

# ANALISIS HASIL UJIAN NASIONAL TAHUN 2016-2019 TINGKAT SEKOLAH MENENGAH PERTAMA DENGAN PENDEKATAN SAINS DATA

Margaretha N.P. Janu<sup>a\*</sup>, Hartono<sup>b</sup>, Ig.Aris Dwiatmoko<sup>b</sup>

<sup>a</sup> Program Studi Pendidikan Matematika, Universitas Sanata Dharma, Yogyakarta

<sup>b</sup> Program Studi Matematika, Universitas Sanata Dharma, Yogyakarta

\*rethajanu@gmail.com

---

## **Abstract**

The Indonesian National Exams was a standard evaluation tool for students in junior high school to measure graduate competencies achievement on some subjects. It had been carried out for more than a decade based on government regulation number 19 of 2005 which was revised to number 19 of 2015. Additionally, it aimed to provide a quality map of Indonesian education. In accordance with the National Exam aim, the objectives of this research are for (1) knowing the result of visual analysis of National Exam average score (2) knowing which the most influential subjects to the National Exam average score (3) analyzing the National Exam score cluster. Data used in this research is National Exam average score from 2016 to 2019 taken from the Center for Educational Assessment of the Ministry of Education and Culture. The result shows that in general there is a statistical significant decrease on average score of National Exam of 2016 and the years after for all subjects. National Exam which was computer based and incorporated higher order thinking skill (hots) questions was suspected to contribute to the decline. Furthermore, after 2016 National Exam was no longer a vital factor to determine graduation so that its role was not as important as before. There are also differences in average National Exam score among regions. There were 7 provinces which their average score were below the national average score. The results of principle component analysis shows some subjects contribute most to the overall average score. Moreover, the result of cluster analysis confirms that there is still a disparity education quality between provinces in Indonesia. Therefore, the regular government policies to reduce this disparity are necessary.

**Keywords:** National Exam; Data Science; Principle Component Analysis; Cluster.

## **Abstrak**

Ujian Nasional yang diselenggarakan di Indonesia adalah sebuah sarana menilai pencapaian kompetensi lulusan secara nasional pada mata pelajaran tertentu. Ujian Nasional yang telah dilaksanakan belasan tahun tersebut bersandar pada Peraturan Pemerintah Nomer 13 tahun 2015 yang merupakan perubahan atas Peraturan Pemerintah Nomer 19 tahun 2005. Salah satu fungsi Ujian Nasional adalah untuk pemetaan mutu Pendidikan di tingkat nasional. Sejalan dengan fungsi tersebut, penelitian ini bertujuan untuk (1) mengetahui hasil analisis visualisasi data capaian rata-rata Ujian Nasional, (2) mengetahui nilai mata pelajaran mana yang paling berpengaruh terhadap capaian nilai rata-rata Ujian Nasional, dan (3) menganalisis hasil klasterisasi wilayah berdasarkan capaian Ujian Nasional. Data yang digunakan adalah data capaian nilai rata-rata Ujian Nasional tahun 2016-2019 yang bersumber dari laman puspendik kemdikbud. Hasil penelitian menunjukkan bahwa secara nasional terdapat perbedaan rata-rata antara tahun 2016 dengan tahun 2017, 2018, dan 2019 untuk semua mata pelajaran. Pemberlakuan soal HOTS, pelaksanaan UNBK dan pelaksanaan USBN diduga turut berpengaruh terhadap turunnya nilai rata-rata Ujian Nasional. Selain itu terdapat perbedaan rata-rata nilai UN antar daerah setiap tahunnya. Terdapat tujuh provinsi yang menunjukkan hasil capaian nilai rata-rata berada di bawah capaian nilai rata-rata secara nasional. Hasil Analisis Komponen Utama menunjukkan tahun 2016 dan 2019 mata pelajaran yang menyumbang nilai cukup besar terhadap besar atau kecilnya capaian nilai rata-rata UN secara nasional adalah Bahasa Inggris dan IPA, sementara pada tahun 2017 dan 2018 adalah Matematika dan IPA. Hasil analisis klaster menegaskan kembali masih terjadinya disparitas mutu luaran pendidikan di antara provinsi-provinsi di Indonesia. Diperlukan kebijakan pemerintah yang terus menerus untuk mengurangi disparitas mutu pendidikan ini sesuai dengan amanat undang-undang.

**Kata Kunci:** Ujian Nasional; Sains Data; Analisis Komponen Utama; Klaster.

---

## 1. Pendahuluan

Upaya pemerintah untuk meningkatkan mutu pendidikan di Indonesia antara lain melakukan perbaikan baik dalam hal kurikulum, profesionalitas dan kualitas guru, serta infrastruktur. Sektor penting dalam bidang pendidikan selain yang telah disebutkan sebelumnya adalah sistem evaluasi, dimana di dalam proses evaluasi terdapat penilaian pendidikan yang terdiri atas penilaian hasil belajar oleh pendidik, satuan pendidikan, dan pemerintah (PP no.13/2015). Bentuk evaluasi yang dilakukan pemerintah ialah penilaian dalam bentuk Ujian Nasional (UN) yang bertujuan untuk menilai pencapaian kompetensi lulusan secara nasional pada mata pelajaran tertentu.

Kehadiran sains data dalam kurun waktu satu dekade terakhir dapat dimanfaatkan untuk proses pengolahan data dalam jumlah besar dengan metode yang lebih cepat. Penelitian ini bertujuan untuk (1) mengetahui hasil analisis visualisasi data capaian rata-rata Ujian Nasional, (2) mengetahui nilai mata pelajaran mana yang paling berpengaruh terhadap capaian nilai rata-rata Ujian Nasional melalui analisis komponen utama, dan (3) menganalisis hasil klasterisasi hasil UN berdasarkan wilayah yang ada di Indonesia dengan algoritma K-Means Clustering. Dengan demikian dapat diperoleh informasi lain tentang pelaksanaan Ujian Nasional di Indonesia.

## 2. Kajian Literatur

Beberapa penelitian di bidang pendidikan telah menggunakan data ujian nasional sebagai bahan pengolahan datanya. Sejauh yang diketahui penulis, hampir semua penelitian yang dilakukan hanya terbatas menggunakan data ujian nasional pada suatu daerah tertentu, seperti dalam (Dwiatmoko et al, 2015) dan (Miniarti, 2013). Berbeda dengan literatur tersebut, penelitian ini menggunakan data ujian nasional dari seluruh provinsi di Indonesia mulai tahun 2016 sampai tahun 2019.

Seperti dalam (Stevencua & Setiawan, 2018) visualisasi data memegang peranan penting untuk dapat memahami data secara lebih mudah, melihat berbagai kecenderungan data, pencilaan data yang kesemuanya itu dapat memberikan informasi yang sangat berguna untuk pengolahan data selanjutnya. Di dalam penelitian ini juga dilakukan visualisasi data baik dalam bentuk grafik, boxplot maupun histogram dari data ujian nasional.

Selanjutnya dengan menggunakan teori analisis komponen utama dalam (James et al, 2017), dapat ditentukan komponen terpenting apa saja yang menentukan hasil secara keseluruhan dari data yang diolah. Secara garis besar, analisis komponen utama digunakan untuk mereduksi dimensionalitas himpunan data multivariat dengan cara mentransformasi suatu himpunan variabel asal ke himpunan variabel baru yang seringkali disebut komponen utama. Komponen utama dapat dipandang sebagai kombinasi linear dari variabel asal yang tidak berkorelasi dan diurutkan sedemikian rupa sehingga sejumlah variabel urutan pertama menjelaskan sebagian besar variansi dari variabel-variabel asal. Hasil dari analisis komponen utama adalah terbentuknya sejumlah kecil variabel baru (komponen utama).

Secara aljabar, komponen utama adalah kombinasi linear-kombinasi linear tertentu dari  $p$  peubah acak  $y_1, y_2, \dots, y_p$ , sehingga Komponen Utama dapat didefinisikan sebagai berikut:

$$KU = AY$$

Rumus 2.1 komponen utama

dengan

$$KU = \begin{pmatrix} KU_1 \\ KU_2 \\ \vdots \\ KU_k \end{pmatrix} \text{ merupakan komponen utama,}$$

$$A = \begin{pmatrix} a_1^T \\ a_2^T \\ \vdots \\ a_k^T \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \cdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kp} \end{pmatrix} \text{ merupakan}$$

vektor eigen yang ditranspose,

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{pmatrix} \text{ vektor variabel acak } y_1, y_2, \dots, y_p.$$

Keluaran dari analisis komponen utama adalah terbentuknya  $p$  kombinasi linear

$$KU_1 = a_{11}y_1 + a_{12}y_2 + \cdots + a_{1p}y_p$$

$$KU_2 = a_{21}y_1 + a_{22}y_2 + \cdots + a_{2p}y_p$$

:

$$KU_k = a_{k1}y_1 + a_{k2}y_2 + \cdots + a_{kp}y_p$$

dimana  $KU_1, KU_2, \dots, KU_k$  adalah  $k$  komponen utama dan  $a_{ij}$  adalah bobot variabel ke- $j$  untuk komponen utama ke- $i$ . Bobot  $a_{ij}$ , yang juga merupakan entri dari vektor eigen harus ditentukan sedemikian sehingga memenuhi:

- Komponen utama pertama,  $KU_1$  menjelaskan varians maksimum dari data, sedangkan komponen utama kedua,  $KU_2$ , menjelaskan varians maksimum yang tidak dijelaskan oleh komponen utama pertama, dan seterusnya.
- $a_{i1}^2 + a_{i2}^2 + \dots + a_{ip}^2 = 1$ , dimana  $i = 1, \dots, k$ .
- $a_{i1}a_{j1} + a_{i2}a_{j2} + \dots + a_{ip}a_{jp} = 0$ , untuk setiap  $i \neq j$ .

Dalam artikel ini komponen utama ditentukan dengan menggunakan proporsi kumulatif varians terhadap total dengan rumus

$$k = \frac{\sum_{j=1}^k \lambda_j}{\sum_{j=1}^p \lambda_j} \times 100\%$$

Rumus 2.2 proporsi kumulatif

untuk  $k \leq p$  dan menggunakan nilai eigen yang bernilai lebih besar dari satu ( $\lambda > 1$ ).

Selanjutnya, analisis kluster seperti dalam (Ahmar et al, 2018) digunakan untuk membuat klusterisasi provinsi-provinsi di Indonesia berdasarkan nilai ujian nasionalnya. Secara umum, algoritma K-Means Clustering :

- Pilih  $k$  buah titik centroid secara acak.
- Kelompokkan data sehingga terbentuk  $K$  buah kluster dengan titik centroid dari setiap cluster merupakan titik centroid yang telah dipilih sebelumnya.
- Perbaharui nilai titik centroid.
- Ulangi langkah ii dan iii sampai nilai dari titik centroid tidak lagi berubah.

Perbedaan jarak atau besaran angka yang cukup jauh dalam data, dapat menyulitkan proses pengelompokan. Salah satu solusi yang digunakan untuk memperkecil besaran angka antar variabel adalah melakukan normalisasi dengan menggunakan

$$\text{nilai normalisasi} = \frac{\text{nilai awal} - \text{nilai min}}{\text{nilai maksimum} - \text{nilai min}}$$

Rumus 2.3 rumus normalisasi

Selain itu untuk menentukan berapa banyak kluster yang paling optimal dapat digunakan *elbow method* (James et al, 2017).

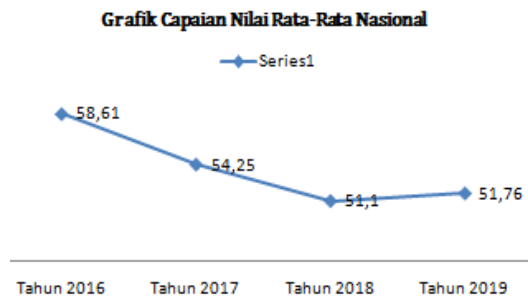
### 3. Metode Penelitian

Penelitian ini menggunakan pendekatan analisis deskriptif kualitatif. Objek penelitian dalam tulisan ini adalah data UN 2016-2019 tingkat Sekolah Menengah Pertama yang diperoleh dari laman puspendik kemdikbud. Teknik analisis data meliputi visualisasi grafis, analisis komponen utama, dan analisis kluster.

## 4. Hasil dan Pembahasan

### 4.1 Visualisasi Data Ujian Nasional

Pada Gambar 4.1 dan Gambar 4.2 berikut ini akan ditampilkan visualisasi grafis capaian nilai Ujian Nasional untuk semua mata pelajaran, dan capaian nilai Ujian Nasional dari masing-masing pelajaran tahun 2016-2019.



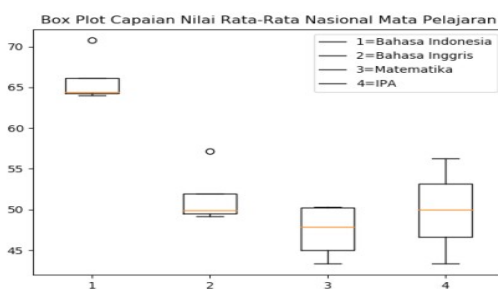
Gambar 4.1 Capaian nilai ujian nasional tahun 2016-2019

Dari Gambar 4.1 di atas diketahui secara nasional nilai rata-rata berkisar pada nilai 51.1 dan 58.61, artinya secara nasional terdapat cukup banyak wilayah yang memperoleh kisaran nilai di atas 50. Dalam skala 100, capaian ini dapat dinilai tidak memuaskan. Selain itu, secara nasional terjadi penurunan secara berturut-turut sekitar 0.04%, 0.06%, dari tahun 2016 sampai tahun 2018, dan kenaikan sebesar 0.01% dari tahun 2018 ke tahun 2019. Secara umum nilai rata-rata UN paling tinggi untuk semua mata pelajaran dicapai pada tahun 2016, kecuali pada mata pelajaran Matematika. Sejauh amatan penulis ada beberapa faktor yang diduga menjadi penyebab turunnya nilai rata-rata ujian nasional setelah tahun 2016 antara lain mulai diperkenalkannya soal HOTS, pelaksanaan UN berbasis komputer, dan pelaksanaan USBN yang menjadikan UN bukan lagi alat utama penentu kelulusan.

Gambar 4.2 memperlihatkan boxplot nilai rata-rata mata pelajaran tahun 2016-2019. Dari boxplot tersebut dapat dilihat perbedaan

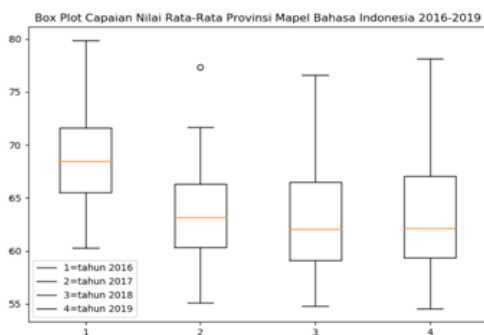
capaian nilai rata-rata yang cukup besar antara mata pelajaran Bahasa Indonesia dengan tiga mata pelajaran lainnya. Fakta ini memunculkan dugaan apakah nilai Bahasa Indonesia mempengaruhi capaian nilai Ujian Nasional atau justru capaian nilai Ujian Nasional dipengaruhi oleh nilai mata pelajaran lainnya.

Pola sebaran data nilai Bahasa Indonesia dan Bahasa Inggris cukup mirip walaupun berada dalam level nilai yang berbeda, sementara sebaran data nilai Matematika dan IPA cukup berimbang. Nilai rata-rata Matematika yang paling rendah mengindikasikan bahwa matapelajaran ini dipandang siswa sebagai matapelajaran tersulit di antara yang lainnya.



Gambar 4.2 Boxplot nilai rata-rata mata pelajaran tahun 2016-2019

Selain melihat bagaimana capaian nilai Ujian Nasional pada tingkat nasional, pada Gambar 4.3 berikut akan ditampilkan capaian nilai rata-rata provinsi untuk 4 mata pelajaran dari tahun 2016-2019.



Gambar 4.3 Boxplot nilai rata-rata mata pelajaran tingkat provinsi tahun 2016-2019

Nilai maksimum pada tahun 2016 dicapai oleh provinsi Sumatera Utara, sementara dari tahun 2017-2019 dicapai oleh provinsi DIY. Secara khusus, pada tahun 2017, provinsi DIY menjadi pencalonan dengan capaian yang jauh melampaui provinsi lainnya. Nilai minimum pada tahun 2016-2017 diperoleh Bengkulu dan dari tahun 2018-2019 oleh Aceh.

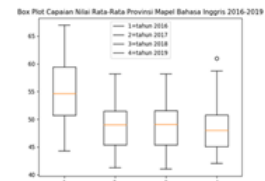
Dengan uji Friedman disimpulkan bahwa terdapat perbedaan rata-rata antara tahun

2016 dan 2017, 2016 dan 2018, juga 2016 dan 2019.

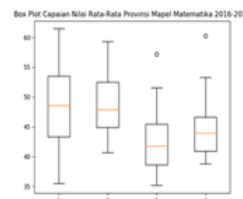
Dengan cara yang sama dapat dilihat capaian nilai setiap mata pelajaran di tingkat provinsi selama tahun 2016-2019 pada Gambar 4.4. Dari boxplot tersebut, terdapat kemiripan perilaku capaian nilai setiap mata pelajaran, yaitu cenderung menurun dari tahun ke tahun kecuali untuk mata pelajaran Matematika yang mengalami kenaikan pada tahun 2019. Provinsi DIY mencapai nilai tertinggi di tingkat nasional setiap tahunnya dan menjadi pencalonan, kecuali dalam mata pelajaran Bahasa Inggris.



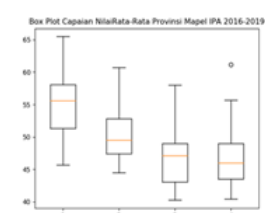
Boxplot Bahasa Indonesia 2016-2019



Boxplot Bahasa Inggris 2016-2019



Boxplot Matematika 2016-2019



Boxplot IPA 2016-2019

Gambar 4.4 Boxplot rata-rata setiap mata pelajaran tingkat provinsi tahun 2016-2019

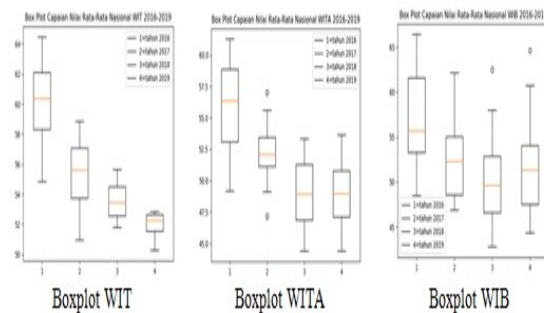
Selain DIY, beberapa provinsi lain yang juga pernah menduduki capaian nilai maksimum adalah provinsi Sumut, Papua Barat, Maluku, dan DKI Jakarta. Sementara bila dilihat dari capaian nilai minimum, perlu ada perhatian khusus bagi provinsi Aceh yang dari tahun 2016 sampai tahun 2019 pada beberapa mata pelajaran terus memperoleh nilai terendah. Selain provinsi Aceh, daerah lain yang juga pernah mendapat capaian nilai minimum adalah Sulbar, Bengkulu, dan NTB.

Dari gambar di atas terlihat adanya overlapping antar boxplot, sehingga perlu diuji apakah ada perbedaan rata-rata yang signifikan antara tahun-tahun pelaksanaan Ujian Nasional di setiap mata pelajaran. Dari hasil uji Friedman pada taraf signifikansi 5%, pada mata pelajaran Bahasa Indonesia dan Bahasa Inggris, rata-rata nilai UN berbeda secara signifikan pada tahun 2016 dan 2017, 2016 dan 2018, juga 2016 dan 2019. Pada mata pelajaran Matematika perbedaan rata-rata yang signifikan terjadi tahun 2016 dan

2018, 2017 dan 2018, 2017 dan 2019, dan 2018 dengan 2019. Sementara pada mata pelajaran IPA terdapat perbedaan rata-rata antara tahun 2016 dan 2017, 2016 dan 2018, 2016 dan 2019, 2017 dan 2018, 2017 dan 2019.

Hal lain yang diamati dari capaian nilai rata-rata provinsi secara nasional yaitu terdapat 7 provinsi yang memperoleh capaian nilai rata-rata di bawah nilai rata-rata nasional untuk keempat mata pelajaran dalam kurun waktu 2016-2019. Ketujuh provinsi tersebut adalah Aceh, Jambi, Sumsel, Lampung, Kalbar, NTB, dan Banten. Seperti yang telah ditunjukkan pada boxplot sebelumnya, provinsi NTB, Kalbar, dan Aceh merupakan 3 provinsi yang juga pernah mendapatkan nilai minimum pada pelaksanaan Ujian Nasional.

Selanjutnya pada Gambar 4.5 dapat dilihat capaian nilai yang diperoleh setiap provinsi di masing-masing zona waktu.



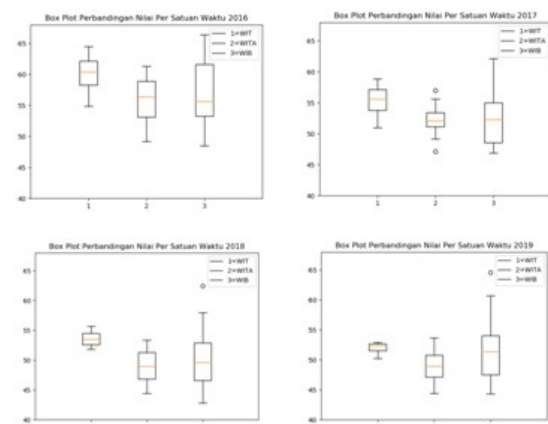
Gambar 4.5 Boxplot rata-rata UN di setiap zona waktu tahun 2016-2019

Dari boxplot terlihat kemiripan perilaku rata-rata nilai UN di ketiga zona waktu yaitu cenderung menurun, kecuali di WIB yang naik pada tahun 2019. Dari 4 provinsi yang ada di zona waktu WIT, diketahui provinsi dengan capaian nilai maksimum dari tahun 2016-2017 adalah provinsi Papua Barat, dan dari tahun 2018-2019 provinsi Maluku. Sementara itu selama 4 tahun terakhir nilai minimum dicapai oleh provinsi Papua. Nilai maksimum di zona waktu WITA yang terdiri dari 12 provinsi dicapai oleh provinsi Sulut pada tahun 2016-2017 dan Bali pada tahun 2018-2019. Pencilan pada tahun 2017 dicapai oleh provinsi Sulut sebagai pencilan atas dan Sulbar sebagai pencilan bawah. Di zona waktu WIB yang terdiri dari 18 provinsi nilai maksimum tahun 2016 dicapai oleh Sumut, sementara tahun 2017-2019 dicapai oleh DIY sekaligus penyebab terjadinya pencilan. Nilai minimum 2016-2017 diperoleh Bengkulu dan tahun 2018-2019 oleh Aceh.

Dengan menggunakan uji Friedman disimpulkan bahwa terdapat perbedaan rata-rata antara tahun 2016 dan 2019 di zona WIT.

Sementara di zona WITA terdapat perbedaan rata-rata tahun 2016 dan 2017, 2016 dan 2018, 2016 dan 2019. Sedangkan di zona WIB terdapat perbedaan rata-rata tahun 2016 dan 2017, 2016 dan 2018, 2016 dan 2019. Tampilan boxplot juga menunjukkan adanya keragaman nilai UN, khususnya pada zona WIT dan WITA. Di zona WIT, rentang nilai yang menyempit menggambarkan capaian nilai UN di antara sekolah-sekolah semakin homogen. Hal ini merupakan suatu pertanda baik karena telah terjadi pemerataan mutu pendidikan. Di zona WITA dan WIB terlihat rentang yang relatif lebar menggambarkan terjadinya disparitas mutu capaian pendidikan. Perhatian pemerintah diperlukan untuk tercapainya pemerataan mutu sebagaimana di zona WIT.

Berikutnya akan pada Gambar 4.6 divisualisasikan juga capaian nilai rata-rata antar zona waktu di setiap tahunnya. Hal ini bertujuan untuk melihat apakah ada perbedaan rata-rata yang signifikan antar ketiga zona waktu, dan apakah benar bahwa nilai Ujian Nasional secara tegas menunjukkan bahwa nilai rata-rata pada zona WIB selalu lebih baik dibanding kedua zona waktu lain sesuai dengan persepsi masyarakat umum selama ini.



Gambar 4.6 Boxplot rata-rata antar setiap zona waktu

Dengan menggunakan Uji Kruskal-Wallis diperoleh kesimpulan bahwa tidak ada perbedaan rata-rata UN yang signifikan antar ketiga zona waktu. Hal ini dapat dipahami karena rata-rata nilai UN merupakan rata-rata gabungan dari keempat mata pelajaran yang perilaku masing-masing dapat saling menutupi sehingga membentuk rata-rata yang relatif sama. Perilaku masing-masing matapelajaran dalam menentukan rata-rata nilai UN dapat dijelaskan dengan analisis komponen utama.

#### 4.2 Analisis Komponen Utama

Analisis komponen utama pada bagian ini dipakai untuk mengetahui nilai apakah yang berpengaruh paling besar terhadap capaian nilai rata-rata UN setiap tahunnya. Misalkan mata pelajaran Bahasa Indonesia dilambangkan dengan  $y_1$ , mata pelajaran Bahasa Inggris dilambangkan dengan  $y_2$ , Matematika dilambangkan dengan  $y_3$  dan IPA dilambangkan dengan  $y_4$ , selanjutnya dicari matriks kovarian, nilai eigen dan vektor eigen, dan komponen utama dari masing-masing tahun pelaksanaan ujian nasional.

Dari perhitungan nilai eigen dari matriks kovarian untuk data tiap tahun, diketahui bahwa hanya terdapat 1 nilai eigen yang besarnya lebih dari satu, sehingga dengan menggunakan satu komponen utama ternyata sudah cukup baik untuk menggambarkan keseluruhan data. Untuk tahun 2016 didapat besar proporsi kumulatif adalah 72.64% dengan

$$KU_1 = -0.2332135 y_1 - 0.5647950 y_2 - 0.5568$$

Adapun komponen utama tersebut didominasi oleh variabel  $y_2$  (Bahasa Inggris) dan variabel  $y_4$  (IPA). Untuk tahun 2017-2019 berturut-turut proporsi kumulatif dari masing-masing tahun adalah 70%, 83% dan 88%. Dengan demikian berturut-turut diperoleh satu komponen utama sebagai berikut untuk masing-masing tahun:

$$KU_1 = -0.3070382 y_1 - 0.5025958 y_2 - 0.5690$$

$$KU_1 = -0.4300257 y_1 - 0.5146113 y_2 - 0.5151$$

$$KU_1 = -0.4641254 y_1 - 0.5051840 y_2 - 0.5021$$

Dari hasil yang diperoleh diketahui pada tahun 2017 dan 2018 mata pelajaran yang berpengaruh pada capaian nilai rata-rata Ujian Nasional adalah mata pelajaran Matematika dan IPA, sementara pada tahun 2019 adalah mata pelajaran Bahasa Inggris dan IPA. Disini kita tahu bahwa sejak tahun 2016 sampai 2019, mata pelajaran IPA terus muncul sebagai salah satu mata pelajaran yang mempengaruhi capaian nilai rata-rata ujian Nasional, sehingga diperlukan perhatian khusus pada mata pelajaran ini.

#### 4.3 Analisis Klaster secara Nasional

Pada bagian ini akan ditampilkan hasil klasterisasi wilayah berdasarkan provinsi-provinsi di Indonesia. Seperti yang telah diketahui sebelumnya terdapat tiga zona waktu di Indonesia yakni WIT, WITA dan WIB. Hal yang akan diamati selain hasil klasterisasinya adalah apakah hasil dari proses klasterisasi tersebut menunjukkan

provinsi-provinsi yang berada dalam satu zona waktu memang terklaster dalam klaster yang sama atau tidak. Pengklasteran ini didasarkan pada capaian nilai rata-rata semua mata pelajaran selama empat tahun terakhir. Selanjutnya dibuat scree plot untuk memperoleh banyak klaster yang paling optimal (yaitu  $k = 4$ ) dengan menggunakan elbow method. Penentuan klasternya menggunakan algoritma K-Means Klastering seperti yang telah dijelaskan sebelumnya.

Setelah melalui proses normalisasi diperoleh pusat data tiap klaster yang dapat dilihat pada Tabel 4.1 berikut:

Tabel.4.1. Pusat data klaster

Klaster	Bhs Indonesia	Bhs Inggris	Matematika	IPA	Rata rata
1	75.6	57.3	54.3	57.6	61
2	62.7	55.5	51.8	52.5	55
3	66.3	51.1	46.5	50.6	53
4	61.6	46.5	41.5	45.8	48

Ternyata setelah melihat rata-rata dari pusat data yang ada, dapat disimpulkan bahwa

$klaster\ 1 > klaster\ 2 > klaster\ 3 > klaster\ 4$ .

Selanjutnya akan dilihat kabupaten mana saja yang menjadi anggota tiap klaster pada Tabel 4.2 berikut :

Tabel 4.2 Provinsi tiap klaster

Klaster	Provinsi
1	DKI Jakarta, DIY
2	Jabar, Sumut, Sulut, Maluku, Malut, Papua Barat
3	Jateng, Jatim, Sumbar, Riau, Kalteng, Kalsel, Kaltim, Sulsul, Sultra, Bali, NTT, Papua, Kepri
4	Aceh, Jambi, Sumsel, Lampung, Kalbar, Sulteng, NTB, Bengkulu, Babel, Gorontalo, Banten, Sulbar, Kalut

Dari hasil klasterisasi tersebut dapat diketahui bahwa daerah-daerah pada suatu klaster tidak selalu berasal dari zona waktu yang sama. Hasil pengklasteran menunjukkan selama 4 tahun terakhir tidak semua provinsi yang ada terklaster menurut zona waktunya. Hal menarik lain dari hasil klasterisasi ini ternyata sesuai dengan hasil yang ditunjukkan pada pembahasan tentang provinsi mana saja yang berada di bawah capaian nilai rata-rata nasional selama 4 tahun terakhir. Hasil klasterisasi ini menunjukkan bahwa ketujuh provinsi yang berada di bawah capaian rata-rata nasional berada dalam 1 klaster. Selain

ketujuh provinsi itu, ternyata ada 6 provinsi lain yakni Sulawesi Tengah, Bengkulu, Bangka Belitung, Sulawesi Barat, Gorontalo dan Kalimantan Utara yang masuk dalam klaster yang sama. Setelah dicek kembali, ternyata pada mata pelajaran tertentu provinsi-provinsi ini berada di bawah nilai rata-rata nasional. Artinya, nilai yang dicapai 6 provinsi tersebut sebenarnya tidak begitu berbeda jauh dengan ketujuh provinsi. Misalnya pada provinsi Bengkulu yang memperoleh nilai Bahasa Indonesia, Bahasa Inggris, dan Matematika di bawah rata-rata nasional, tetapi tidak demikian untuk capaian nilai IPA. Akibatnya provinsi ini tidak masuk dalam irisan ke dalam kelompok yang sama dengan 7 provinsi lainnya. Dengan kata lain, jika diambil rata-rata dari keempat mata pelajaran selama 4 tahun, provinsi Bengkulu sebenarnya memiliki capaian nilai yang cukup dekat dengan capaian ketujuh provinsi lainnya.

Provinsi DKI Jakarta dan D.I Yogyakarta merupakan dua provinsi pencilan dengan capaian nilai rata-rata jauh melebihi provinsi lainnya. Dapat pula kita amati bahwa zona WIB tampil di semua klaster. Artinya di WIB sendiri ada variasi capaian nilai UN dari level terbaik sampai level terendah. Zona WITA tampil di klaster 3 dan 4, artinya variasi nilai UN pada zona waktu ini sebenarnya berada pada rentang nilai rata-rata klaster tersebut. Sama halnya dengan zona WIT yang berada di klaster 2 dan 3.

Dari segi pemetaan mutu, hasil klasterisasi ini menegaskan kembali masih terjadinya disparitas kualitas luaran Pendidikan yang ditunjukkan oleh nilai Ujian Nasional. Sekolah-sekolah di zona WITA tidak pernah menduduki klaster 1 dan 2, sementara itu sekolah-sekolah di zona WIB ada di semua klaster. Artinya, disparitas tinggi terjadi di WIB sebagaimana juga telah ditunjukkan dalam boxplot sebelumnya. Sekolah-sekolah di WIT relatif seragam berada di klaster 2 kecuali provinsi Papua yang ada di zona 3. Fakta ini juga didukung tampilan boxplot yang rentangnya relatif sempit di zona WIT.

## 5. Kesimpulan dan Saran

Dari hasil penelitian ini dapat disimpulkan hal-hal sebagai berikut: (1) terjadi penurunan nilai rata-rata UN yang signifikan dalam kurun waktu 2016-2019, (2) kontribusi nilai tiap matapelajaran yang diujikan dalam UN terhadap nilai keseluruhan tidaklah sama, (3) masih terjadinya disparitas kualitas pendidikan di berbagai daerah di Indonesia sehingga diperlukan campur tangan pemerintah untuk mewujudkan pemerataan mutu pendidikan

sebagaimana diamanatkan dalam undang-undang.

## Daftar Pustaka

- Dwiatmoko, Ig. A., Prima Rosa, P.H. & Gunawan, R. (2015). Analisis Statistis Data Nilai Ujian Nasional dan Nilai Sekolah Menengah Atas di Daerah Istimewa Yogyakarta. *Jurnal Widya Teknik* 14 (2).
- Miniarti, J. (2013). Pengelompokan Sekolah Menengah Pertama Berdasarkan Rata-Rata Nilai Ujian Akhir Nasional di Kota Binjai dengan Analisis Hierrarkhi Clustering. *Skripsi Jurusan Matematika, FMIPA, Universitas Negeri Medan*.
- Ahmar, A.S., Napitulu, D., Rahim, R., Hidayat, R., Sonata, Y. & Azmi, M. (2018). Using K-Means Clustering to Cluster Provinces in Indonesia. *2nd International Conference on Statistics, Mathematics, Teaching, and Research*. IOP Conference Series 1028.
- James, G., Witten, D., Hestie, T. & Tibshirani, R. (2017). *An Introduction to Statistical Learning with Application in R*. Springer, USA.
- Stevencua, S., Setiawan, J. (2018). Data Visualization of Poverty Level at Provinces in Indonesian from The Year 2013-2015. *International Journal of New Media Technology (JNMT)* V(1).