

# **PROPOSAL PENELITIAN**

## **PENERAPAN ALGORITMA K-MEANS PADA KLASTERISASI PENGELUARAN KEUANGAN PT. TASPEN (PERSERO) KANTOR CABANG BOGOR**

**Oleh:**  
**Putri Rezekika Br Ginting**  
**065121117**



**PROGRAM STUDI ILMU KOMPUTER**  
**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM**  
**UNIVERSITAS PAKUAN**  
**BOGOR**  
**2025**

## KATA PENGANTAR

Puji syukur kehadiran Allah SWT, karena rahmat dan hidayah- Nya penulis dapat menyelesaikan skripsi ini yang berjudul **“Penerapan Algoritma K-Means Pada Klasterisasi Pengeluaran Keuangan PT. Taspen (Persero) Kantor Cabang Bogor ”**. Penulisan tugas akhir ini merupakan salah satu syarat memperoleh gelar Sarjana Komputer di Program Studi Ilmu Komputer FMIPA UNPAK Bogor.

Dalam penulisan tugas akhir ini, penulis dengan senang hati ingin mengucapkan terima kasih yang sebesar-besarnya kepada :

1. Dr. Hermawan Taher, selaku Pembimbing Utama yang telah memberikan dorongan moril dan motivasi kepada penulis.
2. Boldson H.S., S.Kom., MMSI., selaku pembimbing Pendamping yang telah memberikan bimbingan, semangat dan motivasi.
3. Arie Qur'ania, M.Kom., Selaku Ketua Program Studi Ilmu Komputer FMIPA Universitas Pakuan Bogor.
4. Keluarga saya yang tanpa henti memberikan perhatian, motivasi, dan dukungan dalam setiap langkah yang saya tempuh, sehingga saya dapat menyelesaikan Proposal ini dengan sebaik-baiknya.
5. Rekan-rekan serta sahabat yang telah memberikan dukungan, semangat, dan kebersamaan dalam menghadapi setiap tantangan akademik.

Penulis menyadari bahwa laporan ini masih jauh dari kesempurnaan. Oleh karena itu, penulis dengan senang hati menerima segala saran dan kritik yang membangun demi penyempurnaan tugas akhir ini. Semoga Allah SWT membalas segala kebaikan semua pihak yang telah membantu dalam penyusunan laporan ini. Akhir kata, semoga laporan ini dapat memberikan manfaat bagi kita semua. Aamiin.

Bogor, 24 Maret 2025

Putri Rezekika Br Ginting  
065121117

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Pada era digital saat ini, sistem informasi berbasis web semakin banyak diadopsi oleh berbagai instansi, termasuk lembaga pemerintahan dan perusahaan, untuk menggantikan serta mengotomatisasi proses kerja yang sebelumnya dilakukan secara manual. PT. TASPEN (Persero) merupakan Badan Usaha Milik Negara (BUMN) yang bergerak di bidang asuransi tabungan hari tua dan dana pensiun bagi Aparatur Sipil Negara (ASN) serta Pejabat Negara. Sebagai entitas yang berfokus pada layanan keuangan, PT. TASPEN memiliki tanggung jawab besar dalam memastikan pengelolaan transaksi keuangan berjalan secara akurat, efisien, dan dapat dipertanggungjawabkan.

Namun, dalam prakteknya, masih terdapat proses manual yang menghambat efisiensi, salah satunya dalam penyusunan Lembar Pertanggungjawaban Transaksi (LPT) yang menjadi tugas divisi Rumah Tangga. Penyusunan LPT ini masih dilakukan menggunakan Microsoft Excel tanpa dukungan sistem terintegrasi, yang mengakibatkan tingginya potensi kesalahan input data dan lamanya waktu penyelesaian laporan. Seiring dengan meningkatnya volume pengeluaran, permasalahan ini menjadi semakin kompleks, terlebih dengan tidak tersedianya informasi analitis yang memadai untuk membantu pimpinan dalam melakukan pengawasan serta pengendalian anggaran secara efektif.

Kondisi tersebut menunjukkan adanya kebutuhan akan sistem yang tidak hanya mampu mencatat transaksi secara digital, tetapi juga dapat mengelola data keuangan secara lebih terstruktur dan informatif. Dalam konteks ini, sistem informasi memiliki peranan penting dalam meningkatkan efisiensi dan efektivitas kerja, terutama dalam pengelolaan data keuangan organisasi. Dengan dukungan teknologi yang tepat, proses pencatatan dan pelaporan dapat dilakukan secara lebih akurat dan efisien. Penelitian oleh Nur dan Maulana (2024) menunjukkan bahwa penerapan sistem informasi berbasis web mampu menurunkan tingkat kesalahan pencatatan serta mempercepat proses verifikasi dokumen keuangan. Sistem seperti ini biasanya dikembangkan menggunakan bahasa pemrograman *Python* dengan *framework FastAPI* untuk menangani *backend* dan *API*, sementara *React* digunakan untuk antarmuka *frontend*, *CSS* untuk desain visual, serta *MySQL* sebagai basis data.

Di sisi lain, kehadiran sistem informasi saja belum cukup jika tidak dilengkapi dengan kemampuan analisis data untuk mendukung pengambilan keputusan. Dalam hal ini, teknik *data mining* seperti klusterisasi dapat digunakan untuk mengelompokkan data berdasarkan karakteristik tertentu, sehingga pola-pola tersembunyi dalam data dapat diungkap. Salah satu metode yang umum digunakan adalah *K-Means Clustering*, yang dapat mengelompokkan data ke dalam beberapa kategori berdasarkan kemiripan nilai. Metode ini efektif untuk menganalisis pengeluaran keuangan dalam berbagai skala, misalnya pengeluaran rendah, sedang, dan tinggi. Penelitian oleh Riyanto dan Akbar (2024) menunjukkan bahwa algoritma *K-Means* mampu mengelompokkan data pengeluaran pribadi secara efektif dan membantu pengguna memahami pola pengeluarannya.

Keefektifan metode K-Means juga didukung oleh berbagai penelitian lainnya. Setianingsih dan Ali (2023) menerapkannya pada data pengeluaran desa dan berhasil membentuk kluster yang valid menggunakan pendekatan KDD dengan evaluasi *Davies-Bouldin Index* sebesar 0,012. Penelitian oleh Arofah dan Hartati (2020) menunjukkan bahwa metode *K-Means* dapat memetakan distribusi anggaran pendidikan ke dalam beberapa kelompok berdasarkan nominal. Sementara itu, Anggraeni dan Kuswanto (2024) juga membuktikan bahwa algoritma ini efektif dalam menyusun data pengeluaran toko menjadi kelompok yang lebih terstruktur dan informatif. Temuan-temuan ini menunjukkan bahwa metode *K-Means* memiliki potensi besar untuk diterapkan pada berbagai konteks pengelolaan keuangan, termasuk dalam skala organisasi.

Berdasarkan permasalahan dan potensi yang telah diuraikan, dikembangkanlah sistem informasi berbasis web bernama *Laporanku* untuk mendukung pencatatan dan pengelolaan Lembar Pertanggungjawaban Transaksi (LPT) di PT. TASPEN (Persero) Kantor Cabang Bogor. Sistem ini dirancang untuk mempermudah proses input data dengan formulir yang terstruktur, memungkinkan pengguna mengunduh laporan dalam format PDF atau Excel, serta mengirimkannya untuk ditandatangani secara digital. Selain mengotomatisasi proses administratif, *Laporanku* juga mengintegrasikan algoritma *K-Means Clustering* untuk mengelompokkan data pengeluaran ke dalam kategori kecil, sedang, atau besar. Dengan demikian, sistem ini tidak hanya meningkatkan efisiensi operasional, tetapi juga memberikan nilai tambah dalam bentuk analisis data guna membantu manajemen dalam mengidentifikasi pola pengeluaran dan mengelola anggaran secara lebih optimal.

## **1.2 Tujuan**

Penelitian ini bertujuan untuk membuat sistem informasi berbasis web untuk mendukung digitalisasi proses penyusunan Lembar Pertanggungjawaban Transaksi (LPT) di PT. TASPEN (Persero) Kantor Cabang Bogor dan menerapkan algoritma *K-Means Clustering* pada data pengeluaran untuk mengelompokkan transaksi sehingga menghasilkan informasi kluster pengeluaran yang dapat dijadikan dasar dalam evaluasi dan perencanaan anggaran secara lebih terstruktur.

## **1.3 Ruang Lingkup**

Untuk memperjelas batasan dan fokus dari penelitian ini, ruang lingkup yang dikaji meliputi beberapa aspek berikut:

1. Data yang digunakan dalam penelitian ini berasal dari PT. TASPEN (Persero) Kantor Cabang Bogor, khususnya bagian Rumah Tangga yang bertugas menyusun Lembar Pertanggungjawaban Transaksi (LPT). Jumlah data yang dianalisis sebanyak 1.234 transaksi, yang dikumpulkan dalam rentang waktu Mei 2020 hingga Oktober 2023.
2. Penelitian ini berfokus pada klasterisasi data Lembar Pertanggungjawaban Transaksi (LPT) berdasarkan akun/deskripsi, tanggal, deskripsi kegiatan, realisasi (RP) dan divisi dengan cakupan wilayah terbatas pada transaksi di PT. TASPEN (Persero) Kantor Cabang Bogor. Sistem yang dikembangkan hanya mencakup digitalisasi proses Lembar Pertanggungjawaban Transaksi (LPT) , tanpa melibatkan keseluruhan sistem keuangan perusahaan.

3. Pengembangan sistem dilakukan menggunakan *Python* dengan *framework FastAPI* untuk *backend*, serta *React.js* dan *CSS* untuk *frontend*. Penyimpanan data dilakukan menggunakan MySQL sebagai basis data. Untuk proses pengolahan dan klasterisasi data, digunakan Google Collab sebagai lingkungan komputasi yang mendukung penerapan algoritma *K-Means*.
4. Sistem ini dirancang untuk digunakan oleh staf bagian Rumah Tangga PT. TASPEN (Persero) Kantor Cabang Bogor dalam menyusun, mengelola, dan mencetak Lembar Pertanggungjawaban Transaksi (LPT) secara digital.

#### **1.4 Manfaat Penelitian**

Penelitian ini dilaksanakan dengan harapan dapat memberikan kontribusi positif, baik secara teoritis maupun praktis, melalui manfaat sebagai berikut:

1. Memberikan solusi teknologi informasi berupa sistem berbasis web yang mendukung digitalisasi proses penyusunan Lembar Pertanggungjawaban Transaksi (LPT), sehingga dapat meningkatkan efisiensi, akurasi, dan akuntabilitas pencatatan LPT di PT. TASPEN (Persero) Kantor Cabang Bogor.
2. Menyediakan kemudahan penggunaan sistem bagi pengguna, khususnya staf Rumah Tangga, melalui antarmuka yang *user-friendly* serta fitur-fitur yang mendukung otomatisasi dalam proses penginputan dan pelaporan data transaksi.
3. Memberikan kontribusi akademik dan praktis berupa penerapan algoritma *K-Means Clustering* yang dapat dijadikan referensi dalam pengembangan sistem analisis data Lembar Pertanggungjawaban Transaksi (LPT) serta menjadi pertimbangan dalam pengambilan keputusan manajerial berbasis data.
4. Membantu divisi Rumah Tangga dalam menentukan Standar Biaya Umum (SBU) untuk setiap jenis transaksi berdasarkan hasil klasterisasi data pengeluaran.
5. Memungkinkan analisis pengeluaran anggaran pada masing-masing divisi, sehingga manajemen dapat menentukan alokasi anggaran operasional tahunan secara lebih tepat dan terencana.

## **BAB II**

### **TINJAUAN PUSTAKA**

#### **2.1 Landasan Teori**

##### **2.1.1 Struktur Biaya Pengeluaran Perusahaan**

Struktur biaya (*cost structure*) dalam konteks perusahaan merujuk pada komposisi dan proporsi berbagai jenis biaya yang dikeluarkan dalam operasional bisnis. Memahami struktur biaya sangat penting bagi manajemen untuk mengendalikan pengeluaran, meningkatkan efisiensi, dan memaksimalkan profitabilitas.

Menurut Tati Sulastri dan Purwanti (2024), struktur biaya terdiri dari biaya tetap dan biaya variabel, yang keduanya memiliki dampak signifikan terhadap kemampuan perusahaan dalam mencapai laba. Biaya tetap tidak berubah dengan fluktuasi volume penjualan, sementara biaya variabel berhubungan langsung dengan tingkat produksi atau penjualan. Pengelolaan yang efisien terhadap kedua jenis biaya ini, bersama dengan peningkatan volume penjualan, dapat meningkatkan profitabilitas perusahaan.

Selain itu, penelitian oleh Ahmad Nur Fuad Chalimi dan Tegar Syaifuloh (2024) menekankan pentingnya struktur organisasi, perencanaan, pelaksanaan, dan sistem pelaporan biaya dalam pengendalian biaya standar. Mereka menemukan bahwa perencanaan memiliki pengaruh dominan terhadap pengendalian biaya, menunjukkan bahwa perencanaan yang efektif dapat membantu perusahaan dalam mengendalikan biaya secara lebih efisien.

Dari pendapat tersebut, dapat disimpulkan bahwa struktur biaya pengeluaran perusahaan adalah komposisi biaya tetap dan variabel yang dikeluarkan selama operasional. Struktur ini penting dalam pengambilan keputusan manajerial karena mempengaruhi efisiensi dan profitabilitas. Biaya tetap tidak berubah meski aktivitas berfluktuasi, sedangkan biaya variabel mengikuti tingkat produksi atau penjualan. Pengelolaan yang efektif bergantung pada perencanaan, struktur organisasi, dan pelaporan biaya yang baik. Pemahaman yang tepat atas struktur biaya membantu perusahaan mengendalikan pengeluaran dan mencapai tujuan finansial secara optimal.

##### **2.1.2 Sistem Informasi**

Sistem informasi merupakan kombinasi dari teknologi informasi dan aktivitas manusia yang digunakan untuk mendukung manajemen dan pengambilan keputusan di dalam suatu organisasi. Menurut Jonny Seah (2020), sistem informasi adalah gabungan dari berbagai komponen teknologi informasi yang saling bekerja sama untuk menghasilkan informasi, sehingga dapat menciptakan jalur komunikasi dalam suatu organisasi atau kelompok. Sementara itu, Wahyudi & Ridho (2020) menyatakan bahwa sistem informasi terdiri dari sejumlah komponen yang saling terhubung dan berfungsi untuk mencapai tujuan yang diharapkan.

Menurut Edbert Kurniawan et al. (2023), sistem informasi memiliki beberapa komponen utama yang mendukung operasionalnya, yaitu:

1. Perangkat Keras (*Hardware*) adalah komponen fisik yang digunakan untuk menjalankan sistem, seperti komputer, server, dan perangkat input/output.

2. Jaringan (*Network*) adalah media penghubung antar perangkat yang memungkinkan pertukaran data secara efisien.
3. Perangkat Lunak (*Software*) adalah program atau aplikasi yang digunakan untuk mengolah dan mengelola informasi.
4. Basis Data (*Database*) adalah tempat penyimpanan data yang memungkinkan pengguna untuk mengelola, mengakses, dan memanipulasi informasi secara sistematis.
5. Protokol (*Protocol*) adalah aturan atau standar komunikasi yang mengatur pertukaran data antar perangkat dalam sistem.

Selain komponen-komponen di atas, sistem informasi juga dapat diklasifikasikan berdasarkan jenis dan fungsinya. Chrissila Jessica (2024) mengidentifikasi beberapa jenis sistem informasi yang umum digunakan dalam organisasi, yaitu:

1. *Transaction Processing System* (TPS) yaitu sistem yang digunakan untuk memproses transaksi rutin seperti penjualan dan pembelian.
2. *Management Information System* (MIS) yaitu sistem yang menyediakan laporan berkala kepada manajer berdasarkan data yang telah diproses dari TPS.
3. *Decision Support System* (DSS) yaitu sistem yang mendukung proses pengambilan keputusan melalui analisis data dan pemodelan.
4. *Executive Information System* (EIS) yaitu sistem yang menyajikan informasi strategis untuk kebutuhan eksekutif tingkat atas dalam organisasi.
5. *Workflow System* yaitu sistem yang mengatur dan mengautomasi alur kerja dalam suatu proses bisnis.
6. *Enterprise Resource Planning* (ERP) yaitu sistem terintegrasi yang mengelola seluruh sumber daya perusahaan dalam satu platform.
7. *Expert System* yaitu sistem yang menggunakan pengetahuan pakar untuk memberikan saran atau keputusan dalam bidang tertentu.

Dengan adanya sistem informasi, organisasi dapat meningkatkan efisiensi, mempercepat pengambilan keputusan, dan mendukung pengelolaan data yang lebih baik dan terstruktur.

### 2.1.3 Algoritma K-Means

Menurut Maulana et al. (2024), algoritma *K-Means* merupakan salah satu metode klusterisasi dalam analisis data yang digunakan untuk mengelompokkan data ke dalam beberapa kelompok (klaster) berdasarkan kesamaan karakteristik. Tujuan dari metode ini adalah meminimalkan jarak antar data dalam satu klaster dan memaksimalkan perbedaan antar klaster. Proses kerja algoritma ini bersifat iteratif dan sangat efektif digunakan pada data dalam jumlah besar. Secara umum, langkah-langkah dalam algoritma *K-Means* adalah sebagai berikut:

1. Menentukan jumlah klaster (*K*) yaitu pengguna menentukan berapa banyak klaster yang diinginkan sebelum proses klusterisasi dimulai.
2. Inisialisasi centroid secara acak yaitu algoritma akan memilih sejumlah titik awal (centroid) secara acak sebanyak jumlah klaster (*K*) yang telah ditentukan.
3. Menghitung jarak antara data dan centroid untuk setiap data, algoritma menghitung jarak ke masing-masing centroid yang telah ditentukan.
4. Mengelompokkan data ke klaster terdekat yaitu setiap data akan dimasukkan ke dalam cluster dengan centroid terdekat berdasarkan hasil perhitungan jarak.
5. Menghitung ulang centroid yaitu setelah semua data dikelompokkan, centroid

- baru dihitung berdasarkan rata-rata posisi data dalam masing-masing klaster.
6. Iterasi hingga konvergensi yaitu dengan langkah 3 sampai 5 diulangi hingga tidak ada lagi perubahan signifikan pada komposisi klaster, atau hingga tercapai jumlah iterasi maksimum.

#### **2.1.4 Klasterisasi**

Menurut Sugiarto dan Wibowo (2020), klasterisasi adalah metode pengelompokan data ke dalam beberapa klaster atau kelompok, di mana data dalam satu klaster memiliki tingkat kemiripan maksimum, sedangkan data antar klaster memiliki kemiripan minimum. Klasterisasi merupakan proses partisi dari suatu himpunan objek data ke dalam kelompok-kelompok yang disebut klaster. Objek-objek dalam satu klaster memiliki karakteristik yang mirip satu sama lain dan berbeda dengan objek di klaster lainnya. Proses partisi ini tidak dilakukan secara manual, melainkan menggunakan algoritma klasterisasi tertentu.

#### **2.1.5 Lembar Pertanggung Jawaban Transaksi (LPT)**

Menurut Malihah, L. (2023), lembar pertanggungjawaban bertujuan untuk menyajikan informasi yang akurat dan jelas mengenai pencapaian tujuan kegiatan serta evaluasi terhadap proses yang telah dilaksanakan. Sementara itu, menurut Muthaqin, R. (2022) Transaksi adalah sebuah kesepakatan antara penjual dan pembeli dalam menukar barang atau jasanya

Dari kedua pendapat tersebut, dapat disimpulkan bahwa Lembar Pertanggungjawaban Transaksi (LPT) merupakan laporan yang memuat rincian mengenai kegiatan, penggunaan dana, serta transaksi yang dilakukan dalam suatu proyek atau aktivitas tertentu. Laporan ini bertujuan untuk memastikan bahwa seluruh pengeluaran dan penerimaan dana tercatat secara jelas dan dapat dipertanggungjawabkan.

#### **2.1.6 Website**

Menurut Yuhefizar (2021), Website adalah metode untuk menampilkan informasi di internet dalam bentuk gambar, video, teks, suara, atau interaktif, yang menghubungkan dokumen satu dengan dokumen lainnya melalui *hyperlink* (*hypertext*) yang dapat diakses melalui browser.

#### **2.1.7 Google Collab**

Menurut Oliver (2022), Google Collab merupakan sebuah *executable document* yang memungkinkan pengguna untuk menyimpan, menulis, dan membagikan program secara langsung melalui Google Drive. Platform ini juga digunakan sebagai lingkungan eksekusi untuk *machine learning* dan pemrograman *Python*.

#### **2.1.8 Python**

*Python* merupakan bahasa pemrograman yang dikembangkan oleh Guido van Rossum dan pertama kali dirilis pada tahun 1991 di Belanda. Dalam beberapa tahun terakhir, *Python* telah menjadi salah satu bahasa pemrograman paling populer karena sintaksisnya yang sederhana dan mudah dipahami. Menurut Suharto (2023), salah satu keunggulan *Python* adalah kemudahan dalam mempelajarinya dibandingkan dengan bahasa pemrograman lainnya, sehingga sangat diminati oleh berbagai kalangan, baik pemula maupun profesional.



### 2.1.9 React JavaScript Library

Menurut Iswari (2021), *React* merupakan *library JavaScript* bersifat *open source* yang bersifat deklaratif, efisien, dan fleksibel, serta dirancang khusus untuk pengembangan antarmuka pengguna (*user interface*). Dengan *React*, pengembang dapat membangun antarmuka pengguna yang kompleks menggunakan kumpulan kode yang ringkas dan terstruktur, yang dikenal sebagai komponen. *React JS* banyak digunakan dalam pengembangan situs website maupun aplikasi berbasis website karena kemampuannya dalam membangun *user interface* yang interaktif dan responsif.

### 2.1.10 Mysql

Menurut Nu'man et al. (2020), MySQL sangat populer dalam pengembangan web karena kecepatan dan ukuran yang kecil, menjadikannya lebih ideal untuk website. MySQL merupakan layanan *database management system* yang bersifat *open source*, *multi-user*, dan *multi-threaded*, serta mampu menyimpan data relasional dengan menggunakan bahasa *SQL*. Selain itu, menurut Arief (2022), MySQL adalah jenis *database server* yang sangat terkenal dan banyak digunakan untuk membangun aplikasi website yang mengandalkan *database* sebagai sumber dan pengolahan datanya.

### 2.1.11 XAMPP

Menurut Aprilian (2020), XAMPP merupakan paket server website yang populer di kalangan pengembang karena menggunakan *PHP* dan MySQL sebagai database utama.

## 2.2 Penelitian Terdahulu

Berikut ini disajikan beberapa penelitian terdahulu yang relevan dengan topik yang diangkat dalam penelitian ini yaitu

1. Nama : Nur Adzilla Aulia dan Muhammad Yasin Simargolang

Judul : Implementasi Algoritma *K-Means Data Mining* untuk *Clustering Data Transaksi Pengeluaran Toko Sepatu*

Tahun : 2024

Isi : Penelitian ini bertujuan mengembangkan sistem berbasis algoritma *K-Means* untuk *clustering* data transaksi pengeluaran di Toko Sepatu Ldr. Secondbrand. Sistem mengolah data seperti biaya sewa gedung, gaji karyawan, air, listrik, dan pengeluaran lainnya, yang kemudian divisualisasikan dalam bentuk website agar informasi lebih terstruktur dan mudah dipahami. Tujuan utama adalah menerapkan teknik *data mining* untuk menganalisis pola pembelian pelanggan dan mendukung pengambilan keputusan bisnis. Dengan algoritma *K-Means*, data transaksi dikelompokkan ke dalam empat cluster: cluster pertama Rp72.000.000 per tahun, cluster kedua Rp6.600.000, cluster ketiga Rp2.940.000, dan cluster keempat Rp105.000.000. Hasil penelitian menunjukkan sistem ini efektif dalam mengatasi masalah clustering data pengeluaran dan memberikan informasi klasifikasi yang berguna.

2. Nama : Widiyanto, M. Rifqy Zakaria, Irvan
- Judul : Penerapan *Data Mining* untuk Klasterisasi Data Anggaran Pendapatan dan Belanja Daerah Menggunakan Algoritma *K-Means*
- Tahun : 2025
- Isi : Penelitian ini membahas penerapan data mining dengan menggunakan Algoritma K-Means untuk meng klasterisasi data anggaran pendapatan dan belanja daerah (APBD). Latar belakangnya adalah kebutuhan menyusun Rencana Pembangunan Daerah (RPD) selama transisi kepemimpinan daerah. Algoritma K-Means digunakan untuk menganalisis potensi pendapatan dan belanja daerah berdasarkan kluster yang terbentuk. Tujuan utama penelitian ini adalah mengelompokkan data anggaran pendapatan dan belanja daerah di Kabupaten Tolikara. Hasil pengelompokan diharapkan memberikan wawasan baru dalam analisis potensi pendapatan dan belanja daerah. Penelitian menunjukkan bahwa data dapat dikelompokkan menjadi dua kluster, dengan kluster 1 memiliki centroid 192973008 dan 167000000, serta kluster 2 dengan centroid 7000000 dan 225000000. Kluster 1 diidentifikasi sebagai potensi belanja tinggi, sedangkan kluster 2 sebagai potensi belanja rendah. Model K-Means yang diterapkan berhasil memberikan pengelompokan data yang informatif bagi Badan Pengelolaan Keuangan dan Aset Daerah (BPKAD) Kabupaten Tolikara.
3. Nama : Akmal Fikri, Benedika Ferdian Hutabarat dan Ulfa Khaira
- Judul : Komparasi *K-Means clustering* Dan *Complete Linkage* Dalam Pengelompokan Penyaluran Pinjaman Financial Technology
- Tahun : 2022
- Isi : Penelitian ini mengembangkan komparasi antara metode K-Means Clustering dan Complete Linkage untuk mengelompokkan penyaluran pinjaman oleh perusahaan Financial Technology (FinTech), dengan tujuan membandingkan efektivitas dan efisiensi kedua metode tersebut dalam mengklasifikasi penyaluran pinjaman. Data pinjaman dari pelanggan FinTech yang diolah oleh OJK digunakan dalam penelitian ini, dan analisis dilakukan menggunakan bahasa pemrograman R yang kemungkinan dijalankan melalui platform Google Collab. Hasil dari pengklasteran menunjukkan bahwa menggunakan Complete Linkage, jumlah cluster optimal adalah 5 dengan nilai DBI sebesar 0,346174, sementara pada K-Means Clustering, jumlah cluster optimal adalah 2 dengan nilai DBI 0,07364345. K-Means Clustering memberikan hasil yang lebih baik dalam mengelompokkan data penyaluran pinjaman FinTech.

4. Nama : Indri Setianingsih dan Irfan Ali
- Judul : Klasterisasi Pengeluaran Kas di Desa Pamengkang Menggunakan Metode *K-Means*
- Tahun : 2023
- Isi : Penelitian ini membahas klasterisasi pengeluaran kas di Desa Pamengkang menggunakan metode *K-Means*. Latar belakangnya adalah masalah pengelolaan pengeluaran kas yang sering mengalami kesalahan perhitungan dan ketidaksesuaian data, yang mengakibatkan keterlambatan laporan. Tujuan penelitian ini adalah untuk mengelompokkan data pengeluaran desa dengan metode *K-Means*. Data yang digunakan adalah 142 data pengeluaran kas tahun 2021, diolah menggunakan aplikasi RapidMiner. Hasil penelitian menunjukkan bahwa klasterisasi menggunakan *K-Means* menghasilkan nilai *Davies Bouldin Index* sebesar 0,012, dengan cluster 1 memiliki nilai 0,014, cluster 2 sebesar 0,017, dan cluster 3 sebesar 0,027. Hasil ini menunjukkan bahwa cluster 1 adalah kelompok pengeluaran paling signifikan. Penelitian ini menyimpulkan bahwa metode *K-Means* dapat memberikan wawasan dalam pengelolaan pengeluaran kas desa.
5. Nama : Agung Riyanto dan Mutaqin Akbar
- Judul : Pengelompokan Data Pengeluaran Bulanan pada Aplikasi Pencatatan Keuangan Pribadi menggunakan Metode *K-Means Clustering*
- Tahun : 2024
- Isi : Penelitian ini membahas pengelompokan data pengeluaran bulanan pada aplikasi pencatatan keuangan pribadi menggunakan metode *K-Means Clustering*. Latar belakangnya adalah pentingnya pengelolaan keuangan pribadi di era modern, di mana gaya hidup berlebihan dan kurangnya pemahaman finansial membuat aplikasi pencatatan keuangan semakin relevan. Penelitian ini bertujuan untuk mengimplementasikan pengelompokan data pengeluaran bulanan menjadi kelompok pengeluaran rendah, sedang, dan tinggi. Data yang digunakan terdiri dari 30 data training dan 25 data uji dari aplikasi Pencatatan Keuangan Pribadi (CaKeP). Proses *clustering* menghasilkan tiga kelompok dengan skor *Silhouette Score* sebesar 72%. Penentuan jumlah cluster menggunakan *Elbow Method* menunjukkan 3 sebagai jumlah cluster optimal. Hasil penelitian menunjukkan bahwa pengelompokan data pengeluaran dengan *K-Means* berhasil dilakukan dengan baik, di mana setiap sampel terklasifikasi sesuai dengan jarak yang baik terhadap klaster lainnya.

Tabel berikut menyajikan perbandingan beberapa penelitian terdahulu yang relevan dan menjadi acuan dalam penelitian ini :

**Tabel 1.** Perbandingan Penelitian Terdahulu

No	Nama Penelitian, Tahun	Platform			Metode		Instansi			
		RapidMiner	Google Collab	Web	K - Means	Complete Linkage	Toko	Pemerintah	Pribadi	Perusahaan
1.	Nur Adzilla Aulia dan Muhammad Yasin Simargolang, (2024)			✓	✓		✓			
2.	Widianto, M. Rifqy Zakaria, Irvan, (2025)			✓	✓			✓		
3.	Akmal Fikri, Benedika Ferdian Hutabarat dan Ulfa Khaira		✓		✓	✓				✓
4.	Indri Setianingsih dan Irfan Ali, (2023)	✓			✓			✓		
5.	Agung Riyanto dan Mutaqin Akbar, (2024)		✓		✓				✓	
6.	Putri Rezekika Br Ginting, (2025)		✓		✓			✓		

## BAB III

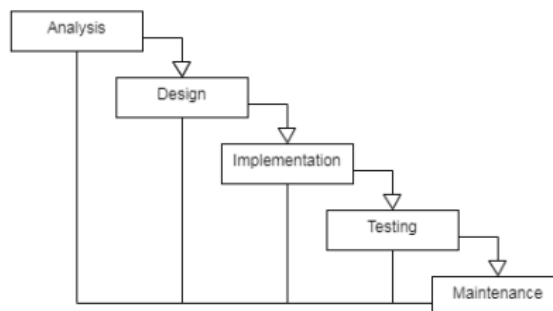
### METODE PENELITIAN

#### 3.1 Metodologi Penelitian

Penelitian ini menggunakan metode pengembangan perangkat lunak *Waterfall*, yang dikenal sebagai pendekatan sistematis dan berurutan. Menurut Wahid (2020), metode *Waterfall* merupakan salah satu model yang paling umum dalam proses pengembangan sistem. Model ini merupakan bentuk awal dari *Software Development Life Cycle* (SDLC) dan sering disebut sebagai model tradisional atau klasik. Disebut juga sebagai model linier sekuensial, karena setiap tahapannya dilakukan secara berurutan, mulai dari analisis, desain, implementasi, pengujian, hingga pemeliharaan.

Pemilihan model *Waterfall* dianggap tepat dalam penelitian ini karena pengembangan sistem membutuhkan alur kerja yang jelas dan terdokumentasi dengan baik. Terlebih, sistem ini mengintegrasikan algoritma *K-Means* untuk proses *clustering* (pengelompokan) data pengeluaran dalam sistem informasi berbasis web bernama Laporanku. Algoritma *K-Means* digunakan untuk mengelompokkan data pengeluaran ke dalam beberapa kategori berdasarkan kesamaan nilai, sehingga menghasilkan informasi yang lebih terstruktur dan mudah dianalisis.

Adapun tahapan pengembangan sistem menggunakan model *Waterfall* dapat dilihat pada Gambar 1 berikut.



**Gambar 1.** Metode Waterfall

Sumber : Titis Ulfa Mustikawati, 2022

##### 3.1.1 Tahap Analisis (*Analysis*)

Tahap analisis merupakan langkah awal dalam proses pengembangan sistem, yang bertujuan untuk mengidentifikasi kebutuhan pengguna serta memahami batasan sistem yang akan dibangun. Proses ini dilakukan melalui wawancara, observasi, dan diskusi langsung dengan pihak terkait di PT. Taspen (Persero) Kantor Cabang Bogor. Berdasarkan hasil analisis, diketahui bahwa proses pembuatan Lembar Pertanggungjawaban Transaksi (LPT) masih dilakukan secara manual menggunakan Microsoft Excel tanpa adanya dukungan sistem yang terintegrasi. Kondisi ini menyebabkan proses pengelolaan data menjadi kurang efisien, memakan waktu, dan rawan terjadi kesalahan pencatatan.

Untuk menjawab permasalahan tersebut, dibutuhkan sebuah sistem informasi berbasis web yang mampu mengotomatisasi proses pengelolaan data LPT secara menyeluruh. Selain itu, sistem ini akan dilengkapi dengan fitur klasterisasi data pengeluaran menggunakan algoritma *K-Means*. Penerapan algoritma ini

memungkinkan pengelompokan data berdasarkan pola tertentu, seperti kategori pengeluaran rendah, sedang, dan tinggi. Dengan adanya pengelompokan ini, pengguna dapat memperoleh gambaran yang lebih jelas mengenai pola pengeluaran, yang pada akhirnya mendukung pengambilan keputusan yang lebih tepat dalam pengelolaan dan perencanaan anggaran.

#### **3.1.1.1 Alur Penelitian**

Alur penelitian ini merupakan tahapan penting dalam proses pengembangan sistem, yang bertujuan untuk menyiapkan data agar siap digunakan dalam proses klusterisasi menggunakan algoritma *K-Means*, serta diintegrasikan ke dalam sistem informasi yang dikembangkan, yaitu Laporanku. Seluruh proses pengolahan data dilakukan di platform Google Collab untuk memanfaatkan lingkungan komputasi yang fleksibel dan mendukung berbagai pustaka pemrosesan data.

Tahapan dimulai dengan mengimpor dataset menggunakan pustaka *pandas*. Setelah data berhasil dimuat, dilakukan pemeriksaan struktur *dataset* untuk memahami kolom-kolom yang tersedia, jenis data, serta jumlah entri yang ada, seperti kolom tanggal, akun/deskripsi, deskripsi kegiatan, dan nilai realisasi (Rp). Tahap berikutnya adalah pembersihan data (*data cleaning*) untuk menghilangkan duplikasi dan nilai-nilai yang tidak relevan. Baris dengan nilai realisasi nol dihapus karena tidak berkontribusi terhadap analisis pengeluaran.

Selanjutnya dilakukan transformasi format tanggal, di mana data pada kolom tanggal diubah ke dalam format *datetime* guna mempermudah analisis berdasarkan periode tertentu. Kemudian, data dinormalisasi agar skala nilai pada kolom realisasi menjadi seimbang dan tidak mendominasi proses klusterisasi.

Tahap utama selanjutnya adalah penerapan algoritma *K-Means*. Penentuan jumlah cluster dilakukan dengan metode *elbow*, dan hasilnya digunakan untuk mengelompokkan data pengeluaran ke dalam beberapa kategori, seperti pengeluaran rendah, sedang, dan tinggi. Keluaran dari proses klusterisasi ini disimpan kembali dalam format .csv atau langsung diintegrasikan ke dalam basis data MySQL.

Hasil tersebut kemudian digunakan dalam sistem Laporanku sebagai bagian dari fitur laporan bulanan dan notifikasi kategori pengeluaran. Dengan mengikuti alur ini, sistem dapat menyajikan informasi keuangan yang lebih terstruktur, efisien, serta memberikan gambaran yang lebih jelas bagi pengambilan keputusan terkait anggaran.

#### **3.1.1.2 Pengumpulan Data**

Data yang digunakan dalam penelitian ini diperoleh dari laporan pengeluaran bulanan PT. Taspen (Persero) Kantor Cabang Bogor, dengan cakupan periode dari Mei 2020 hingga Oktober 2023. Data tersedia dalam format *comma-separated values* (.csv) dan memuat informasi terkait aktivitas pengeluaran pada Lembar Pertanggungjawaban Transaksi (LPT), dengan kolom-kolom utama seperti Tanggal, Akun/Deskripsi, Deskripsi Kegiatan, dan Realisasi (Rp). Pemilihan data ini didasarkan pada kemampuannya dalam merepresentasikan kondisi aktual pengeluaran perusahaan dalam jangka waktu yang cukup panjang.

Hal ini menjadikan dataset tersebut sangat relevan untuk dianalisis menggunakan algoritma *K-Means*, karena pola-pola pengeluaran dapat diidentifikasi secara lebih akurat. Dengan karakteristik tersebut, data ini mendukung proses klusterisasi dalam sistem informasi Laporanku, yang bertujuan untuk menyajikan

laporan LPT berdasarkan kategori pengeluaran. Pendekatan ini diharapkan dapat membantu manajemen dalam memahami tren pengeluaran dan mengambil keputusan yang lebih tepat terkait pengelolaan anggaran.

### 3.1.1.3 Implementasi Algoritma K-Means

Algoritma *K-Means* merupakan salah satu metode *unsupervised learning* yang digunakan untuk mengelompokkan data ke dalam beberapa kluster berdasarkan kemiripan nilai. Dalam penelitian ini, implementasi dilakukan menggunakan platform Google Collab dengan langkah-langkah seperti normalisasi data untuk menyamakan skala nilai realisasi pengeluaran, penentuan jumlah cluster optimal menggunakan metode *elbow*, dan pelatihan model menggunakan algoritma *K-Means*. Model yang terbentuk kemudian digunakan untuk memberi label kluster pada setiap entri data pengeluaran berdasarkan kesamaan nilai realisasinya.

Hasil klusterisasi ini diinterpretasikan ke dalam kategori pengeluaran seperti “Biasa”, “Sedang”, dan “Boros”. Pengelompokan ini tidak hanya membantu memahami pola pengeluaran, tetapi juga menjadi dasar dalam penyajian visualisasi laporan dan fitur analisis pada sistem informasi Laporanku.

Sebagai ilustrasi penerapan algoritma, dilakukan simulasi menggunakan data sampel dari Lembar Pertanggungjawaban Transaksi (LPT). Gambar 2 menunjukkan data awal yang mencakup akun, tanggal, deskripsi kegiatan, dan nilai realisasi pengeluaran. Selanjutnya, Gambar 3 menampilkan hasil klusterisasi yang diperoleh dari algoritma *K-Means*, yang mencakup proses normalisasi, pembentukan kluster, dan penentuan kategori pengeluaran berdasarkan hasil model.

AKUN/DESKRIPSI	TANGGAL	DESKRIPSI KEGIATAN	REALISASI (RP)
612408101/PEMELIHARAAN KANTOR	08/05/2020	PEMBELIAN PERLENGKAPAN KANTOR	1.659.600
612403101/TELEPON/INTERNET/FASIMILI,AIR, LISTRIK & GAS	12/05/2020	TAGIHAN PDAM	172.400
612203106/BIAYA LAYANAN PESERTA	12/05/2020	PEMBELIAN PULSA LAYANAN	124.000
612414101/BBM/TOL/PARKIR	08/05/2020	PEMBELIAN BBM MOBIL DINAS	235.053
612408103/PEMELIHARAAN/PERBAIKAN KENDARAAN	08/05/2020	NITRO-PLUS MOBIL DINAS	8.000

**Gambar 2.** Data Lembar Pertanggungjawaban Transaksi (LPT)

AKUN/DESKRIPSI	TANGGAL	DESKRIPSI KEGIATAN	REALISASI (RP)	TAHUN	BULAN	REALISASI_NORM	KLUSTER	KATEGORI_PENGELUARAN
612408101/PEMELIHARAAN KANTOR	2020-08-05	PEMBELIAN PERLENGKAPAN KANTOR	1659600000	2020	8	215	2	Boros
612403101/TELEPON/INTERNET/FASIMILI,AIR, LISTRIK & GAS	2020-12-05	TAGIHAN PDAM	172400000	2020	12	22	0	Biasa
612203106/BIAYA LAYANAN PESERTA	2020-12-05	PEMBELIAN PULSA LAYANAN	124000000	2020	12	16	0	Biasa
612414101/BBM/TOL/PARKIR	2020-08-05	PEMBELIAN BBM MOBIL DINAS	235053000	2020	8	0.03	0	Biasa
612408103/PEMELIHARAAN/PERBAIKAN KENDARAAN	2020-08-05	NITRO-PLUS MOBIL DINAS	8000000	2020	8	1	0	Biasa

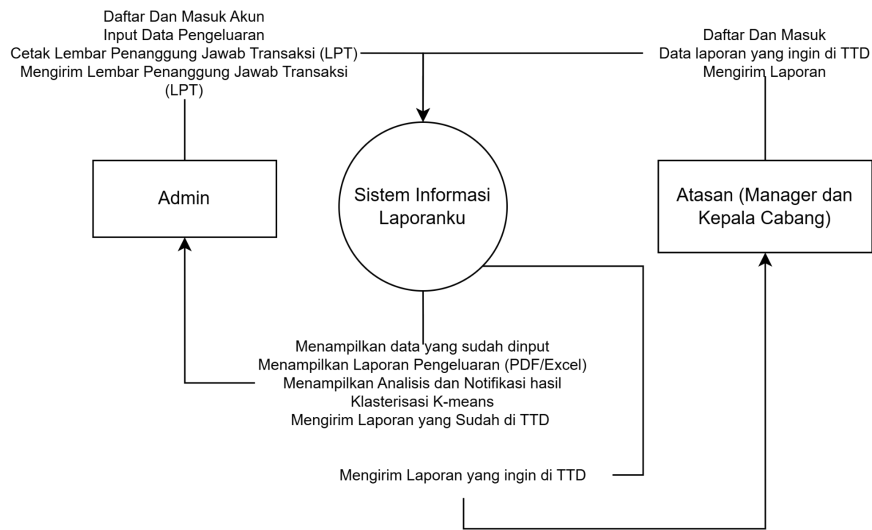
**Gambar 3.** Hasil Penerapan K-Means

### 3.1.2 Tahap Desain (*Design*)

Tahap desain merupakan proses perancangan sistem secara menyeluruh sebelum proses pembangunan dilakukan. Desain sistem ini mencakup Diagram Konteks, *Data Flow Diagram* (DFD), *Entity Relationship Diagram* (ERD), dan desain antarmuka pengguna (UI). Tujuannya adalah memastikan sistem memiliki alur yang jelas, mudah digunakan, serta mampu mendukung proses klusterisasi pengeluaran menggunakan algoritma *K-Means* secara efektif.

#### 3.1.2.1 Diagram Konteks

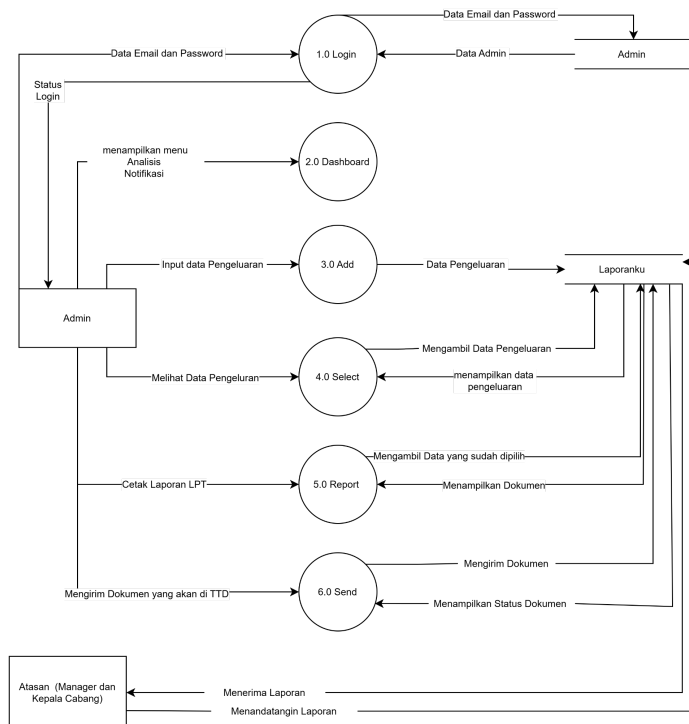
Diagram konteks adalah diagram yang terdiri dari suatu proses dan menggambarkan ruang lingkup suatu sistem. Diagram konteks merupakan level tertinggi dari DFD yang menggambarkan seluruh input ke dalam sistem atau output dari sistem yang memberi gambaran tentang keseluruhan sistem.



**Gambar 4.** Diagram konteks Laporan

### 3.1.2.2 Data Flow Diagram (DFD)

*Data Flow Diagram* (DFD) adalah diagram yang menggambarkan aliran data dari sebuah proses atau sistem informasi.

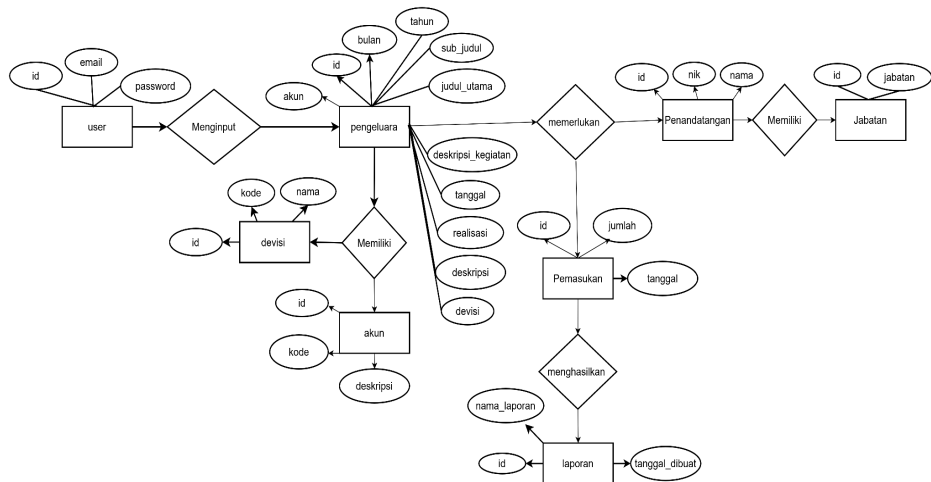


**Gambar 5.** Data Flow Diagram Laporan

### 3.1.2.3 Entity Relationship Diagram (ERD)

*Entity Relationship Diagram* (ERD) berikut ini menggambarkan hubungan antar entitas yang ada dalam sistem. Diagram ini menunjukkan bagaimana entitas-entitas tersebut saling terkait dan berinteraksi satu sama lain dalam mendukung fungsionalitas sistem.





**Gambar 6.** Entity Relationship Diagram

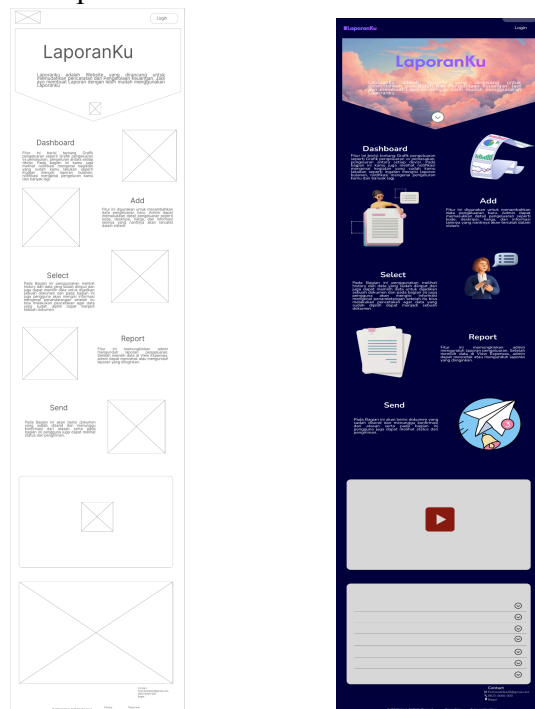
### 3.1.2.4 User Interface (UI)

Tahap perancangan *User Interface* (UI) bertujuan untuk mempermudah penggunaan Sistem Informasi Laporku oleh admin. Dalam tahap ini, pertama kali dilakukan pembuatan wireframe yang berfungsi untuk memberikan gambaran awal tentang tata letak halaman. Selanjutnya, dilakukan perancangan desain antarmuka yang lebih rinci untuk memvisualisasikan tampilan sistem secara keseluruhan.

Berikut ini adalah *wireframe* dan desain antarmuka yang telah dirancang untuk sistem:

#### 1. Halaman Utama

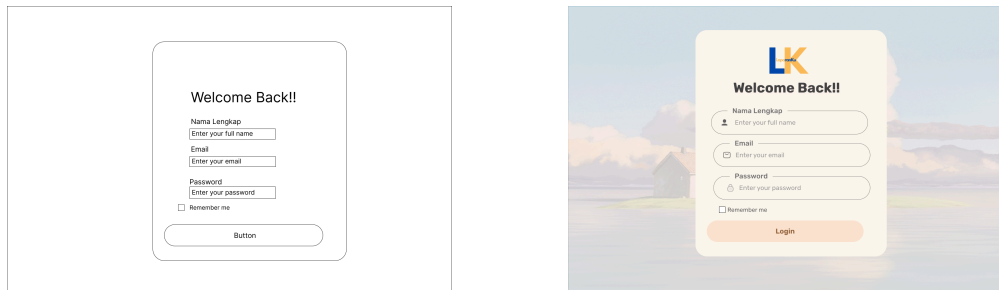
Halaman utama adalah halaman awal yang menampilkan pengenalan singkat mengenai Sistem Informasi Laporku. Pada halaman ini, pengguna dapat melihat deskripsi fitur utama dan juga menyediakan video demo penggunaan sistem serta bagian pertanyaan umum seputar web.



**Gambar 7.** Wireframe dan Desain Halaman Utama

## 2. Halaman Daftar

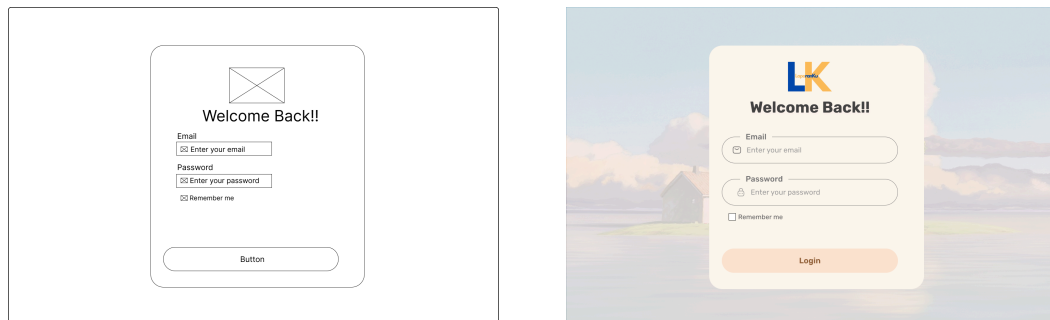
Halaman Daftar adalah halaman yang memungkinkan pengguna untuk membuat akun di Sistem Informasi Laporanku. Pada halaman ini, pengguna diminta mengisi nama lengkap, email, dan *password*. Setelah proses pendaftaran berhasil, pengguna dapat langsung masuk dan mulai menggunakan fitur-fitur yang tersedia.



**Gambar 8.** Wireframe dan Desain Halaman Daftar

## 3. Halaman Masuk

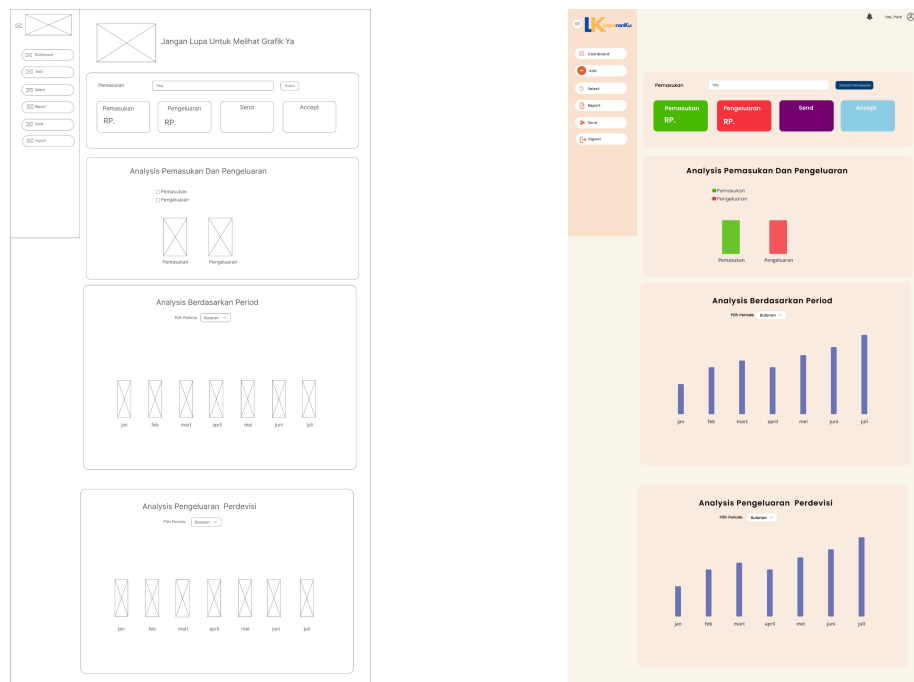
Halaman Masuk adalah halaman yang digunakan oleh pengguna terdaftar untuk mengakses Sistem Informasi Laporanku sesuai dengan hak aksesnya. Jika pengguna memasukkan email yang tidak terdaftar, sistem akan menampilkan notifikasi “Email tidak ditemukan”. Sementara itu, apabila *password* yang dimasukkan salah, sistem akan menampilkan notifikasi “*Password* salah”.



**Gambar 9.** Wireframe dan Desain Halaman Masuk

## 4. Halaman Dashboard

Halaman Dashboard adalah halaman utama yang ditampilkan setelah pengguna berhasil *login* ke Sistem Informasi Laporanku. Pada halaman ini, pengguna dapat melihat grafik perbandingan antara pengeluaran dan pemasukan, serta grafik pengeluaran berdasarkan divisi. Selain itu, dashboard juga menampilkan berbagai notifikasi penting, seperti notifikasi ketika pengeluaran melebihi pemasukan, pengingat untuk mengisi data bulanan, peringatan jika pengeluaran bulan ini lebih besar dari bulan sebelumnya, informasi bahwa dokumen laporan telah ditandatangani, dan peringatan apabila pengeluaran mendekati jumlah pemasukan. Pengguna juga dapat melihat jumlah dokumen yang sudah ditandatangani maupun yang belum ditandatangani oleh atasan.



**Gambar 10.** Wireframe dan Desain Halaman Dashboard

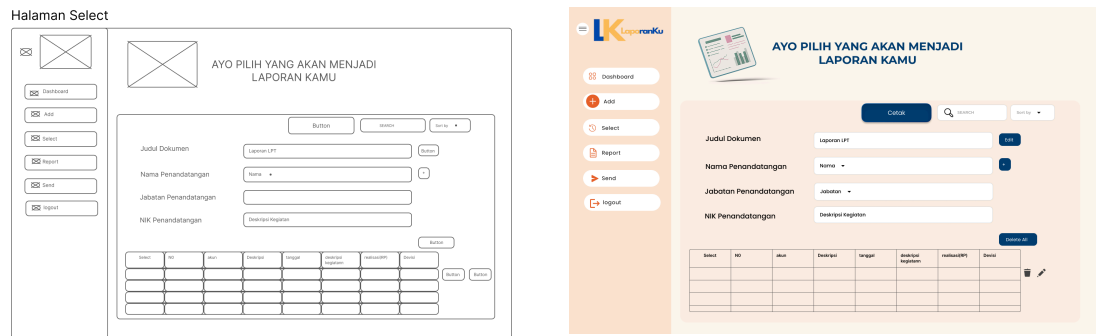
## 5. Halaman Add Data

Halaman *Add Data* adalah halaman yang digunakan oleh admin untuk mencatat data pengeluaran melalui sebuah form. Setelah seluruh kolom pada form diisi dengan lengkap, admin dapat menekan tombol Submit untuk menyimpan data tersebut ke dalam sistem. Data yang telah disubmit akan otomatis tersimpan dan dapat dilihat pada halaman *Select* untuk diproses lebih lanjut

**Gambar 11.** Wireframe dan Desain Halaman Add Data

## 6. Halaman Select

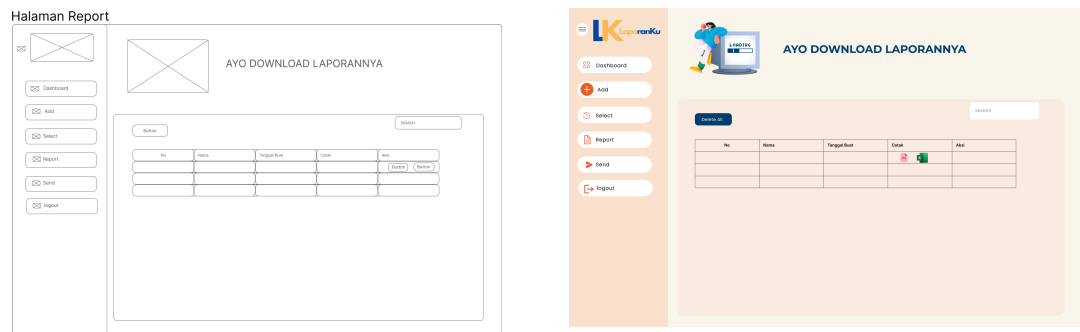
Halaman *Select* adalah halaman yang menampilkan seluruh data pengeluaran yang telah diinput oleh admin melalui halaman *Add Data*. Pada halaman ini, admin dapat memilih data yang akan dijadikan dokumen untuk diproses lebih lanjut. Sebelum data tersebut diubah menjadi dokumen, admin harus mengisi form tambahan yang mencakup judul dokumen dan informasi penandatanganan.



**Gambar 12.** Wireframe dan Desain Halaman Select

## 7. Halaman Report

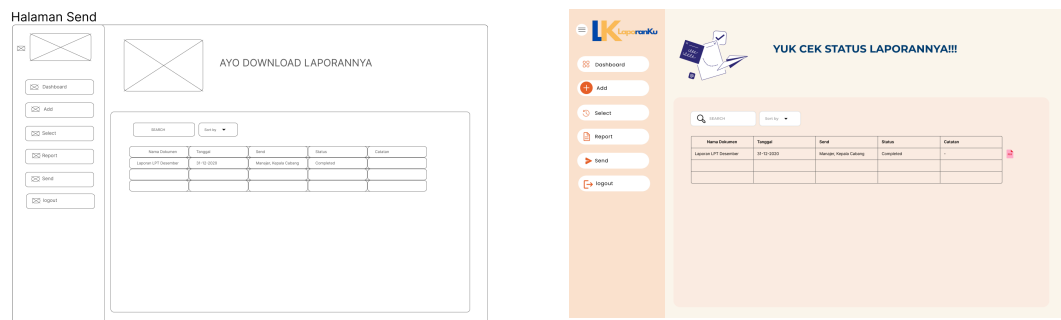
Halaman *Report* adalah halaman di mana admin dapat melihat dokumen yang telah dihasilkan dari data yang dipilih pada halaman *Select*. Dokumen yang sudah diproses dapat diunduh dalam format PDF atau Excel sesuai kebutuhan. Admin juga dapat memfilter dokumen berdasarkan bulan dan tahun untuk memudahkan pencarian. Selain itu, dokumen tersebut dapat langsung dikirim ke atasan untuk ditandatangani secara digital melalui sistem.



**Gambar 13.** Wireframe dan Desain Halaman Report

## 8. Halaman Send

Halaman *Send* adalah halaman yang menampilkan daftar dokumen yang telah dikirim oleh admin dan sedang menunggu proses tanda tangan dari atasan. Pada halaman ini, admin dapat memantau status masing-masing laporan, apakah masih dalam status *Pending*, sudah *Signed by Manager*, atau telah *Complete*. Halaman ini membantu admin dalam melacak progres tanda tangan dokumen secara *real-time*.



**Gambar 14.** Wireframe dan Desain Halaman Send

### 3.1.3 Tahap Implementasi (*Implementation*)

Tahap implementasi merupakan proses di mana sistem mulai direalisasikan berdasarkan rancangan yang telah disusun sebelumnya. Pada tahap ini, pengembangan *backend* dilakukan menggunakan bahasa pemrograman *Python* dengan *framework Flask* atau *FastAPI*. Untuk bagian *frontend*, digunakan *React.js* dan *CSS* guna menciptakan tampilan yang responsif dan *user-friendly*. Sistem penyimpanan data dibangun menggunakan *MySQL* sebagai basis data utama. Sementara itu, untuk proses klasterisasi data pengeluaran, digunakan platform *Google Collab* dengan penerapan algoritma *K-Means* guna memperoleh hasil cluster yang optimal.

### 3.1.4 Tahap Pengujian (*Testing*)

Tahap pengujian dilakukan untuk memastikan bahwa sistem yang dibangun telah berfungsi dengan baik, sesuai dengan kebutuhan pengguna, dan menghasilkan keluaran yang akurat. Pengujian ini mencakup beberapa aspek berikut:

1. Pengujian Struktural

Pengujian ini difokuskan pada struktur internal sistem, meliputi arsitektur kode program, relasi antar komponen, serta efisiensi alur logika sistem. Tujuan dari pengujian ini adalah untuk memastikan bahwa sistem memiliki struktur yang stabil dan mudah untuk dikembangkan di kemudian hari.

2. Pengujian Fungsional

Pada pengujian ini, seluruh fitur dan fungsi dari sistem diuji berdasarkan spesifikasi yang telah ditentukan sebelumnya. Hal ini mencakup proses input data pengeluaran, pemrosesan algoritma *K-Means* untuk klasterisasi, hingga pembuatan dan penyajian output dalam bentuk laporan yang dapat diunduh atau dikirim ke atasan.

3. Pengujian Validasi

Pengujian validasi dilakukan dengan melibatkan pengguna akhir untuk menilai sejauh mana sistem dapat memenuhi kebutuhan mereka. Aspek yang diuji antara lain kemudahan penggunaan (*usability*), kejelasan antarmuka, serta keakuratan dan relevansi hasil yang ditampilkan oleh sistem.

4. Evaluasi Kualitas Klasterisasi dengan *Davies Bouldin Index* (DBI)

Untuk mengukur kualitas hasil klasterisasi yang dihasilkan oleh algoritma *K-Means*, digunakan metrik *Davies Bouldin Index* (DBI). Nilai DBI dihitung berdasarkan rata-rata rasio antara jarak dalam klaster (*intra-cluster*) dan jarak antar klaster (*inter-cluster*). Semakin kecil nilai DBI, maka semakin baik kualitas klasterisasi yang dihasilkan, karena menunjukkan bahwa klaster saling terpisah dengan jelas dan data dalam satu klaster memiliki kemiripan yang tinggi.

### 3.1.5 Tahap Pemeliharaan (*Maintenance*)

Tahap pemeliharaan dilakukan setelah sistem digunakan secara resmi oleh pengguna. Pada tahap ini, sistem dipantau dan dievaluasi secara berkala untuk memastikan kinerjanya tetap optimal serta sesuai dengan kebutuhan yang terus berkembang. Oleh karena itu, kegiatan pemeliharaan mencakup perbaikan bug serta penyempurnaan fitur berdasarkan masukan dari pengguna.