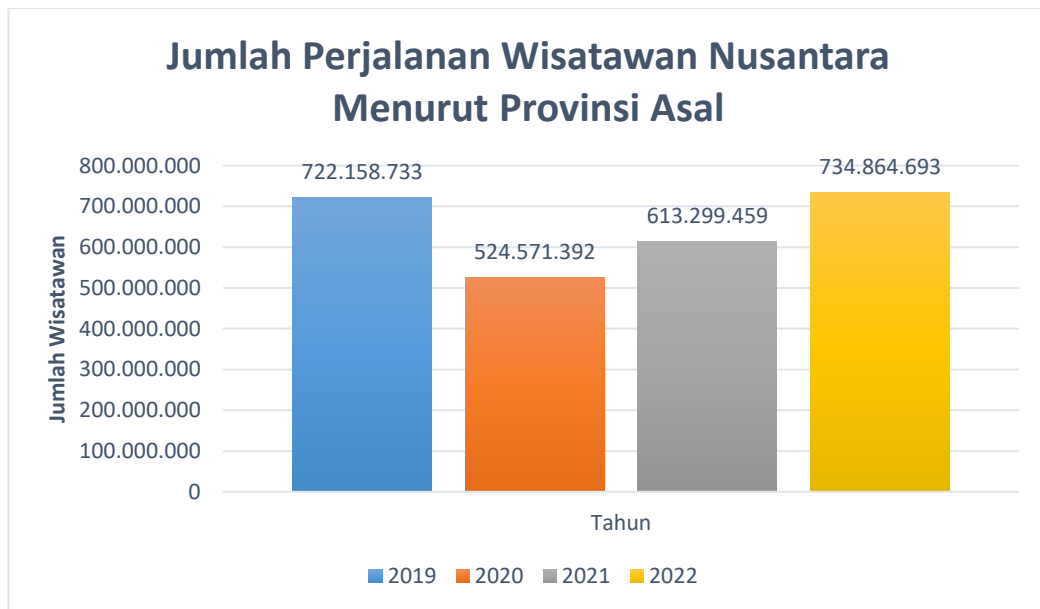


# BAB I PENDAHULUAN

## I.1 Latar Belakang

Pariwisata di Indonesia yang didukung oleh keragaman budaya dan kekayaan alam menjadi daya tarik bagi wisatawan domestik maupun mancanegara. Namun, pandemi COVID-19 pada tahun 2020 menyebabkan penurunan signifikan dalam perjalanan pariwisata (Tantowi, 2022). Badan Pusat Statistik (BPS) mencatat penurunan sebesar 27,36% pada tahun 2020 dibandingkan dengan tahun sebelumnya. Akan tetapi, terdapat peningkatan perjalanan wisatawan pada tahun berikutnya hingga 2022 karena pelonggaran peraturan pembatasan mobilitas masyarakat yang meningkatkan potensi perjalanan wisata.

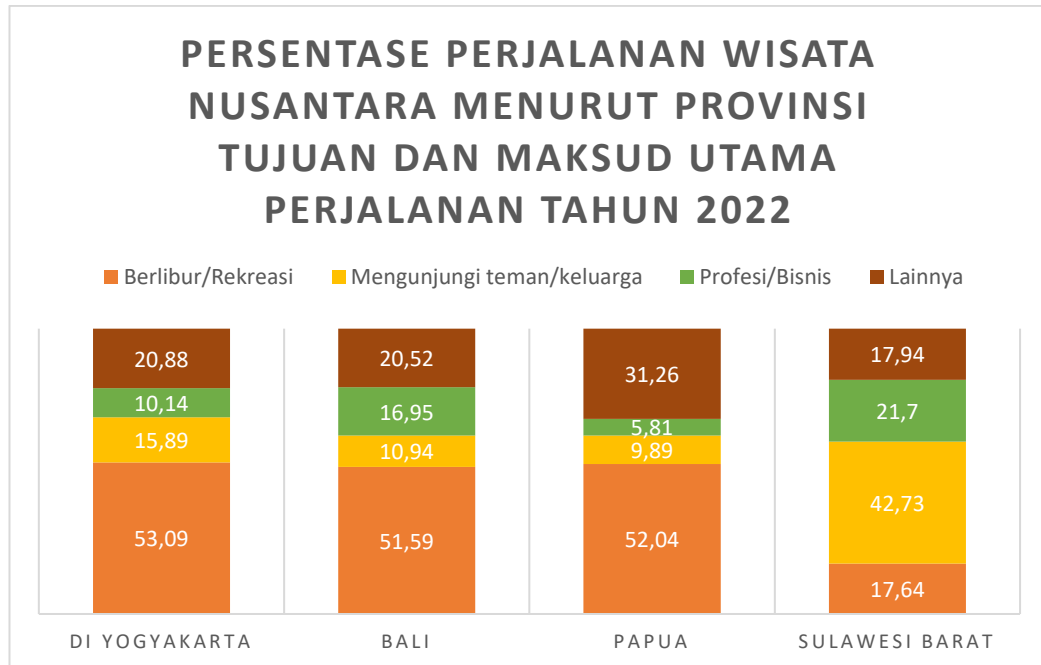


Gambar I.1 Jumlah Perjalanan Wisatawan Nusantara pada tahun 2019-2022

Sumber (Badan Pusat Statistik, 2023)

Pariwisata bukan hanya menjadi sektor utama pendukung ekonomi di tingkat nasional, tetapi juga memiliki peran penting dalam pertumbuhan ekonomi di tingkat regional. Salah satu provinsi yang paling populer untuk menjadi destinasi wisata pada masa kini yaitu Daerah Istimewa Yogyakarta. Berdasarkan data BPS pada tahun 2023, Yogyakarta merupakan destinasi favorit untuk berlibur ataupun rekreasi pada tahun 2022. Hal tersebut didasarkan pada data persentase perjalanan wisatawan nusantara menurut provinsi tujuan dan maksud utama perjalanan.

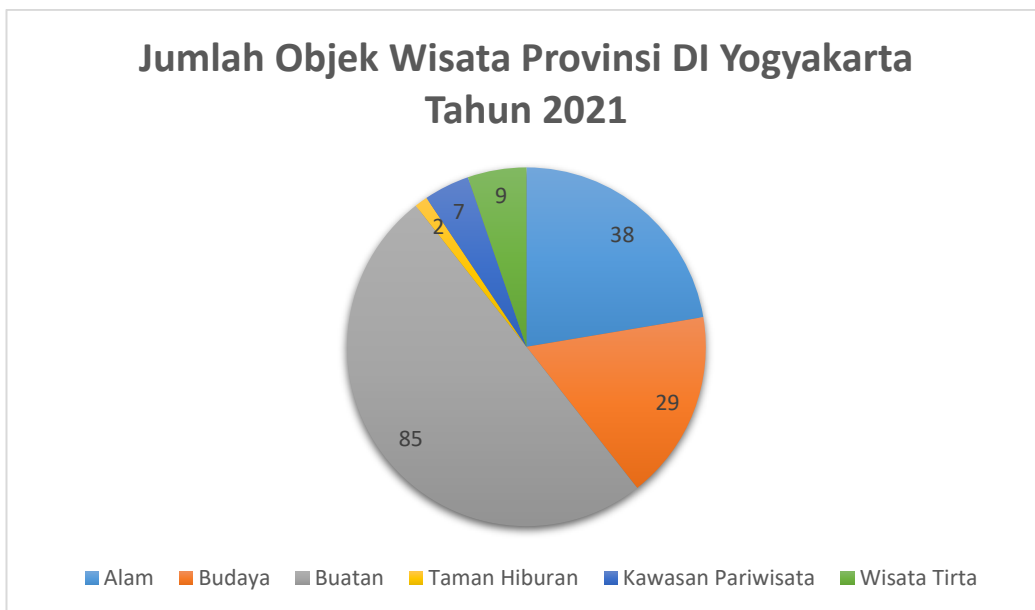
Yogyakarta memiliki persentase tujuan utama perjalanan wisata untuk berlibur/rekreasi sebesar 53,09%, mengalahkan Bali yang memiliki persentase sebesar 51,59%.



Gambar I.2 Persentase Perjalanan Wisata Nusantara Menurut Provinsi Tujuan dan Maksud Utama Perjalanan tahun 2022

Sumber (Badan Pusat Statistik, 2023)

Kekayaan budaya, sejarah, seni, dan keindahan alamnya membuat Yogyakarta menjadi daya tarik bagi para wisatawan dari dalam dan luar negeri. Yogyakarta menawarkan beragam objek wisata yang mencakup Candi Prambanan, Keraton Yogyakarta, serta pemandangan alam yang menakjubkan termasuk pantai dan gunung berapi. Gambar I.3 menunjukkan jumlah persebaran objek wisata Yogyakarta pada tahun 2021.



Gambar I.3 Jumlah Objek Wisata Provinsi DI Yogyakarta tahun 2021

Sumber (Badan Pusat Statistik, 2023)

Sebagaimana halnya dengan sektor pariwisata di Indonesia, Yogyakarta juga mengalami dampak signifikan akibat pandemi COVID-19 yang melanda dunia pada tahun 2020. Hal tersebut menyebabkan kunjungan wisatawan ke objek wisata di Yogyakarta mengalami penurunan yang signifikan selama masa pandemi. Penurunan kunjungan wisatawan tersebut dapat dilihat sebagai cerminan dari penurunan perjalanan wisatawan di seluruh Indonesia, seperti yang telah disebutkan sebelumnya.

Media sosial telah mengubah cara orang mencari informasi, berbagi pengalaman, dan mendapatkan hiburan. Media sosial yang populer di masa kini yaitu YouTube, TikTok, dan lain sebagainya. YouTube merupakan media sosial untuk berbagi video secara online yang telah didirikan sejak tahun 2005. Sementara itu, TikTok merupakan media sosial yang difokuskan pada pengguna *mobile* untuk berbagi video berdurasi pendek. TikTok telah berdiri sejak tahun 2012, tetapi mulai populer di Indonesia pada tahun 2020 (CNN Indonesia, 2020). Berdasarkan sumber Statista pada tahun 2024, Indonesia menempati posisi pertama sebagai pengguna TikTok terbanyak di dunia dengan estimasi sebanyak 127,5 juta pengguna.



Gambar I.4 Jumlah Pengguna TikTok Berdasarkan Negara Pada Juni 2024  
Sumber (Statista, 2024)

Begitu populernya TikTok di Indonesia tentunya memberikan banyak dampak terhadap aktivitas sosial masyarakat. Tren yang sedang banyak dilakukan oleh masyarakat pada saat ini yaitu membuat konten video. Konten video yang dibuat beberapa di antaranya yaitu video kuliner/makanan, film, produk komersial, dan tentunya tempat pariwisata yang sangat menarik perhatian pengguna lainnya. TikTok juga digunakan untuk mendukung pariwisata dengan membuat konten yang strategis (Maesaroh, et al., 2023). Tidak hanya TikTok, YouTube juga memiliki konten-konten yang bermanfaat untuk meningkatkan dan mempromosikan pariwisata (Irkhos & Risdianto, 2022).

Dengan populernya pengguna media sosial seperti TikTok dan YouTube, serta adanya statistik yang menunjukkan Yogyakarta merupakan destinasi berlibur/rekreasi paling favorit, banyak orang telah membuat berbagai konten video destinasi wisata yang ada di Yogyakarta. Pada saat yang sama, tentunya banyak masyarakat yang mencari rekomendasi untuk tujuan berlibur/rekreasi mereka di Yogyakarta melalui konten video yang ada di kanal media sosial seperti YouTube (Afdhaniar, 2021) dan TikTok (Widiawanti, Anom, & Iswadi, 2023). Dari konten tersebut, tentunya penonton juga memberi komentar terhadap video

destinasi wisata tersebut. Komentar tersebut dapat berupa opini atau pengalaman pribadi mereka terhadap wisata di konten tersebut.



Gambar I.5 Video TikTok Tentang Wisata Yogyakarta  
Sumber (alvaro.rival, 2023)

Jumlah komentar dalam satu konten video dapat berjumlah ratusan bahkan ribuan. Dengan begitu banyaknya komentar, penonton menjadi kesusahan untuk mengetahui gambaran umum dari pendapat penonton lainnya (Moalla, Nabli, Bouzguenda, & Hammami, 2016). Hal ini dapat menyebabkan penonton kehilangan informasi penting, bahkan merasa bingung dan tidak puas dengan konten video tersebut. Penonton tidak ingin membangun pola pikirnya hanya berdasarkan video, tetapi juga ingin mengetahui bagaimana pendapat penonton lainnya terhadap video tersebut melalui komentar mereka. Dikarenakan komentar

video yang cenderung berjumlah sangat banyak, maka akan susah untuk mengetahui secara praktis kecenderungan komentar penonton terhadap video tersebut.

Untuk memahami kecenderungan opini masyarakat pada komentar video di YouTube maupun TikTok, dapat dilakukan suatu analisis berupa analisis sentimen. Dari analisis tersebut, dapat diketahui kecenderungan dari sentimen tersebut bersifat positif, negatif, ataupun netral. Analisis sentimen dapat membantu penonton untuk mendapatkan gambaran yang lebih akurat, objektif, dan komprehensif dari pendapat penonton lainnya. Analisis sentimen juga dapat membantu penonton untuk menemukan komentar yang paling relevan, informatif, atau menarik bagi mereka. Analisis sentimen merupakan solusi untuk mengetahui kecenderungan opini masyarakat yang ditulis di media sosial terhadap pariwisata. Hasil dari analisis tersebut yaitu opini-opini tersebut akan diklasifikasikan berdasarkan sifatnya, yaitu positif, negatif, dan netral.

Terdapat penelitian terdahulu yang membandingkan beberapa algoritma klasifikasi *machine learning* untuk melakukan analisis sentimen terhadap wisata kuliner di Bandung dengan judul “Penerapan Analisis Sentimen Melalui Data Instagram Untuk Mengetahui Reputasi Wisata Kuliner di Kota Bandung Menggunakan Metode Klasifikasi Naïve Bayes.” Hasil dari penelitian ini yaitu algoritma *Naïve Bayes* memiliki akurasi yang paling tinggi sebesar 86.67% jika dibandingkan dengan algoritma *decision-tree* dan *K-NN*.

Penelitian tersebut telah menyelesaikan permasalahan banyaknya pendapat masyarakat terhadap kuliner di Kota Bandung menggunakan analisis sentimen. Akan tetapi, penelitian ini masih memiliki beberapa kekurangan. Salah satunya yaitu analisis sentimen hanya mampu mengenali kecenderungan positif, negatif, atau netral, tanpa menyajikan pemahaman mendalam tentang variabilitas dalam opini tersebut. Dengan demikian, penilaian sentimen tidak mempertimbangkan seperti aspek-aspek lain seperti waktu, lokasi, dan jenis wisata. Terlebih lagi, jika data berasal dari lebih dari satu sumber maka pengelolaan data analisis sentimen akan menjadi sulit dikarenakan perbedaan format data antar sumber. Dengan

demikian, perlu dilakukan pengembangan dari penelitian tersebut supaya bisa menjadi lebih bermanfaat bagi *stakeholders* terkait.

Dalam perkembangan teknologi informasi saat ini, fenomena *data warehouse* mencerminkan evolusi signifikan dalam pengelolaan dan analisis data. *Data warehouse* tidak lagi hanya dianggap sebagai gudang penyimpanan data, melainkan sebagai pusat analisis yang kritis untuk mendukung pengambilan keputusan. Teknologi ini menawarkan kemampuan untuk mengintegrasikan, menyajikan, dan menganalisis data secara komprehensif, memberikan wawasan yang diperlukan untuk berbagai disiplin ilmu. Fenomena ini berkorelasi dengan pertumbuhan eksponensial data digital. Selain itu, organisasi dan peneliti semakin menggantungkan *data warehouse* untuk mengatasi kompleksitas dan memanfaatkan potensi penuh data yang mereka miliki.

Terdapat penelitian terdahulu yang membahas pendekatan *data warehouse* berdasarkan data media sosial dengan judul “Data warehouse design from social media for opinion analysis: the case of Facebook and Twitter.” Penelitian ini menyatakan perkembangan dan peningkatan jumlah data media sosial membuat analisis informasi menjadi sulit sehingga pendekatan *data warehouse* perlu dilakukan. Penelitian ini mengajukan langkah-langkah perancangan *data warehouse* dari data media sosial Twitter dan Facebook dan menghasilkan skema multidimensional dari data media sosial tersebut.

Maka dari itu, implementasi *data warehouse* dapat mendukung penelitian ini karena data media sosial yang berukuran besar dan memiliki lebih dari satu sumber. Pendekatan *data warehouse* menjadi sangat penting ketika menangani aspek-aspek atau dimensi tertentu dalam penelitian ini. Pertama, aspek waktu menjadi krusial untuk memahami bagaimana sentimen masyarakat terhadap destinasi wisata berubah seiring waktu. Dengan *data warehouse*, dapat dilakukan analisis tren sentimen secara kronologis, mengidentifikasi perubahan dalam pandangan masyarakat terhadap destinasi wisata dari waktu ke waktu.

Aspek lokasi destinasi juga dapat dianalisis secara lebih mendalam dengan menggunakan *data warehouse*. Melalui integrasi data dari berbagai sumber, *data warehouse* memungkinkan pemahaman yang lebih baik tentang perbedaan

sentimen antar lokasi destinasi wisata. Misalnya, apakah opini positif lebih dominan di area tertentu atau bagaimana variabilitas sentimen dapat terkait dengan karakteristik geografis tertentu. Selain itu, jenis wisata juga dapat dijadikan dimensi penting dalam analisis sentimen. *Data warehouse* memungkinkan pengelompokan dan pemodelan berdasarkan jenis wisata. Hal ini memungkinkan pengguna untuk mengeksplorasi bagaimana sentimen berbeda-beda tergantung pada jenis pengalaman wisata tertentu. Dengan demikian, hasil dari penelitian ini diharap dapat menghasilkan wawasan atau *insight* dari opini masyarakat pada YouTube dan TikTok terhadap wisata yang ada di daerah Yogyakarta sehingga membantu pemangku kepentingan seperti wisatawan dan pengelola wisata untuk mengambil keputusan berdasarkan data yang ada (*data-driven decision making*).

## **I.2 Perumusan Masalah**

Rumusan masalah yang mendasari penelitian ini yaitu:

1. Bagaimana skema multidimensional *data warehouse* yang dapat mengintegrasikan data komentar YouTube dan TikTok terhadap wisata DIY?
2. Bagaimana integrasi model analisis sentimen dengan sistem *data warehouse* untuk data wisata dari TikTok dan YouTube terkait destinasi wisata di DIY?
3. Apa saja wawasan atau *insight* yang dapat diperoleh dari analisis sentimen komentar YouTube dan TikTok terkait destinasi wisata di DIY berdasarkan faktor waktu, lokasi, jenis wisata, dan media sosial?

## **I.3 Tujuan Penelitian**

Dengan masalah yang telah dirumuskan, penelitian ini dilakukan dengan tujuan sebagai berikut:

1. Membangun skema *data warehouse* yang dapat mengintegrasikan dan menyatukan data komentar dari platform YouTube dan TikTok terkait destinasi wisata di Daerah Istimewa Yogyakarta (DIY).
2. Mengintegrasikan model analisis sentimen ke dalam sistem *data warehouse* untuk menganalisis dan memahami sentimen dari komentar-



komentar terkait destinasi wisata DIY yang berasal dari platform TikTok dan YouTube.

3. Mengidentifikasi wawasan atau insight yang dapat diperoleh dari analisis sentimen terhadap komentar-komentar di YouTube dan TikTok mengenai destinasi wisata DIY dengan mempertimbangkan seperti faktor waktu, lokasi, dan media sosial sebagai variabel yang relevan.

#### **I.4 Batasan Penelitian**

Berikut ini merupakan batasan-batasan dari penelitian yang akan dilakukan:

1. Data sentimen yang digunakan bersumber dari fitur komentar pada media sosial YouTube dan TikTok.
2. Data wisata yang digunakan bersumber dari Google Maps.
3. Analisis sentimen yang dilakukan hanya untuk mengklasifikasikan sentimen positif, negatif, dan netral.
4. Algoritma yang digunakan untuk melakukan analisis sentimen dibatasi pada algoritma *Naïve Bayes Classifier*.
5. Analisis dilakukan pada komentar yang dipublikasikan pada tahun 2019 hingga 2024.

#### **I.5 Manfaat Penelitian**

Manfaat yang diharapkan dari penelitian ini yaitu:

1. Memberikan evaluasi berharga bagi pengelola wisata. Dengan menganalisis masukan dan opini masyarakat terhadap pariwisata, pengelola dapat mengambil keputusan yang lebih baik dan meningkatkan kualitas layanan. Selain itu, analisis data berdasarkan waktu, kategori wisata, dan daerah wisata akan membantu pengelola memahami tren dan kebutuhan pengunjung dengan lebih baik. Hasil dari penelitian ini dapat digunakan untuk merancang strategi pemasaran yang lebih efektif dan meningkatkan kepuasan pengunjung.
2. Meningkatkan kualitas pengalaman berwisata untuk calon wisatawan dengan memberikan informasi dan kondisi wisata yang dituju berdasarkan data analisis opini masyarakat. Wisatawan dapat memperoleh wawasan yang lebih akurat tentang destinasi wisata yang akan dikunjungi, termasuk

sentimen umum pengunjung lain terhadap destinasi tersebut. Informasi ini membantu wisatawan dalam membuat keputusan yang lebih baik tentang tempat yang ingin dikunjungi.

## **I.6 Sistematika Penulisan**

Penelitian ini dikerjakan dengan sistematika penulisan sebagai berikut:

- **Bab I Pendahuluan**  
Bab pertama berisi uraian mengenai konteks permasalahan, latar belakang penelitian, perumusan masalah, tujuan penelitian, batasan penelitian, manfaat penelitian, dan sistematika penulisan.
- **Bab II Tinjauan Pustaka**  
Tinjauan pustaka membahas literatur yang relevan dengan permasalahan yang diteliti serta hasil-hasil penelitian terdahulu. Bab ini juga mencakup metodologi/metode/kerangka kerja yang digunakan untuk menyelesaikan permasalahan atau mengurangi gap antara kondisi eksisting dengan target, dan diakhiri dengan analisis pemilihan metodologi/metode/kerangka kerja yang akan digunakan dalam penelitian ini.
- **Bab III Metodologi Penelitian**  
Bab ini menguraikan strategi dan langkah-langkah yang dilakukan dalam penelitian untuk menjawab rumusan masalah. Pada bab ini dijelaskan langkah-langkah penelitian secara rinci meliputi perumusan masalah, pengembangan model penelitian, identifikasi dan operasionalisasi variabel penelitian, perancangan pengumpulan dan pengolahan data, uji instrumen, serta perancangan analisis data.
- **Bab IV Perancangan dan Desain Sistem**  
Bab ini menyajikan perancangan sistem atau model yang akan dibangun, termasuk perancangan tabel, arsitektur sistem, dan proses-proses yang terlibat dalam pengembangan sistem.
- **Bab V Hasil dan Evaluasi**  
Bab ini menyajikan hasil rancangan, temuan, analisis dan pengolahan data. Selain itu, bab ini juga berisi tentang validasi atau verifikasi hasil penelitian untuk memastikan bahwa hasil tersebut dapat menyelesaikan

masalah atau menurunkan gap antara kondisi eksisting dan target yang ingin dicapai. Metode-metode evaluasi lainnya juga dapat diterapkan untuk memvalidasi hasil penelitian sesuai dengan kebutuhan.

- Bab VI Kesimpulan dan Saran

Bab ini menjelaskan kesimpulan dari penelitian yang dilakukan serta jawaban dari pertanyaan penelitian yang disajikan di pendahuluan. Saran penelitian dikemukakan untuk penelitian selanjutnya.