

**ANALISIS DATA SIRKULASI PERPUSTAKAAN UNIVERSITAS  
LAMPUNG MENGGUNAKAN METODE *UNSUPERVISED*  
*LEARNING* DENGAN VISUALISASI DATA**

**(Skripsi)**

**Oleh**

**REISTHA RAMADHANTY**

**NPM 1915061022**



**FAKULTAS TEKNIK  
UNIVERSITAS LAMPUNG  
BANDAR LAMPUNG  
2023**

**ANALISIS DATA SIRKULASI PERPUSTAKAAN UNIVERSITAS  
LAMPUNG MENGGUNAKAN METODE *UNSUPERVISED*  
*LEARNING* DENGAN VISUALISASI DATA**

Oleh

**REISTHA RAMADHANTY**

Skripsi

Sebagai Salah Satu Syarat untuk Mencapai Gelar  
**SARJANA TEKNIK**

Pada

**Program Studi Teknik Informatika  
Jurusan Teknik Elektro  
Fakultas Teknik Universitas Lampung**



**FAKULTAS TEKNIK  
UNIVERSITAS LAMPUNG  
BANDAR LAMPUNG  
2023**

## ABSTRAK

# ANALISIS DATA SIRKULASI PERPUSTAKAAN UNIVERSITAS LAMPUNG MENGGUNAKAN METODE *UNSUPERVISED LEARNING* DENGAN VISUALISASI DATA

Oleh

**REISTHA RAMADHANTY**

Perpustakaan Universitas Lampung merupakan Unit Pelaksana Teknis (UPT) di bidang perpustakaan yang mempunyai kewajiban untuk menyediakan dan melayani kebutuhan informasi bagi seluruh sivitas akademika di lingkungan Universitas Lampung. Saat ini kebanyakan perpustakaan memiliki *database* sirkulasi perpustakaan terkait aktivitas yang dilakukan pada perpustakaan. Data sirkulasi perpustakaan mencakup informasi tentang peminjaman, pengembalian, data pengunjung, dan masih banyak lagi. meskipun data sirkulasi perpustakaan tersedia dalam jumlah yang besar, penggunaan potensi informasi yang terkandung dalam data tersebut masih terbatas. Untuk mendapatkan informasi dari data tersebut dapat dilakukan dengan analisis data. Teknik yang dipakai pada penelitian ini adalah teknik *data mining* yaitu *K-Means Clustering*, serta melakukan visualisasi data menggunakan *Google Data Studio* yang menghasilkan *Dashboard* yang berisi grafik hasil pengolahan data. Metode penelitian yang digunakan adalah *Cross Industry Standart Process for Data Mining* atau CRISP-DM. Pengolahan data sirkulasi perpustakaan menggunakan data yang didapat dari Perpustakaan Universitas Lampung dengan jumlah data 275702. Pada penelitian ini menghasilkan 3 *cluster* pada data member perpustakaan dan data buku yaitu low level, medium level dan high level berdasarkan transaksi peminjaman, serta hasil dari pengklasteran ini divisualisasikan ke dalam dashboard visualisasi data yang bertujuan untuk membuat data lebih mudah dipahami, mengidentifikasi pola dan tren dalam data, serta membantu dalam pengambilan keputusan berdasarkan data.

Kata Kunci : *K-Means, Clustering, CRISP-DM, Dashboard, Visualisasi*

## **ABSTRACT**

# **ANALYSIS OF UNIVERSITY OF LAMPUNG LIBRARY CIRCULATION DATA USING UNSUPERVISED LEARNING METHOD WITH DATA VISUALIZATION**

**By**

**Reistha Ramadhanty**

*The Library of Lampung University is a Technical Implementation Unit (UPT) in the library sector that has the obligation to provide for and serve the information needs of all academics within the University of Lampung. Currently, most libraries have a library circulation database related to activities carried out in the library. Library circulation data includes information on loans, returns, visitor data, and much more. Although library circulation data are available in large quantities, the potential use of the information contained in these data is still limited. Data analysis can be used to get information from the data. The technique used in this study is a data mining technique, namely K-Means Clustering, as well as performing data visualization using Google Data Studio, which produces a Dashboard containing graphs of data processing results. The research method used is the Cross-Industry Standard Process for Data Mining, or CRISP-DM. The library circulation data processing uses information obtained from the University of Lampung Library, with a total of 275,702 data points. This research results in 3 clusters within the library member data and book data, namely low level, medium level, and high level, based on loan transactions. The outcomes of this clustering are then visualized in a data visualization dashboard, aiming to make the data more understandable, identify patterns and trends in the data, and assist in data-driven decision-making.*

**Keywords : K-Means, Clustering, CRISP-DM, Dashboard, Visualization**

Judul Skripsi : **ANALISIS DATA SIRKULASI PERPUSTAKAAN  
UNIVERSITAS LAMPUNG MENGGUNAKAN  
METODE UNSUPERVISED LEARNING  
DENGAN VISUALISASI DATA**

Nama Mahasiswa : **Reistha Ramadhanty**

Nomor Pokok Mahasiswa : **1915061022**

Program Studi : **Teknik Informatika**

Fakultas : **Teknik**



**MENYETUJUI**

1. Komisi Pembimbing

Pembimbing Utama

Pembimbing Pendamping

**Ir. Ing. Hery Dian Septama, S.T.**  
NIP 19850915 200812 1 001

**Paput Budi Wintoro, S.Kom., M.T.I.**  
NIP 19841031 201903 1 004

2. Mengetahui

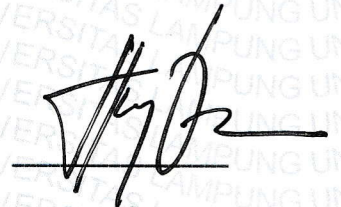
Ketua Jurusan  
Teknik Elektro

**Herlinawati, S.T., M.T.**  
NIP 19710314 199903 2 001

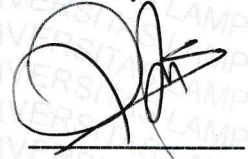
**MENGESAHKAN**

**1. Tim Penguji**

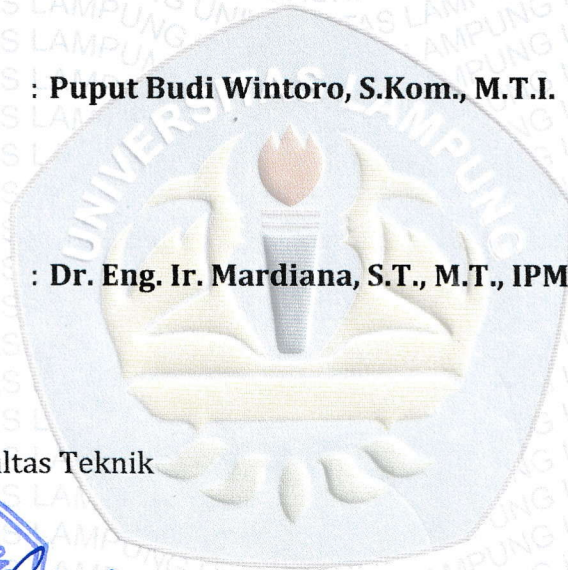
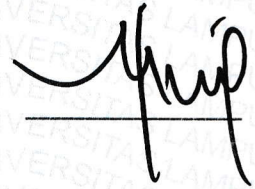
**Ketua : Ir. Ing. Hery Dian Septama, S.T.**



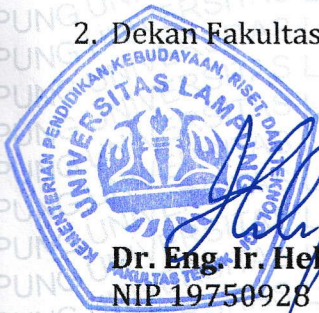
**Sekretaris : Puput Budi Wintoro, S.Kom., M.T.I.**



**Penguji : Dr. Eng. Ir. Mardiana, S.T., M.T., IPM.**



**2. Dekan Fakultas Teknik**



**Dr. Eng. Ir. Helmy Fitriawan, S.T., M.Sc. |  
NIP 19750928 200112 1 002**

**Tanggal Lulus Ujian Skripsi : 29 September 2023**

## SURAT PERNYATAAN

Saya yang bertandatangan di bawah ini, menyatakan bahwa skripsi saya dengan judul “Analisis Data sirkulasi Perpustakaan Universitas Lampung Menggunakan Metode *Unsupervised Learning* dengan Visualisasi Data” dibuat oleh saya sendiri. Semua hasil yang tertuang dalam skripsi ini telah mengikuti kaidah penulisan karya ilmiah Universitas Lampung. Apabila di kemudian hari terbukti bahwa skripsi ini merupakan salinan atau dibuat oleh orang lain, maka saya bersedia menerima sanksi sesuai dengan ketentuan hukum atau akademik yang berlaku.

BandarLampung, 29 September 2023

Pembuat pernyataan,



Reistha Ramadhanty

NPM. 1915061022

## RIWAYAT HIDUP



Penulis dilahirkan di provinsi Lampung, BandarLampung pada tanggal 10 Desember 2000. Penulis merupakan anak kedua dari tiga bersaudara dari pasangan Bapak Suryadi dan Ibu Siwi Suprapti.

Penulis menyelesaikan pendidikannya di SD Negeri 8 Gedong Air pada tahun 2013, SMP Negeri 10 BandarLampung pada tahun 2016 dan SMA Negeri 9 BandarLampung pada tahun 2019. Pada tahun 2019 penulis terdaftar sebagai mahasiswa Program Studi Teknik Informatika, Fakultas Teknik Universitas Lampung melalui jalur SBMPTN. Selama aktif menjadi mahasiswa Teknik Informatika, penulis melakukan beberapa kegiatan, antara lain:

1. Menjadi anggota biasa Himpunan Mahasiswa Teknik Elektro Universitas Lampung, Departemen Komunikasi dan Informasi, Divisi Hubungan Masyarakat periode 2019/2020;
2. Menjadi anggota biasa Himpunan Mahasiswa Teknik Elektro Universitas Lampung, Departemen Komunikasi dan Informasi, Divisi Hubungan Masyarakat periode 2020/2021;
3. Mengikuti program Magang dan Studi Independen Kampus Merdeka dari Kementerian Pendidikan dan Budaya dengan mengambil Kelas *Pengembang Machine Learning* dan *Front-End Web* di Dicoding Academy pada tahun 2021;
4. Mengikuti program Magang dan Studi Independen Kampus Merdeka dari Kementerian Pendidikan dan Budaya dengan mengambil Kelas *Data Analyst* di Generasi Gigih 2.0 X GoTo pada tahun 2022;
5. Melaksanakan Kuliah Kerja Nyata pada bulan Januari sampai dengan Februari 2022 di langkapura, Kecamatan Langkapura, Kota BandarLampung, Provinsi Lampung;



6. Melakukan Kerja Praktik di PT Gawai Mutu Digital pada bulan Agustus sampai dengan September tahun 2022 dengan membuat sistem informasi katalog Taman Kupu-kupu Gita Persada menggunakan *Wordpress*;

## **MOTTO**

“Allah tidak membebani seseorang melainkan sesuai dengan kesanggupannya...”

**(Q.S Al-Baqarah : 286)**

“jadilah baik. Sesungguhnya Allah menyukai orang-orang yang berbuat baik”

**(Q.S Al-Baqarah : 195)**

“Apapun yang menjadi takdirmu, akan mencari jalannya menemukanmu”

**(Ali bin Abi Thalib)**

“Pengetahuan yang baik adalah pengetahuan yang bermanfaat, bukan hanya diingat ”

**(Imam Syafi'i)**

“Jangan jadikan mimpimu hanya sebatas mimpi...”

**(SUGA)**

## PERSEMBAHAN

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

Segala puji syukur kepada Allah SWT. atas segala Rahmat dan karunia-Nya sehingga saya dapat menyelesaikan skripsi ini. shalawat serta salam teriring kepada Nabi Muhammad SAW sebagai suri tauladan *akhlakul karimah*.

### **KUPERSEMBAHKAN KARYA ILMIAH INI KEPADA:**

“Kedua Orang Tuaku Tercinta yang senantiasa selalu memberikan yang terbaik, mengorbankan banyak hal dan terus melantunkan doa yang tak henti-hentinya untukku. Kuucapkan terima kasih sebesar-besarnya karena telah mendidik dan membesarkanku menjadi pribadi yang baik dengan kasih sayang, kecukupan dan kebahagiaan yang akan selalu aku syukuri seumur hidupku. Semoga dengan ilmu yang kudapatkan dari hasil jerih payah kalian menyekolahkanku akan menjadi amal jariah bagi kalian”

“Kakakku Widya Larasati Putri dan adikku Icha Iriani yang berharga, terima kasih yang selalu memberikan keceriaan dan memberikan kebahagiaan dalam hidupku. Semoga kelak kita bisa menjadi saudara yang selalu bersama dan membahagiakan kedua orang tua kita”

“Diriku sendiri yang telah berjuang sampai akhir. Terima kasih diriku sudah berjuang karena, tetap bangkit dan melangkah di tengah keraguan banyak orang sungguh tidak mudah. Tapi kamu mampu. Kamu hebat. Maaf sering dengan sengaja menenggelamkan diri pada zona malas dan berujung menyia-nyiakan waktu.”

“Seluruh Keluarga Besar Teknik Elektro 2019 dan Teman-temanku. Begitu banyak cerita yang telah dilalui bersama baik itu susah, senang, tangis yang akan selalu kukenang. Semoga kita bisa menjadi orang-orang yang sukses di kemudian hari”

“Almamater tercinta, Universitas Lampung dan Jurusan Teknik Elektro”

## SANWACANA

Puji syukur penulis panjatkan kepada kehadirat Allah SWT. yang telah memberikan rahmat dan ridho-Nya sehingga penulis dapat menyelesaikan penelitian yang berjudul “Analisis Data Sirkulasi Perpustakaan Universitas Lampung Menggunakan Metode *Unsupervised Learning* dengan Visualisasi Data”. Selama masa penelitian penulis mendapatkan banyak bantuan, bimbingan dan dukungan dari berbagai pihak. Oleh karena itu penulis mengucapkan terima kasih sebanyak-banyaknya kepada:

1. Allah yang senantiasa memberikan nikmat kemudahan dan kelancaran kepada penulis serta Rasulullah Muhammad yang menjadi suri tauladan selama penelitian berlangsung;
2. Kedua orang tua tercinta dan keluarga penulis yang selalu memberikan dukungan, doa, kasih sayang tiada akhir dan mengingatkan penulis untuk bangkit dalam menyelesaikan penelitian ini;
3. Bapak Dr. Eng. Helmy Fitriawan, S.T., M.Sc., selaku Dekan Fakultas Teknik Universitas Lampung;
4. Ibu Herlinawati, S.T., M.T. selaku Ketua Jurusan Teknik Elektro Universitas Lampung;
5. Bapak Mona Arif Muda, S.T.,M.T. selaku Pembimbing Akademik yang telah membantu proses kelancaran perkuliahan;
6. Bapak Ing. Hery Dian Septama, S.T. selaku Pembimbing Utama yang selalu meluangkan waktunya untuk memberikan bimbingan tugas akhir dan memudahkan penulis dalam menyelesaikan penelitian;
7. Bapak Puput Budi Wintoro, S.Kom., M.T.I., selaku Pembimbing Pendamping yang selalu memberikan dukungan serta bimbingan agar menjadi lebih baik;

8. Ibu Dr. Eng. Ir. Mardiana, S.T., M.T.I., IPM., selaku Penguji yang telah memberikan banyak saran dan masukan terhadap penelitian ini;
9. Pihak Perpustakaan Universitas Lampung yang telah membantu penulis selama melakukan penelitian;
10. Mbak Rika selaku *Admin* Program Studi Teknik Informatika yang telah banyak membantu penulis dalam urusan administrasi selama perkuliahan dan penelitian;
11. Selvia Eldina dan Silvia Naim sebagai teman dekat penulis semasa perkuliahan hingga saat ini yang menemani, memberikan dukungan, dan motivasi.
12. Alfiah Widiyaningsih, Dwi Liliyawati, Husniatun Aini, Meilika Dwi Putri, Yovanta Anjelina dan segenap teman-teman TI 2019 yang telah menjadi rumah bagi penulis selama masa perkuliahan. Menemani, membantu dan memberikan pundak bagi penulis dikala susah dan sedih;
13. Jarina, Elpina May Kasten K., Dianti Eka Putri, Riska Bilgisa Putri, dan Elly Ardiana F.A. sebagai teman semasa SMA sampai sekarang yang selalu mendengarkan keluh kesah dari awal perkuliahan hingga di titik ini dengan memberikan semangat dan motivasi;
14. Keluarga besar Teknik Elektro Angkatan 2019 yang telah menjadi teman seperjuangan sejak mahasiswa baru. Terima kasih telah mewarnai masa perkuliahan penulis dan menulis banyak cerita bersama;

Penulis berharap agar laporan ini dapat menjadi referensi bagi pengembangan keilmuan di bidang Teknik Informatika. Oleh karena itu, semoga penelitian ini bermanfaat bagi yang membacanya.

BandarLampung, 29 September 2023  
Penulis,

Reistha Ramadhanty

## DAFTAR ISI

	Halaman
<b>DAFTAR ISI</b> .....	<b>i</b>
<b>DAFTAR GAMBAR</b> .....	<b>iii</b>
<b>DAFTAR TABEL</b> .....	<b>v</b>
<b>I. PENDAHULUAN</b> .....	<b>1</b>
1.1. Latar Belakang.....	1
1.2. Rumusan Masalah.....	3
1.3. Tujuan Penelitian .....	3
1.4. Batasan Masalah .....	4
1.5. Sistematika Penulisan Skripsi .....	4
<b>II. TINJAUAN PUSTAKA</b> .....	<b>6</b>
2.1. Perpustakaan Universitas Lampung .....	6
2.2. <i>Data Mining</i> .....	6
2.3. <i>Unsupervised Learning</i> .....	8
2.4. <i>K-Means Clustering</i> .....	8
2.5. <i>Google Colaboratory</i> .....	8
2.6. <i>Google Data Studio</i> .....	9
2.7. <i>Python</i> .....	10
2.8. <i>Dashboard dan Visualisasi Data</i> .....	10
2.9. <i>Google Sheet</i> .....	12
2.10. Analisis Data.....	13
2.11. <i>Sum of Square Error (SSE)</i> .....	13
2.12. <i>Cross Industry Standard Process Model for Data Mining (CRISP-DM)</i> 14	14
2.13. Penelitian Terkait .....	16
<b>III. METODE PENELITIAN</b> .....	<b>19</b>
3.1. Waktu dan Tempat.....	19
3.2. Alat dan Bahan .....	20

3.3.	Sumber Data.....	21
3.4.	Tahapan Penelitian Penelitian .....	21
<b>IV.</b>	<b>HASIL DAN PEMBAHASAN.....</b>	<b>27</b>
4.1.	<i>Business Understanding</i> .....	27
4.1.1.	Pernyataan Masalah .....	27
4.1.2.	Tujuan .....	28
4.1.1.	Solusi .....	28
4.2.	<i>Data Understanding</i> .....	28
4.3.	<i>Data Preparation</i> .....	31
4.3.1.	Memilih Data.....	31
4.3.2.	Mengintegrasikan Data.....	31
4.3.2.1.	Penyiapan Data .....	37
4.4.	<i>Modeling</i> .....	42
4.4.1.	<i>Scaling Data</i> .....	42
4.4.2.	<i>K-Means Clustering</i> .....	42
4.4.3.	Hasil <i>Cluster</i> Data Member .....	44
4.4.4.	Hasil <i>Cluster</i> Data Buku .....	45
4.4.5.	Visualisasi Data .....	45
4.5.	<i>Evaluation</i> .....	50
4.6.	<i>Deployment</i> .....	53
4.6.1.	Analisis Data Member .....	53
4.6.2.	Analisis Data Transaksi Peminjaman Buku .....	62
<b>V.</b>	<b>KESIMPULAN DAN SARAN.....</b>	<b>71</b>
5.1.	Kesimpulan.....	71
5.2.	Saran .....	72
	<b>DAFTAR PUSTAKA.....</b>	<b>73</b>
	<b>LAMPIRAN .....</b>	<b>77</b>

## DAFTAR GAMBAR

	Halaman
Gambar 2.1 Bidang Ilmu <i>Data Mining</i> .....	7
Gambar 2.2 <i>Google Colaboratory</i> .....	9
Gambar 2.3 <i>Google Data Studio</i> .....	10
Gambar 2.4 Contoh <i>dashboard</i> .....	11
Gambar 2.5 Google Sheet.....	12
Gambar 2.6 Alur CRISP DM [26] .....	14
Gambar 3.1 Tahapan Metode CRISP-DM.....	22
Gambar 4.1 Memanggil dataset.....	32
Gambar 4.2 Tabel dataset.....	33
Gambar 4.3 Menggabungkan df dan df2.....	33
Gambar 4.4 Menggabungkan df1_2, df3, dan df4.....	33
Gambar 4.5 Hasil Integrasi dataset Buku.....	34
Gambar 4.6 Memanggil dataset.....	34
Gambar 4.7 Tabel dataset.....	35
Gambar 4.8 Menggabungkan df1, df2, dan df3.....	35
Gambar 4.9 Hasil Integrasi Data Member .....	36
Gambar 4.10 Integrasi <i>dataset</i> perpustakaan dan <i>dataset member</i> .....	36
Gambar 4.11 Hasil Integrasi <i>dataset</i> perpustakaan dan <i>dataset member</i> .....	37
Gambar 4.12 Mengganti nilai <i>null</i> dan <i>Drop</i> kolom yang tidak digunakan .....	38
Gambar 4.13 Menghitung Frekuensi peminjaman .....	39
Gambar 4.14 Menghitung Hari Pengembalian Buku.....	39
Gambar 4.15 Menghitung jumlah member pada setiap program studi.....	39
Gambar 4.16 Memilih data yang akan diklaster.....	40
Gambar 4.17 Memilih data yang akan diklaster.....	40
Gambar 4.18 Melihat Outlier pada data.....	41
Gambar 4.19 Menghapus <i>Outlier</i> .....	41
Gambar 4.20 <i>Scaling Data</i> .....	42
Gambar 4.21 Membuat list yang berisi nilai inersia .....	42
Gambar 4.22 Plot nilai inersia .....	43
Gambar 4.23 (a) <i>K-Means Clustering</i> Data Member,.....	43
Gambar 4.24 Hasil klaster data Member (a) <i>Low Level</i> , .....	44
Gambar 4.28 hasil Klaster data Buku (a) <i>low level</i> , .....	45



Gambar 4.31 Data pada Google Sheet .....	47
Gambar 4.32 Proses penambahan sumber data pada google data studio .....	48
Gambar 4.33 Penyiapan grafik visualisasi data.....	48
Gambar 4.34 Bentuk Diagram Visualisasi Data.....	49
Gambar 4.35 Penerapan SSE.....	50
Gambar 4.36 Grafik nilai SSE.....	52
Gambar 4.37 <i>Clustering</i> Transaksi Member .....	53
Gambar 4.38 Perbandingan Jumlah keseluruhan member dengan member yang melakukan transaksi peminjaman buku .....	54
Gambar 4.39 Klaster Low Level pada data member .....	55
Gambar 4.40 Operator, Fakultas/Program Studi, Denda, Tahun Peminjaman .....	55
Gambar 4.41 Perbandingan jumlah member yang meminjam dengan data keseluruhan member .....	56
Gambar 4.42 Klaster Medium Level pada data member .....	57
Gambar 4.43 Operator, Fakultas/Program Studi, Denda, Tahun Peminjaman .....	57
Gambar 4.44 Perbandingan jumlah member yang meminjam dengan data keseluruhan member .....	58
Gambar 4.45 Klaster High Level pada data member.....	59
Gambar 4.46 Operator, Fakultas/Program Studi, Denda, Tahun Peminjaman .....	59
Gambar 4.47 Perbandingan jumlah member yang meminjam dengan data keseluruhan member .....	60
Gambar 4.48 Grafik transaksi per-tahun.....	61
Gambar 4.49 clustering data transaksi peminjaman buku .....	62
Gambar 4.50 Grafik klaster Low Level .....	63
Gambar 4.51 Tabel Judul Buku.....	63
Gambar 4.52 Banyaknya Buku Berdasarkan Tahun Terbit .....	64
Gambar 4.53 Grafik Klaster Medium Level .....	65
Gambar 4.54 Tabel Informasi Terkait Buku .....	66
Gambar 4.55 Grafik Klaster High Level.....	67
Gambar 4.56 Tabel Informasi Terkait Buku .....	68

**DAFTAR TABEL**

Tabel 3.1 Jadwal Penelitian.....	19
Tabel 3.2 Analisis Data dan Tujuan.....	23
Tabel 3.3 tabel-tabel dataset.....	24
Tabel 4.1 Fitur pada dataset data item bibliografi.csv .....	29
Tabel 4.2 Fitur pada dataset data level member.csv .....	29
Tabel 4.3 Fitur pada dataset data member perpustakaan.csv .....	29
Tabel 4.4 Fitur pada dataset data type member.csv .....	29
Tabel 4.5 Fitur pada dataset data_sirkulasi_1321.csv .....	29
Tabel 4.6 Fitur pada dataset id title berdasarkan id biblio.csv .....	30
Tabel 4.7 Fitur pada dataset id_author berdasarkan id_biblio.csv .....	30
Tabel 4.8 Fitur pada dataset nama_author berdasarkan id_author.csv .....	31
Tabel 4.9 Hasil SSE ( <i>Sum of Square Error</i> ) pada data klaster member .....	50
Tabel 4.10 Hasil SSE ( <i>Sum of Square Error</i> ) pada data klaster Buku .....	51

## **I. PENDAHULUAN**

### **1.1. Latar Belakang**

Perpustakaan Universitas Lampung dikembangkan sejak tahun 1980 sebagai pusat informasi dan dokumentasi, Perpustakaan Universitas Lampung merupakan Unit Pelaksana Teknis (UPT) di bidang perpustakaan yang memiliki kewajiban dalam penyediaan dan pelayanan kebutuhan informasi bagi seluruh sivitas akademika di lingkungan Universitas Lampung.

Perpustakaan merupakan salah satu sarana yang sangat penting dalam pendidikan dan penelitian. Sebagai tempat untuk mencari informasi pustaka memiliki peran penting dalam menyediakan akses terhadap pengetahuan dan sumber daya informasi bagi pengguna, perpustakaan yang terorganisir secara baik dan sistematis, secara langsung ataupun tidak langsung dapat memberikan kemudahan bagi proses berlangsungnya aktivitas pendidikan dan penelitian. Untuk memastikan efektivitas dan efisiensi dalam penyelenggaraan layanan perpustakaan, penting untuk melakukan analisis terhadap data sirkulasi perpustakaan[1].

Sirkulasi perpustakaan merupakan layanan tempat keluar dan masuknya koleksi perpustakaan. Sirkulasi perpustakaan memegang peran semua kegiatan yang ada di perpustakaan, dalam pelayanannya sirkulasi perpustakaan lebih sering disebut dengan layanan peminjaman dan pengembalian buku pada perpustakaan. Namun, pengertian secara global dari layanan sirkulasi adalah semua kegiatan dalam pencatatan yang berkaitan dengan kelebihan serta kekurangan koleksi buku pada perpustakaan. Layanan sirkulasi perpustakaan menjadi penghubung antara pustakawan dengan pemustaka yang ikut andil dalam menciptakan layanan yang

berkualitas sesuai dengan aturan standar perpustakaan[24]. Data sirkulasi perpustakaan mencakup informasi tentang peminjaman, pengembalian, data pengunjung, dan masih banyak lagi. Melalui analisis data sirkulasi akan diperoleh informasi – informasi yang berguna dalam pemeliharaan efektivitas layanan perpustakaan dari waktu ke waktu.

Saat ini kebanyakan perpustakaan memiliki *database* sirkulasi perpustakaan terkait aktivitas yang dilakukan pada perpustakaan, *database* yang tersedia dapat berupa *database* sederhana maupun database dalam skala besar. Dalam kumpulan data transaksi yang sangat banyak tersebut kemungkinan terdapat informasi penting yang bisa digunakan dalam mengambil keputusan[2]. Perpustakaan Universitas Lampung memiliki sistem informasi yang mengatur sirkulasi perpustakaan bernama elib (*e-library*). Pada sistem informasi tersebut dapat mengetahui terkait data sirkulasi perpustakaan berupa reporting terkait data peminjaman dan pengembalian berdasarkan *type member*, *level member*, serta operator.

Namun, untuk mengetahui pola peminjaman dan pengembalian buku yang dilakukan oleh member masih belum terdapat pada sistem informasi elib, untuk itu dilakukan analisis menggunakan data mining untuk mendapatkan informasi yang bisa digunakan dalam hal seperti pengambilan keputusan yang lebih informasi, peningkatan efektivitas layanan, perencanaan koleksi buku yang lebih baik, dan pengembangan strategi promosi yang sesuai dengan minat pengguna.

Dalam hal ini, data sirkulasi yang didapatkan sebesar 275702 (dua ratus tujuh puluh lima ribu tujuh ratus dua) diambil dari tahun 2013 hingga 2023, data sirkulasi ini terdiri dari data peminjaman dan pengembalian buku, data member, serta data buku. Untuk menganalisis data sirkulasi perpustakaan tersebut digunakan teknik data mining dan statistik, pada data mining akan dilakukan pemodelan data menggunakan algoritma *Unsupervised Learning* yaitu *K-Means Clustering* untuk memperoleh kluster-kluster dari data yang akan diolah serta memvisualisasikan data menggunakan *Google Data Studio*, dalam penelitian ini menggunakan metode *Cross Industry Standard Process Model for Data Mining* (CRISP-DM), CRISP-DM dapat membantu dalam melakukan analisis kinerja

secara lebih menyeluruh dan terstruktur dengan tahapan *Business Understanding, Data Understanding, Data Preparation, Modeling, Evaluation dan Deployment*[18]. Dengan memanfaatkan teknik-teknik ini, didapatkan informasi-informasi baru yang berguna dari data sirkulasi perpustakaan yang dapat digunakan sebagai dasar pengambilan keputusan yang lebih baik dalam pengelolaan perpustakaan.

Oleh karena itu, penelitian ini bertujuan untuk melakukan analisis data sirkulasi perpustakaan di Perpustakaan Universitas Lampung untuk memperoleh informasi dari data yang diolah dan divisualisasikan ke dalam bentuk *dashboard* agar lebih mudah dipahami. Untuk itu, penelitian ini diharapkan dapat memberikan wawasan yang berharga bagi pengelola perpustakaan dalam meningkatkan kualitas layanan, serta meningkatkan kepuasan pengguna perpustakaan.

## **1.2. Rumusan Masalah**

Adapun rumusan masalah dalam penelitian ini yaitu :

1. Bagaimana cara memperoleh informasi dari data sirkulasi perpustakaan Universitas Lampung menggunakan metode *K-Means Clustering* dan pengetahuan lainnya?
2. Bagaimana penerapan *Google Data Studio* dalam memvisualisasikan data sirkulasi perpustakaan Universitas Lampung untuk mempermudah penyampaian informasi yang dihasilkan?
3. Bagaimana analisis data sirkulasi perpustakaan Universitas Lampung dapat memberikan rekomendasi dan saran bagi pengelola perpustakaan untuk meningkatkan kualitas layanan?

## **1.3. Tujuan Penelitian**

Adapun tujuan dari penelitian ini adalah sebagai berikut :

1. Mendapatkan hasil *cluster* data sirkulasi perpustakaan Universitas Lampung dengan memanfaatkan metode *K-Means Clustering*.
2. Menganalisis data sirkulasi perpustakaan Universitas Lampung untuk memperoleh informasi yang dibutuhkan dengan memvisualisasikan data ke dalam bentuk *dashboard* menggunakan *Google Data Studio*.

3. Memberikan rekomendasi dan saran bagi pengelola perpustakaan Universitas Lampung berdasarkan hasil analisis data sirkulasi, dengan tujuan meningkatkan kualitas layanan.

#### **1.4. Batasan Masalah**

Adapun batasan masalah pada penelitian ini sebagai berikut :

1. Data yang digunakan dalam penelitian ini merupakan data sirkulasi perpustakaan yang diambil dari tahun 2013 - 2023.
2. Metode data mining yang digunakan dalam penelitian ini yaitu *K-Means Clustering*.
3. Analisis data sirkulasi perpustakaan divisualisasikan kedalam bentuk dashboard menggunakan *Google Data Studio*.
4. Analisis yang dilakukan hanya membahas data sirkulasi perpustakaan.

#### **1.5. Sistematika Penulisan Skripsi**

Laporan ini dibagi menjadi beberapa bab untuk memudahkan dalam penguraian, antara lain :

##### **BAB I : PENDAHULUAN**

Bab ini berisi latar belakang, rumusan masalah, tujuan penelitian, batasan masalah, dan sistematika penulisan skripsi.

##### **BAB II : TINJAUAN PUSTAKA**

Bab ini berisi tentang teori-teori dasar dari Penelitian terkait yang digunakan dalam Analisis Data Sirkulasi Perpustakaan Universitas Lampung Menggunakan Metode *Unsupervised Learning* dengan Visualisasi Data.

##### **BAB III : METODOLOGI PENELITIAN**

Bab ini berisi pembahasan mengenai metode penelitian digunakan dalam Analisis Data Sirkulasi Perpustakaan Universitas Lampung Menggunakan Metode *Unsupervised Learning* dengan Visualisasi Data.

##### **BAB IV : HASIL DAN PEMBAHASAN**

Bab ini berisi tentang hasil dan pembahasan yang diperoleh dalam penelitian.

**BAB V : KESIMPULAN DAN SARAN**

Bab ini memuat kesimpulan berdasarkan hasil pembahasan pada penelitian yang diharapkan dapat meningkatkan wawasan dan dijadikan sebagai bahan perbaikan untuk kedepannya.

**DAFTAR PUSTAKA :**

Bab ini berisi daftar sumber kutipan teori - teori yang dijadikan acuan penulis dalam menulis laporan.

## II. TINJAUAN PUSTAKA

### 2.1. Perpustakaan Universitas Lampung

Perpustakaan Universitas Lampung adalah Unit Pelaksana Teknis (UPT) di bidang perpustakaan yang mempunyai kewajiban untuk menyediakan dan melayani kebutuhan informasi bagi seluruh sivitas akademika di lingkungan Universitas Lampung. Perpustakaan merupakan unit kerja dari suatu badan atau lembaga tertentu yang mengelola bahan pustaka, yaitu dapat berupa buku maupun bukan berupa buku (*nonbook material*) yang diatur secara sistematis menurut aturan tertentu sehingga dapat digunakan sebagai sumber informasi oleh setiap pemakainya[5].

### 2.2. *Data Mining*

*Data mining* merupakan aktivitas yang berhubungan dengan pengumpulan data, pemakaian data historis untuk menemukan informasi, pengetahuan, keteraturan serta hubungan dalam data yang berukuran besar. *Data mining* memiliki output yang dapat digunakan sebagai alternatif pengambilan keputusan atau memperbaiki pengambilan keputusan dimasa yang akan datang. *Data mining* memiliki keterkaitan dengan bidang ilmu seperti *database*, statistik, pencarian informasi dan *artificial intelligence*[20].





Gambar 2.1 Bidang Ilmu *Data Mining*

1. *Database – Data Mining* : *Database* merupakan salahsatu sumber data yang digunakan dalam *data mining*.
2. *Statistik – Data Mining* : dalam pengambilan keputusan teknik statistik digunakan untuk menganalisa, mempresentasikan output dan penentuan sampel dalam *data mining*.
3. *Pencarian informasi – Data Mining* : salah satu kegiatan dalam *data mining* yaitu pencarian informasi yang meliputi interpretasi, analisis, dan penyimpanan data.
4. *Artificial Intelligence – Data Mining* : *Machine learning* merupakan salah satu cabang dari *Artificial Intelligence*. Dalam *data mining*, *machine learning* merupakan cabang disiplin ilmu yang penting dimana komputer belajar dari data latih yang digunakan.

Dalam *data mining* terdapat pengelompokan berdasarkan fungsi dan tujuannya sebagai berikut :

1. *Deskripsi* : proses ini bertujuan untuk menemukan atau mengidentifikasi pola yang sering muncul dan merubah pola tersebut menjadi aturan yang dapat digunakan dalam mempermudah suatu aktivitas.
2. *Klasifikasi* : merupakan proses pengelompokan berdasarkan hubungan antara variabel kriteria dengan variabel target.
3. *Prediksi* : secara umum definidi dari prediksi hamper sama dengan klasifikasi, nilai dari hasil prediksi digunakan dimasa yang akan datang berdasarkan data-data sebelumnya.

4. Estimasi : estimasi adalah perkiraan/prediksi, perbedaan dengan klasifikasi berada pada bentuk pengelompokan dimana pengelompokan kearah numerik bukan kategori.
5. Pengklasteran : merupakan pengelompokan data yang memiliki nilai(homogen). Hasil dari pengklasteran berupa hasil pengamatan, kelas-kelas, dan objek-objek yang memiliki kemiripan.
6. Asosiasi : merupakan kumpulan, himpunan, persatuan, atau persekutuan. Dalam data mining asosiasi merupakan proses pencarian atribut yang muncul dalam waktu yang bersamaan.

### ***2.3. Unsupervised Learning***

*Unsupervised Learning* merupakan salah satu teknik yang digunakan dalam *Machine Learning*. *Unsupervised learning* adalah teknik dimana mesin dilatih untuk menjawab pertanyaan yang jawabannya belum dimiliki, dalam *unsupervised learning* dataset yang dimiliki belum terdapat label atau jawaban sehingga keluaran dari hasilnya adalah sebuah label. Beberapa algoritma dari *unsupervised learning* yaitu, *DBSCAN*, *Fuzzy C-Means*, *K-Means*, dan *Self Organizing Map*[25].

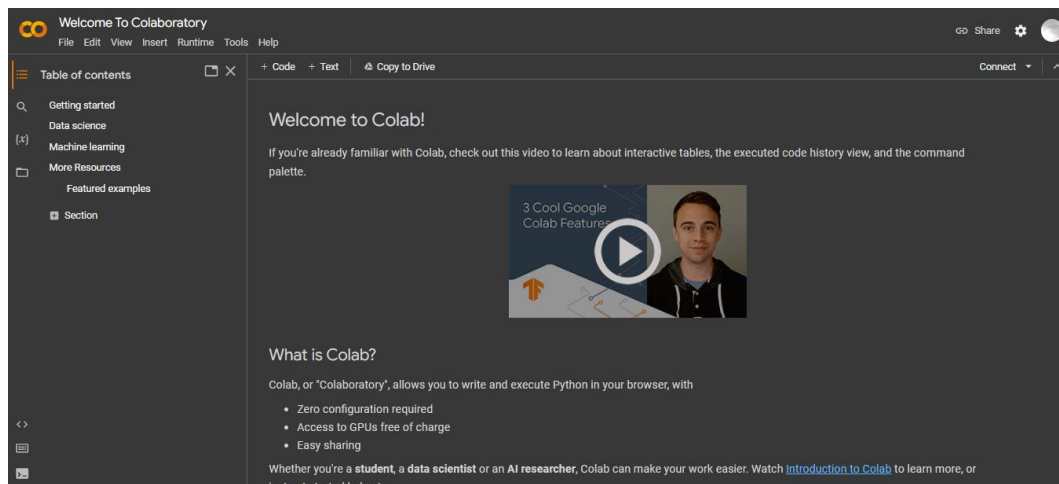
### ***2.4. K-Means Clustering***

*K-Means clustering* merupakan algoritma *nonhierarchical clustering*. Algoritma *K-Means* adalah teknik pengelompokan yang menggunakan jarak untuk mengelompokkan data menjadi beberapa *cluster*[21]. Beberapa kelebihan dari algoritma *K-Means* diantaranya adalah mudah diimplementasikan, memiliki tingkat konvergensi yang tinggi, dan menghasilkan klaster yang lebih padat jika dibandingkan dengan metode hirarki[22].

### ***2.5. Google Colaboratory***

*Google Colaboratory* adalah proyek yang memiliki tujuan menyebarluaskan pembelajaran mengenai *machine learning* dan *research*. *Google colaboratory* didasarkan pada Jupyter dan berfungsi sebagai objek *Google Documents* yaitu,

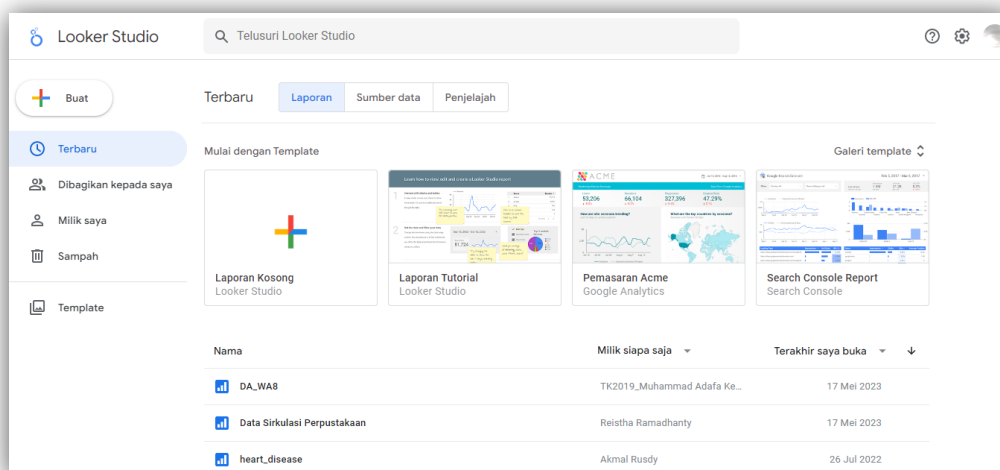
dapat dibagikan dan pengguna dapat berkolaborasi di *notebook* yang sama. *Google Colaboratory* menyediakan *runtime Python* 2 dan 3 yang telah dikonfigurasi sebelumnya dengan esensial *machine learning* dan *library* kecerdasan buatan, seperti *TensorFlow*, *Matplotlib*, dan *Keras*. *Google Colaboratory* ini menyediakan *Runtime* yang dipercepat GPU, juga sepenuhnya dikonfigurasi dengan perangkat lunak yang telah diuraikan sebelumnya. Infrastruktur *Google Colaboratory* dihosting di platform *Google Cloud*[4].



Gambar 2.2 *Google Colaboratory*

## 2.6. *Google Data Studio*

*Google Data Studio* adalah produk visualisasi data kolaboratif gratis, terintegrasi erat dengan komponen lain dari portofolio Google termasuk: *Google Analytics*, *360 Suite*, *DoubleClick Campaign Manager*, *Google AdWords*, *Google BigQuery*, *YouTube*, *Google Spreadsheet*, dan banyak lagi. Produk ini memanfaatkan *Google Cloud Storage* (GCS), akun *Google* / autentikasi dan fungsionalitas *Google Docs*, menawarkan kemampuan keamanan / autentikasi dan berbagi yang canggih. *Google Data Studio* memungkinkan pengguna terhubung ke sumber data, membuat perhitungan, *dashbord*, dan laporan khusus, berbagi dan berkolaborasi dengan yang lain. Produk ini menawarkan *template* dan galeri laporan sampel (yang dikembangkan oleh komunitas) pengguna dapat memanfaatkan alih-alih memulai dari halaman kosong. Beberapa dari mereka secara visual cukup menarik[7].



Gambar 2.3 Google Data Studio

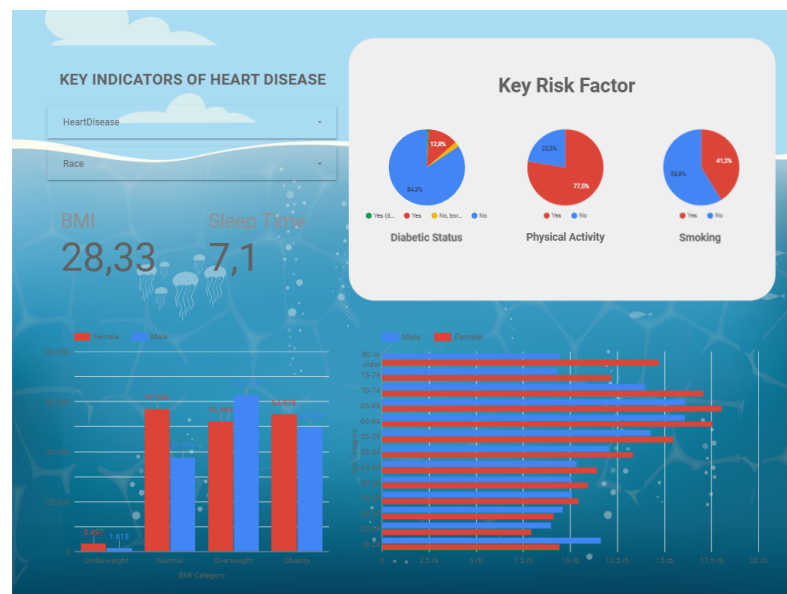
## 2.7. Python

*Python* merupakan bahasa pemrograman interpretatif yang dianggap mudah untuk dipelajari dan berfokus pada keterbacaan kode. *Python* adalah bahasa pemrograman yang dapat digunakan pada pembuatan aplikasi berbasis kecerdasan buatan (*artificial intelligence*) [19]. *Python* adalah bahasa pemrograman yang bersifat open source. Bahasa pemrograman ini dioptimalisasikan untuk *software quality*, *developer productivity*, *program portability*, dan *component integration*. *Python* telah digunakan untuk mengembangkan berbagai macam perangkat lunak, seperti *internet scripting*, *systems programming*, *user interfaces*, *product customization*, *numeric programming* dan lain-lain. *Python* saat ini telah menduduki posisi 4 atau 5 bahasa pemrograman paling sering digunakan di seluruh dunia [6].

## 2.8. Dashboard dan Visualisasi Data

*Dashboard* merupakan hasil visualisasi data yang representatif. *Dashboard* adalah sebuah tampilan visual dari informasi terpenting yang dibutuhkan untuk mencapai satu atau lebih tujuan, digabungkan dan diatur pada sebuah layar, menjadi informasi yang dibutuhkan dan dapat dilihat secara sekilas, sehingga tidak

membutuhkan waktu lama dalam memahami informasi yang ditampilkan[5]. Visualisasi data adalah teknik untuk mempresentasikan data dalam bentuk grafik atau gambar agar dapat memberikan gambaran yang lebih jelas tentang informasi yang terdapat dalam data. Visualisasi data bertujuan untuk membuat data lebih mudah dipahami, mengidentifikasi pola dan tren dalam data, serta membantu dalam pengambilan keputusan yang berbasis data[15].

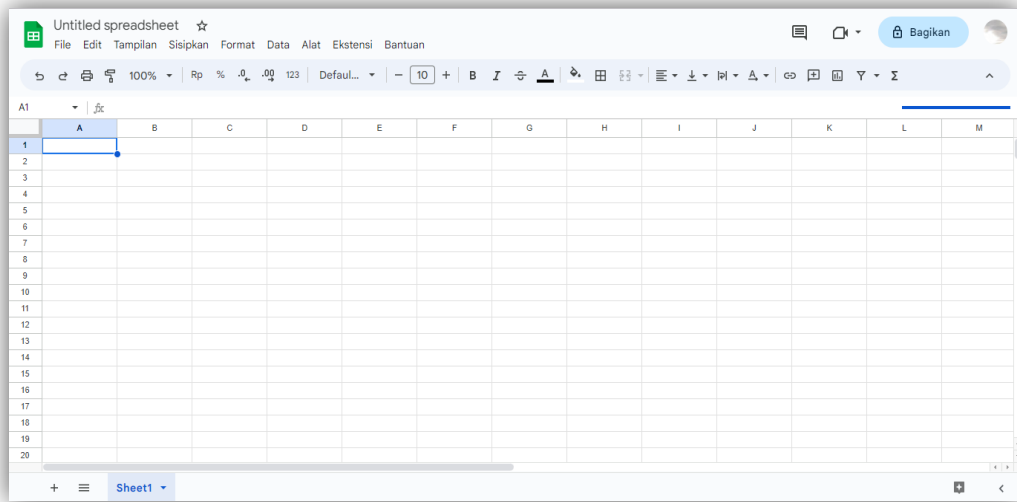


Gambar 2.4 Contoh *dashboard*

Visualisasi data merupakan proses yang mengubah data *numeric* menjadi bentuk grafis atau visual, seperti diagram, grafik, atau peta. Pemanfaatan visualisasi data bertujuan untuk memudahkan pemahaman dan ekspresi informasi serta pola yang tersembunyi dalam data. Dengan adanya visualisasi data dapat ditemukan sebuah tren, perbandingan, pola, serta korelasi antar variabel dalam data secara lebih mudah. Visualisasi data juga dapat membantu dalam menyajikan informasi secara efektif dan mempermudah pemahaman bagi *audiens*[27].

## 2.9. Google Sheet

*Google Spreadsheets* atau yang lebih dikenal dengan *Google Sheet* merupakan aplikasi berbasis Web dengan kegunaan yaitu dapat membuat, memperbaharui dan memodifikasi *spreadsheet* dan berbagi data secara *online*. Produk *Google* ini memiliki fitur *spreadsheet* yang khas, seperti kemampuan untuk menambah, menghapus dan menyortir baris dan kolom. Aplikasi ini juga memungkinkan beberapa pengguna yang tersebar secara geografis untuk bekerja secara baersamaan dalam *spreadsheet* secara waktu nyata dan mengobrol melalui program olah pesan instan bawaan. Pengguna dapat mengunggah *spreadsheet* langsung dari komputer mereka. Di dalam *Google Sheet* juga disediakan bahasa pemrograman yang disebut GAS (*Google App Script*). Kode GAS akan dieksekusi secara *remote* di dalam *google cloud*. Pihak *google* sendiri menyatakan bahwa GAS adalah : “*Google App Script* adalah bahasa *javascript cloud scripting* yang menyediakan kemudahan otomasi tugas untuk seluruh produk *google* dan layanan pihak ketiga.” *Modern cloud-based application* seperti *google sheet* akan mereduksi sulitnya para pemula dalam mengakses aplikasi[7].



Gambar 2.5 Google Sheet

## 2.10. Analisis Data

Analisis data merupakan metode yang digunakan untuk mengetahui bagaimana menggambarkan data, hubungan data, semantik data dan batasan data yang ada pada suatu sistem informasi dengan proses mengumpulkan, membersihkan, menganalisis, dan menafsirkan data yang ada[8]. Pada analisis data terdapat beberapa tahapan yaitu, pengumpulan data, pembersihan dan pemrosesan data, eksplorasi data, melakukan visualisasi data, melakukan analisis dan penarikan kesimpulan serta pengambilan keputusan dan tindakan.

### 1. Analisis Deskriptif

Analisis deskriptif merupakan jenis penelitian yang dilakukan untuk memperoleh deskripsi atau gambaran karakteristik data melalui hasil analisis data yang bersifat apa adanya tanpa membuat kesimpulan secara umum[9].

### 2. Analisis Regresi

Analisis regresi, yaitu analisis yang dapat dilihat hubungan antara dua atau lebih variabel. Secara lebih rinci, tujuan dari analisis deskriptif adalah untuk melihat adanya hubungan antara satu atau lebih variabel independen yang mempengaruhi variabel dependen[10].

## 2.11. *Sum of Square Error (SSE)*

SSE (*Sum of Square Error*) merupakan metode yang digunakan dalam melakukan validasi *cluster* melalui jumlah kuadrat setiap anggota cluster menuju pusatnya. Semakin jauh jarak yang membentuk titik siku, maka jumlah cluster tersebut menjadi yang paling optimal. Rumus SSE adalah sebagai berikut[23] :

$$SSE = \sum_{k=1}^K \sum_{X_i \in S_k} \|X_i - C_k\|_2^2$$

Dimana :

K = jumlah kelompok yang digunakan pada algoritma *K Means*

$X_i$  = banyak nya data ke i

$C_k$  = jumlah cluster pada cluster ke nilai K.

## 2.12. Cross Industry Standard Process Model for Data Mining (CRISP-DM)

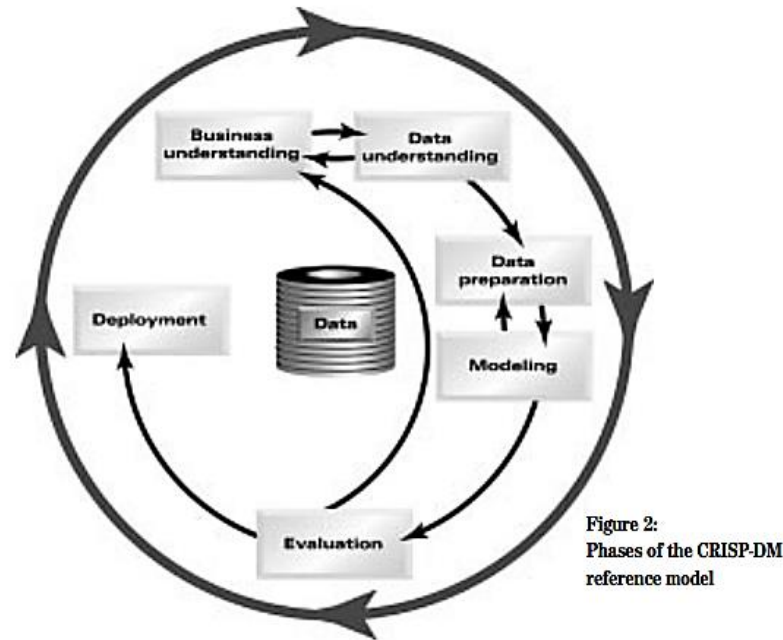


Figure 2:  
Phases of the CRISP-DM  
reference model

Gambar 2.6 Alur CRISP DM [26]

Berikut merupakan alur dari proyek *data mining* yang terdiri dari enam tahapan. Tahapan dari proses data mining menunjukkan fleksibilitas, yaitu dapat berpindah-pindah antara berbagai fase saat dibutuhkan. Hal ini tergantung pada hasil dari setiap tahap, seperti tahap mana yang harus dilakukan selanjutnya. Panah-panah mengindikasikan ketergantungan yang paling penting dan sering terjadi antara tahapan tersebut. Lingkaran pada gambar 2.6 menggambarkan sifat siklikal dari data mining itu sendiri. *Data mining* tidak berakhir begitu sebuah solusi diimplementasikan. Hal yang diperoleh dari implementasi tersebut dapat memicu pertanyaan bisnis baru, yang sering kali lebih fokus. Proses *data mining* selanjutnya akan mendapatkan manfaat dari pengalaman-pengalaman yang sudah dilakukan sebelumnya[26]. *Cross Industry Standard Process Model for Data Mining* (CRISP-DM) merupakan metode yang digunakan sebagai proses standar pengolahan data yang dapat diterapkan pada strategi pemecahan suatu masalah bisnis[16]. Pada *Cross Industry Standard Process for Data Mining* (CRISP-DM) terdapat enam tahapan yang dapat dilakukan dalam pengolahan data yaitu, sebagai berikut *business understanding*, *data understanding*, *data preparation*, *modeling*,



*evaluation* dan *deployment*[11]. Berikut penjelasan dari keenam tahapan tersebut [17]:

1. *Business Understanding*

Pada tahapan *business understanding* terdapat beberapa hal yang dilakukan antara lain, menentukan tujuan bisnis, menilai situasi, dan menentukan tujuan *data mining*.

2. *Data Understanding*

Pada tahapan *data understanding*, terdapat beberapa hal yang dilakukan antara lain, mengumpulkan data awal, mendeskripsikan data, mengeksplorasi data, dan memverifikasi kualitas data.

3. *Data Preparation*

Pada tahapan *data preparation*, ada beberapa hal yang dilakukan antara lain, deskripsi data set, memilih data, membangun data, mengintegrasikan data dan membersihkan data.

4. *Modeling*

Pada tahap ini melibatkan pengembangan dan penerapan model *data mining*, yaitu menggunakan berbagai teknik seperti regresi, klasifikasi, klastering, dan lainnya untuk membangun model yang relevan dengan tujuan bisnis yang ditentukan sebelumnya.

5. *Evaluation*

Pada tahap evaluasi dilakukan pengukuran kinerja dan validitas model. Evaluasi ini melibatkan penggunaan metrik dan teknik evaluasi yang sesuai untuk memastikan model yang dikembangkan memenuhi tujuan bisnis dan berkinerja baik.

6. *Deployment*

Pada tahap *deployment* melibatkan penerapan hasil dari proses *data mining* ke dalam lingkungan produksi. Dapat berupa integrasi model ke dalam sistem yang ada, pelatihan pengguna, dan pembuatan laporan atau presentasi.

### 2.13. Penelitian Terkait

Dalam penyusunan skripsi ini sedikit banyak terinspirasi dari penelitian – penelitian terdahulu. Adapun penelitian yang berkaitan dengan analisis data, *K-Means Clustering*, visualisasi data, *google data studio*, *google colab*, SSE (*Sum of Square Error*) dan penerapan *Cross Industry Standard Process for Data Mining* (CRISP-DM) antara lain :

1. Penelitian yang dilakukan oleh Khamim Surya Hadi Kusuma Al Atros, Abdul Robi Padri<sup>2</sup> , Odi Nurdiawan, Ahmad Faqih, Saeful Anwar yang berjudul “Model Klasifikasi Analisis Kepuasan Pengguna Perpustakaan *Online* Menggunakan *K-Means* dan *Decision Tree*” pada tahun 2021. Pada penelitian ini digunakan *K-Means* dan *Decision tree* di SD Riyadlul Muta'allimin. Hasil evaluasi dari analisis Kepuasan pengguna perpustakaan digital menggunakan *K-Means* dan *decision tree* Di SD Riyadlul Muta'allimin didapati dengan *dataset* berjumlah 150 data menghasilkan persentase akurasi sebesar 95,33%. Kluster terbaik berdasarkan hasil proses data mining adalah Kluster 2[12].
2. Penelitian yang dilakukan oleh Nofita Sari, Hanny Hikmayanti Handayani, Amril Mutoi Siregar dengan judul “Implementasi *Clustering* Data Kasus Covid 19 Di Indonesia Menggunakan Algoritma *K-Means*” pada tahun 2023. Metode yang digunakan dalam analisis data adalah *K-Means Clustering* serta metrik yang digunakan dalam evaluasi *cluster* yang dihasilkan adalah SSE (*Sum of Square Error*), evaluasi data kasus Covid-19 di Indonesia dari beberapa jumlah kasus menggunakan metode elbow yang menghasilkan 3 cluster[21].
3. Penelitian yang dilakukan oleh Winda Willina Sihombing, Himawat Aryadita, Denny Sagita Rusdianto dengan judul “Perancangan Dashboard Untuk *Monitoring* Dan Evaluasi (Studi Kasus : FILKOM UB)” pada tahun 2019. Tujuan dari penelitian ini adalah membuat rancangan *dashboard* untuk untuk *monitoring* dan evaluasi akademik mahasiswa. *Dashboard* merupakan salah satu perangkat visualisasi yang menampilkan visual dari informasi terpenting yang dibutuhkan untuk mencapai satu atau beberapa tujuan. Dashboard pada penelitian dapat digunakan membantu pimpinan program

studi pada FILKOM dalam melakukan monitoring dan evaluasi kinerja terhadap kualitas layanan yang dimiliki oleh FILKOM [13].

4. Penelitian yang dilakukan oleh Tumini, Alni Minatania dengan judul “Visualisasi Data Covid19 Tahun 2021 Di Jawa Barat Menggunakan *Google Data Studio*” pada tahun 2021. Pada penelitian ini visualisasi data dilakukan menggunakan *Google Data Studio* untuk mempermudah orang dalam memahami informasi yang disampaikan[14].
5. Penelitian yang dilakukan oleh Donny Fernando dengan judul “Visualisasi Data Menggunakan *Google Data Studio*” pada tahun 2018. Tujuan penelitian ini mengembangkan visualisasi data penjualan menggunakan *Google Data Studio*, serta terdapat penggunaan *Google Sheet* sebagai sumber data untuk visualisasi data studio[7].
6. Penelitian yang dilakukan oleh Yerik Afrianto Singgalen dengan judul “Analisis Sentimen Pengunjung Pulau Komodo dan Pulau Rinca di Website Tripadvisor Berbasis CRISP-DM” pada tahun 2023. Pada penelitian ini menggunakan metode CRISP-DM dalam melakukan pengolahan data, yaitu dalam analisis sentimen pengunjung Pulau Komodo dan Pulau Rinca di website Tripadvisor digunakan metode *Cross Industry Standard Process for Data Mining (CRISP-DM)* yang terbagi menjadi enam tahapan sebagai berikut *business understanding, data understanding, data preparation, modeling, evaluation dan deployment*. [11].
7. Penelitian yang dilakukan oleh Rangga Gelar Guntara dengan judul “Visualisasi Data Laporan Penjualan Toko *Online* Melalui Pendekatan *Data Science* Menggunakan *Google Colab*” pada tahun 2023. Pada penelitian ini membahas tentang pengolahan dan visualisasi data menggunakan *google colaboratory*. Pada penelitian ini terdapat pembahasan mengenai visualisasi data, *google colaboratory*, serta *library* yang digunakan dalam pengolahan data menggunakan bahasa pemrograman *python*[15].

Berdasarkan penelitian-penelitian terkait tersebut, didapatkan kesimpulan bahwa analisis data dilakukan untuk mendapatkan informasi yang dibutuhkan serta memudahkan dalam pengambilan keputusan. Analisis data dilakukan dengan metode *K-Means Clustering* untuk menghasilkan kluster-kluster data serta

memvisualisasikan data menggunakan beberapa *tools*, salah satunya yaitu *google data studio*, dalam pengolahan data menggunakan *google colab*. Dalam Pengolahan data metode yang sering digunakan yaitu *Cross Industry Standard Process for Data Mining* (CRISP-DM) dengan enam tahapan sebagai berikut *business understanding, data understanding, data preparation, modeling, evaluation dan deployment*.

### III. METODE PENELITIAN

#### 3.1. Waktu dan Tempat

Adapun waktu dan tempat penelitian dan pembuatan tugas akhir ini dilakukan pada :

Waktu : Februari 2023 sampai dengan Juli 2023

Tempat : Gedung Perpustakaan Universitas Lampung Jl. Soemantri Brojonegoro no. 1 Gedong Meneng Bandar Lampung, Lampung.

Tabel 3.1 Jadwal Penelitian

No.	Aktivitas	Waktu					
		Feb	Mar	Apr	Mei	Juni	Juli
1.	Studi Literatur						
2.	<i>Business Understanding</i>						
3.	<i>Data Understanding</i>						
4.	<i>Data Preparation</i>						
5.	<i>Modeling</i>						
6.	<i>Evaluation</i>						
7.	<i>Deployment</i>						
8.	Pembuatan Laporan Skripsi						

### 3.2. Alat dan Bahan

Adapun penelitian dan pembuatan tugas akhir ini menggunakan perangkat keras (*Hardware*) dan perangkat lunak (*Software*) dengan spesifikasi berikut :

#### 3.2.1. Alat

##### 1. Perangkat Keras

Satu *Personal Computer* (PC) dengan Spesifikasi :

- *Processor* : Intel Core i3-7020U CPU
- RAM : 4 Gb
- *Hardisk* : 930 Gb
- Sistem Operasi : *Windows* 10 Pro 64-bit

##### 2. Perangkat Lunak

- *Google Colaboratory*

Digunakan untuk melakukan *coding python* dalam melakukan penarikan data dari *database* untuk dilakukan analisis serta *pre-processing* data seperti *cleaning data*, menghilangkan data yang tidak digunakan, memeriksa kesalahan pada data yang lainnya.

- *Google Data Studio*

*Google Data Studio* digunakan untuk memvisualisasikan data berupa dashboard yang informatif agar data lebih mudah dipahami dan dapat melakukan analisis data untuk menghasilkan informasi yang dibutuhkan.

- *Google Sheet*

*Google Sheet* digunakan untuk menyimpan data yang sudah dilakukan proses *pre-processing* pada aplikasi *Google Colaboratory* untuk dilakukan visualisasi pada *Google Data Studio*.

#### 3.2.2. Bahan

Pada penelitian ini bahan yang digunakan diperoleh dari Perpustakaan Universitas Lampung yaitu data sirkulasi perpustakaan Universitas Lampung dari tahun 2013 hingga 2023 yang akan digunakan dalam analisis data sirkulasi perpustakaan.

### 3.3. Sumber Data

Pada penelitian ini sumber data diperoleh langsung dari Perpustakaan Universitas Lampung. Data dibagi menjadi dua, yaitu:

1. Data Primer

Data primer adalah data yang diperoleh langsung dari Perpustakaan Universitas Lampung berupa data sirkulasi perpustakaan yang terdiri dari data member, data bibliografi, serta data terkait sirkulasi perpustakaan lainnya. Data primer diperoleh dengan cara mengajukan permohonan pengambilan data pada Perpustakaan Universitas Lampung.

2. Data Sekunder

Data sekunder adalah data yang diperoleh dari berbagai macam pustaka yang dapat menunjang dan disamping sebagai alat bantu untuk memecahkan masalah yang dihadapi. Data sekunder diperoleh juga dari dokumen-dokumen, literatur, hasil penelitian terdahulu, bahan pustaka, maupun internet.

### 3.4. Tahapan Penelitian Penelitian

Adapun tahap penelitian ini menggunakan metode *Cross Industry Standard Process for Data Mining* (CRISP-DM) dengan pengolahan data menggunakan Metode *K-Means Clustering*. Berikut tahapan penelitian yang dilakukan yaitu :

#### 3.3.1. Studi Literatur

Tahap ini dilakukan untuk mempelajari ilmu dan penelitian yang terkait sebelumnya dengan sumber yang berasal dari buku, jurnal, artikel dan lain sebagainya. Ilmu yang dipelajari terkait penelitian ini yaitu, analisis data, pengolahan data menggunakan algoritma *K-Means Clustering* pada *google collaboratory*, visualisasi data menggunakan *google data studio*, dan metode *Cross Industry Standard Process for Data Mining* (CRISP-DM).

### 3.3.2. Penggunaan Metode CRISP-DM

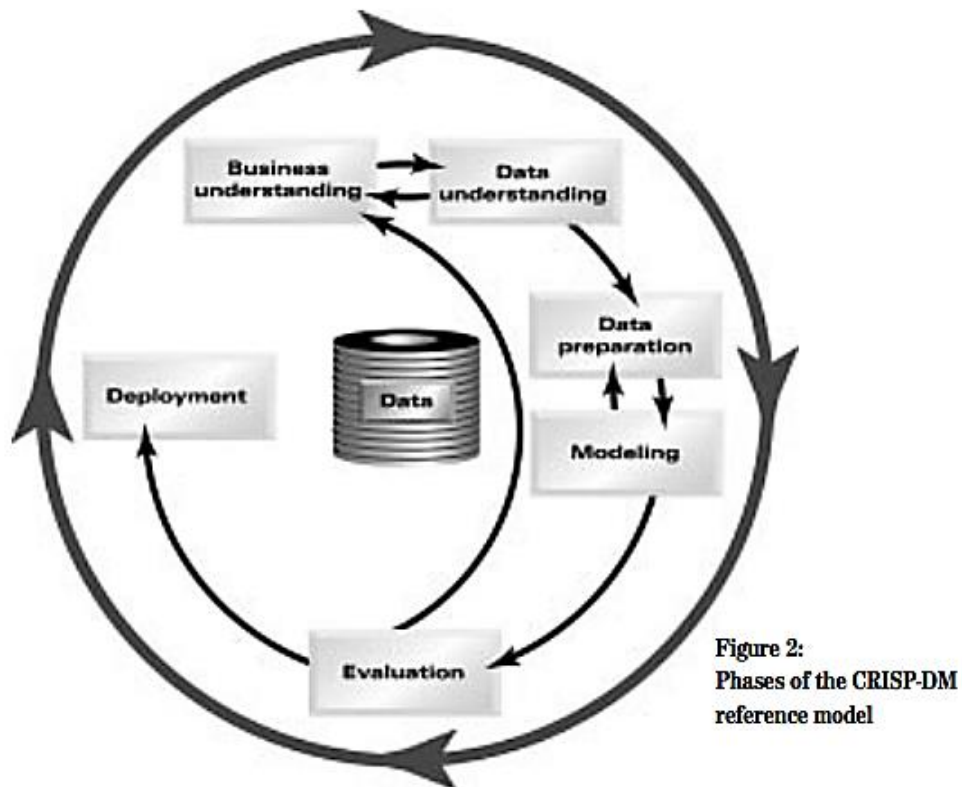


Figure 2:  
Phases of the CRISP-DM  
reference model

Gambar 3.1 Tahapan Metode CRISP-DM[26]

#### 3.3.2.1. *Business Understanding*

Tahap business understanding ini dilakukan beberapa hal seperti menentukan tujuan bisnis, menilai situasi, menentukan tujuan *data mining*, dan menghasilkan rencana *project*. Pada tahapan ini diketahui bahwa tujuan bisnis adalah memberikan *insight* serta rekomendasi strategi peningkatan layanan berdasarkan analisis data sirkulasi perpustakaan dan data lainnya yang dapat meningkatkan pelayanan yang ada pada perpustakaan. Tujuan *data mining* dalam tahapan ini adalah melakukan analisis data menggunakan algoritma *K-Means Clustering* serta visualisasi data menggunakan *google data studio* yang menghasilkan *dashboard* untuk mendapatkan informasi yang berguna bagi perpustakaan rekomendasi strategi peningkatan layanan pada perpustakaan. Berikut tabel analisis dan tujuan yang akan dilakukan :



Tabel 3.2 Analisis Data dan Tujuan

No.	Analisis Data	Tujuan
1.	Data <i>Clustering</i> Transaksi member	Melakukan analisis data sirkulasi perpustakaan yang melakukan peminjaman buku menggunakan teknik <i>data mining</i> yaitu dengan metode <i>K-Means clustering</i> , sehingga menghasilkan kluster-kluster berdasarkan peminjaman yang dilakukan oleh member.
2.	Data transaksi peminjaman buku	Melakukan analisis terhadap Data transaksi peminjaman buku dan melihat jumlah transaksinya setuap tahun.
3.	Data klasifikasi member yang meminjam buku	Melakukan analisis data persebaran member yang meminjam buku, untuk melihat member yang banyak melakukan peminjaman buku.
4.	Data sirkulasi peminjaman dan pengembalian buku	Melakukan analisis data sirkulasi peminjaman buku pada data peminjaman dan pengembalian, untuk mengetahui rata-rata waktu pengembalian buku.
5.	Data Fasilitas Peminjaman dan Pengembalian Buku	Melakukan Analisis data fasilitas peminjaman dan pengembalian buku untuk mengetahui trend peminjaman dan pengembalian setiap tahunnya.
6.	Frekuensi peminjaman buku	Melakukan analisis frekuensi peminjaman buku untuk mengetahui seberapa sering member melakukan peminjaman buku.
7.	Buku dan jenis buku	Menganalisis buku yang sering dipinjam pada perpustakaan Universitas Lampung dan jenis buku yang dimiliki.

### 3.3.2.2. *Data Understanding*

Tahap pemahaman data ini dimulai dengan melakukan pengumpulan data yang dibutuhkan, mengidentifikasi kualitas data yang digunakan, mencari informasi tersembunyi yang dapat membentuk hipotesis baru. Pemahaman data dilakukan pada dataset perpustakaan, dimana pada dataset tersebut terdapat data sirkulasi perpustakaan yang terdiri dari tabel-tabel yaitu data item bibliografi, data level member, data member perpustakaan, data type member, data\_sirkulasi\_1323, id title berdasar id biblio, id\_author berdasar id\_biblio, dan nama\_author berdasarkan id\_author. Berikut tabel dalam dataset yang akan digunakan dalam penelitian :

Tabel 3.3 tabel-tabel dataset

No.	Nama Tabel	Record	Keterangan
1.	data item bibliografi	161478 record	Tabel yang menyimpan informasi terkait item bibliografi.
2.	data level member	114 record	Tabel yang menyimpan informasi terkait data level member.
3.	data member perpustakaan	77942 record	Tabel yang menyimpan informasi terkait data member perpustakaan.
4.	data type member	27 record	Tabel yang menyimpan informasi terkait data tipe member perpustakaan.
5.	data_sirkulasi_1323	275702 record	Tabel yang menyimpan informasi terkait data sirkulasi perpustakaan.
6.	id title berdasar id biblio	29663 record	Tabel yang menyimpan informasi terkait data id judul buku berdasarkan id bibliografi.
7.	id_author berdasar	40765 record	Tabel yang menyimpan

	id_biblio		informasi terkait data id author buku berdasarkan id bibliografi.
8.	nama_author berdasarkan id_author	31037 record	Tabel yang menyimpan informasi terkait data nama author buku berdasarkan id author.

### 3.3.2.3. Data Preparation

Tahap persiapan data ini melakukan seleksi data yang digunakan, mengintegrasikan data yang telah diseleksi, membersihkan data-data yang tidak digunakan, dan memformat data agar dapat diproses dengan baik. Pada fase persiapan ini menggunakan bahasa pemrograman *Python*. Bahasa pemrograman *Python* dengan *library Pandas* digunakan untuk melakukan pembersihan data seperti menghilangkan data yang tidak diinginkan, mengolah dan menganalisa data-data terstruktur. Adapun tahapan pada data preparation yaitu :

#### 1. Memilih data

Pada tahap ini dataset yang dipilih yaitu bibliografi, data level member, data member perpustakaan, data type member, data\_sirkulasi\_1323, id title berdasar id biblio, id\_author berdasar id\_biblio, dan nama\_author berdasarkan id\_author.

#### 2. Mengintegrasikan data

Pada tahap ini dilakukan penggabungan dataset yang telah dipilih menjadi 3 *dataset* yang akan diolah, penggabungan data menggunakan bahasa pemrograman *python* pada *google colab*.

#### 3. Membersihkan Data

Pada tahap ini dilakukan pembersihan data agar saat proses modeling output yang dihasilkan maksimal dan akurat. Hal-hal yang dilakukan berupa mengganti data pada kolom yang terdapat nilai NaN (*Not a Number*) dengan data baru, serta melakukan *drop* kolom pada kolom yang tidak dibutuhkan.

#### 3.3.2.4. *Modeling*

Tahap pemodelan ini merupakan *fase* yang secara langsung melibatkan teknik data mining yaitu dengan memilih teknik data mining dan menentukan algoritma apa yang akan digunakan. Dalam *fase* ini digunakan teknik *K-Means clustering* untuk menghasilkan kluster-kluster data sirkulasi perpustakaan agar mempermudah dalam melakukan analisis data serta melakukan visualisasi data ke dalam bentuk *dashboard* menggunakan *google data studio*.

#### 3.3.2.5. *Evaluation*

Tahap evaluasi merupakan tahap lanjutan yang dilakukan untuk mengevaluasi tujuan data mining secara mendalam sehingga menghasilkan pemodelan yang sesuai dengan yang diinginkan. Beberapa hal yang dilakukan dalam *fase* ini yaitu, mengevaluasi hasil sejauh mana pemodelan yang dibuat telah memenuhi tujuan yang diinginkan, melakukan proses *review* atau pengecekan ulang dengan memastikan seluruh tahapan yang dilakukan tidak terlewatkan, dan menentukan langkah yang dilakukan selanjutnya yaitu melanjutkan ke tahap *deployment* atau kembali ke tahap awal yaitu *business understanding*. Pada tahap ini evaluasi pemodelan data menerapkan metode *Sum of Square Error (SSE)* untuk mengetahui seberapa optimal *cluster* yang dihasilkan dalam *modeling data*.

#### 3.3.2.6. *Deployment*

Pada tahap penyebaran ini, informasi dan pengetahuan yang telah didapatkan kemudian direpresentasikan menjadi bentuk yang lebih mudah dipahami oleh orang awam. Tahapan ini juga menerapkan *Google Data Studio* untuk memvisualisasikan hasil yang didapat. Dengan penerapan *Google Data Studio* akan memberikan kemudahan dalam membaca data, menyebarkan data, berkolaborasi bahkan memungkinkan dalam mengekspor data.

Hasil penelitian ini berupa visualisasi data dalam bentuk *dashboard* yang berisi grafik-grafik data sirkulasi perpustakaan dan analisis data untuk memberikan rekomendasi strategi peningkatan layanan perpustakaan.

## V. KESIMPULAN DAN SARAN

Adapun Kesimpulan dan Saran pada penelitian ini, sebagai berikut :

### 5.1. Kesimpulan

1. Berdasarkan penelitian yang telah dilakukan diperoleh dua dashboard visualisasi data yaitu, dashboard data member dan dashboard data buku perpustakaan Universitas Lampung.
2. Berdasarkan penelitian yang telah dilakukan menggunakan metode *K-Means Clustering* pada data member didapatkan tiga klaster yaitu, low level, medium level dan high level pada data member perpustakaan. Klaster Low level menandakan jumlah peminjaman buku yang dilakukan member berada ditingkat rendah, untuk medium level berada ditingkat sedang dan high level berada di tingkat tinggi.
3. Berdasarkan hasil analisis data yang telah dilakukan, yaitu pada data transaksi, data member, fasilitas yang digunakan, data buku, frekuensi peminjaman serta denda keterlambatan pengembalian dapat dimanfaatkan dalam pengambilan keputusan dan peningkatan kinerja perpustakaan, diantaranya melakukan sosialisasi pada fakultas dengan tingkat peminjaman buku terendah, meningkatkan penyediaan buku dengan minat baca terbanyak, meningkatkan fasilitas pelayanan peminjaman dan pengembalian buku secara mandiri, serta pengambilan keputusan terkait keterlambatan pengembalian buku.
4. Berdasarkan analisis yang telah dilakukan perbandingan data jumlah member keseluruhan dengan member yang melakukan peminjaman buku untuk data dari member angkatan 2017 hingga 2022 menunjukkan bahwa masih banyak member perpustakaan yang belum melakukan peminjaman buku pada perpustakaan dengan selisih antara jumlah member kelesuruhan sebesar 33,5 rb dan member yang melakukan transaksi peminjaman sebesar

30,0 rb member yaitu 3,5 rb member yang belum melakukan peminjaman pada perpustakaan.

5. Berdasarkan hasil penelitian yang telah dilakukan pada data buku menggunakan *K-Means Clustering* didapatkan tiga klaster berupa low level, medium level dan high level berdasarkan jumlah buku yang dipinjam. Low level menandakan buku yang dipinjam dalam keadaan rendah dengan rentang 0 hingga 17 kali dipinjam, pada medium level buku yang dipinjam dalam keadaan sedang dengan rentang 18 hingga 64 kali dipinjam serta, pada high level menandakan buku yang dipinjam dalam keadaan tinggi dengan rentang 65 hingga 294 kali dipinjam.
6. Berdasarkan penelitian yang telah dilakukan dengan visualisasi data dalam bentuk *dashboard* pada data buku didapatkan hasil untuk klaster low level rentang jumlah buku yang dipinjam yaitu 0-17 kali dipinjam yang menandakan buku-buku ini jarang dan bahkan belum pernah dipinjam. Sebanyak 84% data berada pada klaster ini, dan dilihat dari tahun terbitannya sebagian besar buku merupakan buku dengan tahun terbitan yang sudah lama.

## 5.2. Saran

1. Diharapkan pada penelitian berikutnya menggunakan algoritma *clustering* yang lain agar didapatkan perbandingan dan memberikan hasil yang lebih baik.
2. Data sirkulasi perpustakaan diharapkan memiliki lebih lengkap atributnya seperti jumlah pengunjung dan informasi member berupa *gender* sehingga informasi yang dihasilkan lebih maksimal.

## DAFTAR PUSTAKA

- [1] T. Firmansyah , Poningsih , S. R. Andani, “Analisis *Clustering* Algoritma *K-Means* Sebagai Rekomendasi Penambahan Koleksi Buku Di Perpustakaan Madrasah Tsanawiyah Negeri 2 Simalungun”, Jurnal Zahra : Buletin Big Data, *Data Science and Artificial Intelligence*, Vol. 1 No. 1, Juni 2022, Hal. 44-48.
- [2] Supardi, Ratnawati, DE & Mahmudy, WF, “Pengenalan pola transaksi sirkulasi buku pada database perpustakaan menggunakan algoritma *generalized sequential pattern*”, Jurnal Mahasiswa PTIIK Universitas Brawijaya, 2014, vol. 4, no. 11.
- [3] M. Erlita, J. Risanto, “Analisis Pola Peminjaman Buku Perpustakaan Universitas Riau Menggunakan Algoritma *Frequent Pattern Growth* (FP-GROWTH)”, Repository Universitas Riau, 2021
- [4] T. Carneiro, R. V. M. D. Nóbrega, T. Nepomuceno, “*Performance Analysis of Google Colaboratory as a Tool for Accelerating Deep Learning Applications*”, *Jurnal : Special Section On Trends, Perspectives And Prospects Of Machine Learning Applied To Biomedical Systems In Internet Of Medical Things*, Vol. 6, November 9, 2018, hal. 61677-61685.
- [5] M. Silvana, R. Akbar, R. Tifani, “Penerapan *Dashboard System* Di Perpustakaan Universitas Andalas Menggunakan *Tableau Public*”, Jurnal :Prosiding Semantek, November 2017, p- ISSN : 2407 – 1846 e-ISSN : 2460 – 8416.
- [6] A. F. Harismawan, A. P. Kharisma, T. Afirianto, “Analisis Perbandingan Performa Web Service Menggunakan Bahasa Pemrograman Python, PHP, dan Perl pada Client Berbasis Android”, Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer, Vol. 2, No. 1, Januari 2018, hlm. 237-24.

- [7] D. Fernando, “Visualisasi Data Menggunakan *Google Data Studio*”, Seminar Nasional Rekayasa Teknologi Informasi, Vol.8, pp.71-77, Nov. 2018.
- [8] D. Edi , S. Betshani, “Analisis Data dengan Menggunakan ERD dan Model Konseptual *Data Warehouse*”, Jurnal Informatika, Vol.5, No. 1, Juni 2009: 71 – 85.
- [9] A. M. Kusuma, P. Mahardi, “Analisis Deskriptif terhadap Pengembangan media pembelajaran E-Modul Interaktif Berbasis software aplikasi Lectora Inspire”, Jurnal Kajian Pendidikan Teknik Bangunan (JKPTB). Volume 07 Nomor 02 Tahun 2021.
- [10] Y. A. Tampil , H. Komalig, Y. Langi, “Analisis Regresi Logistik Untuk Menentukan Faktor-Faktor Yang Mempengaruhi Indeks Prestasi Kumulatif (IPK) Mahasiswa FMIPA Universitas Sam Ratulangi Manado”, JdC(*Journal of Dedicators Community*), Vol. 6, No. 2, September 2017, hal. 56-62.
- [11] Y. A. Singgalen, “Analisis Sentimen Pengunjung Pulau Komodo dan Pulau Rinca di Website Tripadvisor Berbasis CRISP-DM”, Journal of Information System Research (JOSH), Volume 4, No. 2, Januari 2023, pp 614–625, DOI 10.47065/josh.v4i2.2999
- [12] K. S. H. K. Al Atros , A. R. Padri, dkk, “Model Klasifikasi Analisis Kepuasan Pengguna Perpustakaan *Online* Menggunakan *K-Means* dan *Decision Tree*”, JURIKOM (Jurnal Riset Komputer), Vol. 8 No. 6, Desember 2021, Hal 323–329, DOI 10.30865/jurikom.v8i6.3680
- [13] W. W. Sihombing, H. Aryadita, D. S. Rusdianto, “Perancangan Dashboard Untuk Monitoring Dan Evaluasi (Studi Kasus : FILKOM UB)” Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer, Vol. 3, No. 1, Januari 2019, hlm. 434-441.
- [14] Tumini, A. Minatania, “Visualisasi Data Covid19 Tahun 2021 Di Jawa Barat Menggunakan *Google Data Studio*”, Jurnal informasi dan Komputer Vol: 11 No:1 2023, hal. 44 -51.
- [15] R. G. Guntara, “Visualisasi Data Laporan Penjualan Toko *Online* Melalui Pendekatan *Data Science* Menggunakan *Google Colab*”, ULIL



- ALBAB : Jurnal Ilmiah Multidisiplin Vol.2, No.6, Mei 2023, DOI: <https://doi.org/10.56799/jim.v2i6.1578>.
- [16] D. Astuti, A. R. Iskandar, A. Febrianti, “Penentuan Strategi Promosi Usaha Mikro Kecil Dan Menengah (UMKM) Menggunakan Metode CRISP-DM dengan Algoritma *K-Means Clustering*”, *Journal of Informatics, Information System, Software Engineering and Applications* (J. OF INISTA), VOL. 1, NO.2, PP.060-072, MAY 2019, DOI: 10.20895/INISTA.V1I2.
- [17] A. P. Fadillah, “Penerapan Metode CRISP-DM untuk Prediksi Kelulusan Studi Mahasiswa Menempuh Mata Kuliah (Studi Kasus Universitas XYZ)”, *Jurnal Teknik Informatika dan Sistem Informasi*, Volume 1 Nomor 3 Desember 2015.
- [18] Y. Suhandi, I. Kurniati, S. Norma, “Penerapan Metode Crisp-DM Dengan Algoritma *K-Means Clustering* Untuk Segmentasi Mahasiswa Berdasarkan Kualitas Akademik”, *Jurnal Teknologi Informatika dan Komputer MH Thamrin*, Volume 6 No 2; September 2020.
- [19] J. Enterprise, 2019, “*Python* untuk Programmer Pemula”, Jakarta, PT Elex Media Komputindo
- [20] E. Buulolo, 2020, “*Data Mining* Untuk Perguruan Tinggi”. Yogyakarta, Penerbit Deepublish (Grup Penerbitan CV Budi Utama)
- [21] N. Sari, H. H. Handayani, A. M. Siregar, “Implementasi *Clustering* Data Kasus Covid 19 Di Indonesia Menggunakan Algoritma K-Means”, *Jurnal Komputer dan Informatika Akademi Bina Sarana Informatika Yogyakarta*, Vol. 11 No. 1, Maret 2023
- [22] A. F. Khairati, A. A. Adlina, G. F. Hertono, B. D. Handari, “Kajian Indeks Validitas pada Algoritma *K-Means Enhanced* dan *K-Means MMCA*”, 2019, PRISMA, Prosiding Seminar Nasional Matematika 2, 161-170
- [23] D. Jollyta, S. Efendi, M. Zarlis, H. Mawengkang, “Optimasi *Cluster* Pada Data Stunting: Teknik Evaluasi *Cluster Sum of Square Error* dan *Davies Bouldin Index*”, Prosiding Seminar Nasional Riset Information Science (SENARIS), September 2019 Hal:918-926

- [24] T. J. Husnita, M. el-Khaeri Kesuma, “Pengelolaan Arsip Sebagai Sumber Informasi Bagi Suatu Organisasi Melalui Arsip Manual Dan Arsip Digital”, *Jurnal Ilmu Perpustakaan dan Informasi Islam* 01 (02): 27-41 (2020), DOI: <https://doi.org/10.24042/ei%20pustaka.v1i2.8503>
- [25] E. Retnoningsih, R. Pramudita, “Mengenal Machine Learning Dengan Teknik Supervised dan Unsupervised Learning Menggunakan Python”. *BINA INSANI ICT JOURNAL* Vol. 7, No. 2, Desember 2020, 156-165
- [26] Pete Chapman, 2000, “CRISP-DM 1.0: Step-by-step Data Mining Guide”, SPSS
- [27] I Gede I. S., Ida B., G. S., 2023, “Teknik Visualisasi Data”, Jambi, PT. Sonpedia Publishing Indonesia