



**Laporan Tugas Kelompok  
Mata Kuliah Analisis Data C  
Semester Gasal Tahun Ajaran 2021/2022**

---

**Pengelompokan Kabupaten/Kota di Provinsi Jawa Timur  
Berdasarkan Tenaga Pendidik dengan Metode *Clustering* K-Means**

**Oleh:**

<b>Haiva Qurrota A.</b>	<b>06211840000045</b>
<b>Susilowati Gusinta</b>	<b>06211840000060</b>

**Dosen Pengampu:  
Dr. Kartika Fithriasari, M.Si**

**FAKULTAS SAINS DAN ANALITIKA DATA  
JURUSAN STATISTIKA PRODI SARJANA  
INSTITUT TEKNOLOGI SEPULUH  
NOPEMBER  
SURABAYA  
2021**

## A. Sumber Data

Data yang digunakan adalah data sekunder tentang jumlah guru di Jawa Timur yang diakses dari laman *website* Dinas Pendidikan Jawa Timur pada Selasa, 21 September 2021 pukul 09.45 WIB. Data tersebut menjelaskan jumlah guru secara keseluruhan, jumlah guru berjenis kelamin laki-laki dan perempuan pada tiap lembaga pendidikan, yaitu TK, KB, TPA, SPS, PKBM, SKB, SD, SMP, SMA, SMK dan SLB di seluruh kabupaten/kota di Jawa Timur.

## B. Tujuan Analisis

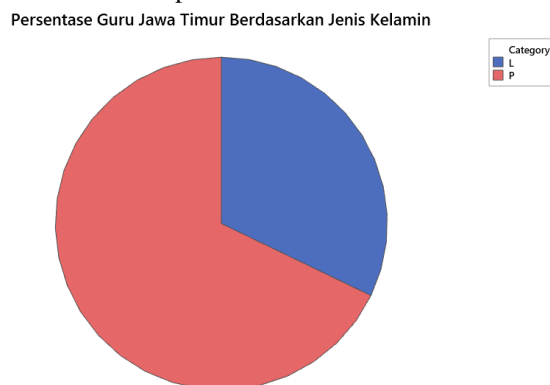
Adapun tujuan dilakukannya analisis ini adalah sebagai berikut.

1. Mendapatkan jumlah kelompok optimum untuk mengelompokkan kabupaten/kota di Jawa Timur Indikator Kesehatan Masyarakat pada metode Clustering K-Means.
2. Mengelompokkan kabupaten/kota di Jawa Timur berdasarkan Indikator Kesehatan Masyarakat dengan menggunakan metode K-Means.
3. Untuk memberikan gambaran bagaimana persebaran tenaga pendidik di Jawa Timur.

## C. Hasil Analisis

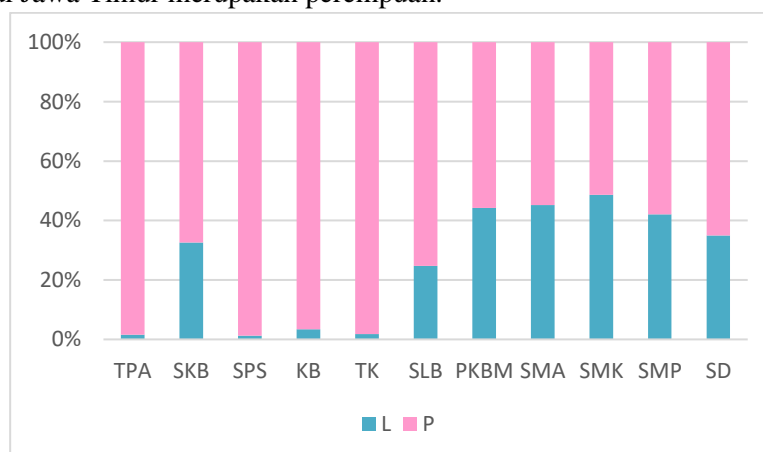
### 1. Analisis Deskriptif

Berikut visualisasi dari analisis deskriptif dari data.



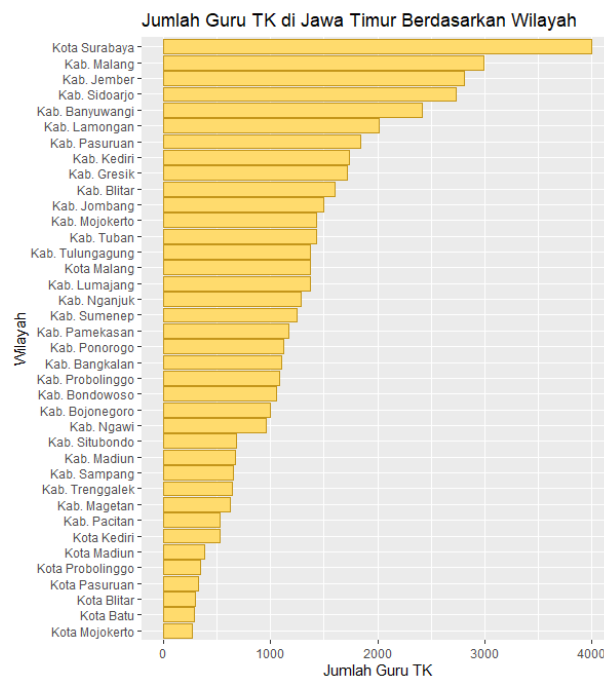
**Gambar 1** Pie Chart Persentase Guru Jawa Timur berdasarkan Jenis Kelamin

Gambar 1 memperlihatkan bahwa dari 415.002 tenaga pendidik di Jawa Timur 67,87% berjenis kelamin perempuan dan 32,12% lainnya berjenis kelamin laki-laki. Sehingga mayoritas dari tenaga pendidik di Jawa Timur merupakan perempuan.



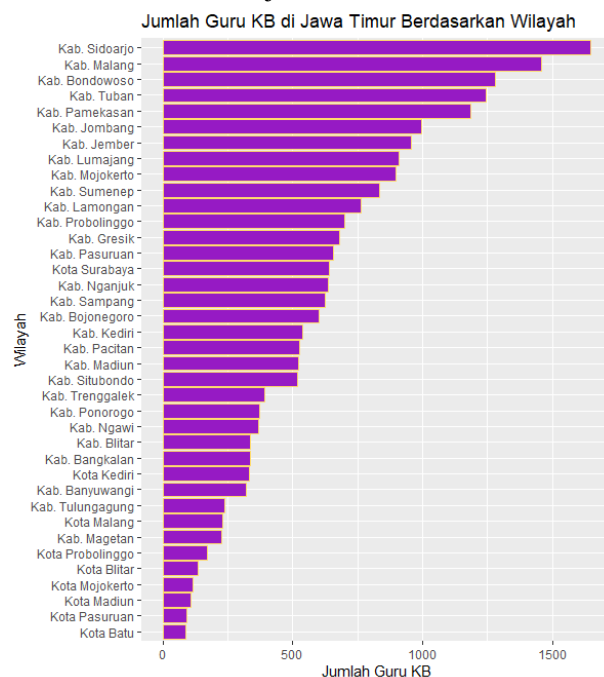
**Gambar 2** Perbandingan Persentase Jumlah Guru Berdasarkan Jenis Kelamin di Beberapa Lembaga Pendidikan

Dari Gambar 2 terlihat bahwa tenaga pendidik perempuan lebih banyak dibanding laki-laki. Jumlah tenaga pendidik laki-laki dan perempuan di SMK hampir seimbang. Lembaga TPA, SPS, dan TK memiliki persentase tenaga pendidik perempuan terbanyak dibandingkan dengan lembaga lainnya.



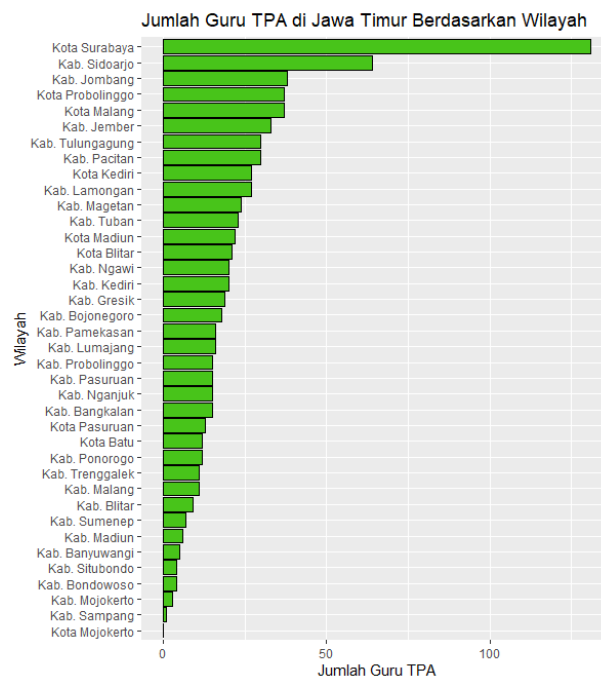
**Gambar 3** Persebaran Jumlah Guru TK di Jawa Timur

Dari gambar 3 dapat disimpulkan bahwa wilayah yang memiliki jumlah guru TK di Jawa Timur yang paling banyak adalah Kota Surabaya yaitu kurang lebih 4000 guru. Sementara wilayah dengan jumlah guru TK paling sedikit adalah Kota Mojokerto.



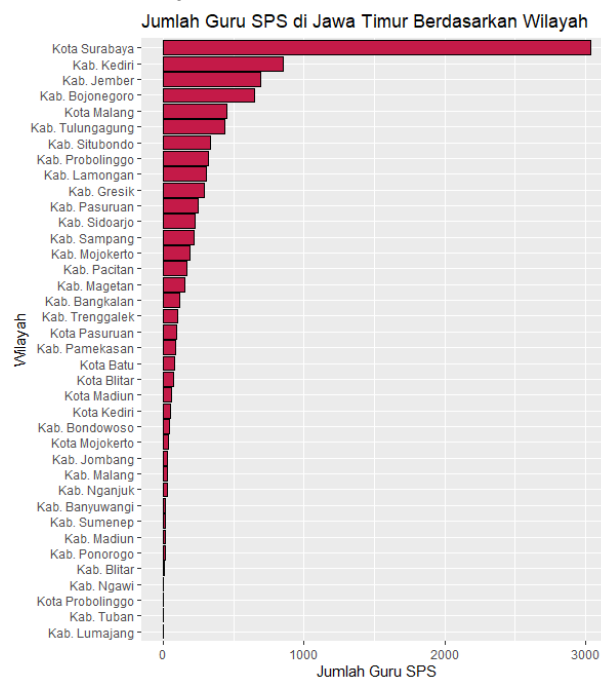
**Gambar 4** Persebaran Jumlah Guru KB di Jawa Timur

Gambar 4 menggambarkan bahwa wilayah dengan jumlah guru KB di Jawa Timur yang paling banyak adalah Kabupaten Sidoarjo yaitu kurang lebih 1600 guru. Sementara wilayah dengan jumlah guru TK paling sedikit adalah Kota Batu.



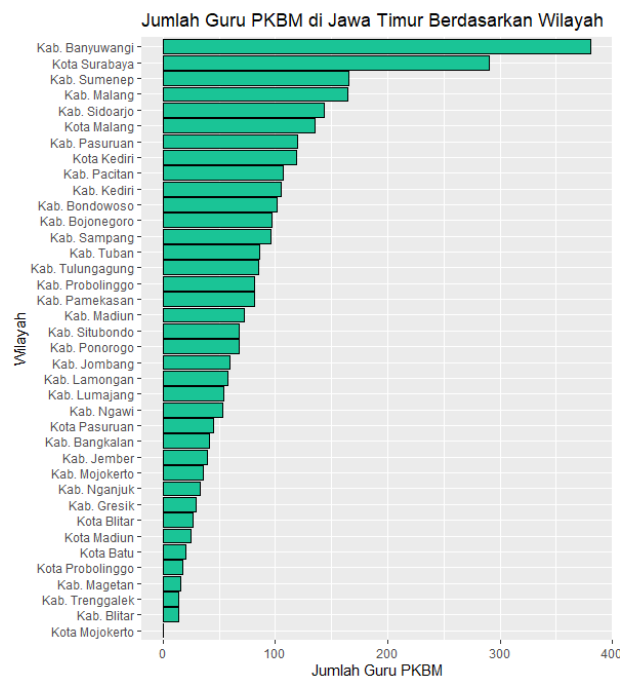
**Gambar 5** Persebaran Jumlah Guru TPA di Jawa Timur

Dari Gambar 5 dapat disimpulkan bahwa wilayah dengan jumlah guru TPA di Jawa Timur yang paling banyak adalah Kota Surabaya yaitu kurang lebih 130 guru. Sementara wilayah dengan jumlah guru TK paling sedikit adalah Kota Mojokerto.



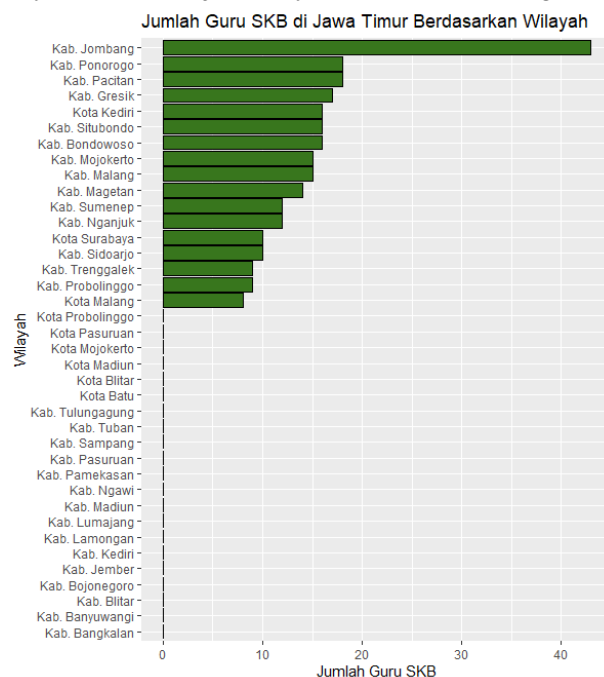
**Gambar 6** Persebaran Jumlah Guru SPS di Jawa Timur

Pada Gambar 6 dapat dilihat bahwa wilayah dengan jumlah guru SPS di Jawa Timur yang paling banyak adalah Kota Surabaya yaitu kurang lebih 3000 guru. Sementara wilayah Kabupaten Ngawi, Probolinggo, Tuban, Lumajang tidak memiliki guru SPS.



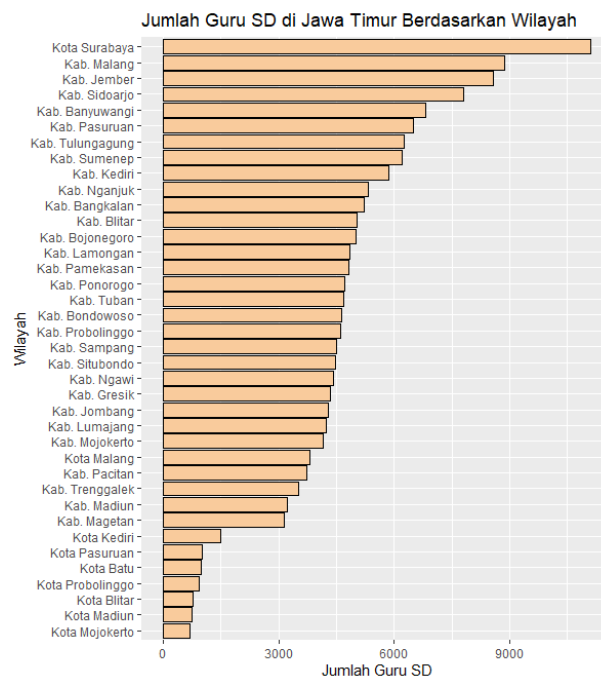
**Gambar 7** Persebaran Jumlah Guru PKMB di Jawa Timur

Gambar 7 memperlihatkan bahwa wilayah dengan jumlah guru PKBM di Jawa Timur yang paling banyak adalah Kabupaten Banyuwangi yaitu kurang lebih 380 guru. Sementara wilayah dengan jumlah guru PKBM paling sedikit yaitu Kota Mojokerto yaitu tidak memiliki guru PKBM.



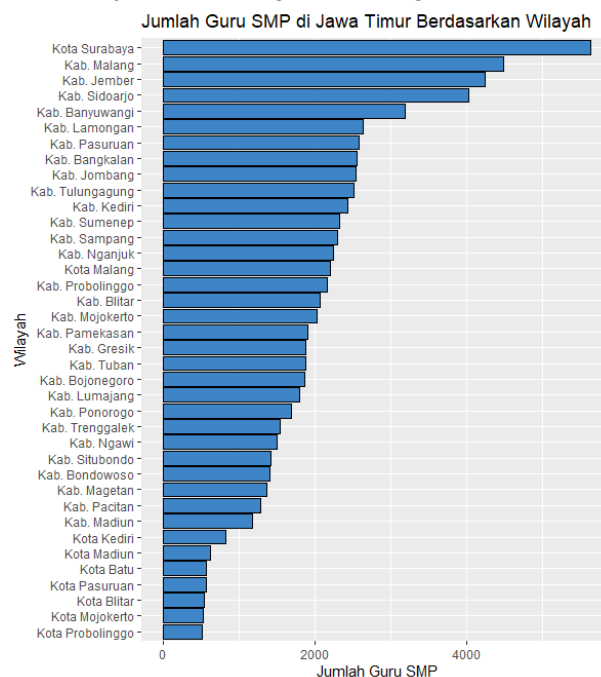
**Gambar 8** Persebaran Jumlah Guru SKB di Jawa Timur

Pada Gambar 8, dapat disimpulkan bahwa wilayah dengan jumlah guru SKB di Jawa Timur yang paling banyak adalah Kabupaten Jombang yaitu kurang lebih 43 guru. Sementara 21 wilayah kabupaten/kota di Jawa Timur tidak memiliki guru SKB.



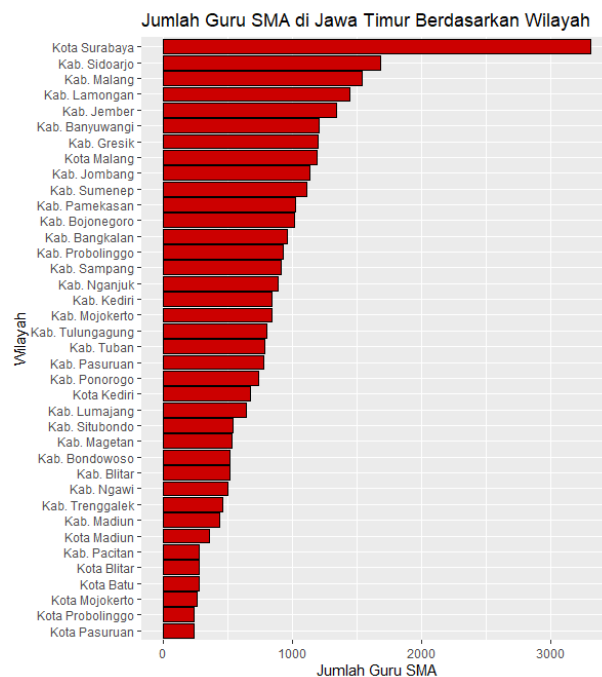
**Gambar 9** Persebaran Jumlah Guru SD di Jawa Timur

Dari Gambar 9 dapat disimpulkan, wilayah dengan jumlah guru SD di Jawa Timur yang paling banyak adalah Kota Surabaya yaitu kurang lebih 11.000 guru. Sementara wilayah dengan jumlah guru SD paling sedikit adalah Kota Mojokerto kurang lebih 500 guru.



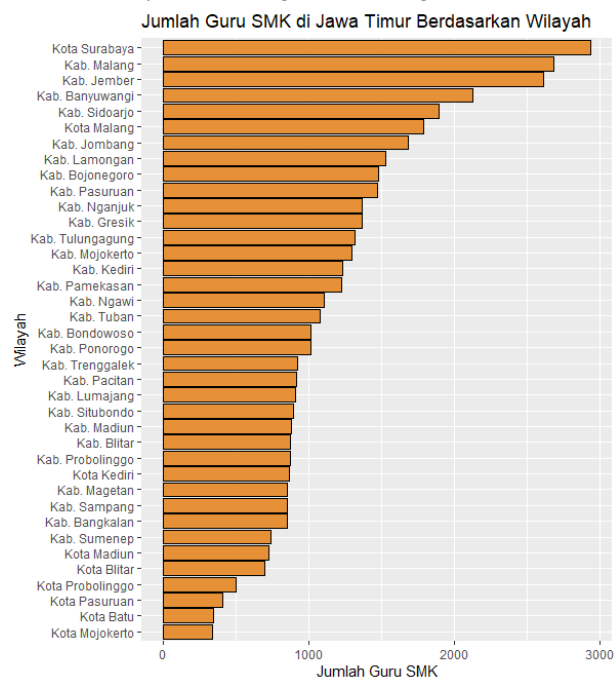
**Gambar 10** Persebaran Jumlah Guru SMP di Jawa Timur

Gambar 10 menggambarkan bahwa wilayah dengan jumlah guru SMP di Jawa Timur yang paling banyak adalah Kota Surabaya yaitu kurang lebih 5.700 guru. Sementara wilayah dengan jumlah guru TK paling sedikit adalah Kota Probolinggo yaitu kurang lebih 500 guru.



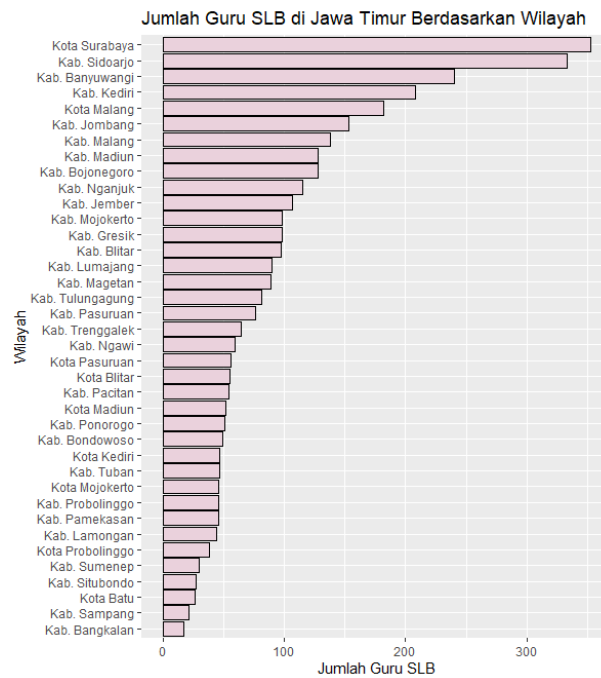
**Gambar 11** Persebaran Jumlah Guru SMA di Jawa Timur

Wilayah dengan jumlah guru SMA di Jawa Timur paling banyak yang ditampilkan pada Gambar 11 adalah Kota Surabaya yaitu kurang lebih 3.300 guru. Sementara wilayah dengan jumlah guru SMA paling sedikit adalah Kota Pasuruan yaitu kurang lebih 250 guru.



**Gambar 12** Persebaran Jumlah Guru SMK di Jawa Timur

Dari Gambar 12 dapat disimpulkan bahwa wilayah dengan jumlah guru SMK di Jawa Timur yang paling banyak adalah Kota Surabaya yaitu kurang lebih 2.900 guru. Sementara wilayah dengan jumlah guru SMK paling sedikit adalah Kota Mojokerto yaitu kurang lebih 300 guru.

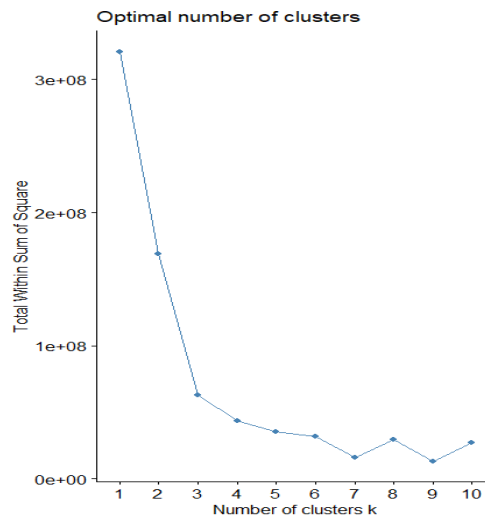


**Gambar 13** Persebaran Jumlah Guru SLB di Jawa Timur

Gambar 13 memperlihatkan bahwa wilayah dengan jumlah guru SLB di Jawa Timur yang paling banyak adalah Kota Surabaya yaitu kurang lebih 350 guru. Sementara wilayah dengan jumlah guru SLB paling sedikit adalah Kabupaten Bangkalan yaitu kurang lebih 20 guru.

## 2. Analisis *Clustering K-Means*

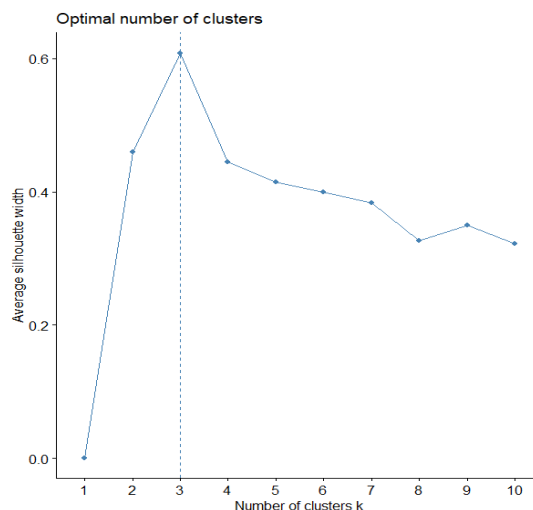
Sebelum dilakukan *clustering k-means*, perlu dihitung jumlah cluster optimum yang dapat terbentuk dari data tersebut. Digunakan dua metode untuk menentukan jumlah cluster optimum, yaitu metode Elbow dan metode *silhouette*. Berikut hasil perhitungan nilai k optimum dengan metode Elbow.



**Gambar 14** Grafik Perhitungan Nilai K Optimum Metode Elbow

Berdasarkan Gambar 14 dapat diartikan bahwa nilai k yang optimal untuk pengolahan data dengan menggunakan metode Elbow adalah sebanyak 3 cluster. Hal ini dikarenakan pada saat k sebesar 3, grafik seperti membentuk siku. Maka akan dilakukan clustering sebanyak 3 cluster. Berikut hasil perhitungan nilai k optimum dengan metode *silhouette*.





**Gambar 15** Grafik Perhitungan Nilai k Optimal Metode *Silhouette*

Berdasarkan Gambar 15 dapat diartikan bahwa nilai k yang optimal untuk pengolahan data menggunakan metode *Silhouette* adalah sebanyak 3 cluster. Hal ini dapat dilihat dari garis yang ditunjukkan oleh gambar. Maka dari kedua metode penentu tersebut didapatkan nilai k optimum yang dapat dibentuk adalah 3. Kemudian dilakukan pengelompokkan menggunakan clustering k-means dengan k=3.

Pada *clustering* k-means menggunakan SPSS, pertama diambil nilai pusat awal yang dipilih secara acak dari 3 observasi. Lalu untuk membentuk 3 *cluster*, perlu dilakukan 2 kali tahap iterasi yang kemudian akan membentuk nilai akhir pusat *cluster* yang digunakan untuk mengelompokkan. Tabel 1 menunjukkan anggota pada ketiga *cluster* yang terbentuk.

**Tabel 1** Persebaran Anggota Tiap *Cluster*

<i>Cluster</i>	Jumlah Anggota	Anggota
1	1	Kota Surabaya
2	5	Kab. Malang, Kab. Jember, Kab. Sidoarjo, Kab. Banyuwangi, dan Kab. Jombang
3	32	Kab. Pasuruan, Kab. Kediri, Kab. Lamongan, Kab. Tulungagung, Kab. Jombang, Kab. Nganjuk, Kab. Bojonegoro, Kab. Gresik, Kab. Pamekasan, Kota Malang, Kab. Tuban, Kab. Bangkalan, Kab. Mojokerto, Kab. Probolinggo, Kab. Blitar, Kab. Sampang, Kab. Bondowoso, Kab. Lumajang, Kab. Ponorogo, Kab. Ngawi, Kab. Situbondo, Kab. Trenggalek, Kab. Pacitan, Kab. Madiun, Kab. Magetan, Kota Kediri, Kota Madiun, Kota Blitar, Kota Pasuruan, Kota Probolinggo, Kota Batu, dan Kota Mojokerto

Kemudian dilakukan validasi cluster dengan uji F parsial (ANOVA) dengan taraf signifikansi ( $\alpha$ ) sebesar 0,05. Hipotesis, statistik uji dan daerah kritis dari pengujian ANOVA F parsial sebagai berikut.

**Hipotesis:**

$H_0$  : Variabel  $i$  bukan variabel pembeda dalam pengelompokkan.

$H_1$  : Variabel  $i$  merupakan variabel pembeda dalam pengelompokkan.

**Statistik Uji:**

$$F = \frac{\text{mean square cluster}}{\text{mean square error}}$$

**Daerah Kritis:**

Tolak  $H_0$  jika  $F > F$  tabel

Diketahui bahwa  $\alpha=0,05$ ,  $n=11$ , dan  $k=2$ , sehingga dalam tabel distribusi F diperoleh nilai F tabel sebesar 5,12. Pengujian dibantu dengan aplikasi SPSS dengan *output* pada Tabel 2.

**Tabel 2** Perhitungan ANOVA untuk Uji F Parsial

ANOVA						
Variabel	Cluster		Error		F	Sig.
	Mean Square	df	Mean Square	df		
TK	11.967	2	0.373	35	32.057	0.000
KB	4.182	2	0.818	35	5.112	0.011
TPA	13.275	2	0.299	35	44.463	0.000
SPS	15.399	2	0.177	35	86.917	0.000
PKBM	7.231	2	0.644	35	11.230	0.000
SKB	1.671	2	0.962	35	1.738	0.191
SD	8.583	2	0.567	35	15.146	0.000
SMP	11.895	2	0.377	35	31.516	0.000
SMA	13.049	2	0.312	35	41.890	0.000
SMK	12.444	2	0.346	35	35.959	0.000
SLB	10.871	2	0.436	35	24.938	0.000

Dari Tabel 2 didapatkan bahwa dua variabel yang memiliki nilai F hitung kurang dari nilai F tabel (gagal menolak  $H_0$ ) yaitu variabel KB dan SKB, sehingga kedua variabel tersebut bukan pembeda dalam pengelompokkan. Karena nilai F hitung selain kedua variabel tersebut lebih dari F tabel, maka variabel TK, TPA, SPS, PKBM, SD, SMP, SMA, SMK, dan SLB merupakan pembeda dalam pengelompokkan.

Kemudian untuk melihat karakteristik dari ketiga kelompok, dilihat dari mean tiap variabel berdasarkan kelompoknya.

**Tabel 3** Tabel Perhitungan Mean Setiap Cluster

Variabel	Cluster	Mean	Variabel	Cluster	Mean
TK	1	3999,0	SD	1	11108
	2	2489		2	7270
	3	1004,2		3	3871
KB	1	640,00	SMP	1	5633,0
	2	1075		2	3692
	3	520,2		3	1652
TPA	1	131,00	SMA	1	3310,0
	2	30,2		2	1382
	3	16,53		3	694,0
SPS	1	3040,0	SMK	1	2942,0
	2	199		2	2201
	3	170,5		3	993,3
PKBM	1	290,00	SLB	1	353,00
	2	157,2		2	194,2
	3	64,63		3	70,78
SKB	1	10,000			
	2	13,60			
	3	5,63			

Dari perhitungan diatas, dapat disimpulkan *cluster* 1 merupakan kelompok wilayah yang memiliki jumlah tenaga pendidik terbanyak. *Cluster* 2 merupakan kelompok wilayah yang memiliki jumlah tenaga pendidik terbanyak kedua, sedangkan *cluster* ke-3 merupakan kelompok wilayah yang memiliki jumlah tenaga pendidik terkecil.

#### D. Kesimpulan

Dari hasil analisis yang sudah kami lakukan, dapat disimpulkan bahwa:

1. Berdasarkan analisis secara statistika deskriptif terhadap jumlah tenaga pendidik berdasarkan lembaga pendidikan dan wilayah di Jawa Timur didapatkan hasil berikut.
  - Mayoritas dari tenaga pendidik di Jawa Timur merupakan perempuan. Lembaga TPA, SPS, dan TK memiliki persentase tenaga pendidik perempuan terbanyak,
  - Wilayah dengan jumlah guru TK di Jawa Timur yang paling banyak adalah Kota Surabaya sedangkan yang paling sedikit adalah Kota Mojokerto,
  - Wilayah dengan jumlah guru KB di Jawa Timur yang paling banyak adalah Kabupaten Sidoarjo dan yang paling sedikit adalah Kota Batu,
  - Wilayah dengan jumlah guru TPA di Jawa Timur yang paling banyak adalah Kota Surabaya dan yang paling sedikit adalah Kota Mojokerto,
  - Wilayah dengan jumlah guru SPS di Jawa Timur yang paling banyak adalah Kota Surabaya, sedangkan wilayah Kabupaten Ngawi, Probolinggo, Tuban, Lumajang tidak memiliki guru SPS,
  - Wilayah dengan jumlah guru PKBM di Jawa Timur yang paling banyak adalah Kabupaten Banyuwangi dan yang paling sedikit yaitu Kota Mojokerto.
  - Wilayah dengan jumlah guru SKB di Jawa Timur yang paling banyak adalah Kabupaten Jombang, sedangkan 21 kabupaten/kota di Jawa Timur tidak memiliki guru SKB.
  - Wilayah dengan jumlah guru SD di Jawa Timur yang paling banyak adalah Kota Surabaya dan yang paling sedikit adalah Kota Mojokerto.
  - Wilayah dengan jumlah guru SMP di Jawa Timur yang paling banyak adalah Kota Surabaya dan yang paling sedikit adalah Kota Probolinggo.
  - Wilayah dengan jumlah guru SMA di Jawa Timur yang paling banyak adalah Kota Surabaya dan yang paling sedikit adalah Kota Pasuruan.
  - Wilayah dengan jumlah guru SMK di Jawa Timur yang paling banyak adalah Kota Surabaya dan yang paling sedikit adalah Kota Mojokerto
  - Wilayah dengan jumlah guru SLB di Jawa Timur yang paling banyak adalah Kota Surabaya dan yang paling sedikit adalah Kabupaten Bangkalan.
2. Hasil analisis cluster jumlah tenaga pendidik dari 38 wilayah berdasarkan lembaga pendidikan dengan metode non-hirarki yaitu k-means *clustering* menunjukkan bahwa:
  - Jumlah *cluster* optimum untuk pengelompokkan adalah 3 *cluster* berdasarkan metode Elbow dan *silhouette*,
  - *Cluster* 1 merupakan kelompok dengan jumlah guru terbesar, berisi satu wilayah yaitu Kota Surabaya,
  - *Cluster* 2 merupakan kelompok dengan jumlah guru terbesar kedua, berisi lima wilayah yaitu Kab. Malang, Kab. Jember, Kab. Sidoarjo, Kab. Banyuwangi, Kab. Jombang,
  - *Cluster* 3 merupakan kelompok dengan jumlah guru paling kecil, berisi 32 wilayah kab/kota di Jawa Timur yang tidak termasuk di cluster 1 dan 2,
  - Variabel pembeda dalam pengelompokkan adalah TK, TPA, SPS, PKBM, SD, SMP, SMA, SMK, dan SLB.

## **Daftar Pustaka**

- [1] <https://dapo.kemdikbud.go.id/guru/1/050000>
- [2] Sitepu, Robinson, Irmeilyana, & Gultom, Berry (2011). Analisis Cluster terhadap Tingkat Pencemaran Udara pada Sektor Industri di Sumatera Selatan. *Jurnal Penelitian Sains*, Vol. 14, No. 3(A), 11 – 17