

OUTLINE

INTRODUCTION

COLLECT DATA

VISUALISASI DATA

KESIMPULAN

INSIGHT

APA ITU WEBSITE LAMBETURAH



"Lambe Turah" adalah sebuah akun media sosial yang terkenal di Indonesia, khususnya di Instagram. Akun ini dikenal karena sering mengunggah gosip, berita selebriti, dan isu-isu terkini yang sedang hangat dibicarakan oleh publik. Akun ini sering kali menjadi sumber informasi pertama mengenai berbagai skandal atau peristiwa yang melibatkan tokoh-tokoh terkenal di Indonesia.

COLLECT DATA



MONGO DB

Digunakan untuk menyimpan dan mengelola data hasil scraping secara aman dan terstruktur di cloud.



R Studio

Digunakan untuk menulis, mengembangkan, dan menjalankan skrip R yang melakukan scraping, pembersihan, dan analisis data dari Kalibrr.

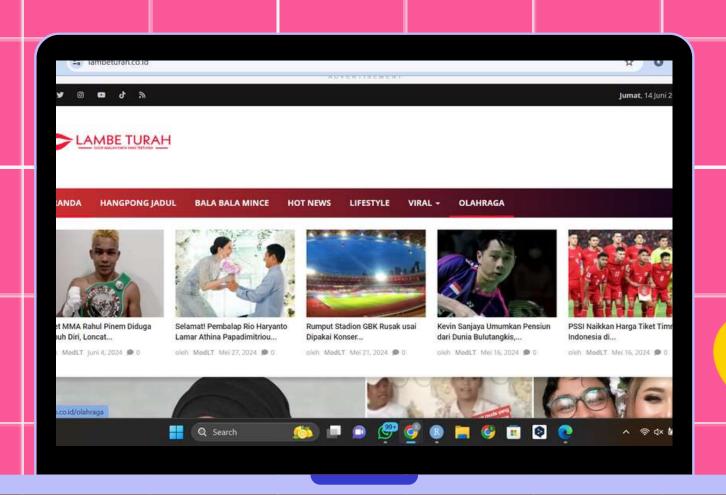


Github

Digunakan untuk mengotomatisasi scraping data setiap hari pada jam 01.00 dengan menjalankan skrip R secara terjadwal. Penelitian ini menggunakan teknik web scraping dari situs "Lambe Turah" untuk mengumpulkan data secara efektif dan akurat. Data yang terdiri dari judul berita utama, tanggal publikasi, dan tautan artikel berita dikumpulkan menggunakan bahasa pemrograman R dan disimpan di MongoDB. Proses scraping diatur secara otomatis untuk mengumpulkan data setiap hari, mulai dari 8 Juni 2024 hingga 13 Juni 2024, memastikan dataset yang selalu diperbarui dan konsisten.

Untuk memastikan dataset yang konsisten dan selalu diperbarui, proses scraping ini diotomatisasi untuk mengumpulkan data setiap hari selama periode tersebut. Penelitian ini juga memanfaatkan GitHub sebagai alat kontrol versi, kolaborasi, pencadangan, dokumentasi, dan otomatisasi, yang membuat proses scraping menjadi lebih efisien, dapat diandalkan, dan mudah dikelola. Dengan pendekatan ini, penelitian berhasil mengumpulkan data yang kaya dan relevan dari situs "Lambe Turah" secara terus-menerus.

SCRAPING DATA LAMBETURAH







Lambe Turah" memiliki banyak pengikut dan sering kali postingannya menjadi viral karena konten yang menarik dan kontroversial. Namun, karena sering kali berisi gosip, informasi yang disebarkan oleh akun ini perlu diterima dengan hati-hati dan sebaiknya diverifikasi lebih lanjut sebelum dianggap sebagai kebenaran. maka disini dilakukan visualisasi selama 1 minggu pertama dibulan juni 2024

KATEGORI BERITA POPULER

TREND JUMLAH ARTIKEL

WORDCLOUD

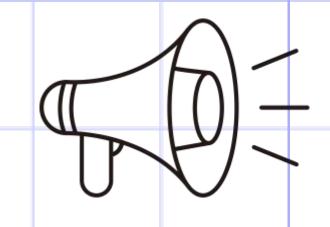
POPULARITAS KATEGORI BERITA SEIRING WAKTU

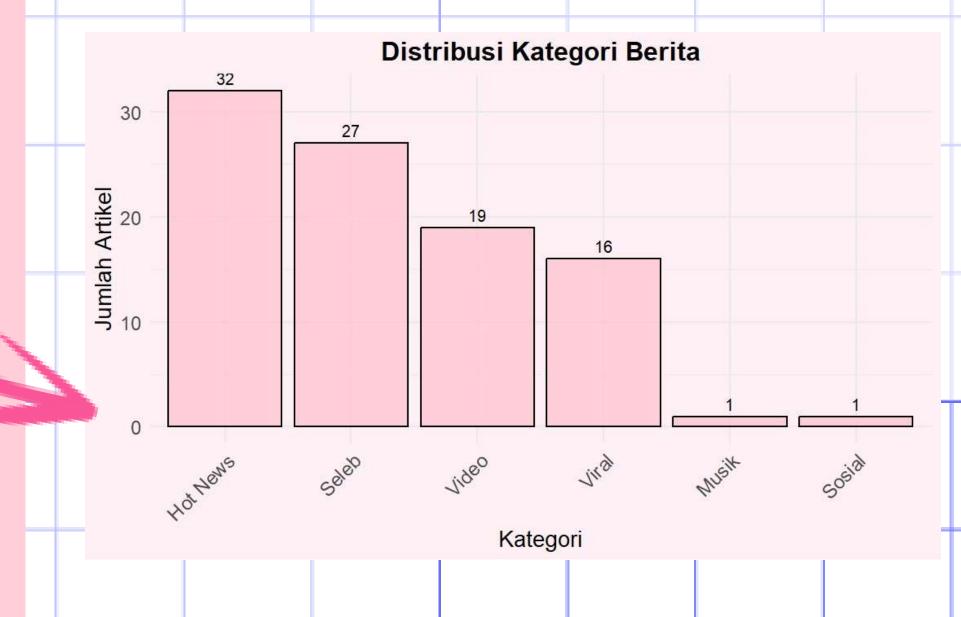
ANALISIS SENTIMEN

DISTRIBUSI SENTIMEN DALAM JUDUL BERITA PER KATEGORI

KATEGORI BERITA POPULER

Grafik distribusi kategori berita dari "Lambe Turah"
menunjukkan bahwa kategori paling populer adalah "Hot
News" dengan 32 artikel, diikuti oleh kategori "Seleb" dengan 27
artikel. Konten video juga cukup sering diposting, dengan 19
artikel, dan kategori "Viral" memiliki 16 artikel. Sementara itu,
kategori seperti "Musik" dan "Sosial" hanya memiliki satu artikel
masing-masing, menunjukkan bahwa topik ini jarang dibahas
atau kurang menarik perhatian dibandingkan kategori lainnya.

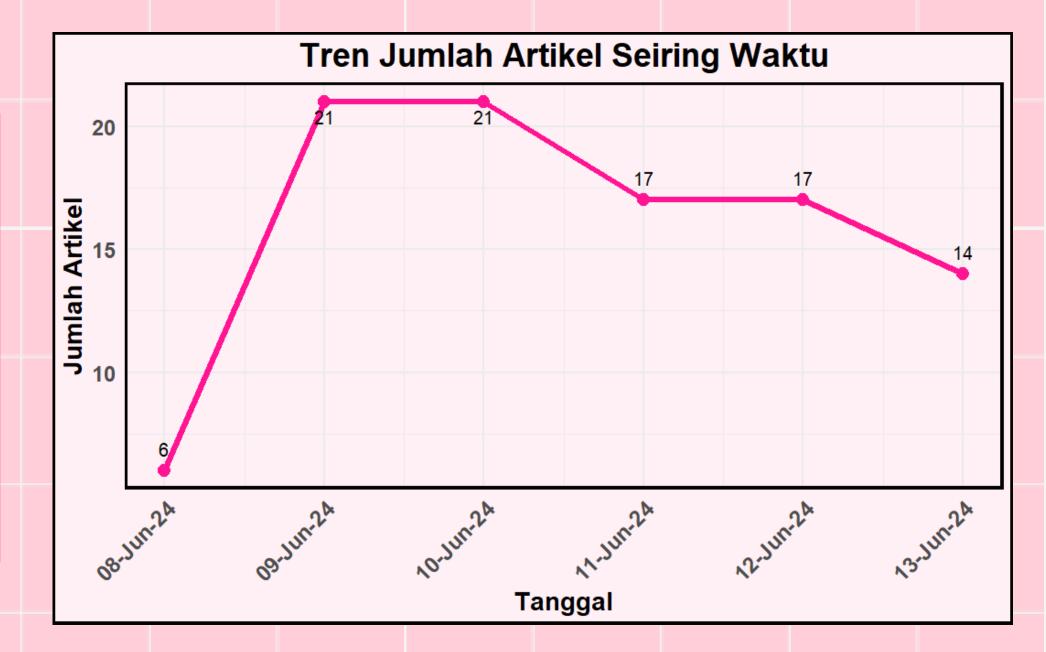








Grafik "Tren Jumlah Artikel Seiring Waktu" dari "Lambe Turah" menunjukkan perubahan jumlah artikel yang diposting dari tanggal 8 Juni 2024 hingga 13 Juni 2024. Pada tanggal 8 Juni 2024, jumlah artikel yang diposting adalah 6. Jumlah ini meningkat tajam menjadi 21 artikel pada tanggal 9 Juni 2024 dan tetap stabil hingga tanggal 10 Juni 2024. Namun, mulai tanggal 11 Juni 2024, jumlah artikel menurun menjadi 17 dan tetap konstan hingga tanggal 12 Juni 2024. Pada tanggal 13 Juni 2024, jumlah artikel kembali menurun menjadi 14. Dari grafik ini, terlihat bahwa aktivitas posting artikel di "Lambe Turah" mengalami peningkatan signifikan pada awal periode, diikuti oleh penurunan secara bertahap di hari-hari berikutnya.

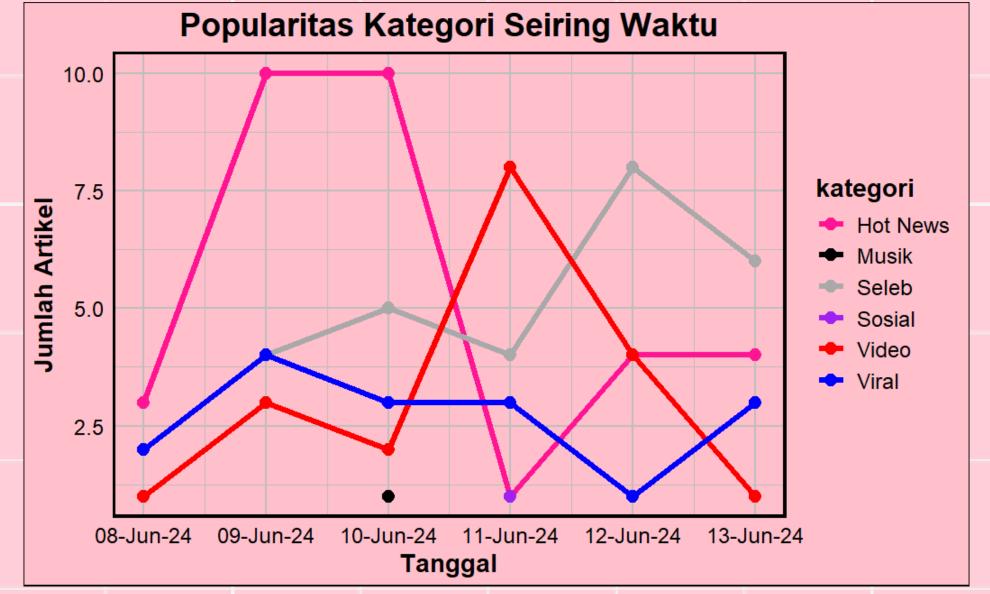


WORD CLOUD

Berdasarkan word cloud dari data web scraping "Lambe Turah" selama seminggu dari tanggal 8-13 Juni 2024, terlihat bahwa kata-kata yang paling sering muncul adalah "viral", "polisi", "terkait", "usai", "anak", dan "mobil". Kata "viral" paling menonjol, menunjukkan bahwa banyak artikel membahas topik yang sedang viral. Kata "polisi" dan "terkait" juga sering muncul, yang mungkin menunjukkan banyak berita yang melibatkan pihak kepolisian atau kasus-kasustertentu.

Selain itu, kata-kata seperti "anak", "usai", dan "mobil" menunjukkan bahwa berita yang melibatkan anak-anak, kejadian yang baru saja selesai, dan kendaraan juga sering dibahas. Word cloud ini memberikan gambaran umum tentang topik dan kata kunci yang dominan dalam artikelartikel "Lambe Turah" selama periode tersebut, mencerminkan fokus konten mereka pada kejadian viral dan berita yang melibatkan pihak berwenang serta berbagai peristiwa sehari-hari.

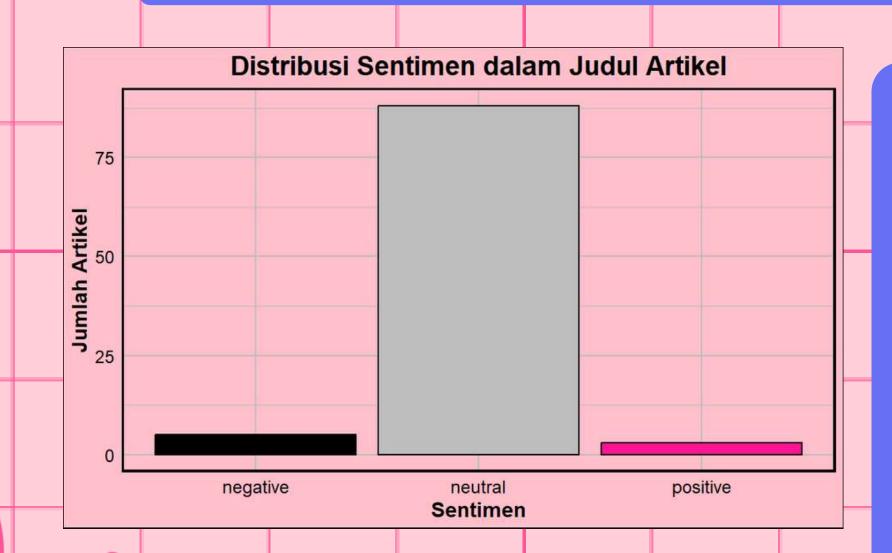
POPULARITAS KATEGORI SEIRING WAKTU







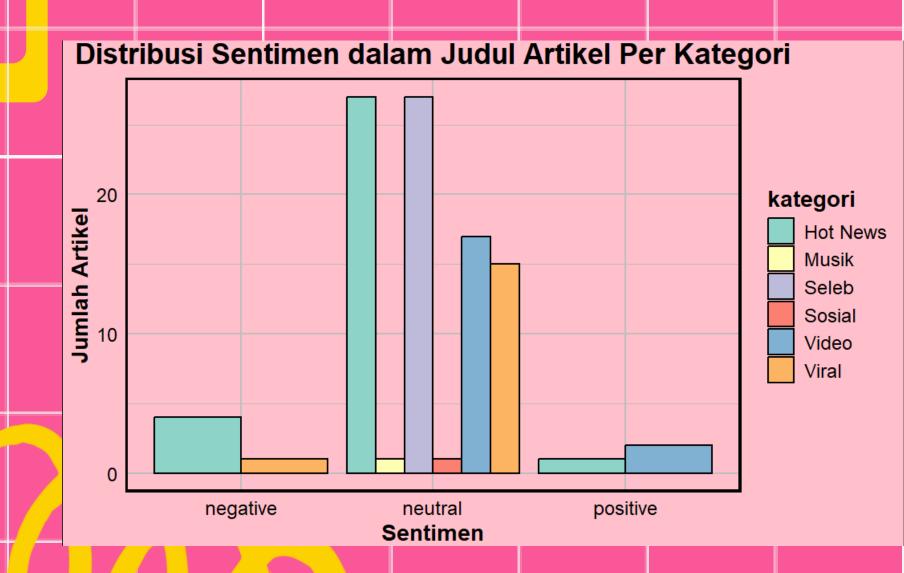
DISTRIBUSI SENTIMEN DALAM ARTIKEL



Grafik distribusi sentimen dalam judul artikel dari "Lambe Turah" menunjukkan mayoritas artikel memiliki sentimen netral, jauh lebih tinggi dibandingkan artikel dengan sentimen positif atau negatif. Artikel netral mendominasi lebih dari 75 judul, sementara sentimen negatif dan positif hanya sedikit, dengan negatif sedikit lebih banyak daripada positif. Ini menunjukkan bahwa "Lambe Turah" cenderung menyajikan berita secara informatif dan objektif, menghindari bias emosional yang ekstrem. Pendekatan ini mungkin bertujuan untuk menarik audiens yang lebih luas dan menjaga kepercayaan pembaca terhadap akurasi informasi yang disajikan.

SENTIMEN JUDUL ARTIKEL PER KATEGORI

Grafik distribusi sentimen dalam judul artikel per kategori dari "Lambe Turah" menunjukkan bahwa mayoritas artikel dalam semua kategori memiliki sentimen netral. Kategori "Hot News" dan "Seleb" didominasi oleh artikel netral, dengan sedikit artikel negatif dan hampir tidak ada yang positif. Kategori "Video" dan "Viral" menunjukkan variasi sentimen yang lebih besar, dengan beberapa artikel positif dan negatif. Sementara itu, kategori "Musik" dan "Sosial" memiliki jumlah artikel yang sangat sedikit dan umumnya bersifat netral. Ini menunjukkan bahwa "Lambe Turah" cenderung menyajikan berita secara informatif dan objektif, dengan sedikit konten yang memancing reaksi emosional yang kuat dari pembaca.

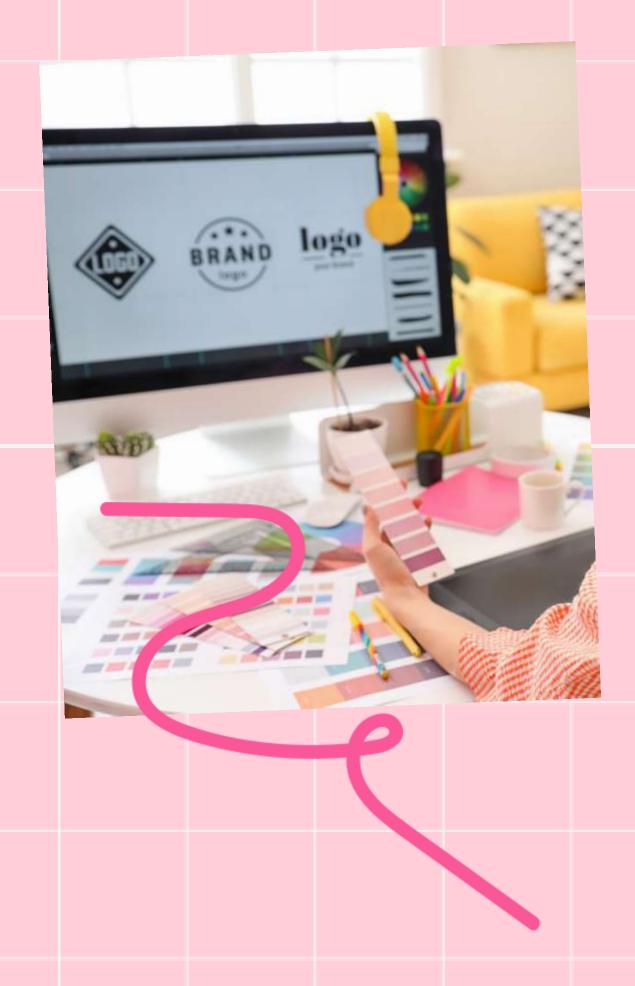


KESIMPULAN

Berdasarkan analisis data web scraping dari "Lambe Turah" selama periode 8-13 Juni 2024, dapat disimpulkan bahwa kategori berita yang paling sering dibahas adalah "Hot News" dan "Seleb", menunjukkan fokus pada berita terbaru dan gosip selebriti. Aktivitas posting mencapai puncaknya pada 9 dan 10 Juni, kemudian menurun hingga 13 Juni. Word cloud menunjukkan kata "viral" paling menonjol, mencerminkan fokus pada berita trending. Sebagian besar judul artikel memiliki sentimen netral, menunjukkan pendekatan informatif dan objektif, dengan sangat sedikit artikel yang memiliki sentimen negatif atau positif. Kategori "Hot News" dan "Seleb" didominasi oleh sentimen netral, sementara kategori "Video" dan "Viral" menunjukkan variasi sentimen yang lebih besar. Kategori "Musik" dan "Sosial" memiliki jumlah artikel yang sangat sedikit dan umumnya bersifat netral. Secara keseluruhan, "Lambe Turah" menyajikan berita yang seimbang dan informatif, menarik audiens luas dan menjaga kepercayaan pembaca.

INSIGHT

Proyek scraping data di situs "Lambe Turah" memberikan beberapa insight penting. Pertama, kategori berita "Hot News" dan "Seleb" adalah yang paling populer, menunjukkan tren minat audiens. Kedua, mayoritas artikel memiliki sentimen netral, mencerminkan pendekatan informatif dan objektif. Ketiga, analisis kata kunci menunjukkan topik yang sering dibahas, seperti "viral" dan "polisi". Keempat, data interaksi pengguna membantu memahami artikel yang paling menarik bagi pembaca. Terakhir, dengan memahami waktu puncak publikasi, "Lambe Turah" dapat mengoptimalkan jadwal posting untuk meningkatkan engagement dan kepercayaan pembaca.



https://lambeturah.co.id/

sitihafsahh12@gmail.com