

## ABSTRAK

**Janu, Margaretha Nobilio Pasia.(2020).Analisis Hasil Ujian Nasional Tahun 2016-2019 Tingkat Sekolah Menengah Pertama dengan Pendekatan Sains Data.**

Penelitian ini bertujuan untuk: (1) mengetahui hasil analisis hasil Ujian Nasional semua provinsi di Indonesia yang diperoleh melalui visualisasi data, (2) mengetahui nilai mana yang paling berpengaruh terhadap capaian nilai ujian nasional, dan (3) mengetahui hasil analisis klaster pada data Ujian Nasional.

Jenis penelitian yang dipakai adalah deskriptif-kuantitatif dengan analisis. Objek penelitian dalam tulisan ini adalah data UN 2016-2019 tingkat Sekolah Menengah Pertama. Langkah analisis data dimulai dengan (1) mengumpulkan dan membaca berbagai literatur Sains Data, (2) mengumpulkan data hasil Ujian Nasional dari laman puspendik.kemdikbud.go.id/hasilun/, (3) membuat visualisasi, (4) menganalisis hasil visualisasi, (5) melakukan Analisis Komponen Utama dengan bantuan perangkat lunak R, dan (6) melakukan analisis hasil klasterisasi wilayah.

Dengan menggunakan uji Friedmann yang bertujuan untuk melihat ada tidaknya perbedaan rata-rata yang signifikan terhadap capaian UN, diketahui 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 diprediksi turut berpengaruh terhadap turunnya rata-rata Ujian Nasional. Pembagian wilayah berdasarkan zona waktu menunjukkan terdapat perbedaan rata-rata di zona WIT, WITA, dan WIB setiap tahunnya, sementara berdasarkan tahun pelaksanaannya disimpulkan tidak ada perbedaan rata-rata yang signifikan antar ketiga zona waktu. Terdapat tujuh provinsi yang menunjukkan hasil capaian nilai rata-rata berada di bawah capaian nilai rata-rata secara nasional yaitu provinsi Aceh, Jambi, Sumatera Selatan,Kalimantan Barat, Lampung, NTB, dan Banten. Hasil Analisis Komponen Utama menunjukkan tahun 2016 dan 2019 mata pelajaran yang menyumbang nilai cukup besar terhadap tinggi atau rendahnya 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 pada tingkat nasional dengan K-Means Clustering menunjukkan dari empat klaster yang terbentuk ada enam provinsi lain yang masuk dalam klaster yang sama dengan ketujuh provinsi yang mencapai nilai rata-rata terendah selama empat tahun terakhir untuk semua mata pelajaran. Proses klasterisasi pada provinsi Nusa Tenggara Timur menghasilkan tiga klaster. Klaster dengan capaian rata-rata rendah adalah klaster 3. Kabupaten-kabupaten di pulau Flores masuk dalam kelompok klaster dengan capaian nilai yang cukup baik. Proses klasterisasi di Kabupaten Manggarai menghasilkan empat klaster. Daerah di desa lebih banyak berada pada klaster terbaik.

**Kata Kunci :** Sains Data, *Unsupervised Learning*, Analisis Komponen Utama, Klaster

## ABSTRACT

**Janu, Margaretha Nobilio Pasia. (2020). Analysis of the 2016-2019 National Exam Results for Junior High Schools Using Data Science Approach.**

This study aims to: (1) find out an analysis result of National Examination result of all provinces in Indonesia obtained through of all provinces in Indonesia, which are obtained through data visualization, (2) find out which scores influence the national exam result the most, and (3) acknowledge the results of cluster analysis on the National exam data.

This is a descriptive-quantitative with analysis research. The object of this research is junior high school's 2016-2019 national examination result. The data was analysed by (1) collecting and reading various Data Science literature, (2) collecting the National Examination results as a data from puspendik.kemdikbud.go.id/hasilun/, (3) creating visualizations, (4) analyzing the results of the visualization, (5) analyzing the Principle Component using R software, and (6) performing an analysis of the regional clustering results.

By using the Friedmann test which aims to see whether there is a significant difference in the average on National Examination results. It is acknowledged that there is a national average difference within 2016 to 2017 and 2018 to 2019 for all subjects. The implementation of HOTS questions, the implementation of UNBK and the implementation of USBN are predicted to have an effect on the decline in the National Exam average. The division of regions based on time zones shows that there are average differences in the WIT, WITA, and WIB zones each year. However there is no significant difference in the average between the three time zones based on the year of the implementation. There are seven provinces which results are below the national average, namely the provinces of Aceh, Jambi, South Sumatra, West Kalimantan, Lampung, NTB, and Banten. The results of the main component analysis show that in 2016 and 2019 the subjects that contributed significantly to the fluctuation in the national examination average were English and Science, while in 2017 and 2018 were Mathematics and Science. The results of the cluster analysis at the national level with K-Means Clustering show that out of the four clusters formed, there are six other provinces that are included in the same cluster with those which achieved the lowest average score over the last four years for all subjects. The clustering process in the province of East Nusa Tenggara resulted in three clusters. The cluster with low average performance is the third cluster. The districts on the island of Flores are included in the cluster group with fairly good scores. The clustering process in Manggarai Regency produces four clusters. Most of the areas in the village are in the best cluster.

**Keywords:** Sains Data, *Unsupervised Learning*, Analisis Komponen Utama, Klaster