

# KLASIFIKASI USAHA INDUSTRI DI KABUPATEN LUWU UTARA BERDASARKAN INDIKATOR INDUSTRI KECIL DAN INDUSTRI MENENGAH MENGGUNAKAN METODE *AVERAGE LINKAGE CLUSTERING*

Riani Sintia Dewi<sup>i</sup>

Risnawati Ibtnas<sup>ii</sup>

M. Ichsan Nawawi<sup>iii</sup>

---

<sup>i</sup> Mahasiswa Program Studi Matematika-FST, UINAM

<sup>ii</sup> Prodi Matematika FST, UINAM

<sup>iii</sup> Prodi Matematika FST, UINAM

---

**ABSTRAK.** Penelitian ini membahas tentang klasifikasi usaha industri kecil dan industri menengah di Kabupaten Luwu Utara menggunakan metode *Average Linkage Clustering*. Metode *Average Linkage Clustering* dianggap lebih stabil, dimana nilai rasio simpangan baku yang dihasilkan lebih kecil dibandingkan dengan metode-metode *Cluster Hierarky* lainnya. Kabupaten Luwu Utara merupakan suatu daerah dengan pertumbuhan ekonomi yang terus meningkat setiap tahunnya dan terdiri dari 48 bidang usaha industri yang berbeda-beda. Adapun penelitian ini bertujuan untuk mengklasifikasikan usaha industri di kabupaten Luwu Utara berdasarkan indikator industri kecil dan industri menengah menggunakan metode *Average Linkage Clustering*. Hasil penelitian diperoleh, dari pengklasifikasian 48 usaha industri di kabupaten Luwu Utara yang dibentuk dalam dua *cluster*, yaitu *cluster A* merupakan *cluster* industri kecil yang terdiri dari 43 usaha industri dan *cluster B* merupakan *cluster* industri menengah yang terdiri dari 5 usaha industri.

---

**Kata Kunci:** *Usaha Industri, Average Linkage Clustering*

## 1. PENDAHULUAN

Industri kecil dan industri menengah merupakan suatu bidang usaha yang mempunyai peran penting di setiap wilayah dalam pembangunan ekonomi. Terbukti, pada tahun 2016 kemampuan industri kecil dan industri menengah untuk menyerap tenaga kerja sangat tinggi yaitu mencapai 97.22%.

Perkembangan penduduk dari tahun ke tahun di kabupaten Luwu Utara menyebabkan pemerintah harus terus menambah jumlah fasilitas hidup layak bagi masyarakat termasuk dengan meningkatkan perekonomian daerah Luwu Utara. Berdasarkan data Badan Pusat Statistik (BPS) yang terangkum dalam buku "Luwu Utara dalam Angka Tahun 2018"

menunjukkan pertumbuhan ekonomi Luwu Utara sebesar 7,60% yang berarti mengalami peningkatan sebesar 0,11% dari tahun sebelumnya[1]. Sektor pertanian memegang peran penting dalam peningkatan perekonomian di kabupaten Luwu Utara. Namun, selain sektor pertanian usaha industri juga berperan dalam proses pembangunan ekonomi. Sehingga, untuk lebih meningkatkan perekonomian di kabupaten Luwu Utara diperlukan perlakuan khusus untuk usaha industri dengan cara mengklasifikasikan usaha industri untuk mengetahui industri mana yang masuk dalam kategori industri kecil atau industri menengah yang ditinjau melalui variabel-variabel pengamatan.

Analisis multivariat mempunyai banyak metode khususnya dalam bidang analisis. Adapun permasalahan yang berkaitan dengan hal pengelompokkan objek-objek yang mempunyai kesamaan karakteristik analisis yang dapat digunakan yaitu Analisis *Cluster*. Dalam analisis *cluster* memiliki dua metode, yaitu metode hirarki dan non-hirarki. Dalam metode hirarki terdapat beberapa teknik analisis, yaitu *Complete Linkage*, *Average Linkage*, *Single Linkage*, dan metode *Ward's*. Sedangkan pada metode non-hirarki terdapat satu teknik analisis yaitu metode *K-Means*.

Pada penelitian ini digunakan metode *Average Linkage* dalam proses peng-*cluster*-an. Metode *Average Linkage* lebih stabil dimana nilai rasio simpangan baku yang dihasilkan lebih kecil dibandingkan dengan tiga metode lainnya yaitu *Single Linkage*, *Complete Linkage*, dan *Ward's*. Sementara itu, untuk perhitungan jaraknya peneliti menggunakan jarak *Euclidean* sebagai

ukuran jarak antar objek, dikarenakan ukuran jarak *Euclidean* paling umum digunakan dan mudah dalam pengaplikasiannya serta tingkat akurasi kemiripan yang lebih tinggi dibanding dengan ukuran jarak yang lain.

## 2. TINJAUAN PUSTAKA

### ANALISIS MULTIVARIAT

Analisis Multivariat adalah salah satu jenis analisis statistik yang digunakan untuk menganalisis data yang terdiri dari banyak variabel baik variabel bebas (*independent variables*) maupun variabel tak bebas (*dependent variables*). Dengan menggunakan teknik analisis ini peneliti bisa menganalisis pengaruh beberapa variabel terhadap variabel lainnya dalam waktu yang bersamaan. Data multivariat merupakan data yang dikumpulkan dari dua atau lebih observasi dengan mengukur observasi tersebut dengan beberapa karakteristik. Analisis Multivariat dibagi dalam dua kategori metode, yaitu metode dependensi dan interdependensi. Metode dependensi terdiri dari dua jenis variabel, yaitu variabel bebas dan variabel tergantung. Sedangkan, pada metode interdependensi hanya terdiri dari satu jenis variabel, yaitu variabel bebas. Analisis dependensi dapat digunakan untuk menerangkan atau memprediksi variabel tergantung (*dependent variables*) dengan menggunakan dua atau lebih variabel bebas. Analisis yang termasuk dalam klasifikasi ini yaitu Analisis Regresi Linear Berganda, Analisis Diskriminan, Analisis Korelasi Kanonikal, dan Analisis Varian Multivariat (MANOVA). Adapun Analisis interdependensi berfungsi untuk memberikan makna terhadap seperangkat variabel atau membuat kelompok-kelompok secara bersamaan. Analisis yang termasuk dalam klasifikasi ini adalah Analisis Faktor, Analisis *Cluster*, dan *Multidimensional Scaling*[2].

### ANALISIS CLUSTER

*Clustering* merupakan proses pembuatan kelompok, sehingga semua anggota kelompok dari setiap partisi memiliki kesamaan berdasarkan matriks tertentu. Objek data yang berada dalam *cluster* harus mempunyai kemiripan, sedangkan yang tidak berada dalam

satu *cluster* tidak mempunyai kemiripan. Analisis *Cluster* merupakan teknik yang digunakan untuk mengklasifikasikan objek atau kasus (responden) kedalam kelompok yang relatif homogen (yang dinamakan *cluster*). Setiap objek dalam pengelompokan, hanya bisa masuk kedalam satu *cluster* saja, sehingga tidak terjadi tumpang tindih (*overlapping/interaction*)[3].

Analisis *cluster* yaitu teknik multivariat yang mempunyai tujuan utama untuk mengelompokkan objek-objek berdasarkan karakteristik yang dimiliki. Dalam proses pengelompokan tersebut terdapat dua metode yang dapat digunakan untuk proses peng-*clusteran*, yaitu metode hirarki dan metode non-hirarki[4]. Adapun dalam penelitian ini hanya berfokus pada penggunaan metode *Clustering Hirarki*.

### METODE HIRARKI

Metode hirarki merupakan metode yang memulai pengelompokkannya dengan dua atau lebih objek yang mempunyai kemiripan paling dekat. Keuntungan metode hirarki yaitu cepat dalam proses pengolahan sehingga dapat meminimalisir waktu dalam proses pengerjaannya.

Berikut merupakan tahap-tahap peng-*clusteran* data dengan menggunakan metode hirarki, yaitu:[5]

1. Menentukan  $k$  yang merupakan jumlah *cluster* yang akan dibentuk.
2. Anggap setiap data objek sebagai *cluster*. Dimana, jika  $n$  = jumlah data dan  $c$  = jumlah *cluster*, maka  $c = n$ .
3. Selanjutnya, menghitung jarak antar *cluster*.
4. Setelah jarak didapatkan, maka cari dua data yang mempunyai jarak paling minimal dan gabungkan kedalam satu *cluster* ( $n = c - 1$ ).
5. Jika  $n > 3$ , maka ulangi langkah 3.

### AVERAGE LINKAGE

Metode *Average Linkage* yaitu salah satu metode *Agglomerative* dalam proses *clustering* hirarki yang didasarkan pada jarak rata-rata antar objeknya. Adapun untuk menghitung jarak dua *cluster* pada metode *Average Linkage* yang merupakan jarak rata-rata dapat dihitung menggunakan persamaan rumus berikut :[6]

$$d_{(UV)W} = \frac{\sum_i \sum_k d_{ik}}{N_{(UV)} N_W}$$

Dengan :

$d_{ik}$  = Merupakan jarak antar objek  $i$  dalam  $(UV)$  dan objek  $k$  dalam  $cluster$ .

$N_{(UV)}$  dan  $N_W$  = Merupakan jumlah objek dalam  $cluster (UV)$  dan  $(W)$ .

### INDUSTRI KECIL DAN INDUSTRI MENENGAH

Berdasarkan Badan Pusat Statistik industri kecil dan industri menengah adalah usaha rumah tangga yang memiliki kegiatan usaha yang sama yaitu kegiatan dalam bidang produksi. Industri kecil dan industri menengah bisa dibedakan dari jumlah tenaga kerja dan nilai investasi yang digunakan untuk melakukan suatu produksi.

Selain itu, Dinas Perindustrian dan Perdagangan (Disperindag) dalam RIP-IKM (2002-2004) mendefinisikan industri kecil sebagai kegiatan ekonomi yang dilakukan oleh perindividu atau rumah tangga maupun suatu badan, bertujuan untuk memproduksi barang maupun jasa untuk diperdagangkan secara komersial, yang mempunyai nilai kekayaan bersih paling banyak 200 juta rupiah dan mempunyai nilai penjualan pertahun sebesar 1 milyar rupiah atau kurang. Sedangkan Industri menengah adalah kegiatan ekonomi yang dilakukan oleh perindividu atau badan, bertujuan untuk memproduksi barang ataupun jasa untuk di perdagangkan secara komersial yang mempunyai nilai penjualan pertahun lebih besar dari 1 milyar rupiah namun kurang dari 50 milyar rupiah[7].

### 3. METODOLOGI

Data usaha industri dari Dinas Perindustrian Provinsi Sulawesi Selatan tahun 2018 menjadi objek dari analisis *Clustering* dengan indikator industri kecil dan industri menengah. Adapun variabel-variabel yang diperhatikan yaitu Jumlah Tenaga Kerja (Orang) ( $X_1$ ) dan Nilai Investasi (Rp.) ( $X_2$ ).

### Prosedur Penelitian

Langkah-langkah analisis *clustering* dari usaha industri, dijelaskan sebagai berikut :

1. Mengumpulkan data usaha industri yaitu data jumlah tenaga kerja dan nilai investasi serta melakukan pengkategorian variabel.
2. melakukan teknis analisis dengan langkah-langkah sebagai berikut :
  - a. Merancang penelitian dalam analisis *cluster*.

1. Standarisasi data, dapat dilakukan dengan rumus berikut :

$$z = \frac{x - \bar{x}}{s}$$

dengan :

$x$  = Nilai data

$\bar{x}$  = Nilai rata-rata

$s$  = Standar deviasi

2. Mendeteksi *outlier*, data dikatakan outlier jika nilai  $z$  tidak berada pada selang  $(-2.5; +2.5)$

- b. Asumsi-asumsi

1. Uji kecukupan sampel, *kaiser-Meyer Olkin (KMO)* yang kurang dari 0.5 menyatakan bahwa sampel yang diteliti tidak dapat mewakili populasi yang ada atau sampel tidak layak untuk digunakan
2. Uji multikolinearitas, jika tidak ada nilai korelasi yang melebihi 0.80 maka model tersebut dinyatakan tidak terjadi multikolinearitas.

- c. Menentukan ukuran kemiripan objek menggunakan jarak *Euclidean*, dapat dihitung menggunakan persamaan rumus berikut :

$$d_{(xy)}^2 = \sqrt{\sum_{j=1}^p (x_j - y_j)^2}$$

Dengan :

$d_{(xy)}^2$  = Kuadrat jarak *euclidean* antara objek X dan Y

$p$  = Banyaknya variabel yang diamati

$x_j$  = Nilai  $j$  pada objek x

$y_j$  = Nilai  $j$  pada objek y

- d. Proses analisis *clustering* menggunakan metode *Average Linkage*, dapat dihitung menggunakan persamaan rumus berikut :

$$d_{(UV)W} = \frac{\sum_i \sum_k d_{ik}}{N_{(UV)}N_W}$$

Dengan :

$d_{ik}$  = Merupakan jarak antar objek  $i$  dalam  $(UV)$  dan objek  $k$  dalam *cluster*.

$N_{(UV)}$  dan  $N_W$  = Merupakan jumlah objek dalam *cluster*  $(UV)$  dan  $(W)$

- e. Menentukan Jumlah *cluster* dan anggotanya.
- f. Melakukan interpretasi *cluster*, Proses ini dimulai dengan suatu ukuran yang sering digunakan yaitu *centroid cluster*. Adapun untuk membandingkan *centroid* pada tiap-tiap *cluster*, dapat dihitung menggunakan rumus berikut :

$$C = \frac{\sum_{j=1}^n \sum_{k=1}^p x_{jk}}{N}$$

#### 4. PEMBAHASAN

##### *Profil Data*

Terdapat 48 usaha industri yang akan dilakukan proses *clustering*, dimana akan dibentuk kedalam dua kelompok yaitu *cluster* industri kecil dan *cluster* industri menengah.

##### *Statistika Deskriptif*

untuk mengetahui gambaran secara umum karakteristik dari setiap indikator atau variabel-variabel yang digunakan maka ditampilkan statistika deskriptif berikut :

Tabel 1 Statistika Deskriptif Jumlah Tenaga Kerja ( $X_1$ )

| Min. | Mean | Maks. |
|------|------|-------|
| 2    | 145  | 1165  |

Tabel 2 Statistika Deskriptif Nilai Investasi ( $X_2$ )

| Min.    | $Q_1$    | Median    |
|---------|----------|-----------|
| 1000000 | 30750000 | 183000000 |

| Mean       | $Q_3$     |
|------------|-----------|
| 1317000000 | 927000000 |

##### *Merancang Penelitian dalam Analisis Cluster Standarisasi Data*

Standarisasi data dilakukan jika terdapat perbedaan satuan yang tidak signifikan antar variabel-variabel yang diteliti. Adapun proses standarisasi data dapat dilakukan menggunakan persamaan rumus berikut :

$$z = \frac{x - \bar{x}}{s}$$

##### *Mendeteksi Data Outlier*

Berdasarkan hasil standarisasi data yang diperoleh, jika terdapat nilai data yang tidak berada diantara rentang  $\pm 2.5$  maka data tersebut merupakan data *outlier*. Berikut merupakan hasil pengamatan dari data *outlier* hasil standarisasi yang dapat di lihat pada Tabel 3 :

Tabel 3 Data *Outlier*

| USAHA INDUSTRI                                  | Variabel | Nilai |
|---|----------|-------|
| Industri Penggajian Kayu                        | $ZX_2$   | 2.99  |
| Industri Gula Merah                             | $ZX_1$   | 3.85  |
| Industri Penggilingan Padi dan Penyosohan Beras | $ZX_1$   | 3.64  |
|   | $ZX_2$   | 3.71  |
| Industri Minyak Atsiri                          | $ZX_2$   | 3.56  |

Pada Tabel 3 diatas telah diperoleh data *outlier*. Cara mengatasi data *outlier* yaitu dengan menghilangkan atau mempertahankan data *outlier* tersebut. Dalam penelitian ini langkah yang diambil untuk mengatasi data *outlier* yaitu dengan cara tetap mempertahankan data *outlier* karena data *outlier* tersebut merupakan representasi dari populasi yang diteliti.

##### *Asumsi-Asumsi dalam Analisis Cluster*

##### *Asumsi Kecukupan Sampel*

Hipotesis :

$H_0$  : Sampel belum memadai untuk dianalisis lebih lanjut

$H_1$  : Sampel telah memadai untuk dianalisis lebih lanjut

Adapun rumus yang dapat digunakan untuk menghitung nilai *Kaiser Mayer Olkin (KMO)*, yaitu :

Statistik Uji :

$$KMO = \frac{\sum_{i=1}^p \sum_{j=1}^p r_{ij}^2}{\sum_{i=1}^p \sum_{j=1}^p r_{ij}^2 + \sum_{i=1}^p \sum_{j=1}^p \rho_{ij}^2}$$

Kriteria Uji : jika nilai  $KMO > 0.5$ , maka tolak  $H_0$  atau sampel layak untuk dianalisis lebih lanjut.

Berikut merupakan perhitungan  $KMO$  menggunakan  $R$ -Programming.

Tabel 4 Nilai  $KMO$  dan  $Bartlett Test$

| Uji                          | Nilai |
|------------------------------|-------|
| Kaiser Mayer Olkin ( $KMO$ ) | 0.5   |

Berdasarkan Tabel 4., diperoleh nilai  $Kaiser Mayer Olkin (KMO)$  yaitu sebesar 0.5. Nilai tersebut tidak kurang dari 0.5 yang berarti sampel telah cukup dan layak untuk dianalisis

#### Asumsi Multikolinearitas

Dengan perhitungan menggunakan  $R$ -Programming diperoleh nilai-nilai korelasi antar variabel dapat dilihat pada tabel 5 berikut :

Tabel 5 Korelasi Antar Variabel

| Variabel | $X_1$ | $X_2$ |
|----------|-------|-------|
| $X_1$    | 1     | 0.73  |
| $X_2$    | 0.73  | 1     |

Berdasarkan Tabel 5 dapat dilihat bahwa variabel  $X_1$  dan  $X_2$  tidak mengalami multikolinearitas karena nilai korelasi yang dihasilkan tidak lebih dari 0.80.

#### Menentukan Ukuran Kemiripan Objek Menggunakan Jarak Euclidean

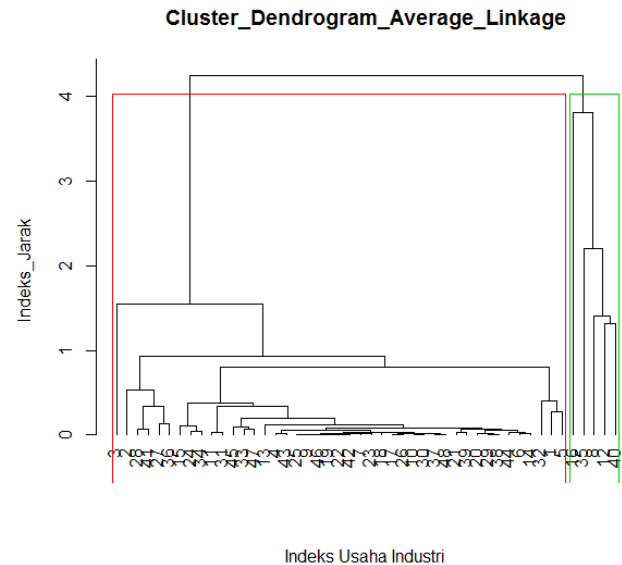
Dengan perhitungan menggunakan  $R$ -Programming didapat hasil perhitungan jarak tiap objek yang dapat dilihat pada tabel 6 berikut :

Tabel 6 Jarak  $Euclidean$  Usaha Industri

| USAHA INDUSTRI | 1    | 2    | 3    | ... | 22   |
|----------------|------|------|------|-----|------|
| 1              | 0.00 | 1.16 | 1.01 | ... | 0.76 |
| 2              | 1.16 | 0.00 | 0.97 | ... | 1.45 |
| 3              | 1.01 | 0.97 | 0.00 | ... | 1.78 |
| ⋮              | ⋮    | ⋮    | ⋮    | ⋮   | ⋮    |
| 22             | 0.76 | 1.45 | 1.78 | ... | 0.00 |

#### Proses Cluster dengan Metode Average Linkage Clustering

Berikut merupakan dendrogram yang terbentuk dari metode  $Average Linkage Clustering Hierarky$  yang dapat dilihat pada gambar 1 berikut :



Gambar 1 Dendrogram Metode  $Average Linkage Clustering Hierarky$

Dari Gambar 1 terlihat bahwa terdapat dua cluster yang terbentuk yaitu cluster A dan cluster B dari hasil  $R$ -Programming menggunakan metode  $Average Linkage Clustering Hierarky$ .

#### Menentukan Jumlah Cluster Beserta Anggotanya

Dalam menentukan jumlah cluster dapat dilihat dari hasil output dendrogram proses cluster dengan metode  $Average Linkage Clustering Hierarky$  yang telah diperoleh sebelumnya. Dari hasil yang diperoleh tersebut maka dapat ditentukan banyaknya cluster atau kelompok yang terbentuk adalah sebanyak 2 cluster, yaitu cluster A dan cluster B. Dimana, cluster A terdiri dari 43 industri dan cluster B terdiri dari 5 industri.

#### Interpretasi Cluster

Hasil cluster dapat diinterpretasikan melalui nilai-nilai  $centroid$  yang dihasilkan pada tiap cluster. Adapun nilai-nilai  $centroid$  yang dihasilkan tiap cluster dapat dilihat pada Tabel 7 berikut :

Tabel 7 *Centroid* Tiap *Cluster*

| Variabel | Cluster A   | Cluster B  |
|----------|-------------|------------|
| $X_1$    | 66.9        | 815.6      |
| $X_2$    | 481194767.4 | 8503509200 |

Berdasarkan Tabel 7, maka *cluster* dapat diinterpretasikan sebagai berikut :

1. Untuk *cluster* A variabel Jumlah Tenaga Kerja ( $X_1$ ) dan variabel Nilai Investasi ( $X_2$ ) mempunyai nilai *centroid* yang lebih kecil dari *cluster* B. Hal ini menunjukkan bahwa *Cluster* A merupakan *cluster* untuk **Industri Kecil**. Sedangkan,
2. Untuk *cluster* B, variabel Jumlah Tenaga Kerja ( $X_1$ ) dan variabel Nilai Investasi ( $X_2$ ) mempunyai nilai *centroid* yang lebih tinggi dari *cluster* A. Hal ini menunjukkan bahwa *Cluster* B merupakan *cluster* untuk **Industri Menengah**.

## 5. KESIMPULAN

kesimpulan yang diperoleh dari hasil penelitian klasifikasi usaha industri dengan indikator industri kecil dan industri menengah menggunakan metode *Average Linkage Clustering*, yaitu *cluster* A terdiri dari 43 usaha industri yang merupakan industri kecil dan *cluster* B terdiri dari 5 usaha industri yang merupakan industri menengah yang masing-masing dipengaruhi oleh Jumlah Tenaga Kerja ( $X_1$ ) dan Nilai Investasi ( $X_2$ ).

## 6. DAFTAR PUSTAKA

- [1] Badan Pusat Statistik (BPS) Luwu Utara. 2018. “Kabupaten Luwu Utara dalam Angka Tahun 2018”. Luwu Utara : Badan Pusat Statistik (BPS) Luwu Utara.
- [2] Wijaya, Tony. 2010. “Analisis Multivariat Teknik Olah Data untuk Skripsi, Tesis, dan Disertasi Menggunakan SPSS”. Yogyakarta : Universitas Atma Jaya Yogyakarta (UAJY), h. 3-4.
- [3] Gunawan, Imam. 2017. “Pengantar Statistika Inferensial”. Jakarta : PT. Rajagrafindo Persada, h. 326.
- [4] Gunawan, Imam. 2017. “Pengantar Statistika Inferensial”. Jakarta : PT. Rajagrafindo Persada, h. 335 & 342.
- [5] Handoyo, Rendy. dkk. 2014. “Perbandingan Metode *lustering* Menggunakan Metode *Single Linkage* dan *K-Means* pada Pengelompokan Dokumen”. Bandung : Fakultas Teknik Elektro Universitas Telkom, h. 75-76.
- [6] Nafisah, Qonitatin dan Novita Eka C. 2017. “Analisis *Cluster Average Linkage* Berdasarkan Faktor-Faktor Kemiskinan di Provinsi Jawa Timur”. Lamongan : Jurusan Matematika Fakultas MIPA Universitas Islam Darul Ulum Lamongan, h. 32.
- [7] Ratnasari, Andri dan Drs. H. Kirwani SE, MM. 2013. “Peranan *Industri Kecil Menengah (Ikm)* Dalam Penyerapan Tenaga Kerja Di Kabupaten Ponorogo. (Vol 1, No 3 (2013)), h. 5-6.