



Implementasi *Average Linkage Clustering* pada Identifikasi Karakteristik Data Osteometri Manusia *Indigenous*

Toni Adam Purba¹, Triyana Muliawati², Danni Gathot Harbowo³

^{1,2}Program Studi Matematika, Institut Teknologi Sumatera

¹toni.120160051@student.itera.ac.id

²triyana.muliawati@ma.itera.ac.id

³Program Studi Teknik Geologi, Institut Teknologi Sumatera

³danni.gathot@gl.itera.ac.id

Corresponding author email: triyana.muliawati@ma.itera.ac.id

Abstract: Cluster analysis is a data grouping technique found in mathematics. The purpose of cluster analysis is to group data according to the characteristics of each data based on the distance between the data. This research uses a data clustering technique in the form of a hierarchical method, namely Average Linkage. The method is used to conduct cluster analysis of indigenous human bone size data (specimens) based on 43 long bone measurements and then determine the characteristics formed after the clustering process is carried out. Based on the results of the cluster analysis, specimens found in Africa, America, Europe and Australia have a size pattern of 43 types of long bone measurements that are longer than other subregions. In addition, specimens found in the Americas, Africa, Europe and Australia tend to have similar characteristics. Meanwhile, specimens found in the Asian continent tend to have similarities with specimens found in the Melanesian, Polynesian and South African subregions. The significant variable in the grouping process is the RHAPD variable which is located on the upper arm. These results indicate that humans originally came from Africa and were nomadic creatures that eventually created the current genetic, cultural and environmental diversity.

Keywords: Average Linkage, characteristics, cluster analysis, indigenous human, specimen

Abstrak: Analisis cluster merupakan teknik pengelompokan data yang terdapat dalam keilmuan matematika. Tujuan dari analisis cluster adalah mengelompokkan data sesuai karakteristik tiap datanya berdasarkan jarak antar datanya. Penelitian ini menggunakan teknik pengelompokan data berupa metode hierarki yaitu Average Linkage. Metode tersebut digunakan untuk melakukan analisis cluster terhadap data (spesimen) ukuran tulang manusia *indigenous* berdasarkan 43 pengukuran tulang panjang yang kemudian mengetahui karakteristik yang terbentuk setelah proses pengelompokan dilakukan. Berdasarkan hasil analisis cluster, spesimen-spesimen yang ditemukan di Benua Afrika, Amerika, Eropa dan Australia memiliki pola ukuran 43 jenis pengukuran tulang panjang yang lebih panjang dari subregion lain. Selain itu, spesimen-spesimen yang ditemukan di Benua Amerika, Afrika, Eropa, dan Australia cenderung memiliki karakteristik yang mirip. Sementara spesimen-spesimen yang ditemukan di Benua Asia cenderung memiliki kemiripan dengan spesimen-spesimen yang ditemukan di subregion Melanesia, Polinesia, dan Afrika Selatan. Adapun variabel yang signifikan dalam proses pengelompokan tersebut adalah variabel RHAPD yang letaknya berada di lengan atas. Hasil ini menunjukkan terdapat indikasi bahwa manusia pada awalnya berasal dari Afrika dan merupakan makhluk yang nomaden yang akhirnya membuat keragaman genetik, budaya, dan lingkungan pada saat ini.

Kata kunci: analisis cluster, Average Linkage, karakteristik, manusia *indigenous*, spesimen

I. PENDAHULUAN

Secara etimologis, manusia adalah makhluk yang memiliki akal dan budi, serta memiliki kemampuan untuk mengendalikan makhluk lainnya. Manusia juga merupakan makhluk sosial yang dapat berkomunikasi dan berinteraksi satu sama lain, serta membentuk kebiasaan yang menghasilkan budaya tertentu [1]. Kebudayaan itu menghasilkan keberagaman yang mengubah manusia yang awalnya sedikit dalam hal kualitas, menjadi beragam karena perkembangan kebudayaan. Dari segi biologis, manusia adalah makhluk dengan variasi dalam struktur dan fungsi tubuh, yang mencerminkan adaptasi mereka terhadap lingkungan dan berbagai faktor lainnya.

Berdasarkan sejarah, manusia telah melalui proses evolusi untuk mencapai bentuknya yang sekarang. Manusia dalam genus *Homo* berevolusi hingga menjadi spesies *Homo sapiens*. Evolusi ini mencakup perubahan dalam struktur kerangka, kebudayaan, lingkungan, serta variasi gen atau DNA.



Revolusi tersebut terjadi sekitar 20.000 tahun yang lalu, dikenal sebagai masa Holosen. Banyak ahli antropologi dan paleontologi percaya bahwa fenomena sejarah ini menunjukkan bahwa manusia modern awal (pramodern) berasal dari benua Afrika berdasarkan studi filogenetik [2]. Ada juga penelitian lain yang membuktikan rute-rute potensial yang ditempuh oleh manusia *indigenous* saat itu dari benua Afrika ke benua-benua lainnya [3]. Namun, ada teori alternatif yang dikenal sebagai Multiregional Evolution Model (MRE), yang menyatakan bahwa manusia *indigenous* (pramodern) tidak hanya berasal dari benua Afrika, tetapi juga dari benua Eropa dan Asia. Oleh karena itu, fenomena perdebatan ini menjadi menarik untuk diteliti berdasarkan sudut pandang struktur kerangka tulang manusia pramodern dan subregion tempat ditemukannya struktur kerangka tersebut yang dikombinasikan dengan teknik bermatematika khususnya analisis *cluster*. Nantinya, data osteometri pada penelitian ini akan dikelompokkan menjadi 2 *cluster*, 4 *cluster*, 6 *cluster*, 8 *cluster*, 10 *cluster*, 14 *cluster*, dan 16 *cluster* sesuai dengan banyaknya subregion tempat ditemukannya tulang manusia *indigenous* tersebut pada *dataset* yaitu sebanyak 16 subregion. Keenam belas subregion tersebut mencakup Amerika Utara, Amerika Selatan, Afrika Timur, Afrika Tengah, Afrika Selatan, Afrika Utara, Eropa Barat, Eropa Timur, Eropa Selatan, Eropa Utara, Asia Selatan, Asia Tenggara, Asia Timur, Melanesia, Polinesia, dan Australia. Adapun penelitian ini memiliki urgensi tingkat tinggi terutama pada poin 16 yaitu “Perdamaian, Keadilan dan Kelembagaan yang Tangguh”. Penelitian ini diharapkan mampu menghargai dan menghormati pengetahuan dan kebudayaan masyarakat adat (manusia *indigenous*) sehingga mendorong inklusi dan keadilan sosial sebagai dasar institusi yang kuat dan damai.

II. METODE PENELITIAN

2.1. Data dan Sumber Data

Data yang digunakan dalam penelitian ini terdiri dari 43 jenis pengukuran tulang panjang manusia *indigenous*, mencakup ukuran tulang lengan atas, tulang lengan bawah, tulang pinggul, tulang paha, dan tulang kering. Data sekunder ini berasal dari penelitian yang dilakukan oleh Dr. Benjamin M. Auerbach, Ph.D., yang disebut The Goldman Dataset [4].

2.2. Inisiasi Variabel

Pengukuran jenis tulang pada data osteometri manusia *indigenous* merupakan variabel dalam penelitian ini. Terdapat 43 jenis pengukuran tulang panjang yang diukur yang menyusun kerangka manusia. Adapun 43 variabel yang dimaksud, antara lain:

Tabel 1. Variabel-variabel pada data *The Goldman Dataset*

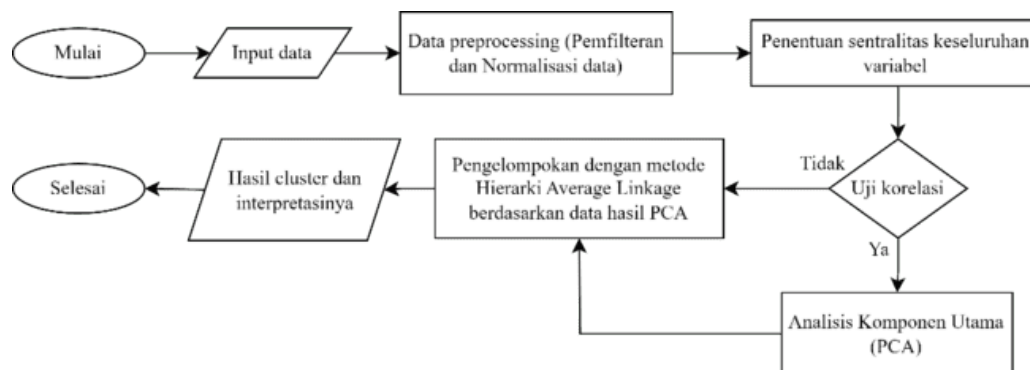
LHML	RRAPD	LTML
LHEB	LFML	LTPB
LHHD	LFBL	LTMLD
LHMLD	LFEB	LTAPD
LHAPD	LFAB	RTML
RHML	LFHD	RTPB
RHEB	LFMLD	RTMLD
RHHD	LFAPD	RTAPD
RHMLD	RFML	BIB
RHAPD	RFBL	LIBL
LRML	RFEB	RIBL
LRMLD	RFAB	LAcH
LRAPD	RFHD	RAcH
RRML	RFMLD	-
RRMLD	RFAPD	-

Variabel LHML, LHEB, LHHD, LHMLD, LHAPD, RHML, RHEB, RHHD, RHMLD, dan RHAPD merupakan jenis pengukuran yang dilakukan pada tulang lengan atas (*humerus*). Variabel LRML, LRMLD, LRAPD, RRML, RRMLD, dan RRAPD merupakan jenis pengukuran yang dilakukan pada tulang lengan bawah (*radius*). Variabel LFML, LFBL, LFEB, LFAB, LFHD, LFMLD, LFAPD, RFML, RFBL, RFEB, RFAB, RFHD, RFMLD, dan RFAPD merupakan jenis pengukuran yang dilakukan pada tulang paha (*femur*). Sementara variabel LTML, LTPB, LTMLD, LTAPD, RTML, RTPB, RTMLD, dan RTAPD merupakan jenis pengukuran yang dilakukan pada tulang kering (*tibia*). Sedangkan variabel BIB, LIBL, dan RIBL adalah jenis pengukuran yang dilakukan pada tulang pinggul (*iliac*) dan variabel LAcH dan RAcH adalah jenis pengukuran pada tulang panggul (*acetabulum*).

2.3. Metode Penelitian

Penelitian ini menggunakan metode *Average Linkage Clustering* untuk mengelompokkan ukuran tulang panjang manusia *indigenous* dan sentralitas untuk menentukan variabel paling signifikan di antara 43 variabel yang tercantum pada **Tabel 1**. Metode *Average Linkage Clustering* dapat dilakukan dengan menggunakan *software R Studio*, sedangkan penentuan sentralitas dilakukan dengan pemrograman *Python*.

Untuk mencapai hasil yang diinginkan dalam penelitian ini, beberapa langkah perlu dilakukan. Analisis dilakukan menggunakan metode *Average Linkage Clustering*, yang dapat dilihat pada diagram berikut ini.



Gambar 1. Diagram alir penelitian

Gambar 1 merupakan diagram alir dari penelitian ini dan dijabarkan secara sederhana pada langkah-langkah di bawah ini.

1. Mulai
2. Tahapan memasukkan data 43 jenis pengukuran tulang panjang ke dalam *software*
3. Melakukan pembersihan data kosong dan normalisasi data untuk mengatasi data pencilan
4. Penentuan sentralitas untuk menentukan variabel yang paling signifikan [5], dengan cara:

- Membuat matriks korelasi dari data pada tahapan 2 dengan persamaan:

$$r_{ij} = \frac{Cov(X_i, X_j)}{\sqrt{Var(X_i)}\sqrt{Var(X_j)}} = \frac{\sum_{n=1}^p (x_{ni} - \bar{x}_i)(x_{nj} - \bar{x}_j)}{\sqrt{\sum_{n=1}^p (x_{ni} - \bar{x}_i)^2} \sqrt{\sum_{n=1}^p (x_{nj} - \bar{x}_j)^2}} \quad (1)$$

- Menentukan matriks jarak (D) untuk menghitung jarak antar variabel dengan persamaan:

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

- Menentukan matriks ketetanggaan melalui matriks jarak. Adapun langkah-langkah untuk mengubah matriks jarak menjadi matriks ketetanggaan yaitu:



1. Misal $k = 2$, hitung $D^k = D \cdot D^{k-1}$ dengan perkalian matriks pada umumnya, tetapi perkalian 2 buah bilangan real a dan b didefinisikan $\max\{a, b\}$ dan penjumlahannya didefinisikan sebagai $\min\{a, b\}$
2. Jika $D^k = D^{k-1}$, maka proses dilanjutkan ke tahap 3. Jika tidak, maka kembali ke tahap 1 dengan nilai $k = k + 1$
3. Hitung $D - D^k$ untuk menentukan matriks ketetanggaan
4. Konversi matriks pada tahap 3 menjadi matriks ketetanggaan berentri 0 dan 1 dengan syarat: ubah entri matriks tersebut menjadi 1 jika hasil $D - D^k = 0$, ubah entri matriks tersebut menjadi 0 jika hasil $D - D^k > 0$, sedangkan pada diagonal utama matriks tersebut tidak dilakukan perubahan apapun
- Setelah menentukan matriks ketetanggaan berentri 0 dan 1, selanjutnya menghitung sentralitas derajat [6] menggunakan persamaan:

$$C_G(i) = \frac{\sum_{j=1}^N g_{ij}}{N-1} \quad (3)$$

dengan $C_G(i)$ adalah sentralitas derajat pada simpul ke- i , N adalah banyaknya simpul, dan g_{ij} adalah entri-entri matriks ketetanggaan

5. Kemudian, dilakukan pemeriksaan multikolinearitas (uji korelasi). Karena data pada tahap kedua menunjukkan adanya multikolinearitas, maka untuk mengatasi masalah ini dilakukan Analisis Komponen Utama (PCA).
6. Setelah itu melakukan pengelompokan data menggunakan algoritma *Average Linkage Clustering* dengan cara:
 - Menentukan jarak tiap data menggunakan metrik *Euclidean*

$$o_{ij} = \sqrt{\sum_{k=1}^p (x_{ik} - x_{jk})^2} \quad (4)$$

dengan o_{ij} merupakan jarak antara data i dan data j , x_{ik} merupakan nilai data i pada variabel ke- k , x_{jk} merupakan nilai data j pada variabel ke- k , dan p merupakan banyaknya variabel pada data.

- Menggabungkan dua *cluster* yang paling dekat. Jika jarak antara data a dan data b lebih kecil dibandingkan dengan jarak antara data lainnya, maka pada tahap pertama, dua *cluster* yang digabungkan adalah *cluster* (ab).
- Perbaharui matriks jarak antara *cluster* (ab) dengan *cluster* lainnya, seperti *cluster* c , dengan menggunakan rumus berikut:

$$o_{(ab)c} = \frac{\sum_i \sum_k o_{ik}}{N_{ab} N_c} \quad (5)$$

dengan o_{ik} adalah jarak antar data i pada *cluster* (ab) dan data k dalam *cluster* c , N_{ab} merupakan banyaknya data di *cluster* (ab) dan N_c merupakan banyaknya data di *cluster* c .

- Ulangi dari langkah 2 hingga seluruh data tergabung ke dalam sebuah *cluster*.

7. Membuat interpretasi dari hasil pengelompokan
8. Selesai.

III. HASIL DAN PEMBAHASAN

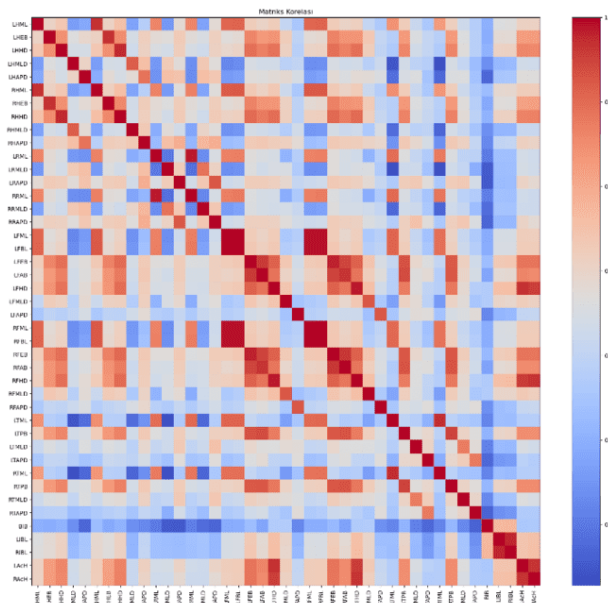
3.1 Data Preprocessing

Dalam *The Goldman Dataset*, terdapat 69 kolom yang berisi 1538 baris data dimana 43 kolom merupakan data yang akan dijadikan variabel pada penelitian ini. Data tersebut memiliki banyak

missing value atau data kosong untuk beberapa variabel. Kekurangan pada data seperti itu dapat menyebabkan kesalahan pada analisis lebih lanjut. Oleh karena itu, data perlu dilakukan pemfilteran untuk memisahkan data kosong tersebut dan tidak memasukkannya pada hasil interpretasi nantinya untuk menghindari *noise* atau kesalahan pada data [7]. Setelah melakukan pemfilteran data, selanjutnya yaitu menormalisasi data menggunakan standardisasi *z-score*. Hal ini dilakukan agar data memiliki rentang yang sama di setiap variabel [8]. Adapun tujuan dilakukannya hal tersebut yaitu untuk memastikan tiap variabel memiliki peranan yang optimal pada analisis data yang dilakukan [9]. Selain itu, normalisasi data *z-score* dilakukan untuk mengatasi data pencilan yang dapat mempengaruhi hasil pengelompokan.

3.2 Penentuan Sentralitas Keseluruhan Variabel

Pada tahap menentukan variabel mana yang paling berpengaruh (signifikan) berdasarkan besarnya hubungan variabel tersebut dengan variabel lain dapat diukur dengan menggunakan sentralitas. Sebelum menentukan sentralitas, terlebih dahulu menentukan matriks korelasi menggunakan (1) sehingga akan terbentuk matriks korelasi data 43 jenis pengukuran tulang panjang seperti pada Gambar 2.



Gambar 2. Diagram *heatmap* matriks korelasi keseluruhan variabel

Setelah menentukan matriks korelasi, kemudian akan dihitung jarak antar variabel menggunakan (2) sehingga akan terbentuk matriks jarak yang setiap entrinya menyatakan jarak antar variabel. Selanjutnya, akan ditentukan matriks ketetangaan berdasarkan matriks jarak untuk menghitung sentralitas derajat menggunakan algoritma pada tahapan 4. Setelah sentralitas derajat dihitung menggunakan (3), berikut merupakan hasil nilainya tiap variabel pada Tabel 2.

Tabel 2. Nilai sentralitas derajat keseluruhan variabel

Variabel	Sentralitas Derajat	Variabel	Sentralitas Derajat	Variabel	Sentralitas Derajat
LHML	0.02381	RRAPD	0.04762	LTML	0.07143
LHEB	0.02381	LFML	0.02381	LTPB	0.04762
LHHD	0.02381	LFBL	0.07143	LTMLD	0.04762
LHMLD	0.04762	LFEB	0.02381	LTAPD	0.04762
LHAPD	0.02381	LFAB	0.07143	RTML	0.02381



Variabel	Sentralitas Derajat	Variabel	Sentralitas Derajat	Variabel	Sentralitas Derajat
RHML	0.07143	LFHD	0.04762	RTPB	0.02381
RHEB	0.04762	LFMLD	0.02381	RTMLD	0.04762
RHHD	0.07143	LFAPD	0.02381	RTAPD	0.02381
RHMLD	0.02381	RFML	0.02381	BIB	0.02381
RHAPD	0.14286	RFBL	0.07143	LIBL	0.02381
LRML	0.02381	RFEB	0.11905	RIBL	0.07143
LRMLD	0.02381	RFAB	0.04762	LAcH	0.04762
LRAPD	0.04762	RFHD	0.09524	RAcH	0.02381
RRML	0.04762	RFMLD	0.04762	-	-
RRMLD	0.04762	RFAPD	0.04762	-	-

Berdasarkan **Tabel 2**, diperoleh bahwa variabel RHAPD memiliki ukuran sentralitas derajat terbesar dengan nilai 0.14286. Secara umum, sentralitas derajat digunakan untuk mengidentifikasi struktur penting dari suatu graf. Semakin besar nilai sentralitas derajat pada suatu simpul (variabel), maka interaksi simpul tersebut dengan simpul lainnya semakin banyak. Hal ini karena sentralitas derajat didasarkan pada jumlah koneksi suatu simpul pada graf [10]. Berdasarkan hal tersebut, maka variabel RHAPD memiliki peran yang sangat penting dalam mempengaruhi dan dipengaruhi oleh variabel lain sehingga menjadi variabel yang paling signifikan pada proses pengelompokan data.

3.3 Proses Pengelompokan Data 43 Jenis Pengukuran Tulang Panjang

Proses pengelompokan diawali dengan pemeriksaan apakah data saling berkorelasi atau tidak. Jika data saling berkorelasi (multikolinearitas), maka akan mempengaruhi hasil *cluster*. Untuk menentukan data berkorelasi atau tidak, dapat dilakukan menggunakan uji *Bartlett*. **Tabel 3** di bawah ini merupakan hasil dari uji *Bartlett* dari data tersebut menggunakan *software R Studio*.

Tabel 3. Nilai uji multikolinearitas keseluruhan variabel

Nilai uji <i>Bartlett</i>	
χ^2_{hitung}	53735
χ^2_{tabel}	903

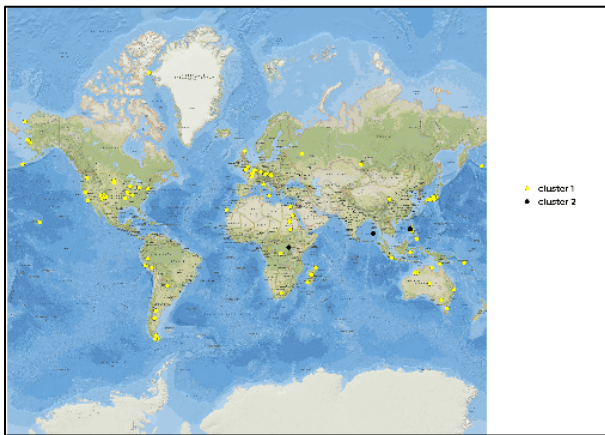
Berdasarkan **Tabel 3**, diperoleh bahwa $\chi^2_{hitung} = 53735 > \chi^2_{tabel} = 903$ sehingga dapat disimpulkan bahwa terdapat multikolinearitas pada data. Oleh karena itu, sebelum dilakukan pengelompokan, data terlebih dahulu dilakukan analisis faktor dengan menggunakan *Principal Component Analysis* (PCA).

Principal Component Analysis (PCA) atau analisis komponen utama adalah suatu teknik reduksi dimensi data sehingga dimensinya menjadi lebih kecil yang dapat mewakili informasi dari variabel asli [11]. Adapun dimensi data yang sebelumnya terdapat 43 kolom (variabel) direduksi menjadi 4 komponen baru yang mewakili 84.004% variansi dari data asli.

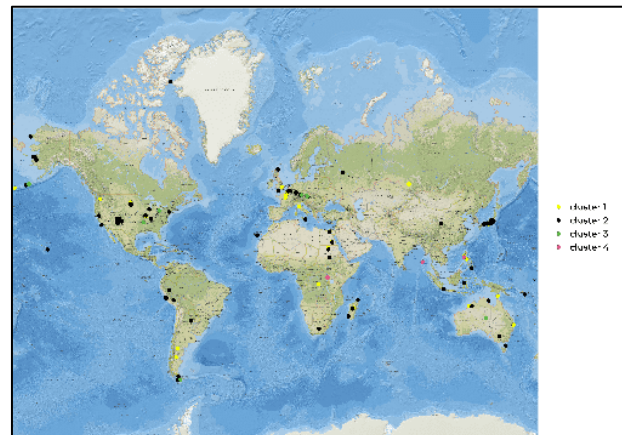
Proses selanjutnya yaitu pengelompokan data hasil PCA menggunakan algoritma *Average Linkage*. Algoritma *Average Linkage* merupakan algoritma pengklasteran berdasarkan jarak rata-rata antar data [12]. Adapun perhitungan jarak antar baris data hasil PCA dilakukan menggunakan (4) dan pembaruan jarak akibat terbentuknya *cluster* baru dilakukan menggunakan (5). Proses pengelompokan dilakukan sehingga data dikelompokkan menjadi 2 *cluster*, 4 *cluster*, 6 *cluster*, 8 *cluster*, 10 *cluster*, 14 *cluster*, dan 16 *cluster* sesuai dengan banyaknya subregion tempat ditemukannya tulang manusia *indigenous* tersebut pada *dataset* yaitu sebanyak 16 subregion berdasarkan langkah-langkah pada tahapan 6 menggunakan bantuan *software R Studio*.

3.4 Interpretasi Hasil Pengelompokan pada Data Hasil PCA

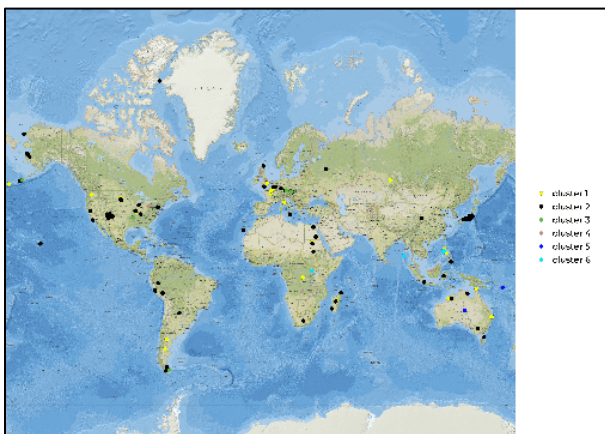
Setelah data dikelompokkan menjadi 2 *cluster*, 4 *cluster*, 6 *cluster*, 8 *cluster*, 10 *cluster*, 14 *cluster*, dan 16 *cluster* maka langkah selanjutnya adalah melakukan interpretasi terhadap hasil pengelompokan tersebut. Untuk melihat sebaran anggota tiap *cluster* ke dalam peta dunia, akan dibentuk plot titik koordinat yang sebelumnya ditentukan secara manual melalui *Google Maps* berdasarkan kolom “*Location*” pada *The Goldman Dataset*. Berikut merupakan plot titik koordinat seluruh anggota (spesimen) tiap *cluster* ke dalam peta dunia dengan bantuan *QGIS*.



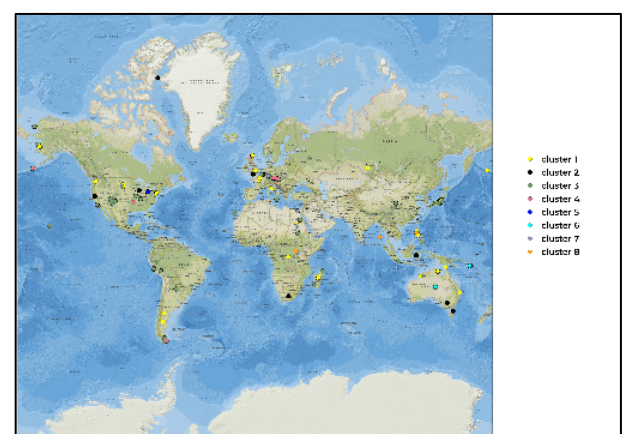
Gambar 4. Sebaran spesimen pada 2 *cluster*



Gambar 3. Sebaran spesimen pada 4 *cluster*



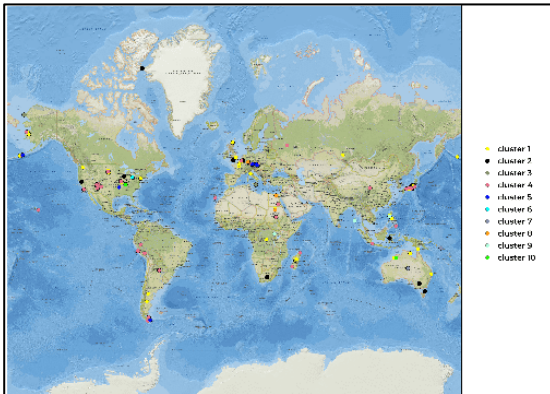
Gambar 5. Sebaran spesimen pada 6 *cluster*



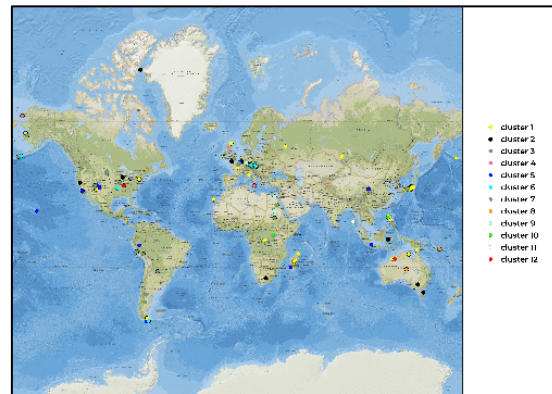
Gambar 6. Sebaran spesimen pada 8 *cluster*

Berdasarkan **Gambar 3-10**, terdapat karakteristik yang terbentuk dari 8 kali pengelompokan yaitu spesimen-spesimen yang ditemukan di subregion Amerika Utara, Amerika Selatan, Afrika Utara, Afrika Tengah, Afrika Timur, Eropa, dan Australia memiliki kemiripan yang sama. Sementara itu, spesimen-spesimen yang ditemukan di Benua Asia, Melanesia, dan Polinesia cenderung memiliki kemiripan dengan spesimen yang ditemukan di subregion Afrika Selatan. Untuk melihat karakteristik spesimen tiap *cluster* di 16 subregion berdasarkan ukurannya, dapat ditentukan melalui resume numerik berupa grafik *boxplot*. Setelah grafik *boxplot* dibuat, karakteristik spesimen yang ditemukan di Benua Afrika, Amerika, Eropa, dan Australia memiliki karakteristik ukuran tulang yang cenderung lebih panjang atau lebih besar dibanding subregion lain. Hal tersebut menandakan bahwa terdapat pembentukan karakteristik ukuran tulang pada lokasi geografis tertentu sehingga mengindikasikan

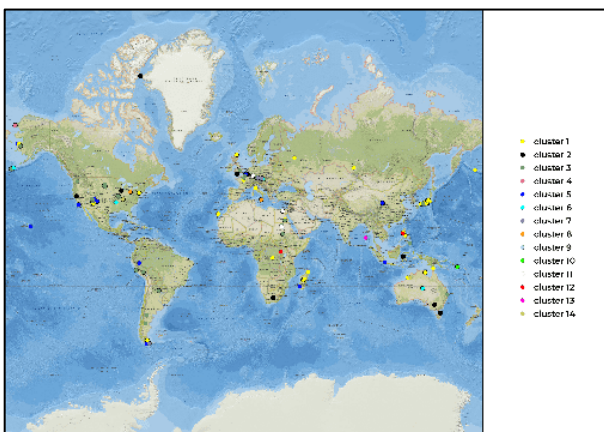
bahwa manusia pramodern berasal dari benua Afrika dan makhluk yang nomaden yang akhirnya membuat keragaman genetik, budaya, dan lingkungan pada saat ini.



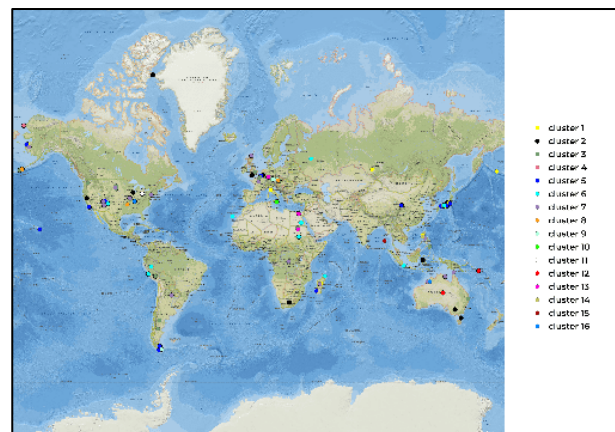
Gambar 7. Sebaran spesimen pada 10 *cluster*



Gambar 8. Sebaran spesimen pada 12 *cluster*



Gambar 9. Sebaran spesimen pada 14 *cluster*



Gambar 10. Sebaran spesimen pada 16 *cluster*

IV. KESIMPULAN

Berdasarkan hasil dan pembahasan yang telah dijabarkan, disimpulkan sebagai berikut:

1. Setelah dilakukan 8 kali pengelompokan, karakteristik spesimen-spesimen yang ditemukan di subregion Amerika Utara, Amerika Selatan, Afrika Utara, Afrika Tengah, Afrika Timur, Eropa, dan Australia memiliki kemiripan yang sama. Sementara itu, spesimen-spesimen yang ditemukan di Benua Asia, Melanesia, dan Polinesia cenderung memiliki kemiripan dengan spesimen yang ditemukan di subregion Afrika Selatan.
2. Berdasarkan hasil resume numerik tiap *cluster* berupa grafik *boxplot* di 16 subregion, karakteristik spesimen yang ditemukan di Benua Afrika, Amerika, Eropa, dan Australia memiliki karakteristik ukuran tulang yang cenderung lebih panjang atau lebih besar dibanding subregion lain. Hal tersebut menandakan bahwa terdapat pembentukan karakteristik ukuran tulang pada lokasi geografis tertentu sehingga mengindikasikan bahwa manusia pramodern berasal dari benua Afrika dan makhluk yang nomaden yang akhirnya membuat keragaman genetik, budaya, dan lingkungan pada saat ini. Dengan hasil ini diharapkan mengetahui dan menghormati masyarakat adat (manusia *indigenous*) sebagai dasar dari poin 15 dan 16 SDGs.



UCAPAN TERIMA KASIH

Ucapan terima kasih diberikan kepada Dr. Benjamin M. Auerbach, Ph.D. karena telah memberikan akses bebas untuk data hasil penelitiannya bernama *The Goldman Dataset* di laman resmi University of Tennessee – Knoxville (<https://web.utk.edu/~auerbach/DATA>).

REFERENSI

1. M. Mahdayeni, M. R. Alhaddad, and A. S. Saleh, “Manusia dan Kebudayaan (Manusia dan Sejarah Kebudayaan, Manusia dalam Keanekaragaman Budaya dan Peradaban, Manusia dan Sumber Penghidupan),” *Tadbir: Jurnal Manajemen Pendidikan Islam*, vol. 7, no. 2, pp. 154–165, Aug. 2019, doi: 10.30603/tjmpi.v7i2.1125.
2. M. Ingman, H. Kaessmann, S. Pääbo, and U. Gyllensten, “Mitochondrial genome variation and the origin of modern humans,” *Nature*, vol. 408, no. 6813, pp. 708–713, Dec. 2000, doi: 10.1038/35047064.
3. S. López, L. Van Dorp, and G. Hellenthal, “Human Dispersal Out of Africa: A Lasting Debate,” *Evolutionary Bioinformatics*, vol. 11s2, p. EBO.S33489, Jan. 2015, doi: 10.4137/EBO.S33489.
4. B. M. Auerbach and C. B. Ruff, “Human body mass estimation: A comparison of ‘morphometric’ and ‘mechanical’ methods,” *Am J Phys Anthropol*, vol. 125, no. 4, pp. 331–342, Dec. 2004, doi: 10.1002/ajpa.20032.
5. M. A. Djauhari and S. L. Gan, “Minimal spanning tree problem in stock networks analysis: An efficient algorithm,” *Physica A: Statistical Mechanics and its Applications*, vol. 392, no. 9, pp. 2226–2234, May 2013, doi: 10.1016/j.physa.2012.12.032.
6. K. Vignery and W. Laurier, “A methodology and theoretical taxonomy for centrality measures: What are the best centrality indicators for student networks?,” *PLoS One*, vol. 15, no. 12, p. e0244377, Dec. 2020, doi: 10.1371/journal.pone.0244377.
7. S. Kaka, S. D.I. Mau, and A. Purnami Setiawati, “PENERAPAN SISTEM INFORMASI AKADEMIK PENGOLAHAN DATA PRESTASI PEMBELAJARAN SISWA SDN WERY BERBASIS WEBSITE,” *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 7, no. 4, pp. 2647–2651, Jan. 2024, doi: 10.36040/jati.v7i4.7794.
8. D. A. Nasution, H. H. Khotimah, and N. Chamidah, “Perbandingan Normalisasi Data untuk Klasifikasi Wine Menggunakan Algoritma K-NN,” *Computer Engineering, Science and System Journal*, vol. 4, no. 1, p. 78, Jan. 2019, doi: 10.24114/cess.v4i1.11458.
9. R. G. Whendasmoro and J. Joseph, “Analisis Penerapan Normalisasi Data Dengan Menggunakan Z-Score Pada Kinerja Algoritma K-NN,” *JURIKOM (Jurnal Riset Komputer)*, vol. 9, no. 4, p. 872, Aug. 2022, doi: 10.30865/jurikom.v9i4.4526.
10. S. Muazir, L. Lestari, M. Nurhamsyah, and M. R. Alhamdani, “Komparasi sentralitas pusat kegiatan pada kota dan kabupaten di Provinsi Kalimantan Barat,” *Region: Jurnal Pembangunan Wilayah dan Perencanaan Partisipatif*, vol. 19, no. 1, p. 1, Jan. 2024, doi: 10.20961/region.v19i1.61885.
11. A. RACHMAWARDANI, S. K. WIJAYA, P. PRAWITO, and A. SOPAHELWAKAN, “Prediksi Banjir menggunakan ANFIS-PCA sebagai Peringatan Dini Bencana Banjir,” *ELKOMIKA: Jurnal Teknik Energi Elektrik, Teknik Telekomunikasi, & Teknik Elektronika*, vol. 12, no. 2, p. 335, Apr. 2024, doi: 10.26760/elkomika.v12i2.335.
12. S. I. Maiyanti, Oki Dwipurwani, Azizah, and Nurul Aisyah, “Pengelompokan Provinsi Di Indonesia Berdasarkan Konsumsi Kalori Per Kapita Sehari Menurut Kelompok Komoditas / Makanan Menggunakan Average Linkage Dan Ward’s Method,” *MULTIPLE: Journal of Global and Multidisciplinary*, vol. 2, no. 5, pp. 1698–1713, May 2024, [Online]. Available: <https://journal.institercom-edu.org/index.php/multiple/article/view/361>