

Analisis Kluster dan Prediksi Indikator Ekonomi Makro Indonesia Menggunakan Algoritma *K-Means* dan Model Regresi

Ahmad Basith¹, Arfan Yusuf², Chaerunisa³, Dimas Satrio Wicaksono⁴, Nathan Sanjaya⁵, Syaeful Machfud⁶

^{1,2,3,4,5,6} Universitas Pamulang, Indonesia

E-mail: dsatrio268@gmail.com

Article History:

Received: 24 Juni 2025

Revised: 10 September 2025

Accepted: 23 September 2025

Keywords: *K-Means, Regresi Linier, Pengelompokan Data, Indikator Makroekonomi, Pendapatan Negara.*

Abstract: *Penelitian ini bertujuan untuk menganalisis pola dan memprediksi perkembangan indikator ekonomi makro Indonesia dengan pendekatan penambangan data. Dua metode, yaitu algoritma K-Means Clustering untuk pengelompokan data dan Regresi Linier, memprediksi tren masa depan. Data yang dianalisis meliputi neraca perdagangan Indonesia (ekspor dan impor), penerimaan negara (perpajakan dan PNBPN), peran jenis penggunaan Produk Domestik Bruto (PDB), data Anggaran Pendapatan dan Belanja Daerah (APBD) per kabupaten/kota di Provinsi Daerah Istimewa Yogyakarta, serta tren ekspor industri nonmigas. Hasil klusterisasi menunjukkan segmentasi waktu atau wilayah dengan karakteristik ekonomi yang berbeda, seperti kelompok tahun dengan surplus atau defisit neraca perdagangan. Model regresi menunjukkan tren kenaikan yang konsisten pada sektor-sektor tertentu, seperti konsumsi rumah tangga dan penerimaan pajak. Evaluasi dilakukan dengan menggunakan metrik inersia, skor siluet, dan Indeks Davies-Bouldin, yang menunjukkan kualitas kluster yang baik. Temuan ini diharapkan dapat menjadi dasar bagi perencanaan kebijakan fiskal dan strategi pembangunan ekonomi yang terarah, terutama pada sektor-sektor yang berkontribusi signifikan terhadap pertumbuhan nasional dan daerah.*

PENDAHULUAN

Kondisi ekonomi makro memiliki peran yang vital dalam menentukan kestabilan dan arah perkembangan suatu negara (Zulaikah, 2025). Di Indonesia, pemantauan indikator ekonomi makro seperti neraca perdagangan, Produk Domestik Bruto (PDB), inflasi, dan pengangguran penting untuk mengukur kinerja ekonomi secara keseluruhan. Setiap perubahan dalam indikator-indikator ini tidak hanya mencerminkan kondisi ekonomi saat ini, tetapi memberikan gambaran mengenai potensi pertumbuhan atau risiko krisis di masa depan. Indonesia, dengan tantangan globalnya, seperti ketergantungan pada ekspor komoditas, fluktuasi harga energi, serta dampak dari krisis keuangan global (Hermawan dkk., 2025), ini memerlukan pemahaman yang mendalam terhadap hubungan antar-indikator ekonomi makro untuk merancang kebijakan yang lebih tepat sasaran.

Menghadapi tantangan tersebut, semakin banyak negara yang memanfaatkan pendekatan

berbasis data untuk mampu merancang kebijakan ekonomi. Pada ranah Indonesia, penerapan metode analisis data semakin relevan, mengingat ketersediaan data ekonomi yang besar dan beragam. Teknik analisis berbasis data, seperti *data mining* ini, memberikan kemampuan untuk mengeksplorasi data secara lebih mendalam dan menemukan pola-pola yang tersembunyi, yang tidak bisa dilihat dengan analisis tradisional (Syam dkk., 2024). Salah satu dari metode yang telah terbukti efektif dalam hal ini adalah penggunaan *K-Means Clustering* untuk pengelompokan data dan regresi linier untuk memprediksi tren ekonomi di masa depan. Dengan menggabungkan kedua pendekatan, diharapkan diperoleh wawasan komprehensif tentang kondisi ekonomi Indonesia.

K-Means Clustering ini adalah algoritma memungkinkan pengelompokan data berdasarkan kesamaan karakteristik (Hendrastuty, 2024), yang berguna untuk bisa mengidentifikasi pola atau kelompok dalam data ekonomi yang mungkin menunjukkan hubungan tertentu. Dalam ranah ekonomi makro, *K-Means* sendiri dapat digunakan untuk mengelompokkan wilayah atau sektor ekonomi yang berdasarkan indikator-indikator seperti tingkat pendapatan, belanja pemerintah, atau neraca perdagangan. Dengan pemahaman yang lebih baik mengenai struktur ekonomi ini, pembuat kebijakan dapat merancang intervensi yang lebih tepat sasaran, misalnya dengan meningkatkan investasi di sektor-sektor yang mengalami penurunan kinerja atau mendorong ekspor dari wilayah-wilayah yang belum optimal (Suparmoko, 1997).

Selain itu, regresi linier memberikan alat yang kuat untuk memprediksi nilai-nilai ekonomi di masa depan berdasarkan tren historis yang ada (Wilda dkk., 2024). Dengan menggunakan data historis mengenai variabel-variabel ekonomi makro, regresi linier membantu dalam memodelkan hubungan antara berbagai faktor ekonomi dan mengidentifikasi kemungkinan arah perkembangan yang akan datang. Dalam penelitian ini, regresi linier akan digunakan untuk memproyeksikan indikator-indikator ekonomi makro Indonesia, dengan fokus pada variabel-variabel kunci seperti neraca perdagangan, PDB, dan anggaran daerah. Dengan menggabungkan kedua metode ini, maka diharapkan penelitian ini yang dapat memberikan wawasan yang lebih dalam mengenai dinamika ekonomi Indonesia serta memfasilitasi perencanaan kebijakan yang lebih berbasis data dan lebih responsif terhadap perubahan kondisi ekonomi.

LANDASAN TEORI

K-Means Clustering adalah salah satu dari teknik dalam *data mining* yang digunakan untuk mengelompokkan data ke dalam beberapa kelompok berdasarkan kemiripan karakteristik antar data (Hendrastuty, 2024). Algoritma ini yang berguna untuk menemukan pola dalam data yang kompleks dan besar tanpa memerlukan informasi sebelumnya tentang struktur data tersebut. Dalam konteks ekonomi makro, *K-Means Clustering* dapat digunakan untuk mengidentifikasi kelompok-kelompok wilayah atau sektor dengan karakteristik serupa, seperti pola-pola neraca perdagangan, tingkat pendapatan daerah, dan belanja pemerintah. Pengelompokan ini memungkinkan peneliti dan pembuat kebijakan untuk menganalisis dan memahami perbedaan antar kelompok ekonomi yang mungkin tidak tampak jelas pada analisis konvensional.

Sementara itu, regresi linier merupakan metode statistik yang digunakan untuk memodelkan hubungan antara variabel independen (penyebab) dan variabel dependen (hasil). Dalam analisis ekonomi, regresi linier digunakan untuk memprediksi nilai-nilai variabel ekonomi di masa depan berdasarkan pola yang terlihat pada data historis. Model regresi linier sederhana mengasumsikan adanya hubungan linear antara dua variabel, sedangkan regresi linier berganda mengakomodasi lebih dari dua variabel (Priyatno, 2023). Dalam penelitian ini, regresi linier akan digunakan untuk memprediksi perkembangan indikator ekonomi makro Indonesia, yang seperti

neraca perdagangan, pendapatan negara, dan pengeluaran pemerintah, berdasarkan data historis yang ada. Model ini berguna untuk meramalkan tren ekonomi dan memberikan gambaran perubahan satu variabel ekonomi dapat mempengaruhi variabel lainnya.

Selain itu, teori tentang ekonomi makro juga memberikan dasar penting bagi penelitian ini, terutama dalam pemahaman terhadap hubungan antara indikator-indikator ekonomi yang saling berinteraksi (Huda, 2018). Indikator seperti Produk Domestik Bruto (PDB), neraca perdagangan, tingkat inflasi, dan belanja pemerintah daerah memiliki pengaruh signifikan terhadap stabilitas ekonomi negara. Teori ekonomi makro menyatakan fluktuasi dalam indikator-indikator ini dapat mempengaruhi keputusan investasi, kebijakan fiskal, serta daya beli masyarakat. Dalam penelitian ini, indikator ekonomi yang dianalisis akan digunakan untuk mengidentifikasi pola dan hubungan antar variabel ekonomi, serta memprediksi perkembangan ekonomi Indonesia di masa depan.

METODE PENELITIAN

Metode penelitian ini terdiri dari beberapa tahapan yang sistematis untuk mengumpulkan, memproses, dan menganalisis data ekonomi makro Indonesia. Pada bagian pertama, yaitu sumber data, penelitian ini menggunakan lima sumber utama yang diperoleh dari buku referensi dan dokumen resmi pemerintah. Data yang digunakan mencakup neraca perdagangan (ekspor, impor, dan selisih) yang diambil dari buku Statistik Teori dan Aplikasi (Supranto, 2000), data pendapatan negara yang diperoleh dari buku Hukum Pajak Indonesia, serta data struktur penggunaan Produk Domestik Bruto (PDB) yang diambil dari buku Keuangan Negara dalam Teori dan Praktik. Selain itu, data sektor ekonomi per tahun dari buku Ekonomi Industri, dan data Anggaran Pendapatan dan Belanja Daerah (APBD) per kabupaten/kota di Provinsi Daerah Istimewa Yogyakarta didapat dari buku Dinamika Pengawasan Dana Otonomi Khusus dan Istimewa.

Pada tahap pra-pemrosesan data, dilakukan beberapa langkah untuk memastikan bahwa data yang digunakan bersih dan siap untuk analisis. Tahapan ini mencakup penghapusan baris kosong dan entri total yang tidak relevan dengan analisis (`df = df[df['Uraian'] != 'Total']`), serta mengubah format data dari bentuk lebar ke panjang menggunakan fungsi `pd.melt`. Selain itu, angka yang terpisah oleh koma digabungkan menjadi satu nilai numerik yang utuh, dan kolom tahun diubah menjadi format integer. Langkah terakhir dalam tahap ini adalah normalisasi fitur menggunakan `StandardScaler()` memastikan skala antar variabel seimbang sebelum dilakukan proses klusterisasi. Tahapan *pre-processing* untuk memastikan bahwa analisis yang dilakukan tidak terdistorsi data yang tidak konsisten atau tidak relevan. Contoh kode normalisasi:

```
from sklearn.preprocessing import StandardScaler
scaler = StandardScaler()
X_scaled = scaler.fit_transform(df[features])
```

Metode utama yang digunakan dalam penelitian ini adalah algoritma *K-Means Clustering* dan *Linear Regression*. Untuk klusterisasi, algoritma *K-Means* digunakan untuk mengelompokkan data berdasarkan kesamaan karakteristik. Klusterisasi ini dilakukan pada indikator ekonomi seperti neraca perdagangan (ekspor-impor), pendapatan negara berdasar kategori penerimaan, penggunaan PDB berdasarkan jenis konsumsi, dan APBD antar wilayah.

Evaluasi klaster dilakukan dengan beberapa metrik seperti *inertia*, yang mengukur kompaknya klaster, *silhouette score* mengukur pemisahan antar klaster, dan *Davies-Bouldin Index* ini yang mengevaluasi kualitas pembentukan klaster. Contoh kode implementasi *K-Means*:

```
from sklearn.cluster import KMeans
kmeans = KMeans(n_clusters=3, random_state=42, n_init=10)
clusters = kmeans.fit_predict(scaled_features_csv)
df_csv["Cluster"] = clusters
```

Kode implementasi *K-Means* ini yang mencakup langkah di dalam *Python*, memungkinkan peneliti untuk memvisualisasikan dan mengevaluasi hasil klasterisasi. Sedangkan untuk mampu memprediksi tren masa depan, *linear regression* yang digunakan untuk memodelkan hubungan antara variabel dan menghasilkan proyeksi nilai indikator ekonomi, seperti total penerimaan negara dan sektor-sektor PDB antara tahun 2017 hingga 2020. Data historis dari tahun 2012 hingga 2016 digunakan untuk melatih model ini. Contoh kode:

```
from sklearn.linear_model import LinearRegression
model = LinearRegression()
model.fit(df[['Tahun']], df['Nilai'])
prediksi = model.predict(tahun_prediksi)
```

Terakhir, pada tahap evaluasi model ini, yang dilakukan penilaian untuk mengukur seberapa baik model *K-Means* dan regresi linier bekerja. Untuk klasterisasi, evaluasi menggunakan *inertia* (semakin rendah semakin baik), *silhouette score* (nilai > 0.5 menunjukkan klaster yang baik), dan *Davies-Bouldin Index* (semakin kecil semakin baik). Contoh kode:

```
from sklearn.metrics import silhouette_score, davies_bouldin_score
sil_score = silhouette_score(X_scaled, df['Cluster'])
db_index = davies_bouldin_score(X_scaled, df['Cluster'])
```

Untuk regresi linier ini, model dievaluasi dengan membandingkan tren aktual dan prediksi melalui visualisasi grafik garis. Selain itu, untuk metrik tambahan seperti nilai R^2 , *Mean Squared Error* (MSE), atau *Mean Absolute Error* (MAE) ini yang dapat digunakan untuk mengukur akurasi model. Semua analisis ini dilakukan dengan menggunakan *Python 3.10* dan berbagai pustaka seperti *pandas* dan *numpy* untuk pengolahan data, *matplotlib* dan *seaborn* untuk visualisasi, serta *sklearn* untuk implementasi algoritma *K-Means* dan *Linear Regression*. Editor yang digunakan adalah *Visual Studio Code*.

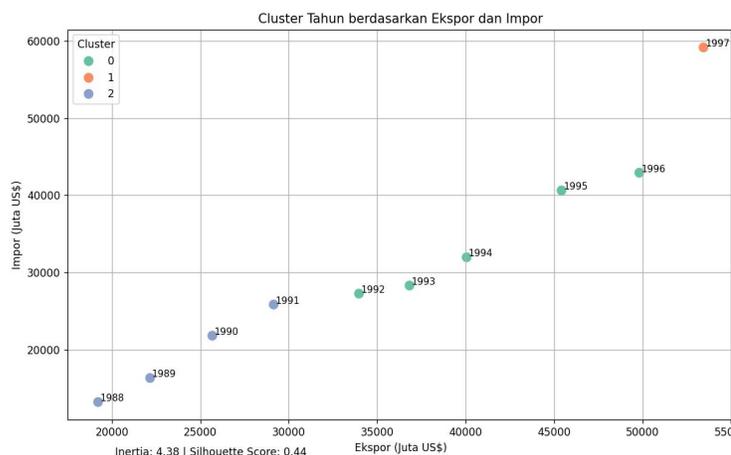
HASIL DAN PEMBAHASAN

1. Klasterisasi Neraca Perdagangan

Data neraca perdagangan Indonesia, yang mencakup nilai ekspor, impor, dan selisih neraca perdagangan, dianalisis menggunakan algoritma *K-Means* untuk membentuk klaster berdasarkan kinerja ekspor-impor setiap tahun. Tujuan dari klasterisasi ini adalah mengelompokkan tahun-tahun tersebut berdasarkan kesamaan karakteristik perdagangan luar negeri, khususnya untuk mengidentifikasi apakah suatu tahun menunjukkan surplus atau defisit. Proses klasterisasi ini

bertujuan untuk memberikan gambaran yang lebih jelas tentang tren perdagangan Indonesia dalam periode waktu yang berbeda, serta untuk membantu pembuat kebijakan dalam merumuskan strategi ekonomi yang lebih efektif sesuai kondisi pasar yang berbeda-beda di setiap periode.

Dalam hal ini, *K-Means Clustering* yang mengelompokkan tahun-tahun dengan pola serupa dalam hal nilai ekspor dan impor, yang memungkinkan analisis lebih lanjut tentang kinerja perdagangan Indonesia. Hasil dari klasterisasi ini menghasilkan tiga klaster utama: Klaster 0, yang terdiri dari tahun-tahun dengan nilai ekspor dan impor relatif rendah, mencerminkan aktivitas perdagangan yang terbatas, seperti yang terjadi pada periode 1990–1993; Klaster 1, yang berisi tahun-tahun dengan nilai ekspor dan impor sedang, contohnya tahun 1994–1996; dan Klaster 2, yang mana mencakup tahun dengan aktivitas ekspor-impor tinggi dan neraca perdagangan yang menonjol, seperti yang terlihat pada tahun 1997. Hasil klasterisasi yang ini dapat divisualisasikan dalam Gambar 1.



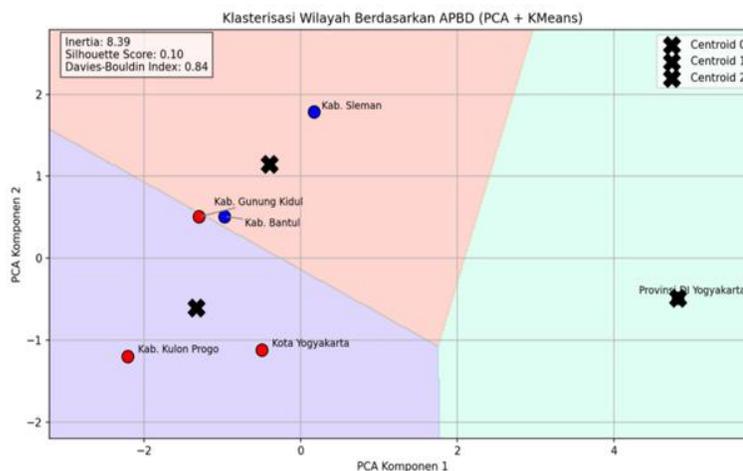
Gambar 1. Hasil Klasterisasi Berdasarkan Tahun Ekspor dan Impor

Klaster 0 menggambarkan periode dengan aktivitas perdagangan yang relatif terbatas, yang mungkin menunjukkan ketergantungan yang lebih rendah terhadap perdagangan internasional atau kondisi ekonomi global tidak menguntungkan. Klaster 1 menunjukkan periode dengan aktivitas perdagangan yang moderat, yang mungkin mencerminkan fase transisi perekonomian Indonesia di mana perdagangan luar negeri yang mengalami pertumbuhan yang stabil namun tidak signifikan. Sementara itu, Klaster 2 menonjolkan tahun dengan tingkat ekspor-impor yang sangat tinggi.

2. Klasterisasi APBD Kabupaten/Kota Yogyakarta

Analisis klasterisasi dilakukan terhadap data Anggaran Pendapatan dan Belanja Daerah (APBD) per kabupaten/kota di Provinsi Daerah Istimewa Yogyakarta (DIY) mengelompokkan wilayah berdasarkan kemiripan pola belanja dan pendapatan. Tujuan utama dari analisis ini adalah mengidentifikasi disparitas fiskal antar daerah yang dapat membantu dalam merancang kebijakan yang lebih tepat sasaran. Dalam analisis ini, digunakan algoritma *K-Means* dengan dukungan reduksi dimensi melalui PCA (*Principal Component Analysis*) untuk memudahkan pemahaman terhadap pola data yang kompleks. Hasil klasterisasi menunjukkan tiga klaster utama. Klaster 0 terdiri kabupaten/kota dengan nilai APBD yang rendah, seperti Kulon Progo dan Kota Yogyakarta. Klaster 1 mencakup wilayah dengan belanja sedang dan struktur APBD yang moderat, seperti Gunung Kidul dan Bantul. Sementara itu, Klaster 2 berisi wilayah APBD relatif tinggi, seperti Kabupaten Sleman dan Provinsi DIY. Hasil klasterisasi ini divisualisasikan

Gambar 2.

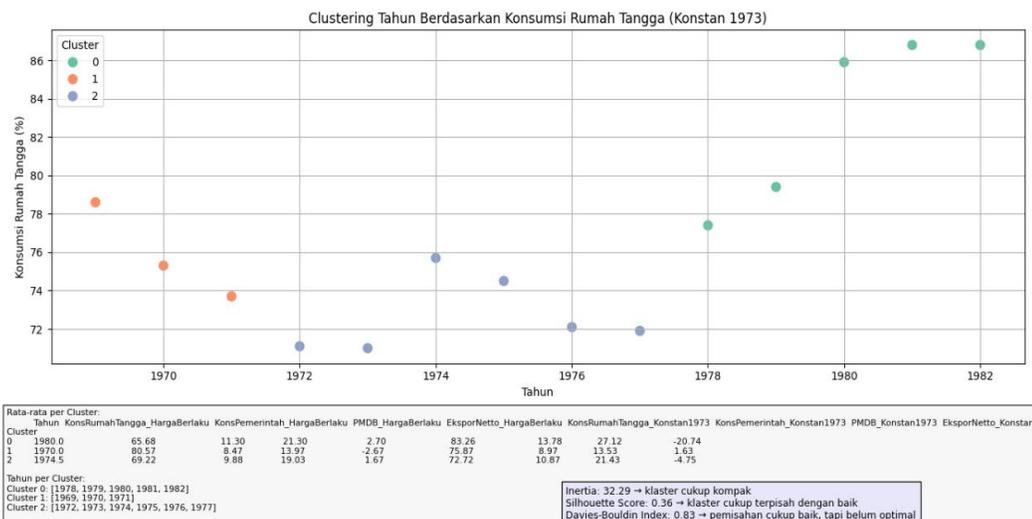


Gambar 2. Hasil Klasterisasi Berdasarkan APBD Wilayah Kab/Kota Yogyakarta

Hasil klasterisasi menunjukkan disparitas yang jelas dalam distribusi APBD di Provinsi DIY. Klaster 0, mencakup kabupaten/kota dengan nilai APBD rendah, seperti Kulon Progo dan Kota Yogyakarta, mungkin mencerminkan daerah keterbatasan sumber daya fiskal atau ketergantungan yang tinggi terhadap dana transfer dari pusat. Sementara Klaster 1, yang mencakup wilayah dengan belanja sedang, seperti Gunung Kidul dan Bantul, menunjukkan wilayah memiliki pengelolaan fiskal yang lebih moderat, di mana belanja daerah masih berada pada tingkat yang cukup tetapi tidak berlebihan. Klaster 2, mencakup wilayah dengan APBD tinggi, seperti Kabupaten Sleman dan Provinsi DIY, mungkin menggambarkan daerah kapasitas fiskal lebih baik, memungkinkan mereka untuk melakukan investasi dan pengeluaran publik yang lebih besar. Evaluasi klasterisasi menunjukkan bahwa meskipun klaster cukup kompak (dengan *inertia* 8.39), pemisahan antar klaster masih perlu ditingkatkan, seperti yang terlihat dari nilai *Silhouette Score* yang relatif rendah (0.10). Ini menunjukkan meskipun ada perbedaan antar klaster, ada tumpang tindih antar beberapa wilayah yang perlu diperhatikan. *Davies-Bouldin Index* yang bernilai 0.84 menunjukkan kualitas klaster cukup baik, meskipun masih ada ruang untuk perbaikan agar pemisahan antar klaster dapat lebih tajam dan lebih representatif. Hasil ini memberikan gambaran lebih jelas ketimpangan fiskal antar kabupaten/kota di Provinsi DIY.

3. Klasterisasi Penggunaan PDB

Data struktur penggunaan Produk Domestik Bruto (PDB) Indonesia pada tahun 1969–1982 dianalisis untuk mengidentifikasi pola konsumsi dan belanja berdasarkan komponen utama seperti konsumsi rumah tangga, konsumsi pemerintah, dan ekspor neto. Untuk mengelompokkan tahun-tahun berdasarkan kesamaan pola ekonomi, algoritma *K-Means* digunakan. Hasil klasterisasi ini membentuk tiga klaster utama. Klaster 0 mencakup tahun-tahun dengan konsumsi rumah tangga tinggi dan konsumsi pemerintah menengah, yang terdiri dari tahun 1978, 1979, 1980, 1981, dan 1982. Klaster 1 ini mewakili periode awal (1969, 1970, 1971), di mana konsumsi rumah tangga dan konsumsi pemerintah relatif rendah. Sedangkan Klaster 2 mencakup tahun 1972 hingga 1977, yang ditandai dengan tingkat konsumsi rumah tangga yang lebih rendah dan pola belanja yang bervariasi. Hasil klasterisasi ini divisualisasikan dalam Gambar 3.

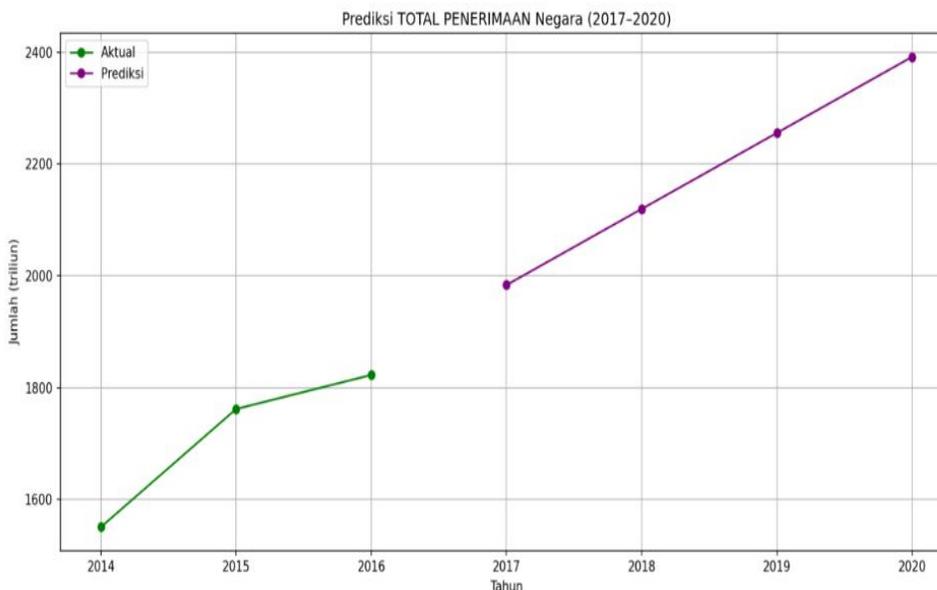


Gambar 3. Hasil Klasterisasi Berdasarkan Konsumsi Rumah Tangga

Hasil klasterisasi ini memberikan gambaran yang jelas mengenai perbedaan pola ekonomi Indonesia selama periode 1969–1982. Klaster 0, yang terdiri dari tahun-tahun dengan konsumsi rumah tangga tinggi dan konsumsi pemerintah menengah, mencerminkan periode di mana ekonomi Indonesia mengalami peningkatan konsumsi domestik cukup signifikan, mungkin dipengaruhi oleh stabilitas ekonomi dan kebijakan fiskal yang mendukung konsumsi rumah tangga dan belanja pemerintah. Klaster 1, yang mewakili periode awal dengan konsumsi rumah tangga dan pemerintah yang relatif rendah, mungkin menggambarkan masa awal setelah Indonesia merdeka, di mana kondisi ekonomi masih dalam tahap pengembangan dan belanja publik belum terlalu tinggi. Klaster 2 menunjukkan periode antara 1972 dan 1977, di mana meskipun konsumsi rumah tangga lebih rendah, terdapat variasi pola belanja yang mencerminkan ketidakstabilan ekonomi atau perubahan kebijakan ekonomi pada masa itu. Evaluasi klasterisasi menunjukkan bahwa klasterisasi cukup kompak dengan nilai *inertia* 32.29, yang berarti bahwa data dikelompokkan dengan cukup baik. Nilai *Silhouette Score* sebesar 0.36 menunjukkan pemisahan antar klaster cukup baik, meskipun masih ada ruang untuk perbaikan. Nilai *Davies-Bouldin Index* yang sebesar 0.83 menunjukkan bahwa kualitas klasterisasi cukup baik, meskipun masih bisa dioptimalkan agar pemisahan antar klaster lebih jelas dan representatif.

4. Prediksi Pendapatan Negara

Analisis terhadap data aktual menunjukkan bahwa penerimaan negara Indonesia mengalami tren pertumbuhan positif selama periode tahun 2014 hingga 2016 (Sekretariat Negara Republik Indonesia, 2016). Pada tahun 2014, total penerimaan negara yang tercatat sebesar 1.550,49 triliun rupiah dan mengalami peningkatan signifikan pada tahun 2015 menjadi 1.761,64 triliun rupiah, dengan laju pertumbuhan sebesar 13,62%. Namun demikian, pada tahun 2016, laju pertumbuhan melambat menjadi 3,46%, dengan total penerimaan mencapai 1.822,54 triliun rupiah.



Gambar 4. Hasil Prediksi Pendapatan Negara

Berdasarkan model regresi linier yang dibangun dari data historis tersebut, maka dilakukan proyeksi penerimaan negara untuk periode tahun 2017 hingga 2020. Hasil prediksi menunjukkan tren pertumbuhan yang masih stabil, meskipun dengan laju yang cenderung menurun secara bertahap setiap tahunnya. Prediksi untuk tahun 2017 ini yang menunjukkan nilai sebesar 1.983,61 triliun rupiah, atau meningkat 8,84% dibanding tahun sebelumnya. Tahun 2018 diproyeksikan mencapai 2.119,63 triliun rupiah (tumbuh 6,86%), diikuti oleh 2.255,66 triliun rupiah pada tahun 2019 (tumbuh 6,42%), dan juga 2.391,68 triliun rupiah pada tahun 2020 (tumbuh 6,03%).

Tahun	Pendapatan Negara (Triliun IDR)	Persentase Perubahan Tahunan
2014	1.550,49	—
2015	1.761,64	13,62%
2016	1.822,54	3,46%
2017	1.940,40	6,46%
2018	2.165,11	11,53%
2019	1.960,64	-9,45%
2020	1.633,60	-16,68%

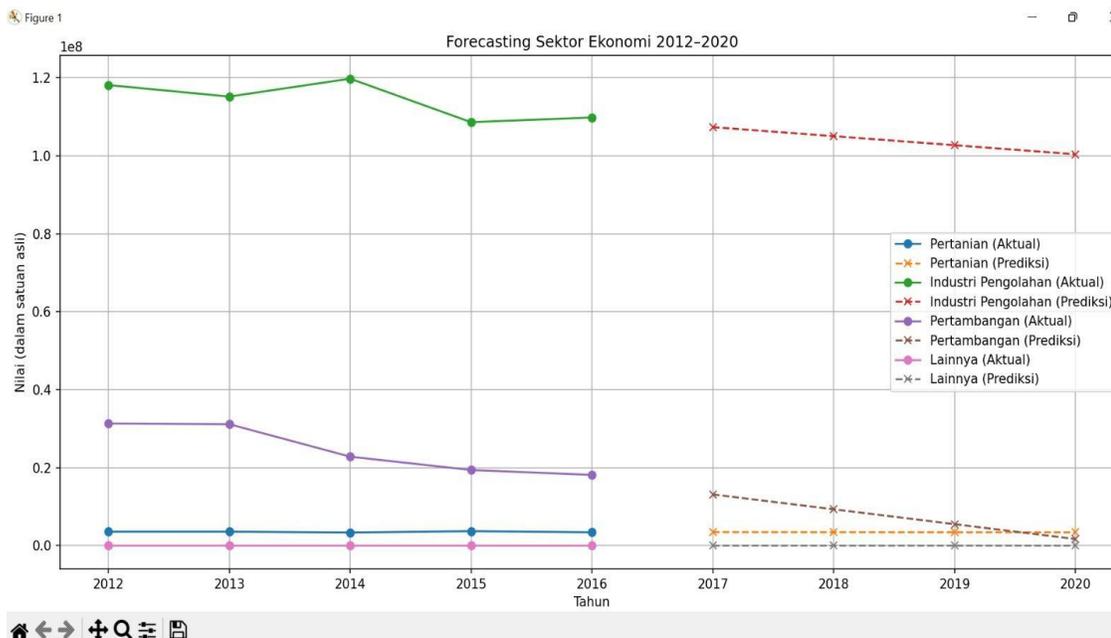
Gambar 5. Laporan Keuangan Pemerintahan Pusat (LKPP)

Meskipun model regresi linier yang menunjukkan performa yang cukup baik pada periode awal prediksi, hasil evaluasi terhadap realisasi penerimaan negara pada tahun 2017 hingga 2020 mengindikasikan adanya ketidaktepatan yang signifikan, khususnya pada dua tahun terakhir. Pada tahun 2017, hasil prediksi sebesar 1.983,61 triliun rupiah hanya meleset sekitar 2,18% dari realisasi yang tercatat sebesar 1.940,40 triliun rupiah. Meskipun terdapat sedikit deviasi, tren pertumbuhan sebesar 6,46% masih selaras dengan proyeksi optimis model. Tahun 2018 bahkan menunjukkan realisasi yang mana melebihi ekspektasi, yakni sebesar 2.165,11 triliun rupiah, atau 2,14% di atas prediksi. Ini mencerminkan kinerja ekonomi yang lebih kuat dibandingkan asumsi dasar model. Namun, pada tahun 2019 dan 2020, prediksi gagal menangkap perubahan arah ekonomi secara drastis. Prediksi tahun 2019 sebesar 2.255,66 triliun rupiah meleset jauh dari realisasi 1.960,64 triliun rupiah, mencerminkan deviasi negatif sebesar -13,06%. Keadaan semakin memburuk pada tahun 2020, di mana realisasi penerimaan hanya mencapai 1.633,60 triliun rupiah, jauh di bawah proyeksi 2.391,68 triliun rupiah, atau deviasi sebesar -31,68%. Kondisi ini menunjukkan bahwa model prediksi linier gagal mengantisipasi adanya *shock* ekonomi besar. Tahun 2019 ditandai dengan perlambatan ekonomi global, tekanan terhadap penerimaan perpajakan, dan meningkatnya ketegangan perdagangan internasional. Sementara itu, tahun 2020 mengalami *shock* ekonomi besar akibat pandemi Covid-19 yang berdampak pada hampir seluruh sektor ekonomi nasional, termasuk penurunan aktivitas usaha, pelonggaran pajak, serta pengalihan anggaran negara untuk penanganan krisis kesehatan.

Hasil analisis terhadap penerimaan negara menunjukkan adanya tren pertumbuhan yang stabil antara 2014 hingga 2016, tetapi model prediksi menggunakan regresi linier tidak sepenuhnya dapat menangkap perubahan mendalam dalam perekonomian pada tahun-tahun berikutnya. Pada tahun 2017 dan 2018, prediksi yang dihasilkan model cukup akurat dan bahkan sedikit melebihi ekspektasi, mencerminkan stabilitas dan kinerja ekonomi yang lebih baik dari yang diperkirakan. Namun, pada tahun 2019 dan 2020, prediksi meleset jauh dari realisasi. Hal ini mengindikasikan bahwa model regresi linier yang digunakan tidak mampu mengakomodasi faktor eksternal yang besar, seperti krisis ekonomi global pada 2019 dan dampak luar biasa dari pandemi Covid-19 pada 2020. Kejadian-kejadian ini menyebabkan penurunan penerimaan negara yang signifikan, yang jauh di bawah proyeksi. Prediksi ini yang tidak mampu mengantisipasi dampak besar dari krisis kesehatan global dan gejolak ekonomi global menekankan perlunya pendekatan prediksi yang lebih fleksibel dan dapat memasukkan faktor-faktor yang tak terduga atau kejutan ekonomi. Sehingga, untuk prediksi jangka panjang yang lebih akurat, model lebih kompleks dan mampu menangani variabel eksternal perlu dipertimbangkan.

5. Klasterisasi Sektor Ekonomi

Data sektor ekonomi Indonesia dari tahun 2012 hingga 2016 digunakan sebagai data latih untuk membangun model regresi linier dalam memprediksi kontribusi tiap sektor terhadap Produk Domestik Bruto (PDB) hingga tahun 2020. Sektor ini yang dianalisis meliputi pertanian, industri pengolahan, pertambangan, dan sektor lainnya. Hasil prediksi menunjukkan beberapa pola yang signifikan. Sektor industri pengolahan mengalami tren penurunan kontribusi meski tetap menjadi penyumbang utama PDB. Sektor pertambangan, merupakan salah satu sektor penting Indonesia, menunjukkan penurunan cukup tajam dari tahun ke tahun, mencerminkan kemungkinan penurunan produksi dan harga komoditas. Sementara itu, sektor pertanian dan sektor lainnya relatif stabil, namun tidak menunjukkan peningkatan signifikan dalam kontribusinya terhadap PDB Indonesia. Visualisasi prediksi untuk masing-masing sektor dilihat pada Gambar 6., yang menggambarkan perubahan kontribusi sektor-sektor tersebut hingga tahun 2020.



Gambar 5. Laporan Keuangan Pemerintahan Pusat (LKPP)

Hasil prediksi berdasarkan model regresi linier memberikan gambaran yang cukup jelas tentang arah perkembangan sektor ekonomi Indonesia pada periode 2012 hingga 2020. Penurunan kontribusi sektor industri pengolahan menunjukkan adanya pergeseran dalam struktur ekonomi Indonesia, mungkin disebabkan perubahan permintaan industri manufaktur atau peningkatan ketergantungan terhadap sektor lain, seperti sektor jasa. Penurunan tajam di sektor pertambangan mencerminkan dampak negatif dari penurunan harga komoditas global dan semakin berkurangnya cadangan sumber daya alam, yang menjadi tantangan bagi perekonomian yang sangat bergantung pada sektor ini. Sementara itu, stabilitas sektor pertanian dan sektor lain mencerminkan ketahanan sektor-sektor ini yang meskipun tidak ada lonjakan signifikan dalam kontribusinya terhadap PDB. Meskipun stabil, sektor-sektor ini yang tetap menjadi komponen penting perekonomian Indonesia, terutama dalam menyediakan lapangan pekerjaan dan memenuhi kebutuhan pangan domestik. Hasil analisis ini menunjukkan meskipun ada sektor-sektor yang mengalami penurunan, masih ada peluang untuk merumuskan kebijakan yang dapat mendorong pertumbuhan sektor-sektor lainnya, serta menciptakan kebijakan yang lebih berkelanjutan untuk mengatasi tantangan yang muncul, terutama yang berkaitan dengan ketergantungan pada sektor sumber daya alam.

KESIMPULAN

Penelitian ini berhasil menerapkan pendekatan *data mining* melalui kombinasi algoritma *K-Means Clustering* dan *Linear Regression* dalam menganalisis serta memprediksi tren indikator ekonomi makro Indonesia. Dengan menggunakan data neraca perdagangan, pendapatan negara, struktur penggunaan Produk Domestik Bruto (PDB), APBD wilayah Daerah Istimewa Yogyakarta, serta ekspor industri non-migas, penelitian ini mampu mengidentifikasi pola tersembunyi dan kecenderungan jangka panjang dalam sistem ekonomi nasional. Proses klusterisasi menggunakan *K-Means* berhasil mengelompokkan wilayah dan periode waktu berdasarkan karakteristik ekonomi yang serupa, seperti tahun-tahun dengan neraca perdagangan yang defisit atau surplus. Evaluasi kluster menunjukkan hasil yang baik, didukung nilai *inertia*

yang rendah, *silhouette score* yang tinggi, dan *Davies-Bouldin Index* yang kecil. Sementara itu, model regresi linier menunjukkan tren positif yang konsisten sektor utama seperti konsumsi rumah tangga dan pendapatan perpajakan. Model regresi linier dalam penelitian ini bersifat prediktif berdasarkan data historis, namun tidak sepenuhnya mencerminkan dinamika ekonomi kompleks, terutama pengaruh faktor eksternal seperti kebijakan pemerintah, krisis global, atau perubahan harga komoditas. Temuan ini yang menjadi dasar pertimbangan dalam proses penyusunan kebijakan-kebijakan fiskal maupun strategi pembangunan ekonomi yang lebih tepat sasaran, baik di tingkat nasional maupun daerah.

DAFTAR REFERENSI

- Direktorat Jenderal Perbendaharaan Kementerian Keuangan RI. (2025). *Laporan keuangan pemerintah pusat (LKPP)*. DJPb Kemenkeu. Available: <https://djpb.kemenkeu.go.id/portal/id/laporan-keuangan-pemerintah-pusat-lkpp.html>. [Accessed: Jun. 2025].
- Hendrastuty, N. (2024). Penerapan Data Mining Menggunakan Algoritma K-Means Clustering Dalam Evaluasi Hasil Pembelajaran Siswa. *Jurnal Ilmiah Informatika dan Ilmu Komputer (JIMA-ILKOM)*, 3(1), 46-56. <https://doi.org/10.58602/jima-ilkom.v3i1.26>
- Hermawan, M. S., Judijanto, L., Kusumastuti, S. Y., Purnamaningrum, T. K., Suparyati, A., Pracoyo, A., Prabandari, A. I., Atmaja, U., & Ilma, A. F. N. (2025). *Pengantar Ekonomi Indonesia: Perkembangan dan tantangan perekonomian dalam rangka percepatan pembangunan nasional menuju Indonesia Emas 2045*. PT. Sonpedia Publishing Indonesia.
- Huda, N. (2018). *Ekonomi Makro Islam: Pendekatan Teoritis*. Prenada Media.
- Priyatno, D. (2023). *Olah data sendiri analisis regresi linier dengan SPSS dan analisis regresi data panel dengan Eviews*. Penerbit Andi.
- Sekretariat Negara Republik Indonesia. (2016). *Buku II: Nota keuangan dan RAPBN tahun anggaran 2016*. Kementerian Keuangan RI.
- Suparmoko, M. (1997). *Keuangan negara dalam teori dan praktik* (4th ed.). BPFE–Yogyakarta.
- Supranto, J. (2000). *Statistik: Teori dan aplikasi*. Erlangga.
- Syam, S., Tokoro, Y., Judijanto, L., Garonga, M., Sinaga, F. M., Umar, N., Handika, I. P. S., Nur Iin, J., Apriyanto, A., & Sitanggang, A. T. (2024). *Data mining: Teori dan penerapannya dalam berbagai bidang*. PT. Sonpedia Publishing Indonesia.
- Wilda, R. W., Sukmarini, M. A., & Mahar, A. C. (2024). Prediksi Harga Saham PT. Unilever Indonesia TBK Dengan Metode Regresi Linier Sederhana. *Balance: Media Informasi Akuntansi dan Keuangan*, 16(2). <https://doi.org/10.52300/blnc.v16i2.14249>
- Zulaikah, Z. (2025). Peran Kebijakan Fiskal dan Moneter dalam Menjaga Stabilitas Ekonomi Makro. *Journal of Mandalika Literature*, 6(1), 95-108. <https://doi.org/10.36312/jml.v6i1.3900>