



**ANALISIS CLUSTERING ANGKA HARAPAN HIDUP
BERDASARKAN JENIS KELAMIN MENGGUNAKAN
ALGORITMA K-MEANS DI WILAYAH INDONESIA**

LAPORAN TUGAS AKHIR

BAGAS MAULANA IBRAHIM

41520010177

**UNIVERSITAS
MERCU BUANA**

PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS ILMU KOMPUTER

UNIVERSITAS MERCU BUANA

JAKARTA

2024



**ANALISIS CLUSTERING ANGKA HARAPAN HIDUP
BERDASARKAN JENIS KELAMIN MENGGUNAKAN
ALGORITMA K-MEANS DI WILAYAH INDONESIA**

LAPORAN TUGAS AKHIR

BAGAS MAULANA IBRAHIM

41520010177

Diajukan sebagai salah satu syarat untuk memperoleh gelar sarjana

PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS ILMU KOMPUTER

UNIVERSITAS MERCU BUANA

JAKARTA

2024

HALAMAN PENYATAAN KARYA SENDIRI

Saya yang bertanda tangan di bawah ini:

Nama : Bagas Maulana Ibrahim

NIM : 41520010177

Program Studi : Teknik Informatika

Judul Laporan Skripsi : Analisis Clustering Angka Harapan Hidup

Berdasarkan Jenis Kelamin Menggunakan Algoritma K-Means di Wilayah
Indonesia

Menyatakan bahwa Laporan Skripsi ini adalah hasil karya saya sendiri dan bukan plagiat, serta semua sumber baik yang dikutip maupun dirujuk telah saya nyatakan dengan benar. Apabila ternyata ditemukan di dalam Laporan Skripsi saya terdapat unsur plagiat, maka saya siap mendapatkan sanksi akademis yang berlaku di Universitas Mercu Buana.

Jakarta, 14 Juli 2024

UNIVERSITAS
MERCU BUANA



Bagas Maulana Ibrahim

HALAMAN PENGESAHAN

Laporan Skripsi ini diajukan oleh:

Nama : Bagas Maulana Ibrahim
NIM : 41520010177
Program Studi : Teknik Informatika
Judul Laporan Skripsi : Analisis Clustering Angka Harapan Hidup Berdasarkan Jenis Kelamin Menggunakan Algoritma K-Means di Wilayah Indonesia

Telah berhasil dipertahankan pada sidang di hadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Strata 1 pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Disahkan oleh:

Pembimbing : Dr. Ir. Eliyani.
NIDN : 0321026901
Ketua Penguji : Dr. Ir. Eliyani.
NIDN : 0321026901
Penguji 1 : Dwi Anindyani Rochman, S.T., M.TI.
NIDN : 011057801
Penguji 2 : Dr. Muhammad Syaukani, S.T., M.Cs., M.Kom
NIDN : 0317047309



(
(
(
(
(

Jakarta, 30 Juli 2024

Mengetahui,

Dekan

Ketua Program Studi



Dr. Bambang Jekonowo, S.Si, M.T.I

NIDN: 0320037002



Dr. Hadi Santoso, S. Kom, M.Kom

NIDN: 0225067701

KATA PENGANTAR

Puji syukur saya panjatkan kepada Tuhan Yang Maha Esa, karena atas berkat dan rahmat-Nya, saya dapat menyelesaikan Proposal Penelitian ini. Penulisan Proposal Penelitian ini dilakukan da

lam rangka memenuhi salah satu syarat untuk disidangkan pada sidang Proposal Penelitian. Saya menyadari bahwa, tanpa bantuan dan bimbingan dari berbagai pihak, dari masa perkuliahan sampai pada penyusunan Proposal Penelitian ini, sangatlah sulit bagi saya untuk menyelesaikan Proposal Penelitian. Oleh karena itu, saya mengucapkan terima kasih kepada:

1. Prof. Dr. Ir. Andi Adriansyah, M.Eng selaku Rektor Universitas Mercu Buana.
2. Dr. Bambang Jokonowo, S.Si., MTI selaku Dekan Fakultas Ilmu komputer /Direktur ProgramPascasarjana.
3. Dr. Hadi Santoso, S.Kom., M.Kom selaku Ketua Program Studi Teknik Informatika.
4. Dr. Ir. Eliyani. selaku Dosen Pembimbing yang telah menyediakan waktu, tenaga, dan pikiran untuk mengarahkan saya dalam penyusunan Skripsi ini.
5. Dwi Anindyani Rochman, S.T., M.TI. dan Dr. Muhammad Syaukani, S.T., M.Cs., M.Kom., selaku Dosen Penguji Skripsi atas koreksi dan arahan serta masukannya.
6. Keluarga & Teman dekat, yang selalu mendukung saya dalam mengerjakan Skripsi ini.

Akhir kata, saya berharap Tuhan Yang Maha Esa berkenan membalas segala kebaikan semua pihak yang telah membantu. Semoga Proposal Penelitian ini membawa manfaat bagi pengembangan ilmu.

Jakarta, 14 Juli 2024



Bagas Maulana Ibrahim

**HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS
AKHIR UNTUK KEPENTINGAN AKADEMIS**

Sebagai sivitas akademik Universitas Mercu Buana, saya yang bertanda tangan di bawah ini:

Nama : Bagas Maulana Ibrahim
NIM : 41520010177
Program Studi : Teknik Informatika
Judul Proposal Penelitian : Analisis Clustering Angka Harapan Hidup Berdasarkan Jenis Kelamin Menggunakan Algoritma K-Means di Wilayah Indonesia

Demi pengembangan ilmu pengetahuan, dengan ini memberikan izin dan menyetujui untuk memberikan kepada Universitas Mercu Buana **Hak Bebas Royalti Non-Eksklusif (Non-exclusive Royalty-Free Right)** atas karya ilmiah saya yang berjudul di atas beserta perangkat yang ada (jika diperlukan).

Dengan Hak Bebas Royalti Non-Eksklusif ini Universitas Mercu Buana berhak menyimpan, mengalihmedia/format-kan, mengelola dalam bentuk pangkalan data (database), merawat, dan mempublikasikan Laporan Magang/Skripsi/Tesis/Disertasi saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 14 Juli 2024

Yang Menyatakan,



Bagas Maulana Ibrahim

ABSTRAK

Nama : Bagas Maulana Ibrahim
NIM : 41520010177
Program Studi : Teknik Informatika
Judul Proposal Penelitian : Analisis Clustering Angka Harapan Hidup Berdasarkan Jenis Kelamin Menggunakan Algoritma K-Means di Wilayah Indonesia
Pembimbing : Dr. Ir. Eliyani.

Angka harapan hidup merupakan salah satu indikator utama yang digunakan untuk menilai kualitas kesehatan dan kesejahteraan suatu wilayah. Penelitian ini bertujuan untuk menganalisis angka harapan hidup di Indonesia berdasarkan jenis kelamin menggunakan algoritma K-Means. Data yang digunakan mencakup angka harapan hidup dari berbagai kabupaten/kota di Indonesia untuk periode 2017 hingga 2023, yang diperoleh dari website Badan Pusat Statistik (BPS). Metode clustering K-Means digunakan untuk mengelompokkan wilayah berdasarkan kesamaan karakteristik angka harapan hidup. Penelitian ini juga menggunakan tiga metode evaluasi untuk menentukan jumlah kluster optimal, yaitu Elbow Method, Silhouette Score, dan Davies-Bouldin Index. Hasil dari penelitian ini menunjukkan bahwa terdapat tiga kluster utama untuk angka harapan hidup laki-laki dan perempuan. Kluster pertama terdiri dari wilayah-wilayah dengan angka harapan hidup yang tinggi, kluster kedua dengan angka harapan hidup sedang, dan kluster ketiga dengan angka harapan hidup yang rendah.

Kata Kunci: Angka Harapan Hidup, Clustering, K-Means, Elbow Method, Silhouette Score, Davies-Bouldin Index

ABSTRACT

Nama : Bagas Maulana Ibrahim
NIM : 41520010177
Program Studi : Teknik Informatika
Judul Proposal Penelitian : Analisis Clustering Angka Harapan Hidup Berdasarkan Jenis Kelamin Menggunakan Algoritma K-Means di Wilayah Indonesia
Pembimbing : Dr. Ir. Eliyani.

Life expectancy is a key indicator used to assess the quality of health and welfare in a region. This study aims to analyze life expectancy in Indonesia based on gender using the K-Means algorithm. The data used includes life expectancy figures from various regencies/cities in Indonesia for the period 2017 to 2023, obtained from the official website of the Central Statistics Agency (BPS). The K-Means clustering method was employed to group regions based on similarities in life expectancy characteristics. This study also used three evaluation methods to determine the optimal number of clusters, namely the Elbow Method, Silhouette Score, and Davies-Bouldin Index. The results of this study indicate that there are three main clusters for male and female life expectancy. The first cluster consists of regions with high life expectancy, the second cluster with medium life expectancy, and the third cluster with low life expectancy.

Keywords: Life Expectancy, Clustering, K-Means, Elbow Method, Silhouette Score, Davies-Bouldin Index

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PENYATAAN KARYA SENDIRI	ii
HALAMAN PENGESAHAN	iii
KATA PENGANTAR.....	iv
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS	v
ABSTRAK	vi
ABSTRACT	vii
DAFTAR ISI.....	viii
DAFTAR TABEL	x
DAFTAR GAMBAR.....	xi
DAFTAR LAMPIRAN.....	xiii
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang	1
1.2. Perumusan Masalah.....	3
1.3. Tujuan Penelitian.....	3
1.4. Manfaat Penelitian.....	3
1.5. Batasan Penelitian	4
1.6. Sistematika Penulisan.....	4
BAB II TINJAUAN PUSTAKA.....	6
2.1. Penelitian Terdahulu.....	6
2.2. Teori Pendukung	19
BAB III METODE PENELITIAN	22
3.1. Jenis Penelitian	22

3.2. Tahapan Penelitian	23
BAB IV HASIL DAN PEMBAHASAN	29
4.1. Pengumpulan Data	29
4.2. Preprocessing Data	30
4.3. Persiapan Data dan Standarisasi	39
4.4. Modeling Algoritma K-Means	40
4.5. Tahap Clustering dan Hasil Clustering	48
4.6. Analisis Visualisasi Angka Harapan Hidup Tertinggi dan Terendah	66
BAB V KESIMPULAN DAN SARAN	69
5.1. Kesimpulan.....	69
5.2. Saran.....	70
DAFTAR PUSTAKA	72
LAMPIRAN.....	75



DAFTAR TABEL

Table 1. Penelitian Terkait	6
-----------------------------------	---



DAFTAR GAMBAR

Gambar 1. Data AHH Tahun 2022 - 2023	1
Gambar 2. Proses Clustering K-Means	20
Gambar 3. Penentuan jumlah kluster dengan Elbow Method	20
Gambar 4. Penentuan jumlah kluster dengan Silhouette.....	21
Gambar 5. Penentuan jumlah kluster dengan Davies-Bouldin Index	21
Gambar 6. Rumus Interquartile Range.....	24
Gambar 7. Rumus Batas Bawah dan Batas Atas	24
Gambar 8. Rumus Within-Cluster Sum of Squares	26
Gambar 9. Rumus Silhouette Score	27
Gambar 10. Rumus Davis-Bouldin Index.....	27
Gambar 11. Dataset Asli	29
Gambar 12. Data Format Wide	30
Gambar 13. Data Format Long	31
Gambar 14. Data Statistik Deskriptif.....	31
Gambar 15. Code plot Histogram dan Boxplot.....	32
Gambar 16. Distribusi Angka Harapan Hidup.....	33
Gambar 17. Perbandingan AHH berdasarkan Jenis Kelamin	34
Gambar 18. Menampilkan Informasi DataFrame.....	35
Gambar 19. Mengecek Missng Values	36
Gambar 20. Code Menghitung IQR.....	36
Gambar 21. Hasil Data Outlier.....	37
Gambar 22. Code Menghapus Outlier	37
Gambar 23. Code Mereduksi Data.....	38
Gambar 24. Code Menghitung rata-rata AHH.....	39
Gambar 25. Hasil Data setelah Preprocessing	39
Gambar 26. Code Persiapan dan Standarisasi Data	40
Gambar 27. Code Modelling Data	41
Gambar 28. Code Evaluasi Modeling Laki-laki	43
Gambar 29. Visualisasi Evaluasi Modeling Laki-laki	43
Gambar 30. Code Evaluasi Modeling Perempuan	45

Gambar 31. Visualisasi Evaluasi Modeling Perempuan	45
Gambar 32. Code Evaluasi Modeling Keseluruhan	47
Gambar 33. Visualisasi Evaluasi Modeling Keseluruhan.....	47
Gambar 34. Tahap Clustering dan Penambahan Hasil ke Dataframe	49
Gambar 35. Code Menghasilkan Clustering Tabel Laki-laki	50
Gambar 36. Tabel Hasil Clustering Laki-laki	50
Gambar 37. Code Menghasilkan Visual Clustering Laki-laki.....	52
Gambar 38. Visualisasi Clustering Laki-laki	53
Gambar 39. Code Menghasilkan Clustering Tabel Perempuan	55
Gambar 40. Tabel Hasil Clustering Perempuan.....	55
Gambar 41. Code Menghasilkan Visual Clustering Perempuan.....	57
Gambar 42. Visualisasi Clustering Perempuan.....	58
Gambar 43. Code Menghasilkan Clustering Tabel Keseluruhan.....	60
Gambar 44. Tabel Hasil Clustering Keseluruhan	60
Gambar 45. Code Menghasilkan Visual Clustering Keseluruhan	62
Gambar 46. Visualisasi Clustering Keseluruhan	63
Gambar 47. Visualisasi Clustering Keseluruhan dengan Nilai $K = 2$	65
Gambar 48. Kota dengan AHH Tertinggi dan Terendah Laki-laki	66
Gambar 49. Kota dengan AHH Tertinggi dan Terendah Perempuan	67
Gambar 50. Kota dengan AHH Tertinggi dan Terendah Keseluruhan	68

UNIVERSITAS
MERCU BUANA

DAFTAR LAMPIRAN

Lampiran 1 Kartu Asistensi Tugas Akhir	75
Lampiran 2 Curriculum Vitae	76
Lampiran 3 Surat Pernyataan HAKI.....	77
Lampiran 4 Sertifikat BNSP	79
Lampiran 5 Hasil Cek Turnitin	80
Lampiran 6 Lembar Persetujuan	81
Lampiran 7 Form Revisi Dosen Penguji 1	82
Lampiran 8 Form Revisi Dosen Penguji 2.....	84

