

## KLASTERISASI LITERATUR MAHASISWA MENGGUNAKAN METODE AHC DI DINAS KEARSIPAN DAN PERPUSTAKAAN PROVINSI JAWA TENGAH

Annisa Nur Fadhilah<sup>1</sup>, Arif Jananto<sup>2</sup>

Program Studi Sistem Informasi, Fakultas Teknologi Informasi, Universitas Stikubank Semarang  
e-mail: [1annisanurfadhilah@gmail.com](mailto:1annisanurfadhilah@gmail.com), [2ajananto09@edu.unisbank.ac.id](mailto:2ajananto09@edu.unisbank.ac.id)

### ABSTRAK

Dinas Kearsipan dan perpustakaan Provinsi Jawa Tengah adalah suatu instansi pemerintahan yang menjadi pusat pengelolaan buku dan arsip yang ada di Provinsi Jawa Tengah. pelajar dan mahasiswa untuk memenuhi kebutuhan literasi yang diperlukan atau untuk sekedar membaca buku. Namun, seringkali buku yang diperlukan oleh mahasiswa persediaannya terbatas dan terkadang statusnya masih dipinjam. Untuk mengelompokkan buku yang sering dipinjam agar pihak perpustakaan dapat lebih efektif dalam menambahkan koleksi buku yang ada pada perpustakaan dapat menggunakan metode Single Linkage dan Complete Linkage pada Agglomerative Hierarchical Clustering (AHC).

Hasil yang didapatkan pada penelitian ini yaitu untuk metode Single Linkage menggunakan data sebanyak 1890 data dengan cluster 1 sebanyak 1840 data, cluster 2 sebanyak 34 data, dan cluster 3 sebanyak 16 data. Untuk metode Complete Linkage menggunakan 1890 data dengan hasil cluster 1 sebanyak 1592 data, cluster 2 sebanyak 248 data, dan cluster 3 sebanyak 3 data.

Dari hasil yang didapat tersebut, dapat diketahui bahwa atribut pekerjaan yang mendominasi adalah mahasiswa dan atribut kategori buku yang mendominasi adalah Ilmu Umum, sehingga pihak perpustakaan dapat menambahkan koleksi Ilmu Umum dan mengurangi koleksi dengan kategori Fiksi.

**Kata Kunci:** *Agglomerative Hierarchical Clustering (AHC), Single Linkage, Complete Linkage, Koleksi Buku Perpustakaan*

### 1. PENDAHULUAN

Dinas Kearsipan dan perpustakaan Provinsi Jawa Tengah adalah suatu instansi pemerintahan yang menjadi pusat pengelolaan buku dan arsip yang ada di Provinsi Jawa Tengah. Buku dan arsip ini dikelola menggunakan komputer yang saling terkoneksi antara satu dan lainnya. Bagian layanan sirkulasi yang menangani peminjaman buku dan data setiap daerah sehingga banyak menggunakan komputer dan sistem informasi sebagai alat agar pekerjaannya lebih rapi dan lebih tertata. Sistem informasi yang digunakan di Dinas Kearsipan dan Perpustakaan Provinsi Jawa Tengah sudah berbasis Web yang memudahkan berbagai pekerjaan yang dapat menguras waktu sehingga lebih efektif. Contohnya seperti peminjaman dan pengembalian buku. Dinas Kearsipan dan Perpustakaan Jawa Tengah menjadi pusat bagi pelajar dan mahasiswa untuk memenuhi kebutuhan literasi yang diperlukan atau untuk sekedar membaca buku. Namun, seringkali buku yang diperlukan oleh mahasiswa persediaannya terbatas dan terkadang statusnya masih dipinjam.

Salah satu metode *clustering* atau pengelompokkan data mining dapat membantu mengetahui pola pengelompokkan buku yang sering dipinjam sehingga menciptakan kluster dan lainnya akan berada pada kluster yang berbeda. Hal ini dapat membantu pihak Perpustakaan menambahkan koleksi buku yang dibutuhkan oleh mahasiswa.

Dalam penelitian ini, peneliti tertarik untuk mengelompokkan minat dan kebutuhan literasi mahasiswa menggunakan *Agglomerative Hierarchical Clustering (AHC)*. Maka itu, judul yang diambil dari penelitian ini adalah Klasterisasi Kebutuhan Literatur Mahasiswa Menggunakan Metode AHC.

### 2. TINJAUAN PUSTAKA

Penelitian yang dilakukan oleh Ginanjar Abdurrahman dari Universitas Muhammadiyah Jember pada tahun 2018 dengan menggunakan 822 *record* data menggunakan membahas tentang teknik *clustering* data kredit bank menggunakan *Agglomerative Hierarchical Clustering* untuk mempermudah menentukan *clustering* dataset kredit bank. Penelitian ini memiliki hasil akurat, *cluster* 1 sebanyak 98% , *cluster* 2 sebanyak 1%, dan sisanya 1%.

Penelitian yang dilakukan oleh Widyawati, mahasiswi STMIK Sinar Nusantara Surakarta pada tahun 2020 yang mengambil tema penerapan metode *Agglomerative Hierarchical Clustering (AHC)* untuk segmentasi pelanggan. Penelitian ini bertujuan untuk membentuk *clustering* pelanggan agar penyediaan barang lebih efektif. Pada penelitian ini peneliti menggunakan *Average Linkage* dan *Ward Linkage*,

dengan hasil yang diperoleh adalah *cluster* 1 sebanyak 20 anggota, *cluster* 2 sebanyak 43 anggota, *cluster* 3 sebanyak 75 anggota, *cluster* 4 sebanyak 158 anggota, *cluster* 5 sebanyak 9 anggota, *cluster* 6 sebanyak 2 anggota, *cluster* 7 sebanyak 43 anggota. Penelitian ini mempunyai hasil akhir pengelompokkan beberapa *cluster* dan memiliki strategi promosi yang berbeda pada tiap hasil *cluster*.

Penelitian yang dilakukan oleh Aloysius Ari Kurniawan mahasiswa Universitas Sanata Dharma pada tahun 2017 dengan menggunakan *record* data nilai sekolah dari SD Model Kabupaten Sleman. Peneliti mendapatkan hasil akhir menggunakan metode *Agglomerative Hierarchical Clustering (AHC)* mendapatkan *cluster* 1 sebanyak 14, *cluster* 2 sebanyak 22, dan *cluster* 3 sebanyak 31.

Penelitian yang dilakukan Lisna Zahrotun mahasiswi Universitas Ahmad Dahlan pada tahun 2015 menggunakan *record* data penumpang bus Trans Jogja. Penelitian ini menggunakan metode *Agglomerative Hierarchical Clustering (AHC)* dan mendapatkan hasil akhir pengelompokkan penumpang bus kedalam 3 *cluster*. Hasil dari penelitian ini diperoleh *cluster* 0 sebanyak 2 data, *cluster* 1 sebanyak 57 data, dan *cluster* 2 sebanyak 54 data.

### 3. METODE PENELITIAN

#### 3.1 Metode Penelitian

Obyek penelitian yang digunakan adalah data peminjaman dan pengembalian buku pada Dinas Kearsipan dan Perpustakaan Provinsi Jawa Tengah.

#### 3.2 Metode Pengumpulan Data

##### 1. Metode Observasi

Suatu metode pengumpulan data yang dilakukan secara langsung kelapangan untuk memperoleh data yang dibutuhkan.

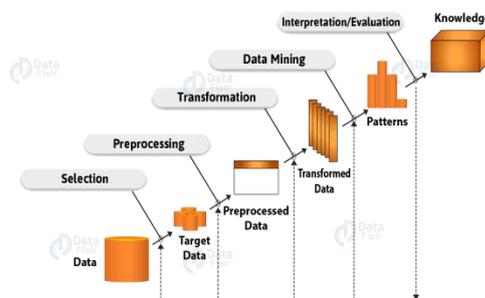
##### 2. Studi Kepustakaan

Suatu metode pengumpulan data dengan membaca dan mempelajari buku- buku yang ada dipustaka maupun buku diluar perpustakaan yang berhubungan dengan penelitian.

#### 3.3 Metode Analisis Data

Metode analisa yang digunakan adalah *Agglomerative Hierarchical Clustering (AHC)* dengan menggunakan metode *Knowledge Discovery and Data Mining (KDD)*. Menurut (Santoso, 2007) KDD adalah kegiatan yang meliputi pengumpulan, pemakaian data, historis, untuk menemukan keteraturan, pola, atau hubungan dalam set data berukuran besar. Rangkaian proses KDD digambarkan sebagai berikut:

1. *Data cleaning*, menghilangkan noise dan data yang inkonsisten.
2. *Data integration*, menggabungkan data dari berbagai sumber data yang berbeda.
3. *Data selection*, mengambil data yang relevan dengan tugas analisis dari database.
4. *Data transformation*, mentransformasi atau mengubah data ke dalam bentuk yang lebih sesuai untuk penggalian lewat operasi summary atau aggregation.
5. *Data mining*, proses esensial untuk mengekstrak pola dari data dengan metode cerdas.
6. *Pattern evaluation*, mengidentifikasi pola yang menarik dan merepresentasikan pengetahuan berdasarkan interestingness measures.
7. berdasarkan interestingness measures.
8. *Knowledge presentation*, penyajian pengetahuan yang digali kepada pengguna dengan menggunakan visualisasi dan teknik representasi pengetahuan.



Gambar 3.1 Proses Data Mining (Shapiro, 2017)

#### 4. HASIL DAN PEMBAHASAN

##### 4.1 Perhitungan Manual *Agglomerative Clustering Hierarchial (AHC)*

Algoritma yang digunakan pada penelitian ini adalah *Agglomerative Clustering Hierarchial* yang mengelompokkan data kedalam beberapa cluster. Namun, sebelum mengelompokkan data kedalam cluster, matriks jarak harus dihitung terlebih dahulu. Rumus yang digunakan peneliti dalam penelitian ini adalah *Euclidean Distance*. Data yang akan diolah menggunakan empat atribut, yaitu jenis kelamin, pekerjaan, kategori buku, dan status buku.

##### 1. Menghitung Jarak

Menghitung semua jarak data menggunakan metode *Euclidean Distance*.

- Data (1,1)
 
$$= \sqrt{(2-2)^2 + (2-2)^2 + (1-1)^2 + (1-1)^2}$$

$$= \sqrt{0^2 + 0^2 + 0^2 + 0^2}$$

$$= 0$$
- Data (1,2)
 
$$= \sqrt{(2-2)^2 + (2-1)^2 + (1-1)^2 + (1-1)^2}$$

$$= \sqrt{0^2 + 1^2 + 0^2 + 0^2}$$

$$= 1$$
- Data (1,3)
 
$$= \sqrt{(2-1)^2 + (2-4)^2 + (1-2)^2 + (2-1)^2}$$

$$= \sqrt{1^2 + 2^2 + 1^2 + 1^2}$$

$$= 2,449$$
- Data (1,4)
 
$$= \sqrt{(2-2)^2 + (2-1)^2 + (1-2)^2 + (1-1)^2}$$

$$= \sqrt{0^2 + 1^2 + 1^2 + 0^2}$$

$$= 1,414$$
- Data (1,5)
 
$$= \sqrt{(2-2)^2 + (2-2)^2 + (1-2)^2 + (1-1)^2}$$

$$= \sqrt{0^2 + 0^2 + 1^2 + 0^2}$$

$$= 1$$
- Data (2,3)
 
$$= \sqrt{(2-1)^2 + (1-4)^2 + (1-2)^2 + (1-1)^2}$$

$$= \sqrt{1^2 + 3^2 + 1^2 + 0^2}$$

$$= 3,162$$
- Data (2,4)
 
$$= \sqrt{(2-2)^2 + (1-1)^2 + (1-2)^2 + (1-1)^2}$$

$$= \sqrt{0^2 + 0^2 + 1^2 + 0^2}$$

$$= 1$$
- Data (2,5)
 
$$= \sqrt{(2-2)^2 + (1-2)^2 + (1-2)^2 + (1-1)^2}$$

$$= \sqrt{0^2 + 1^2 + 1^2 + 0^2}$$

$$= 1,414$$
- Data (3,4)
 
$$= \sqrt{(1-2)^2 + (4-1)^2 + (2-2)^2 + (1-1)^2}$$

$$= \sqrt{1^2 + 3^2 + 0^2 + 0^2}$$

$$= 3,162$$
- Data (3,5)
 
$$= \sqrt{(1-2)^2 + (4-2)^2 + (2-2)^2 + (1-1)^2}$$

$$= \sqrt{1^2 + 2^2 + 0^2 + 0^2}$$

$$= 2,236$$
- Data (4,5)
 
$$= \sqrt{(2-2)^2 + (1-2)^2 + (2-2)^2 + (1-1)^2}$$

$$= \sqrt{0^2 + 1^2 + 0^2 + 0^2}$$

$$= 1$$

Tabel 4.1 Hasil Perhitungan Matriks menggunakan *Euclidean Distance*

Data	1	2	3	4	5
1	0	1	2,449	1,414	1
2	1	0	3,162	1	1,414
3	2,449	3,162	0	3,162	2,236
4	1,414	1	3,162	0	1
5	1	1,414	2,236	1	0

## 2. Perhitungan Manual *Agglomerative Hierarchical Clustering (AHC)* menggunakan metode *Single Linkage*

Setelah mendapatkan matriks jarak, perhitungan AHC menggunakan *Single Linkage* mempunyai langkah-langkah sebagai berikut:

Dari matriks yang sudah didapatkan tersebut, gabungkan dua matrik terdekat.

$$d_{1,2} = \min \{1\} = 1$$

$$d_{4,5} = \min \{1\} = 1$$

Setelah mendapat dua kelompok terdekat selanjutnya menggabungkan dengan kelompok yang tersisa.

$$\begin{aligned} d_{(12)(45)} &= \min \{d_{14}, d_{15}, d_{24}, d_{25}\} \\ &= \min \{1,414, 1, 1, 1,414\} \\ &= 1 \end{aligned}$$

$$\begin{aligned} d_{(12)(3)} &= \min \{d_{13}, d_{23}\} \\ &= \min \{1, 2,449\} \\ &= 1 \end{aligned}$$

$$\begin{aligned} d_{(45)(3)} &= \min \{d_{43}, d_{53}\} \\ &= \min \{3,162, 2,236\} \\ &= 2,236 \end{aligned}$$

Kolom dan baris (1 dan 2) dan (4 dan 5) yang sebelumnya telah disatukan selanjutnya dihapus dan kemudian ditambahkan baris dan kolom baru yaitu kolom dan baris (12) dan (45).

Tabel 4.2 Hasil Pengelompokkan ke-1  
(*Single Linkage*)

Data	1,2	3	4,5
1,2	0	1	1
3	1	0	2,236
4,5	1	2,236	0

Setelah mendapat dua kelompok terkecil selanjutnya menggabungkan dengan kelompok yang tersisa.

$$\begin{aligned} d_{(123)(45)} &= \min \{d_{14}, d_{24}, d_{34}, d_{15}, d_{25}, d_{35}\} \\ &= \min \{1,414, 1, 3,162, 1, 1,414, 2,236\} \\ &= 1 \end{aligned}$$

Setelah menggabungkan (1, 2 dan 3) dan (4 dan 5), selanjutnya kolom dan baris tersebut dihapus dan kemudian ditambahkan baris dan kolom baru yaitu kolom dan baris (123) dan (45).

Tabel 4.3 Tabel Hasil Pengelompokkan ke-2  
(*Single Linkage*)

Data	1,2,3	4,5
1,2,3	0	1
4,5	1	0

3. Perhitungan Manual *Agglomerative Hierarchial Clustering (AHC)* menggunakan metode *Complete Linkage*

Complete Linkage merupakan metode menghitung jarak matriks terjauh. Dari matriks yang sudah didapatkan tersebut, gabungkan dua jarak terdekat.

$$d_{1,5} = \min \{1\} = 1$$

$$d_{2,4} = \min \{1\} = 1$$

Setelah mendapat dua kelompok terdekat selanjutnya menggabungkan dengan kelompok yang tersisa.

$$\begin{aligned} d_{(15)(24)} &= \max \{d_{12}, d_{52}, d_{14}, d_{54}\} \\ &= \max \{1, 1,414, 1,414, 1\} \\ &= 1,414 \end{aligned}$$

$$\begin{aligned} d_{(15)(3)} &= \max \{d_{15}, d_{13}\} \\ &= \max \{1, 2,449\} \\ &= 2,449 \end{aligned}$$

$$\begin{aligned} d_{(24)(3)} &= \max \{d_{24}, d_{23}\} \\ &= \max \{1, 3,162\} \\ &= 3,162 \end{aligned}$$

Kolom dan baris (1, 5) dan (2, 4) yang sebelumnya telah disatukan selanjutnya dihapus dan kemudian ditambahkan baris dan kolom baru yaitu kolom dan baris (15) dan (24).

Tabel 4.4 Hasil Pengelompokkan ke-1  
(*Complete Linkage*)

Data	1,5	2,4	3
1,5	0	1,414	2,449
2,4	1,414	0	3,162
3	2,449	3,162	0

Setelah mendapat dua kelompok terbesar selanjutnya menggabungkan dengan kelompok yang tersisa.

$$\begin{aligned} d_{(1524)(3)} &= \min \{d_{13}, d_{53}, d_{23}, d_{43}\} \\ &= \min \{2,449, 2,236, 3,162, 3,162\} \\ &= 2,236 \end{aligned}$$

Setelah menggabungkan (1, 5, 2, 4) dan (3) selanjutnya kolom dan baris tersebut dihapus dan kemudian ditambahkan baris dan kolom baru yaitu kolom dan baris (1524) dan (3).

Tabel 3.5 Tabel Hasil Pengelompokkan Ke-2  
(*Complete Linkage*)

Data	1,5,2,4	3
1,5,2,4	0	3,162
3	3,162	0

4.2 Implementasi Menggunakan RStudio

1. Data Preparation

Tahap pertama dalam mengolah data menggunakan Rstudio adalah mengimport data. Data yang akan diolah diimport dalam format Excel maupun CSV.

```
skripsi<-read_excel("D/Skripsi/Data/Skripsifix.xlsx")
view(Skripsifix)
```

Gambar

Data yang digunakan untuk selanjutnya diproses pada Rstudio sebanyak 1.890 data setelah melalui proses cleaning data. Proses selanjutnya adalah meng-*install packages*. Packages *car* merupakan suatu package yang berfungsi untuk fungsi regresi.

```
library(car)
```

Langkah kedua adalah menghilangkan kolom pertama atau kolom yang tidak diproses yaitu kolom 'No'.

```
skripsi <- skripsifix [,-c(1)]
skripsi
```

Source code ini mempunyai fungsi menyalin data 'skripsifix' yang sebelumnya diimpor. Data akan menjadi duplikasi dengan menghilangkan kolom 'No' dengan nama 'skripsi'. Tujuan dari proses ini adalah agar data asli tidak hilang.

Langkah ketiga adalah melakukan uji korelasi. Uji Korelasi digunakan untuk menampilkan hubungan korelasi antar variabel dengan rentang nilai -1 sampai 1.

```
multikol <- cor(skripsi)
multikol
```

Source code multikol mempunyai fungsi mengetahui hubungan antar variabel dengan nilai korelasi harus = 1. Multikolinearitas merupakan hubungan antar variabel independen.

```
> multikol
      Jenis Kelamin  Pekerjaan Kategori Buku Status Buku
Jenis Kelamin  1.000000000 -0.22541200  0.000262221 -0.02072263
Pekerjaan     -0.225412001  1.00000000  0.074364960 -0.04597123
Kategori Buku  0.000262221  0.07436496  1.000000000 -0.01369366
Status Buku   -0.020722635 -0.04597123 -0.013693660  1.00000000
> |
```

Gambar 4.1. Tampilan Hasil Uji Korelasi

Langkah keempat yaitu menghitung matriks jarak antar data menggunakan fungsi 'dist'.

```
jarak=dist(skripsi)
jarak
```

Perhitungan jarak matriks digunakan untuk menentukan kelompok berdasarkan metode yang digunakan selanjutnya.

```
> jarak
      1      2      3      4      5      6      7
2  1.000000
3  2.449490  3.316625
4  1.414214  1.000000  3.162278
5  1.000000  1.414214  2.236068  1.000000
6  2.236068  3.162278  1.000000  3.000000  2.000000
7  1.000000  0.000000  3.316625  1.000000  1.414214  3.162278
8  1.414214  1.732051  2.449490  1.414214  1.000000  2.236068  1.732051
9  1.414214  1.000000  3.162278  0.000000  1.000000  3.000000  1.000000
10 1.732051  2.449490  1.000000  2.236068  1.414214  1.414214  2.449490
11 1.000000  1.414214  2.236068  1.000000  0.000000  2.000000  1.414214
```

Gambar 4.2. Tampilan Hasil Perhitungan Jarak

Diketahui jarak data pertama dan kedua adalah 1,00 sedangkan jarak data pertama dan ketiga adalah 2,449. Sedangkan jarak untuk data kedua dan ketiga adalah 3,316 dan seterusnya.

## 2. Implementasi Agglomerative Hierarchical Clustering (AHC)

### Single Linkage

Fungsi hclust() dalam r dapat digunakan untuk menghitung cluster terdekat dengan menambahkan argumen "single" pada method.

```
ahc <- hclust(dist(scale(skripsi)), method = "single")
ahc
```

### Complete Linkage

Fungsi hclust() dalam r dapat digunakan untuk menghitung cluster terjauh dengan menambahkan argumen "complete" pada method.

```
ahc <- hclust(dist(scale(skripsi)), method = "complete")
ahc
```

- Membuat Plot DendrogramPlot(ahc) adalah code yang dapat menampilkan hasil plot diagram dari hasil clustering. Sedangkan code rect.hclust(ahc,3) adalah source code yang mengelompokkan data menjadi kelompok cluster yang sudah ditentukan sebelumnya.

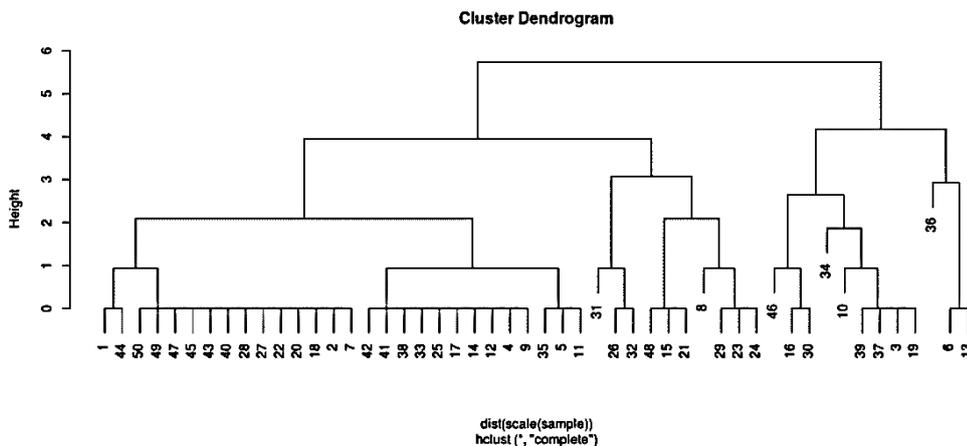
```
plot(ahc)
rect.hclust(ahc,3)
```

*Single Linkage*



Gambar 4.3 Hasil Plot Dendrogram *Single Linkage*

*Complete Linkage*



Gambar 4.4 Hasil Plot Dendrogram *Complete Linkage*

- Menghitung Nilai Korelasi *Cophenetic*  
 Nilai Korelasi *Cophenetic* merupakan nilai korelasi antara nilai yang mempunyai ketidakmiripan dan nilai yang dihasilkan oleh dendrogram. Nilai Korelasi *Cophenetic* berkisar antara -1 sampai 1. Semakin mendekati 1, maka semakin baik metode yang digunakan.

*Single Linkage*

```
d1 <- dist(skripsifix)
hc <- hclust(d1,"Single")
d2 <- cophenetic(hc)
corave=cor(d1,d2)
corave

> corave
[1] 0.5705934
```

Gambar 4.5 Hasil Perhitungan Korelasi *Cophenetic* (*Single Linkage*)

Complete Linkage

```
d1 <- dist(skripsifix)
hc <- hclust(d1, "Complete")
d2 <- cophenetic(hc)
corave=cor(d1,d2)
corave
> corave
[1] 0.7504569
```

Gambar 4.6 Hasil Perhitungan Korelasi *Cophenetic (Complete Linkage)*

Menurut hasil perhitungan nilai korelasi *cophenetic* metode *Complete Linkage* dianggap sebagai metode terbaik karena nilainya lebih mendekati 1.

4.3 Hasil Rekap *Running Data* pada RStudio

Hasil *running* Rstudio untuk Keseluruhan data yang sebelumnya telah dicleaning adalah sebanyak 1890 data.

Single Linkage

Cluster 1	Jenis Kelamin	Laki-Laki	484	1840
		Perempuan	1356	
	Pekerjaan	Mahasiswa	1319	1840
		Pelajar	273	
		Pegawai Negeri	97	
		Pegawai Swasta	119	
		Umum	32	
	Kategori Buku	Pendidikan	715	1840
		Ilmu Umum	1103	
		Fiksi	22	
Status Buku	Dikembalikan	1840	1840	
	Belum Dikembalikan	0		
Cluster 2	Jenis Kelamin	Laki-Laki	0	34
		Perempuan	34	
	Pekerjaan	Mahasiswa	28	34
		Pelajar	5	
		Pegawai Negeri	1	
		Pegawai Swasta	0	
		Umum	0	
	Kategori Buku	Pendidikan	15	34
		Ilmu Umum	19	
		Fiksi	0	
Status Buku	Dikembalikan	0	34	
	Belum Dikembalikan	34		
Cluster 3	Jenis Kelamin	Laki-Laki	16	16
		Perempuan	0	
	Pekerjaan	Mahasiswa	13	16
		Pelajar	2	
		Pegawai Negeri	0	
		Pegawai Swasta	1	
		Umum	0	
	Kategori Buku	Pendidikan	6	16
		Ilmu Umum	10	
		Fiksi	0	
Status Buku	Dikembalikan	0	16	
	Belum Dikembalikan	16		

Gambar 4.7 Hasil Rekap Rstudio *Single Linkage*

Complete Linkage

Cluster 1	Jenis Kelamin	Laki-Laki	352	1592
		Perempuan	1240	
	Pekerjaan	Mahasiswa	1319	1592
		Pelajar	273	
		Pegawai Negeri	0	
		Pegawai Swasta	0	
		Umum	0	
	Kategori Buku	Pendidikan	637	1592
		Ilmu Umum	938	
		Fiksi	17	
Status Buku	Dikembalikan	1592	1592	
	Belum Dikembalikan	0		
Cluster 2	Jenis Kelamin	Laki-Laki	132	248
		Perempuan	116	
	Pekerjaan	Mahasiswa	0	248
		Pelajar	0	
		Pegawai Negeri	97	
		Pegawai Swasta	119	
		Umum	32	
	Kategori Buku	Pendidikan	78	248
		Ilmu Umum	165	
		Fiksi	5	
Status Buku	Dikembalikan	248	248	
	Belum Dikembalikan	0		
Cluster 3	Jenis Kelamin	Laki-Laki	16	3
		Perempuan	34	
	Pekerjaan	Mahasiswa	41	50
		Pelajar	7	
		Pegawai Negeri	1	
		Pegawai Swasta	1	
		Umum	0	
	Kategori Buku	Pendidikan	21	50
		Ilmu Umum	29	
		Fiksi	0	
Status Buku	Dikembalikan	0	50	
	Belum Dikembalikan	50		

Gambar 4.8 Hasil Rekap Rstudio *Complete Linkage*

5. KESIMPULAN

1. Hasil pembagian *cluster* pada metode *Single Linkage* dan metode *Complete Linkage*, keduanya memiliki 3 cluster dan dapat ditarik kesimpulan bahwa anggota yang sering meminjam buku di Dinas Kearsipan dan Perpustakaan Provinsi Jawa Tengah didominasi oleh Mahasiswa berjenis kelamin perempuan. Sedangkan kategori buku yang paling banyak dipinjam oleh Mahasiswa adalah Ilmu Umum kemudian Pendidikan dan 97% buku sudah dikembalikan.
2. Anggota terbanyak dari atribut pekerjaan yaitu Mahasiswa lebih banyak membutuhkan dan berminat pada buku dengan kategori Ilmu Umum sehingga pihak Dinas Kearsipan dan Provinsi

Jawa Tengah dapat menambahkan koleksi Ilmu Umum dan menambahkan beberapa koleksi buku dengan kategori Pendidikan. Pihak perpustakaan juga dapat mengurangi koleksi untuk buku dengan kategori Fiksi.

3. Berdasarkan hasil korelasi *cophenetic* metode *Single Linkage* memiliki nilai korelasi sebesar 0,5705934 sedangkan *Complete Linkage* memiliki nilai korelasi sebesar 0,7504569. Dengan hasil ini, dapat disimpulkan bahwa metode *Complete Linkage* merupakan metode terbaik karena mempunyai nilai korelasi tertinggi.

## 6. SARAN

Saran yang dapat diberikan peneliti adalah sebagai berikut:

1. Penelitian ini memiliki data yang kurang beragam sehingga hasil *cluster* yang dihasilkan terlalu berdekatan. Untuk penelitian selanjutnya, diharapkan dapat menggunakan data yang lebih beragam agar didapatkan hasil pengelompokan *cluster* yang lebih baik.
2. Pada penelitian selanjutnya dapat ditambahkan dengan menggunakan metode *Average Linkage*.
3. Penelitian ini juga dapat dimanfaatkan oleh pihak Dinas Kearsipan dan Perpustakaan Jawa Tengah sehingga penambahan koleksi buku secara efektif dapat dioptimalkan.

## DAFTAR PUSTAKA

- [1] Abdurrahman, G. (2018). Clustering Data Kredit Bank Menggunakan Algoritma Agglomerative Hierarchical Clustering Average Linkage. *Jurnal Sistem & Teknologi Informasi Indonesia*, 13–20.
- [2] Ariska, E. (2018). Implementasi Agglomerative Hierarchical Clustering pada Data Produksi dan Data Penjualan Perusahaan. Universitas Sumatera Utara.
- [3] Kadir, A. (2014). Pengenalan Sistem Informasi. Yogyakarta: Andi Offset.
- [4] Kurniawan, A. A. (2017). Implementasi Algoritma Agglomerative Hierarchical Clustering Untuk Mengelompokkan Capaian Belajar Siswa SD. Universitas Sanata Dharma.
- [5] Shapiro, G. P. (2017). Learn Data Mining. Data Flair. <https://data-flair.training/blogs>. Diakses pada 15 September 2020.
- [6] Widyawati, W., Saptomo, W. L. Y., & Utami, Y. R. W. (2020). Penerapan Agglomerative Hierarchical Clustering Untuk Segmentasi Pelanggan. *Jurnal Ilmiah SINUS*. <https://doi.org/10.30646/sinus.v18i1.448>.
- [7] Zahrotun, L. (2015). Analisa Pengelompokan Jumlah Penumpang Bus Trans Jogja Menggunakan Metode Clustering K-Means dan Agglomerative Hierarchical Clustering (AHC). *Jurnal Informatika*, 9(1), 1039–1047.