middle">Status
65 Pernikahan</td>
66 <td width="246" align="left"><select name="status"
67 id="status">
68 <option value="0">Lajang</option>
69 <option value="1">Cesaka</option>
70 <option value="2">Ditinggal</option>
71 <option value="3">Hidup</option>
72 </select></td>
73 </tr>
74 <tr>
75 <td height="30" align="left" valign="middle">Berat
76 Badan</td>
77 <td width="144" height="30" align="left"
78 valign="middle"><select name="berat" id="berat">
79 <option value="10">&lt;50 kg</option>
80 <option value="11">&lt;55 kg</option>
81 <option value="12">&lt;60 kg</option>
82 <option value="13">&lt;65 kg</option>
83 </select></td>
84 </tr>
85 </tbody>
86 </table>

```



4.2.1.2 Min Max

Implementasi proses min max pada penelitian ini akan meliputi nilai data minimal, nilai daya maksimal, nilai data baru minimal, dan nilai data baru maksimum. Pada proses ini adalah proses menormalisasikan data kedalam range yang lebih spesifik sehingga lebih mudah dilakukan perhitungan. Adapun *source code* pada proses Min Max ditunjukkan oleh *sourcecode 4.2*

```

1 $data = array();
2 $i=1;
3 while($proses = mysql_fetch_array($query))
4     $data[$i][1]= $proses['jk'];
5     $data[$i][2]= $proses['agama'];
6     $data[$i][3]= $proses['umur'];
7     $data[$i][4]= $proses['tinggi'];
8     $data[$i][5]= $proses['berat'];
9     $data[$i][6]= $proses['rambut'];
10    $data[$i][7]= $proses['resol'];
11    $data[$i][8]= $proses['adanya'];
12    $data[$i][9]= $proses['status'];
13    $data[$i][10]= $proses['rokok'];
14
15    $i++;
16 }
17 $norm=array();
18 for($a=1;$a<=91;$a++)
19     $norm[$a][1] = (($data[$a][1]-0)/(3-0)*((1-0)+0));
20     $norm[$a][2] = (($data[$a][2]-0)/(3-0)*((1-0)+0));
21     $norm[$a][3] = (($data[$a][3]-0)/(3-0)*((1-0)+0));
22     $norm[$a][4] = (($data[$a][4]-0)/(3-0)*((1-0)+0));
23     $norm[$a][5] = (($data[$a][5]-0)/(3-0)*((1-0)+0));
24     $norm[$a][6] = (($data[$a][6]-0)/(3-0)*((1-0)+0));
25     $norm[$a][7] = (($data[$a][7]-0)/(10-0)*((1-0)+0));
26     $norm[$a][8] = (($data[$a][8]-0)/(3-0)*((1-0)+0));
27     $norm[$a][9] = (($data[$a][9]-0)/(3-0)*((1-0)+0));
28     $norm[$a][10] = (($data[$a][10]-0)/(3-0)*((1-0)+0));
29 }

```

Sourcecode 4.2 Min Max

Keterangan :

1. Pada baris 1 sampai baris 15 merupakan proses pendeklarasian seluruh kriteria dari data yang di gunakan.
2. Pada baris 17 sampai 28 metupakan proses normalisasi seluruh data yang di ambil dari database berdasarkan kriteria.

4.2.2 Improve K-Means

Algoritma *improve K-Means* merupakan algoritma yang digunakan untuk menentukan centroid awal dari seluruh data yang ada sesuai dengan perancangan yang telah disusun di awal.

4.2.2.1 Euclidean Distance

Algoritma *Euclidean Distance* merupakan algoritma yang digunakan untuk menghitung jarak antar data. *Euclidean Distance* adalah tahapan pertama dalam algoritma *improve K-Means* yang menggunakan seluruh nilai dari kriteria dari masing masing data dan mencari jarak antar data tersebut. Adapun *Sourcecode* yang digunakan ditunjukkan pada *Sourcecode 4.3*

```

1  $totaldistance = array();
2  $distance = array(array());
3  for($b=1;$b<=$jumlah_data;$b++){
4      $sum=0;
5      for($c=1;$c<=$jumlah_data;$c++){
6          $distance[$b][$c]=0;
7          for($d=1;$d<=10;$d++){
8              $distance[$b][$c] = ($distance[$b][$c] +
9                  pow(($norm[$b][$d] - $norm[$c][$d]),2));
10             }
11             $distance[$b][$c] = sqrt($distance[$b][$c]);
12             $sum += $distance[$b][$c];
13             echo "<br> user $b dibandingkan dengan user $c
14             dengan jarak $distance[$b][$c]";
15         }
16     }
17     $totaldistance[$b] = $sum;
18 }

```

Sourcecode 4.3 Euclidean Distance

Keterangan:

1. Pada baris 1, 2, 4, dan 6 merupakan proses inialisasi variable.
2. Pada baris 8 sampai baris 10 merupakan proses perhitungan *Euclidean Distance*.
3. Pada baris 11 adalah proses penjumlahan *Euclidean Distance*.

4.2.2.2 Menghitung Nilai V_j

Menghitung nilai V_j merupakan tahapan selanjutnya setelah perhitungan *Euclidean Distance*. Pada tahapan ini yaitu membagi hasil *Euclidean Distance* data i dengan total data i . Adapun *Sourcecode* yang digunakan ditunjukkan pada *Sourcecode 4.4*

```

1 $v = array();
2 for ($b=1;$b<=$jum_data;$b++)
3     $vttotal = 0;
4     for($c=1;$c<=$jum_data;$c++)
5         $vttotal = $vttotal
6         ($distance[$b][$c]/$totaldistance*$b)
7     }
8     $v[$b] = $vttotal;
9     $q=mysql_query("update response set nilai_v =
10 $vttotal where id = '$b'");
11     echo "<br> hasil total nilai v = " . $v[$b];
12 }
```

Sourcecode 4.4 Menghitung Nilai V_j

Keterangan:

1. Pada baris 1 dan 3 merupakan proses inialisasi variable.
2. Pada baris 5 sampai baris 7 merupakan proses perhitungan nilai V_j .
3. Pada baris 8 adalah proses memasukkan hasil V_j kedalam database.

4.2.2.3 Mengurutkan Nilai V_j dan Menentukan k Awal

Pada tahapan ini yaitu mengurutkan nilai V_j berdasarkan nilai yang paling kecil dan kemudian ditentukan pula jumlah klaster yang akan digunakan. Beberapa

4.2.3 Menghitung Ulang Centroid

Proses menghitung ulang *centroid* merupakan tahapan dimana perhitungan ulang jarak tiap data terhadap rata-rata jarak pada tiap *cluster*. Pada perhitungan ini menggunakan algoritma *Euclidean Distance* untuk mendapatkan hasil perhitungan.

```

1  $iterasi = 1;
2  $flag = 1;
3  $pusatCluster = array();
4  $cluster = tentukancluster($k, $pusatCluster, $norm, $flag, $cluster);
5  $jumlahcluster = jumlahnggakaluster($k, $cluster);
6  $pusatCluster = newcentroid($k, $cluster, $norm, $jumlahcluster);
7  $sama = $flag;
8  $iterasi++;
9  while ( ($iterasi <= 10) && ($sama != 0) )
10 function tentukancluster ($k, $pusatCluster, $norm, $cluster) {
11     $distcluster = array();
12     $clusterlama = array();
13     $clusterlama = $cluster;
14     global $flag;
15     $flag = 1;
16     for ($d=1; $d<=$n_data; $d++) {
17         $cluster[$d]=1;
18         $temp = jarak($pusatCluster[1], $norm[$d]);
19         for ($c=1; $c<=$k; $c++) {
20             $distcluster[$d][$c] =
21                 jarak($pusatCluster[$c], $norm[$d]);
22             if ($distcluster[$d][$c] < $temp) {
23                 $temp = $distcluster[$d][$c];
24                 $cluster[$d] = $c;
25             }
26         }
27     }
28     $sama = $n_data - $iterasi;
29     $flag = 0;

```

```

1  cluster = 0;
2  flag = 0;
3  = set;
4  = set;
5  flag = $flag & 0;
6  }
7  return $cluster;
8
9  function jarak($dstCluster,$norm) {
10     $dist = 0;
11     for ($b=1;$b<=10;$b++) {
12         $dist += pow($dstCluster[$b] - $norm[$b],2);
13     }
14     $dist = sqrt($dist);
15     return $dist;
16 }
17
18 function jumlahAngkaCluster($k,$cluster) {
19     $jumlahcluster = array();
20     for ($c=1;$c<=$k;$c++) {
21         $jumlahcluster[$c] = 0;
22     }
23
24     for ($d=1;$d<=count($cluster);$d++) {
25         for ($e=1;$e<=$k;$e++) {
26             if ($cluster[$d] == $e) {
27                 $jumlahcluster[$e] =
28                 $jumlahcluster[$e]+1;
29             }
30         }
31     }
32
33     return $jumlahcluster;
34 }

```

```

1  for ($c=1; $c<=$k; $c++){
2  for ($d=1; $d<=$n; $d++){
3  for ($kri=1; $kri<=10; $kri++){
4  $centeroidbaru[$c][$kri] = 0;
5
6
7  for ($c=1; $c<=$k; $c++){
8  for ($d=1; $d<=$n; $d++){
9  if ($cluster[$d] == $c){
10     for ($kri=1; $kri<=10; $kri++){
11         $centeroidbaru[$c][$kri] +=
12             $sum[$d][$kri];
13     }
14 }
15 }
16 }
17 for ($c=1; $c<=$k; $c++){
18     for ($kri=1; $kri<=10; $kri++){
19         if ($jumlahcluster[$c] > 0){
20             $centeroidbaru[$c][$kri] =
21                 $centeroidbaru[$c][$kri] / $jumlahcluster[$c] ;
22         }
23         else {
24             $centeroidbaru[$c][$kri] =
25                 $centeroidbaru[$c][$kri] + 1;
26         }
27     }
28 }
29 }
30 return $centeroidbaru;

```

Sourcecode 4.6 Menghitung Ulang Centroid

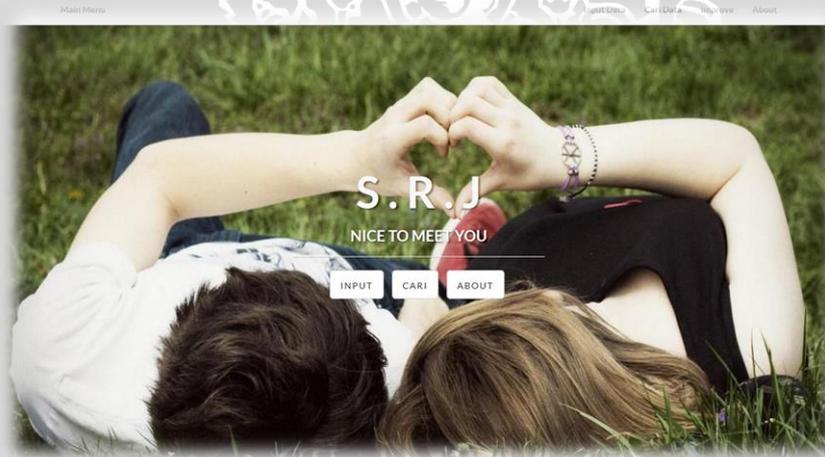
Keterangan:

1. Pada baris 1, 2, 7, 8, 11, 12, 13, 14, dan 15 merupakan proses inisialisasi variabel.

2. Pada baris 16 sampai baris 25 merupakan proses perhitungan ulang centroid dengan iterasinya.
3. Pada baris 27 sampai 35 merupakan proses perhitungan dimana apabila terdapat data yang memiliki nilai V sama maka akan diganti dengan data lain yang memiliki nilai berbeda
4. Pada baris 36 sampai 91 adalah adalah fungsi yang akan di panggil ke dalam perhitungan ulang centroid dengan iterasinya.

4.3 Implementasi *Interface*

Implementasi *interface* diterapkan berdasarkan pada perancangan yang dijelaskan pada Bab III sebelumnya. Implementasi penentuan rekomendasi jodoh ini menggunakan Bahasa pemrograman web. Berikut adalah implementasi antarmuka yang diterapkan untuk dapat berinteraksi dengan *end user* dimana antarmuka halaman utama ditunjukkan pada gambar 4.1 *interface* Halaman Utama.



Gambar 4.1 *interface* Halaman Utama.

4.3.1 Interface *Input Data*

Implementasi halaman *input data* adalah antarmuka yang bertujuan untuk menambahkan data yang akan dikelompokkan pada aplikasi ini. Adapun antarmuka pada input data ditunjukkan pada gambar 4.2 *Interface Input Data*.

Nama	<input type="text"/>	Ras	Banjar
Jenis Kelamin	Laki - Laki	Warna Kulit	Putih
Agama	Islam	Pendidikan Terakhir	SMA Sederajat
Umur	20 - 25 Tahun	Pekerjaan	Pegawai Negeri Sipil
Tinggi Badan	<155 Cm	Status Pernikahan	Lajang
Berat Badan	<50 Kg	Merokok	Tidak
Jenis Rambut	Ikal	Pandangan Poligami	Tidak Setuju

Input

Gambar 4.2 Interface Input Data.

4.3.2 Interface Pencarian Data

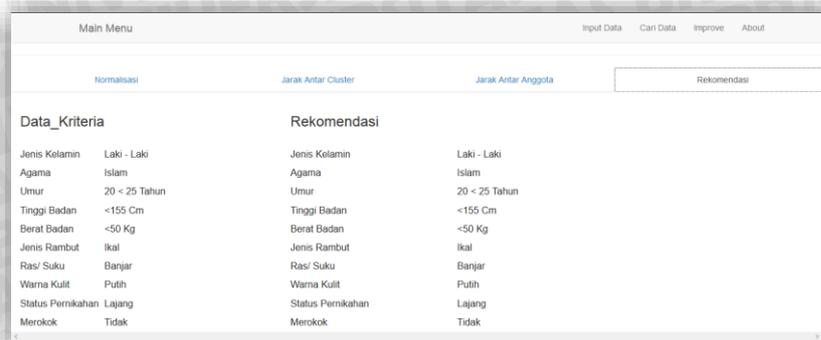
Implementasi halaman pencarian data adalah *interface* yang bertujuan untuk mencari kriteria berdasarkan apa yang diinginkan oleh *end user*. Adapun antarmuka pada input data ditunjukkan pada gambar 4.3 *Interface* Pencarian Data.

Jenis Kelamin	Laki - Laki	Ras	Banjar
Agama	Islam	Warna Kulit	Putih
Umur	20 - 25 Tahun	Pendidikan Terakhir	SMA Sederajat
Tinggi Badan	<155 Cm	Pekerjaan	Pegawai Negeri Sipil
Berat Badan	<50 Kg	Status Pernikahan	Lajang
Jenis Rambut	Ikal	Merokok	Tidak

Find

Gambar 4.3 Interface Pencarian Data.

Kemudian apabila kita menekan tombol cari, maka akan dijalankan proses pencarian kriteria sesuai dengan apa yang telah di inputkan. Adapun antarmuka hasil pencarian ditunjukkan pada gambar gambar 4.3 *Interface* Hasil Pencarian.



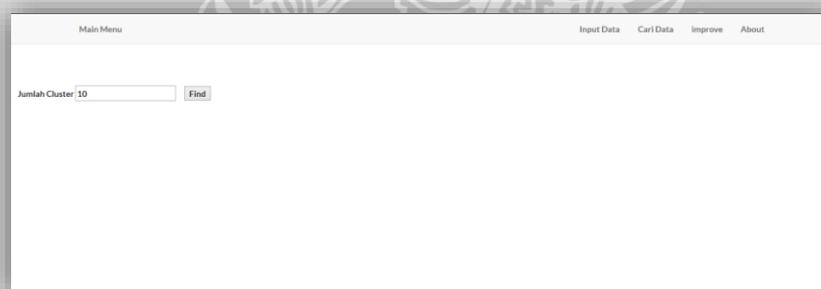
The screenshot shows a web application interface with a navigation bar at the top containing 'Main Menu', 'Input Data', 'Cari Data', 'Improve', and 'About'. Below the navigation bar, there are three tabs: 'Normalisasi', 'Jarak Antar Cluster', and 'Jarak Antar Anggota'. The 'Normalisasi' tab is active, displaying two columns: 'Data_Kriteria' and 'Rekomendasi'. Each column contains a list of criteria and their corresponding values.

Data_Kriteria		Rekomendasi	
Jenis Kelamin	Laki - Laki	Jenis Kelamin	Laki - Laki
Agama	Islam	Agama	Islam
Umur	20 < 25 Tahun	Umur	20 < 25 Tahun
Tinggi Badan	<155 Cm	Tinggi Badan	<155 Cm
Berat Badan	<50 Kg	Berat Badan	<50 Kg
Jenis Rambut	Ikal	Jenis Rambut	Ikal
Ras/ Suku	Banjar	Ras/ Suku	Banjar
Warna Kulit	Putih	Warna Kulit	Putih
Status Pernikahan	Lajang	Status Pernikahan	Lajang
Merokok	Tidak	Merokok	Tidak

Gambar 4.4 Interface Hasil Pencarian

4.3.3 Improve Interface

Implementasi halaman *improve interface* adalah antarmuka yang bertujuan untuk menampilkan hasil klustering yang telah dilakukan. Adapun antarmuka pada *improve interface* ditunjukkan pada gambar 4.5 *Improve Interface*.



The screenshot shows the 'Main Menu' interface with the 'Improve' tab selected. The interface features a text input field labeled 'Jumlah Cluster' with the value '10' entered, and a 'Find' button next to it. The navigation bar at the top includes 'Main Menu', 'Input Data', 'Cari Data', 'Improve', and 'About'.

Gambar 4.5 Interface Improve.

Kemudian apabila nilai klaster sudah ditentukan dan dilakukan proses maka akan ditampilkan hasil klustering sesuai dengan jumlah klaster yang di inputkan. Adapun antarmuka dari tampilnya hasil tersebut ditunjukkan pada gambar 4.6 Hasil *Improve*.

Cluster	ID	Nilai V
Cluster 1	20	0.81004170313028
Cluster 2	1040	0.83621896080205
Cluster 3	17	0.86182175368082
Cluster 4	31	0.86324049254899
Cluster 5	1120	0.8408924570972
Cluster 6	22	0.87087794669126
Cluster 7	6	0.8720561802894
Cluster 8	1380	0.87322648971213
Cluster 9	1170	0.87581923214986
Cluster 10	82	0.8786993897343

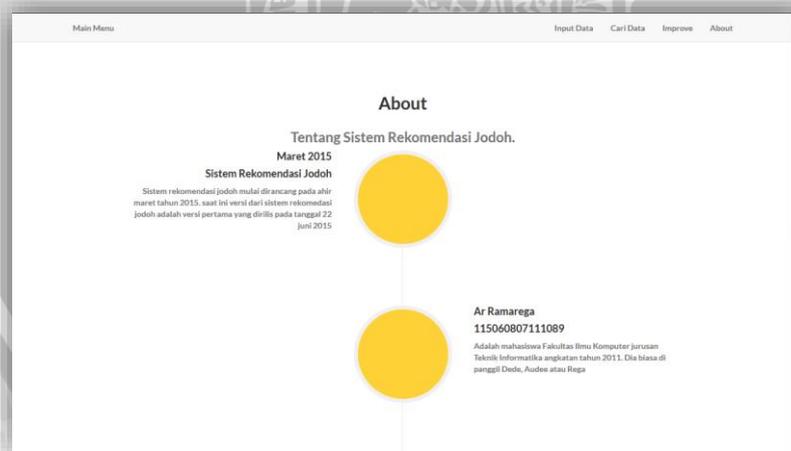
Cluster	Jumlah
Jumlah Cluster 1	14
Jumlah Cluster 2	26
Jumlah Cluster 3	15
Jumlah Cluster 4	9
Jumlah Cluster 5	9
Jumlah Cluster 6	25
Jumlah Cluster 7	29
Jumlah Cluster 8	8
Jumlah Cluster 9	14
Jumlah Cluster 10	8

Jumlah Iterasi = 8

Gambar 4.6 Hasil Improve.

4.3.4 Interface About

Implementasi halaman *about* adalah *interface* yang bertujuan untuk menampilkan segala sesuatu tentang aplikasi ini. Pada halaman *about* akan menampilkan tentang versi rilis, *copyright*, dan tentang pembuat. Adapun antarmuka pada *about* ditunjukkan pada gambar 4.7 *Interface About*



Gambar 4.7 Interface About

BAB V

PENGUJIAN DAN ANALISIS

Pada bab ini akan membahas mengenai proses pengujian dari implementasi metode *Improve K- Means pada inisialisasi centroid* pada penentuan rekomendasi jodoh. Proses pengujian akan dilakukan dengan 3 skenario, yaitu pengujian 150 data dengan menggunakan percobaan 3 sampai 100 *cluster*, pengujian kedua sebanyak 5 *cluster* dengan menggunakan 10, 30, 50, 80, 110, 130, dan 150 data dan pengujian ketiga yaitu percobaan 3 sampai 100 *cluster* pada metode K-Means kemudian dibandingkan dengan hasil *Silhouette* pada pengujian *Improve K-Means*.

5.1 Pengujian Cluster

Pengujian cluster ini dilakukan untuk mengetahui berapa jumlah cluster terbaik yang menghasilkan nilai *Silhouette index* tertinggi untuk digunakan pada rekomendasi jodoh ini. Nilai kualitas cluster memiliki rentang dari -1 sampai 1

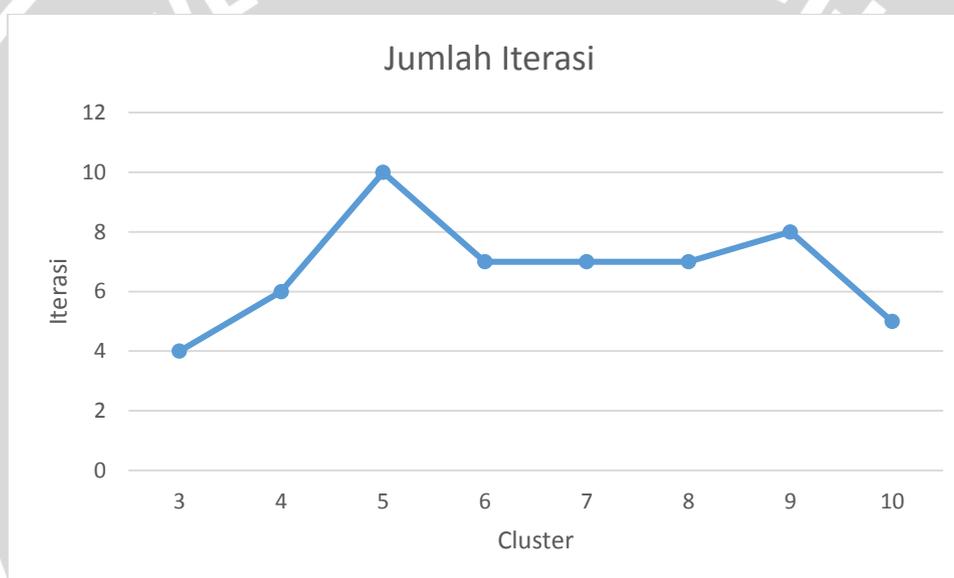
Pengujian jumlah cluster dilakukan dengan menggunakan 150 data responden dengan 10 kriteria. Jumlah cluster yang digunakan pada pengujian ini adalah 3, 4, 5, 6, 7, 8, 9, 10, 20, 25, 50, dan 100. Adapun hasil pengujian ditunjukkan pada tabel 5.1 pengujian jumlah *cluster*

Table 5.1 Pengujian Jumlah Cluster.

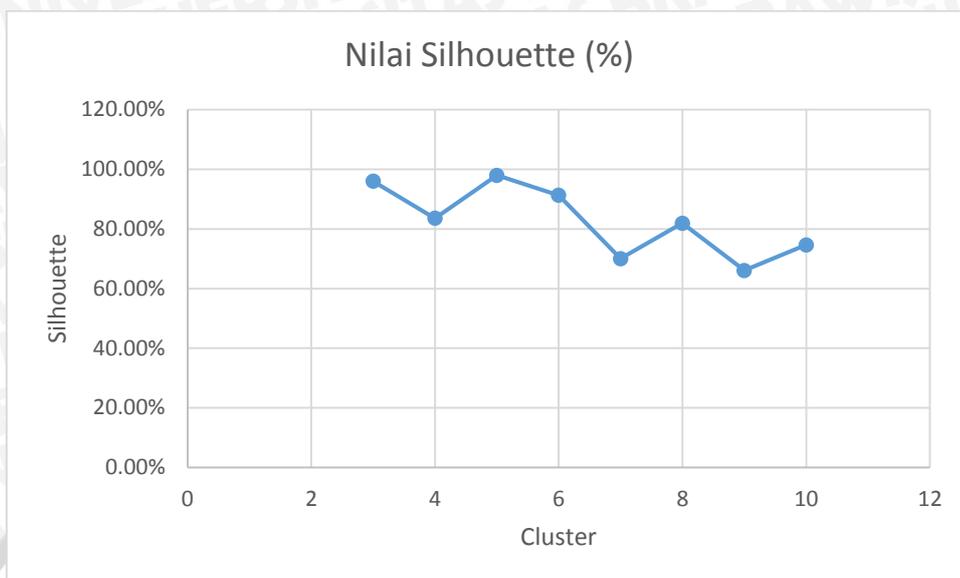
Cluster	Iterasi	Nilai Silhouette	
		Rentang -1 Sampai 1	Dalam % $((x+1)/2 \times 100)$
3	4	0.920255471	96.01%
4	6	0.671534537	83.58%
5	10	0.95959899	97.98%
6	7	0.826251786	91.31%
7	7	0.401467064	70.07%
8	7	0.639059229	81.95%
9	8	0.321679639	66.08%
10	5	0.492930654	74.65%

Cluster	Iterasi	Nilai Silhouette	
		Rentang -1 Sampai 1	Dalam % $((x+1)/2 \times 100)$
20	9	-0.982587751	49.13%
25	5	-0.946666667	47.33%
50	4	-0.593333333	29.67%
100	4	-0.1	5.00%

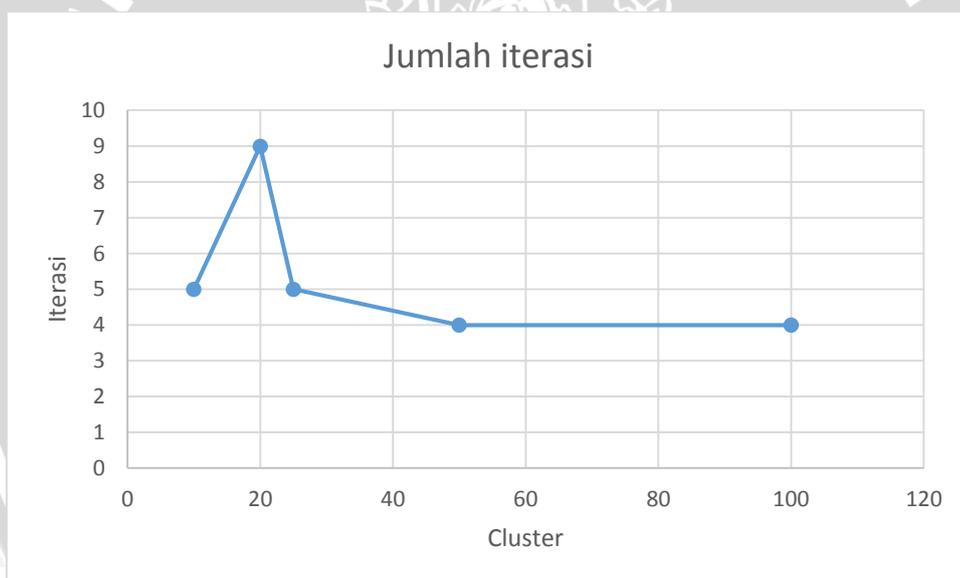
Pada tabel 5.1 di atas dapat dilihat bahwa hasil pengujian yang dilakukan dengan metode *silhouette index* menghasilkan nilai tertinggi pada penggunaan 5 *cluster*. Semakin tinggi jumlah *cluster* semakin rendah nilai *silhouette* yang di dapat. Adapun hasil dari pengujian dalam bentuk statistik ditunjukkan pada gambar 5.1 sampai 5.4:



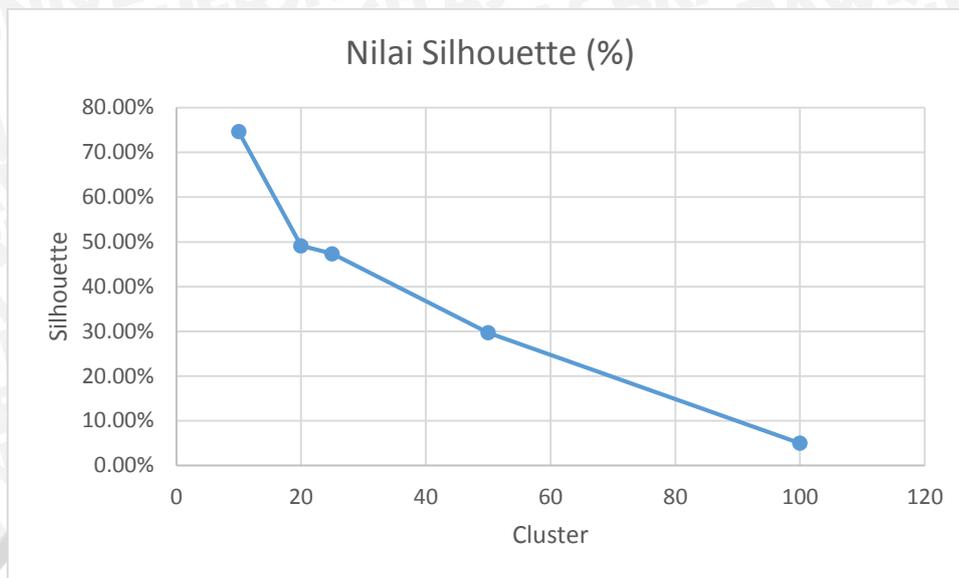
Gambar 5.1 Statistik Jumlah Iterasi Cluster 3 sampai 10



Gambar 5.2 Statistik Nilai Silhouette Cluster 3 sampai 10



Gambar 5.3 Statistik Jumlah Iterasi Cluster 10 sampai 100



Gambar 5.4 Statistik Nilai Silhouette Cluster 10 sampai 100

Berdasarkan grafik hasil pengujian jumlah *cluster* yang ditunjukkan pada gambar 5.1 sampai gambar 5.4 dapat dilihat hasil bahwa pengujian dengan menggunakan jumlah *cluster* = 5 memiliki nilai *silhouette* yang paling tinggi hal tersebut terjadi karena pada penggunaan 5 klaster jarak antar data pada klaster memiliki kerapatan yang cukup baik dan jarak antar klaster memiliki nilai yang cukup besar sehingga menghasilkan nilai *silhouette* yang optimal. Kemudian pada penggunaan 10 klaster ke atas terjadi penurunan kualitas yang disebabkan oleh jarak data pada klaster terlalu jauh dan jarak antar klaster cenderung dekat.

Kemudian untuk jumlah iterasi yang terjadi pada proses perhitungan cenderung memiliki jumlah iterasi yang stabil dan tidak ada peningkatan atau penurunan jumlah iterasi yang sangat signifikan. Pada proses iterasi yang terjadi pada penggunaan klaster 3 sampai 10 memiliki nilai iterasi yang sedikit lebih banyak daripada penggunaan 10 klaster ke atas, hal tersebut disebabkan oleh semakin banyaknya jumlah pusat klaster sehingga data yang sangat dekat dengan pusat klaster akan langsung bergabung pada klaster tersebut sehingga menurunkan kemungkinan data yang berpindah pada iterasi selanjutnya semakin kecil. Sedangkan semakin tinggi jumlah *cluster* yang digunakan pada pengujian ini maka perhitungan nilai *silhouette* yang di hasilkan akan semakin rendah.

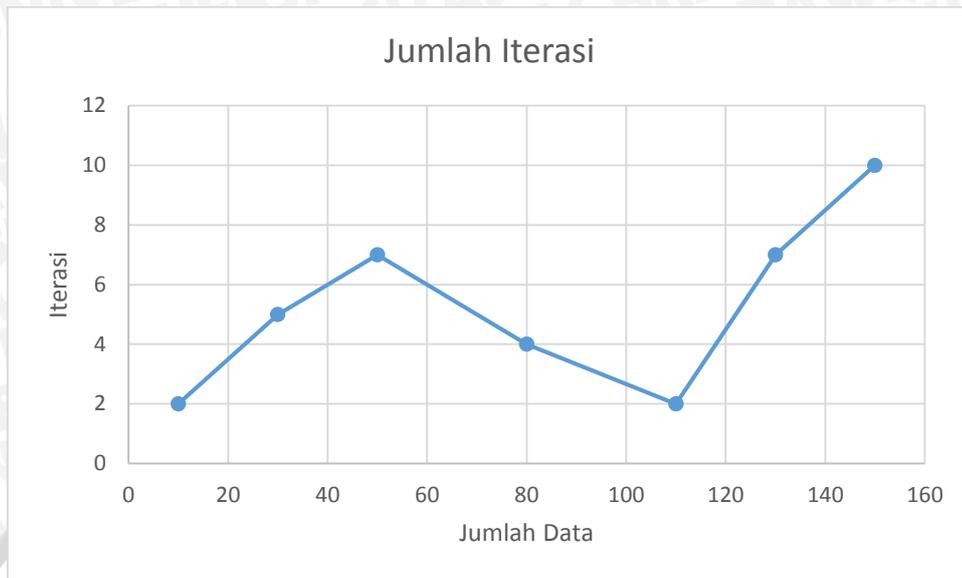
5.2 Pengujian Jumlah Data

Pengujian jumlah data ini dilakukan untuk mengetahui seberapa besar efek dari banyak data terhadap hasil dari pengujian *Shiluite index*. Pengujian jumlah data dilakukan dengan menggunakan 5 *cluster* karena pada pengujian sebelumnya memiliki nilai *Shiluite index* yang paling bagus. Jumlah data yang akan di uji pada pengujian ini adalah 10, 30, 50, 80, 110, 130, dan 150 data. Adapun hasil pengujian ditunjukkan pada tabel 5.2 Pengujian Jumlah Data.

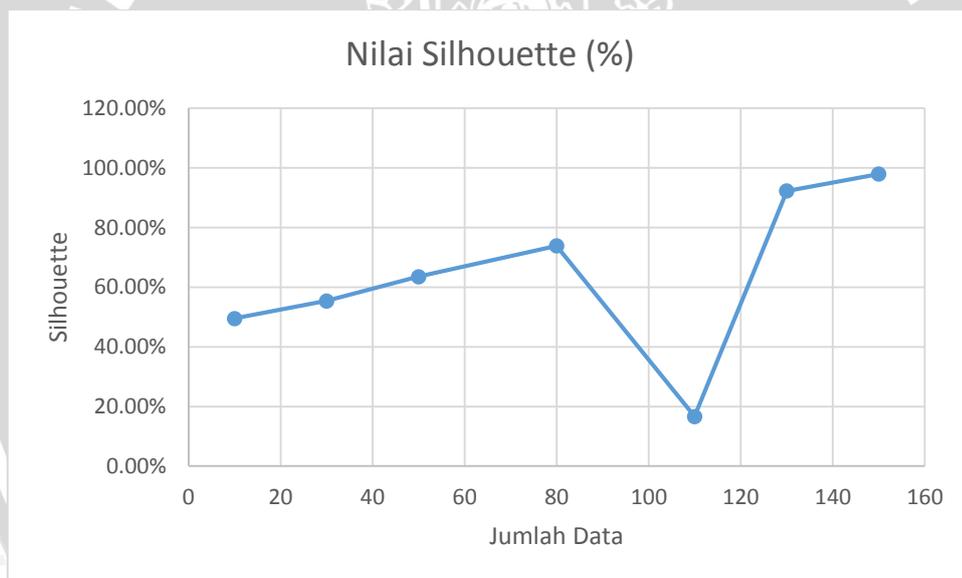
Table 5.2 Pengujian Jumlah Data.

Jumlah Data	Iterasi	Nilai Silhouette	
		Rentang -1 Sampai 1	Dalam % $((x+1)/2 \times 100)$
10	2	-0.009094672	49.55%
30	5	0.108245615	55.41%
50	7	0.271305403	63.57%
80	4	0.477937414	73.90%
110	2	-0.666666667	16.67%
130	7	0.845408815	92.27%
150	10	0.95959899	97.98%

Pada tabel 5.2 di atas dapat dilihat bahwa hasil pengujian yang dilakukan dengan metode *silhouette index* menghasilkan nilai tertinggi pada penggunaan 150 data dengan 5 cluster. Semakin sedikit jumlah data yang diuji maka semakin rendah nilai *silhouette* yang di dapat. Adapun hasil dari pengujian dalam bentuk statistik ditunjukkan pada gambar 5.5 dan gambar 5.6:



Gambar 5.5 Statistik Jumlah Iterasi Data 10 sampai 150



Gambar 5.6 Statistik Nilai Silhouette Data 10 sampai 150

Berdasarkan grafik hasil pengujian jumlah *cluster* yang ditunjukkan pada gambar 5.5 dan gambar 5.6 dapat dilihat hasil bahwa pengujian dengan menggunakan data sebanyak 150 data dengan 5 *cluster* menghasilkan nilai *silhouette* yang paling tinggi, hal tersebut dipengaruhi oleh jarak antar data pada kluster memiliki kerapatan yang cukup baik dan jarak antar kluster memiliki nilai yang cukup besar. Penurunan nilai *silhouette* pada penggunaan data yang sedikit dipengaruhi oleh data yang digunakan dan *centroid* yang terbentuk sehingga

memberikan pengaruh pada kerapatan data pada klaster dan jarak antar klaster satu dengan klaster lainnya. Pada hasil pengujian di atas menunjukkan bahwa penggunaan 150 data pada penelitian ini merupakan jumlah yang tepat karena memiliki kualitas nilai *silhouette* yang paling baik.

5.3 Pengujian Cluster K-Means

Pengujian cluster ini dilakukan untuk mengetahui berapa jumlah cluster terbaik yang menghasilkan nilai *Shiluite index* tertinggi untuk digunakan pada penentuan rekomendasi jodoh ini pada metode K-Means. Pengujian ini dilakukan dengan tujuan untuk mengetahui dan membandingkan hasil dari pengelompokan pada penggunaan improve K-Means dan K-Means. Jumlah cluter yang digunakan pada pengujian ini adalah 3, 4, 5, 6, 7, 8, 9, 10, 20, 25, 50, dan 100. Pada tiap pengujian *cluster* dilakukan 5 kali run program dan dari hasil yang didapat dilakukan rata rata yang merupakan nilai ahir dari pengujian. Adapun hasil pengujian ditunjukkan pada tabel 5.3 pengujian jumlah *cluster*

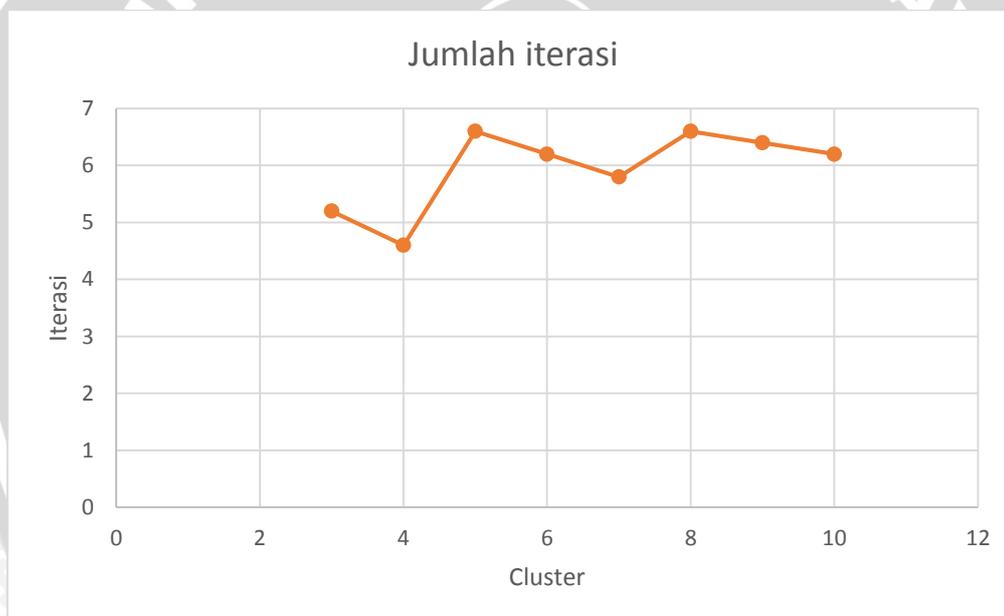
Table 5.3 Pengujian Jumlah cluster K-Means.

Cluster	Iterasi	Silhouette	Rata Rata	Silhouette % $((x+1)/2 \times 100)$
3	3	0.990077	0.972799	98.64%
	5	0.919617		
	5	0.98758		
	5	0.987427		
	8	0.979294		
4	6	0.986212	0.77566	88.78%
	6	0.338763		
	3	0.949099		
	4	0.660141		
	4	0.944084		
5	3	0.777444	0.865919	93.30%
	8	0.81151		
	9	0.968947		
	8	0.963707		
	5	0.807988		

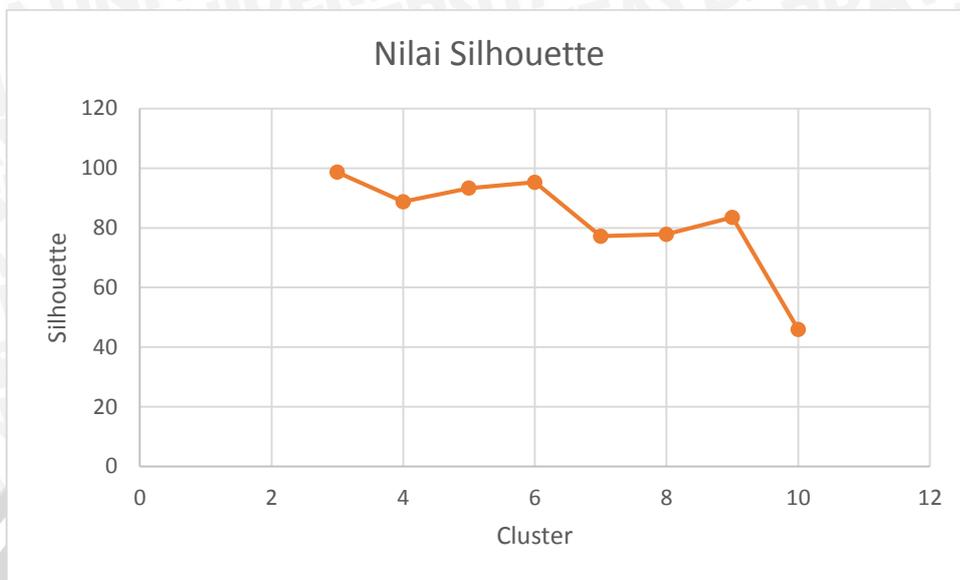
Cluster	Iterasi	Silhouette	Rata Rata	Silhouette % $((x+1)/2 \times 100)$
6	6	0.909827	0.905526	95.28%
	4	0.831688		
	5	0.937368		
	7	0.90844		
	9	0.940307		
7	4	0.948655	0.543853	77.19%
	6	-0.487226		
	6	0.919714		
	5	0.488662		
	8	0.849457		
8	6	0.588667	0.556738	77.84%
	4	0.877817		
	8	0.107498		
	8	0.289427		
	7	0.920281		
9	5	0.784571	0.670205	83.51%
	9	0.640333		
	7	0.854088		
	5	0.639282		
	6	0.432752		
10	6	-0.826714	-0.08103	45.95%
	7	0.411616		
	9	0.875073		
	5	-0.094522		
	4	-0.770611		
20	7	-0.93736	-0.46647	26.68%
	4	0.521477		
	4	-0.908453		
	5	-0.939545		
	4	-0.068485		
25	5	-0.94	-0.73353	13.32%
	7	-0.906667		
	4	-0.893333		
	3	0.019004		
	5	-0.946667		
50	3	-0.786667	-0.81867	9.07%
	4	-0.82		
	4	-0.846667		
	2	-0.786667		
	3	-0.853333		

Cluster	Iterasi	Silhouette	Rata Rata	Silhouette % $((x+1)/2 \times 100)$
100	3	-0.54	-0.46	27.00%
	3	-0.44		
	2	-0.46		
	3	-0.46		
	3	-0.4		

Pada tabel 5.3 di atas dapat dilihat bahwa hasil pengujian yang dilakukan dengan metode *silhouette index* pada metode K-Means menunjukkan bahwa penggunaan 3 *cluster* memiliki nilai *silhouette* paling tinggi disbanding dengan menggunakan jumlah *cluster* yang lain. Adapun hasil dari pengujian dalam bentuk statistik ditunjukkan pada gambar 5.7 sampai gambar 5.10:



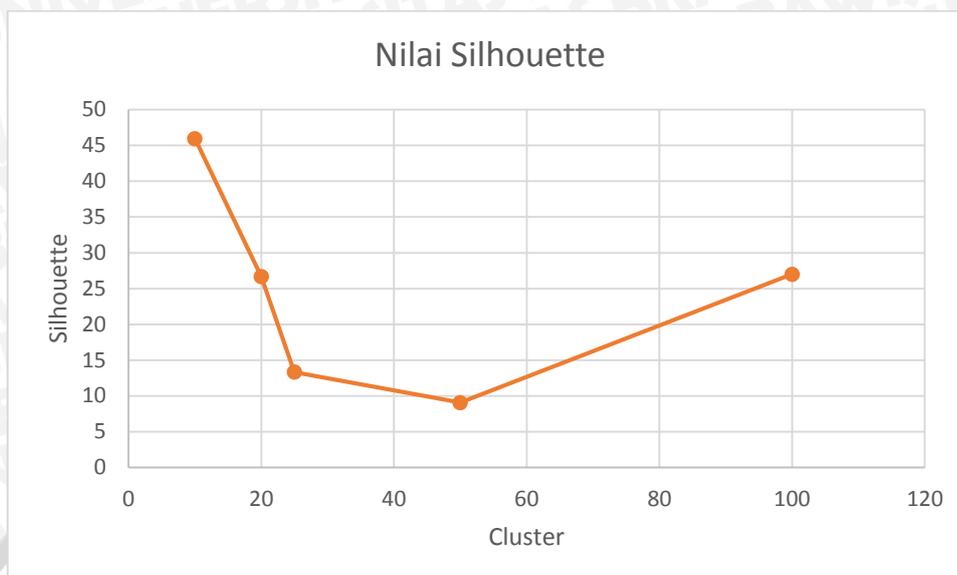
Gambar 5.7 Statistik jumlah iterasi cluster 3 sampai 10 K-Means



Gambar 5.8 Statistik nilai silhouette cluster 3 sampai 10 K-Means



Gambar 5.9 Statistik jumlah iterasi cluster 10 sampai 100 K-Means



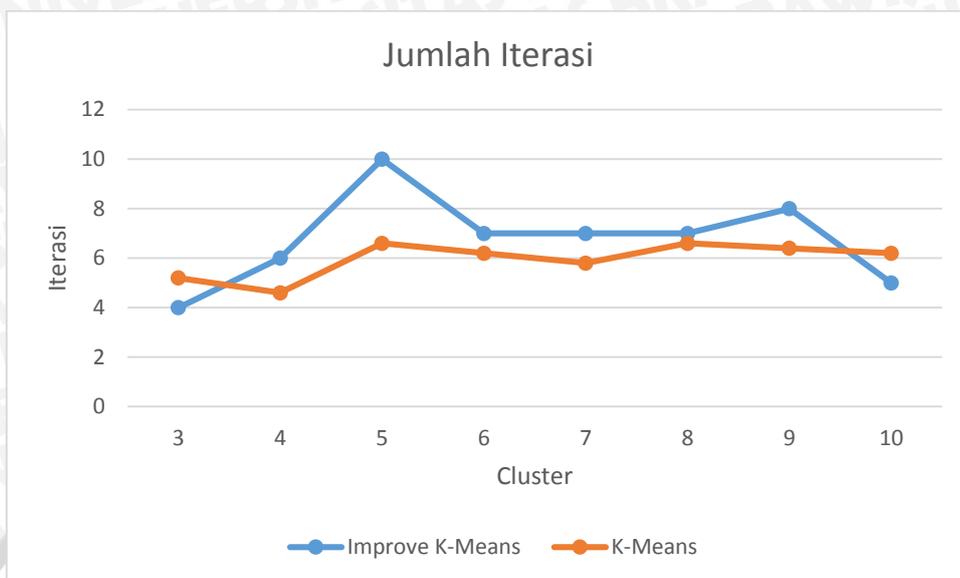
Gambar 5.10 Statistik nilai silhouette cluster 10 sampai 100 K-Means

Berdasarkan grafik hasil pengujian yang ditunjukkan pada gambar 5.6 sampai gambar 5.10 dapat dilihat hasil bahwa pengujian dengan menggunakan jumlah *cluster* = 3 memiliki nilai *silhouette* yang paling tinggi. Hal tersebut menunjukkan dengan penggunaan 3 *cluster* pada metode K-Means, jarak kerapatan data pada satu klaster memiliki jarak yang cukup baik dan jarak antar klaster satu dengan klaster lainnya cukup besar.

Penggunaan metode *K-Means* tidak memiliki nilai *silhouette* yang pasti dikarenakan pada metode *K-Means* inialisasi pusat klaster dilakukan secara *random* sehingga nilai *silhouette* yang dihasilkan tidak dapat diprediksi. Pada percobaan yang sudah dilakukan jumlah iterasinya memiliki kecenderungan nilai yang stabil, tidak terjadi peningkatan atau penurunan nilai iterasi yang sangat signifikan. Besar kecil nya banyak iterasi yang terjadi tidaklah dapat diprediksi karena pada *K-Means* inialisasi klaster awal dilakukan secara *random*.

5.4 Perbandingan pengujian K-Means dan Improve K-Means

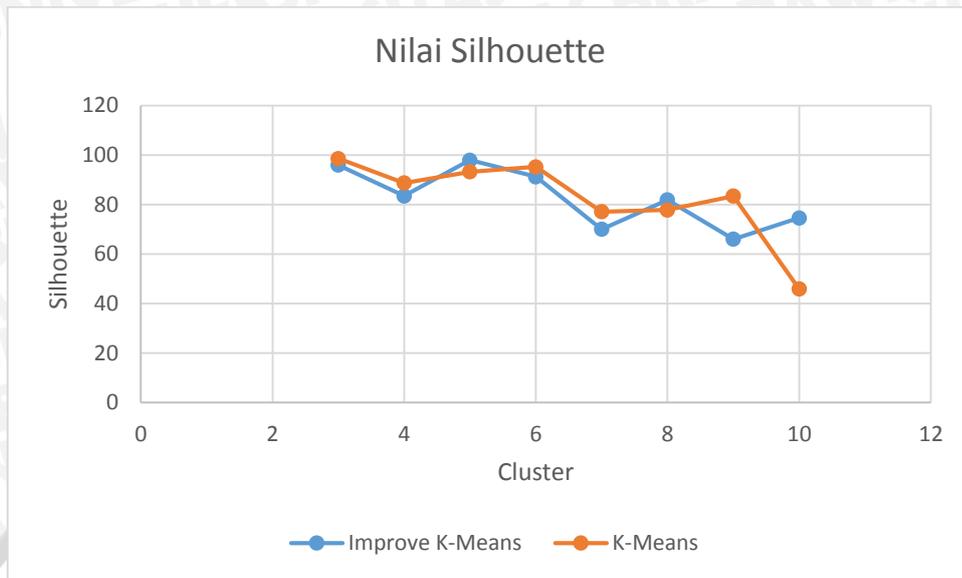
Pada sub bab ini akan membandingkan hasil dari pengujian yang dilakukan oleh metode Improve K-Means dengan metode K-Means. Adapun yang poin yang dilakukan perbandingan pada sub bab ini yaitu pada jumlah iterasi dan jumlah *cluster*. Adapun perbandingan yang diperoleh dari pengujian yang telah dilakukan sebelumnya ditunjukkan pada gambar 5.11 sampai gambar 5.14



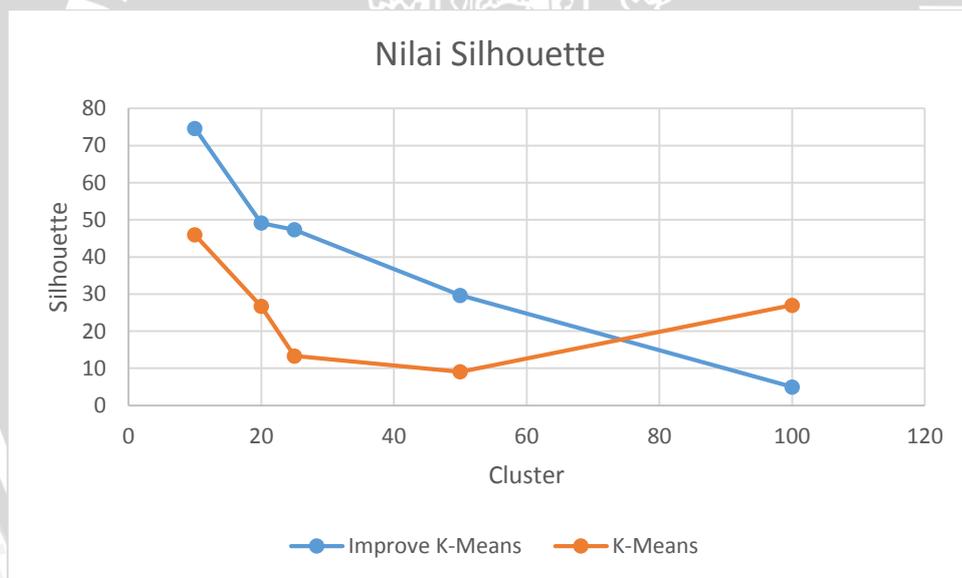
Gambar 5.11 Statistik Perbandingan iterasi cluster 3 – 10



Gambar 5.12 Statistik Perbandingan iterasi cluster 10 – 100



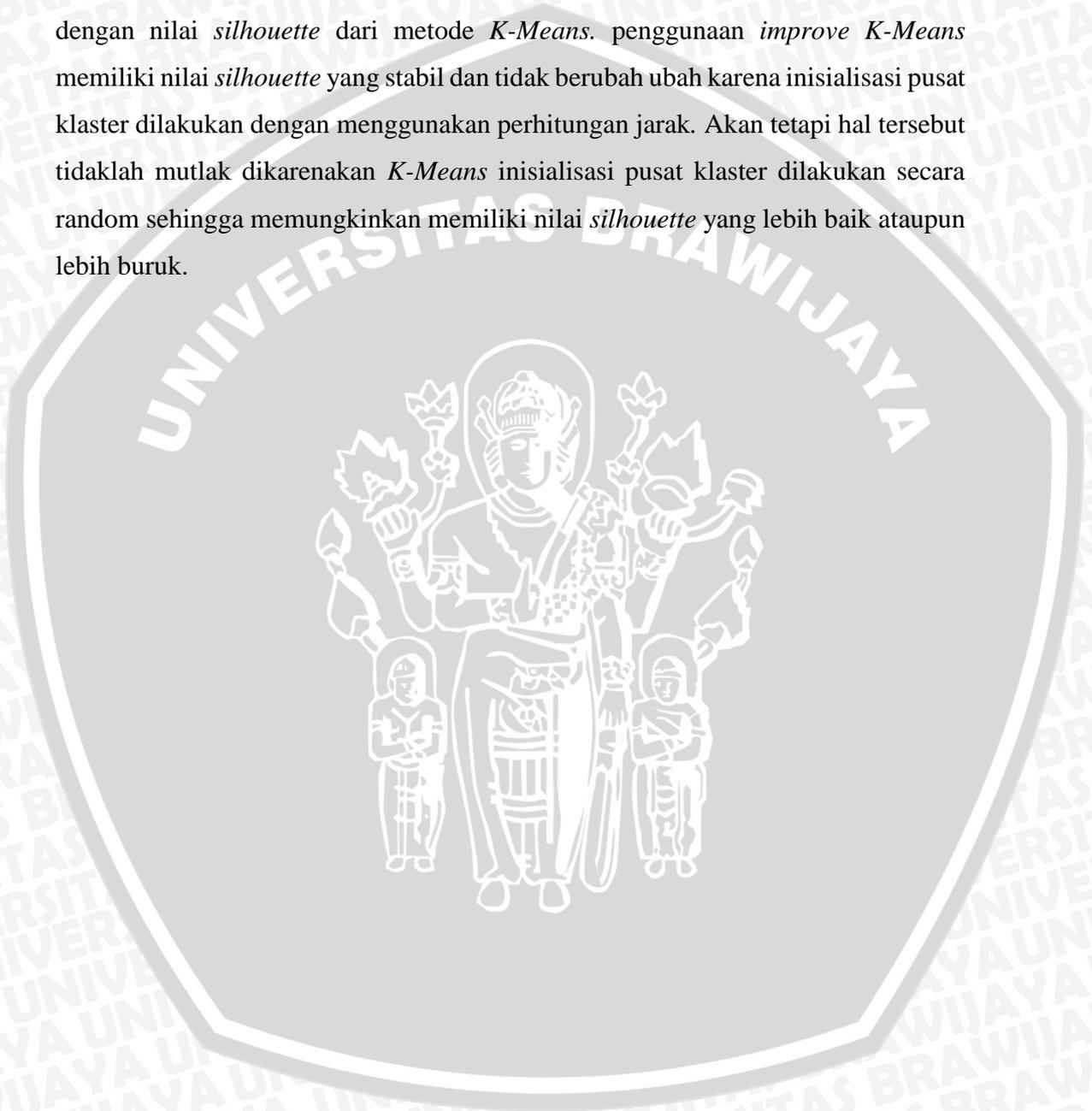
Gambar 5.13 Statistik Perbandingan nilai silhouette cluster 3 – 10



Gambar 5.14 Statistik Perbandingan nilai silhouette cluster 10 – 100

Dari proses pengujian di atas dilakukan perbandingan antara kedua hasil yang didapat dari pengujian masing masing metode. Pada jumlah iterasi yang terjadi, metode *improve K-Means* cenderung memiliki jumlah iterasi yang lebih banyak dibanding dengan iterasi yang terjadi pada metode *K-Means*, akan tetapi hal tersebut tidaklah mutlak karena pada *K-Means* inisialisasi pusat kluster dilakukan secara *random* sehingga memungkinkan jumlah iterasi yang terjadi lebih banyak ataupun lebih sedikit. Pada percobaan yang dilakukan sebelumnya, nilai *silhouette*

dari kedua metode nilai *silhouette* pada metode *improve K-Means* dan *K-Means* memiliki kecenderungan nilai yang hampir sama pada penggunaan 3 sampai 10 *cluster*. Sedangkan pada penggunaan 10 sampai 100 *cluster* penggunaan metode *improve K-Means* cenderung memiliki nilai *silhouette* yang lebih baik dibanding dengan nilai *silhouette* dari metode *K-Means*. penggunaan *improve K-Means* memiliki nilai *silhouette* yang stabil dan tidak berubah ubah karena inisialisasi pusat klaster dilakukan dengan menggunakan perhitungan jarak. Akan tetapi hal tersebut tidaklah mutlak dikarenakan *K-Means* inisialisasi pusat klaster dilakukan secara random sehingga memungkinkan memiliki nilai *silhouette* yang lebih baik ataupun lebih buruk.



BAB VI

PENUTUP

6.1 Kesimpulan

Berdasarkan hasil perancangan, implementasi dan pengujian yang telah dilakukan sebelumnya, maka didapatkan kesimpulan antara lain sebagai berikut :

1. Metode *improve K-Means pada inisialisasi centroid awal* dapat diterapkan pada penentuan rekomendasi jodoh dimana memiliki hasil yang lebih stabil dibandingkan dengan metode *K-Means*.
2. Metode *Improve K- Means pada inisialisasi centroid awal* dapat di implementasikan untuk rekomendasi jodoh dengan nilai kualitas *cluster* terbaik yang didapatkan adalah 0.959598 (97.98%) untuk 150 data dengan penggunaan 5 klaster.
3. Penambahan dan pengurangan jumlah *cluster* pada metode *Improve K- Means pada inisialisasi centroid awal* memiliki pengaruh terhadap nilai dari kualitas *cluster*. Pada penggunaan 5 *cluster* memiliki nilai kualitas yang paling baik daripada penggunaan jumlah *cluster* lain.
4. Penambahan dan pengurangan data pada *Improve K- Means pada inisialisasi centroid awal* memiliki pengaruh terhadap meningkat atau menurunnya kualitas *cluster*. Pada penggunaan 150 data dengan 5 *cluster* memiliki nilai kualitas yang lebih baik daripada penggunaan data yang lebih sedikit.
5. Pada pengujian jumlah cluster pada metode *K-Means* memiliki kecenderungan nilai *silhouette* yang hampir sama terhadap nilai *silhouette* yang dihasilkan pada metode *improve K-Means* pada penggunaan 3 sampai 10 cluster. Sedangkan pada penggunaan 10 – 100 cluster memiliki kecenderungan nilai yang kurang baik dibanding dengan nilai yang di hasilkan pada penggunaan metode *improve K-Means*. Akan tetapi hal tersebut tidaklah mutlak karena pada metode *K-Means* inisialisasi pusat klaster awal ditentukan secara *random*.
6. Dari hasil pengujian yang telah dilakukan sebelumnya, penggunaan metode *improve K-Means* memiliki kestabilan pada nilai *silhouette* dibandingkan dengan penggunaan metode *K-Means*. Penggunaan metode *K-Means* memiliki

kecenderungan nilai *silhouette* yang berubah ubah karena *centroid* ditentukan secara *random*.

6.2 Saran

Saran yang diberikan untuk pengembangan berikutnya yaitu hasil penelitian tentang implementasi metode *Improve K- Means pada inisialisasi centroid awal* untuk penentuan rekomendasi jodoh yaitu dengan menambahkan pengujian akurasi agar dapat diketahui tingkat keberhasilan dari pengujian yang dilakukan. Kemudian pada sistem agar didesain lebih *user friendly* agar mempermudah user untuk menggunakan sistem ini.



Daftar Pustaka

- [MSN-05] Murti, Darlis H., Suciati, Nanik., & Nanjaya, Daru. (2005) Clustering Data Non-Numerik Dengan Pendekatan Algoritma K-Means Dan Hamming Distance Studi Kasus Biro Jodoh. JUTI Vol .4 No.1, Januari(2005) 46-53 .
- [HAS-14] Harshanti, Septi (2014). Aplikasi Pendukung Keputusan Pencarian Jodoh. Semarang: Fakultas Ilmu Komputer Universitas Dian Nuswantoro Semarang .
- [YDW-13] Yao, Hong., Duan, Qingling., Li, Daoliang., & Wang, Jianping. (2013) An improved K-means clustering algorithm for fish image segmentation. *Mathematical and Computer Modelling*, 58 (2013) Pp. 790–798
- [GAH-10] Gao, Jinxin., & Hitchcock, David B. (2010) James–Stein shrinkage to improve k-means cluster analysis. *Computational Statistics and Data Analysis*, 54 (2010) Pp. 2113–2127
- [QZH-10] Qin, Xiaoping., Zheng, Shijue., Huang, Ying., & Deng, Guangsheng. (2010) Improved K-Means Algorithm and Application in Customer Segmentation Wearable. *Computing Systems (APWCS), (2010) Asia-Pacific Conference on*. Pp.224 – 227
- [ZHT-05] Zhong, W., Altun, G., Harrison, R., Tai, P.C., & Pan, Yi. (2005) Improved K-means clustering algorithm for exploring local protein sequence motifs representing common structural property. *NanoBioscience, IEEE Transactions on* (2005) Vol .4 Issue .3 Pp. 255 - 265
- [JHZ-10] Jin-hong, Zhang. (2010). A short-term prediction for QoS of Web Service based on RBF neural networks including an improved K-means algorithm. *Computer Application and System Modeling (ICCASM), International Conference on* Vol .5 (2010) Pp. V5-633 - V5-637
- [LLL-14] Li, Zhenkun., Lou, Yuehuan., Liu, Jun., Liu, Shu., Yang, Fan., Fu, Yang. (2014). Optimal partition of power distribution network

service areas based on improved K-Means algorithm. Power System Technology (POWERCON), International Conference on (2014) Pp. 883 - 889

- [DTS-13] Dalila, Shabrina Mardhi., Tjandrasa, Handayani., dan Suciati, Nanik. (2013). Modifikasi Algoritma Pengelompokan K-Means untuk Segmentasi Citra Ikan Berdasarkan Puncak Histogram. Surabaya: Institut Teknologi Sepuluh Nopember (ITS) JURNAL TEKNIK POMITS Vol. 1 No. 1 (2013) Pp. 1-5
- [PHC-09] Park , Hae-Sang dan Jun , Chi-Hyuck (2009). A simple and fast algorithm for K-medoids clustering. International Journal of Expert Systems with Applications 39. pp. 3336-3341
- [UMF-09] Umran, Munzir dan Fuadi Abidin, Taufik (2009). Pengelompokan Dokumen Menggunakan K-Means Dan Singular Value Decomposition: Studi Kasus Menggunakan Data Blog. SESINDO (2009) Jurusan Sistem Informasi ITS.
- [WBS-01] <http://www.apapengertianahli.com/2014/09/pengertian-data-mining-apa-itu-data-mining.html> di akses pada tanggal 14/5/2015 jam 12.46
- [WBS-02] <https://yudiagusta.wordpress.com/clustering/> di akses pada tanggal 14/5/2015 jam 13.58
- [WBS-03] <http://id.wikipedia.org/wiki/K-means> di akses pada 15/5/2015 jam 20.55
- [CNS-14] Cholid ,Nella Safitri (2014). Faktor utama seseorang sulit mencari jodoh. <http://lifestyle.okezone.com/read/2014/09/15/198/1039528/faktor-utama-seseorang-sulit-cari-jodoh>. Di akses pada 20 Mei 2015
- [SUT-03] Sutanto (2003). Analisis Data. Depok: Fakultas Kesehatan Masyarakat Universitas Indonesia.
- [SSN-10] Singih Santoso (2010). Statistkik Parametrik Konsep dab Aplikasi dengan SPSS. PT Alex Media Komputindo (2010)
- [DEG-08] Degenova (2008). Intimate Relationships, Marriages, and Families 7TH EDITION. United States of America. McGraw-Hil.

- [JBM-11] Junaedi, Hartarto., Budianto, Herman., Maryati, Indra., dan Melani, Yuliana (2011). Data Transformation Pada Mining. Prosiding Konferensi Nasional. Inovasi dalam Desain dan Teknologi IDEaTech(2011).
- [WBS-04] <http://www.cikgudahlia.com/2011/11/jenis-datanumerik-dan-bukan-numerik.html>. Jenis Data-Numerik dan bukan Numerik. Diakses pada tanggal 26 mei 2015
- [SAN-04] Santoso (2004). SPSS Versi 10. Mengolah Data Statistik Secara Profesional, Jakarta: Elex Media Komputindo.
- [WBS-05] http://mines.humanoriented.com/classes/2010/fall/csci568/portfolio_exports/mvo_get/cluster/cluster.html. Diakses pada tanggal 28 mei 2015
- [WBS-06] <http://informatika.web.id/definisi-dan-tujuan-normalisasi.htm> diakses pada tanggal 30 mei 2015
- [HAR-13] Handoyo, Rendy (2013). Perbandingan Metode Clustering Menggunakan Metode Single Linkage dan K-Means pada Pengelompokan Dokumen. Program Studi Sistem Komputer Fakultas Elektro dan Komunikasi Institut Teknologi Telkom Bandung (2013)
- [CZS-10] Chen, Zhang dan Shixiong , Xia. K-means Clustering Algorithm with improved Initial Center. Second International Workshop on Knowledge Discovery and Data Mining

DAFTAR LAMPIRAN

LAMPIRAN 1 FORM KUISIONER

SURVEY CALON PASANGAN HIDUP IDAMAN

Kuisisioner ini bertujuan untuk mensurvei data secara umum pada usia produktif masyarakat terkait kriteria calon pasangan hidup yang diinginkan dan data ini akan digunakan untuk membuat sistem rekomendasi jodoh idaman yang tepat sesuai dengan kriteria kriteria yang telah ditentukan.

Dimohon berjasamanya untuk mengisikan data sebaik-baiknya demi menunjang kelancaran pencarian data responden dalam menyelesaikan tugas akhir. Terima Kasih

Continue » 33% completed

Powered by Google Forms. This content is neither created nor endorsed by Google. Report Abuse - Terms of Service - Additional Terms

SURVEY CALON PASANGAN HIDUP IDAMAN

* Required

Silahkan Isikan Data Diri Anda
Dimohon isilah data dengan sebenar-benarnya karena data yang anda isikan tidak akan dipublikasikan.

Silahkan Isi Nama Lengkap anda? *

Apa Jenis Kelamin anda? *

Laki - Laki
 Perempuan

Berapa Umur anda sekarang? *

20 tahun < 25 tahun
 25 tahun < 30 tahun
 30 tahun < 35 tahun
 > 35 tahun





Berapa tinggi badan anda? *

< 155 cm
 155 cm < 165 cm
 165 cm < 175 cm
 > 175 cm

Berapa berat badan anda? *

< 50 kg
 50 kg < 60 kg
 60 kg < 70 kg
 > 70 kg

Bagaimana jenis rambut anda? *

Lurus
 Keriting
 Ikal
 Botak

Berasal dari suku manakah anda? *

Banjar
 Satak
 Betawi
 Bugis
 Dayak
 Jawa
 Minang
 Papua
 Sunda
 Tionghoa
 Lainnya

Apa warna kulit anda? *

Putih
 Sawo Matang
 Kuning Langsat
 Hitam

Apa agama anda? *

Islam
 Kristen Protestan
 Khatolik
 Hindu
 Kong Huchu
 Budha

Apa pendidikan terakhir anda? *

SMA Sederajat
 Sarjana
 Pasca Sarjana (S2/S3)
 Lainnya

Apa pekerjaan anda? *

Pegawai Negeri Sipil
 Swasta
 Wiraswasta/Entrepreneur
 BLUMN
 Olahragawan
 Lainnya

Apa status pernikahan anda sekarang? *

Lajang
 Cera
 Ditinggal Mati
 Pisah Ranjang/Rumah

Apakah anda Perokok? *

Ya
 Tidak

66% completed

Powered by Google Forms

This content is neither created nor endorsed by Google.
[Report Abuse](#) - [Terms of Service](#) - [Additional Terms](#)



SURVEY CALON PASANGAN HIDUP IDAMAN

Your response has been recorded.

[Submit another response](#)

This form was created using Google Forms.
[Create your own](#) 



LAMPIRAN 2 HASIL KUISIONER (GOOGLE DRIVE)

SURVEY CALON PASANGAN HIDUP IDAMAN (Responses) ☆

File Edit View Insert Format Data Tools Form Add-ons Help All changes saved in Drive

Comments Share

fx Silahkan isi Nama Lengkap anda?

	B	C	D	E	F	G	H	I	J
1	Silahkan isi Nama Lengkap anda?	Apa Jenis Kelamin anda?	Berapa Umur anda sekarang?	Berapa tinggi badan anda?	Berapa berat badan anda?	Bagaimana jenis rambut?	Berasal dari suku manakah?	Apa warna kulit anda?	Apa agama anda?
2	yunar pribadi	Laki - Laki	20 tahun < 25 tahun	165 cm < 175 cm	60 kg < 70 kg	Lurus	Jawa	Sawo Matang	Islam
3	Chasandra Puspitasari	Perempuan	20 tahun < 25 tahun	155 cm < 165 cm	50 kg < 60 kg	Lurus	Jawa	Sawo Matang	Khatolik
4	Dwi Kurnia Ayuningrum	Perempuan	20 tahun < 25 tahun	155 cm < 165 cm	50 kg < 60 kg	Ikhal	Jawa	Kuning Langsat	Islam
5	Ilham Andy Wira Utama	Laki - Laki	20 tahun < 25 tahun	165 cm < 175 cm	60 kg < 70 kg	Lurus	Jawa	Sawo Matang	Islam
6	Anonymous	Laki - Laki	20 tahun < 25 tahun	165 cm < 175 cm	> 70 kg	Keriting	Jawa	Sawo Matang	Islam
7	Dany Muhammad	Laki - Laki	20 tahun < 25 tahun	155 cm < 165 cm	50 kg < 60 kg	Lurus	Jawa	Sawo Matang	Islam
8	Briandana Riznov	Laki - Laki	20 tahun < 25 tahun	165 cm < 175 cm	> 70 kg	Lurus	Jawa	Sawo Matang	Islam
9	Dwi Hardyanto	Laki - Laki	20 tahun < 25 tahun	155 cm < 165 cm	60 kg < 70 kg	Ikhal	Sunda	Sawo Matang	Islam
10	ria rizki wardani	Perempuan	20 tahun < 25 tahun	155 cm < 165 cm	50 kg < 60 kg	Ikhal	Jawa	Kuning Langsat	Islam
11	Lolipop	Laki - Laki	20 tahun < 25 tahun	165 cm < 175 cm	60 kg < 70 kg	Ikhal	Sunda	Sawo Matang	Islam
12	Erwin Rizki Ariyanto	Laki - Laki	20 tahun < 25 tahun	165 cm < 175 cm	60 kg < 70 kg	Lurus	Jawa	Kuning Langsat	Kristen Protesta
13	adryan deryck	Laki - Laki	25 tahun < 30 tahun	> 175 cm	> 70 kg	Lurus	Tionghoa	Putih	Budha
14	Setyan Pamungkas	Laki - Laki	20 tahun < 25 tahun	155 cm < 165 cm	50 kg < 60 kg	Lurus	Jawa	Putih	Islam
15	Dimas Ferdiansyah	Laki - Laki	20 tahun < 25 tahun	> 175 cm	50 kg < 60 kg	Lurus	Jawa	Sawo Matang	Islam
16	Novan Agung Wasikto	Laki - Laki	20 tahun < 25 tahun	165 cm < 175 cm	60 kg < 70 kg	Lurus	Tionghoa	Kuning Langsat	Kristen Protesta
17	Ilham Setiawan	Laki - Laki	20 tahun < 25 tahun	< 155 cm	< 50 kg	Lurus	Jawa	Sawo Matang	Islam
18	Faisal Rofda Rohman	Laki - Laki	20 tahun < 25 tahun	165 cm < 175 cm	50 kg < 60 kg	Lurus	Jawa	Sawo Matang	Islam
19	Agata	Perempuan	20 tahun < 25 tahun	155 cm < 165 cm	< 50 kg	Ikhal	Jawa	Kuning Langsat	Islam
20	Ferdi Alvanda	Laki - Laki	20 tahun < 25 tahun	165 cm < 175 cm	60 kg < 70 kg	Lurus	Jawa	Sawo Matang	Islam
21	Renno Andika Syawaludi	Laki - Laki	20 tahun < 25 tahun	165 cm < 175 cm	50 kg < 60 kg	Keriting	Jawa	Sawo Matang	Islam
22	FAIZAL ABDI	Laki - Laki	20 tahun < 25 tahun	165 cm < 175 cm	60 kg < 70 kg	Lurus	Jawa	Sawo Matang	Islam
23	Henda	Laki - Laki	20 tahun < 25 tahun	165 cm < 175 cm	60 kg < 70 kg	Ikhal	Jawa	Sawo Matang	Islam
24	Muhammad Hafizh	Laki - Laki	20 tahun < 25 tahun	165 cm < 175 cm	50 kg < 60 kg	Ikhal	Minang	Sawo Matang	Islam
25	Febri ardianto	Laki - Laki	20 tahun < 25 tahun	155 cm < 165 cm	< 50 kg	Lurus	Jawa	Sawo Matang	Islam
26	azam khoirutsani	Laki - Laki	25 tahun < 30 tahun	> 175 cm	60 kg < 70 kg	Lurus	Jawa	Sawo Matang	Islam
27	Ariztyawan	Laki - Laki	20 tahun < 25 tahun	165 cm < 175 cm	60 kg < 70 kg	Ikhal	Jawa	Sawo Matang	Islam

Form Responses 1

SURVEY CALON PASANGAN HIDUP IDAMAN (Responses) ☆

File Edit View Insert Format Data Tools Form Add-ons Help All changes saved in Drive

Comments Share

fx Silahkan isi Nama Lengkap anda?

	B	C	D	E	F	G	H	I	J
1	Silahkan isi Nama Lengkap anda?	Apa Jenis Kelamin anda?	Berapa Umur anda sekarang?	Berapa tinggi badan anda?	Berapa berat badan anda?	Bagaimana jenis rambut?	Berasal dari suku manakah?	Apa warna kulit anda?	Apa agama anda?
27	Ariztyawan	Laki - Laki	20 tahun < 25 tahun	165 cm < 175 cm	60 kg < 70 kg	Ikhal	Jawa	Sawo Matang	Islam
28	Zulfa Az-Zakiyah R	Perempuan	20 tahun < 25 tahun	155 cm < 165 cm	< 50 kg	Keriting	Jawa	Kuning Langsat	Islam
29	tri wahyu sulistyawati	Perempuan	20 tahun < 25 tahun	155 cm < 165 cm	60 kg < 70 kg	Ikhal	Jawa	Sawo Matang	Islam
30	riza aris prayudhi	Laki - Laki	20 tahun < 25 tahun	165 cm < 175 cm	50 kg < 60 kg	Lurus	Jawa	Sawo Matang	Islam
31	ricka amalia praptiwi	Perempuan	20 tahun < 25 tahun	155 cm < 165 cm	50 kg < 60 kg	Lurus	Jawa	Kuning Langsat	Islam
32	alif fachrony	Laki - Laki	20 tahun < 25 tahun	165 cm < 175 cm	50 kg < 60 kg	Ikhal	Jawa	Sawo Matang	Islam
33	yulian asri ansoni	Laki - Laki	20 tahun < 25 tahun	> 175 cm	60 kg < 70 kg	Lurus	Jawa	Sawo Matang	Islam
34	Anas Rachmadi Priambo	Laki - Laki	20 tahun < 25 tahun	165 cm < 175 cm	50 kg < 60 kg	Lurus	Jawa	Putih	Islam
35	Fajar Amullah	Laki - Laki	20 tahun < 25 tahun	165 cm < 175 cm	50 kg < 60 kg	Lurus	Jawa	Sawo Matang	Islam
36	Aini Wahyu Tyastari	Perempuan	20 tahun < 25 tahun	155 cm < 165 cm	50 kg < 60 kg	Keriting	Jawa	Putih	Islam
37	rachmawati	Perempuan	20 tahun < 25 tahun	< 155 cm	50 kg < 60 kg	Ikhal	Jawa	Sawo Matang	Islam
38	Dinda Novitasari	Perempuan	20 tahun < 25 tahun	155 cm < 165 cm	50 kg < 60 kg	Ikhal	Jawa	Putih	Islam
39	FFF	Perempuan	20 tahun < 25 tahun	155 cm < 165 cm	50 kg < 60 kg	Lurus	Jawa	Sawo Matang	Islam
40	adu	Laki - Laki	20 tahun < 25 tahun	155 cm < 165 cm	50 kg < 60 kg	Lurus	Jawa	Sawo Matang	Islam
41	Fitarina Azmi Syafiq	Perempuan	20 tahun < 25 tahun	155 cm < 165 cm	< 50 kg	Keriting	Jawa	Kuning Langsat	Islam
42	I Gusti Wahyu W	Laki - Laki	20 tahun < 25 tahun	165 cm < 175 cm	50 kg < 60 kg	Lurus	Jawa	Sawo Matang	Islam
43	Nimas Ayu Sekamingrum	Perempuan	20 tahun < 25 tahun	155 cm < 165 cm	50 kg < 60 kg	Ikhal	Jawa	Kuning Langsat	Islam
44	Fauziatul Munawaroh	Perempuan	20 tahun < 25 tahun	155 cm < 165 cm	< 50 kg	Lurus	Jawa	Sawo Matang	Islam
45	Samaher	Perempuan	20 tahun < 25 tahun	155 cm < 165 cm	< 50 kg	Ikhal	Jawa	Putih	Islam
46	Mira Susanti	Perempuan	20 tahun < 25 tahun	155 cm < 165 cm	< 50 kg	Keriting	Jawa	Sawo Matang	Islam
47	Verawati Arifin	Perempuan	20 tahun < 25 tahun	155 cm < 165 cm	50 kg < 60 kg	Lurus	Jawa	Sawo Matang	Islam
48	Agandaning Ilimi Hamis	Perempuan	20 tahun < 25 tahun	< 155 cm	< 50 kg	Lurus	Jawa	Sawo Matang	Islam
49	Rainey Merdana	Laki - Laki	20 tahun < 25 tahun	165 cm < 175 cm	> 70 kg	Lurus	Minang	Putih	Islam
50	Muhammad Gighi Wicak	Laki - Laki	20 tahun < 25 tahun	165 cm < 175 cm	> 70 kg	Ikhal	Jawa	Sawo Matang	Islam
51	Marlia Sari Pangestuti	Perempuan	20 tahun < 25 tahun	< 155 cm	< 50 kg	Lurus	Jawa	Kuning Langsat	Islam
52	The only Kuniamon	Laki - Laki	20 tahun < 25 tahun	165 cm < 175 cm	50 kg < 60 kg	Ikhal	Jawa	Sawo Matang	Islam

Form Responses 1



SURVEY CALON PASANGAN HIDUP IDAMAN (Responses) ☆

File Edit View Insert Format Data Tools Form Add-ons Help All changes saved in Drive

Comments Share

fx Silahkan isi Nama Lengkap anda?

	B	C	D	E	F	G	H	I	J
1	Silahkan isi Nama Lengkap anda?	Apa Jenis Kelamin anda?	Berapa Umur anda sekarang?	Berapa tinggi badan anda?	Berapa berat badan anda?	Bagaimana jenis rambut?	Berasal dari suku mana?	Apa warna kulit anda?	Apa agama anda?
53	nur	Perempuan	20 tahun < 25 tahun	155 cm < 165 cm	50 kg < 60 kg	Lurus	Jawa	Kuning Langsat	Islam
54	Khusna Indra P	Laki - Laki	20 tahun < 25 tahun	165 cm < 175 cm	> 70 kg	Lurus	Jawa	Sawo Matang	Islam
55	Setya Uswatun Hasanah	Perempuan	20 tahun < 25 tahun	165 cm < 175 cm	< 50 kg	Keriting	Jawa	Sawo Matang	Islam
56	zahro mairuna	Perempuan	20 tahun < 25 tahun	155 cm < 165 cm	50 kg < 60 kg	Lurus	Jawa	Sawo Matang	Islam
57	Durotul Fakhriroh	Perempuan	20 tahun < 25 tahun	155 cm < 165 cm	< 50 kg	Ikhal	Jawa	Sawo Matang	Islam
58	ziel	Laki - Laki	20 tahun < 25 tahun	165 cm < 175 cm	60 kg < 70 kg	Lurus	Jawa	Sawo Matang	Islam
59	Eureka Kurnia Sari	Perempuan	20 tahun < 25 tahun	155 cm < 165 cm	< 50 kg	Lurus	Jawa	Putih	Islam
60	N.M Rashid	Laki - Laki	20 tahun < 25 tahun	> 175 cm	> 70 kg	Ikhal	Jawa	Sawo Matang	Islam
61	Rizki	Laki - Laki	20 tahun < 25 tahun	> 175 cm	60 kg < 70 kg	Lurus	Jawa	Sawo Matang	Islam
62	OLYVIA BETA PUTRI PE	Perempuan	20 tahun < 25 tahun	165 cm < 175 cm	50 kg < 60 kg	Keriting	Jawa	Kuning Langsat	Islam
63	Monel Lindu Sagala	Laki - Laki	20 tahun < 25 tahun	165 cm < 175 cm	50 kg < 60 kg	Ikhal	Batak	Putih	Kristen Protesta
64	LDC	Perempuan	20 tahun < 25 tahun	155 cm < 165 cm	50 kg < 60 kg	Lurus	Jawa	Sawo Matang	Islam
65	Ibrahim kusuma	Laki - Laki	20 tahun < 25 tahun	165 cm < 175 cm	50 kg < 60 kg	Ikhal	Jawa	Putih	Islam
66	Mega Ratri	Perempuan	20 tahun < 25 tahun	155 cm < 165 cm	< 50 kg	Ikhal	Jawa	Sawo Matang	Islam
67	Fajar Zazuli	Laki - Laki	20 tahun < 25 tahun	155 cm < 165 cm	< 50 kg	Keriting	Jawa	Sawo Matang	Islam
68	welly cyndy pitaningrum	Perempuan	20 tahun < 25 tahun	< 155 cm	60 kg < 70 kg	Lurus	Jawa	Kuning Langsat	Islam
69	Nur Alfiyah	Perempuan	20 tahun < 25 tahun	155 cm < 165 cm	50 kg < 60 kg	Ikhal	Jawa	Kuning Langsat	Islam
70	zella	Perempuan	20 tahun < 25 tahun	155 cm < 165 cm	50 kg < 60 kg	Lurus	Tionghoa	Sawo Matang	Khatolik
71	Detri Avinia Kusnanto	Perempuan	20 tahun < 25 tahun	155 cm < 165 cm	50 kg < 60 kg	Ikhal	Jawa	Sawo Matang	Islam
72	iko ahiyama	Laki - Laki	20 tahun < 25 tahun	165 cm < 175 cm	50 kg < 60 kg	Lurus	Sunda	Sawo Matang	Islam
73	Arif Sebastian	Laki - Laki	25 tahun < 30 tahun	165 cm < 175 cm	50 kg < 60 kg	Keriting	Banjar	Putih	Islam
74	N. Said	Laki - Laki	20 tahun < 25 tahun	155 cm < 165 cm	> 70 kg	Ikhal	Banjar	Sawo Matang	Islam
75	Rini	Perempuan	20 tahun < 25 tahun	155 cm < 165 cm	< 50 kg	Ikhal	Jawa	Kuning Langsat	Islam
76	Riza nindiana	Perempuan	25 tahun < 30 tahun	155 cm < 165 cm	50 kg < 60 kg	Keriting	Jawa	Kuning Langsat	Islam
77	Zahida arga	Perempuan	20 tahun < 25 tahun	155 cm < 165 cm	< 50 kg	Lurus	Jawa	Putih	Islam
78	alfiana ainur rohma	Perempuan	25 tahun < 30 tahun	165 cm < 175 cm	50 kg < 60 kg	Lurus	Jawa	Putih	Islam

Form Responses 1

SURVEY CALON PASANGAN HIDUP IDAMAN (Responses) ☆

File Edit View Insert Format Data Tools Form Add-ons Help All changes saved in Drive

Comments Share

fx Silahkan isi Nama Lengkap anda?

	B	C	D	E	F	G	H	I	J
1	Silahkan isi Nama Lengkap anda?	Apa Jenis Kelamin anda?	Berapa Umur anda sekarang?	Berapa tinggi badan anda?	Berapa berat badan anda?	Bagaimana jenis rambut?	Berasal dari suku mana?	Apa warna kulit anda?	Apa agama anda?
79	gino carlos	Laki - Laki	20 tahun < 25 tahun	165 cm < 175 cm	60 kg < 70 kg	Lurus	Jawa	Putih	Islam
80	dieded	Laki - Laki	20 tahun < 25 tahun	165 cm < 175 cm	60 kg < 70 kg	Ikhal	Jawa	Sawo Matang	Islam
81	Supriyadi	Laki - Laki	25 tahun < 30 tahun	165 cm < 175 cm	50 kg < 60 kg	Ikhal	Jawa	Sawo Matang	Islam
82	Angelina Restiawan	Perempuan	25 tahun < 30 tahun	155 cm < 165 cm	50 kg < 60 kg	Lurus	Jawa	Putih	Kristen Protesta
83	Rika Niesha	Perempuan	25 tahun < 30 tahun	< 155 cm	50 kg < 60 kg	Ikhal	Jawa	Sawo Matang	Islam
84	Saecripto	Laki - Laki	25 tahun < 30 tahun	165 cm < 175 cm	50 kg < 60 kg	Keriting	Bugis	Sawo Matang	Islam
85	Ahmad ariz prabawa	Laki - Laki	20 tahun < 25 tahun	165 cm < 175 cm	60 kg < 70 kg	Ikhal	Jawa	Sawo Matang	Islam
86	Rani Kurnia	Perempuan	25 tahun < 30 tahun	155 cm < 165 cm	< 50 kg	Lurus	Jawa	Kuning Langsat	Islam
87	sofyana ali pradana	Laki - Laki	20 tahun < 25 tahun	155 cm < 165 cm	50 kg < 60 kg	Lurus	Jawa	Kuning Langsat	Islam
88	Vronika Ervina	Perempuan	25 tahun < 30 tahun	155 cm < 165 cm	50 kg < 60 kg	Ikhal	Jawa	Putih	Islam
89	Naksa Garnida Arfie	Perempuan	25 tahun < 30 tahun	155 cm < 165 cm	50 kg < 60 kg	Ikhal	Sunda	Kuning Langsat	Islam
90	Merin Indarto Putri	Perempuan	25 tahun < 30 tahun	155 cm < 165 cm	60 kg < 70 kg	Lurus	Jawa	Putih	Islam
91	Destyana Puspitasari	Perempuan	20 tahun < 25 tahun	155 cm < 165 cm	< 50 kg	Keriting	Jawa	Putih	Islam
92	Eva Shand Hadji	Perempuan	25 tahun < 30 tahun	155 cm < 165 cm	50 kg < 60 kg	Lurus	Jawa	Putih	Islam
93	Igusti Wahyu Wardana	Laki - Laki	25 tahun < 30 tahun	165 cm < 175 cm	50 kg < 60 kg	Ikhal	Jawa	Sawo Matang	Islam
94	Atin Nurmayasanti	Perempuan	20 tahun < 25 tahun	165 cm < 175 cm	50 kg < 60 kg	Ikhal	Jawa	Putih	Islam
95	Na Mira	Perempuan	20 tahun < 25 tahun	155 cm < 165 cm	50 kg < 60 kg	Ikhal	Jawa	Sawo Matang	Islam
96	Anis S	Perempuan	25 tahun < 30 tahun	155 cm < 165 cm	50 kg < 60 kg	Ikhal	Jawa	Sawo Matang	Islam
97	rendyan	Laki - Laki	20 tahun < 25 tahun	165 cm < 175 cm	60 kg < 70 kg	Ikhal	Sunda	Sawo Matang	Islam
98	Era Nurria	Perempuan	20 tahun < 25 tahun	155 cm < 165 cm	< 50 kg	Lurus	Jawa	Sawo Matang	Islam
99	Dedi	Laki - Laki	20 tahun < 25 tahun	165 cm < 175 cm	60 kg < 70 kg	Ikhal	Jawa	Putih	Islam
100	Wanda Igro Fatra	Perempuan	25 tahun < 30 tahun	165 cm < 175 cm	50 kg < 60 kg	Ikhal	Lainya	Kuning Langsat	Islam
101	Alim budono	Laki - Laki	20 tahun < 25 tahun	> 175 cm	> 70 kg	Ikhal	Lainya	Sawo Matang	Islam
102	Rini Rosidah	Perempuan	20 tahun < 25 tahun	< 155 cm	< 50 kg	Ikhal	Jawa	Sawo Matang	Islam
103	ayub tatya admaja	Laki - Laki	20 tahun < 25 tahun	165 cm < 175 cm	> 70 kg	Lurus	Jawa	Sawo Matang	Islam
104	Rona Azura Mukti	Perempuan	20 tahun < 25 tahun	155 cm < 165 cm	50 kg < 60 kg	Ikhal	Sunda	Sawo Matang	Islam

Form Responses 1



SURVEY CALON PASANGAN HIDUP IDAMAN (Responses) ☆

File Edit View Insert Format Data Tools Form Add-ons Help All changes saved in Drive

Comments Share

Siapkan isi Nama Lengkap anda?

	B	C	D	E	F	G	H	I	J
1	Siapkan isi Nama Lengkap anda?	Apa Jenis Kelamin anda?	Berapa Umur anda sekarang?	Berapa tinggi badan anda?	Berapa berat badan anda?	Bagaimana jenis rambut?	Berasal dari suku mana?	Apa warna kulit anda?	Apa agama anda?
105	Aji Fatchul Huda	Laki - Laki	20 tahun < 25 tahun	165 cm < 175 cm	60 kg < 70 kg	Keriting	Jawa	Sawo Matang	Islam
106	Muhammad Fadhil Rezka	Laki - Laki	20 tahun < 25 tahun	> 175 cm	> 70 kg	Lurus	Minang	Kuning Langsat	Islam
107	Ummul Latifah	Perempuan	20 tahun < 25 tahun	155 cm < 165 cm	< 50 kg	Ikall	Jawa	Sawo Matang	Islam
108	Alda Alfa Maghthroh	Perempuan	20 tahun < 25 tahun	155 cm < 165 cm	< 50 kg	Lurus	Jawa	Sawo Matang	Islam
109	Don Stark Khal Drogo	Laki - Laki	30 tahun < 35 tahun	> 175 cm	> 70 kg	Keriting	Papua	Hitam	Islam
110	Happysia Virganadine	Perempuan	25 tahun < 30 tahun	< 155 cm	50 kg < 60 kg	Lurus	Jawa	Putih	Kristen Proteste
111	Nova Dynad	Perempuan	25 tahun < 30 tahun	155 cm < 165 cm	50 kg < 60 kg	Keriting	Jawa	Sawo Matang	Islam
112	Dede	Laki - Laki	20 tahun < 25 tahun	165 cm < 175 cm	50 kg < 60 kg	Ikall	Jawa	Kuning Langsat	Islam
113	Amiruddin Afif	Laki - Laki	25 tahun < 30 tahun	155 cm < 165 cm	50 kg < 60 kg	Keriting	Jawa	Sawo Matang	Islam
114	Mandha d'Aizcha	Perempuan	20 tahun < 25 tahun	155 cm < 165 cm	< 50 kg	Lurus	Jawa	Putih	Islam
115	Nindhita Cahya	Perempuan	20 tahun < 25 tahun	155 cm < 165 cm	< 50 kg	Ikall	Jawa	Kuning Langsat	Islam
116	deli	Laki - Laki	20 tahun < 25 tahun	155 cm < 165 cm	> 70 kg	Keriting	Jawa	Sawo Matang	Islam
117	Aji	Laki - Laki	20 tahun < 25 tahun	165 cm < 175 cm	50 kg < 60 kg	Ikall	Jawa	Sawo Matang	Islam
118	Alpecka	Laki - Laki	25 tahun < 30 tahun	165 cm < 175 cm	60 kg < 70 kg	Keriting	Jawa	Sawo Matang	Islam
119	Jino	Laki - Laki	25 tahun < 30 tahun	165 cm < 175 cm	60 kg < 70 kg	Lurus	Tionghoa	Kuning Langsat	Kristen Proteste
120	Billy Astian Putra	Laki - Laki	20 tahun < 25 tahun	155 cm < 165 cm	50 kg < 60 kg	Ikall	Lainya	Putih	Islam
121	NJonathan Novandi	Laki - Laki	25 tahun < 30 tahun	165 cm < 175 cm	> 70 kg	Lurus	Tionghoa	Putih	Kristen Proteste
122	Azhar Mareto	Laki - Laki	25 tahun < 30 tahun	165 cm < 175 cm	60 kg < 70 kg	Keriting	Jawa	Sawo Matang	Islam
123	Satria	Laki - Laki	20 tahun < 25 tahun	165 cm < 175 cm	50 kg < 60 kg	Ikall	Jawa	Sawo Matang	Islam
124	Vincentius Daniel	Laki - Laki	25 tahun < 30 tahun	165 cm < 175 cm	60 kg < 70 kg	Lurus	Tionghoa	Putih	Khatolik
125	Anom Harya Wicaksana	Laki - Laki	20 tahun < 25 tahun	165 cm < 175 cm	60 kg < 70 kg	Lurus	Lainya	Kuning Langsat	Hindu
126	Albertus Yeri	Laki - Laki	25 tahun < 30 tahun	165 cm < 175 cm	50 kg < 60 kg	Lurus	Jawa	Sawo Matang	Kristen Proteste
127	Amridio Zulhimi Saifinnul	Laki - Laki	20 tahun < 25 tahun	> 175 cm	60 kg < 70 kg	Lurus	Jawa	Sawo Matang	Islam
128	Ellyia Lestari	Perempuan	20 tahun < 25 tahun	155 cm < 165 cm	60 kg < 70 kg	Keriting	Jawa	Putih	Islam
129	Elok Fatma Anjanwati	Perempuan	20 tahun < 25 tahun	155 cm < 165 cm	50 kg < 60 kg	Ikall	Jawa	Kuning Langsat	Islam
130	Atan Farhan	Laki - Laki	20 tahun < 25 tahun	155 cm < 165 cm	60 kg < 70 kg	Ikall	Jawa	Putih	Islam

Form Responses 1

SURVEY CALON PASANGAN HIDUP IDAMAN (Responses) ☆

File Edit View Insert Format Data Tools Form Add-ons Help All changes saved in Drive

Comments Share

Siapkan isi Nama Lengkap anda?

	B	C	D	E	F	G	H	I	J
1	Siapkan isi Nama Lengkap anda?	Apa Jenis Kelamin anda?	Berapa Umur anda sekarang?	Berapa tinggi badan anda?	Berapa berat badan anda?	Bagaimana jenis rambut?	Berasal dari suku mana?	Apa warna kulit anda?	Apa agama anda?
131	ryan hendy septianto	Laki - Laki	20 tahun < 25 tahun	155 cm < 165 cm	50 kg < 60 kg	Ikall	Jawa	Sawo Matang	Islam
132	Desi Mayuyu Saitama	Perempuan	20 tahun < 25 tahun	155 cm < 165 cm	50 kg < 60 kg	Lurus	Jawa	Putih	Islam
133	Krisdhamara W. P.	Perempuan	20 tahun < 25 tahun	155 cm < 165 cm	< 50 kg	Ikall	Jawa	Kuning Langsat	Islam
134	Mita Sevenfold	Laki - Laki	25 tahun < 30 tahun	165 cm < 175 cm	60 kg < 70 kg	Ikall	Betawi	Kuning Langsat	Islam
135	Bim Bim	Laki - Laki	20 tahun < 25 tahun	165 cm < 175 cm	> 70 kg	Keriting	Sunda	Kuning Langsat	Islam
136	Drun Travian 3	Laki - Laki	25 tahun < 30 tahun	165 cm < 175 cm	50 kg < 60 kg	Ikall	Jawa	Kuning Langsat	Islam
137	ermawati yuningtyas	Perempuan	20 tahun < 25 tahun	< 155 cm	< 50 kg	Lurus	Jawa	Sawo Matang	Islam
138	Navi Dendi	Laki - Laki	25 tahun < 30 tahun	165 cm < 175 cm	50 kg < 60 kg	Ikall	Tionghoa	Putih	Kristen Proteste
139	alex	Laki - Laki	20 tahun < 25 tahun	165 cm < 175 cm	50 kg < 60 kg	Keriting	Jawa	Putih	Islam
140	Adi Yahya	Laki - Laki	20 tahun < 25 tahun	165 cm < 175 cm	50 kg < 60 kg	Ikall	Jawa	Sawo Matang	Islam
141	Kosaki Onodera	Laki - Laki	20 tahun < 25 tahun	165 cm < 175 cm	50 kg < 60 kg	Ikall	Jawa	Sawo Matang	Islam
142	Hans	Laki - Laki	20 tahun < 25 tahun	165 cm < 175 cm	50 kg < 60 kg	Ikall	Jawa	Sawo Matang	Islam
143	Yogi Bayu Avian	Laki - Laki	20 tahun < 25 tahun	165 cm < 175 cm	50 kg < 60 kg	Keriting	Batak	Kuning Langsat	Islam
144	firdausil hawa	Perempuan	20 tahun < 25 tahun	155 cm < 165 cm	< 50 kg	Lurus	Jawa	Putih	Islam
145	Kerang ajaib	Laki - Laki	20 tahun < 25 tahun	165 cm < 175 cm	50 kg < 60 kg	Lurus	Betawi	Putih	Islam
146	septian adi wijaya	Laki - Laki	20 tahun < 25 tahun	165 cm < 175 cm	60 kg < 70 kg	Lurus	Jawa	Sawo Matang	Islam
147	Familia Yulianti	Perempuan	20 tahun < 25 tahun	155 cm < 165 cm	50 kg < 60 kg	Ikall	Jawa	Putih	Islam
148	Rani Andriani	Perempuan	20 tahun < 25 tahun	155 cm < 165 cm	< 50 kg	Ikall	Jawa	Sawo Matang	Islam
149	rahmawati purwantingis	Perempuan	20 tahun < 25 tahun	155 cm < 165 cm	50 kg < 60 kg	Lurus	Jawa	Sawo Matang	Islam
150	adhyta teddy lukman	Laki - Laki	20 tahun < 25 tahun	155 cm < 165 cm	60 kg < 70 kg	Lurus	Jawa	Kuning Langsat	Islam
151	Afni Wiliujeng Setyorini	Perempuan	20 tahun < 25 tahun	155 cm < 165 cm	< 50 kg	Lurus	Jawa	Kuning Langsat	Islam
152	firda ika pratwi	Perempuan	20 tahun < 25 tahun	155 cm < 165 cm	50 kg < 60 kg	Ikall	Jawa	Sawo Matang	Islam
153	Qoirul Kotimah	Perempuan	20 tahun < 25 tahun	< 155 cm	< 50 kg	Ikall	Jawa	Sawo Matang	Islam
154	Azizah Nurya Dianing Pa	Perempuan	20 tahun < 25 tahun	< 155 cm	50 kg < 60 kg	Lurus	Jawa	Sawo Matang	Islam
155	adhi tama	Laki - Laki	20 tahun < 25 tahun	165 cm < 175 cm	60 kg < 70 kg	Ikall	Jawa	Sawo Matang	Islam
156	budiono	Laki - Laki	30 tahun < 35 tahun	165 cm < 175 cm	60 kg < 70 kg	Lurus	Jawa	Kuning Langsat	Islam

Form Responses 1



LAMPIRAN 3 DATA KUISIONER

NAMA	JK	Agama	Umur	Tinggi	Berat	Rambut	Ras	Kulit	Pendidikan	Pekerjaan	Status	Perokok
Data 1	1	0	0	2	2	0	5	1	1	5	0	0
Data 2	0	2	0	1	1	0	5	1	1	5	0	0
Data 3	0	0	0	1	1	2	5	2	0	5	0	0
Data 4	1	0	0	2	2	0	5	1	1	5	0	0
Data 5	1	0	0	2	3	1	5	1	0	2	0	0
Data 6	1	0	0	1	1	0	5	1	0	5	0	0
Data 7	1	0	0	2	3	2	5	1	1	3	0	0
Data 8	1	0	0	1	2	0	8	1	0	2	0	0
Data 9	0	0	0	1	1	2	5	2	0	2	0	0
Data 10	1	0	0	2	2	2	8	1	1	5	0	0
Data 11	1	1	0	2	2	0	5	2	1	5	0	0
Data 12	1	5	1	3	3	0	9	0	0	2	0	1
Data 13	1	0	0	1	1	0	5	0	1	5	0	0
Data 14	1	0	0	3	1	0	5	1	0	1	0	0
Data 15	1	1	0	2	2	0	9	2	1	1	0	0
Data 16	1	0	0	0	0	0	5	1	3	1	0	0
Data 17	1	0	0	2	1	0	5	1	0	5	0	0
Data 18	0	0	0	1	0	2	5	2	1	2	0	0
Data 19	1	0	0	2	2	0	5	1	1	5	0	1
Data 20	1	0	0	2	1	1	5	1	0	5	0	0
Data 21	1	0	0	2	2	0	5	1	1	2	0	0
Data 22	1	0	0	2	2	2	5	1	0	5	0	0
Data 23	1	0	0	2	1	2	6	1	0	5	0	0
Data 24	1	0	0	1	0	0	5	1	0	5	0	0
Data 25	1	0	1	3	2	0	5	1	0	1	0	1
Data 26	1	0	0	2	2	2	5	1	0	5	0	0
Data 27	0	0	0	1	0	1	5	2	0	5	0	0
Data 28	0	0	0	1	2	2	5	1	0	2	0	0
Data 29	1	0	0	2	1	0	5	1	0	5	0	0
Data 30	0	0	0	1	1	0	5	2	0	5	0	0
Data 31	1	0	0	2	1	2	5	1	0	5	0	0
Data 32	1	0	0	3	2	0	5	1	1	5	0	0
Data 33	1	0	0	2	1	0	5	0	0	5	0	0
Data 34	1	0	0	2	1	0	5	1	0	1	0	1
Data 35	0	0	0	1	1	1	5	0	1	1	0	0
Data 36	0	0	0	0	1	2	5	1	0	5	0	0
Data 37	0	0	0	1	1	2	5	0	0	5	0	0



Data 38	0	0	0	1	1	0	5	1	1	5	0	0
Data 39	1	0	0	1	1	0	5	1	0	5	0	0
Data 40	0	0	0	1	0	1	5	2	1	5	0	0
Data 41	1	0	0	2	1	0	5	1	1	5	0	0
Data 42	0	0	0	1	1	2	5	2	0	5	0	0
Data 43	0	0	0	1	0	0	5	1	1	5	0	0
Data 44	0	0	0	1	0	2	5	0	0	5	0	0
Data 45	0	0	0	1	0	1	5	1	3	2	0	0
Data 46	0	0	0	1	1	0	5	1	1	5	0	0
Data 47	0	0	0	0	0	0	5	1	0	5	0	0
Data 48	1	0	0	2	3	0	6	0	0	5	0	0
Data 49	1	0	0	2	3	2	5	1	1	5	0	0
Data 50	0	0	0	0	0	0	5	2	0	5	0	0
Data 51	1	0	0	2	1	2	5	1	0	5	0	0
Data 52	0	0	0	1	1	0	5	2	0	5	0	0
Data 53	1	0	0	2	3	0	5	1	0	5	0	0
Data 54	0	0	0	2	0	1	5	1	0	5	0	0
Data 55	0	0	0	1	1	0	5	1	0	2	0	0
Data 56	0	0	0	1	0	2	5	1	1	5	0	0
Data 57	1	0	0	2	2	0	5	1	0	4	0	1
Data 58	0	0	0	1	0	0	5	0	0	5	0	0
Data 59	1	0	0	3	3	2	5	1	0	5	0	0
Data 60	1	0	0	3	2	0	5	1	1	2	0	0
Data 61	0	0	0	2	1	1	5	2	3	3	0	0
Data 62	1	1	0	2	1	2	1	0	1	5	0	0
Data 63	0	0	0	1	1	0	5	1	0	5	0	0
Data 64	1	0	0	2	1	2	5	0	1	5	0	0
Data 65	0	0	0	1	0	2	5	1	0	5	0	0
Data 66	1	0	0	1	0	1	5	1	1	5	0	0
Data 67	0	0	0	0	2	0	5	2	0	5	0	0
Data 68	0	0	0	1	1	2	5	2	0	5	0	0
Data 69	0	2	0	1	1	0	9	1	3	5	0	0
Data 70	0	0	0	1	1	2	5	1	1	5	0	0
Data 71	1	0	0	2	1	0	8	1	0	5	0	1
Data 72	1	0	1	2	1	1	0	0	1	1	0	0
Data 73	1	0	0	1	3	2	0	1	1	5	0	1
Data 74	0	0	0	1	0	2	5	2	1	5	0	0
Data 75	0	0	1	1	1	1	5	2	1	3	0	0
Data 76	0	0	0	1	0	0	5	0	1	5	0	0



Data 77	0	0	1	2	1	0	5	0	1	1	0	0
Data 78	1	0	0	2	2	0	5	0	1	1	0	0
Data 79	1	0	0	2	2	2	5	1	0	5	0	0
Data 80	1	0	1	2	1	2	5	1	1	0	0	1
Data 81	0	1	1	1	1	0	5	0	1	1	0	0
Data 82	0	0	1	0	1	2	5	1	1	1	0	0
Data 83	1	0	1	2	1	1	3	1	1	1	0	0
Data 84	1	0	0	2	2	2	5	1	1	5	0	1
Data 85	0	0	1	1	0	0	5	2	1	1	0	0
Data 86	1	0	0	1	1	0	5	2	3	1	0	1
Data 87	0	0	1	1	1	2	5	0	1	1	0	0
Data 88	0	0	1	1	1	2	8	2	1	1	0	0
Data 89	0	0	1	1	2	0	5	0	1	0	0	0
Data 90	0	0	0	1	0	1	5	0	1	1	0	0
Data 91	0	0	1	1	1	0	5	0	0	3	0	0
Data 92	1	0	1	2	1	2	5	1	1	2	0	0
Data 93	0	0	0	2	1	2	5	0	0	1	0	0
Data 94	0	0	0	1	1	2	5	1	1	1	0	0
Data 95	0	0	1	1	1	2	5	1	1	1	0	0
Data 96	1	0	0	2	2	2	8	1	1	5	0	0
Data 97	0	0	0	1	0	0	5	1	0	5	0	0
Data 98	1	0	0	2	2	2	5	0	1	2	0	0
Data 99	0	0	1	2	1	2	10	2	1	0	0	0
Data 100	1	0	0	3	3	2	10	1	0	1	0	0
Data 101	0	0	0	0	0	2	5	1	1	5	0	0
Data 102	1	0	0	2	3	0	5	1	1	4	0	0
Data 103	0	0	0	1	1	2	8	1	1	5	0	0
Data 104	1	0	0	2	2	1	5	1	1	5	0	0
Data 105	1	0	0	3	3	0	6	2	1	2	0	0
Data 106	0	0	0	1	0	2	5	1	1	0	0	0
Data 107	0	0	0	1	0	0	5	1	0	5	0	0
Data 108	1	0	2	3	3	1	7	3	2	4	0	0
Data 109	0	1	1	0	1	0	5	0	1	0	0	0
Data 110	0	0	1	1	1	1	5	1	1	1	0	0
Data 111	1	0	0	2	1	2	5	2	1	5	0	0
Data 112	1	0	1	1	1	1	5	1	1	5	0	0
Data 113	0	0	0	1	0	0	5	0	1	5	0	0
Data 114	0	0	0	1	0	2	5	2	0	5	0	0
Data 115	1	0	0	1	3	1	5	1	0	5	0	0

Data 116	1	0	0	2	1	2	5	1	1	5	0	0
Data 117	1	0	1	2	2	1	5	1	1	5	0	0
Data 118	1	1	1	2	2	0	9	2	1	2	0	0
Data 119	1	0	0	1	1	2	10	0	0	5	0	1
Data 120	1	1	1	2	3	0	9	0	3	2	0	0
Data 121	1	0	1	2	2	1	5	1	0	2	0	0
Data 122	1	0	0	2	1	2	5	1	1	5	0	0
Data 123	1	2	1	2	2	0	9	0	1	2	0	1
Data 124	1	3	0	2	2	0	10	2	1	2	0	0
Data 125	1	1	1	2	1	0	5	1	0	2	0	0
Data 126	1	0	0	3	2	0	5	1	0	5	0	0
Data 127	0	0	0	1	2	1	5	0	1	2	0	0
Data 128	0	0	0	1	1	2	5	2	1	5	0	0
Data 129	1	0	0	1	2	2	5	0	1	2	0	0
Data 130	1	0	0	1	1	2	5	1	1	5	0	1
Data 131	0	0	0	1	1	0	5	0	0	5	0	0
Data 132	0	0	0	1	0	2	5	2	0	5	0	0
Data 133	1	0	1	2	2	2	2	2	1	2	0	0
Data 134	1	0	0	2	3	1	8	2	1	2	0	0
Data 135	1	0	1	2	1	2	5	2	1	1	0	0
Data 136	0	0	0	0	0	0	5	1	1	1	0	0
Data 137	1	1	1	2	1	2	9	0	1	5	0	0
Data 138	1	0	0	2	1	1	5	0	0	5	0	0
Data 139	1	0	0	2	1	2	5	1	1	4	0	0
Data 140	1	0	0	2	1	2	5	1	0	5	0	0
Data 141	1	0	0	2	1	2	5	1	1	2	0	0
Data 142	1	0	0	2	1	1	1	2	1	1	0	0
Data 143	0	0	0	1	0	0	5	0	0	1	0	0
Data 144	1	0	0	2	1	0	2	0	1	2	0	0
Data 145	1	0	0	2	2	0	5	1	1	5	0	0
Data 146	0	0	0	1	1	2	5	0	1	1	0	0
Data 147	0	0	0	1	0	2	5	1	0	5	0	0
Data 148	0	0	0	1	1	0	5	1	0	5	0	0
Data 149	1	0	0	1	2	0	5	2	1	2	0	0
Data 150	0	0	0	1	0	0	5	2	1	5	0	0

LAMPIRAN 4 LABELING

Kriteria	Poin	Value
Jenis Kelamin	Perempuan	0
	Laki - laki	1
Umur	20 tahun < 25 tahun	0
	25 tahun < 30 tahun	1
	30 tahun < 35 tahun	2
	> 35 tahun	3
Tinggi	< 155 cm	0
	155 cm < 165 cm	1
	165 cm < 175 cm	2
	> 175 cm	3
Berat	< 155 cm	0
	155 cm < 165 cm	1
	165 cm < 175 cm	2
	> 175 cm	3
Jenis Rambut	Lurus	0
	Keriting	1
	Ikal	2
	Botak	3
Suku	Banjar	0
	Batak	1
	Betawi	2
	Bugis	3
	Dayak	4
	Jawa	5
	Minang	6
	Papua	7
	Sunda	8
	Tionghoa	9
Lainya	10	

Warna Kulit	Putih	0
	Sawo Matang	1
	Kuning Langsat	2
	Hitam	3
Agama	Islam	0
	Kristen Protestan	1
	Khatolik	2
	Hindu	3
	Kong Huchu	4
	Budha	5
Status	Lajang	0
	Cerai	1
	Ditinggal Mati	2
	Pisah Ranjang	3
Perokok	Ya	0
	Tidak	1