

ANALISIS TREN WISATA BALI

TUGAS BESAR DATA MINING

Oleh

Difta Alghifari Aryanto	714220015
Wafid Adzka Haifan M	714220014
Muh Dany Ma'ruf Riadi	714220025



DIPLOMA IV TEKNIK INFORMATIKA
SEKOLAH VOKASI
UNIVERSITAS LOGISTIK DAN BISNIS INTERNASIONAL
BANDUNG
2025

HALAMAN PERNYATAAN ORISINALITAS

Tugas besar ini adalah hasil karya saya sendiri, dan semua sumber baik yang dikutip maupun dirujuk telah saya nyatakan dengan benar. Bilamana di kemudian hari ditemukan bahwa karya tulis ini menyalahi peraturan yang ada berkaitan etika dan kaidah penulisan karya ilmiah yang berlaku, maka saya bersedia dituntut dan diproses sesuai dengan ketentuan yang berlaku

Yang menyatakan,

Nama : Difta Alghifari Aryanto

NIM : 714220015

Tanda Tangan :

Tanggal : 10 Juli 2025

Mengetahui,

Ketua :..... (.....tanda tangan.)

Dosen Pengampu Mata Kuliah :..... (.....tanda tangan.)

KATA PENGANTAR

Puji syukur kami panjatkan kehadirat Allah SWT, karena berkat rahmat dan karunia-Nya, kami dapat menyelesaikan laporan tugas besar ini dengan baik. Tak lupa shalawat serta salam senantiasa tercurah kepada Nabi Muhammad SAW, yang merupakan suri tauladan bagi umat islam. Tugas besar ini, dengan topik “Analisis Tren Wisata Bali” bertujuan untuk mempraktikkan konsep dan keterampilan yang telah kami pelajari dalam melakukan pengolahan data.

Laporan ini disusun untuk memenuhi tugas besar mata kuliah Data Mining. Tentunya kami menyadari bahwa laporan ini masih memiliki banyak kekurangan. Oleh karena itu, kami mengucapkan terima kasih sebesar-besarnya kepada :

1. Nisa Hanum Harani, S.Kom., M.T.,CDSP, SFPC, selaku dosen mata kuliah Data Mining. Yang telah memberikan ilmu dan bimbingannya selama perkuliahan.
2. Rekan 1 kelompok atas kerja sama nya dalam menyelesaikan tugas besar ini.

Kami sadar bahwa laporan tugas besar yang kamu buat masih memiliki banyak kekurangan. Kami mengharapkan kritik dan saran yang membangun agar kami dapat membuat laporan yang lebih baik lagi di masa yang akan datang.

Bandung, Juli 2025

Penulis

HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademik Universitas Logistik Bisnis Internasional, saya yang bertanda tangan di bawah ini:

Nama : Difta Alghifari Aryanto

NIM : 714220015

demikian demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Logistik Bisnis Internasional, Hak Bebas Royalti Noneksklusif (*Non-exclusive Royalti Free Right*) atas karya ilmiah saya yang berjudul:

“Analisis Tren Wisata Bali”

Beserta perangkat yang ada (jika diperlukan). Dengan Hak ini Universitas Logistik Bisnis Internasional berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan tugas akhir saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Bandung

Pada tanggal : 10 Juli 2025

Yang menyatakan

Difta Alghifari Aryanto

Ketua Kelompok

ABSTRAK

Pariwisata merupakan salah satu industri skala global terbesar yang menunjukkan pertumbuhan yang konsisten dari tahun ke tahun. Di Indonesia, sektor ini berkontribusi signifikan terhadap perekonomian nasional, dengan peningkatan jumlah wisatawan asing dari 5,88 juta pada tahun 2022 menjadi 13,9 juta pada tahun 2024. Selain sebagai sarana rekreasi, pariwisata juga berperan strategis dalam pembangunan ekonomi, penciptaan lapangan kerja, dan pelestarian budaya. Perkembangan teknologi digital turut mengubah perilaku wisatawan dalam merencanakan dan membagikan pengalaman berwisata, sehingga memunculkan tren – tren baru yang mengharuskan pelaku industri pariwisata untuk beradaptasi. Tren seperti wisata berbasis pengalaman, personalisasi, dan keberlanjutan menuntut pendekatan strategis berbasis data. Analisis tren wisata menjadi kunci dalam memahami preferensi wisatawan dan merancang strategi pemasaran serta pengembangan destinasi yang lebih efektif dan relevan.

Kata Kunci : *Pariwisata, Tren, Regresi Linear, Ekonomi Nasional, Digitalisasi, Personalisasi*

ABSTRACT

Tourism is one of the largest global-scale industries that shows consistent growth from year to year. In Indonesia, this sector contributes significantly to the national economy, with the number of foreign tourists increasing from 5.88 million in 2022 to 13.9 million in 2024. Apart from being a means of recreation, tourism also plays a strategic role in economic development, job creation, and cultural preservation. The development of digital technology has also changed tourist behavior in planning and sharing travel experiences, giving rise to new trends that require tourism industry players to adapt. Trends such as experiential tourism, personalization, and sustainability demand a data-driven strategic approach. Travel trend analysis is key in understanding traveler preferences and designing more effective and relevant marketing and destination development strategies.

Keywords : Tourism, Trends, Linear Regression, National Economy, Digitalization, Personalization

DAFTAR ISI

HALAMAN PERNYATAAN ORISINALITAS	ii
KATA PENGANTAR	iii
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS.....	iv
ABSTRAK	v
ABSTRACT.....	vi
DAFTAR ISI	vii
DAFTAR TABEL	ix
DAFTAR GAMBAR	x
DAFTAR RUMUS	xi
DAFTAR NOTASI	xii
BAB 1 PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Tujuan Penelitian	2
1.4 Manfaat Penelitian	2
1.5 Ruang Lingkup.....	2
BAB II TINJAUAN PUSTAKA	3
2.1 Kajian Teori	3
2.2 Diagram Alur Konsep Metodologi	4
2.3 State Of The Art	4
BAB III METODOLOGI PENELITIAN	6
3.1 Tahapan Penelitian	6
3.2 Deskripsi Dataset	7
3.3 Algoritma	8
3.4 Evaluasi Kinerja	10
BAB 4 Hasil dan Pembahasan	11
4.1 Visualisasi Eksploratif (EDA)	11
4.2 Hasil Preprocessing dan Pemodelan	11
4.3 Tabel Hasil Eksperimen	12
4.4 Interpretasi Hasil	12
4.5 Analisis Keunggulan dan Keterbatasan	12
BAB 5	13
Kesimpulan dan Saran	13

5.1 Ringkasan Temuan Utama	13
5.2 Jawaban Atas Rumusan Masalah	13
5.3 Saran Untuk Pengembangan Lebih Lanjut	14
DAFTAR PUSTAKA.....	15
LAMPIRAN	16

DAFTAR TABEL

DAFTAR GAMBAR

Gambar 1 Diagram Alur Konsep Metodologi.....	4
--	---

DAFTAR RUMUS

DAFTAR NOTASI

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Pariwisata merupakan salah satu industri yang paling besar dan konsisten di dunia dengan pertumbuhan yang terus meningkat dari tahun ke tahun [1]. Menurut barometer pariwisata *UN Tourism*, pada tahun 2024 jumlah wisatawan internasional diseluruh dunia mencapai angka 1.4 miliar, yang meningkat sebesar 11% dari tahun sebelumnya. Di Indonesia, pariwisata merupakan salah satu sektor industri yang sangat penting dalam membantu perekonomian nasional. Pada tahun 2022 jumlah wisatawan asing yang ke Indonesia berjumlah 5.88 juta orang, pada tahun 2023 ini meningkat menjadi 11.6 juta, dan pada tahun 2024 mencapai 13.9 juta. Ini menunjukkan bahwa industri pariwisata di Indonesia terus mengalami peningkatan yang menandakan bahwa sektor pariwisata juga termasuk dalam salah satu sektor yang krusial dalam berkontribusi untuk pendapatan negara. Di Indonesia, pariwisata bukan hanya menjadi sarana rekreasi, tetapi juga memiliki peran strategis dalam pembangunan ekonomi, penciptaan lapangan kerja, dan pelestarian budaya lokal [2].

Kemajuan teknologi digital telah mengubah cara masyarakat mencari informasi, merencanakan perjalanan, hingga berbagi pengalaman berwisata. Media sosial, platform ulasan wisata, dan aplikasi pemesanan daring seperti Traveloka, Tiket.com, dan Agoda mendorong transparansi informasi serta kemudahan akses layanan pariwisata. Hal ini memicu perubahan preferensi wisatawan yang semakin mengarah pada wisata berbasis pengalaman, personalisasi, dan keberlanjutan [3].

Namun, hal tersebut juga menciptakan tantangan bagi pelaku industri pariwisata. Preferensi wisatawan yang cepat berubah membuat para pelaku industri harus melakukan adaptasi yang berkelanjutan terhadap tren baru. Contohnya, peningkatan minat terhadap wisata alam, wisata budaya, serta wisata berbasis komunitas (*community-based tourism*) yang menunjukkan perlunya strategi yang tepat dalam pengembangan dan promosi destinasi wisata [4].

Dalam hal ini, analisis tren wisata menjadi sangat penting untuk memahami perilaku

wisatawan. Dengan menganalisis data seperti destinasi populer, jumlah kunjungan, musim kunjungan, ulasan pengunjung, hingga interaksi digital, pelaku industri dapat merancang paket wisata yang lebih relevan dan menarik [5]. Pendekatan berbasis data (*data-driven*) akan membantu penyusunan strategi pemasaran, inovasi produk wisata, dan peningkatan pelayanan wisatawan [6].

1.2 Rumusan Masalah

1. Bagaimana tren destinasi wisata berkembang dari waktu ke waktu?
2. Faktor apa saja yang mempengaruhi popularitas suatu destinasi wisata?
3. Bagaimana agen perjalanan dapat memanfaatkan analisis tren untuk menyusun paket wisata yang lebih menarik?

1.3 Tujuan Penelitian

1. Mengetahui perkembangan tren destinasi wisata .
2. Mengetahui faktor apa saja yang mempengaruhi tren suatu tempat wisata
3. Memberikan gambaran tren pariwisata yang ada pada masyarakat untuk membantu pihak agen perjalanan menyusun paket wisata yang menarik sesuai tren

1.4 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan kontribusi dalam memahami perubahan perilaku dan preferensi wisatawan di era digital melalui analisis tren berbasis data. Dengan pemahaman ini, pelaku industri pariwisata, termasuk pemerintah, pengelola destinasi, dan pelaku usaha, dapat merumuskan strategi yang lebih tepat dalam hal pemasaran, inovasi produk wisata, serta peningkatan kualitas layanan. Selain itu, penelitian ini juga dapat menjadi referensi akademik dalam pengembangan kajian pariwisata berbasis teknologi dan data.

1.5 Ruang Lingkup

Penelitian ini berfokus pada analisis tren wisata di Indonesia khususnya di Pulau Bali dengan memanfaatkan data digital sebagai dasar pengambilan keputusan. Cakupan analisis meliputi identifikasi destinasi wisata yang paling populer berdasarkan frekuensi kunjungan, dan jumlah kunjungan wisatawan dalam kurun waktu tertentu. Melalui data tersebut, penelitian ini akan memberikan gambaran tren preferensi wisatawan dan membantu agen perjalanan dalam merancang paket wisata yang relevan, personal, dan sesuai dengan kebutuhan pasar saat ini.

BAB II

TINJAUAN PUSTAKA

2.1 Kajian Teori

2.1.1 Beberapa penelitian terkini, telah memanfaatkan penggunaan data mining untuk melakukan analisis terhadap tren wisata di masyarakat, untuk menghasilkan data tren yang lebih variatif dan akurat :

2.1.2 Sentiment Analysis dengan Deep Learning

[9] melakukan penelitian terkait pengembangan sistem klasifikasi ulasan wisata berbasis LSTM (*Long Short-Term Memory*) dan *Attention Mechanism* yang mampu mengklasifikasikan sentimen pengunjung secara akurat. Pendekatan ini membantu para pengurus destinasi wisata untuk mengetahui tanggapan pengunjung terhadap suatu destinasi melalui ulasan *online*.

2.1.3 Integrasi NLP dan Visualisasi Interaktif

[11] melakukan penelitian dalam kajian literturnya bahwa integrasi *Natural Language Processing* (NLP) dengan alat visualisasi seperti *Tableau* dan *Power BI* dapat mempercepat proses pemahaman tren ulasan wisata. Teknik ini mendukung penyusunan strategi yang berbasis data dan *visual insight*.

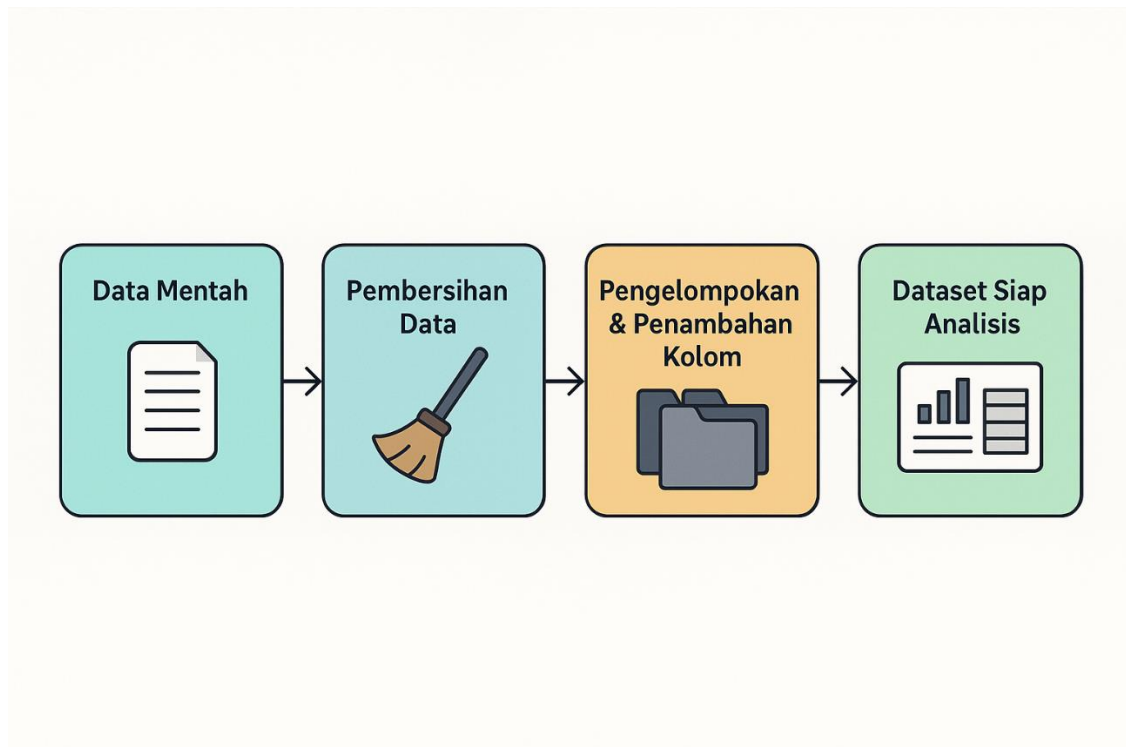
2.1.4 Deteksi Preferensi Pasca Pandemi

[2] melakukan penelitian terkait dengan preferensi wisatawan milenial di masa pandemi dan menemukan bahwa wisata alam menjadi pilihan utama. Pendekatan *data-driven* digunakan untuk memantau perubahan minat secara *real-time*, khususnya dengan melihat volume pencarian dan hasil ulasan digital.

2.1.5 Klasifikasi Opini Multibahasa

[10] melakukan penelitian menggunakan pendekatan *Bidirectional RNN* dengan *Attention* dan *Word Vector Enrichment* untuk mengklasifikasikan opini wisatawan dalam berbagai bahasa. Model ini dapat menangkap makna mendalam dari teks ulasan dan relevan untuk platform global seperti *TripAdvisor*.

2.2 Diagram Alur Konsep Metodologi



Gambar 1 Diagram Alur Konsep Metodologi

2.3 State Of The Art

Penelitian yang relevan dengan topik ini dilakukan oleh Lestari dan Gunawan (2020) dengan judul “Pengaruh Media Sosial terhadap Keputusan Berwisata di Kalangan Generasi Z” yang dipublikasikan dalam Jurnal Pariwisata Nusantara. Penelitian ini menyoroti peran media sosial sebagai salah satu faktor utama dalam membentuk keputusan berwisata, khususnya pada kalangan generasi muda. Dalam studi ini, Lestari dan Gunawan mengamati bahwa generasi Z cenderung lebih responsif terhadap konten visual seperti foto, video, dan testimoni yang dibagikan melalui platform seperti Instagram, TikTok, dan YouTube[2].

Penelitian tersebut menggunakan pendekatan kuantitatif dengan menyebarkan kuesioner kepada responden berusia 17–25 tahun yang pernah melakukan perjalanan wisata dalam satu tahun terakhir. Hasilnya menunjukkan bahwa lebih dari 70% responden mengakui bahwa mereka mendapatkan inspirasi perjalanan dari konten yang mereka lihat di media sosial. Selain itu, 62% dari mereka menjadikan ulasan dari pengguna lain sebagai referensi utama dalam menentukan destinasi yang akan dikunjungi.

Beberapa penelitian menunjukkan bahwa digitalisasi memainkan peran yang penting dalam membentuk perilaku dan preferensi wisatawan, khususnya di era pasca-pandemi. [3] membahas tentang dampak media sosial dalam mempengaruhi keputusan wisata generasi Z, dengan lebih dari 70% responden mengaku terinspirasi oleh konten visual yang mereka lihat secara daring dan 62% diantaranya menjadikan ulasan pengguna sebagai referensi utama. Hasil ini memperkuat pandangan bahwa ulasan digital tidak hanya menjadi sumber informasi, tetapi juga media promosi yang lebih dipercaya daripada iklan formal.

Hal yang sama juga dikemukakan oleh [6] yang menekankan pentingnya pemanfaatan teknologi digital dalam pemasaran wisata lokal. Mereka menunjukkan bahwa penggunaan peta digital, ulasan daring, serta platform media sosial dapat meningkatkan daya tarik destinasi, terutama bagi wisatawan domestik yang mengandalkan akses informasi secara cepat dan praktis.

Dalam konteks perubahan perilaku akibat pandemi, [4] menemukan bahwa tren perjalanan di Jawa Barat mengalami pergeseran signifikan, dengan wisatawan lebih tertarik pada destinasi terbuka dan alam. [4] menekankan bahwa data digital seperti review dan pencarian daring dapat digunakan sebagai indikator real-time untuk memantau minat wisatawan yang cepat berubah.

Sementara itu, pendekatan jangka panjang terhadap pengembangan pariwisata, disorot oleh [5] yang menekankan pentingnya informasi kunjungan dan ulasan digital dalam merancang strategi pariwisata berkelanjutan. Mereka berpendapat bahwa pemanfaatan data historis dari interaksi pengguna dapat membantu merancang pembangunan destinasi yang tidak hanya sesuai dengan tren saat ini, tetapi juga tahan terhadap perubahan jangka panjang.

Dari berbagai studi di atas, dapat disimpulkan bahwa pemanfaatan data digital, baik dalam bentuk ulasan, konten media sosial, maupun data kunjungan telah menjadi salah satu hal penting dalam memahami wisatawan. Namun demikian, masih terdapat celah dalam integrasi pendekatan data mining secara langsung, terutama dalam melakukan analisis sentimen dan prediksi tren wisata berdasarkan berbagai sumber data sekaligus. Oleh karena itu, penelitian ini bertujuan untuk mengisi kekosongan tersebut melalui pendekatan berbasis CRISP-DM yang terintegrasi dengan analisis sentimen dan visualisasi data interaktif.

BAB III

METODOLOGI PENELITIAN

3.1 Tahapan Penelitian

Penelitian ini dilaksanakan melalui beberapa tahapan yang sistematis guna mencapai tujuan penelitian secara efektif dan efisien. Adapun tahapan-tahapan penelitian yang dilakukan meliputi:

1. Pengumpulan Data yang digunakan dalam penelitian ini diperoleh dari file berformat CSV yang berisi informasi mengenai tempat-tempat wisata di berbagai wilayah. Data tersebut mencakup jumlah kunjungan wisatawan nusantara (wisnus) dan wisatawan mancanegara (wisman) serta tahun kunjungan.
2. Pra-pemrosesan Data (Preprocessing) Tahapan ini mencakup proses pembersihan data dari nilai kosong (missing values), penghapusan data duplikat, serta transformasi data ke dalam format yang sesuai untuk analisis. Proses ini juga mencakup konversi tipe data dan encoding terhadap data kategorikal.
3. Eksplorasi Data (Exploratory Data Analysis) Analisis eksplorasi dilakukan untuk memahami karakteristik data melalui statistik deskriptif dan visualisasi. Tahapan ini bertujuan untuk melihat pola-pola distribusi data, hubungan antar fitur, dan mendeteksi anomali atau outlier dalam dataset.
4. Pemodelan Data Setelah data dipersiapkan, langkah selanjutnya adalah membangun model analisis menggunakan algoritma yang sesuai. Pemodelan dilakukan untuk mencapai tujuan penelitian, seperti pengelompokan data menggunakan algoritma K-Means atau klasifikasi menggunakan algoritma Decision Tree.
5. Evaluasi Model yang telah dibangun dievaluasi kinerjanya menggunakan metrik evaluasi yang sesuai dengan jenis algoritma yang digunakan. Evaluasi bertujuan untuk menilai sejauh mana model dapat menghasilkan hasil yang akurat dan relevan dengan tujuan analisis.

Melalui tahapan-tahapan tersebut, penelitian ini diharapkan dapat menghasilkan proses analisis data yang sistematis dan terukur. Setiap langkah yang dilakukan saling berkaitan dan membentuk alur kerja yang runtut, mulai dari pengumpulan data hingga interpretasi hasil. Dengan demikian, proses penelitian ini tidak hanya berfokus pada hasil akhir, tetapi juga pada validitas dan akurasi setiap tahapan yang dilalui. Tahapan selanjutnya akan dijelaskan lebih rinci melalui uraian deskripsi dataset yang digunakan sebagai dasar dalam penelitian ini.

3.2 Deskripsi Dataset

Penelitian ini menggunakan dataset yang berisi informasi mengenai berbagai tempat wisata yang tersebar di sejumlah wilayah administrative. Dataset tersebut terdiri dari 1058 baris data, yang masing-masing barisnya mewakili satu entitas tempat wisata. Setiap entri dilengkapi dengan 5 atribut atau fitur yang memberikan gambaran komprehensif mengenai karakteristik dari tempat wisata tersebut. Atribut-atribut tersebut meliputi identifikasi unik (id), nama tempat wisata (nama), jumlah kunjungan wisatawan nusantara (wisnus), jumlah kunjungan wisatawan mancanegara (wisman) dan tahun kunjungan.

3.3 Algoritma

Penelitian ini menggunakan algoritma Linear Regression untuk menganalisis tren kunjungan wisatawan pada masing-masing tempat wisata dari tahun ke tahun. Algoritma ini termasuk dalam metode supervised learning dan digunakan untuk memodelkan hubungan antara variabel bebas (tahun) dan variabel target (jumlah pengunjung).

Linear Regression atau regresi linier merupakan metode statistik yang memodelkan hubungan linier antara variabel independen dan dependen. Dalam konteks penelitian ini, variabel independen adalah tahun, sedangkan variabel dependen adalah jumlah pengunjung wisata, baik wisatawan nusantara (Wisnus) maupun wisatawan mancanegara (Wisman).

Model regresi linear memiliki bentuk umum:

$$y = aX + b$$

dengan:

- y : jumlah pengunjung (Wisnus/Wisman)
- X : tahun
- a : *slope* (kemiringan garis regresi)
- b : *intercept* (titik potong terhadap sumbu y)

Untuk setiap tempat wisata, dilakukan pelatihan model regresi linier menggunakan data historis yang telah difilter. *Koefisien slope (a)* dari masing-masing model digunakan untuk menentukan arah tren sebagai berikut:

- Jika $a > 100$, maka tren dianggap Naik
- Jika $a < -100$, maka tren dianggap Turun
- Jika $-100 \leq a \leq 100$, maka tren dianggap Stabil

Pendekatan ini dipilih karena sifat linier dari perubahan jumlah pengunjung dari tahun ke tahun yang relatif dapat direpresentasikan secara sederhana namun cukup efektif menggunakan garis lurus. Implementasi algoritma dilakukan melalui fungsi *analisis_tren_regresi()* yang secara otomatis menghitung slope untuk setiap tempat wisata berdasarkan data tahunan dan kemudian mengklasifikasikan tren berdasarkan ambang batas (*threshold*). Fungsi ini dipanggil dua kali: untuk wisatawan nusantara dan wisatawan mancanegara, kemudian disimpan ke dalam file *hasil_tren_wisnus.csv* dan *hasil_tren_wisman.csv*.

3.4 Evaluasi Kinerja

Evaluasi kinerja dalam konteks regresi linier ini tidak difokuskan pada akurasi prediksi numerik, melainkan pada validitas arah tren berdasarkan *slope* hasil pelatihan model. Evaluasi dilakukan melalui tiga pendekatan berikut:

a. Interpretasi Koefisien (*Slope*)

Nilai *slope* dari hasil pelatihan model menunjukkan kecepatan dan arah perubahan jumlah pengunjung dari tahun ke tahun:

- *Slope* positif besar menunjukkan peningkatan tajam.
- *Slope* negatif besar menunjukkan penurunan tajam.
- *Slope* mendekati nol menunjukkan tren yang cenderung stagnan.

b. Threshold-based Trend Classification

Evaluasi klasifikasi tren dilakukan menggunakan pendekatan *threshold* sebesar ± 100 . Threshold ini ditetapkan berdasarkan observasi terhadap variasi umum jumlah pengunjung dan diuji pada beberapa sampel tempat wisata. Jika *slope* melebihi *threshold* tersebut, maka perubahan tren dianggap signifikan.

c. Visualisasi Data dan Model

Sebagai bentuk validasi tambahan, dilakukan visualisasi regresi linier pada beberapa tempat wisata menggunakan scatter plot dan garis prediksi:

- *Titik biru menunjukkan data aktual dari tahun ke tahun.*
- *Garis merah menunjukkan garis regresi hasil model.*

Visualisasi ini membantu memastikan bahwa klasifikasi tren sesuai dengan pola data aktual. Visualisasi dibuat dengan fungsi `plot_regresi_tempat()` untuk memverifikasi hasil regresi secara visual.

BAB 4

Hasil dan Pembahasan

4.1 Visualisasi Eksploratif (EDA)

Setelah dilakukan EDA, didapatkan beberapa hasil, yaitu sebagai berikut :

- **Statistik Deskriptif:** Analisis awal menggunakan `.describe()` menunjukkan adanya rentang nilai yang sangat besar pada fitur jumlah pengunjung (Wisnus & Wisman). Terlihat jurang pemisah yang signifikan antara nilai kuartil atas (75%) dan nilai maksimum (max), yang merupakan indikasi awal adanya nilai ekstrem.
- **Distribusi Data Sangat Miring (Skewed):** Visualisasi **histogram** menunjukkan distribusi yang tidak simetris. Mayoritas data terkonsentrasi pada nilai-nilai kecil (terutama nol), yang kemungkinan besar disebabkan oleh dampak pandemi COVID-19. Tampilan grafik yang menumpuk di satu sisi ini secara visual mengonfirmasi adanya outlier.
- **Kesimpulan EDA:** Data tidak terdistribusi normal dan memiliki outlier signifikan. Hal ini mengindikasikan bahwa model linear kemungkinan akan kesulitan, dan diperlukan preprocessing yang cermat sebelum pemodelan.

4.2 Hasil Preprocessing dan Pemodelan

Preprocessing :

- **Seleksi Fitur:** Fitur ID dan kolom yang tidak relevan lainnya dibuang. Fitur yang digunakan adalah Wilayah, Tempat Wisata, Tahun, Wisnus, dan Wisman.
- **Encoding Fitur Kategorikal:** Fitur Wilayah dan Tempat Wisata diubah menjadi format numerik menggunakan One-Hot Encoding agar dapat diproses oleh model.
- **Pembagian Data:** Dataset dibagi menjadi data latih (...) dan data uji (...) untuk melatih dan mengevaluasi model secara objektif.
- **Penskalaan Fitur:** Fitur numerik (Tahun dan Wisnus) diskalakan menggunakan StandardScaler untuk menyamakan rentang nilainya, yang penting untuk performa model seperti Regresi Linear.

4.3 Tabel Hasil Eksperimen

Setelah dilakukan eksperimen, evaluasi model menggunakan R2-Score, MAE, dan RMSE untuk model regresi linear, dan didapatkan hasil sebagai berikut:

Nama Model	R2-Score	MAE	RMSE
Regresi Linear Wisnus	0.94	801874.88	953916.20
Regresi Linear Wisman	0.85	1312224.76	1470743.56

4.4 Interpretasi Hasil

Berdasarkan hasil evaluasi model, dapat disimpulkan sebagai berikut :

1. Model dapat memprediksi dengan baik untuk kedua variabel (Wisnus & Wisman), dengan hasil akhir prediksi model terhadap wisnus lebih unggul dibandingkan wisman, ini menunjukkan bahwa model cukup akurat dalam memprediksi jumlah total wisnus (Wisatawan Nusantara), yang dapat dibuktikan dengan nilai R2-Score yang cukup tinggi yaitu 0.95 dan nilai MAE (*Mean Squared Error*) yang lebih rendah dibandingkan wisman, yaitu 801874.88 lebih rendah daripada nilai wisman yang mencapai 1312224.76.
2. Wisman (Wisatawan Mancanegara) lebih sulit untuk diprediksi dibandingkan dengan Wisnus (Wisatawan Nusantara), yang mungkin dapat dipengaruhi oleh berbagai faktor, seperti kondisi ekonomi global, kebijakan, keamanan, dan faktor lainnya.

4.5 Analisis Keunggulan dan Keterbatasan

Berdasarkan hasil evaluasi dan eksperimen yang sudah dilakukan, dapat diketahui beberapa keunggulan dan keterbatasan dari data dan model yaitu :

1. Model menunjukkan performa yang sangat baik dalam memprediksi pergerakan tren wisata yang ada di daerah Bali, ini bisa membantu para penyedia jasa wisata khususnya dalam menentukan tujuan dari paket wisata yang mereka akan jual kepada calon wisatawan.
2. Adapun keterbatasan model, karena model ini hanya menggunakan dataset destinasi wisata pada Pulau Bali, maka model ini hanya bisa digunakan di daerah yang spesifik, yaitu hanya di pulau bali.

BAB 5

Kesimpulan dan Saran

5.1 Ringkasan Temuan Utama

Berdasarkan eksperimen yang sudah dilakukan, data pariwisata menunjukkan bahwa tren kunjungan tidak tumbuh secara linear, melainkan sangat dipengaruhi oleh berbagai peristiwa eksternal seperti pandemi. Tren wisata juga dipengaruhi oleh faktor internal, seperti tahun dan jumlah wisatawan nusantara yang berkunjung, turut mempengaruhi jumlah wisatawan asing yang berkunjung ke tempat wisata tersebut. Hal ini menandakan bahwa penyedia jasa wisata lokal, seringkali melihat tren atau data tempat wisata yang ering dikunjungi oleh wisatawan nusantara untuk dijadikan referensi untuk destinasi wisata para penyedia jasa tur wisata, khususnya kepada wisatawan mancanegara (Wisman).

5.2 Jawaban Atas Rumusan Masalah

1. Bagaimana tren destinasi wisata berkembang?

Tren destinasi wisata berkembang tidak secara linear, akan tetapi, secara dinamis, dan dipengaruhi oleh berbagai faktor internal dan eksternal.

2. Faktor apa yang mempengaruhi popularitas suatu destinasi wisata?

Faktor yang mempengaruhi popularitas suatu destinasi wisata, khususnya pada pulau bali, meliputi faktor internal, seperti tanggapan wisatawan nusantara (lokal) terhadap tempat wisata tersebut, lokasi, hingga faktor eksternal seperti kebijakan yang berlaku, peristiwa eksternal seperti Covid-19 juga mempengaruhi perkembangan destinasi wisata.

3. Bagaimana agen perjalanan dapat memanfaatkan analisis tren untuk menyusun paket wisata yang lebih menarik?

Dengan cara, memanfaatkan analisis tren yang berlaku di wisatawan nusantara (Lokal) sebagai patokan untuk menentukan apakah sebuah tempat wisata itu ramai atau tidak, agen perjalanan atau penyedia jasa wisata dapat membuat paket wisata yang relevan dan sejalan dengan tren yang sedang berlaku.

5.3 Saran Untuk Pengembangan Lebih Lanjut

1. Peningkatan indikator penentu suatu tren wisata. Untuk mendapatkan hasil yang lebih akurat terkait dengan tren wisata, penambahan variabel seperti sentimen masyarakat terhadap suatu tempat pariwisata juga dapat membantu menentukan tren wisata yang lebih baik.
2. Perluasan dataset. Karena pada eksperimen yang kami buat hanya mencakup wilayah pulau Bali saja, peningkatan dataset ke ranah yang lebih luas dapat membantu menganalisis tren destinasi wisata dengan lebih jelas, penambahan daerah dataset dapat membantu menganalisis tren suatu wilayah dengan lebih jelas.

DAFTAR PUSTAKA

- [1] I. Utama, *Pengantar industri pariwisata*. books.google.com, 2015..
- [2] M. Wijaya, “Analisis Preferensi Wisatawan Milenial terhadap Wisata Alam di Masa Pandemi COVID-19,” *J. Kepariwisata Indonesia*, vol. 15, no. 2, pp. 88–97, 2021.
- [3] D. Lestari and A. Gunawan, “Pengaruh Media Sosial terhadap Keputusan Berwisata di Kalangan Generasi Z,” *J. Pariwisata Nusantara*, vol. 12, no. 1, pp. 45–56, 2020.
- [4] T. A. Ramadhan, “Tren Perjalanan Wisata Pasca-Pandemi: Studi Kasus pada Destinasi Wisata di Jawa Barat,” *J. Riset Pariwisata Indonesia*, vol. 7, no. 2, pp. 112–125, 2023.
- [5] I. Yuliana and R. Syahputra, “Peran Data dan Informasi dalam Pengembangan Pariwisata Berkelanjutan di Indonesia,” *J. Pembangunan Pariwisata*, vol. 5, no. 1, pp. 30–40, 2019.
- [6] H. Prasetyo and S. Wahyuni, “Penerapan Teknologi Digital dalam Pemasaran Destinasi Wisata Lokal di Indonesia,” *J. Ilmu Sos. dan Humaniora*, vol. 11, no. 3, pp. 102–115, 2022
- [7] Few, S. (2009). *Now you see it: simple visualization techniques for quantitative analysis*. Analytics Press.
- [8] Han, J., Kamber, M., & Mining, D. (2006). Concepts and techniques. *Morgan Kaufmann*, 340(1), 94104–103205.
- [9] Li, K., Xie, Y., Liu, Z., Li, S., & Liu, D. (2023). Tourist Attraction Reviews Based on Deep Learning Sentiment Analysis System. *Journal of Computing and Electronic Information Management*, 24(1), 45–52. <https://doi.org/10.54097/066t1j26>
- [10] Li, Q., Li, S., Hu, J., Zhang, S., & Hu, J. (2018). Tourism Review Sentiment Classification Using a Bidirectional Recurrent Neural Network with an Attention Mechanism and Topic-Enriched Word Vectors. *Sustainability*, 10(9), 3313. <https://doi.org/10.3390/su10093313>
- [11] Panjaitan, R. (2023). Systematic Literature Review of Sentiment Analysis on Various Review Platforms in the Tourism Sector. *Journal of Advanced Computer Knowledge and Algorithms*, 2(1), 15–25. <https://ojs.unimal.ac.id/index.php/jacka/article/view/20287>

LAMPIRAN

1. Lampiran A – Dataset dan Informasi Terkait

Sumber Data: Data Statistik Pariwisata Dinas Pariwisata Kota Bali

Jumlah Data: 1058 Data

Jumlah Atribut: 7

Deskripsi Atribut:

- Atribut 1 : Nama Tempat Wisata
- Atribut 2 : Jumlah Pengunjung 2020
- Atribut 3 : Jumlah Pengunjung 2021
- Atribut 4 : Jumlah Pengunjung 2022
- Atribut 5 : Jumlah Pengunjung 2023
- Atribut 6 : Jumlah Pengunjung 2024
- Atribut 7 : Lokasi Tempat Wisata

2. Lampiran A2 – Contoh Dataset Mentah (Raw)

OBYEK DAN DATA TARIK WISATA DI BALI TAHUN 2015-2024																		
NO	THE VISITED PLACES OF INTEREST	2015	2016	2017	2018		2019		2020		2021		2022		2023		2024	
					wisnus	wisman	wisnus	wisman	wisnus	wisman	wisnus	wisman	wisnus	wisman	wisnus	wisman	wisnus	wisman
1	DENPASAR CITY																	
1	Museum Bali	37.532	38.437	31.690	16.927	16.704	18.015	13.457	4.248	3.082	1.991	86	11.335	6.553	18.230	11.803	24796	11835
2	Museum Le Mayeur	5.265	6.194	5.229	1.799	3.324	1.106	2.476	1.945	413	131	1	945	707	1.584	1.948	3826	1466
3	Taman Budaya art centre	15.704	14.345	10.300	5.078	2.286	6.357	1.106	717	93	1.903	12	9.895	439	3.720	917	621	212
4	Pulau Serangan	125.643	16.303	28.000	9.799	28.626	12.566	35.709	2.395	2.595	1.404	214	-	-	-	-	-	-
5	Pirasthi Blanjong	463	313	211	121	133	0	0	-	-	73	-	42	14	8	6	69	10
6	Pasar Kumbasari	13.592	11.123	10.464	0	7.896	0	6.035	-	1.319	-	-	-	576	4.775	2.478	10162	7670
7	Pasar Badung	29.533	4.114	0	0	0	0	9.836	-	2.660	-	-	157	6.769	18.737	13.924	19705	15058
8	Museum Lukisan Sidik Jari	6.937	426	429	696	153	705	80	147	14	223	23	199	5	325	10	204	19
9	Monumen Perjuangan Rakyat Bali	135.020	251.438	237.251	170.410	50.529	159.322	33.787	37.443	4.736	6.325	198	98.482	7.222	101.998	15.404	184783	12096
10	Hutan Bakau Mangrove	17.088	17.064	22.642	17.449	21	12.278	48	831	1.946	-	-	-	-	-	-	-	-
11	Desa Budaya Kertalangu	74.311	78.170	223.572	58.781	7.489	67.684	1.511	4.027	95	4.770	-	107.013	-	-	-	-	-
12	Pura Dalem Sakenan	2.973	2.276	440	1.635	507	0	0	0	4.027	95	-	-	-	-	-	-	-
13	Pura Maspahit Grenceng	-	-	-	-	-	-	-	-	-	68	5	235	93	154	1.402	108	509
14	Dewata Oleh-Oleh Bali	-	-	-	862.623	22.164	490	0	-	-	-	-	-	-	-	-	-	-
15	Syarbandar Serangan	-	-	-	724.693	3.910	529.050	1.205.128	-	-	-	-	-	-	-	-	-	-
16	Koris Dance	-	-	-	8.200	57.151	3.582	44.760	375	6.300	-	-	28	-	-	-	-	-
17	Pura Maspahit Toraja	-	-	-	-	-	-	-	-	-	20	2	135	-	105	-	-	-
18	Taman Inspirasi (Munlig Siokan Mertasari)	-	-	-	-	-	-	-	-	-	61.991	949	60.922	4.655	19.196	663	11353	4155
19	Pantai Matahari Terbit Dan Bali Beach	-	-	-	-	-	-	-	-	-	201.967	7.189	446.469	96.553	126.705	12.071	664903	18519
20	Pantai Duryug	-	-	-	-	-	-	-	-	-	12.732	542	26.543	1.780	49.803	6.860	59712	8815
21	Pantai Kusumasanari	-	-	-	-	-	-	-	-	-	11.855	513	27.877	1.585	49.454	5.912	57999	7764
22	Pantai Cemara Alt	-	-	-	-	-	-	-	-	-	3.088	99	8.744	1.098	23.057	5.833	30981	8220

Gambar 2 Contoh Dataset Mentah Sebelum PreProcessing

A. Lampiran B1 – Data Cleaning

Langkah-langkah pembersihan:

Penanganan nilai kosong: Nilai Kosong Dihapuskan, Dominan Data Dibawah Tahun 2020

Duplikasi data : Tidak Ada Data Duplikat

Outlier :

B. Lampiran B2 – Transformasi Data

Jenis transformasi:

Normalisasi/Standarisasi:

Encoding:

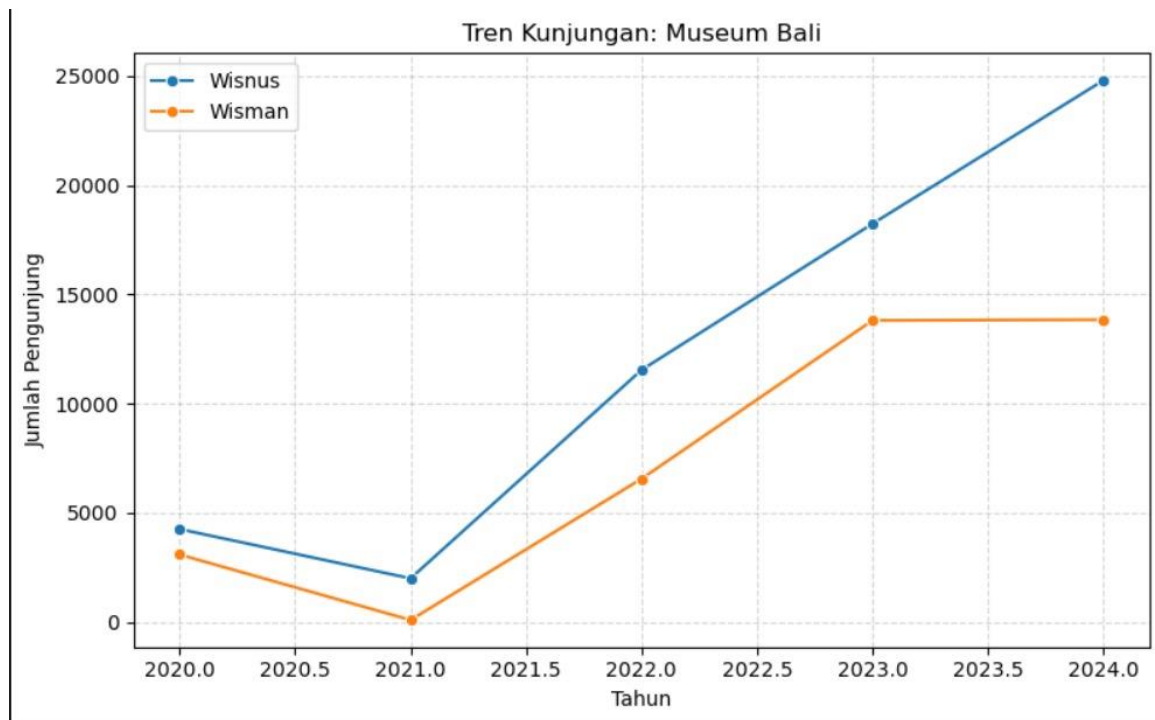
Binning/Discretization:

3. Lampiran A3 – Eksplorasi Data dan Visualisasi (EDA)

A. Lampiran C1 – Statistik Deskriptif

```
=== Statistik Deskriptif ===  
count    ID    Tahun    Wisnus    Wisman  
mean     525.500000  2022.000000  3.210434e+04  2.210195e+04  
std      303.253195    1.414887  1.420546e+05  9.683834e+04  
min       1.000000  2020.000000  0.000000e+00  0.000000e+00  
25%      263.250000  2021.000000  0.000000e+00  0.000000e+00  
50%      525.500000  2022.000000  8.110000e+02  1.110000e+02  
75%      787.750000  2023.000000  1.172725e+04  4.305250e+03  
max     1050.000000  2024.000000  2.075152e+06  1.328680e+06  
  
=== Median ===  
ID        525.5  
Tahun     2022.0  
Wisnus    811.0  
Wisman    111.0  
dtype: float64  
  
=== Jumlah Nilai NaN ===  
ID        0  
Tempat Wisata  0  
Tahun     0  
Wisnus    0  
Wisman    0  
dtype: int64
```

B. Lampiran C2 – Grafik & Visualisasi



4. Lampiran D1 – Pemodelan dan Evaluasi

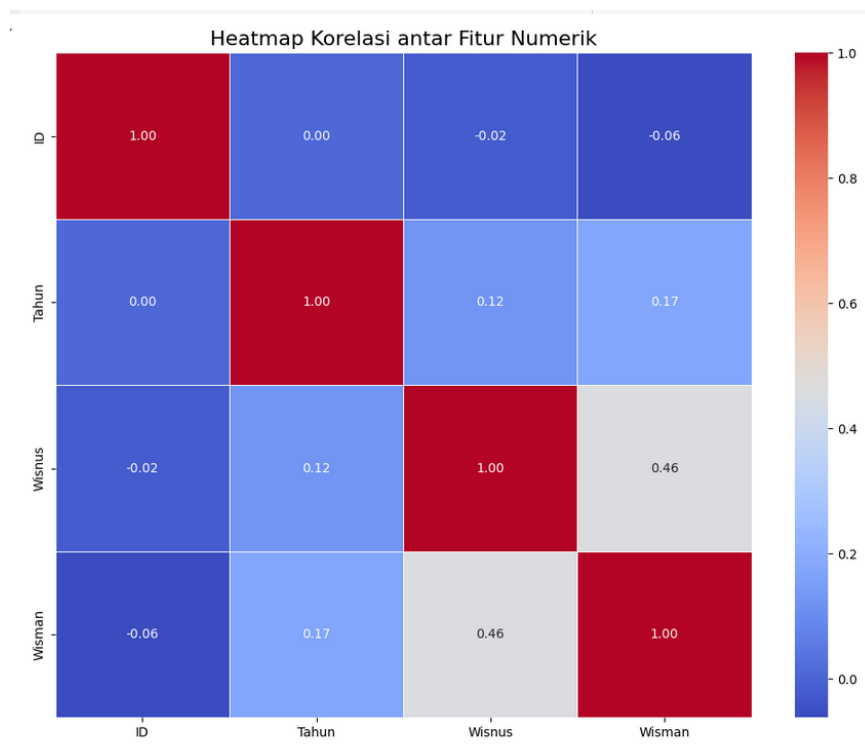
A. Lampiran D1 – Rincian Model

Model Yang Digunakan : Linear Regression

Parameter Model :

B. Lampiran D2 – Hasil Evaluasi Model

- Confusion Matrix



- Classification Report (Accuracy, Precision, Recall, F1)
- ROC Curve / AUC / RMSE (Jika Regresi)

```
Evaluasi Regresi untuk Total Wisnus  
R² Score      : 0.94  
MAE           : 801874.88  
MSE           : 909956125191.47  
RMSE          : 953916.20  
Evaluasi Regresi untuk Total Wisman  
R² Score      : 0.85  
MAE           : 1312224.76  
MSE           : 2163086632965.08  
RMSE          : 1470743.56
```

5. Lampiran E – Kode Program

A. Lampiran E1 – Script Python/R/Notebook

```
import pandas as pd

# Baca file Excel dengan dua baris header
df_raw = pd.read_excel('[CLASSIFIED] Data Table-17-Final-2-1.xlsx', sheet_name=0, header=[5, 6]) # perhatikan header=[5,6]

# Ambil kolom tempat wisata + 2020-2024 (wisnus & wisman)
cols_to_use = [col for col in df_raw.columns if col[0] in ['THE VISITED PLACES of INTEREST'] or col[0] in [2020, 2021, 2022, 2023, 2024]]
df = df_raw[cols_to_use]

# Rename kolom untuk keperluan selanjutnya
df.columns = ['Tempat Wisata' + f'_{col[1]}' for col in df.columns if col[0] != 'THE VISITED PLACES of INTEREST']

# Hapus baris kosong dan baris kategori wilayah (seperti 'DENPASAR CITY')
df.dropna(subset=['Tempat Wisata'], inplace=True)
df = df[~df['Tempat Wisata'].str.contains('CITY|REGENCY|KABUPATEN|KOTA', case=False, na=False)]

# Bersihkan format angka
for col in df.columns:
    if col.startswith('20'):
        df[col] = df[col].astype(str).replace('.', '', regex=False)
        df[col] = pd.to_numeric(df[col], errors='coerce').fillna(0).astype(int)

# Ubah ke format final
rows = []
for _, row in df.iterrows():
    tempat = row['Tempat Wisata']
    for year in range(2020, 2025):
        wisnus = row.get(f'{year}_wisnus', 0)
        wisman = row.get(f'{year}_wisman', 0)
        rows.append([tempat, year, wisnus, wisman])

# Buat DataFrame final
clean_df = pd.DataFrame(rows, columns=['Tempat Wisata', 'Tahun', 'Wisnus', 'Wisman'])
clean_df.insert(0, 'ID', range(1, len(clean_df)+1))

# Preview
print(clean_df.head())
```

	ID	Tempat Wisata	Tahun	Wisnus	Wisman
0	1	Museum Bali	2020	0	0
1	2	Museum Bali	2021	0	0
2	3	Museum Bali	2022	0	0
3	4	Museum Bali	2023	0	0
4	5	Museum Bali	2024	0	0

B. Lampiran E2 – Struktur Folder Proyek

