

Grouping community reading interests using the k-means clustering method (case study: Magelang district library and archive service)

Achmat Mujafar^{1*}, Mukhtar Hanafi², Maimunah³

^{1,3} Teknik Informatika, Fakultas Teknik, Universitas Muhammadiyah Magelang, Indonesia

² Teknologi Informasi, Fakultas Teknik, Universitas Muhammadiyah Magelang, Indonesia

*email: jafarlesprivate@gmail.com

DOI: <https://doi.org/10.31603/binr.6810>

Abstract

So far, the problem of gathering information at the Library and Archives Office of Magelang Regency is relatively low. To increase reading interest, policies are needed to determine reading interest. So the data used is book transaction data. This study aimed to classify people's reading interests according to the number of borrowed books at the Magelang Regency Library and Archives Service using the K-Means Clustering method and to find out which book categories are most in demand by the public at the Magelang Regency Library and Archives Service. One way to manage this data is to use data mining using the K-Means method. The results of this study are low reading interest, evidenced by using 2 clusters and the category of books that are most in-demand Literature with a high cluster strength value, namely with a Silhouette Coefficient value of 0.7354.

Keywords: Reading Interest; Data Mining; K-Means; Clustering.

Abstrak

Selama ini permasalahan penggalian informasi di Dinas Perpustakaan dan Kearsipan Kabupaten Magelang masih tergolong rendah. Untuk meningkatkan minat baca dilakukan kebijakan yang diperlukan untuk menentukan minat baca tersebut. Tujuan penelitian ini untuk mengelompokkan minat baca masyarakat sesuai jumlah peminjaman buku di Dinas Perpustakaan dan Kearsipan Kabupaten Magelang menggunakan metode *K-Means Clustering* dan mengetahui kategori buku yang paling banyak diminati oleh masyarakat di Dinas Perpustakaan dan Kearsipan Kabupaten Magelang. Salah satu cara untuk mengelola data tersebut yaitu menggunakan data mining dengan memanfaatkan metode *K-Means*. Hasil dari penelitian ini adalah minat baca rendah, hal ini dibuktikan dengan menggunakan 2 *cluster* dan dengan kategori buku yang paling banyak diminati Kesusasteraan dengan nilai kekuatan *cluster* tinggi yaitu dengan nilai *Silhouette Coefficient* sebesar 0.7354.

Kata Kunci: Minat Baca; Data Mining; K-Means; Klastering.



This work is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/).

1. Pendahuluan

Perkembangan teknologi informasi saat ini menjadikan suatu informasi sebagai elemen yang penting dalam perkembangan masyarakat. Penyajian informasi tidak sepadan dengan kebutuhan informasi yang sangat tinggi, sehingga informasi tersebut perlu digali lebih dalam dari data yang jumlahnya besar. Penggalan suatu informasi atau pola yang penting atau menarik dari data dalam jumlah besar digunakan para pengambil keputusan dalam memanfaatkan gudang data. Proses penggalan ini menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai basis data besar disebut juga sebagai *data mining* ([Aronson et al., 2005](#)).

Selama ini permasalahan penggalan informasi di Dinas Perpustakaan dan Kearsipan Kabupaten Magelang masih tergolong rendah. Hal ini dibuktikan dengan banyaknya kategori buku yang berada di rak-rak buku perpustakaan. Untuk meningkatkan minat baca dilakukan kebijakan yang diperlukan untuk menentukan minat baca tersebut. Maka petugas perpustakaan bisa mengetahui kategori buku yang menjadi minat baca masyarakat. Sehingga data yang digunakan adalah data transaksi buku. Dalam hal ini data transaksi buku dapat dijadikan sebagai basis data untuk dasar kebijakan pengelolaan data perpustakaan yang tepat dan akurat. Sehingga dengan adanya *data mining* dengan menggunakan data *cluster Algoritma K-Means* memudahkan petugas perpustakaan dalam menentukan minat baca masyarakat atau pengunjung perpustakaan tersebut.

Penelitian yang dilakukan Widya Safira Aziz dan Deni Atmajaya dengan judul “Pengelompokan Minat Baca Mahasiswa Menggunakan Metode K-Means”, menyatakan bahwa permasalahan yang dihadapi yaitu penerapan sistem yang diterapkan di Perpustakaan Utsman Bin Affan Universitas Muslim Indonesia yang berjalan kurang efektif. Sehingga cara yang digunakan untuk mengelola data tersebut secara komputerisasi. Salah satu cara mengelola data tersebut yaitu dengan memanfaatkan metode *Algoritma K-means*, setiap data buku akan dikelompokkan menjadi 3 cluster, yaitu prioritas tinggi, dipertimbangkan dan bukan prioritas untuk perencanaan penambahan koleksi buku ([Azis & Atmajaya, 2016](#)).

Penelitian yang dilakukan Agus Perdana Windarto dengan judul “Implementation of Data Mining on Rice Imports by Major Country of Origin Using Algorithm Using K-Means Clustering Method”, menyatakan bahwa Produksi beras Indonesia memang begitu tinggi sehingga tidak dapat memenuhi kebutuhan penduduknya, akibatnya Indonesia masih harus mengimpor beras dari penghasil pangan lainnya negara. Hasilnya adalah penilaian berdasarkan indeks impor beras dengan 2 negara cluster impor tinggi yaitu Vietnam dan Thailand, 4 cluster tingkat sedang negara importir yaitu China, India, Pakistan dan 4 Lainnya serta rendah negara cluster yang diimpor yaitu USA, Taiwan, Singapore dan Myanmar ([Windarto, 2017](#)).

Penelitian yang dilakukan Yuda Irawan dengan Judul “Implementation Of Data Mining For Determining Majors Using K-Means Algorithm In Students Of SMA Negeri 1 Pangkalan Kerinci”. menyatakan bahwa permasalahan yang dihadapi yaitu pemilihan jurusan yang tidak tepat sehingga bisa merugikan mahasiswa untuk minat dan karir mereka di masa depan. Dengan jurusan tersebut diharapkan dapat memaksimalkan potensi, bakat atau bakat individu, sehingga dapat memaksimalkan

nilai akademik. Sehingga cara yang digunakan untuk menentukan jurusan dengan menerapkan teknik *data mining* diharapkan dapat membantu siswa dalam menentukan jurusan yang benar sesuai dengan kriteria yang ditetapkan. Teknik *data mining* digunakan dalam penentuan jurusan ini menggunakan *Algoritma K-Means* ([Yuda Irawan, 2019](#)).

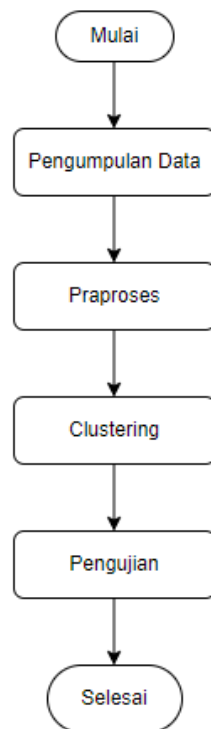
Penelitian yang dilakukan Eduardus Hardika Sandy Atmaja dengan judul “Implementation of K-Medoids Clustering Algorithm to Cluster Crime Patterns in Yogyakarta”. Menyatakan bahwa meningkatnya angka kriminalitas dari hari ke hari perlu menjadi perhatian polisi, sebagai pihak yang bertanggung jawab atas keamanan di masyarakat. Pencegahan kriminalitas usaha harus dilakukan dengan sungguh-sungguh dengan segala ilmu yang dimilikinya. Sehingga cara yang digunakan untuk mengelompokkan data yaitu ke dalam kelompok yang memiliki kemiripan data yang digunakan dalam penelitian ini adalah pola kejahatan yang kemudian dikelompokkan menggunakan *algoritma clustering K-medoids*. Hasil yang diperoleh dalam hal ini yaitu dilakukan pada tiga kelompok kejahatan yaitu tingkat kejahatan tinggi dengan 4 anggota, tingkat kejahatan sedang dengan 6 anggota dan tingkat kejahatan rendah dengan 8 anggota ([Atmaja, 2019](#)).

Penelitian yang dilakukan Essau Baker dengan judul “Implementasi Data Mining Clustering Data Perpustakaan Menggunakan Algoritma K-Means untuk Menentukan Penambahan Koleksi Buku di Perpustakaan UPY”, menyatakan bahwa permasalahan yang dihadapi yaitu kesulitan dalam menentukan jenis koleksi apa yang seharusnya perlu menjadi prioritas untuk diperbanyak atau tidak, harus menjadi pengetahuan penting bagi para pengelola perpustakaan. Pengelola perpustakaan harus memiliki pengetahuan berdasarkan data pengunjung, sehingga tidak terjadi kesalahan dalam pengadaan. Sehingga cara yang diperlukan adalah sistem yang dapat membantu menentukan sebuah keputusan yang tepat dan cepat. Sehingga akan merancang sebuah sistem implementasi *data mining clustering* data perpustakaan menggunakan *algoritma K-Means* untuk menentukan penambahan koleksi buku di perpustakaan UPY berbasis website untuk mengatasi masalah yang ada ([Bakker, 2020](#)).

Berdasarkan dari penelitian sebelumnya dapat disimpulkan bahwa penelitian terdahulu membahas tentang hal yang sejenis dalam proses pengelompokan/cluster dengan metode *Algoritma K-Means Clustering*. Pada penelitian ini digunakan metode *Algoritma K-Means Clustering*, sehingga cocok untuk pengolahan data yang ada pada Dinas Perpustakaan dan Kearsipan Kabupaten Magelang dimana menggunakan data transaksi buku. Perbedaan penelitian ini dengan penelitian relevan yaitu menggunakan jumlah kategori buku berdasarkan kategori buku dan buku yang di pinjam dalam jangka waktu setahun.

2. Metode

Tahapan penelitian seperti yang terlihat pada [Gambar 1](#) dimulai dari tahapan pengumpulan data, dilanjutkan tahapan praproses, setelah itu berlanjut pada tahapan clustering, dan pada tahapan akhir adalah pengujian.



Gambar 1. Tahapan Penelitian

a. Pengumpulan data

Pada tahap pengumpulan data adalah meminta data. Data yang di dapatkan adalah data transaksi peminjaman dan pengembalian buku dari Januari 2020 sampai Desember 2020 dari Aplikasi Inlislite di di Dinas Perpustakaan dan Kearsipan Kabupaten Magelang kumpulkan terlebih dahulu.

b. Praproses

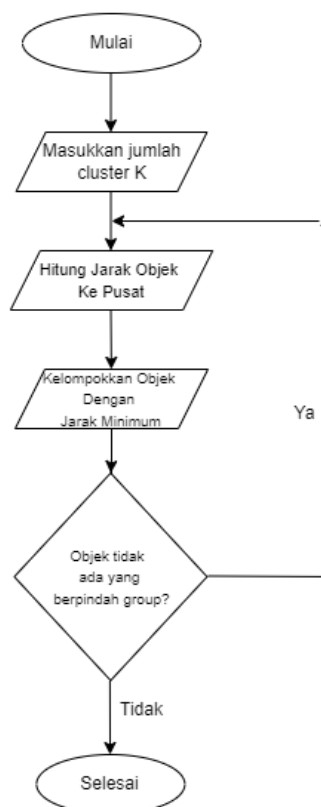
Setelah data didapat, kemudian di lakukan tahap praproses. Pada tahapan praproses ini yaitu proses *pre-processing* mencakup *data outlier*. Dimana data yang menyimpang terlalu jauh dari data yang lainnya dalam suatu rangkaian data. Setelah tahapan *data outlier* dilakukan *scaled* agar data dapat kembali seperti jenis data diawal sebelum di *array* kan. *Min-Max Scaler* bekerja dengan *scaling data* menyesuaikan data dalam rentang tertentu (rentang nilai minimum hingga nilai maksimum), dengan rentang yang biasa digunakan adalah 0 hingga 1.

c. Clustering

Pada tahap ini mengelompokkan data yang ada ke dalam beberapa kelompok. Dalam tahapan ini akan dibuat melakukan *clustering* minat baca. Untuk langkah-langkah melakukan *clustering* dengan metode Algoritma K-Means adalah sebagai berikut:

1. Menentukan atribut yang akan dibuat *cluster*. Pada tahap ini bertujuan untuk menentukan atribut sebelum dibuat *cluster* yaitu untuk melihat atribut mana saja yang masuk tahap *clustering*.
2. Memvisualkan persebaran data. Pada tahap ini bertujuan untuk untuk melihat persebaran data dari data yang divisualkan.

3. Mengubah variabel *data frame* menjadi *array*. Tujuan dari menstandarkan kembali ukuran variabel yaitu agar data dapat kembali seperti jenis data diawal sebelum diubah menjadi *array*.
4. Menentukan nilai K. Pada tahap ini bertujuan untuk mencari nilai K atau banyak nya "*n*" *cluster* didapatkan dari *elbow method*.
5. Menstandarkan ukuran variabel. Pada tahap ini bertujuan untuk menstandarkan kembali ukuran variabel yaitu agar ukuran variabel menyesuaikan data dalam rentang tertentu (range nilai minimum hingga nilai maksimum), dengan rentang yang biasa digunakan adalah 0 hingga 1.
6. Menentukan K-Means. Pada tahap ini bertujuan untuk memasukkan banyaknya nilai k pada *k-means*. Yang didapat dari menentukan nilai K menggunakan *elbow method* sehingga bertujuan mencari nilai *cluster* terbaik.
7. Menampilkan hasil *cluster*. Pada tahap ini bertujuan untuk menampilkan hasil *cluster* setelah menentukan banyaknya nilai k pada *k-means*.
8. Menambahkan kolom *cluster*. Pada tahap ini bertujuan untuk menambahkan kolom yang bernama kolom *cluster* setelah diketahui hasil *cluster* dari tahap sebelumnya.
9. Visual hasil *cluster*. Pada tahap ini bertujuan untuk mengetahui persebaran hasil *cluster* setelah dilakukan visualisasi menggunakan diagram lingkaran atau kartesian dengan nilai x dan y. Adapun tahapan pengerjaan metode *K-Means Clustering* dalam penelitian ini seperti dalam *flowchart* pada [Gambar 2](#).



Gambar 2. Algoritma *K-Means Clustering*

d. Pengujian

Pada tahap ini merupakan proses memberikan pengujian atas hasil *clustering*. Dimana pada penelitian ini menggunakan *Silhouette Coefficient* yang digunakan untuk melihat kualitas dan kekuatan *cluster*, dan seberapa baik suatu objek ditempatkan dalam suatu *cluster*.

3. Hasil dan pembahasan

Pada penelitian ini data transaksi buku terdiri dari 2065 peminjaman buku yang melakukan transaksi peminjaman yang ada di Dinas Perpustakaan Dan kearsipan Kabupaten Magelang pada tahun 2020. Data transaksi buku yang merupakan jumlah peminjaman buku terdiri dari nomor anggota yang melakukan transaksi peminjaman di Dinas Perpustakaan Dan Kearsipan Kabupaten Magelang. Data yang digunakan merupakan data yang diminta langsung dari Dinas Perpustakaan Dan Kearsipan Kabupaten Magelang. Contoh sampel data yang digunakan untuk clustering dapat dilihat pada [Gambar 3](#). Untuk contoh dari data peminjaman buku untuk clustering dapat dilihat pada output program pada [Gambar 4](#).

	Tanggal Pinjam	Tanggal Jatuh Tempo	Tanggal Dikembalikan	Jumlah Hari Telat	Judul Buku	Kategori Buku	Kode Buku	Data Bibliografis	Nomer Anggota	Nama Anggota	Nama Petugas	Nama Petugas Pengambilan	No
0	1/2/2020 0:00	8/2/2020 0:00	8/2/2020 0:00	0	SARANGHAEYO / KARUMI IVAGI	Kesusasteraan	800	981/h	1.804080e+10	A.MEYSSHERLY FERAYU W	Inisite	Inisite	0
1	1/26/20 12:00 AM	2/2/2020 0:00	2/2/2020 0:00	0	KOMIK 101 HADIS / HENDRA WIBAWA	Agama	200	50096	1.910060e+10	ACHMAD FANDI M	Inisite	Inisite	1
2	11/1/2020 0:00	1/18/20 12:00 AM	1/19/20 12:00 AM	-1	MY STUPID BOSS : SERIOUSLY BOSSMAN PASTI BERL...	Kesusasteraan	800	58991	1.910060e+10	ACHMAD FANDI M	Inisite	Inisite	2
3	11/1/2020 0:00	1/18/20 12:00 AM	1/19/20 12:00 AM	-1	STEPHANIE THE BAKER	Kesusasteraan	800	15.MTL.030	1.910060e+10	ACHMAD FANDI M	Inisite	Inisite	3
4	9/2/2020 0:00	2/16/20 12:00 AM	2/16/20 12:00 AM	0	KOMIK 25 NABI & RASUL / TRIANI RETNO	Kesusasteraan	800	59250	1.910060e+10	ACHMAD FANDI M	Layanan	Inisite	4

Gambar 3. Sampel Data Peminjaman

Kode Buku	Kategori Buku	Total Pinjam
100	Psikologi	20
200	Agama Islam	94
297	Agama	121
300	Ilmu Ilmu Sosial	44
400	Bahasa	14
500	Ilmu Ilmu Mumi	52
600	Ilmu Terapan	128
800	Kesusasteraan	1552
900	Psikologi Dan Sejarah	40

Gambar 4. Tabel kategori buku

Setelah *load* data kemudian dilakukan tahap *pre-processing* dimana yang dilakukan adalah mengecek *missing value* dan *data outlier*, pada proses ini dilakukan berbagai cara untuk memproses data sebelum data tersebut di lakukan *clustering* seperti pada [Gambar 5](#).

```
In [1423]: df.isnull().sum()

Out[1423]: Tanggal Pinjam      0
Tanggal Jatuh Tempo      0
Tanggal Dikembalikan      0
Jumlah Hari Telat      0
Judul Buku      0
Kategori Buku      0
Kode Buku      0
Data Bibliografis      0
Nomer Anggota      0
Nama Anggota      0
Nama Petugas      0
Nama Petugas Pengambilan      0
No      0
dtype: int64
```

[Gambar 5](#). Cek Missing Value

Setelah dilakukan tahap data *pre-processing* kemudian dilakukan tahap pemilihan atribut. Berdasarkan luaran program untuk pemilihan atribut nilai x terdiri dari *dataframe* yang berisi atribut nama anggota dan total pinjam. Atribut yang digunakan dalam pemilihan atribut ini adalah atribut total pinjam, seperti luaran program pada [Gambar 6](#).

Peminjaman Buku	
BRIGITA JEANNY	28
FIRDA OCTAVIA INDRADEWI	20
AHMAD DAHRI	18
CATUR NURHIDAYATI	18
JONI SUTOMO	18
...	...
ZAKIYA LABIBA	1
AGUNG SULISTİYANA HIDAYAT	1
NAFA NAILI FAUZIAH	1
SITI NUR AISYAH	1
MAULITA ASTUTI	1

[Gambar 6](#). Output Pemilihan Atribut

Setelah dilakukan tahap pemilihan atribut, kemudian dilakukan tahap *scaling* menggunakan *Min-Max Scaler*. *Min-Max Scaler* bekerja dengan *scaling* data/menyesuaikan data dalam rentang tertentu (range nilai minimum hingga nilai maksimum), dengan rentang yang biasa digunakan adalah 0 hingga 1. Untuk perintah *Min Max Scaler* bisa dilihat pada [Gambar 7](#).

```
Out[572]: array([[1.,  
 [0.7037037 ],  
 [0.62962963],  
 [0.62962963],  
 [0.62962963],  
 [0.62962963],  
 [0.55555556],  
 [0.55555556],  
 [0.55555556],  
 [0.51851852],  
 [0.51851852],  
 [0.51851852],  
 [0.51851852],  
 [0.51851852],  
 [0.48148148],  
 [0.48148148],  
 [0.48148148],  
 [0.48148148],  
 [0.48148148],  
 [0.48148148],
```

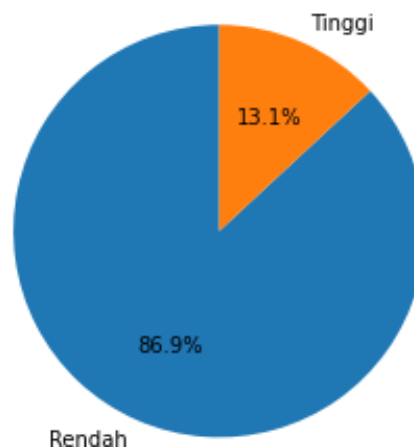
98

Dari hasil proses perhitungan *elbow method* pada Gambar untuk peminjaman buku, maka hasil yang didapat yaitu 2, Maka menggunakan cluster 2 untuk mencari *cluster* terbaik dari cluster peminjaman buku. Berdasarkan perhitungan *elbow method* diperoleh *cluster* rendah dengan 2 kali pinjaman, menunjukkan bahwa yang meminjam buku lebih dari 2 kali dengan total 444 anggota *cluster* dan bahwa *cluster* tinggi dengan 11 kali pinjaman, menunjukkan bahwa yang meminjam buku lebih dari



11 kali dengan total 67 anggota *cluster*. Selanjutnya visualisasi hasil *clustering* peminjaman buku seperti yang terlihat pada [Gambar 9](#).

Visualisasi Cluster Ketertarikan Minat



[Gambar 9](#). Visualisai Cluster Ketertarikan Minat

Pada [Gambar 9](#) menunjukkan bahwa minat baca dengan *cluster* tinggi mempunyai persentase sebesar 13.1% dengan anggota melakukan peminjaman sebanyak 67 kali dan minat baca dengan *cluster* rendah mempunyai persentase sebesar 86.9% dengan anggota melakukan peminjaman sebanyak 444 kali.

Minat baca dikatakan rendah apabila anggota yang melakukan peminjaman lebih dari 2 kali sehingga mempunyai persentase sebesar 86,9%. Sedangkan minat baca dikatakan tinggi apabila anggota yang melakukan peminjaman lebih dari 11 kali sehingga mempunyai persentase sebesar 13,1%. Untuk hasil *cluster* peminjaman buku bisa dilihat pada [Gambar 10](#).

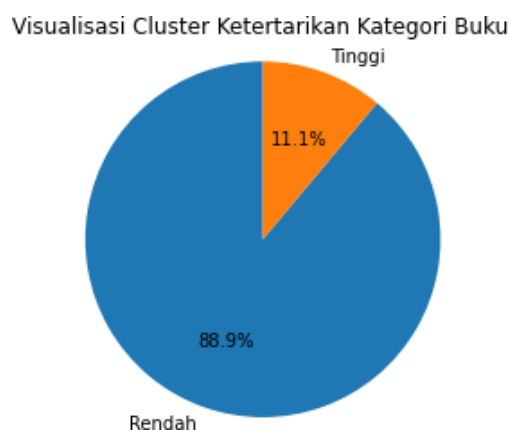
	Peminjaman Buku	cluster
BRIGITA JEANNY	28	Cluster Tinggi
FIRDA OCTAVIA INDRADEWI	20	Cluster Tinggi
JONI SUTOMO	18	Cluster Tinggi
AHMAD DAHRI	18	Cluster Tinggi
WAYA KURNIA WATI	18	Cluster Tinggi
...
FAISHOL ADISWARA SALAMAN SYAUQI	1	Cluster Rendah
EKA FATMAWATI	1	Cluster Rendah
VIDIA FADILAH ROSID	1	Cluster Rendah
ARISKA SEPTIYANI SUSANTI	1	Cluster Rendah
PUTRI DIAH ANANDA	1	Cluster Rendah

511 rows × 2 columns

[Gambar 10](#). Cluster Peminjaman Buku

Cluster rendah dengan 64 kali pinjaman, menunjukkan bahwa yang meminjam buku lebih dari 64 kali dengan jumlah *cluster* sebanyak 8 anggota kategori buku dan bahwa *cluster* tinggi dengan 1552

kali pinjaman, menunjukkan bahwa yang meminjam buku lebih dari 1552 kali dengan jumlah *cluster* sebanyak 1 anggota. Selanjutnya akan dilakukan visualisasi hasil *clustering* peminjaman buku. Untuk visualisasi diagram lingkaran bisa dilihat pada [Gambar 11](#).



[Gambar 11](#). Visualisasi *Cluster* Ketertarikan Kategori Buku

Pada [Gambar 11](#) menunjukkan bahwa minat baca dengan *cluster* tinggi mempunyai persentase sebesar 11,1% dengan anggota melakukan peminjaman sebanyak 1552 kali, minat baca dengan *cluster* rendah mempunyai persentase sebesar 88,9% dengan anggota melakukan peminjaman sebanyak 64 kali

Minat baca dikatakan rendah apabila anggota yang melakukan peminjaman lebih dari 64 kali sehingga mempunyai persentase sebesar 11,1% dengan anggota *cluster* yaitu kategori buku Agama, Agama Islam, Ilmu Ilmu Sosial, Psikologi, Bahasa, Ilmu Terapan, Ilmu Ilmu Murni dan Geografi Sejarah. Minat baca dikatakan tinggi apabila anggota yang melakukan peminjaman lebih dari 1552 kali sehingga mempunyai persentase sebesar 88,9% dengan anggota *cluster* yaitu kategori buku Kesusasteraan. Untuk hasil dari *cluster* bisa dilihat pada [Gambar 12](#).

Kode Buku	Total Pinjam	cluster
800	1552	Cluster Tinggi
600	128	Cluster Rendah
297	121	Cluster Rendah
200	94	Cluster Rendah
500	52	Cluster Rendah
300	44	Cluster Rendah
900	40	Cluster Rendah
100	20	Cluster Rendah
400	14	Cluster Rendah

[Gambar 12](#). Hasil Cluster Peminjaman Buku

Setelah dilakukan visualisasi kemudian melakukan pengujian hasil *clustering* peminjaman menggunakan *Silhouette Coefficient*, dilakukan pada hasil *clustering* setiap *distance measure* pada metode *K-Means Clustering*. Untuk hasil dari *Silhouette Coefficient* peminjaman buku bisa dilihat pada Gambar 13.

```
In [46]: #Calculating the silhouette coefficient
from sklearn.metrics import silhouette_score
from sklearn.cluster import KMeans

for n_cluster in range(2,7):
    kmeans = KMeans(n_clusters=n_cluster).fit(x)
    label = kmeans.labels_
    sil_coeff = silhouette_score(x, label, metric='euclidean')
    print("For n_clusters={}, The Silhouette Coefficient is {}".format(n_cluster, sil_coeff))

For n_clusters=2, The Silhouette Coefficient is 0.735494977375434
For n_clusters=3, The Silhouette Coefficient is 0.7021917363193213
For n_clusters=4, The Silhouette Coefficient is 0.7110847105449074
For n_clusters=5, The Silhouette Coefficient is 0.7123789163748092
For n_clusters=6, The Silhouette Coefficient is 0.7131374513988719
```

Gambar 13. Hasil Silhouette Coefficient

Pada Gambar 13 menunjukkan bahwa *cluster* dua mempunyai *Silhouette Coefficient* sebesar 0.7354 dan *cluster* tiga mempunyai *Silhouette Coefficient* sebesar 0.7021. Karena dari nilai *Silhouette Coefficient* dua dan tiga terletak diantara 0.7 sampai 1, maka nilai *Silhouette Coefficient* tersebut dapat dikatakan mempunyai struktur/kekuatan yang kuat. Namun *cluster* terbaik yang didapatkan antara *cluster* dua dan tiga adalah *cluster* dua karena mempunyai nilai *Silhouette Coefficient* sebesar 0.7354. Sehingga nilai *Silhouette Coefficient cluster* dua lebih besar nilainya daripada *cluster* tiga. Untuk hubungan minat baca dengan kategori buku, tidak ada hubungan nya, karena peminatan buku berdasarkan dari jumlah peminjaman.

4. Kesimpulan

Berdasarkan pembahasan yang sudah dijelaskan sebelumnya diperoleh kesimpulan bahwa Minat baca masyarakat di Kabupaten Magelang berdasarkan pengelompokan menggunakan *K-Means Clustering* tergolong rendah, kondisi tersebut terjadi baik ketika dikelompokkan menjadi dua atau tiga kelompok. Berdasarkan pengujian menggunakan *Silhouette Coefficient*, *clustering* dengan dua kelompok lebih kuat dari *clustering* menggunakan tiga kelompok. Berdasarkan pengelompokan kategori buku, ketertarikan minat masyarakat yaitu kategori buku kesusasteraan dengan 1552 kali pinjaman.

Referensi

Aronson, J. E., Liang, T.-P., & MacCarthy, R. v. (2005). *Decision support systems and intelligent systems* (Vol. 4). Pearson Prentice-Hall Upper Saddle River, NJ, USA:

- Atmaja, E. H. S. (2019). Implementation of k-Medoids Clustering Algorithm to Cluster Crime Patterns in Yogyakarta. *International Journal of Applied Sciences and Smart Technologies*, 1(1). <https://doi.org/10.24071/ijasst.v1i1.1859>
- Azis, W. S., & Atmajaya, D. (2016). Pengelompokan Minat Baca Mahasiswa menggunakan Metode K-Means. *ILKOM Jurnal Ilmiah*, 8(2). <https://doi.org/10.33096/ilkom.v8i2.51.89-94>
- Bakker, E. (2020). Implementasi Data Mining Clustering Data Perpustakaan Menggunakan Algoritma K-Means untuk Menentukan Penambahan Koleksi Buku di Perpustakaan UPY. *SEMINAR NASIONAL Dinamika Informatika*.
- Manulang, D., Abdillah, L. A., & Kurniawan, K. (2017). Sistem Informasi Penjualan Online (E-Commerce) Menggunakan CMS Wordpress pada Toko Soraya Shop dengan Menerapkan Metode RAD. *Seminar Hasil Penelitian Sistem Informasi Dan Teknik Informatika Ke-3 (SHaP-SITI2017)*.
- Windarto, A. P. (2017). Implementation of Data Mining on Rice Imports by Major Country of Origin Using Algorithm Using K-Means Clustering Method. *International Journal of Artificial Intelligence Research*, 1(2). <https://doi.org/10.29099/ijair.v1i2.17>
- Yuda Irawan. (2019). Implementation Of Data Mining For Determining Majors Using K-Means Algorithm In Students Of SMA Negeri 1 Pangkalan Kerinci. *Journal of Applied Engineering and Technological Science (JAETS)*, 1(1). <https://doi.org/10.37385/jaets.v1i1.18>
-