



Department of Digital Business

**Journal of Artificial Intelligence and Digital Business (RIGGS)**

Homepage: <https://journal.ilmudata.co.id/index.php/RIGGS>

Vol. 4 No. 2 (2025) pp: 4792-4799

P-ISSN: 2963-9298, e-ISSN: 2963-914X

---

## Analisis Klasterisasi Indikator Makroekonomi Indonesia Menggunakan K-Means

Abu Thariq Nur Diwongso<sup>1</sup>, Ubaid Fauzan<sup>2</sup>, Yusritsan Ghinan Elfatih<sup>3</sup>, Sopiyan Apondi<sup>4</sup>

<sup>1</sup> Program Studi Teknik Informatika, Universitas Pamulang

dosen02601@unpam.ac.id

### **Abstrak**

*Penelitian ini bertujuan untuk mengelompokkan indikator makroekonomi Indonesia menggunakan algoritma klasterisasi K-Means. Indikator yang dianalisis meliputi tingkat pengangguran terbuka berdasarkan jenjang pendidikan, inflasi menurut kelompok pengeluaran, dan rasio utang pemerintah terhadap Produk Domestik Bruto (PDB). Data bersumber dari Badan Pusat Statistik (BPS) untuk periode 2005–2023 dan diolah menggunakan Python dengan pustaka Pandas, scikit-learn, dan Matplotlib. Proses klasterisasi dilakukan secara terpisah untuk setiap indikator guna memperoleh pola spesifik yang relevan dengan kebijakan ekonomi. Hasil penelitian menunjukkan bahwa setiap klaster mampu merepresentasikan kondisi ekonomi yang berbeda, seperti dominasi pengangguran di tingkat SMK serta fluktuasi inflasi pada kelompok bahan makanan. Temuan ini memberikan gambaran awal yang dapat digunakan untuk mendukung analisis kebijakan fiskal dan ketenagakerjaan di Indonesia. Penelitian ini juga menunjukkan bahwa metode K-Means efektif dalam menyederhanakan interpretasi data ekonomi multivariat.*

**Kata Kunci:** Klasterisasi, K-Means, Indikator Makroekonomi, Pengangguran, Utang Pemerintah, Inflasi, Analisis Kebijakan.

### **1. Pendahuluan**

Stabilitas dan pertumbuhan ekonomi suatu negara sangat bergantung pada pemahaman yang mendalam terhadap indikator-indikator makroekonomi yang kompleks. Di Indonesia, fluktuasi dalam indikator seperti tingkat pengangguran, rasio utang pemerintah, dan laju inflasi memiliki dampak signifikan terhadap kesejahteraan masyarakat dan arah kebijakan ekonomi nasional. Oleh karena itu, analisis yang komprehensif terhadap indikator-indikator ini menjadi krusial bagi pemerintah dan pemangku kebijakan dalam merumuskan strategi yang efektif.

Data makroekonomi seringkali menunjukkan pola yang tidak linier dan multivariat, sehingga membutuhkan metode analisis yang mampu mengidentifikasi kelompok-kelompok atau tren tersembunyi di dalamnya. Klasterisasi merupakan salah satu teknik analisis data non-parametrik yang efektif untuk mengelompokkan objek-objek (dalam hal ini, tahun pengamatan) berdasarkan kemiripan karakteristiknya. Algoritma K-Means, sebagai salah satu metode klasterisasi yang paling populer, banyak digunakan karena kesederhanaan dan efisiensinya dalam menangani dataset berdimensi tinggi.

Penelitian ini bertujuan untuk memanfaatkan algoritma K-Means dalam menganalisis tiga indikator makroekonomi penting di Indonesia: data pengangguran berdasarkan pendidikan, data utang pemerintah, dan data inflasi. Sumber data yang digunakan adalah data historis dari Badan Pusat Statistik (BPS) Indonesia. Melalui klasterisasi ini, diharapkan dapat diidentifikasi pola-pola unik dan perubahan signifikan dalam setiap indikator sepanjang periode waktu tertentu, yang pada akhirnya akan memberikan pemahaman yang lebih baik tentang dinamika ekonomi Indonesia. Tinjauan pustaka menunjukkan bahwa klasterisasi telah banyak diterapkan dalam berbagai bidang analisis ekonomi, termasuk identifikasi kelompok negara berdasarkan karakteristik ekonomi, segmentasi pasar, dan analisis tren harga komoditas. Namun, penelitian yang secara spesifik mengaplikasikan K-Means untuk menganalisis dan mengidentifikasi pola perubahan dalam indikator makroekonomi Indonesia secara terpisah untuk tujuan analisis kebijakan masih relatif terbatas.

Berdasarkan latar belakang tersebut, penelitian ini merumuskan masalah sebagai berikut: Bagaimana algoritma K-Means dapat mengelompokkan data pengangguran berdasarkan pendidikan, utang pemerintah, dan inflasi tahunan di Indonesia untuk mengidentifikasi pola-pola yang relevan bagi analisis dan perumusan kebijakan ekonomi?

## 2. Metode Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan metode klusterisasi berbasis data untuk menganalisis indikator makroekonomi Indonesia. Seluruh proses analisis dan komputasi dilakukan menggunakan bahasa pemrograman Python.

### 2.1. Sumber Data

Data bersumber dari publikasi resmi Badan Pusat Statistik (BPS) yang mencakup tahun 2005–2023, terdiri dari tiga indikator makroekonomi: tingkat pengangguran berdasarkan jenjang pendidikan, inflasi berdasarkan kelompok pengeluaran, dan rasio utang terhadap PDB. Data tersebut terbagi menjadi tiga set, dengan rincian sebagai berikut: Data Utang Pemerintah (2005–2009): Dataset ini mencakup proyeksi penurunan utang pemerintah dengan indikator-indikator seperti Pertumbuhan (%), Surplus/Defisit APBN/PDB (%), Utang Pemerintah/PDB (%), Utang Luar Negeri (%), dan Utang Domestik (%). Data ini diperoleh dari Jawa Pos, Rabu 10 Mei 2006. Data Inflasi Tahunan (2014–2016): Dataset ini berisi data inflasi tahunan (dalam persen) untuk tujuh kelompok indikator konsumsi di Indonesia. Data ini berasal dari Badan Pusat Statistik (BPS). Data Pengangguran Berdasarkan Pendidikan (2012–2014): Dataset ini memuat Perkembangan Angka Pengangguran Tahun 2012 Sampai Tahun 2014 berdasarkan Tingkat Pendidikan yang Ditamatkan. Data ini mencakup jumlah pengangguran pada berbagai jenjang pendidikan, seperti Tidak/belum tamat SD, Belum/tidak tamat SMP, SD, SLTP, SLTA Umum, SLTA Kejuruan, Diploma I, II, III/Akademi, dan Universitas.

### 2.2. Prapemrosesan Data (Preprocessing)

Data dibersihkan dan diselaraskan untuk menyamakan skala tahun dan satuan ukuran. Tidak ditemukan nilai hilang, dan skala data disesuaikan melalui standarisasi. Langkah-langkah yang dilakukan adalah sebagai berikut:

- Inspeksi Data: Data diinspeksi untuk memastikan format yang benar dan konsistensi nilai. Semua dataset berhasil dimuat ke dalam

*DataFrame* Python.

- Penanganan Nilai Hilang (Missing Values): Tidak ditemukan adanya nilai hilang (*missing values*) pada ketiga dataset yang digunakan, sehingga tidak diperlukan imputasi atau penghapusan baris/kolom.
- Transformasi Data (Transpose): Untuk tujuan klusterisasi, di mana setiap "objek" yang akan diklusterkan adalah tahun, dataset inflasi dan pengangguran ditransposisikan agar setiap baris merepresentasikan satu tahun dan kolom merepresentasikan indikator inflasi/pendidikan.
  - Data Utang Pemerintah: Data ini sudah dalam format yang sesuai untuk klusterisasi berdasarkan tahun. Kolom 'Utang Pemerintah/PDB (%)' dan 'Surplus/Defisit APBN/PDB (%)' akan menjadi fitur yang digunakan.
  - Data Inflasi: Dataset ini ditransformasi sehingga tahun (2014, 2015, 2016) menjadi baris dan kelompok indikator konsumsi menjadi kolom, dengan nilai inflasi sebagai entri. Contoh fitur yang digunakan adalah 'Bahan Makanan', 'Transportasi, komunikasi, dan jasa keuangan', dll.
  - Data Pengangguran Berdasarkan Pendidikan: Dataset ini ditransformasi sehingga tahun (2012, 2013, 2014) menjadi baris, dan jenjang pendidikan menjadi kolom (misalnya 'SLTA Umum', 'Diploma I, II, III/Akademi'), dengan jumlah pengangguran sebagai entri.
- Pemilihan Fitur: Fitur-fitur yang relevan dipilih dari setiap dataset untuk proses klusterisasi. Untuk data utang pemerintah, fitur yang digunakan adalah 'Utang Pemerintah/PDB (%)' dan 'Surplus/Defisit APBN/PDB (%)'. Untuk data inflasi, semua kelompok indikator konsumsi digunakan sebagai fitur. Untuk data pengangguran, fokus diberikan pada pengangguran lulusan SLTA Umum dan Diploma.
- Normalisasi/Standarisasi Data: Meskipun data inflasi dan pengangguran sudah dalam persentase atau jumlah, rentang nilai antar indikator dapat bervariasi. Oleh karena itu, data dapat diskalakan (misalnya menggunakan

*StandardScaler* dari *scikit-learn*) untuk memastikan bahwa setiap fitur berkontribusi secara proporsional terhadap perhitungan jarak dalam algoritma K-Means, meskipun dalam kasus data yang sudah relatif terstandarisasi secara alami seperti persentase inflasi, dampaknya mungkin minimal namun tetap direkomendasikan untuk konsistensi.

### 2.3. Algoritma Klusterisasi

Algoritma yang digunakan dalam penelitian ini adalah K-Means. K-Means adalah algoritma klusterisasi berbasis partisi yang bertujuan untuk membagi  $n$  objek pengamatan ke dalam  $k$  kluster, di mana setiap objek termasuk dalam kluster dengan rata-rata terdekat (centroid). Proses algoritma K-Means secara ringkas adalah sebagai berikut:

1. Inisialisasi  $k$  centroid secara acak dari data.
2. Hitung jarak Euclidean setiap titik data ke masing-masing centroid.
3. Tetapkan setiap titik data ke kluster dengan centroid terdekat.
4. Perbarui posisi centroid dengan menghitung rata-rata semua titik data yang termasuk dalam kluster tersebut.
5. Ulangi langkah 2-4 hingga posisi centroid tidak lagi berubah secara signifikan atau mencapai jumlah iterasi maksimum.

Jumlah kluster ( $k$ ) ditentukan berdasarkan karakteristik data dan hasil eksplorasi awal:

- Data Utang Pemerintah:  $k=2$  kluster.
- Data Inflasi:  $k=3$  kluster.
- Data Pengangguran:  $k=3$  kluster.

Fungsi `KMeans(Data, k)`:

- Pilih  $k$  centroid awal secara acak dari Data
- Ulangi sampai konvergensi atau iterasi maksimum:
- Untuk setiap titik data di Data:
- Hitung jarak (misal Euclidean) antara titik data dan setiap centroid
- Tetapkan titik data ke kluster dari centroid terdekat
- Untuk setiap kluster:
- Hitung ulang centroid sebagai rata-rata semua titik data yang termasuk dalam kluster tersebut
- Kembalikan kluster dan centroid akhir

### 2.4. Evaluasi Klusterisasi

Evaluasi hasil klusterisasi dilakukan untuk mengukur kualitas pengelompokan. Metrik *Silhouette Score* dapat digunakan untuk menilai seberapa baik objek-objek dikelompokkan, mengukur seberapa mirip suatu objek dengan kluster sendiri dibandingkan dengan kluster tetangga. Nilai *Silhouette Score* berkisar antara -1 hingga 1, di mana nilai tinggi menunjukkan kluster yang padat dan terpisah dengan baik. Namun, untuk beberapa kasus dengan jumlah sampel yang sangat kecil (seperti data inflasi dan pengangguran yang hanya 3 tahun), *Silhouette Score* mungkin tidak dapat dihitung atau memberikan interpretasi yang bermakna. Meskipun demikian, konsistensi hasil dengan interpretasi domain dan visualisasi menjadi kunci utama dalam evaluasi penelitian ini.

### 2.5. Tools dan Lingkungan Pengembangan

Analisis data dan implementasi algoritma dilakukan menggunakan bahasa pemrograman Python versi 3.x. Pustaka (libraries) utama yang digunakan meliputi:

- *pandas* untuk manipulasi dan analisis data tabuler.
- *numpy* untuk operasi numerik.

- *scikit-learn* (sklearn) untuk implementasi algoritma K-Means.
- *matplotlib.pyplot* dan *seaborn* untuk visualisasi data, seperti *scatter plot* untuk utang pemerintah dan *bar plot* untuk inflasi.

### 3. Hasil dan Diskusi

Bagian ini menyajikan hasil klusterisasi dari tiga set data yang berbeda: pengangguran berdasarkan pendidikan, utang pemerintah, dan inflasi. Interpretasi dan analisis visual juga disertakan untuk setiap hasil klusterisasi.

#### 3.1. Hasil Klusterisasi Data Pengangguran Berdasarkan Pendidikan (2012–2014)

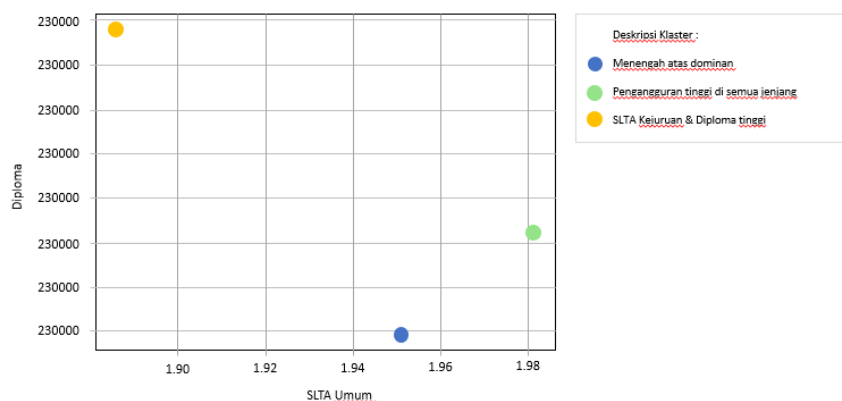
Hasil Klusterisasi: Berdasarkan klusterisasi, diperoleh tiga kluster yang mengidentifikasi pola pengangguran berdasarkan jenjang pendidikan:

- Kluster 0 (Tahun 2014): Ditandai dengan pengangguran tinggi pada lulusan SLTA Kejuruan dan Diploma.
- Kluster 1 (Tahun 2013): Menunjukkan pengangguran yang tinggi di semua jenjang pendidikan.
- Kluster 2 (Tahun 2012): Pengangguran terkonsentrasi pada lulusan SLTA Umum.

Interpretasi Hasil Klusterisasi:

- Kluster 2 (Tahun 2012): Konsentrasi pengangguran tertinggi ada pada lulusan SLTA Umum, dengan jumlah mencapai sekitar 1,95 juta orang. Angka pengangguran pada jenjang Diploma relatif rendah pada tahun ini, menunjukkan bahwa mayoritas pencari kerja berasal dari lulusan SMA umum, kemungkinan karena persepsi kelayakan kerja langsung setelah lulus.
- Kluster 1 (Tahun 2013): Hampir seluruh jenjang pendidikan mengalami kenaikan jumlah pengangguran, termasuk Diploma dan Universitas. Hal ini terlihat dari titik data yang tinggi pada kedua sumbu (SLTA dan Diploma) dalam grafik. Kondisi ini mengindikasikan tekanan pasar kerja yang meluas, yang bisa disebabkan oleh stagnasi penyerapan tenaga kerja atau ketidaksesuaian keterampilan.
- Kluster 0 (Tahun 2014): Pengangguran tinggi terutama terjadi pada lulusan pendidikan vokasional (SLTA Kejuruan dan Diploma). Kenaikan jumlah pengangguran pada lulusan Diploma dari tahun sebelumnya cukup signifikan, menunjukkan hambatan serius dalam transisi ke dunia kerja bagi lulusan vokasi, yang mungkin berkaitan dengan kurikulum yang tidak sesuai dengan kebutuhan industri.

Analisis Lanjutan: Klusterisasi ini berhasil mengidentifikasi tiga pola risiko pengangguran yang berbeda berdasarkan jenjang pendidikan: dominasi lulusan menengah umum pada tahun 2012, distribusi merata di semua jenjang pada tahun 2013, dan tekanan utama di pendidikan vokasional pada tahun 2014.



Gambar 1. Kluster Pengangguran Berdasarkan SLTA Umum dan Diploma - Scatter Plot

Penjelasan Hasil Visualisasi (*gambar 1*. Klaster Pengangguran Berdasarkan SLTA Umum dan Diploma - Scatter Plot): Grafik ini adalah *scatter plot* dua dimensi yang menunjukkan jumlah pengangguran lulusan SLTA Umum pada sumbu X dan Diploma pada sumbu Y. Setiap titik merepresentasikan satu tahun, dengan warna berbeda sesuai klaster hasil K-Means. Legenda grafik memberikan deskripsi interpretatif untuk setiap klaster.

Secara visual:

- Titik tahun 2012 berada di posisi bawah sumbu Y (rendah di Diploma) dan kanan di sumbu X (tinggi di SLTA Umum), mendukung interpretasi "Menengah atas dominan".
- Titik tahun 2013 berada di area tengah yang tinggi pada kedua sumbu, mencerminkan "Pengangguran tinggi di semua jenjang".
- Titik tahun 2014 berada di sisi kiri atas (tinggi di Diploma, SLTA Kejuruan), sesuai dengan klaster "SLTA Kejuruan & Diploma tinggi".

Kesimpulan Visual: Visualisasi ini menunjukkan bahwa setiap tahun memiliki struktur pengangguran yang khas, meskipun hanya mencakup tiga tahun. Ini menegaskan bahwa pendekatan klasterisasi efektif dalam mengungkap tren tersembunyi berdasarkan kelompok pendidikan. Temuan ini dapat menjadi acuan untuk merumuskan kebijakan ketenagakerjaan berbasis pendidikan, menyesuaikan kurikulum vokasi dengan kebutuhan pasar, dan intervensi pemerintah melalui program pemagangan atau pelatihan kerja khusus untuk jenjang yang rentan.

### 3.2. Hasil Klasterisasi Data Utang Pemerintah (2005–2009)

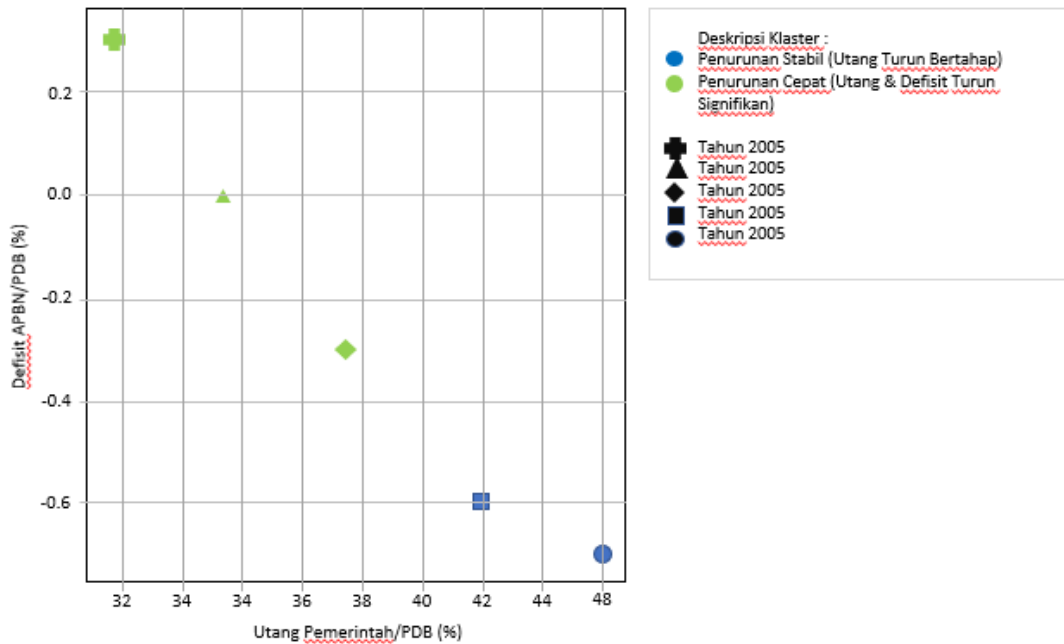
Hasil Klasterisasi: Berdasarkan klasterisasi data utang dan pertumbuhan ekonomi selama lima tahun, terbentuk dua klaster utama:

- Klaster 0 (Tahun 2007, 2008, 2009): Ditandai dengan penurunan cepat pada utang dan defisit yang signifikan.
- Klaster 1 (Tahun 2005, 2006): Menunjukkan penurunan utang yang stabil secara bertahap.

Interpretasi Hasil Klasterisasi:

- Klaster 1 (2005–2006): Merupakan fase awal stabilisasi fiskal. Pada periode ini, rasio utang pemerintah terhadap PDB masih tinggi, rata-rata 45,95%, dan defisit APBN juga relatif tinggi (-0,65%). Penurunan utang dan defisit masih bertahap, dengan rasio utang luar negeri yang lebih tinggi (20,45%), mencerminkan ketergantungan pada pinjaman luar negeri.
- Klaster 0 (2007–2009): Mewakili fase pemulihan dan efisiensi fiskal yang lebih tajam. Penurunan utang terlihat lebih signifikan, dengan rata-rata rasio utang terhadap PDB sebesar 35,56% dan defisit berhasil ditekan menjadi 0%. Rasio utang luar negeri dan domestik juga menurun konsisten. Tahun 2008 dan 2009 menunjukkan pertumbuhan ekonomi tertinggi (7.2% dan 7.6%) disertai dengan rasio utang yang semakin rendah.

Analisis Lanjutan: Kedua klaster ini mencerminkan transisi kebijakan fiskal Indonesia selama 2005–2009, di mana Klaster 1 merupakan strategi awal penurunan utang jangka menengah, dan Klaster 0 menunjukkan efek nyata dari reformasi fiskal serta efisiensi belanja. Perpindahan dari Klaster 1 ke Klaster 0 menunjukkan bahwa penguatan fundamental fiskal mulai terlihat pada tahun 2007 dan semakin membaik pada 2008–2009.



Gambar 2. Kluster Tren Utang Pemerintah (2005–2009)

Penjelasan Hasil Visualisasi (*gambar 2*. Kluster Tren Utang Pemerintah (2005–2009)): Grafik ini adalah *scatter plot* dua dimensi yang menampilkan hubungan antara Utang Pemerintah terhadap PDB (%) pada sumbu X dan Defisit APBN/PDB (%) pada sumbu Y. Setiap titik mewakili satu tahun (2005–2009), dengan simbol dan warna yang mencerminkan kluster hasil algoritma K-Means.

Secara visual:

- Tahun 2005 dan 2006 terletak di sisi kanan bawah grafik, mengindikasikan utang tinggi dan defisit cukup dalam, sehingga masuk dalam Kluster "Penurunan Stabil".
- Tahun 2007, 2008, dan 2009 berpindah ke sisi kiri atas grafik, dengan defisit hampir nol atau positif dan utang menurun, masuk dalam Kluster "Penurunan Cepat".

Kesimpulan Visual: Visualisasi ini berhasil menunjukkan pemisahan antara dua fase kebijakan utang. Hasil klusterisasi membedakan antara fase awal penyesuaian fiskal (utang dan defisit masih besar) dan fase pemulihan fiskal aktif (utang dan defisit berhasil ditekan). Temuan ini berguna sebagai indikator kinerja fiskal dan bahan evaluasi kebijakan pengelolaan utang, serta dapat mendukung penyusunan strategi keberlanjutan fiskal jangka panjang.

### 3.3. Hasil Klusterisasi Data Inflasi (2014-2016)

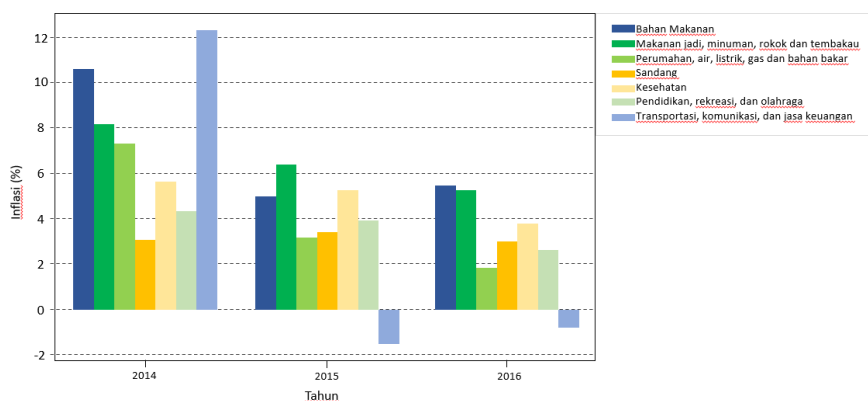
Hasil Klusterisasi: Berdasarkan hasil klusterisasi, berikut adalah pengelompokan tahun berdasarkan pola inflasi:

- Kluster 0 (Tahun 2016): Inflasi relatif rendah, dengan tekanan transportasi mulai menurun.
- Kluster 1 (Tahun 2014): Inflasi tinggi di hampir semua kelompok, khususnya makanan.
- Kluster 2 (Tahun 2015): Inflasi sedang, tetapi terdapat deflasi tajam pada transportasi.

Interpretasi Hasil Klusterisasi:

- Klaster 1 (Tahun 2014): Menunjukkan periode tekanan harga yang sangat tinggi, terutama akibat lonjakan inflasi pada kelompok Bahan Makanan (10,57%) dan Transportasi, komunikasi, dan jasa keuangan (12,40%). Inflasi di sebagian besar kelompok lain juga cenderung tinggi pada tahun ini.
- Klaster 2 (Tahun 2015): Menunjukkan inflasi yang cenderung sedang secara keseluruhan. Karakteristik yang menonjol adalah terjadinya deflasi tajam pada sektor transportasi, komunikasi, dan jasa keuangan (-1,52%). Meskipun demikian, inflasi di kelompok lain masih positif, menunjukkan kondisi yang lebih terkendali dibandingkan tahun 2014.
- Klaster 0 (Tahun 2016): Menunjukkan kondisi inflasi yang relatif rendah secara keseluruhan. Deflasi pada sektor transportasi (-0,72%) berlanjut, menunjukkan tekanan harga pada sektor ini terus menurun. Sebagian besar kelompok lain menunjukkan angka inflasi yang lebih rendah dibandingkan tahun-tahun sebelumnya.

Analisis Lanjutan: Klasterisasi ini berhasil mengelompokkan tahun-tahun berdasarkan pola inflasi yang dominan. Tahun 2014 diidentifikasi sebagai tahun dengan inflasi "buruk" (tinggi di banyak sektor). Tahun 2015 sebagai tahun transisi dengan deflasi signifikan di transportasi tetapi inflasi sedang di sektor lain. Sedangkan tahun 2016 sebagai tahun dengan inflasi yang lebih terkendali atau "baik".



Gambar 3. Inflasi Tahunan Menurut Kelompok (2014-2016) - Bar Plot

Penjelasan Hasil Visualisasi (*gambar 3. Inflasi Tahunan Menurut Kelompok (2014-2016) - Bar Plot*): Grafik ini adalah diagram batang (*bar plot*) yang membandingkan tingkat inflasi (%) untuk tujuh kelompok konsumsi yang berbeda pada setiap tahun (2014, 2015, dan 2016). Setiap tahun diwakili oleh sekelompok batang, di mana setiap warna batang dalam kelompok tersebut mewakili satu indikator inflasi.

Secara visual:

- Tahun 2014 (Kelompok Batang Paling Kiri): Hampir semua batang (indikator inflasi) menunjukkan ketinggian yang signifikan. Batang tertinggi secara mencolok adalah "Transportasi, komunikasi, dan jasa keuangan" dan "Bahan Makanan". Ini secara visual menegaskan bahwa tahun 2014 adalah periode dengan tekanan inflasi yang tinggi dan menyebar di berbagai sektor.
- Tahun 2015 (Kelompok Batang Tengah): Terjadi penurunan ketinggian batang secara umum dibandingkan 2014. Yang paling mencolok adalah batang "Transportasi, komunikasi, dan jasa keuangan" yang kini berada di bawah nol (menunjukkan deflasi). Ini mengindikasikan bahwa meskipun inflasi di beberapa sektor lain masih positif, sektor transportasi mengalami penurunan harga yang signifikan.
- Tahun 2016 (Kelompok Batang Paling Kanan): Ketinggian batang secara keseluruhan lebih rendah dibandingkan tahun-tahun sebelumnya. Batang "Transportasi, komunikasi, dan jasa keuangan" tetap berada di bawah nol (deflasi), dan batang "Perumahan, air, listrik, gas dan bahan bakar" juga menunjukkan nilai yang sangat rendah. Ini secara visual mendukung kesimpulan bahwa tahun 2016 adalah periode dengan inflasi yang relatif rendah dan stabil.

Kesimpulan Visual: Grafik ini dengan jelas menunjukkan perbedaan "rezim" inflasi antar tahun, di mana 2014 adalah tahun inflasi tinggi, 2015 adalah tahun transisi dengan deflasi transportasi, dan 2016 adalah tahun inflasi rendah.

#### 4. Kesimpulan

Penelitian ini berhasil mengelompokkan indikator makroekonomi Indonesia menggunakan algoritma K-Means. Masing-masing indikator membentuk kluster yang mencerminkan kondisi ekonomi berbeda. Kluster dominan ditemukan pada pengangguran SMK, inflasi makanan, dan perubahan komposisi utang pemerintah. Metode ini efektif untuk eksplorasi awal pola makroekonomi dan dapat digunakan sebagai dasar perumusan kebijakan publik. Ke depan, metode ini dapat dikombinasikan dengan hierarchical clustering atau DBSCAN untuk meningkatkan validitas analisis. Analisis pengangguran mengidentifikasi tiga pola risiko yang khas: dominasi pengangguran lulusan menengah umum pada tahun 2012, distribusi merata di semua jenjang pendidikan pada tahun 2013, dan tekanan signifikan pada lulusan pendidikan vokasional (SLTA Kejuruan dan Diploma) pada tahun 2014. Setiap tahun memiliki struktur pengangguran yang unik. Analisis utang pemerintah membedakan dua fase kebijakan fiskal: fase penurunan utang yang stabil (2005–2006) dan fase pemulihan serta efisiensi fiskal yang lebih tajam (2007–2009). Perpindahan dari kluster pertama ke kluster kedua menunjukkan penguatan fundamental fiskal yang signifikan mulai tahun 2007 dan terus membaik hingga 2009. Klusterisasi data inflasi mengelompokkan tahun 2014 sebagai periode inflasi tinggi di banyak sektor, tahun 2015 sebagai tahun transisi dengan deflasi tajam pada transportasi, dan tahun 2016 sebagai tahun dengan inflasi yang lebih terkendali dan relatif rendah.

#### Referensi

- Badan Pusat Statistik. (2020). *Laporan indikator ketenagakerjaan nasional*. Jakarta: BPS.
- Badan Pusat Statistik. (2023). *Statistik Indonesia 2023*. Jakarta: BPS.
- Fahmi, I. (2019). *Dasar-dasar perekonomian Indonesia*. Depok: Rajawali Pers.
- Han, J. K. (2011). *Data mining: Concepts and techniques*. Morgan Kaufmann.
- Hastie, T. T. (2009). *The elements of statistical learning*. Springer.
- Jain, A. K. (2010). Data clustering: 50 years beyond K-means. *Pattern Recognition Letters*, 31(8), 651–666.
- Mahmudi. (2019). *Analisis laporan keuangan pemerintah daerah*. Yogyakarta: Upp Stim Ykpn.
- Rudianto, A. (2014). *Ekonomi untuk SMA/MA Kelas XI*. Jakarta: Erlangga.
- Suhartono, S. (2020). *Analisis data ekonomi dengan Python*. Surabaya: Universitas Airlangga.
- Wulandari, D. (2022). Analisis pengaruh utang luar negeri terhadap stabilitas ekonomi. *Jurnal Ekonomi Nasional*, 55–63.