

Pengelompokan Lapangan Usaha Berdasarkan Level Skill Menggunakan *Multidimensional Scaling* dan *K-Means Clustering*

Aulia Nur Joviandi¹, Ayundyah Kesumawati¹

¹ Program Studi Statistika, Universitas Islam Indonesia, Jalan Kaliurang KM 14,5, Kabupaten Sleman, Daerah Istimewa Yogyakarta, 55584, Indonesia

*Corresponding author: 20611037@students.uii.ac.id



P-ISSN: 2986-4178
E-ISSN: 2988-4004

Riwayat Artikel

Dikirim: 02 September 2023
Direvisi: 04 Maret 2024
Diterima: 19 Maret 2024

ABSTRAK

Kementerian Ketenagakerjaan atau Kemnaker adalah salah satu kementerian atau penyelenggara urusan administrasi pemerintahan dalam bidang ketenagakerjaan atau tenaga kerja. Seseorang yang bekerja atau mampu melakukan kegiatan guna menghasilkan barang dan/atau jasa baik untuk memenuhi kebutuhan sendiri maupun untuk keperluan kelompok lain disebut tenaga kerja. Penempatan tenaga kerja pada lapangan usaha yang tersedia menjadi salah satu faktor pendukung pembangunan ekonomi yang dilakukan oleh pemerintah. Sehingga tenaga kerja berkeahlian perlu memperhatikan kelompok lapangan usaha sesuai *skill* yang dimiliki. Analisis *Multidimensional Scalling* dan *K-Means Clustering* digunakan untuk mengetahui kondisi tenaga kerja berkeahlian di *cluster* lapangan usaha yang ada. Dari penelitian ini, didapatkan bahwa pengelompokan menjadi tiga *cluster*, dengan *cluster* 1 beranggotakan 13 lapangan usaha, *cluster* 2 beranggotakan 2 lapangan usaha, dan *cluster* 3 beranggotakan 2 lapangan usaha.

Kata Kunci: *Multidimensional scaling*, *K-Means clustering*, tenaga kerja, lapangan usaha, level skill

ABSTRACT

Ministry of Manpower as called as Kemnaker is one of The Ministry that organize about government administrative which focus on manpower issues. A worker or labor is a person who works or is able to work to produce goods and/or services both to meet his own needs and for the community. Placement of labor in available business fields is one of the supporting factors for economic development carried out by the government. So that skilled workers need to pay attention to the business field groups according to the skills they have. *Multidimensional Scalling* and *K-Means Clustering* analysis are used to determine the condition of skilled workers in existing business field clusters. From this study, it was found that the grouping into three clusters, with cluster 1 consisting of 13 business fields, cluster 2 consisting of 2 business fields, and cluster 3 consisting of 2 business fields

Keywords: *Multidimensional scaling*, *K-Means clustering*, labor, business fields, skilled level

1. Pendahuluan

Pada perkembangan zaman saat ini, jumlah penduduk usia kerja di Indonesia semakin meningkat. Badan Pusat Statistik (BPS) mencatat bahwa ada lebih dari 135 juta penduduk bekerja pada tahun 2022. Angka ini menunjukkan kenaikan dibanding tahun sebelumnya, yaitu sebesar 131 juta penduduk bekerja [1]. Sedangkan persentase jumlah angkatan kerja terhadap penduduk usia kerja menurut kelompok umur mengalami kenaikan sebesar 0,83%, yaitu 67,8% pada tahun 2021 dan 68,63% pada tahun 2022 [2]. Hal ini tentunya menjadi kesempatan besar bagi Negara Indonesia untuk memperbaiki keadaan para tenaga kerja. Perbaikan keadaan tenaga kerja di sektor-sektor perkantoran dapat dimulai dari hal dasar, yaitu penempatan tenaga kerja yang sesuai dengan *skill* yang dimilikinya.

Sebagaimana yang disepakati pada *Strategic Planning Meeting* IMT-GT, ketiga negara yang tergabung dalam IMT-GT yaitu Indonesia, Malaysia, dan Thailand menyepakati menggunakan asumsi yang didekati dari jenis pekerjaan untuk mengkategorikan tenaga kerja berdasarkan *skill* yang dimiliki. Diambil dari Buku Ketenagakerjaan Dalam Data edisi 4 Tahun 2021, pekerja yang digolongkan sebagai *skilled labor* adalah pekerja sebagai manager, profesional, serta teknisi dan asisten profesional. Pekerja dengan golongan *semi-skilled* adalah pekerja sebagai tenaga tata usaha, tenaga usaha jasa dan tenaga penjualan, pekerja terampil pertanian, kehutanan dan perikanan, pekerja pengolahan, kerajinan dan yang berhubungan dengan itu serta operator dan perakitan mesin. Sementara pekerja dengan golongan *basic skilled* adalah pekerja sebagai pekerja kasar. Sementara itu, penduduk yang bekerja dengan jenis pekerjaan sebagai TNI/POLRI dianggap sebagai penduduk yang bekerja dengan keahlian khusus sehingga tidak diikutkan dalam pengelompokan penduduk bekerja berdasarkan level *skill* [3].

Masalah ketenagakerjaan masih menjadi salah satu masalah yang belum dapat terselesaikan. Hal ini terjadi karena adanya ketimpangan antara tenaga kerja dengan level *skill* tertentu yang bekerja tidak sesuai dengan bidang yang seharusnya. Fakta dilapangan mengatakan bahwa ada beberapa lapangan usaha yang seharusnya diisi oleh tenaga kerja dengan tingkat pendidikan yang lebih tinggi dan level *skill* yang setara, tetapi lebih banyak diisi oleh tenaga kerja dengan tingkat pendidikan dan level *skill* yang lebih rendah. Contohnya adalah lapangan usaha Pertambangan, Real Estate, dst [4]. Kondisi ini menjadi salah satu faktor meningkatnya angka pengangguran juga kualitas tenaga kerja yang tidak setara dengan kelas pekerjaan yang dimilikinya.

Dalam upaya membantu pemerataan tenaga kerja Indonesia, perlu ada kajian untuk mengetahui situasi tenaga kerja di lapangan usaha berdasarkan level *skill* yang dimiliki oleh tenaga kerja tersebut. Salah satu cara untuk mengetahui situasi tersebut, dapat dilakukan pengelompokan lapangan usaha yang idealnya memiliki kemiripan jumlah tenaga kerja dengan level *skill* tertentu. Dengan menggunakan analisis *Multidimensional scaling* dan *K-Means Clustering* terhadap pengelompokan lapangan usaha berdasarkan level *skill* akan diketahui keadaan kelompok-kelompok lapangan usaha yang berdekatan atau memiliki kemiripan antara jumlah tenaga kerja berdasarkan level *skill* tenaga kerja tersebut.

Penelitian Tsaniya Ahda Salsabila (2022) [5] tentang analisis *multidimensional scaling* pada pemetaan kabupaten/kota di Jawa Barat berdasarkan indikator kesejahteraan rakyat. Penelitian ini menghasilkan pemetaan dengan nilai STRESS 6,49% terkategori baik dan juga nilai RSQ sebesar 99,21% terkategori data yang digunakan dapat dipetakan dengan baik. Penelitian M.W. Talakua et al., (2017) [6] tentang analisis *cluster* dengan menggunakan metode *k-means* untuk pengelompokan kabupaten/kota di Provinsi Maluku berdasarkan indikator indeks pembangunan manusia tahun 2014. Berdasarkan kemiripan

karakteristik dari lima ukuran Indeks Pembangunan Manusia (IPM), didapatkan tiga *cluster* dan disimpulkan terdapat perbedaan yang signifikan pada tahun 2014.

Oleh karena itu, pada penelitian ini akan dilakukan pengelompokan lapangan usaha berdasarkan level *skill* tenaga kerja berkeahlian. Dengan menggunakan metode analisis *Multidimensional Scaling* dan *cluster* non hirarki yaitu *k-means clustering*. Tujuan penelitian ini adalah mengetahui gambaran umum dari data Tenaga Kerja Berkeahlian pada tahun 2022, mengetahui hasil pemetaan lapangan usaha pada peta *multidimensional scaling*, dan mengetahui perbedaan keadaan *cluster* yang didapatkan dari analisis *multidimensional scaling* dan *k-means clustering*.

2. Metodologi Penelitian

Penelitian ini menggunakan metode analisis *Multidimensional Scaling* dan *K-Means Clustering* dengan mengambil data Tenaga Kerja Berkeahlian Menurut Lapangan Usaha dan Level *Skill* tahun 2022. Analisis *Multidimensional Scaling* adalah suatu metode menggambarkan atau memetakan posisi-posisi obyek yang ditelitinya menggunakan persepsi kemiripan. Pemetaan ini dilakukan dengan menggunakan peta persepsi atau *perceptual map* [7]. Tahapan-tahapan dalam analisis *Multidimensional Scaling* adalah sebagai berikut [8].

- a. Menghitung matriks jarak dengan menggunakan jarak Euclidean. Kedekatan antar objek pada *perceptual map* dapat dihitung dengan menggunakan jarak Euclidean antara objek satu dengan objek lainnya sampai objek terakhir (ke-j) dengan rumus berikut ini:

$$d_{ij} = \sqrt{\sum_{k=1}^p (x_{ik} - x_{jk})^2} \quad (1)$$

atau [9]

$$d_{ij} = \|x_i - x_j\| \text{ untuk } i, j, \leq n \quad (2)$$

dengan d_{ij} = jarak antar objek ke-i dan objek ke-j, x_{ih} = hasil pengukuran objek ke-i pada peubah h, x_{jh} = hasil pengukuran objek ke-j pada peubah h

- b. Mencari nilai eigen dan eigen *vector* dengan rumus berikut ini:

$$\det(B - \lambda I) \text{ dan } \det(B - \lambda I) X \quad (3)$$

matriks B berupa elemen-elemen

$$b_{ij} = -\frac{1}{2}(d_{ij}^2 - d_{i.}^2 - d_{.j}^2 + d_{..}^2) \quad (4)$$

keterangan:

$$d_{i.}^2 = \frac{1}{n} \sum_i d_{ij}^2, \quad d_{.j}^2 = \frac{1}{n} \sum_j d_{ij}^2, \quad \text{dan} \quad d_{..}^2 = \frac{1}{n^2} \sum_{ii} d_{ij}^2 \quad (5)$$

- c. Membentuk koordinat objek berdasarkan eigen *vector* $X = [x_1 \ x_2]$, kemudian menghitung \hat{D} yang merupakan jarak Euclidean dari koordinat terbentuk
- d. Menghitung nilai *stress* dengan rumus berikut ini:

$$S = \left(\frac{\sum_{i=j}^n (d_{ij} - \hat{d}_{ij})^2}{\sum_{i=j}^n d_{ij}^2} \right) \quad (6)$$

Nilai *stress* digunakan untuk penentuan kecocokan model *Multidimensional Scaling*. Nilai *stress* yang kecil menunjukkan nilai *error* yang kecil juga. Tabel kriteria *stress* sebagai berikut [5].

Tabel 1. Tabel Kriteria Stress.

No	Stress	Kriteria
1.	$\geq 20\%$	Buruk
2.	$10\% - 20\%$	Cukup
3.	$5\% - 10\%$	Baik
4.	$2.5\% - 5\%$	Sangat baik
5.	$\leq 2.5\%$	sempurna

K-Means Clustering adalah salah satu metode *clusterisasi* non hirarki untuk mengelompokkan item, bukan variabel, yang dikelompokkan pada k *cluster* [10]. Pengelompokkan ini dilakukan dengan membagi sejumlah item ke dalam kelompok atau k kelompok. Analisis *cluster* ini bertujuan untuk mengelompokkan anggota item yang memiliki kemiripan sifat atau homogen, sedangkan antar kelompoknya bersifat heterogen. Analisis *K-Means clustering* harus memenuhi asumsi berikut [11]:

- a. Sampel yang diambil harus benar-benar bisa mewakili populasi.
- b. Sebaiknya tidak terjadi multikolinieritas yaitu korelasi antar objek, nilai maksimal dari multikolinieritas adalah 0.5.

Langkah melakukan pengelompokan pada metode *K-Means Clustering* adalah sebagai berikut [12]:

- a. Menginputkan item: $L = \{x_i, i = 1, 2, \dots, n\}$, dan menentukan besarnya k (jumlah *cluster*)
- b. Melakukan salah satu dari hal berikut:
 - o Mengalokasikan data ke dalam *cluster* secara random dan untuk k *cluster* akan dihitung nilai *centroid* saat ini, $\bar{x}_k, k = 1, 2, \dots, k$
 - o Sebelum menentukan *centroid* k *cluster*, menghitung nilai rata-rata, $\bar{x}_k, k = 1, 2, \dots, k$
- c. Menghitung jarak kuadrat Euclidean dari setiap item ke *centroid cluster* saat ini dengan rumus:

$$ESS = \sum_{k=1}^K \sum_{c(i)=k} (x_i - \bar{x}_k)^T (x_i - \bar{x}_k) \quad (7)$$

dengan: ESS adalah *Error Sum of Square*, \bar{x}_k adalah *centroid cluster* k , dan $c(i)$ adalah *cluster* yang berisi x_i .

- d. Menetapkan Kembali setiap item pengamatan ke pusat massa *cluster* terdekat sehingga nilai ESS berkurang. Perbarui *centroid cluster* setelah setiap penugasan kembali.
- e. Ulangi Langkah 3 dan 4 sampai tidak ada penugasan kembali item lebih lanjut, yang berarti tidak ada lagi objek yang berpindah *cluster*.

Ada beberapa metode optimalisasi jumlah *cluster* dalam analisis *cluster* non hirarki yang dapat digunakan sebagai acuan dalam menentukan banyak *cluster* yaitu metode *silhouette*, *elbow*, dan *gap-statistics*.

a. Metode *Silhouette*

Metode *silhouette* menunjukkan kualitas dan kekuatan *cluster*, mengevaluasi baik atau buruk obyek tersebut ditempatkan ke dalam suatu *cluster*. Metode ini merupakan gabungan dari metode *separation* dan *cohesion* [13]. Untuk menghitung nilai *silhouette* diperlukan perhitungan nilai *silhouette index* dari masing-masing objek. Nilai *silhouette* didapatkan dengan mencari nilai maksimal dari nilai *silhouette index* global dari jumlah *cluster* 2 sampai *cluster* $n-1$, dengan rumus [14]:

$$SC = \max_k SI(k) \quad (8)$$

dengan: SC = *silhouette coefficient*, SI = *silhouette index global*, dan k = jumlah *cluster*

b. Metode *Elbow*

Metode *elbow* merupakan salah satu metode untuk menentukan jumlah *cluster* yang tepat melalui persentase hasil perbandingan antara jumlah *cluster* yang akan membentuk siku pada suatu titik. Jika nilai *cluster* pertama dengan nilai *cluster* kedua memberikan sudut dalam grafik atau nilainya mengalami penurunan paling besar maka jumlah nilai *cluster* tersebut yang tepat. Untuk mendapatkan perbandingannya adalah dengan menghitung *error sum of square* (ESS) dari masing-masing nilai *cluster* [14].

c. Metode *Gap-Statistics*

Metode *gap-statistic* membandingkan nilai $\log W$ dan $\log E$ untuk masing-masing k *cluster*. Perkiraan *cluster* optimal akan menjadi nilai yang memaksimalkan statistik *gap* (yaitu, yang menghasilkan statistik *gap* terbesar). Penentuan jumlah *cluster* dapat dilihat dari hasil grafik *gap-statistic* yang terbentuk. [15], [16]

Setelah *cluster-cluster* terbentuk, profilisasi atau *profiling cluster* perlu dilakukan untuk mengetahui kondisi atau profil setiap *cluster* yang terbentuk. Profilisasi akan menghasilkan nilai rata-rata setiap *cluster*.

3. Hasil dan Pembahasan

3.1. Analisis Deskriptif

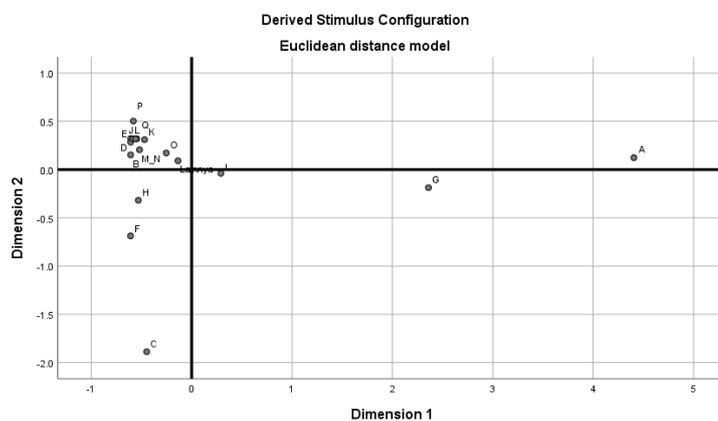
Tabel 2. Analisis deskriptif

	Minimal	Mean	Maksimal
<i>Skilled</i>	17776	681822	5533840
<i>Semi Skilled</i>	95171	4752735	37936925
<i>Basic Skilled</i>	125169	2524073	16837361

Dari hasil *output* di atas, dapat dilihat bahwa nilai minimum dari jumlah tenaga kerja berkeahlian di *skilled* level sebesar 17776 di lapangan usaha Sektor E dan nilai tertinggi berada pada nilai sebesar 5533840 yaitu jumlah tenaga kerja berkeahlian di *skilled* level lapangan usaha Sektor P. Pada tenaga kerja berkeahlian *semi skilled* level, memiliki jumlah tenaga kerja minimum sebesar 95171 di lapangan usaha Sektor D dan jumlah maksimumnya ada pada 37936925 di lapangan usaha Sektor A. Sedangkan tenaga kerja berkeahlian di *basic skilled* level memiliki jumlah minimum tenaga kerja sebesar 125169 di lapangan usaha sektor D dan nilai maksimumnya sebesar 16837361 di lapangan usaha sektor C.

3.2. *Multidimensional Scaling*

Diperoleh peta posisi dari masing-masing lapangan usaha dengan komposisi berikut di bawah ini:



Gambar 1. Peta *Multidimensional*

Plot pada **Gambar 1** menunjukkan posisi dari 17 golongan lapangan usaha berdasarkan level skill yaitu *skilled level*, *semi skilled level*, dan *basic skilled level* pada tahun 2022. Lapangan usaha yang jaraknya berdekatan dapat dinyatakan memiliki kemiripan berdasarkan indikator level *skill*.

Dari hasil pemetaan di atas, dapat dilihat hasil peta tersebut menunjukkan 5 kelompok lapangan usaha yang memiliki kedekatan jarak antar titik nya, tetapi berbeda dengan lainnya. Kelima kelompok tersebut, yaitu:

- Kelompok I : Kategori A
- Kelompok II : Kategori G
- Kelompok III : Kategori B, Kategori D, Kategori E, Kategori I, Kategori J, Kategori K, Kategori L, Kategori M dan N, Kategori O, Kategori P, Kategori Q, serta Kategori R, S, T, U dan lainnya
- Kelompok IV : Kategori H dan Kategori F
- Kelompok V : Kategori C

Dari hasil pengolahan pada *software* SPSS diperoleh nilai stress sebesar 0.05811. Nilai stress yang telah didapat dapat digunakan untuk menentukan kecocokan model hasil analisis *Multidimensional Scaling*. Berdasarkan Gambar 1 Tabel kriteria stress, 0.05811 termasuk ke dalam kategori sangat baik. Nilai tersebut juga dapat memberi indikasi bahwa *error* antara jarak dan nilai kemiripan dari ruang yang disajikan adalah kecil.

Setelah didapatkan *cluster-cluster* pengelompokannya, perlu dilakukan profilisasi menggunakan nilai rata-rata sehingga ditemukan kondisi setiap *cluster*:

Tabel 3 Profilisasi *cluster*

	<i>Skilled Level</i>	<i>Semi-Skilled Level</i>	<i>Basic Skilled Level</i>
<i>Cluster I</i>	53.946	37.936.925	713.125
<i>Cluster II</i>	227.455	22.557.201	3.409.234
<i>Cluster III</i>	838.634,33	1.459.138,33	780.541,75
<i>Cluster IV</i>	330.538,50	521.276,50	6.291.513,50
<i>Cluster V</i>	584.887	1.750.149	16.837.361

Berdasarkan Tabel 3, warna merah merupakan kelompok yang memiliki karakteristik rata-rata jumlah tenaga kerja yang tinggi, sedangkan warna hijau merupakan kelompok dengan karakteristik rata-rata jumlah tenaga kerja yang rendah. Dapat disimpulkan bahwa rata-rata jumlah tenaga kerja yang bekerja pada lapangan usaha di *cluster* I, II, IV dan V dengan nilai terendah dimiliki oleh tenaga kerja berkeahlian *Skilled level*. Sedangkan rata-rata jumlah tenaga kerja yang bekerja pada lapangan usaha di *cluster* I, II, dan III dengan nilai tertinggi dimiliki oleh tenaga kerja berkeahlian *Semi-Skilled level*.

Hasil *cluster* menggunakan analisis *Multidimensional Scaling* bersifat subjektif. Tidak terdapat teori mutlak untuk menentukan jumlah *cluster* dan pengelompokan objek pada *cluster* tertentu. Oleh karena itu, hasil profilisasi pada kelima *cluster* di **Tabel 2** terlihat tidak cukup rapi. Didapatkan pula nilai koordinat sebagai berikut:

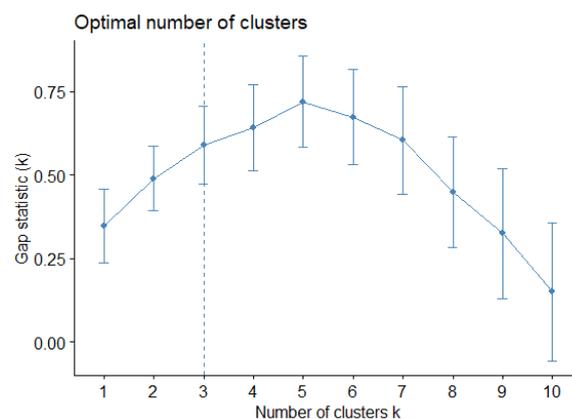
Lapangan Usaha	x	y
1. Kategori A Pertanian, Kehutanan, dan Perikanan	4.4046	0.1226
2. Kategori B Pertambangan dan Penggalian	-0.6071	0.1518
3. Kategori C Industri Pengolahan	-0.448	-1.8878
4. Kategori D Pengadaan Listrik dan Gas	-0.6064	0.321
5. Kategori E Pengadaan Air, Pengelolaan Sampah, Limbah, dan Daur Ulang	-0.6073	0.2838
6. Kategori F Kontruksi	-0.6081	-0.6874
7. Kategori G Perdagangan Besar dan Eceran; Reparasi dan Perawatan Mobil dan Sepeda Motor	2.3591	-0.1867
8. Kategori H Transportasi dan Pergudangan	-0.5309	-0.3179
9. Kategori I Penyediaan Akomodasi dan Makan Minum	0.2916	-0.0386
10. Kategori J Informasi dan Komunikasi	-0.5464	0.3194
11. Kategori K Jasa Keuangan dan Asuransi	-0.4685	0.3106
12. Kategori L Real Estat	-0.5869	0.3191
13. Kategori M, N Jasa Perusahaan	-0.5191	0.2049
14. Kategori O Administrasi Pemerintahan, Pertahanan dan Jaminan Sosial Wajib	-0.253	0.1726
15. Kategori P Jasa Pendidikan	-0.5803	0.503
16. Kategori Q Jasa Kesehatan dan Kegiatan Sosial	-0.5572	0.3179
17. Kategori R, S, T, U Jasa Lainnya	-0.1363	0.0915

Gambar 2. Koordinat multidimensional

Hasil koordinat ini akan digunakan untuk membuat pengelompokan menggunakan *K-Means clustering*.

3.2.1. K-Means Clustering

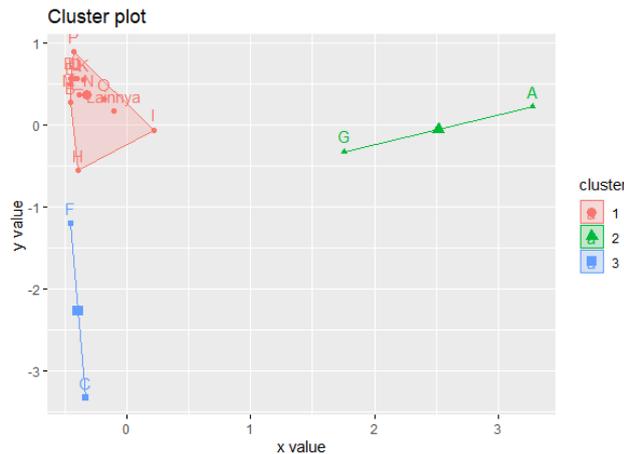
Analisis *K-Means Clustering* dimulai dengan penentuan jumlah *cluster* terlebih dahulu. Pada analisis ini, penentuan jumlah *cluster* menggunakan metode *gap-statistic*.



Gambar 3. Cluster optimal gap-statistic

Pada grafik di atas terlihat bahwa terdapat garis vertikal putus-putus pada *cluster* 3. Dengan adanya ciri tersebut, dapat diambil kesimpulan bahwa jumlah *cluster* yang optimal adalah 3 *cluster*.

Setelah jumlah *cluster* ditentukan, langkah selanjutnya adalah melakukan *K-Means Clustering* dengan menggunakan *software R*, didapatkan:



Gambar 4. Cluster lapangan usaha

Berdasarkan *output* dibentuk pengelompokan sebanyak tiga *cluster* untuk 17 kategori lapangan usaha. Hasil pengclusteran menunjukkan bahwa terdapat tiga belas kategori lapangan usaha pada *cluster* I, 2 kategori lapangan usaha pada *cluster* II, dan 2 kategori lapangan usaha pada *cluster* III. Rincian dari ketiga *cluster* tersebut adalah sebagai berikut.

- *Cluster* I meliputi Kategori B, Kategori D, Kategori E, Kategori H, Kategori I, Kategori J, Kategori K, Kategori L, Kategori M dan N, Kategori O, Kategori P, Kategori Q, serta Kategori R, S, T, U dan lainnya
- *Cluster* II meliputi Kategori A dan Kategori G
- *Cluster* III meliputi Kategori C dan Kategori F

Setelah didapatkan *cluster-cluster* pengelompokannya, perlu dilakukan profilisasi sehingga ditemukan kondisi setiap *cluster*.

Tabel 4. Profilisasi *cluster*

	<i>Cluster</i> I	<i>Cluster</i> II	<i>Cluster</i> III
<i>Skilled</i> level	787,978.62	140,700.50	532,927.00
<i>Semi Skilled</i> Level	1,406,283.08	30,247,063.00	1,010,341.00
<i>Basic Skilled</i> Level	1,093,821.46	2,061,179.50	12,283,605.00

Berdasarkan Tabel 3, warna merah merupakan kelompok yang memiliki karakteristik rata-rata jumlah tenaga kerja yang tinggi, sedangkan warna hijau merupakan kelompok dengan karakteristik rata-rata jumlah tenaga kerja yang rendah. Dapat disimpulkan bahwa rata-rata jumlah tenaga kerja yang bekerja pada lapangan usaha di *cluster* I, II, dan III dengan nilai terendah dimiliki oleh tenaga kerja berkeahlian *Skilled* level. Sedangkan rata-rata jumlah tenaga kerja yang bekerja pada lapangan usaha di *cluster* I dan II dengan nilai tertinggi dimiliki oleh tenaga kerja berkeahlian *Semi-Skilled* level. Rata-rata jumlah tenaga kerja yang bekerja pada lapangan usaha di *cluster* III dengan nilai tertinggi dimiliki oleh tenaga kerja berkeahlian *Basic-Skilled* level.

Sehingga lapangan usaha pada *cluster* I dan lapangan usaha pada *cluster* II cukup memiliki kemiripan. Hasil ini juga dipengaruhi oleh jumlah tenaga kerja berkeahlian pada setiap sektor atau kategori. *Cluster* II dan *Cluster* III memiliki jumlah kategori lapangan usaha yang sama yaitu 2 kategori di setiap *Cluster*, tetapi rata-rata jumlah tenaga kerja berkeahlian *Semi-Skilled* pada *Cluster* I jauh berbeda dengan rata-rata jumlah tenaga kerja berkeahlian *Semi-Skilled* pada *Cluster* II. Fakta ini menunjukkan bahwa terdapat perbedaan yang signifikan pada jumlah tenaga kerja berkeahlian *Semi-Skilled* level dan *Basic-Skilled* level pada Kategori Lapangan Usaha A Pertanian, Kehutanan, dan Perikanan dan Kategori G Perdagangan Besar dan Eceran; Reparasi dan Perawatan Mobil dan Sepeda Motor terhadap Kategori Lapangan Usaha C Industri Pengolahan dan Kategori F Kontruksi.

4. Kesimpulan

Berdasarkan analisis *Multidimensional Scalling* dan *K-Means Clustering* tentang data Tenaga Kerja Berkeahlian Menurut Lapangan Usaha Berdasarkan Level Skill maka dapat ditarik kesimpulan sebagai berikut :

1. Kategori E Pengadaan Air, Pengelolaan Sampah, Limbah, dan Daur Ulang memiliki jumlah tenaga kerja dengan keahlian *Skilled* level terendah dibandingkan dengan tenaga kerja berkeahlian lainnya. Sedangkan Kategori A Pertanian, Kehutanan, dan Perikanan memiliki jumlah tenaga kerja dengan keahlian *Semi-Skilled* level tertinggi dibandingkan dengan tenaga kerja berkeahlian lainnya.
2. Hasil analisis *Multidimensional Scalling* pada penelitian ini menunjukkan nilai *stress* sebesar 0.05811. Nilai ini dapat dikategorikan sebagai hasil analisis yang sangat baik sesuai pada tabel kriteria *stress*. Oleh karena itu, hasil titik koordinat atau posisi lapangan usaha pada peta *Multidimensional Scalling* dapat digunakan untuk langkah analisis selanjutnya.
3. Berdasarkan analisis *Multidimensional Scalling*, diperoleh 5 kelompok lapangan usaha, sedangkan berdasarkan analisis *K-Means Clustering* diperoleh 3 kelompok atau *cluster* lapangan usaha. Kelompok III berdasarkan *Multidimensional Scalling* dan *Cluster I* berdasarkan *K-Means Clustering* memiliki anggota lapangan usaha yang sama. Pada kelompok atau *cluster* lainnya terdapat perbedaan. Perbedaan ini disebabkan oleh jumlah kelompok dan pengelompokan yang dilakukan pada analisis *Multidimensional Scalling* bersifat subjektif sehingga akan sangat berbeda dengan hasil pengelompokan menggunakan analisis *K-Means Clustering* dengan jumlah *cluster* ditentukan dengan metode *gap-statistic* dan pengelompokan secara objektif dengan memperhatikan jarak antar objek dan jumlah *K cluster*.

5. Daftar Pustaka

- [1] “Jumlah dan Persentase Penduduk Bekerja dan Pengangguran,” Badan Pusat Statistik. Diakses: 19 Maret 2024. [Daring]. Tersedia pada: <https://www.bps.go.id/id/statistics-table/2/MTk1MyMy/jumlah-dan-persentase-penduduk-bekerja-dan-pengangguran.html>
- [2] “Persentase Angkatan Kerja Terhadap Penduduk Usia Kerja (TPAK) Menurut Golongan Umur,” Badan Pusat Statistik. Diakses: 19 Maret 2024. [Daring]. Tersedia pada: <https://www.bps.go.id/id/statistics-table/2/NzE0IzI%3D/persentase-angkatan-kerja-terhadap-penduduk-usia-kerja--tpak--menurut-golongan-umur.html>
- [3] Zulfiyandi, F. A. Wirawan, N. P. S. Tanjung, R. Yolanda, dan M. Zaini, *Ketenagakerjaan Dalam Data*, Edisi 4. Jakarta Selatan: Pusat Data dan Teknologi Informasi Ketenagakerjaan, 2021. [Daring]. Tersedia pada: <https://satudata.kemnaker.go.id>
- [4] “Penduduk Berumur 15 Tahun Ke Atas yang Bekerja Selama Seminggu yang Lalu Menurut Lapangan Pekerjaan Utama (17 Sektor) dan Pendidikan Tertinggi yang Ditamatkan,” Badan Pusat Statistik. Diakses: 19 Maret 2024. [Daring]. Tersedia pada: <https://www.bps.go.id/id/statistics-table/2/MTk3MSMy/penduduk-berumur-15-tahun-ke-atas-yang-bekerja-selama-seminggu-yang-lalu-menurut-lapangan-pekerjaan-utama--17-sektor--dan-pendidikan-tertinggi-yang-ditamatkan.html>
- [5] T. A. Salsabila dan L. Wachidah, “Analisis *Multidimensional Scalling* pada Pemetaan Kabupaten/Kota di Jawa Barat Berdasarkan Indikator Kesejahteraan Rakyat,” *Bandung Conference Series: Statistics*, vol. 2, no. 2, hlm. 173–179, Jul 2022, doi: 10.29313/bcss.v2i2.3701.
- [6] M. W. Talakua, Z. A. Leleury, dan A. W. Talluta, “Analisis Cluster dengan Menggunakan Metode *K-Means* untuk Pengelompokan Kabupaten/Kota di

- Provinsi Maluku Berdasarkan Indikator Indeks Pembangunan Manusia Tahun 2014,” *Jurnal Ilmu Matematika dan Terapan*, vol. 11, no. 2, hlm. 119–128, Des 2017.
- [7] H. Nadya Novita, “Analisis Multidimensional Scaling dan Cluster Fuzzy C-Means Pada Data Indikator Kesejahteraan Rakyat di Indonesia”.
- [8] J. Nahar, “Penerapan Metode Multidimensional Scaling dalam Pemetaan Sarana Kesehatan di Jawa Barat,” vol. 12, no. 1, hlm. 43–50, 2016, [Daring]. Tersedia pada: <http://jabar.bps.go.id>
- [9] L. Liberti, C. Lavor, N. Maculan, dan A. Mucherino, “Euclidean Distance Geometry and Applications,” *SIAM Review*, vol. 56, no. 1, hlm. 3–69, 2014, doi: 10.1137/120875909.
- [10] R. A. Johnson dan D. W. Winchern, *Applied Multivariate Statistical Analysis*. New Jersey: Prentice Hall, 2007.
- [11] J. F. Hair, W. C. Black, B. J. Babin, dan R. Anderson, *Multivariate Data Analysis*, 7 ed. New Jersey: Pearson Prentice Hall, 2010.
- [12] M. Muhajir, *Modul Praktikum Statistika Multivariat Terapan*, I. Yogyakarta: Universitas Islam Indonesia, 2021.
- [13] S. Paembonan dan H. Abduh, “Penerapan Metode Silhouette Coeficient Untuk Evaluasi Clustering Obat,” *PENA TEKNIK*, vol. 6, no. 2, Sep 2021, [Daring]. Tersedia pada: <https://ojs.unanda.ac.id/index.php/jiit/index>
- [14] A. I. Cahya Dewi dan A. K. Pramita, “Analisis Perbandingan Metode Elbow dan Silhouette pada Algoritma Clustering K-Medoids dalam Pengelompokan Produksi Kerajinan Bali,” *Jurnal Matrix*, vol. Vol. 9 No. 3, 2019.
- [15] F. A. Dewa dan M. T. Jatipaningrum, “Segmentasi E-Commerce dengan Cluster K-Means dan Fuzzy C-Means (Studi Kasus: Media Sosial di Indonesia yang diunduh di Play Store),” *Jurnal Statistika Industri dan Komputasi*, vol. 4, no. 1, hlm. 53–67, 2019.
- [16] Data Novia, “Determining TheOptimal Number Of Clusters: 3 Must Know Methods.” Diakses: 14 Mei 2023. [Daring]. Tersedia pada: <https://www.datanovia.com/en/lessons/determining-the-optimal-number-of-clusters-3-must-know-methods/>