

**PENERAPAN MODEL CRISP-DM UNTUK MENGELOMPOKAN OBJEK  
WISATA DI KABUPATEN SUBANG BERDASARKAN FASILITAS DAN  
JUMLAH PENGUNJUNG DENGAN ALGORITMA K-MEANS**

**Proposal Tugas Akhir  
Program Studi S-1 Sistem Informasi**



diajukan oleh:  
Agisni Amalia Lugina  
212103002

**FAKULTAS TEKNIK & TEKNOLOGI INFORMASI  
UNIVERSITAS JENDERAL ACHMAD YANI YOGYAKARTA  
2024**

**Proposal Tugas Akhir**

**PENERAPAN MODEL CRISP-DM UNTUK MENGELOMPOKAN OBJEK  
WISATA DI KABUPATEN SUBANG BERDASARKAN FASILITAS DAN  
JUMLAH PENGUNJUNG DENGAN ALGORITMA K-MEANS**

diajukan oleh  
Agisni Amalia Lugina  
212103002

telah disetujui pada tanggal:  
2024

Pembimbing

Aris Wahyu Murdiyanto, S.Kom., M.Cs  
NPP: 0502098501

**Pengesahan Proposal Tugas Akhir**

**PENERAPAN MODEL CRISP-DM UNTUK MENGELOMPOKAN OBJEK  
WISATA DI KABUPATEN SUBANG BERDASARKAN FASILITAS DAN  
JUMLAH PENGUNJUNG DENGAN ALGORITMA K-MEANS**

dipersiapkan dan disusun oleh  
Agisni Amalia Lugina  
212103002

telah disahkan pada tanggal:  
2024

Penguji I

Penguji II

Kharisma, S.T., M.Cs.  
NIDN : 0502108201

Ulfi Saidata Aesyi, S.Kom., M.Cs.  
NIDN : 0515129002

Mengetahui  
Ketua Program Studi S-1 Sistem Informasi  
Fakultas Teknik & Teknologi Informasi  
Universitas Jenderal Achmad Yani Yogyakarta

Ahmad Hanafi, S.T., M.Eng.  
NPP : 2008.13.0020

## DAFTAR ISI

<b>1</b>	<b>Latar Belakang.....</b>	<b>1</b>
1.1	Perumusan Masalah.....	2
1.2	Manfaat Penelitian.....	3
1.3	Pertanyaan Penelitian .....	3
<b>2</b>	<b>Tujuan Penelitian.....</b>	<b>3</b>
<b>3</b>	<b>Tinjauan Pustaka.....</b>	<b>3</b>
<b>4</b>	<b>Landasan Teori .....</b>	<b>6</b>
4.1	Teori .....	Error! Bookmark not defined.
4.1.1	Penulisan Tabel .....	Error! Bookmark not defined.
4.1.2	Penulisan Gambar .....	Error! Bookmark not defined.
<b>5</b>	<b>Metode Penelitian.....</b>	<b>11</b>
5.1	Bahan Penelitian.....	11
5.2	Alat Penelitian .....	12
5.3	Jalan Penelitian.....	13
<b>6</b>	<b>Jadwal Penelitian .....</b>	<b>14</b>
	<b>Lampiran .....</b>	Error! Bookmark not defined.

## **1 LATAR BELAKANG**

Salah satu sektor penting yang berperan dalam pembangunan perekonomian suatu daerah adalah pariwisata. Kabupaten Subang sebagai salah satu daerah tujuan wisata di Jawa Barat mempunyai banyak sekali potensi daya tarik wisata sehingga dapat dikembangkan. Meski demikian, proses pembangunan dan pengelolaan objek wisata di Kabupaten Subang masih belum maksimal (Primanty, 2023). Mungkin hal tersebut merupakan salah satu permasalahan yang dihadapi oleh pengembangan teknologi smart pariwisata dimana bagaimana cara mengklasifikasikan suatu objek wisata berdasarkan fasilitas dan jumlah wisatawan yang mengunjungi tempat wisata tersebut sehingga pengembangan dapat dilakukan sesuai dengan karakteristik masing-masing wisatawan. daya tarik (Sulistyo et al., 2022).

Pendahuluan Pengelompokan objek wisata berdasarkan fasilitas yang dibatasi dan jumlah pengunjung dapat dilakukan dengan teknik data mining, salah satu teknik yang digunakan dalam implementasinya adalah algoritma K-Means. K-Means merupakan salah satu algoritma clustering yang paling banyak digunakan dengan hasil yang baik dalam clustering data yang memiliki karakteristik yang sama (Wahyudi et al., 2020). Kedua, CRISP-DM (Cross-Industry Standard Process for Data Mining) merupakan model berbasis proses yang dapat diterapkan untuk melakukan rangkaian proses data mining secara terorganisir dan sistematis (Normah et al. 2020).

Oleh karena itu, penelitian ini akan menggunakan model CRISP-DM untuk mengelompokkan objek wisata di Kabupaten Subang berdasarkan fasilitas yang tersedia dan jumlah pengunjung dengan menggunakan algoritma K-Means. Hasil penelitian ini diharapkan dapat menjadi sumber informasi bagi pemerintah daerah pinggiran dan pengelola tempat wisata di wilayah Kabupaten Subang untuk lebih mengembangkan dan mengelola tempat wisatanya secara lebih optimal.

Penelitian sebelumnya telah menganalisis potensi dan pengembangan objek wisata di Kabupaten Subang serta strategi pengembangan objek wisata di daerah tersebut. Namun, penelitian tersebut belum fokus pada pengelompokan objek wisata berdasarkan fasilitas dan jumlah pengunjung menggunakan algoritma data mining (Lesmana, 2023). Hal ini menunjukkan adanya kebutuhan untuk

menerapkan teknik data mining, seperti algoritma K-Means, dalam mengklasifikasikan objek wisata agar pengembangan dapat dilakukan secara lebih efektif dan sesuai dengan karakteristik masing-masing objek wisata (Hidayati et al., 2021).

Selain itu, penelitian terkait penerapan model CRISP-DM untuk analisis data di bidang pariwisata juga telah dilakukan, seperti pada analisis perilaku wisatawan berdasarkan data ulasan di destinasi wisata (Singgalen, 2023). Penelitian tersebut menunjukkan bahwa model CRISP-DM dapat digunakan untuk menghasilkan rekomendasi pengembangan destinasi yang kontekstual sesuai dengan preferensi wisatawan. Namun, belum ada penelitian yang secara khusus menerapkan model CRISP-DM untuk mengelompokkan objek wisata di Kabupaten Subang.

Dengan demikian, penelitian ini diharapkan dapat memberikan kontribusi yang signifikan dalam pengembangan dan pengelolaan objek wisata di Kabupaten Subang melalui penerapan model CRISP-DM dan algoritma K-Means untuk mengelompokkan objek wisata berdasarkan fasilitas dan jumlah pengunjung.

Selain itu, penelitian ini juga diharapkan dapat menjadi referensi bagi penelitian-penelitian selanjutnya yang ingin mengkaji penerapan teknik data mining dalam pengembangan pariwisata di daerah lain.

Dengan berbagai pertimbangan tersebut, penelitian ini perlu dilakukan untuk memberikan solusi terkait permasalahan pengelompokan objek wisata di Kabupaten Subang berdasarkan fasilitas dan jumlah pengunjung.

## **1.1 PERUMUSAN MASALAH**

Berdasarkan latar belakang yang telah dipaparkan, permasalahan utama yang akan diteliti dalam penelitian ini adalah bagaimana mengelompokkan objek wisata di Kabupaten Subang berdasarkan fasilitas dan jumlah pengunjung menggunakan model CRISP-DM dan algoritma K-Means.

## **1.2 MANFAAT PENELITIAN**

Hasil penelitian ini diharapkan dapat memberikan manfaat sebagai berikut:

1. Memberikan informasi yang bermanfaat bagi pemerintah daerah dan pengelola objek wisata dalam mengembangkan dan mengelola objek wisata di Kabupaten Subang secara lebih optimal.
2. Memberikan kontribusi dalam pengembangan penelitian di bidang pariwisata, khususnya dalam penerapan model CRISP-DM dan algoritma data mining untuk pengelompokan objek wisata.
3. Menjadi referensi bagi penelitian-penelitian selanjutnya yang ingin mengkaji penerapan teknik data mining dalam pengembangan pariwisata di daerah lain.

## **1.3 PERTANYAAN PENELITIAN**

1. Bagaimana proses penerapan model CRISP-DM dalam mengelompokkan objek wisata di Kabupaten Subang berdasarkan fasilitas dan jumlah pengunjung?
2. Bagaimana hasil pengelompokan objek wisata di Kabupaten Subang berdasarkan fasilitas dan jumlah pengunjung menggunakan algoritma K-Means?

## **2 TUJUAN PENELITIAN**

1. Menerapkan model CRISP-DM dalam mengelompokkan objek wisata di Kabupaten Subang berdasarkan fasilitas dan jumlah pengunjung.
2. Menganalisis hasil pengelompokan objek wisata di Kabupaten Subang berdasarkan fasilitas dan jumlah pengunjung menggunakan algoritma K-Means.

## **3 TINJAUAN PUSTAKA**

### **3.1 PENELITIAN TENTANG POTENSI DAN PENGEMBANGAN OBJEK WISATA DI KABUPATEN SUBANG**

Penelitian terkait analisis potensi objek wisata di Kabupaten Subang telah dilakukan oleh beberapa peneliti. Salah satu studi yang menonjol adalah penelitian oleh Sulistyo et al. (2023), yang mengkaji pengembangan potensi Curug Cileat sebagai objek wisata alam yang berkelanjutan. Penelitian ini menggunakan metode

campuran dan Analisa SWOT-8K untuk mengevaluasi kelebihan, kelemahan, peluang, dan ancaman dalam pengelolaan wisata di Curug Cileat. Hasil penelitian menunjukkan bahwa dengan pengelolaan yang tepat, Curug Cileat dapat menjadi destinasi wisata unggulan yang tidak hanya meningkatkan perekonomian lokal tetapi juga mempertahankan kelestarian alam sekitar. Sulistyو et al. menekankan pentingnya partisipasi masyarakat setempat dalam setiap tahapan pengembangan, mulai dari perencanaan hingga pelaksanaan, untuk memastikan keberlanjutan jangka panjang (Sulistyو et al., 2023).

Selain itu, penelitian tentang strategi pengembangan objek wisata di Kabupaten Subang juga telah dilakukan oleh beberapa peneliti. Farida et al. (2023) mengidentifikasi potensi desa wisata alternatif di Desa Pringkasap, Kecamatan Pabuaran. Penelitian ini berfokus pada strategi pengembangan wisata yang berbasis pada potensi lokal dan kearifan budaya setempat. Mereka menggunakan pendekatan *participatory rural appraisal* (PRA) untuk melibatkan masyarakat dalam proses identifikasi dan pengembangan potensi wisata. Hasil penelitian menunjukkan bahwa Desa Pringkasap memiliki potensi besar untuk dikembangkan sebagai desa wisata dengan atraksi budaya dan alam yang unik. Strategi pengembangan yang diusulkan meliputi peningkatan infrastruktur, pelatihan bagi masyarakat setempat, dan promosi pariwisata yang terintegrasi (Farida et al., 2023).

Penelitian yang mengkaji peran pariwisata dalam pembangunan ekonomi daerah di Kabupaten Subang juga telah banyak dibahas dalam jurnal-jurnal Indonesia terbaru. Krisdina et al. (2021) menyebutkan bahwa pengembangan desa wisata tidak hanya berperan dalam pelestarian budaya tetapi juga memiliki dampak signifikan terhadap perekonomian lokal. Studi mereka menunjukkan bahwa dengan pengembangan desa wisata yang fokus pada budaya lokal, masyarakat setempat dapat meningkatkan pendapatan dan kualitas hidup mereka. Krisdina et al. menekankan bahwa kolaborasi antara pemerintah, swasta, dan masyarakat sangat penting untuk keberhasilan pengembangan pariwisata. Mereka juga menyarankan adanya kebijakan yang mendukung pengembangan pariwisata berkelanjutan untuk memastikan bahwa manfaat ekonomi dapat dirasakan oleh seluruh lapisan masyarakat (Krisdina et al., 2021).



### **3.2 PENELITIAN TENTANG PENERAPAN ALGORITMA K-MEANS UNTUK PENGELOMPOKAN OBJEK WISATA**

Penelitian tentang penggunaan algoritma K-Means untuk mengelompokkan objek wisata di Indonesia telah banyak dilakukan oleh beberapa peneliti. Salah satu penelitian yang penting adalah oleh Helia et al (2023), yang membahas pengelompokan tempat wisata berdasarkan jenisnya menggunakan metode K-Means pada data dari Provinsi Jawa Barat. Mereka menemukan bahwa pengelompokan ini bisa membantu menyebarkan kunjungan wisatawan lebih merata dan meningkatkan ekonomi lokal.

Selain itu, Lusianah et al (2023) meneliti pengelompokan jumlah wisatawan berdasarkan jenis akomodasi di Jawa Barat dengan algoritma K-Means. Hasil penelitian mereka menunjukkan bahwa pengelompokan ini dapat membantu merencanakan strategi pemasaran yang lebih efektif, sehingga penyedia layanan wisata bisa mengembangkan layanan yang sesuai dengan kebutuhan berbagai segmen wisatawan.

Penelitian lain oleh Al-Fahmi et al (2023) juga menggunakan algoritma K-Means untuk mengelompokkan objek wisata di Kabupaten Bojonegoro. Penelitian ini menemukan bahwa pengelompokan ini dapat membantu dalam membuat keputusan strategi pemasaran yang lebih efisien, sehingga daya tarik wisata dan keuntungan ekonomi bagi daerah tersebut bisa meningkat.

### **3.3 PENELITIAN TENTANG PENERAPAN MODEL CRISP-DM DALAM ANALISIS DATA DI BIDANG PARIWISATA**

CRISP-DM adalah metode yang sangat terstruktur dan terorganisir yang digunakan dalam proyek data science, termasuk dalam analisis data wisatawan. Pramono et al (2023) menjelaskan bagaimana model ini dapat digunakan untuk mengelola data wisatawan, seperti menganalisis ulasan pengunjung dan mengelompokkan wisatawan berdasarkan preferensi mereka. Model ini membantu dalam memahami kebutuhan wisatawan dan meningkatkan layanan pariwisata (Pramono et al., 2023).

Penelitian lainnya oleh Purnama et al (2023) menunjukkan bahwa penerapan model CRISP-DM pada sistem informasi eksekutif Dinas Kelautan dan Perikanan Provinsi Jawa Tengah dapat menghasilkan laporan yang lebih informatif dan bermanfaat untuk pengambilan keputusan. Dengan menggunakan CRISP-DM, data yang ada diubah menjadi informasi yang mudah dimengerti dan digunakan untuk membuat keputusan yang lebih baik dalam pengelolaan pariwisata dan kelautan (Purnama et al., 2023).

Yerik Afrianto (2023) juga mengaplikasikan model CRISP-DM untuk klasifikasi ulasan pengunjung di Danau Toba. Hasil penelitiannya menunjukkan bahwa metode ini dapat membantu dalam mengidentifikasi kebutuhan wisatawan dan meningkatkan kualitas layanan pariwisata. Dengan menganalisis ulasan pengunjung, penyedia layanan wisata dapat memahami apa yang disukai dan tidak disukai wisatawan, sehingga mereka dapat meningkatkan kualitas dan daya tarik destinasi wisata (Afrianto, 2023).

Penelitian oleh Dwi Astuti dan Surya (2023) meneliti penerapan CRISP-DM dalam analisis data wisatawan di Bali dan menemukan bahwa model ini dapat membantu dalam merencanakan strategi pemasaran yang lebih efektif. Surya dan Astuti (2023) juga menerapkan model CRISP-DM untuk mengelompokkan data wisatawan berdasarkan preferensi dan perilaku, yang dapat membantu penyedia layanan wisata dalam mengembangkan produk yang lebih sesuai dengan kebutuhan wisatawan. Penelitian ini menunjukkan bahwa CRISP-DM adalah alat yang sangat berguna dalam mengelola dan menganalisis data wisatawan untuk meningkatkan pariwisata di berbagai daerah (Astuti & Surya, 2023).

## **4 LANDASAN TEORI**

### **4.1 KONSEP PARIWISATA DAN PENGEMBANGAN OBJEK WISATA**

Pariwisata, atau *tourism*, didefinisikan sebagai aktivitas perjalanan di luar lingkungan yang dilakukan seseorang dengan berbagai tujuan, yang biasanya berlangsung kurang dari satu tahun (Prakoso, & Zafrullah, 2020). Aktivitas ini dapat berupa perjalanan rekreasi, bisnis, atau pendidikan, dan berperan penting dalam mendorong pembangunan ekonomi dan meningkatkan kesejahteraan masyarakat

(Novandi & Rukminto, 2021). Dengan meningkatnya mobilitas dan aksesibilitas, pariwisata tidak hanya memberikan pengalaman yang berharga bagi individu, tetapi juga berkontribusi pada pendapatan nasional melalui sektor industri yang terus berkembang (Maharani et al, (2024).

Pengembangan pariwisata merupakan upaya terkoordinasi yang bertujuan untuk menarik wisatawan dengan menyediakan sarana dan prasarana yang memadai, serta melayani kebutuhan mereka secara optimal (Gustin et al, 2021). Langkah ini mencakup perencanaan strategis dan pelaksanaan program yang dapat meningkatkan daya tarik destinasi wisata. Selain itu, pengembangan pariwisata juga harus melibatkan partisipasi aktif dari komunitas lokal untuk memastikan keberlanjutan dan manfaat jangka panjang bagi mereka (Rusyidi & Ferdiansyah, 2019).

Menurut Rosidah et al (2024), Tujuan utama dari pengembangan objek wisata adalah untuk memberikan keuntungan baik bagi wisatawan maupun pelaku bisnis pariwisata, sekaligus meningkatkan perekonomian masyarakat setempat. Pendekatan ini mencakup investasi dalam infrastruktur, promosi destinasi, serta peningkatan kualitas layanan (Rosidah et al, 2024). Dengan demikian, pariwisata dapat menjadi pendorong utama pertumbuhan ekonomi di daerah yang berpotensi.

Namun, dalam pengembangan pariwisata terdapat berbagai tantangan, termasuk kekurangan fasilitas, infrastruktur yang kurang memadai, dan keterbatasan sumber daya manusia (Yuda, 2023). Agar pengembangan pariwisata dapat berhasil, penting untuk mempertimbangkan kondisi lingkungan dan mendapatkan dukungan dari pemerintah daerah serta stakeholder lainnya (Pratiwi & Ruspianda, 2023).

## **4.2 DATA MINING**

Menurut Mardieni et al (2023), Data mining merupakan proses yang kompleks dan penting dalam dunia analisis data. Proses ini bertujuan untuk menggali pengetahuan berharga atau menemukan pola tersembunyi dari kumpulan data yang besar dan kompleks. Menurut tujuan utama dari data mining adalah untuk mengungkap informasi yang mungkin tidak terlihat dengan jelas atau yang tidak mudah diidentifikasi melalui analisis yang rutin (Mardiani et al, 2023).

Dalam praktiknya, data mining melibatkan penggunaan berbagai metode statistik, teknik matematika, serta kecerdasan buatan untuk menganalisis dan menginterpretasikan data yang ada (Mardiani et al, 2023). Pendekatan ini memungkinkan pengambilan keputusan yang lebih baik dan lebih informasional dalam *business inteligent*. Seperti yang dinyatakan oleh Tantowi (2021), data mining berfungsi sebagai komponen krusial dalam proses pengambilan keputusan strategis di berbagai organisasi.

Metode yang umum digunakan dalam data mining mencakup klasifikasi, pengelompokan, dan asosiasi (Tantowi, 2021). Berikut ini merupakan penjelasan dari masing-masing metode:

1. Klasifikasi

Menurut Aryani dan Wijayanto (2021) dalam penelitiannya yang menyebutkan bahwa klasifikasi merupakan proses yang digunakan untuk mengelompokkan data ke dalam kategori yang sudah ditentukan. Metode ini biasanya melibatkan algoritma pembelajaran mesin, seperti Decision Trees atau Random Forests, yang membantu mengidentifikasi pola dari data pelatihan untuk membuat prediksi pada data baru (Setiono, 2022).

2. Pengelompokan (*clustering*)

Penelitian yang dilakukan oleh Annisa et al (2022) menyebutkan bahwa *Clustering* adalah metode yang digunakan untuk membagi sekumpulan data menjadi kelompok-kelompok yang disebut cluster. Dalam setiap *cluster*, objek-objek memiliki karakteristik yang serupa satu sama lain, sedangkan perbedaan terlihat antara *cluster* yang berbeda. Teknik ini sangat bermanfaat karena dapat mengidentifikasi kelompok-kelompok yang tidak terduga dalam suatu dataset.

3. Asosiasi

Metode Asosiasi berfokus pada menemukan aturan atau hubungan antara item dalam data (Wijaya et al, 2022). Contohnya adalah analisis keranjang belanja, di mana dapat diidentifikasi bahwa pelanggan yang membeli roti sering juga membeli mentega. Teknik ini biasanya melibatkan algoritma seperti Apriori atau FP-Growth (Rachman & Hunaifi,

2020). Asosiasi diukur menggunakan parameter support dan Confidence, di mana support menunjukkan tingkat dominasi item dalam transaksi, sedangkan Confidence menggambarkan kekuatan hubungan antar-item (Wijaya et al, 2022). Hasil analisis metode asosiasi dapat digunakan untuk pengambilan keputusan bisnis dan strategi penjualan (Fahrudin, 2019)

#### **4.3 ALGORITMA K-MEANS**

Algoritma K-Means adalah salah satu teknik pembelajaran tanpa pengawasan (unsupervised learning) yang sering digunakan untuk mengelompokkan data ke dalam beberapa kelompok berdasarkan kemiripan (Ramadhani & Megawati, 2023). Teknik ini terbukti sangat efektif dalam berbagai bidang, termasuk pariwisata. Misalnya, algoritma K-Means diterapkan untuk menganalisis dan mengelompokkan data pariwisata dalam berbagai konteks. Penelitian yang dilakukan oleh Nurjanah dan Arifin (2021) mengelompokkan ulasan perjalanan dari platform TripAdvisor untuk mengidentifikasi objek wisata populer dan memberikan peringkat berdasarkan preferensi pengguna. Dengan menggunakan algoritma ini, informasi berharga mengenai perilaku dan pilihan wisatawan dapat diekstrak, yang bertujuan untuk mendukung pengambilan keputusan dalam industri pariwisata (Seimahuira, 2021).

K-Means bekerja dengan cara mempartisi titik-titik data ke dalam k-kluster berdasarkan kemiripannya (Nasution & Hasibuan, 2020). Setiap kluster direpresentasikan oleh centroid-nya, yang merupakan titik tengah dari semua data yang berada dalam kluster tersebut. Algoritma ini secara iteratif menetapkan titik-titik data ke centroid terdekat dan menghitung ulang posisi centroid hingga mencapai konvergensi (Abidin et al, 2022). Untuk menentukan jumlah cluster yang optimal, dapat digunakan metode Elbow (Umargono et al, 2020).

K-Means seringkali mengungguli metode lain seperti K-Medoids dalam hal kecepatan pemrosesan dan kualitas kluster, sebagaimana diukur dengan Indeks Davies-Bouldin (DBI) (Abidin et al, 2022). Dalam hasil penelitian perbandingan algoritma yang dilakukan oleh Abidin et al (2022), K-Means secara konsisten memperoleh nilai DBI yang lebih rendah daripada K-Medoids, yang mengindikasikan hasil klasterisasi yang lebih baik.

Meskipun banyak digunakan untuk pengelompokan, algoritma K-Means memiliki beberapa kelemahan. Salah satu masalah utama adalah sensitivitasnya terhadap pemilihan centroid awal, yang dapat menyebabkan solusi lokal suboptimal (Orisa, 2022). Keacakan dalam inisialisasi centroid ini sering kali memerlukan beberapa kali pengujian untuk mencapai hasil yang optimal (Orisa, 2022). K-means juga rentan terhadap outlier, yang dapat secara signifikan mendistorsi distribusi kluster (Kamila et al, 2019). Untuk mengatasi keterbatasan ini, berbagai teknik optimasi telah diusulkan. Metode evaluasi internal seperti Elbow, Davies-Bouldin Index, dan Silhouette Index dapat membantu menentukan jumlah kluster yang optimal (Orisa, 2022).

#### **4.4 MODEL CROSS-INDUSTRY STANDARD PROCESS FOR DATA MINING (CRISP-DM)**

Model CRISP-DM adalah sebuah kerangka kerja yang sering digunakan dalam proyek data mining. Kerangka kerja ini terdiri dari enam tahap yang berurutan, di mana masing-masing tahap memiliki tujuan dan proses yang spesifik. Berdasarkan penelitian yang dilakukan oleh Nodeh et al. (2019), enam tahap tersebut adalah:

1. *Business Understanding*

Fase pertama ini bertujuan untuk memahami tujuan bisnis yang ingin dicapai melalui analisis data dan bagaimana hasil dari proyek data mining dapat memberikan solusi yang sesuai dengan kebutuhan organisasi.

2. *Data Understanding*

Dalam fase ini, peneliti melakukan pengumpulan dan eksplorasi data untuk memahami karakteristik, kualitas, dan struktur data yang akan digunakan.

3. *Data Preparation*

Fase ini melibatkan persiapan data untuk analisis lebih lanjut, termasuk pembersihan, transformasi, dan penggabungan data dari berbagai sumber.

4. *Modeling*

Pada tahap ini, berbagai teknik pemodelan diterapkan untuk menciptakan model yang dapat menjelaskan pola dalam data.

## 5. *Evaluation*

Setelah model dibuat, fase evaluasi dilakukan untuk menilai efektivitas model dalam memenuhi tujuan bisnis yang ditetapkan sebelumnya.

## 6. *Deployment*

Fase terakhir adalah implementasi model ke dalam sistem nyata untuk mulai digunakan dalam pengambilan keputusan.

Metodologi CRISP-DM sangat penting untuk pengambilan keputusan berbasis data karena memberikan panduan yang sistematis dalam setiap langkah proses data mining (Nodeh et al., 2019). Model ini telah digunakan di berbagai industri, termasuk keuangan, ritel, dan pariwisata (Nodeh et al., 2019). Pendekatan terstruktur dari CRISP-DM membantu organisasi mengembangkan model prediktif yang memberikan nilai tambah (Pambudi, 2023). Dengan CRISP-DM, organisasi bisa lebih efektif dalam mengelola dan menganalisis data besar, serta menghasilkan wawasan yang berharga (Pambudi, 2023).

# 5 METODE PENELITIAN

## 5.1 BAHAN PENELITIAN

Dalam melaksanakan penelitian ini, penulis akan memanfaatkan data sekunder yang diperoleh dari platform Open Data Subang. Platform ini merupakan sumber data terbuka yang disediakan oleh Pemerintah Provinsi Kabupaten Subang. Data yang digunakan dalam penelitian ini berasal dari Dinas Pariwisata, Kebudayaan, dan Olah Raga Kabupaten Subang, yang mencakup informasi-informasi penting terkait objek wisata. Informasi ini meliputi lokasi kelurahan atau kecamatan objek wisata, jenis pengunjung, jumlah pengunjung, serta fasilitas yang tersedia seperti jumlah akomodasi, restoran, dan rumah makan.

Penggunaan data sekunder dari Open Data Subang ini diharapkan dapat memberikan gambaran yang komprehensif mengenai kondisi objek wisata di Kabupaten Subang. Data tersebut akan dianalisis untuk mengidentifikasi tren dan pola kunjungan wisatawan, serta untuk menilai ketersediaan dan kualitas fasilitas pendukung pariwisata. Dengan demikian, penelitian ini dapat memberikan

rekomendasi yang lebih akurat dan relevan untuk pengembangan dan pengelolaan objek wisata di Kabupaten Subang.

## 5.2 ALAT PENELITIAN

Pada penelitian ini, penulis menggunakan alat *hardware* dan *software* untuk mendukung proses pengolahan dan analisis data secara optimal. Adapun alat yang digunakan sebagai berikut:

- 1) Perangkat Keras (*Hardware*)
  - a) Laptop Dell Inspiron 3000
  - b) RAM 6 GB
  - c) AMD Ryzen 5
  - d) Sistem Operasi Windows 11 64-bit

- 2) Perangkat Lunak (*Software*)

- a) *Google Collab*

Google Colab akan digunakan untuk mengolah dan menganalisis data. Platform ini mendukung berbagai pustaka Python dan menyediakan notebook interaktif yang memudahkan proses kerja. Dengan fasilitas ini, peneliti dapat melakukan pembersihan data, analisis statistik, dan visualisasi hasil dengan lebih efisien.

- b) Bahasa Pemrograman Python

Bahasa pemrograman Python akan digunakan untuk mengolah dan menganalisis data. Python dipilih karena memiliki berbagai alat (pustaka) yang tersedia, seperti Pandas untuk mengatur data dan Matplotlib untuk membuat grafik. Pandas akan membantu peneliti dalam proses pengolahan data, termasuk pembersihan dan transformasi data, sementara Matplotlib akan digunakan untuk visualisasi data, memungkinkan peneliti untuk melihat pola dan informasi penting dari data yang ada. Dengan memanfaatkan pustaka-pustaka ini, peneliti dapat melakukan analisis data secara mendalam dan menghasilkan visualisasi yang informatif.



### **5.3 JALAN PENELITIAN**

Dalam penelitian ini, penulis akan mengikuti beberapa langkah penting untuk mencapai tujuan yang diinginkan, seperti:

1. **Kajian teori**

Pada tahap ini, peneliti akan melakukan kajian teori untuk memahami konsep-konsep dan prinsip-prinsip yang relevan dengan topik penelitian. Kajian ini mencakup studi literatur dari berbagai sumber, seperti buku, artikel, dan jurnal ilmiah. Dengan memahami teori yang ada, peneliti dapat membangun landasan yang kuat untuk analisis data yang akan dilakukan.

2. **Metodologi Penelitian**

Setelah kajian teori, peneliti akan menetapkan metodologi penelitian yang akan digunakan. Ini mencakup pemilihan metode dan pendekatan yang tepat untuk mengumpulkan dan menganalisis data. Metodologi yang jelas akan membantu peneliti dalam merancang langkah-langkah penelitian secara sistematis dan terarah.

3. **Pengumpulan Data**

Pada tahap ini, peneliti akan mengumpulkan data yang diperlukan untuk mendukung analisis. Pengumpulan data harus dilakukan dengan hati-hati agar informasi yang diperoleh berkualitas baik dan dapat diandalkan.

4. **Data Mining**

Pada tahap data mining, peneliti akan menerapkan metode CRISP-DM untuk menganalisis data yang telah dikumpulkan dan dipersiapkan sebelumnya. Proses ini dimulai dengan pemahaman yang mendalam tentang data yang akan dianalisis, serta tujuan dari analisis tersebut. Selain itu juga penulis akan mengimplementasikan algoritma K-Means untuk mengelompokkan data ke dalam beberapa kluster berdasarkan kesamaan karakteristiknya.

5. **Kesimpulan**

Peneliti akan menyusun kesimpulan berdasarkan temuan-temuan yang diperoleh. Kesimpulan ini akan memberikan gambaran jelas tentang hasil penelitian dan menjawab pertanyaan penelitian yang telah diajukan.

sebelumnya. Selain itu, kesimpulan juga dapat mencakup rekomendasi untuk penelitian lebih lanjut atau aplikasi praktis dari hasil penelitian.

#### 6. Penyelesaian laporan

Tahap terakhir adalah penyusunan laporan penelitian. Peneliti akan mendokumentasikan seluruh proses penelitian, mulai dari kajian teori hingga kesimpulan, dalam bentuk laporan yang sistematis dan terstruktur. Laporan ini penting untuk menyampaikan hasil penelitian kepada pihak-pihak yang berkepentingan dan memberikan kontribusi pada pengembangan ilmu pengetahuan di bidang yang diteliti.

### 6 JADWAL PENELITIAN

Penelitian dilaksanakan mulai semester genap tahun akademik 2025/2026 dengan tabel waktu sebagai berikut:

No	Deskripsi Kegiatan	2025						
		Jan	Feb	Mar	Apr	Mei	Jun	Jul
1	Kajian Teori	■	■					
2	Metodologi Penelitian		■	■				
3	Pengupulan Data			■	■			
4	Data mining				■	■		
5	Kesimpulan						■	
6	Penyelesaian Laporan						■	■

## DAFTAR PUSTAKA

- Abidin, N. A. S. Z., Avila, R. D., Hermatyar, A., & Rismayani, R. (2022). Perbandingan Algoritma K-Means dan K-Medoids untuk Pengelompokan Daerah Produksi Kakao. *Jurnal Teknik Informatika dan Sistem Informasi*, 8(2), 383-391.
- Afrianto, Y. (2023). Penerapan metode CRISP-DM dalam klasifikasi data ulasan pengunjung destinasi Danau Toba. *Jurnal Universitas Katolik Indonesia Atma Jaya*.
- Astuti, D., & Surya, A. (2023). Penerapan model CRISP-DM dalam analisis data wisatawan di Bali. *Jurnal Teknologi Informasi dan Pariwisata*.
- Surya, A., & Astuti, D. (2023). Pengelompokan data wisatawan berdasarkan preferensi dan perilaku menggunakan model CRISP-DM. *Jurnal Pariwisata dan Teknologi Informasi*.
- Al-Fahmi, B., Mahardhika, B., Rahmawati, E., & Sagirani, T. (2023). Penerapan K-Means clustering pada pariwisata Kabupaten Bojonegoro untuk mendukung keputusan strategi pemasaran. *Jurnal Nasional Teknologi dan Sistem Informasi*.
- Annisa, K., Ginting, B. S., & Syari, M. A. (2022). Penerapan Data Mining Pengelompokan Data Pengguna Air Bersih Berdasarkan Keluhannya Menggunakan Metode Clustering Pada PDAM Langkat. *ALGORITMA: Jurnal Ilmu Komputer dan Informatika*, 6(1).
- Aryani, Y., & Wijayanto, A. W. (2021). Klasifikasi Pengembalian Radar dari Ionosfer Menggunakan SVM, Naive Bayes dan Random Forest. *Komputika: Jurnal Sistem Komputer*, 10(2), 111-117.
- Fahrudin, N.F. (2019). Penerapan Algoritma Apriori untuk Market Basket Analysis. *MIND Journal*.
- Farida, F., Djuhardi, L., Trisnawati, N., Mahanani, E., Marhalinda, & Rahayu, E. A. (2023). Identifikasi dan strategi pengembangan potensi desa wisata alternatif di Desa Pringkasap, Kecamatan Pabuaran, Kabupaten Subang. *IKRAITH-ABDIMAS*, 2(1), 1-20.
- Gustin, G.M., Umam, M.K., Khatomy, H., Karantina, T., & Syukur, A. (2021).

- Pengembangan Potensi Wisata Untuk Meningkatkan Daya Tarik Wisatawan di Desa Ketapang Raya Kecamatan Keruak Kabupaten Lombok Timur. *Jurnal Pengabdian Magister Pendidikan IPA*.
- Helia, A., Mulyawan, M., Rohmat, C. L., & Fathurrohman. (2023). Analisis pengelompokan daya tarik obyek wisata berdasarkan jenisnya menggunakan metode K-Means pada data Provinsi Jawa Barat. *Jurnal Informatika Sunan Kalijaga*.
- Hidayati, N., Prabowo, A., & Setiawan, B. (2021). Penerapan Algoritma K-Means dalam Pengelompokan Objek Wisata di Jawa Barat. *Jurnal Teknologi dan Sistem Informasi*, 5(2), 15-27.
- Kamila, I., Khairunnisa, U., & Mustakim, M. (2019). Perbandingan Algoritma K-Means dan K-Medoids untuk Pengelompokan Data Transaksi Bongkar Muat di Provinsi Riau. *Jurnal Ilmiah Rekayasa dan Manajemen Sistem Informasi*, 5(1), 119-125.
- Krisdina, H., Hadiwijoyo, J., Jimura, T., Munir, M., & Fitanto, N. (2021). Pengembangan desa wisata untuk meningkatkan ketahanan budaya di Kabupaten Subang. *UPI Repository*.
- Lesmana, R. (2023). Pengembangan Objek Wisata Alam Melalui Analisa PEKA di Kabupaten Subang. *Jurnal Pemasaran Kompetitif*, 7(1), 61-68.
- Lusianah, N., Purnamasari, A. I., & Nurhakim, B. (2023). Implementasi algoritma K-Means dalam pengelompokan jumlah wisatawan akomodasi di Jawa Barat. *EBISMEN Journal*.
- Maharani, A. N. A., Sulisty, A., Kristianto, D. A., Suharyono, E., & Sudanang, E. A. (2024). Pengembangan Dark Tourism Sebagai Salah Satu Daya Tarik Wisata Tana Toraja. *PANUNTUN (Jurnal Budaya, Pariwisata, dan Ekonomi Kreatif)*, 1(1), 19-28.
- Mardiani, E., Rahmansyah, N., Kurniati, I., Matondang, N., Tesalonika, T., Zanitha, D.A., & Romzy, I. (2023). Membandingkan Algoritma Data Mining Dengan Tools Orange untuk Social Economy. *Digital Transformation Technology*.
- Nasution, M. Z., & Hasibuan, M. S. (2020). Pendekatan Initial Centroid Search Untuk Meningkatkan Efisiensi Iterasi Klustering K-Means. *Techno. com*,

19(4).

- Nodeh, M.J., Calp, M.H., & Sahin, I. (2019). Analyzing and Processing of Supplier Database Based on the Cross-Industry Standard Process for Data Mining (CRISP-DM) Algorithm. *Artificial Intelligence and Applied Mathematics in Engineering Problems*.
- Normah, N., Nurajizah, S., & Salbinda, A. (2021). Penerapan data mining metode k-means clustering untuk analisa penjualan pada toko fashion hijab Banten. *Jurnal Khatulistiwa Informatika*, 7(2), 158-163.
- Novandi, H.R., & Rukminto Adi, I. (2021). Dampak Pengembangan Ekonomi Lokal melalui Community Based Tourism terhadap Kesejahteraan Ekonomi Masyarakat di Desa Tamansari, Banyuwangi, Jawa Timur. *Salus Cultura: Jurnal Pembangunan Manusia dan Kebudayaan*.
- Nurjanah, M., & Arifin, T. (2021). Penerapan Algoritma K-Means Untuk Analisis Data Ulasan Di Situs Tripadvisor. *Jurnal Responsif: Riset Sains dan Informatika*, 3(1), 75-82.
- Orisa, M. (2022). Optimasi Cluster pada Algoritma K-Means. *Prosiding SENIATI*.
- Prakoso, A.B., & Zafrullah, A. (2020). Review Tugas Individu Perekonomian Indonesia Boosting Indonesia's Tourism Sector to be Competitive. Penulis, AZ. Tayibnapis; MS. Sundari. 130218339 Abimanyu Bagas Prakoso KP A.
- Pramono, A. R., Abdul Majid, N. W., & Fauzi, A. (2023). CRISP-DM: Metodologi proyek data science. *Jurnal Universitas Duta Bangsa*.
- Pratiwi, R., & Ruspianda, R. (2023). Tantangan Sumberdaya Manusia dalam Pariwisata Berkelanjutan (Studi Kasus: Objek Wisata Pemandian Air Panas Desa Sungai Pinang). *Seminar Nasional Pariwisata dan Kewirausahaan (SNPK)*.
- Primanty, A. (2023). Strategi Pengembangan Wisata Alam Air Terjun (Curug) di Desa Curugrendeng Kecamatan Jalancagak, Subang, Jawa Barat. *Masyarakat Pariwisata: Journal of Community Services in Tourism*, 38-45.
- Purnama, I., Saputra, R., & Wibowo, A. (2023). Implementasi data mining menggunakan CRISP-DM pada sistem informasi eksekutif Dinas Kelautan

- dan Perikanan Provinsi Jawa Tengah. Jurnal Universitas Diponegoro.
- Rachman, R., & Hunaifi, N. (2020). Penerapan Metode Algoritma Apriori dan FP-Tree Pada Penentuan Pola Pembelian Obat. Jurnal Khatulistiwa Informatika, 22(2), 175-182.
- Ramadhani, I., & Megawati, M. (2023). Implementasi Algoritma K-Means Untuk Klustering Data Produktivitas Kelapa Sawit: Implementation Of K-Means Algorithm For Palm Oil Productivity Data Clustering. Indonesian Journal of Informatic Research and Software Engineering (IJIRSE), 3(1), 56-64.
- Rosidah, Latifaturrahmah, & Wahyuni, N.T. (2024). Strategi Pengembangan Objek Wisata dan Pengelolaan Wisata Berkelanjutan Terhadap Perekonomian Masyarakat Garut. Indonesian Journal of Tourism Business and Entrepreneurship.
- Rusyidi, B., & Fedryansah, M. (2019). Pengembangan Pariwisata Berbasis Masyarakat. Focus : Jurnal Pekerjaan Sosial.
- Seimahuira, S. (2021). Implementasi datamining dalam menentukan destinasi unggulan berdasarkan online reviews tripadvisor menggunakan algoritma K-Means. Technologia: Jurnal Ilmiah, 12(1), 53-58.
- Setiono, M. H. (2022). Komparasi Algoritma Decision Tree, Random Forest, Svm Dan K-Nn Dalam Klasifikasi Kepuasan Penumpang Maskapai Penerbangan. INTI Nusa Mandiri, 17(1), 32-39.
- Singgalen, Y. A. (2023). Analisis Perilaku Wisatawan Berdasarkan Data Ulasan di Website Tripadvisor Menggunakan CRISP-DM: Wisata Minat Khusus Pendakian Gunung Rinjani dan Gunung Bromo. Journal of Computer System and Informatics (JoSYC), 4(2), 326-338.
- Sulistyo, A., Nurhaji, A., Yudiandri, T. E., Ernawati, H., & Buchori, A. (2022). Pengembangan Wisata Alam Unggulan Dalam Menciptakan Pariwisata Berkelanjutan Di Kabupaten Subang (Studi Kasus: Curug Cileat). Kepariwisata: Jurnal Ilmiah, 17(1), 1-17.
- Tantowi, A.A. (2021). Literasi Ringkas Data Mining. Journal of Academia Perspectives.
- Umargono, E., Suseno, J.E., & Vincensius Gunawan, S. (2020). K-Means

Clustering Optimization Using the Elbow Method and Early Centroid Determination Based on Mean and Median Formula. Proceedings of the 2nd International Seminar on Science and Technology (ISSTEC 2019).

Wijaya, H. O. L., Yuliansyah, C., & Armanto, A. (2022). Implementasi Asosiasi Rule Mining pada Data Transaksi Penjualan Menggunakan Algoritma Apriori. *Technologia: Jurnal Ilmiah*, 13(1), 30-35.