

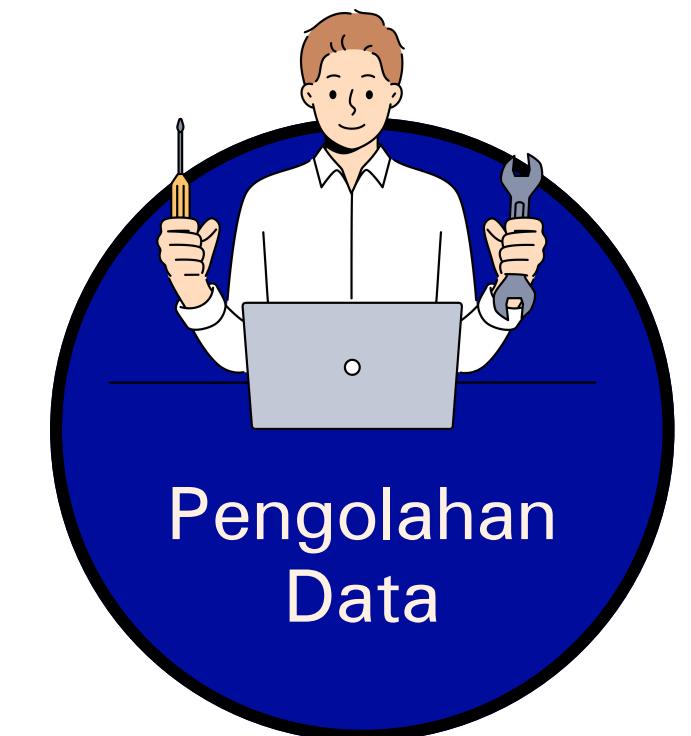
Scraping 'Cek Fakta' From :



Aktual, Tajam, dan Terpercaya

Dwi Fitrianti (G1501231013)

Outline Content



Apa itu? Liputan6 : 'Cek Fakta'



The screenshot shows the homepage of Liputan6.com with a blue header. The top navigation bar includes categories like NEWS, PEMILU, BISNIS, BOLA, TV, SHOWBIZ, TEKNO, FOTO, HOT, CEK FAKTA, ISLAMI, CRYPTO, and LAINNYA. Below the header is a search bar and a 'MASUK' button. A red banner at the top says 'HEADLINE HARI INI' followed by 'cek dan ukir Sejarah?'. The main content area features several news articles under the 'CEK FAKTA' category, each with a thumbnail image, title, and a timestamp. To the right, there's a sidebar with numbered items 4 through 8, each with a 'CEK FAKTA' link and a brief description.

4	CEK FAKTA Kumpulan Hoaks yang Mencatut Dana Haji, Simak Faktanya
5	CEK FAKTA Literasi Digital Kunci Cegah Perempuan Jadi Korban di Ruang Digital
6	CEK FAKTA Cek Fakta: Tidak Benar Dalam Video Ini Gunung Anak Krakatau Sedang
7	CEK FAKTA Hoaks Ancam Stabilitas Publik, Polri Ikut Wujudkan Ekosistem Digital Sehat
8	CEK FAKTA Hoaks Menag Minta Masyarakat Ikhlaskan Dana Haji untuk IKN

Cek Fakta : Kebenaran di Balik Berita yang Viral dari Liputan 6
Di era digital ini, berita palsu atau hoaks semakin marak dan sering kali sulit dibedakan dari berita asli. Inilah mengapa Liputan6 hadir dengan segmen "Cek Fakta" untuk memastikan Anda mendapatkan informasi yang benar dan dapat dipercaya.

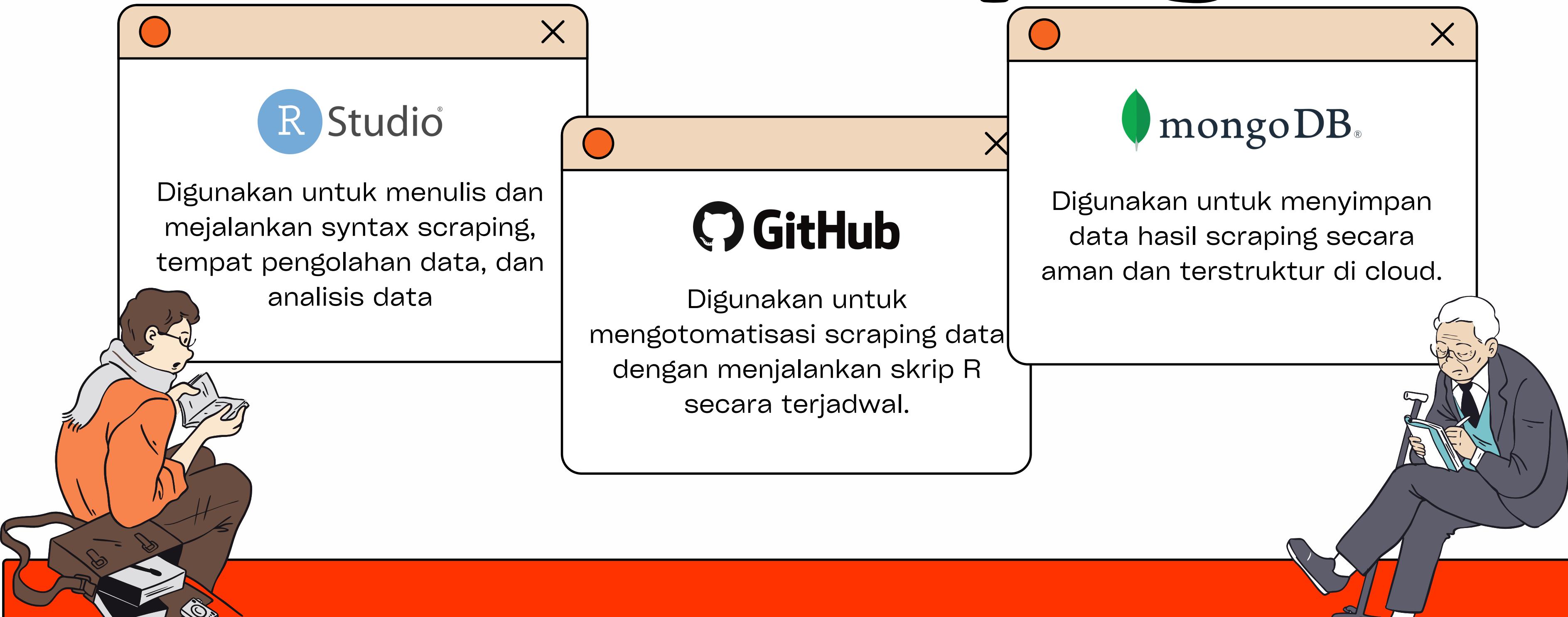


NEWS

UPDATE ➤

Alat

Scraping Data



Hasil Scraping



_id	ObjectId	time_scraped	Date	titles	String	dates	String
1	ObjectId('666053e9d5c7826...')	2024-06-05T12:02:46.613+0...		"Cek Fakta: Tidak Benar I...		"05 Jun 2024 19:00"	
2	ObjectId('666053e9d5c7826...')	2024-06-05T12:02:46.613+0...		"Cek Fakta: Tidak Benar G...		"05 Jun 2024 17:00"	
3	ObjectId('666053e9d5c7826...')	2024-06-05T12:02:46.613+0...		"Hoaks Undiah dari Bank T...		"04 Jun 2024 21:00"	
4	ObjectId('666053e9d5c7826...')	2024-06-05T12:02:46.613+0...		"Tokoh Terkenal Dicatut H...		"04 Jun 2024 19:00"	
	ObjectId('666053e9d5c7826...')	2024-06-05T12:02:46.613+0...		"Waspada Praktik Calo dan...		"04 Jun 2024 17:00"	
	ObjectId('666053e9d5c7826...')	2024-06-05T12:02:46.613+0...		"Cek Fakta: Tidak Benar P...		"04 Jun 2024 15:55"	
7	ObjectId('6661188ef317c0a...')	2024-06-06T02:01:47.601+0...		"BSI Ingatkan Nasabah Was...		"06 Jun 2024 07:00"	
	ObjectId('6661188ef317c0a...')	2024-06-06T02:01:47.601+0...		"Ragam Penipuan Lewat Wha...		"05 Jun 2024 21:00"	
	ObjectId('6661188ef317c0a...')	2024-06-06T02:01:47.601+0...		"Cek Fakta: Tidak Benar I...		"05 Jun 2024 19:00"	
	ObjectId('6661188ef317c0a...')	2024-06-06T02:01:47.601+0...		"Cek Fakta: Tidak Benar G...		"05 Jun 2024 17:00"	
	ObjectId('6661188ef317c0a...')	2024-06-06T02:01:47.601+0...		"Hoaks Undiah dari Bank T...		"04 Jun 2024 21:00"	
	ObjectId('6661188ef317c0a...')	2024-06-06T02:01:47.601+0...		"Tokoh Terkenal Dicatut H...		"04 Jun 2024 19:00"	



Mulai

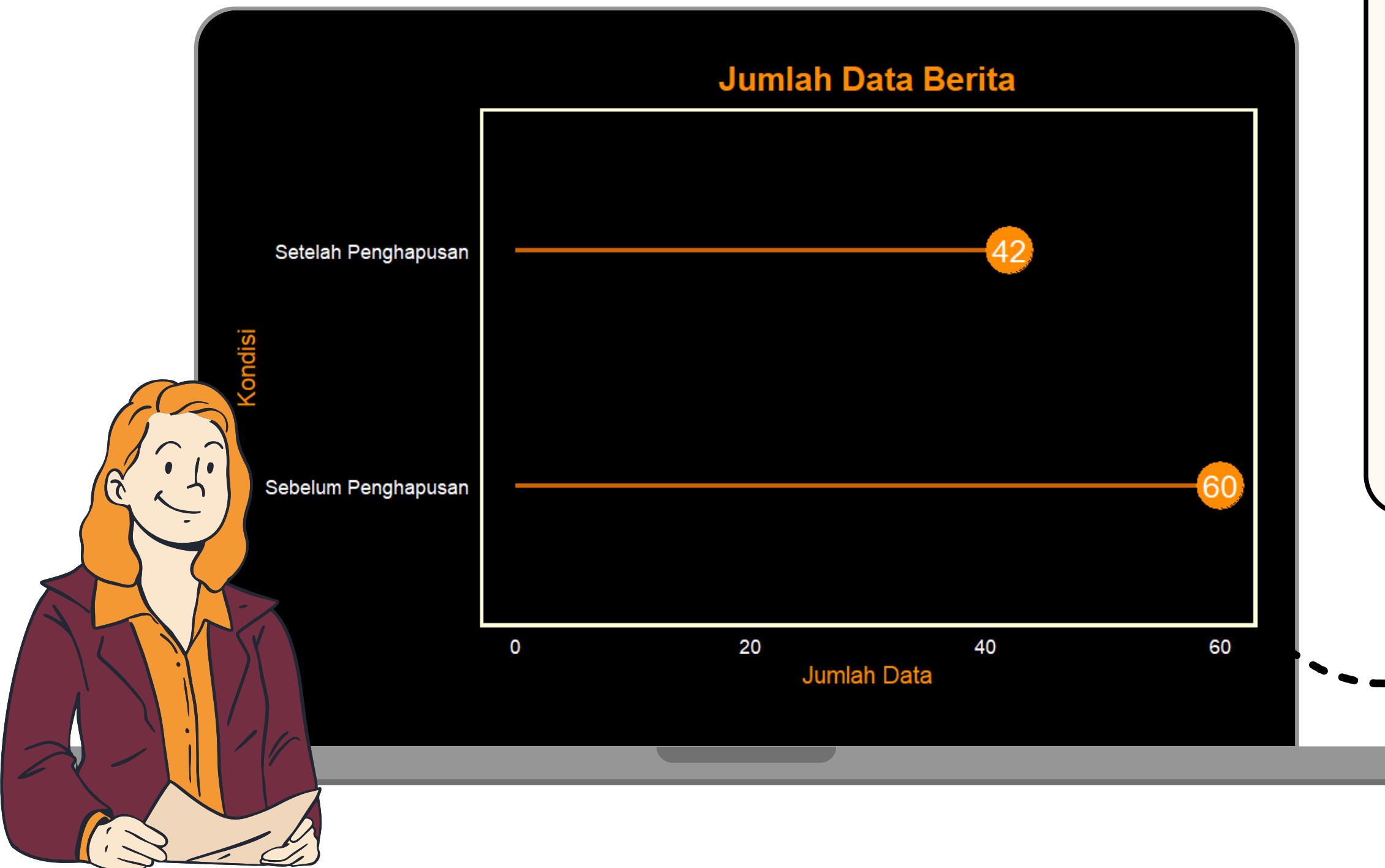
04 Juni 2024

Selesai

14 Juni 2024

Data yang didapatkan
60

BREAKING NEWS //

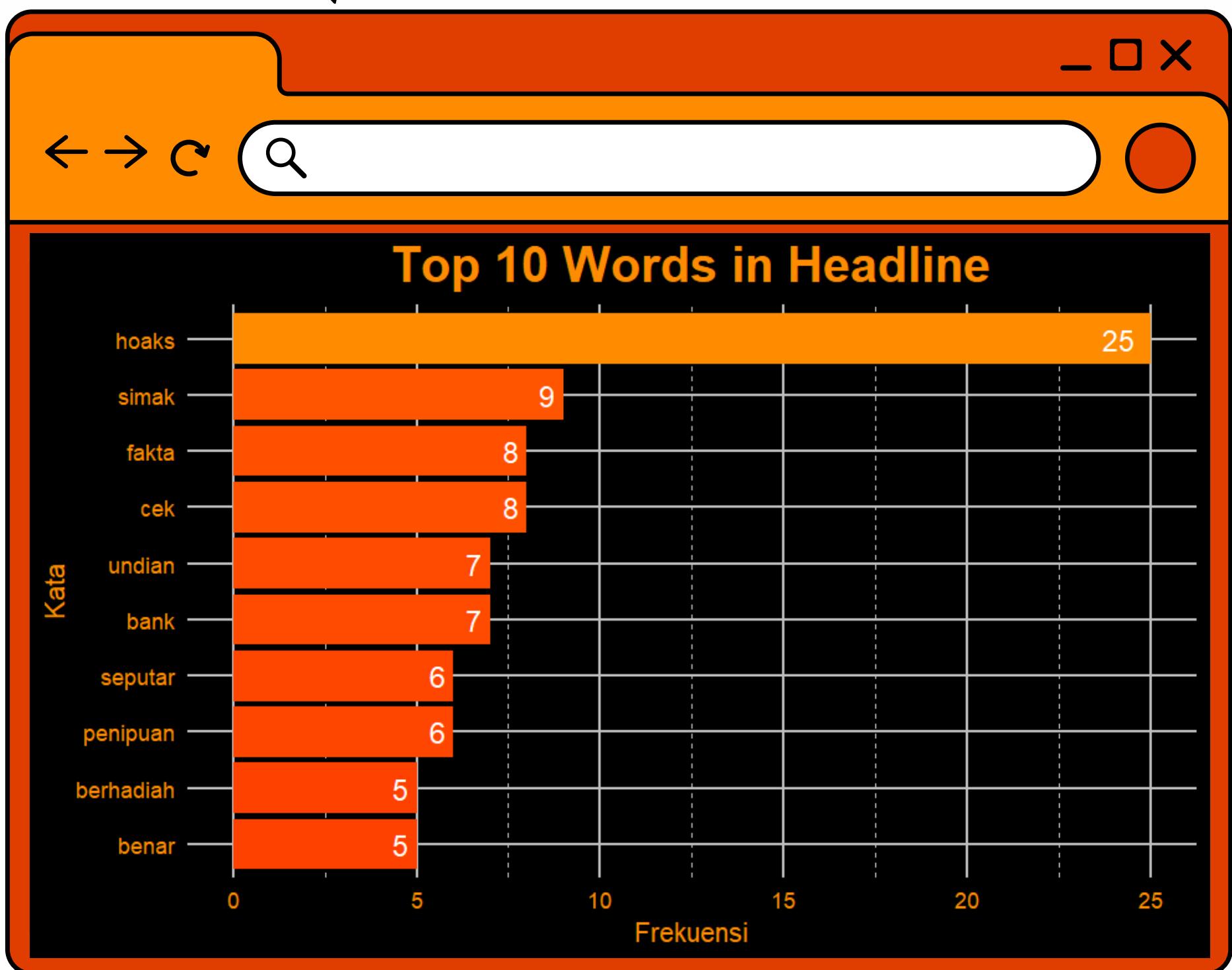


Dari 60 Data menjadi 41 Data setelah dilakukannya penghapusan terhadap data yang duplikasi. Duplikasi terjadi karena banyak data yang terambil sama setiap harinya namun pada hari tertentu tidak adanya update sejumlah data yang diambil sehingga terambilah data hari sebelumnya.

**Duplikat
Data**

Top 10 Words In Headline

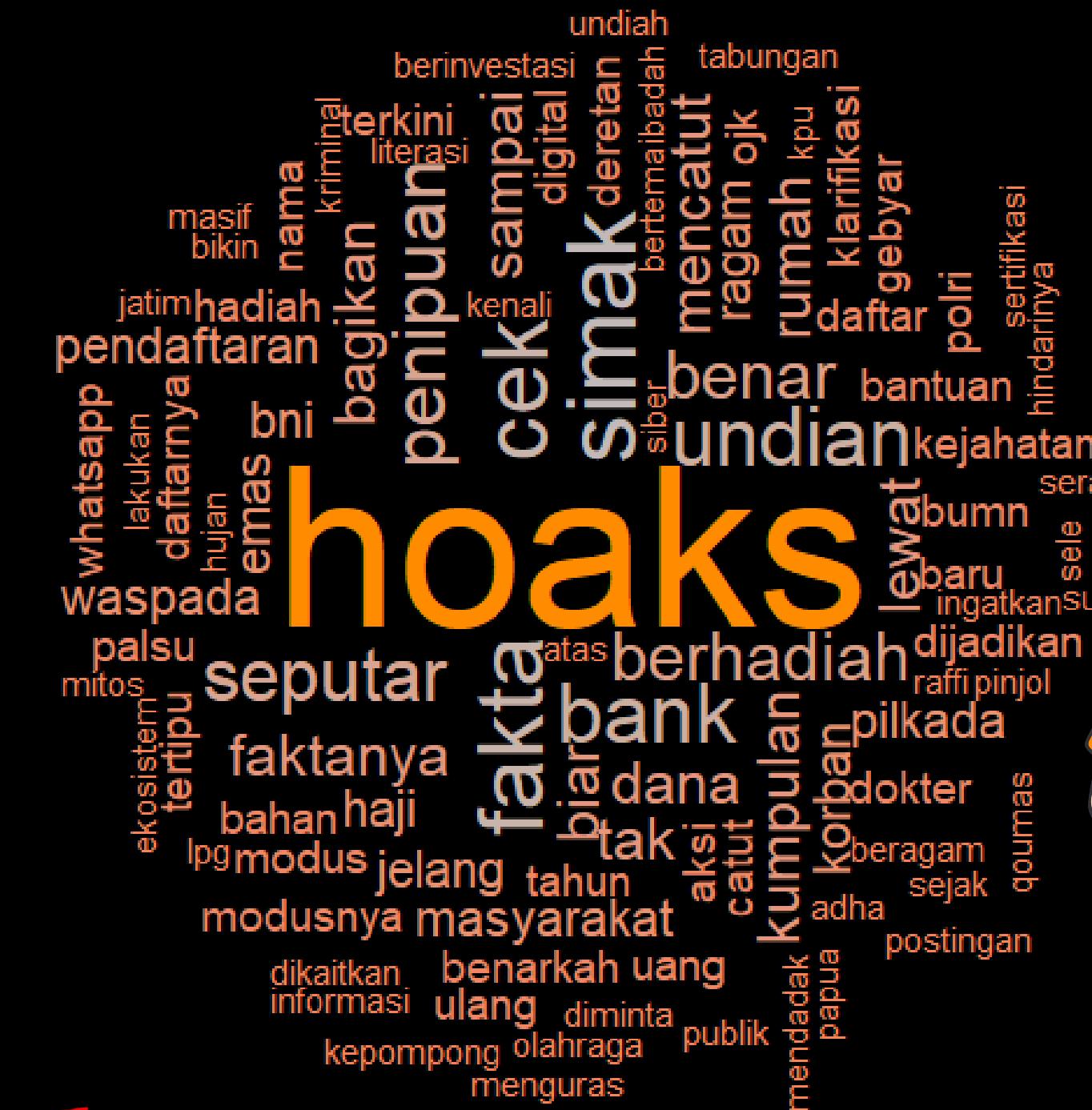
"**hoaks**" sebagai kata yang paling dominan atau yang paling sering muncul dengan frekuensi sebesar **25**. Kemudian dilanjutkan dengan "simak", "fakta", "cek", dan lain-lain. Hal ini menunjukkan tren atau fokus utama dalam berita yang dipublish terkait dengan verifikasi fakta atau berita palsu.



Words Cloud!

Wordcloud ini memberikan gambaran komprehensif tentang fokus dan tren dalam judul berita. Kekhawatiran terhadap hoaks dan penipuan jelas terlihat, disertai dengan penekanan pada verifikasi fakta dan keakuratan informasi. Media tampaknya berusaha keras untuk menyajikan berita yang dapat dipercaya.

dapat | **FAKE NEWS**



Trigrams

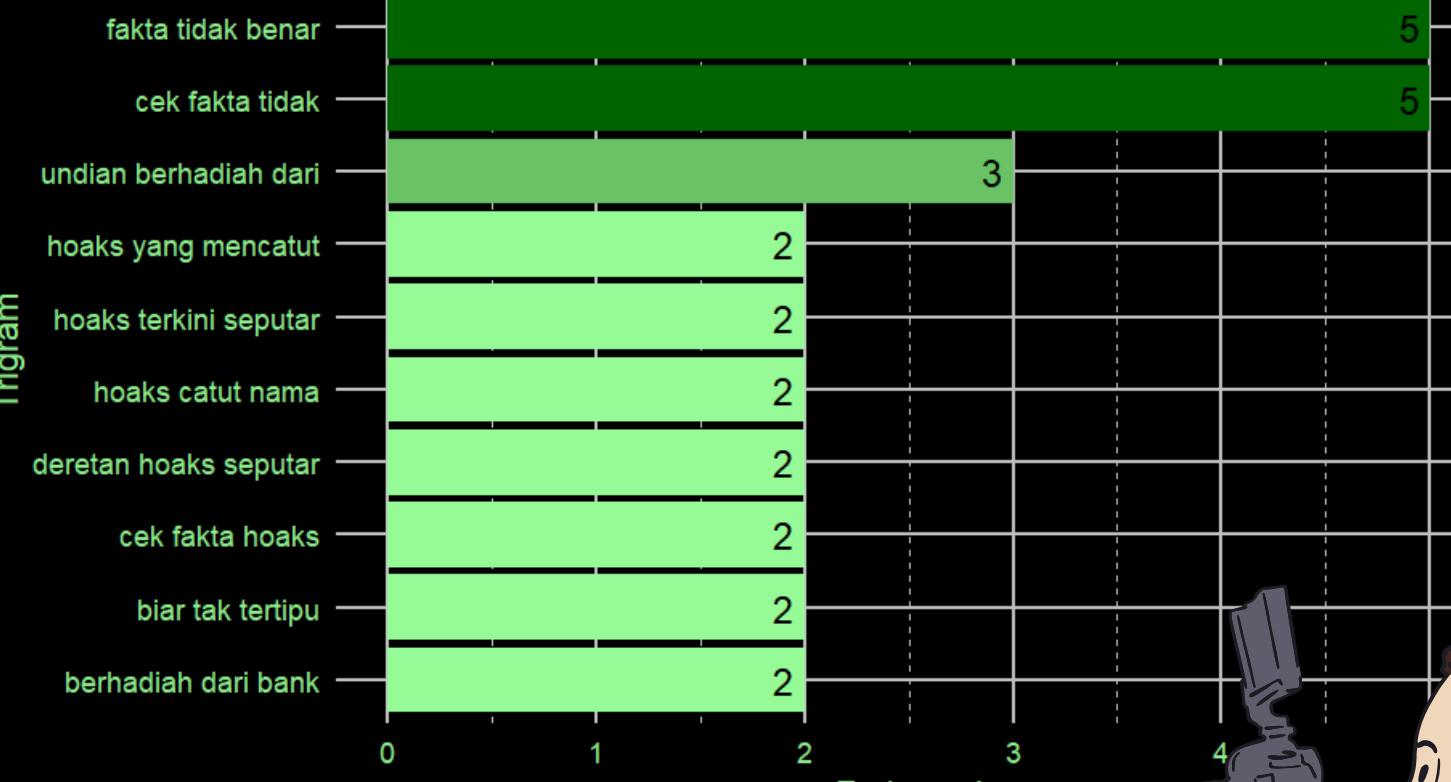
Document2 - Macrostuff Board

File Edit Format View Help

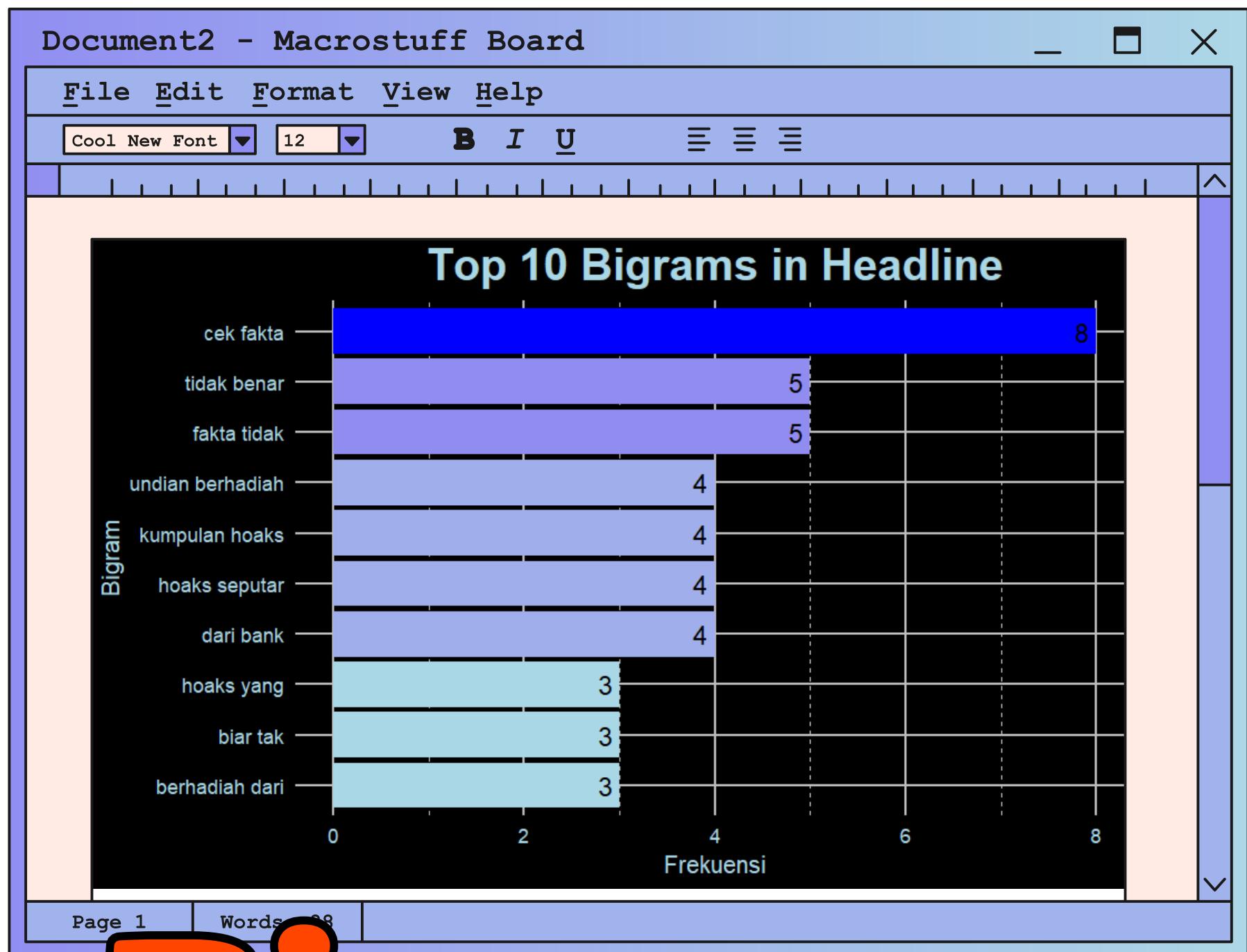
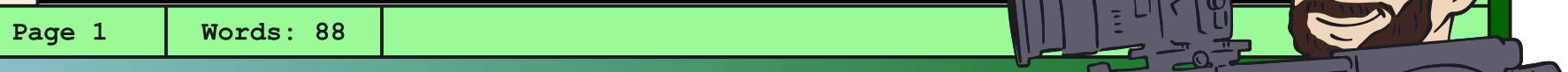
Cool New Font 12 B I U

Cool New Font 12 B I U

Top 10 Trigrams in Headline



Page 1 Words: 88

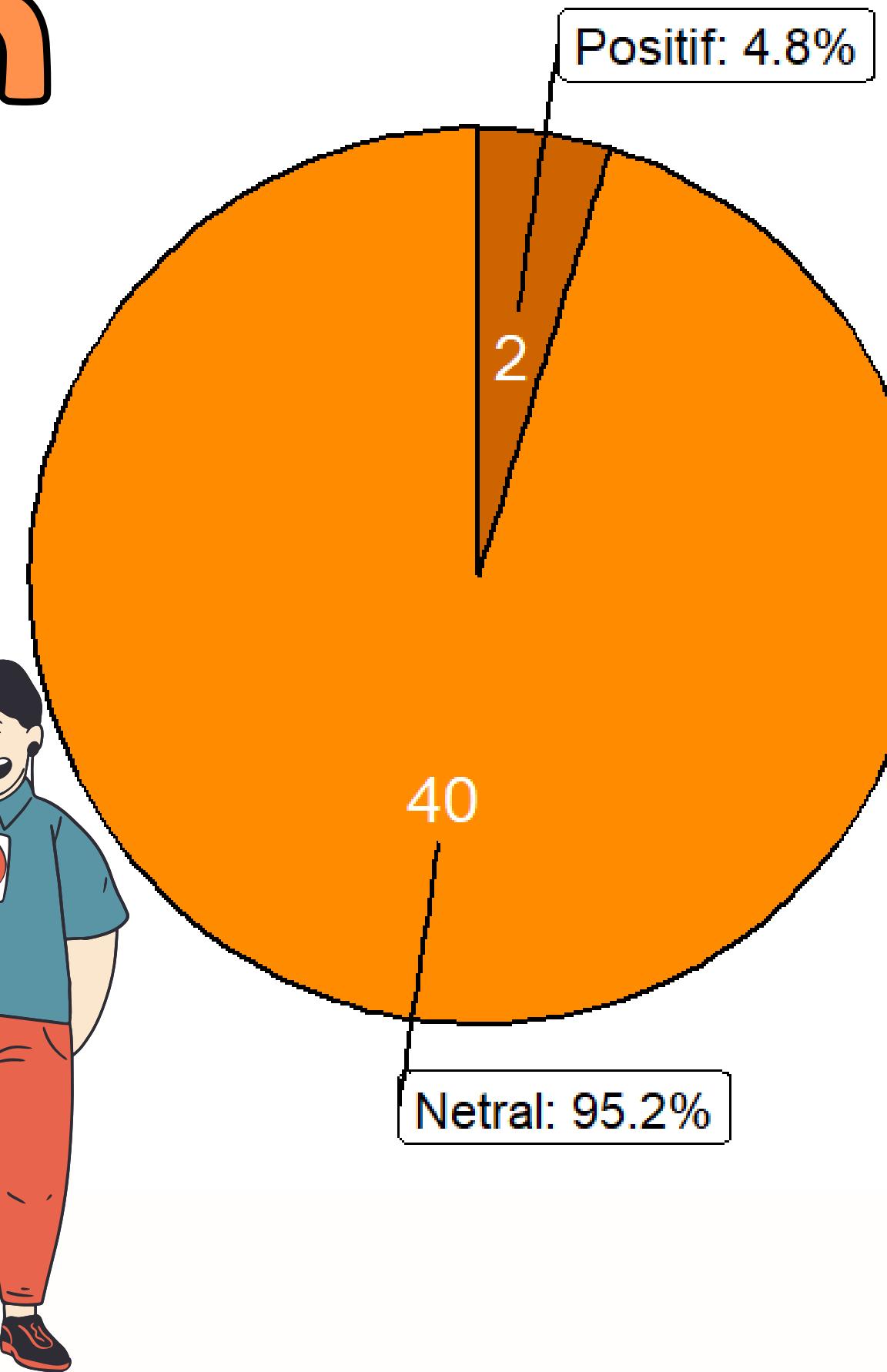


Bigrams

Page 1 Words: 88

Analisis Sentimen

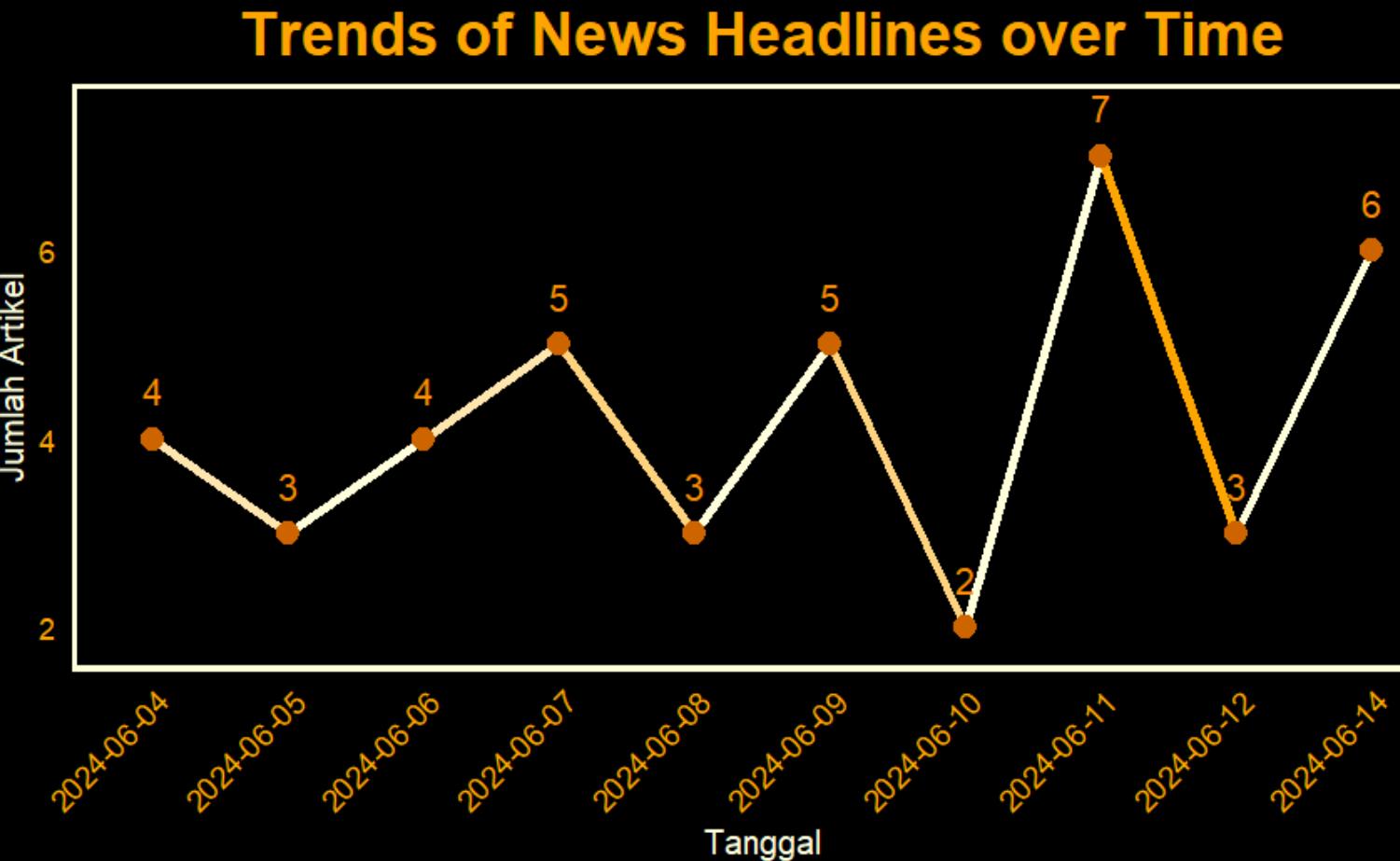
Data menunjukkan bahwa sebagian besar informasi yang diverifikasi oleh **Liputan6** cenderung **netral**, dengan hanya sebagian kecil yang memiliki konotasi **positif**. Tidak ada indikasi sentimen **negatif** dalam data yang ditampilkan pada diagram ini.



Angka **95.2%** ini menggambarkan bahwa dari total **42** data cek fakta, **40** di antaranya bersifat netral.

**BREAKING
NEWS**

Angka **4.8%** berarti dari **42** data cek fakta, hanya **2** yang memiliki sentimen positif.



Tren jumlah artikel berita cek fakta dari Liputan6 dalam rentang waktu 04 Juni 2024 hingga 14 Juni 2024. Terdapat fluktuasi signifikan dalam jumlah artikel yang diterbitkan setiap hari, dengan puncak tertinggi 7 artikel pada 11 Juni 2024 dan titik terendah 2 Juni 2024 artikel pada 10 Juni 2024 .



Tren Setiap Hari

Kesimpulan

Dalam proyek ini, telah berhasil mengotomatisasi proses pengambilan data dari situs Liputan6 untuk berita cek fakta yang dipublikasikan dari tanggal 4 Juni 2024 hingga 14 Juni 2024. Data yang diperoleh disimpan di MongoDB dan dianalisis menggunakan R Studio. Melalui proyek ini, tidak hanya berhasil mendapatkan wawasan berharga dari data cek fakta Liputan6, tetapi juga menunjukkan bagaimana teknologi dapat digunakan secara efektif untuk mengotomatisasi pengambilan dan analisis data.





Dwi Fitrianti
G1501231013

LIPUTAN⁶
.COM

THANK YOU

THE END!

BREAKING NEWS