



# Eksplorasi Data *Craniometry* Manusia *Indigenous* dalam Upaya Pelacakan Pola Biogeografis pada *Homo sapiens* dengan Menggunakan Metode *Complete Linkage Clustering*

Ulfa Diana<sup>1</sup>, Triyana Muliawati<sup>2</sup>, Danni Gathot Harbowo<sup>3</sup>

<sup>1,2</sup>Program Studi Matematika, Institut Teknologi Sumatera

<sup>1</sup>[ulfa.120160101@student.itera.ac.id](mailto:ulfa.120160101@student.itera.ac.id)

<sup>2</sup>[triyana.muliawati@ma.itera.ac.id](mailto:triyana.muliawati@ma.itera.ac.id)

<sup>3</sup>Program Studi Teknik Geologi, Institut Teknologi Sumatera

<sup>3</sup>[danni.gathot@gl.itera.ac.id](mailto:danni.gathot@gl.itera.ac.id)

Corresponding autor email: [triyana.muliawati@ma.itera.ac.id](mailto:triyana.muliawati@ma.itera.ac.id)

**Abstract:** *Indigenous humans are indigenous people who come from various parts of the world and have diverse characteristics. The characteristics of indigenous humans can be seen from the diversity of skull shapes. So it is necessary to measure data on the dimensions of the skull or craniometry which is used to analyse biological diversity and evolution in humans, including indigenous humans in various regions. To analyse the characteristics of craniometry in each region, mathematics can be used, namely cluster analysis. One of the methods in cluster analysis is complete linkage clustering which is used to cluster the distribution of craniometry based on the measured part of the skull. In this study used data as many as 2524 samples with 82 measurement variables of skull parts and spread over 30 world populations. Based on the results of grouping 2, 4, 6, 8, 10, and 12 clusters, it is found that almost all specimens in all subregions are in each cluster, which means that the specimens have similar skull size characteristics. This shows that there are migration routes that cause humans to interact with each other so that the size and shape of the skull in each subregion tend to be more diverse. In addition, the variable JUB (Bijugal Breadth) or a measure of facial width was found to be a significant or main variable.*

**Keywords:** *cluster analysis, complete linkage, craniometry, indigenous human, centrality*

**Abstrak:** Manusia *indigenous* merupakan masyarakat adat yang berasal dari berbagai belahan dunia dan memiliki karakteristik yang beragam. Karakteristik dari manusia *indigenous* dapat dilihat dari keberagaman bentuk tengkorak. Sehingga diperlukan data pengukuran dimensi tengkorak atau *craniometry* yang digunakan untuk menganalisis tentang keragaman biologis dan evolusi pada manusia, termasuk manusia *indigenous* diberbagai wilayah. Untuk menganalisis karakteristik *craniometry* di setiap wilayah, dapat digunakan ilmu matematika yaitu analisis *cluster*. Salah satu metode pada analisis *cluster* adalah *complete linkage clustering* yang digunakan untuk melakukan pengelompokan terhadap persebaran dari *craniometry* berdasarkan pada bagian tengkorak yang diukur. Pada penelitian ini digunakan data sebanyak 2524 sampel dengan 82 variabel pengukuran bagian tengkorak dan tersebar pada 30 populasi dunia. Berdasarkan hasil pengelompokan 2, 4, 6, 8, 10, dan 12 *cluster* diperoleh hasil bahwa hampir seluruh spesimen di seluruh subregional berada pada setiap *cluster* yang artinya spesimen memiliki karakteristik ukuran tengkorak yang mirip. Hal ini menunjukkan bahwa terdapat jalur migrasi yang menyebabkan manusia saling berinteraksi sehingga ukuran dan bentuk tengkorak pada setiap subregional cenderung lebih beragam. Selain itu, diperoleh variabel JUB (*Bijugal Breadth*) atau ukuran dari lebar wajah yang menjadi variabel signifikan atau utama.

**Kata kunci:** *analisis cluster, complete linkage, craniometry, manusia indigenous, sentralitas*

## I. PENDAHULUAN

Manusia selalu melakukan revolusi untuk mengubah dan memperbaiki kehidupannya. Manusia pertama kali berevolusi di Afrika Timur kemudian bermigrasi menuju Afrika Utara, Eropa lalu Asia yang selanjutnya menetap di daerah-daerah tersebut [1]. Sejarah migrasi pada manusia dapat dilacak menggunakan catatan arkeologi, perubahan bahasa, budaya serta faktor biologis manusia seperti gen dan faktor fisik manusia itu sendiri. Peninggalan fosil tulang belulang dapat menunjukkan kehidupan manusia zaman dahulu. Bagian fosil-fosil tersebut diantaranya yaitu tengkorak, tulang badan dan tulang kaki. Salah satu bagian fosil yang berperan penting pada penelitian yang berhubungan dengan karakteristik dari makhluk hidup adalah tengkorak. Ukuran tempurung kepala pada tengkorak dapat menunjukkan sejauh mana kemampuan berpikir manusia zaman dulu dibandingkan dengan manusia

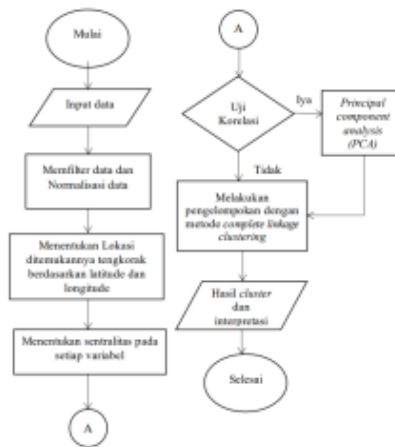


modern [2]. Bentuk tengkorak yang ditemukan dapat memberikan informasi terkait dengan asal usul nenek moyang termasuk juga karakteristik pada manusia *indigenous* atau masyarakat adat. Manusia *indigenous* atau masyarakat adat berasal dari berbagai belahan dunia dan memiliki karakteristik yang beragam. Keberagaman dari karakteristik tersebut bergantung pada wilayah geografis, sejarah genetik dan budaya mereka. Perbedaan manusia *indigenous* dapat dilihat pada keunikan bentuk dan ukuran dari tengkorak, yang digunakan untuk mendeskripsikan keragaman antar populasi. Perkembangan ilmu pengetahuan sangat bermanfaat di berbagai bidang, salah satunya pada bidang antropologi yang berguna untuk menjelaskan ciri khas dari suatu masyarakat. Salah satu ilmu yang berfokus di bidang ini yaitu *craniometry* atau pengukuran dimensi tulang tengkorak. Dimensi tulang tengkorak manusia dapat digunakan untuk menganalisis tentang keragaman biologis dan evolusi pada manusia termasuk manusia *indigenous*. Pengukuran pada tulang tengkorak ini mencakup pada pengukuran linier, sudut, volume, kedalaman dan juga busur tengkorak manusia. *Craniometry* digunakan sebagai metode pengukuran tulang tengkorak dapat menyediakan data yang dibutuhkan untuk melakukan analisis morfometri. Morfometri merupakan analisis pada bentuk dan ukuran dari suatu objek yang bertujuan untuk mengukur variasi morfologi serta untuk memahami perbedaan pada pola bentuk antar individu ataupun populasi tertentu, objek pada penelitian ini adalah makhluk hidup.

Keberagaman bentuk dan ukuran dari tengkorak dapat memberikan informasi terkait dengan perbedaan spesimen pada subregional atau wilayah tertentu. Perbedaan ukuran tengkorak pada spesimen yang berada di suatu subregional tersebut dapat dieksplor dengan menggunakan data *craniometry*. Namun, perlu dilakukan proses pengelompokan terlebih dahulu dengan tujuan untuk mengelompokkan objek ke dalam suatu *cluster*. Sehingga, objek yang berada pada *cluster* yang sama cenderung memiliki kemiripan dibandingkan dengan *cluster* lainnya. Analisis *cluster* dibagi menjadi 2 yaitu metode hierarki dan non-hierarki [3]. Pada metode hierarki jumlah *cluster* tidak ditentukan terlebih dahulu, sementara itu untuk metode non-hierarki perlu diasumsikan jumlah *cluster* yang diperoleh diawal. Fokus penelitian ini untuk mengetahui hasil pengelompokan dengan menggunakan metode hierarki yaitu metode *complete linkage* serta mengetahui karakteristik dari suatu spesimen dan menentukan bagian tengkorak yang menjadi pusat serta berperan penting pada data penelitian.

## II. METODELOGI PENELITIAN

Pada penelitian ini, proses penelitian dimulai dengan menginput data, memfiter data, menentukan lokasi ditemukannya tengkorak berdasarkan *latitude* dan *longitude*, menentukan sentralitas pada setiap variabel, uji korelasi, melakukan pengelompokan dengan metode *complete linkage clustering*, dan melakukan interpretasi hasil *cluster*.



Gambar 1. Alur Penelitian

## 2.1 Data

Data yang digunakan pada penelitian ini bersumber dari penelitian Dr. William Howells yang membahas terkait dengan pengukuran tengkorak manusia atau *Craniometry* [4] [5] [6]. Terdapat 86 variabel pada data dengan 82 variabel merupakan pengukuran bagian tengkorak dan 4 variabel lainnya yaitu *ID*, *Sex*, *Populasi Numerik*, dan *Populasi* dari lokasi ditemukannya tengkorak manusia.

## 2.2 Filter Data dan Normalisasi Data

*Filter* data dilakukan untuk menghilangkan/menghapus data kosong atau data yang tidak lengkap karena dapat mempengaruhi hasil yang akan di peroleh nantinya. Normalisasi merupakan teknik analisa data dengan cara mentransformasikan sekumpulan data menjadi bentuk yang normal. Perubahan pada normalisasi data tidak menyebabkan berubahnya informasi dari data asli. Normalisasi dilakukan untuk menyetarakan nilai (sangat besar atau sangat kecil) yang terdapat pada data [7]. Pada penelitian ini akan digunakan teknik normalisasi *Z-score* atau standardisasi dengan menggunakan rumus sebagai berikut:

$$Z = \frac{x_i - \bar{x}}{\sigma} \quad (1)$$

dengan  $Z$  merupakan nilai peubah baku,  $x_i$  merupakan data ke- $i$ ,  $\bar{x}$  merupakan nilai rata-rata keseluruhan data dan  $\sigma$  merupakan nilai standar deviasi [8].

## 2.3 Menentukan Lokasi Ditemukannya Tengkorak

Lokasi penemuan tengkorak manusia berdasarkan pada variabel *Population* pada data penelitian dengan menggunakan garis lintang dan garis bujur. Garis lintang (*Latitude*) merupakan garis yang menunjukkan lokasi di sebelah utara atau selatan ekuator dan diukur dari titik  $-90^0$  kutub utara hingga  $90^0$  kutub timur. Sedangkan garis bujur (*Longitude*) merupakan garis yang menunjukkan lokasi di sebelah barat atau timur garis meridian dan diukur dari titik  $0^0$  hingga  $180^0$ .

## 2.4 Menentukan Sentralitas pada Setiap Variabel

Sentralitas digunakan untuk mengidentifikasi suatu objek terpenting dalam suatu penelitian. Untuk menentukan sentralitas langkah-langkah yang harus dilakukan yaitu:

- Matriks Korelasi

Matriks korelasi merupakan matriks bujur sangkar atau matriks persegi dengan entri  $r_{ij}$ , dimana  $r_{ij}$  menyatakan koefisien korelasi antara baris ke- $i$  dan kolom ke- $j$ . Matriks korelasi memiliki rentang



nilai korelasi atau koefisien korelasi yaitu terletak pada interval  $[-1,1]$  [9]. Matriks korelasi data diperoleh dengan cara mencari nilai dari matriks kovarians terlebih dahulu.

$$s_{ij} = cov(x, y)$$

$$cov(x, y) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$$

$$r_{ij} = \frac{s_{ij}}{\sqrt{s_{ii}}\sqrt{s_{jj}}}$$

$$R = \begin{bmatrix} r_{11} & r_{12} & \dots & r_{1p} \\ r_{21} & r_{21} & \dots & r_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ r_{p1} & r_{p2} & \dots & r_{pp} \end{bmatrix}$$

dengan koefisien kovarians data antara variabel ke- $i$  dan variabel ke- $j$ .

- Matriks Jarak

Matriks jarak merupakan matriks bujur sangkar yang entri-entrinya menunjukkan jarak dari satu titik ke titik yang lain. Matriks jarak dinotasikan dengan  $D = [d_{ij}]$  yang didefinisikan sebagai berikut:

$$d_{ij} = \sqrt{2(1 - r_{ij})} \quad (2)$$

dengan  $d_{ij}$  merupakan koefisien jarak antara baris ke- $i$  dan kolom ke- $j$  [10].

- Matriks Ketetanggaan

Matriks ketetanggaan merupakan matriks bujur sangkar yang dinotasikan dengan  $B = [b_{ij}]$  dimana  $b_{ij}$  adalah koefisien matriks ketetanggaan yang akan bernilai bernilai 1 jika titik  $i$  dan  $j$  saling bertetangga dan akan bernilai 0 jika titik  $i$  dan  $j$  tidak saling bertetangga. Menurut [10] matriks ketetanggaan dapat dicari dengan menggunakan matriks jarak dengan cara sebagai berikut:

- Misalkan  $k = 2$
- Hitung  $D^k = D \cdot D^{(k-1)}$  menggunakan perkalian matriks biasa. Namun perkalian dua buah bilangan real  $a$  dan  $b$  dapat didefinisikan sebagai  $\max\{a, b\}$  dan penjumlahan didefinisikan sebagai  $\min\{a, b\}$ .
- Jika  $D^k = D^{(k-1)}$ , maka lanjut menghitung matriks ketetanggaan, namun jika tidak kembali ke langkah 2 dengan  $k = k + 1$ .
- Hitung  $D - D^k$  untuk menemukan matriks ketetanggaan.
- Hasil perhitungan  $D - D^k$  diubah menjadi matriks ketetanggaan yang memiliki entri 0 dan 1, dengan ketentuan yaitu ubah entri matriks menjadi 1 jika hasil perhitungan  $D - D^k = 0$  dan ubah entri matriks menjadi 0 jika hasil perhitungan  $D - D^k > 0$  serta nilai pada diagonal utama pada matriks ketetanggaan tidak diubah.

- Graf

Graf adalah pasangan terurut dari himpunan titik-titik dan himpunan sisi-sisi yang memiliki keterhubungan satu sama lain, dimana titik merupakan bagian dari graf yang merepresentasikan objek tertentu. Berdasarkan arahnya graf dapat dibedakan menjadi 2 yaitu graf berarah yang memiliki arah pada setiap sisi dan graf tak berarah yang tidak memiliki arah pada setiap sisi-sisinya [11].

- Sentralitas Derajat

Sentralitas pada graf merupakan salah satu konsep yang digunakan untuk mengidentifikasi titik terpenting pada suatu graf. Sentralitas juga digunakan untuk melihat pengaruh dari suatu titik pada graf. Terdapat beberapa jenis sentralitas yaitu, sentralitas derajat, sentralitas kedekatan dan



sentralitas keantaraan. Sentralitas derajat merupakan sentralitas yang mendeskripsikan jumlah interaksi yang dilakukan sebuah titik dalam graf. Semakin besar nilai sentralitas derajat yang dimiliki oleh suatu titik, maka akan semakin banyak titik yang dapat mempengaruhi atau dipengaruhi oleh titik tersebut. Sentralitas derajat didefinisikan sebagai berikut [12]:

$$C_D(i) = \frac{\sum_i^N b_{ij}}{N - 1} \quad (3)$$

dengan  $C_D(i)$  merupakan nilai sentralitas pada titik ke- $i$ ,  $b_{ij}$  merupakan koefisien ketetangaan pada baris ke- $i$  dan kolom ke- $j$  dan  $N$  adalah banyaknya titik pada graf.

## 2.5 Uji Korelasi

Uji korelasi merupakan prosedur yang digunakan untuk mengecek ada tidaknya hubungan antar variabel penelitian.

### - Uji Kecukupan Data

Uji kecukupan data digunakan untuk mengetahui data dari sampel yang diambil sudah cukup untuk dilakukan uji selanjutnya atau belum. Uji kecukupan data dapat dianalisis dengan menggunakan uji *Kaiser Mayer Olkin* (KMO). Uji KMO digunakan untuk mengukur kecukupan dari seluruh data dan elemen populasi untuk setiap indikator. Uji KMO memiliki kisaran nilai antara 0 sampai dengan 1 [13].

Hipotesis:

$H_0$  : data telah memadai untuk dianalisis lebih lanjut ( $KMO \geq 0.5$ )

$H_1$  : data tidak memadai untuk dianalisis lebih lanjut ( $KMO < 0.5$ )

Statistik uji:

$$KMO = \frac{\sum_{i=1}^p \sum_{j=1}^p r_{ij}^2}{\sum_{i=1}^p \sum_{j=1}^p r_{ij}^2 + \sum_{i=1}^p \sum_{j=1}^p a_{ij}^2}, \text{ untuk } i \neq j \quad (3)$$

Selain itu, dilakukan juga uji *Maesure of Sampling Adequacy* (MSA) untuk mengukur kecukupan sampel dari setiap variabel. Hipotesis yang digunakan sama seperti hipotesis pada uji KMO dengan statistik uji yaitu sebagai berikut:

$$MSA_i = \frac{\sum_{i=1}^p \sum_{j=1}^p r_{ij}^2}{\sum_{i=1}^p \sum_{j=1}^p r_{ij}^2 + \sum_{i=1}^p \sum_{j=1}^p a_{ij}^2}, \text{ untuk } i \neq j \quad (4)$$

Jika nilai MSA di bawah 0.5, maka variabel tersebut tidak layak untuk dianalisis lebih lanjut sehingga harus dieliminasi [14]. Dengan  $r_{ij}$  merupakan koefisien korelasi antara variabel  $i$  dan  $j$  serta  $a_{ij}$  merupakan koefisien korelasi parsial/kovarians antara variabel  $i$  dan  $j$ .

### - Uji Bartlett

Uji *Bartlett* merupakan uji yang dilakukan untuk menguji homogenitas varians dari data yang digunakan. Uji *Bartlett* digunakan juga untuk mengecek ada tidaknya hubungan yang signifikan antar variabel penelitian. Langkah yang dilakukan untuk melakukan uji *Bartlett* yaitu sebagai berikut [15]:

Hipotesis:

$H_0$  :  $R = I$  (tidak ada hubungan/korelasi antar variabel)

$H_1$  :  $R \neq I$  (ada hubungan/korelasi antar variabel)

Statistika uji:

$$\chi_{hitung}^2 = - \left( n - 1 - \frac{2p + 5}{6} \right) \ln |R|$$
$$\chi_{tabel}^2 = \chi_{\alpha/2, 1(p(p-1))}^2 \quad (5)$$



$$I = \begin{bmatrix} 1 & 0 & 0 & & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}$$

dengan,  $I$  merupakan matriks identitas,  $n$  banyaknya data,  $p$  banyaknya variabel dan  $|R|$  merupakan determinan matriks korelasi.  $H_0$  akan ditolak jika  $\chi^2_{hitung} > \chi^2_{tabel}$  atau nilai dari  $p - value < \alpha = 0.05$  yang artinya terdapat hubungan/korelasi antar variabel.

### 2.6 Principal Component Analysis (PCA)

PCA atau analisis komponen utama merupakan analisis yang dilakukan dengan tujuan untuk mereduksi dimensi dari suatu data yang memiliki korelasi sehingga menghasilkan data yang tidak berkorelasi (saling bebas) tanpa mengurangi karakteristik dari data tersebut. Variabel baru yang dihasilkan tersebut merupakan komponen utama dari data [16]. Langkah-langkah dalam melakukan analisis komponen utama diantaranya yaitu:

- Nilai Eigen

Nilai Eigen diperoleh dari matriks korelasi. Nilai eigen ini akan digunakan untuk menganalisis banyaknya komponen utama yang terdapat pada data. Persamaan yang digunakan untuk mengitung nilai eigen dari matriks korelasi adalah sebagai berikut:

$$\det(R - \lambda I) = 0 \tag{6}$$

dengan  $R$  merupakan matriks korelasi dan  $\lambda$  merupakan nilai eigen.

- Menentukan Banyaknya Komponen Utama

Untuk menentukan banyaknya komponen utama dapat dilakukan berdasarkan pada nilai eigen yang diperoleh dari matriks korelasi dengan batasan nilai eigen lebih besar dari 1 yang akan digunakan sebagai patokan banyaknya komponen utama.

- Vektor Eigen

Vektor eigen dari matriks korelasi dihitung dengan menggunakan persamaan berikut:

$$R\mathbf{v} = \lambda\mathbf{v} \tag{7}$$

dengan  $R$  adalah matriks korelasi,  $\lambda$  adalah nilai eigen dan  $\mathbf{v}$  adalah vektor eigen.

- Persamaan Komponen Utama

Vektor eigen digunakan sebagai koefisien dari masing-masing variabel yang pada proses pembentukan komponen utama, sehingga dapat diperoleh persamaan dari komponen utama sebagai berikut:

$$PC_i = v_{1p}X_1 + v_{2p}X_2 + v_{3p}X_3 + \dots + v_{ip}X_p \tag{8}$$

dengan  $i$  merupakan banyaknya komponen utama,  $p$  banyaknya komponen asli dan  $PC_i$  merupakan komponen utama ke- $i$ .

### 2.7 Pegelompokan dengan Metode Complete Linkage Clustering

Complete linkage merupakan salah satu metode pada analisis cluster dengan menggunakan metode hierarki. Metode hierarki merupakan suatu pendekatan analisis yang digunakan untuk mengelompokkan objek menjadi sebuah cluster berdasarkan pada tingkat kemiripannya sampai terbentuk sebuah pohon yang terdapat tingkatan dari yang paling mirip sampai dengan yang paling tidak mirip, pohon ini biasa disebut dengan dendogram [17]. Metode pengelompokan dengan menggunakan hierarki yang paling sering digunakan adalah metode Agglomerative. Metode Agglomerative (Penggabungan) merupakan metode yang menggunakan asumsi bahwa setiap objek yang digunakan adalah sebuah cluster. Objek yang memiliki jarak terdekat digabungkan menjadi satu

*cluster*, objek lainnya akan digabungkan ke dalam *cluster* yang lain sehingga membentuk sebuah *cluster* baru [18]. *Complete Linkage* (Pautan Lengkap) merupakan pengelompokan yang dilakukan berdasarkan objek yang memiliki jarak terjauh. Langkah yang harus dilakukan yang pertama adalah mencari objek yang memiliki jarak terdekat dan digabungkan sehingga didapatkan sebuah *cluster xy*. Kemudian, pada langkah selanjutnya dicari jarak maksimum atau jarak terjauh dari objek *xz* dan *yz* terhadap *xy* sehingga dapat diperoleh hasilnya dengan menggunakan rumus berikut [19]:

$$d_{(xy)z} = \max \{d_{xz}, d_{yz}\}$$

dengan  $d_{(xy)z}$  merupakan jarak maksimum antara objek *xy* dengan objek *z*,  $d_{xz}$  merupakan jarak antara objek *x* dengan objek *z* dan  $d_{yz}$  merupakan jarak antara objek *y* dengan objek *z*.

### III. HASIL DAN PEMBAHASAN

#### 3.1 Menentukan Lokasi Ditemukannya Tengkorak Manusia

Penentuan lokasi ditemukannya tengkorak manusia perlu dicari terlebih dahulu titik koordinat pada permukaan bumi yaitu dengan mencari garis lintang (*Latitude*) dan garis bujur (*Longitude*). Titik koordinat dicari berdasarkan lokasi penemuan tengkorak yang terdapat pada *The William W. Howells Craniometry dataset* yang berada pada kolom “*population*”. *Latitude* dan *Longitude* dari penemuan *craniometry* seluruh spesimen akan diplotkan ke dalam peta dunia seperti pada Gambar 2 dibawah ini.



**Gambar 2.** Plot *Latitude* dan *Longitude* pada peta dunia

Berdasarkan plot *Latitude* dan *Longitude* pada peta diatas, lokasi ditemukannya *craniometry* pada seluruh spesimen ditunjukkan oleh titik berwarna merah serta spesimen banyak tersebar di benua Eropa, Asia, Afrika, Australia, Amerika dan Oseania.

#### 3.2 Menentukan Sentralitas dari Seluruh Variabel

Untuk menentukan sentralitas dapat menggunakan matriks korelasi yang diperoleh dari hasil matriks kovarians. Matriks korelasi dihitung berdasarkan pada data *craniometry* sebanyak 82 pengukuran bagian tengkorak dan diperoleh hasil yang ditunjukkan pada Tabel 1.



**Tabel 1.** Matriks korelasi

Korelasi	GOL	NOL	BNL	BBH	XCB	...	TBA
<b>GOL</b>	1.000	0.984	0.718	0.518	0.280	...	-0.097
<b>NOL</b>	0.984	1.000	0.720	0.516	0.307	...	-0.098
<b>BNL</b>	0.718	0.720	1.000	0.724	0.263	...	-0.321
<b>BBH</b>	0.518	0.516	0.724	1.000	0.376	...	-0.481
<b>XCB</b>	0.280	0.307	0.263	0.376	1.000	...	0.173
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
<b>TBA</b>	-0.097	-0.098	-0.321	-0.481	0.173	...	1.000

Setelah diperoleh matriks korelasi, langkah selanjutnya yaitu menghitung jarak setiap variabel dan diperoleh hasil matriks jarak yang ditunjukkan pada Tabel 2.

**Tabel 2.** Matriks jarak

Jarak	GOL	NOL	BNL	BBH	XCB	...	TBA
<b>GOL</b>	0.000	0.181	0.751	0.982	1.200	...	1.481
<b>NOL</b>	0.181	0.000	0.749	0.984	1.177	...	1.482
<b>BNL</b>	0.751	0.749	0.000	0.743	1.214	...	1.625
<b>BBH</b>	0.982	0.984	0.743	0.000	1.117	...	1.721
<b>XCB</b>	1.200	1.177	1.214	1.117	0.000	...	1.286
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
<b>TBA</b>	1.481	1.482	1.625	1.721	1.286	...	0.000

Setelah diperoleh matriks jarak, tahapan selanjutnya yaitu menghitung matriks ketetanggaan dan diperoleh hasil seperti yang ditunjukkan pada Tabel 3.

**Tabel 3** Matriks ketetanggaan

Ketetanggaan	GOL	NOL	BNL	BBH	XCB	...	TBA
GOL	0	1	0	0	0	...	0
NOL	1	0	0	0	0	...	0
BNL	0	0	0	1	0	...	0
BBH	0	0	1	0	0	...	0
XCB	0	0	0	0	0	...	0
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
TBA	0	0	0	0	0	...	0

Berdasarkan matriks ketetanggaan yang sudah diperoleh, dapat dilihat bahwa nilai entri dari matriks ketetanggaan adalah 1 dan 0. Nilai entri 1 merepresentasikan bahwa variabel saling bertetangga atau memiliki hubungan ketetanggaan dan entri 0 merepresentasikan bahwa variabel tidak saling bertetangga atau tidak memiliki hubungan ketetanggaan. Selanjutnya yaitu membentuk graf tidak berarah dari matriks ketetanggaan yang sudah diperoleh seperti yang ditunjukkan pada Gambar 3.



**Gambar 3.** Graf korelasi berdasarkan matriks ketetanggaan





Selanjutnya, akan ditentukan pusat dari graf korelasi dengan menggunakan ukuran sentralitas derajat. Berikut merupakan hasil perhitungan sentralitas derajat yang ditunjukkan pada Tabel 4.

**Tabel 4.** Sentralitas derajat

Variabel	Sentralitas Derajat	Variabel	Sentralitas Derajat	Variabel	Sentralitas Derajat
GOL	0.0494	EKB	0.0494	AVR	0.0123
NOL	0.0247	DKS	0.0123	BRR	0.0370
⋮	⋮	⋮	⋮	⋮	⋮
JUB	0.0741	PAS	0.0123	NFA	0.0494
⋮	⋮	⋮	⋮	⋮	⋮
SSS	0.0123	EKR	0.0247	SLA	0.0247
FMB	0.0370	ZMR	0.0247	TBA	0.0247

Berdasarkan hasil perhitungan sentralitas derajat yang ditampilkan pada Tabel 4, diperoleh bahwa variabel yang memiliki ukuran paling besar adalah variabel *Bijugal breadth* (JUB), maka variabel tersebut menjadi pusat dari graf korelasi pada Gambar 3. Sehingga variabel JUB memiliki peranan yang sangat peting dalam mempengaruhi atau dipengaruhi oleh variabel lain.

### 3.3 Uji Korelasi

#### Uji Kecukupan Data

Uji kecukupan data dilakukan dengan menggunakan uji *Kaiser-Mayer Olkin* (KMO) dan *Measure of Sampling Adequacy* (MSA) yang berfungsi untuk mengetahui kelayakan dari variabel-variabel yang diteliti.

**Tabel 5.** Hasil Uji KMO

Uji KMO	
Nilai KMO	0.875

Berdasarkan hasil KMO diatas diperoleh bahwa nilai KMO sudah lebih dari 0.5 sehingga memenuhi uji kecukupan data. Selain itu diperoleh juga hasil MSA bahwa terdapat 3 variabel yang nilai MSA nya kurang dari 0.5, sehingga ketiga variabel tersebut dihapuskan dan tidak dilanjutkan ke tahap selanjutnya. Ketiga variabel tersebut yaitu WNB, ROA, dan BSA, sehingga hanya 79 variabel yang akan dilakukan uji selanjutnya.

#### Uji Barlett

Uji *barlett* digunakan untuk melihat apakah data memiliki korelasi antar variabel atau tidak.

**Tabel 6.** Uji barlett

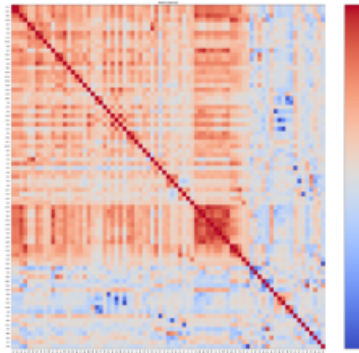
Nilai Uji	
$\chi^2_{hitung}$	309776.02
<i>p - value</i>	0.00

Berdasarkan Tabel 6 menunjukkan bahwa nilai *p - value* ( $0.00 < \alpha(0.05)$ ), yang artinya bahwa terdapat korelasi antar variabel, sehingga perlu dilakukannya proses PCA atau *Principal Component Analysis* terlebih dahulu sebelum melakukan proses pengelompokan.

### 3.4 *Principal Component Analysis* (PCA)

#### Matriks Korelasi

Matriks korelasi diperoleh dari matriks kovarian, sehingga langkah awal yang harus dilakukan untuk mencari matriks korelasi adalah menentukan matriks kovarian terlebih dahulu, sehingga diperoleh hasil yang divisualisasikan seperti Gambar 4 berikut.



**Gambar 4.** Matriks korelasi

Semakin merah warna yang ditampilkan artinya variabel semakin memiliki hubungan positif yang kuat, begitupun sebaliknya semakin biru warna yang ditampilkan maka variabel memiliki hubungan negatif yang kuat.

### Menghitung Nilai Eigen

Tujuan menghitung nilai eigen ini adalah untuk menjelaskan keseragaman data pada setiap komponen.

**Tabel 7.** Nilai Eigen

Komponen	Nilai Eigen	<i>Proportion of Variance</i>	<i>Cumulative Proportion</i>
1	22.475	0.284	0.284
2	7.356	0.093	0.377
3	5.589	0.071	0.448
4	4.625	0.058	0.507
⋮	⋮	⋮	⋮
79	0.001	$1.64 \times 10^{-5}$	1

### Menentukan Banyaknya Komponen Utama

Berdasarkan hasil nilai eigen yang dihasilkan, diperoleh bahwa terdapat 16 komponen yang memiliki nilai eigen lebih besar dari 1 yang artinya data tersebut dapat direduksi menjadi 16 komponen utama dengan nilai kumulatifnya yaitu 83.5%.

### Menghitung Vektor Eigen

Vektor eigen yang dihitung merupakan vektor eigen yang bersesuaian dengan nilai eigen yang sudah diperoleh sebelumnya. Vektor eigen yang diperoleh akan digunakan sebagai koefisien untuk persamaan komponen utama.

### Membentuk Persamaan Komponen Utama

Setelah diperoleh hasil dari vektor eigen, maka dapat dibentuk persamaan sebanyak jumlah dari komponen utama yang sudah ditentukan sebelumnya yaitu 16 persamaan komponen utama dengan vektor eigen sebagai koefisien dan hasil normalisasi data sebagai variabel pada persamaan tersebut.

**Tabel 8.** Hasil PCA

PC1	PC2	PC3	PC4	⋯	PC16
1,403	-1,031	-5,062	-3,201	⋯	0,503
1,045	-0,797	-0,798	-1,787	⋯	0,192
4,033	-0,722	-2,637	-2,609	⋯	-0,168

PC1	PC2	PC3	PC4	...	PC16
-0,728	-3,216	-1,355	-4,464	...	-0,885
⋮	⋮	⋮	⋮	⋮	⋮
-13,032	1,656	0,830	3,013	...	2,708

Data baru yang sudah diperoleh ini akan digunakan untuk proses selanjutnya yaitu analisis *cluster*.

### 3.5 Pengelompokan dengan menggunakan Metode *Complete Linkage*

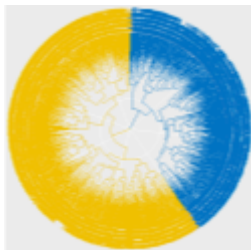
#### Menentukan Ukuran Kedekatan Objek

Untuk menentukan ukuran kedekatan objek data akan digunakan *euclidean distance* atau menghitung jarak setiap objek berdasarkan pada jarak minimum. Sehingga diperoleh hasil sebagai berikut.

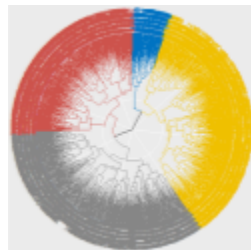
**Tabel 9.** *Euclidean distance*

ID	111	112	113	114	115	...	3190
111	0					...	
112	6.20	0				...	
113	6.47	6.64	0			...	
114	9.77	9.47	9.25	0		...	
115	7.31	6.22	7.92	9.73	0	...	
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
3190	18.37	16.75	19.66	17.23	16.84	...	0

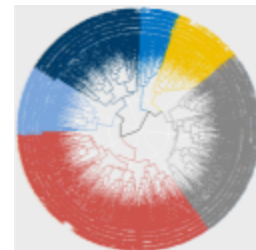
#### Pengelompokan 2 hingga 12 *cluster*



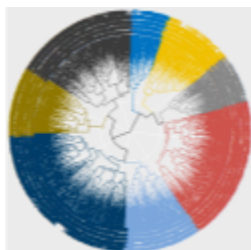
**Gambar 5.** Dendrogram pengelompokan 2 *cluster*



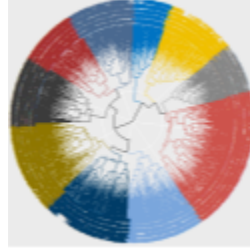
**Gambar 6.** Dendrogram pengelompokan 4 *cluster*



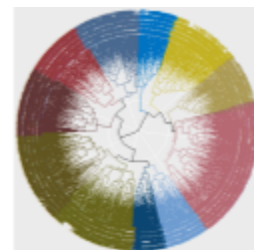
**Gambar 7.** Dendrogram pengelompokan 6 *cluster*



**Gambar 8.** Dendrogram pengelompokan 8 *cluster*



**Gambar 9.** Dendrogram pengelompokan 10 *cluster*



**Gambar 10.** Dendrogram pengelompokan 12 *cluster*

Berdasarkan hasil pengelompokan sampai dengan 12 *cluster* diperoleh bahwa setiap subregional berada pada masing-masing *cluster*. Hal ini menunjukkan bahwa karakteristik tengkorak manusia pada masing-masing subregional memiliki kemiripan, artinya bentuk dan ukuran tengkorak manusia di 12 subregional dunia memiliki keberagaman dan saling mirip satu sama lainnya. Berdasarkan hasil tersebut dapat diambil kesimpulan bahwa keberagaman dari bentuk tengkorak ini menunjukkan



bahwa terdapat jalur migrasi sehingga manusia saling berinteraks. Hal ini dapat menyebabkan keberagaman dari bentuk dan ukuran tengkorak serta dapat menghilangkan ciri khas khusus bentuk tengkorak pada setiap daerah.

Selain itu, dapat dilihat juga ukuran tengkorak pada setiap subregional dengan memanfaatkan variabel penelitian yaitu 79 pengukuran bagian tengkorak dengan memperhatikan variabel yang akan dianalisis merupakan variabel pengukuran panjang, lebar, dan tinggi dari bagian tengkorak. Untuk variabel pengukuran sudut dapat diabaikan saja karena hanya ingin melihat ukuran besar kecilnya tengkorak manusia. Dengan menggunakan variabel-variabel tersebut diperoleh bahwa subregional Asia Selatan memiliki ukuran tengkorak yang cenderung lebih kecil dibandingkan dengan subregional lainnya. Sementara itu, subregional lainnya yaitu subregional yang berada pada benua Eropa, Asia, Amerika, Afrika, dan Oseania memiliki ukuran tengkorak yang cenderung mirip dan beragam.

#### IV. KESIMPULAN

Berdasarkan hasil pengelompokan 79 pengukuran bagian tengkorak menjadi 2 *cluster* hingga 12 *cluster* diperoleh hasil bahwa terdapat pola dari ukuran tengkorak manusia pada lokasi geografis tertentu. Subregional Asia Selatan cenderung memiliki ukuran tengkorak yang lebih kecil, sedangkan subregional lainnya memiliki ukuran tengkorak yang cenderung mirip. Seluruh spesimen terbagi hampir pada setiap subregional dunia yang artinya spesimen-spesimen pada setiap *cluster* memiliki karakteristik ukuran tengkorak yang mirip. Hal ini menunjukkan bahwa terdapat jalur migrasi yang menyebabkan manusia saling berinteraksi sehingga ukuran dan bentuk tengkorak pada setiap subregional cenderung lebih beragam.

Berdasarkan hasil perhitungan sentralitas derajat menunjukkan bahwa *Bijugal Breadth* (JUB) memiliki nilai sentralitas derajat yang paling besar. Variabel JUB merupakan bagian pengukuran lebar wajah manusia yang secara horizontal dari pipi. Dikarenakan JUB merupakan variabel yang memiliki nilai sentralitas derajat yang paling besar, maka variabel JUB memiliki pengaruh yang paling besar pada hubungan antar spesimen. Sehingga dapat dikatakan bahwa variabel JUB dapat diidentifikasi sebagai variabel utama atau variabel paling signifikan pada pengelompokan antar spesimen. Jika variabel JUB mengalami masalah maka hasil analisis data menjadi kurang valid.

#### UCAPAN TERIMA KASIH

Penulis sangat berterimakasih kepada semua pihak yang sudah membantu semua proses penelitian ini sehingga dapat selesai dengan baik. Semoga penelitian ini dapat bermanfaat untuk semua yang membacanya.

#### REFERENSI

1. Yuval Noah Harari, *Sapiens; Riwayat Singkat Umat Manusia.*: Kepustakaan Polpuler Gramedia, 2017.
2. Mariana, *Sejarah Indonesia X; Kehidupan Manusia Purba dan Asal Usul Nenek Moyang.* Bekasi: Kementerian Pendidikan dan Kebudayaan, 2020
3. Nur Mutmainnah Djafar, "Pengelompokan Produksi Perkebunan Menurut Kabupaten/Kota Jawa Tengah Tahun 2020 Menggunakan Hierarchical Clustering," *Journal of Mathematics Education and Science*, vol. Vol.4, no. <https://doi.org/10.32665/james.v4i2.230>, pp. 59-66, oktober 2021.
4. William W. Howells, *Cranial Variation in Man: A Study by Multivariate Analysis of Patterns of Differences Among Recent Human Populations.* Cambridge, Mass., Amerika Serikat: Peabody Museum, 1973.
5. William W. Howells, *Skull Shapes and the Map: Craniometric Analyses in the Dispersion of Modern Homo.* Cambridge, Mass., Amerika Serikat: Peabody Museum, 1989.



6. William W. Howells, *Who's Who in Skulls: Ethnic Identification of Crania from Measurements*. Cambridge, Mass., Amerika Serikat: Peabody Museum, 1995.
7. Dwi Febby Haryati, "Klasifikasi Jenis Batu Bara Menggunakan Jaringan Syaraf Tiruan dengan Algoritma Backpropagation," *Seminar Nasional Teknologi Informasi dan Komunikasi 2016*, pp. 557-562, 2016.
8. Irwan, *Statistika Multivariat*. Sulawesi Selatan: Alauddin University, 2021.
9. Fanklin A. Graybill, *Matrices with Applications in Statistics*, Second Edition, Ed. Taipei, Taiwan: Wadsworth Publishing company, 1983.
10. Maman Abdurachman Djauhari, "Minimal Spanning Tree Problem in Stock Networks Analalysus : An Efficient Algorithm," *Physica A*, pp. 2227-2229, 2013.
11. P.S. Ria Rahadi Nasution, "Aplikasi Pewarnaan Graf pada Penyusunan Jadwal Mata Kuliah Jurusan Matematika di FMIPA Universitas Negeri Medan," *KARISMATIKA*, pp. 11-13, 2020.
12. Rian Febrian Umbara, dan Indwiarti Dika Rizky Nurcholis, "Analisis Hasil Tracer Study Terhadap Alumni Universitas Telkom dengan menggunakan Minimum Spanning Tree(MST)," *e-Proceeding of Engineering* , pp. 8093-8104, 2018.
13. Sigit Nugroho, *Statistika Multivariat Terapan*. Bengkulu: UNIB Press, 2008.
14. A. Widarjono, *Analisis Statistika Multivariat Terapan*. Yogyakarta: Unit Penerbit dan Percetakan Sekolah Tinggi Ilmu Manajemen YKPN, 2010.
15. A.Z. Wattimena, S.D. Saputri M.S. Nayo Van Delsen, "Penggunaan Metode Analisis Komponen Utama untuk Mereduksi Faktor-Faktor Inflasi di Kota Ambon," *Ilmu Matematika dan Terapan*, pp. 109-118, 2017.
16. Mariana, "Analisis Komponen Utama," *Jurnal Matematika dan Pembelajarannya*, pp. 99-114, 2013.
17. Anang Kurnia, Ristiyanti Cici Suhaeni, "Perbandingan Hasil Pengelompokan menggunakan Analisis Cluster Berhirarki, K-Means Cluster, dan Cluster Ensemble (Studi Kasus Data Indikator Pelayanan Kesehatan Ibu Hamil)," *Jurnal Media Infotama*, pp. 31-38, 2018.
18. Suhirman & Hero Wintolo, "System for Determining Public Health Level Using The Agglomerative Hierarchical Clustering Method," pp. 95-104, 2019.
19. Sukarna, Aswi Muhammad Arif Tiro, *Statistika Deskriptif Peubah Banyak*. Makassar: Andira Publisher, 2010.