

# Big Data Analytics



Big Data  
STF1724

Rini Nuraini Sukmana, M.T  
0020087901 - 08882024236  
rini.nuraini@usbypkp.ac.id



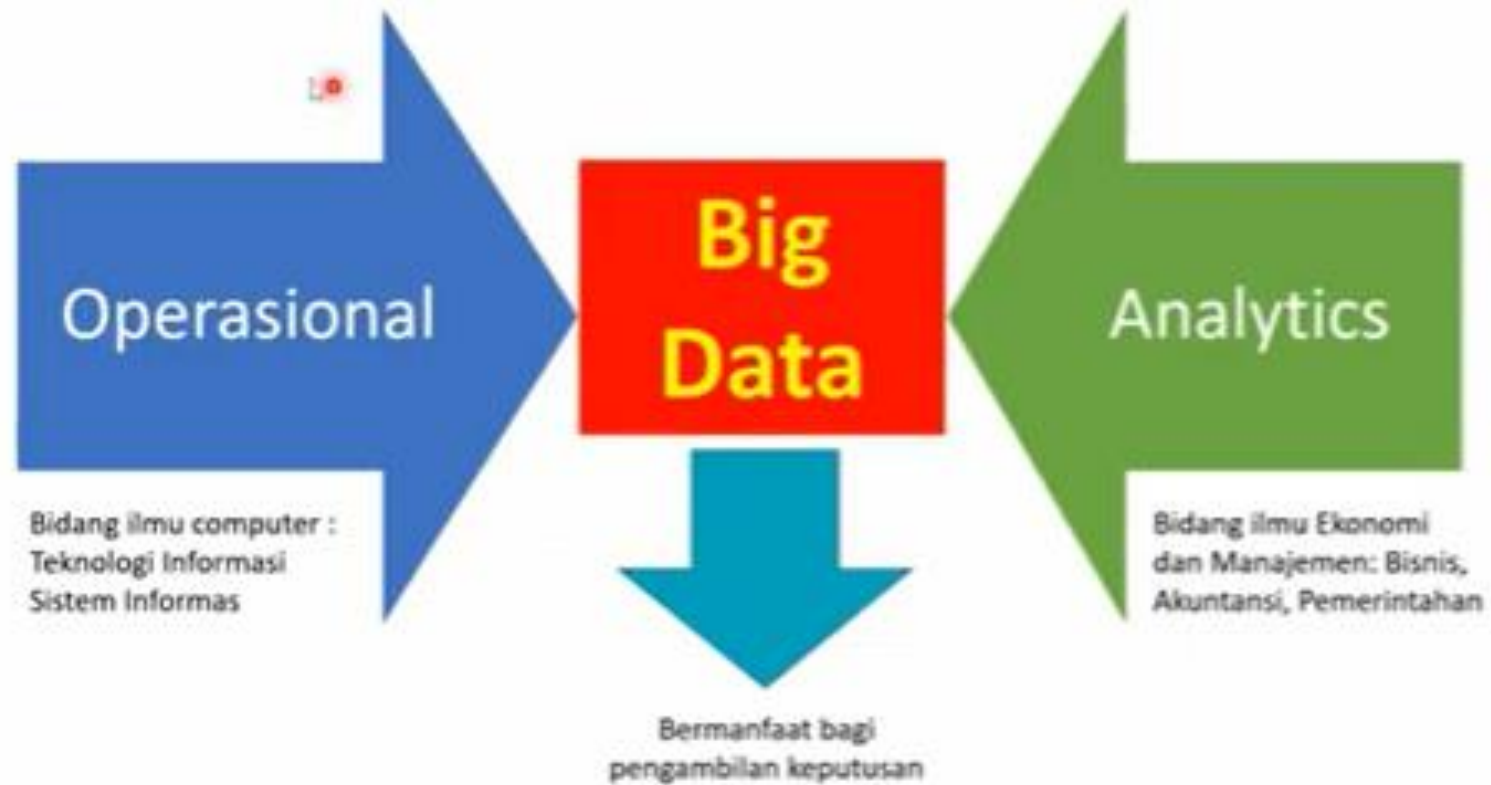
# Bahasan



- Big Data Analytics
- Cara Kerja Big Data Analytics
- Langkah-Langkah Penerapan Big Data Analytics
- Data Driven
- Contoh Implementasi Pengumpulan Data Sederhana



# Overview



# Teknologi Big Data







# Overview

- *Big data* disiapkan oleh perusahaan-perusahaan besar, firma ataupun organisasi → contoh : Bursa Efek Indonesia, facebook, Instagram, dll
- Data-data ini diambil, diolah, dan digunakan oleh perusahaan untuk kepentingan tertentu → analisis pasar, saham, keuangan perusahaan, profiling, marketing
- Keseluruhan proses untuk mengumpulkan, merapikan, menganalisis *big data* yang disebut dengan *big data analytics*.
- Banyak manfaat dari *big data analytics* khususnya untuk perusahaan. Salah satunya adalah untuk mengidentifikasi kesempatan-kesempatan baru.
- Contohnya, jika ada *sponsored post* yang lewat di *timeline* Instagram kamu dengan preferensi yang sesuai denganmu. Pasti kamu akan langsung membukanya kan?
- Kesempatan baru ini yang dimanfaatkan oleh perusahaan dengan menggunakan *big data*. Selain mereka diuntungkan, *customer* pun juga senang karena ada rekomendasi yang sesuai dengan kebutuhan dan kesukaan mereka.



# Cara Kerja Big Data Analytics

## 1. *Machine learning*

- Untuk mengumpulkan data, mesin yang berbasis Artificial Intelligent (Kecerdasan Buatan) digunakan sebagai mesin pencarian. Mesin ini dengan cepat mencari serta mempelajari data yang akan diambil → **Contoh : Google**
- Secara otomatis mesin akan menghasilkan model lain yang bisa menganalisis data lebih besar, kompleks, akurat, serta penyampaian yang lebih cepat lagi.

## 2. *Data management*

- Sebelum memberikan data ke dalam perusahaan, data harus dikaji ulang dan dipastikan kepada instansi terkait → **Contoh : Validitas data**
- Hal ini diperlukan agar data yang digunakan merupakan data yang berkualitas tinggi dan bukan merupakan data palsu yang dibuat-buat.



# Cara Kerja Big Data Analytics

## 3. *Data Mining*

- Teknologi [data mining](#) berfungsi agar para *data analyst* memeriksa data berskala besar untuk menemukan pola-pola di dalam sebuah data. Hasil dari analisis ini dapat digunakan untuk menjawab pertanyaan-pertanyaan rumit perusahaan.
- Dengan adanya teknologi *data mining*, para *analyst* dapat masuk ke dalam berbagai data, menandai hal-hal penting, serta membuat data menjadi salah satu solusi untuk mempengaruhi pengambilan keputusan.



# Cara Kerja Big Data Analytics

## 4. In-memory analytics

- Dengan menganalisis data menggunakan teknologi memori dalam sistem, para *data analyst* bisa mendapatkan *insight* sebuah data secara cepat.
- Teknologi ini dapat menganalisis secara cepat, membuat algoritma baru, menciptakan model baru serta menghapus analisa yang dianggap keliru.
- Teknologi ini diklaim bukan hanya dapat mempengaruhi pengambilan keputusan sebuah perusahaan, tetapi juga menciptakan berbagai scenario sebagai bahan pembelajaran.





# Cara Kerja Big Data Analytics

## 5. *Predictive analytics*

- Teknologi prediksi ini menggunakan data, algoritma statistik, teknik machine learning untuk mengidentifikasi outcomes berdasarkan riwayat data yang pernah digunakan.
- Predictive analytics akan menghasilkan prediksi-prediksi yang akan terjadi di masa depan, sehingga para perusahaan akan lebih percaya diri dengan keputusan apa yang akan mereka ambil nantinya.

## 6. *Text mining*

- Dengan teknologi ini, data analyst dapat menganalisa dari hasil tulisan yang berada di web, kolom komentar, buku, dan bagian web lainnya yang berbasis teks.
- Biasanya text mining akan dipasang dalam blog, Twitter, survey, bahkan email untuk menemukan topik terhangat yang dapat menciptakan relasi perusahaan kepada (calon) pelanggannya.



# Langkah-Langkah Penerapan Big Data Analytics

## 1. *Data Mining dan data extraction*

- Ada dua hal yang difokuskan dalam *big data analytics* yaitu *data mining* dan *data extraction*.
- *Data mining* adalah sebuah proses identifikasi dari *insight* yang berharga dari *database*.
- *Data extraction* adalah sebuah proses pengumpulan data dari halaman web ke dalam *database*.

## 2. *Data Collection*

- Big data tidak memiliki tombol “End”, sehingga data yang masuk ke dalam *database* akan terus bertambah seiring dengan pertumbuhan dunia.
- Tidak hanya bertambah karena adanya data baru, *data extraction* harus terus dilakukan untuk mengumpulkan perubahan data yang terjadi dari tiap orang.
- *Data extraction* akan memberikan info sedetail-detailnya dari setiap orang dan menciptakan berbagai macam skenario.



# Langkah-Langkah Penerapan Big Data Analytics

## 3. *Data storing*

- Menyimpan sebuah data, apalagi data yang besar tentunya tidak bisa sembarangan.
- *Storage* untuk penyimpanan data yang baik menyediakan infrastruktur yang memiliki mesin analisis data terbaru. Tak hanya itu, *storage* yang baik juga serta ruang penyimpanan dengan ukuran besar.
- Banyak software yang digunakan untuk menyimpan data berskala besar. Beberapa contohnya adalah Excel, Hadoop, Cloudera, dan Talend.





# Langkah-Langkah Penerapan Big Data Analytics

## 4. *Data Cleaning*

- Data yang didapat dari proses *big data analytics* didapatkan secara keseluruhan melalui internet. Dari 100% data yang sudah didapatkan, kemungkinan ada 30%-40% data yang tidak akurat dan tidak dibutuhkan oleh perusahaan.
- Maka dari itu, dibutuhkan *data cleaning* alias pembersihan data untuk menyaring data mana yang dibutuhkan atau tidak. Dari hal ini, *data analyst* tak perlu repot menganalisa dan mengira-ngira lagi data mana yang harus dipakai.
- Melalui langkah ini, *data analyst* akan langsung mendapatkan data yang sesuai dengan kemauan perusahaan karena sudah tersortir secara otomatis.





# Langkah-Langkah Penerapan Big Data Analytics

## 5. *Data Analysis*

- Bagian terbesar dari *big data analytics* tentu saja analisis data. Ketika menganalisis data, *data analyst* akan masuk ke dalam pola sampai kebiasaan para audiens dan mencari mana yang paling dibutuhkan oleh klien.
- Analisis merupakan proses mempertanyakan pertanyaan yang spesifik dan mencari jawaban yang benar. Qubole dan Statwing diklaim menjadi alat analisis yang sangat *powerful* untuk proses ini.

## 6. *Data consumption*

- Data digunakan untuk keperluan dan kebutuhan yang berbeda oleh perusahaan, pemerintah, instansi bahkan organisasi.
- Pertanyaannya, apakah semua orang dapat mengakses *big data* dan membuka data lewat internet? **Tentu saja tidak bisa. Untuk melakukannya dibutuhkan *data analyst* yang handal dan mengerti bagaimana mengolah sebuah data → PERLU BELAJAR, tidak hanya mendengar namun harus mencari ....**



# Data Driven

- Data dapat menjadi salah satu faktor terbentuknya suatu keputusan. Dan, data juga bisa menjadi acuan suatu pekerjaan.
- Dalam pemanfaatannya, data biasanya dimanfaatkan dengan beberapa cara, salah satunya dengan *data driven*.
- *Data driven* merupakan suatu cara melakukan pekerjaan berdasarkan data yang ada. Jadi, setiap pekerjaan yang kita lakukan haruslah berdasarkan data yang dimiliki



# Cara Menerapkan Data Driven

## 1. Tentukanlah Tujuan

- Agar dapat menerapkan *data driven* dengan baik, kamu harus tahu dulu apa tujuanmu menerapkan hal tersebut. Dengan menentukan tujuan sejak awal sehingga dapat menerapkan *data driven* sesuai dengan apa yang diinginkan.

## 2. Ketahui Sumber-Sumber Data

- Data tak disimpan di sembarang tempat. Tentu ada sumber-sumber tertentu di mana data berasal. Kalau ingin sukses dalam menerapkan *data driven*, maka diketahui di mana sumber-sumber data berada. Dengan melakukan cara ini, dapat membantu dalam pencarian data yang akan digunakan.





# Cara Menerapkan Data Driven

## 3. Pilih Data yang Menjadi Acuan atau Referensi

- Sumber data sangatlah banyak, oleh karena itu perlu dipilih data yang sesuai dengan tujuan. Dengan demikian dapat lebih focus dalam menerapkan *data driven*





# Contoh Implementasi Pengumpulan Data Sederhana

## *Data Mining dan data extraction*

Data mining adalah sebuah proses identifikasi dari *insight* yang berharga dari *database*.

- Anda sedang melakukan riset pasar tentang mobil : family car yang sedang dicari banyak orang di seluruh Indonesia : **Avanza, Xenia, Innova, expander, ertiga**

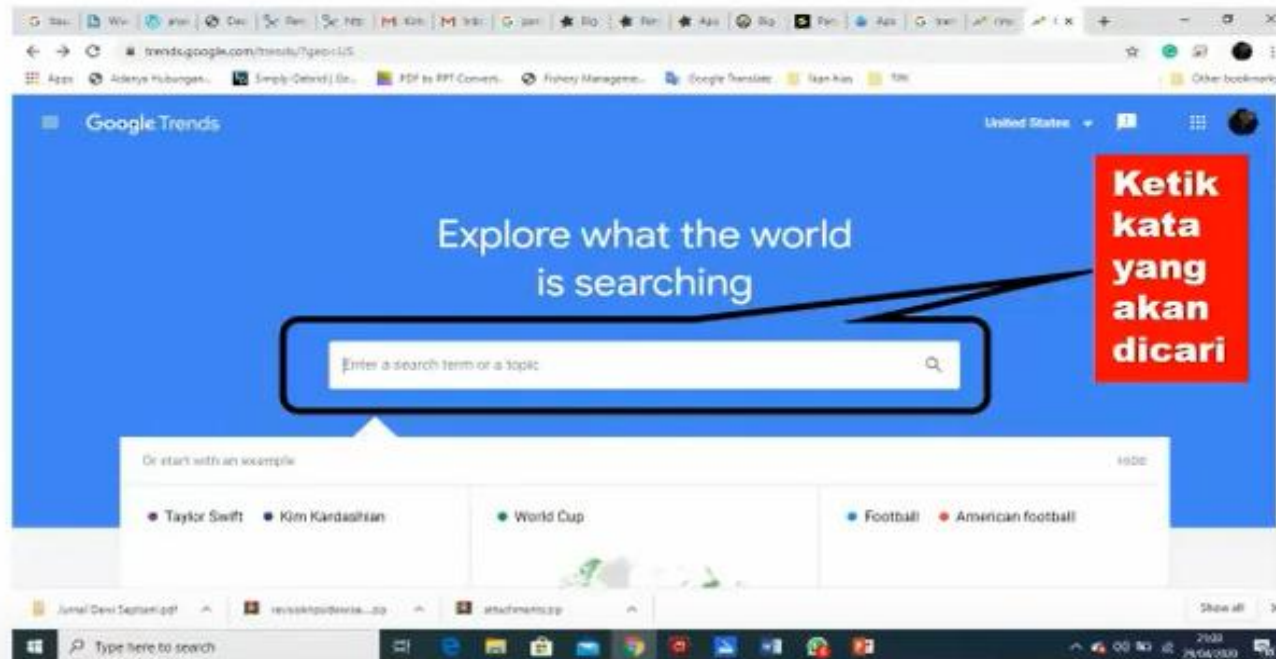
Data extraction adalah sebuah proses pengumpulan data dari halaman web ke dalam *database*.

- Dengan cara biasa → Anda harus mencari data dengan survei ke seluruh Indonesia karena membutuhkan dana yang sangat besar dan waktu yang sangat lama, kesulitan mencari responden → kalah bersaing → kegagalan karena pesaing sudah melakukan penjualan
- Pemanfaatan big data → cepat, efisien dan efektif → cukup buka google trend → <https://trends.google.com/>





# Contoh Implementasi Pengumpulan Data Sederhana



- Buka di laptop <https://trends.google.com/>
- Explore what the world is searching
- Ketik kata yang ingin dicari
- contohnya : ketik avanza



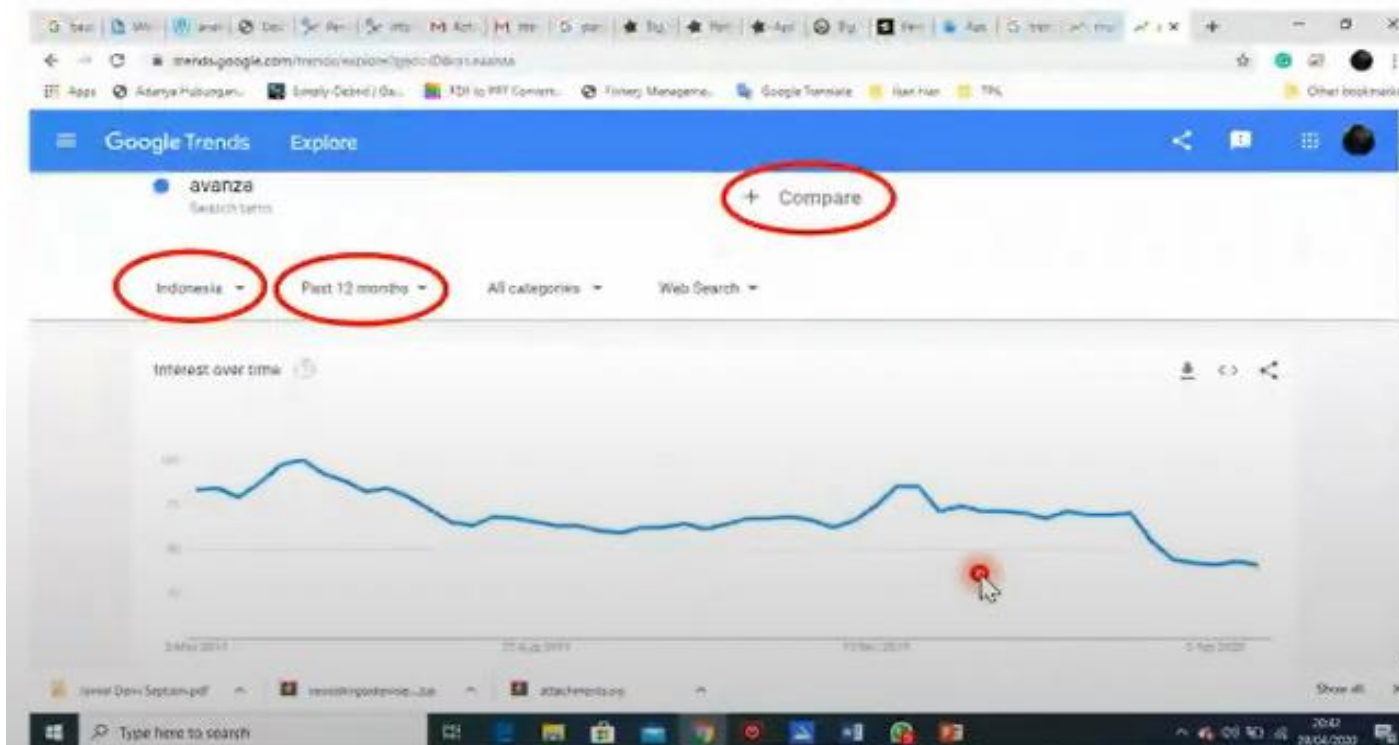
Tampilan di smartphone → biasanya berbeda... Tapi prinsipnya adalah sama



# Contoh Implementasi Pengumpulan Data Sederhana

## Data Collection :

Masukkan kata yang ingin kita lihat contohnya mobil avanza, pilih Negara : Indonesia, pilih waktu : 12 bulan yang lalu (lingkaran merah), maka akan muncul seperti pada gambar. Klik compare untuk membandingkan (lingkaran merah).

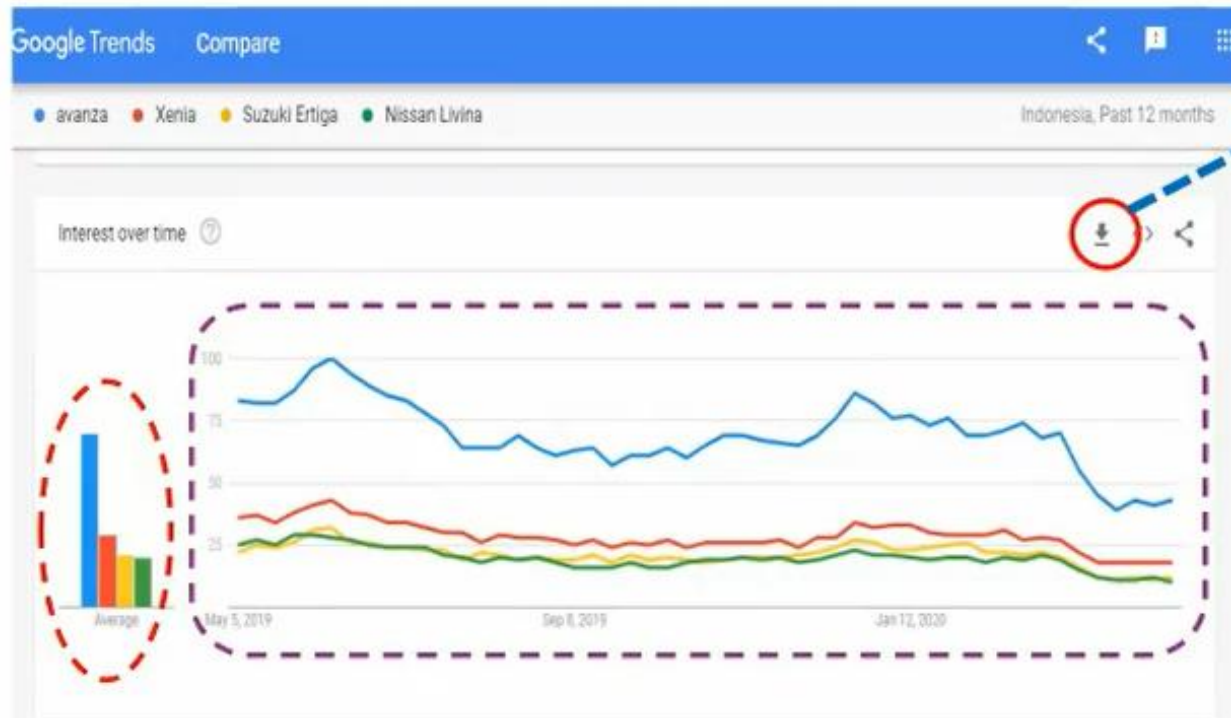


Dapat dilihat dalam kurun waktu 12 bulan, maka trend pencarian untuk topic avanza mengalami fluktuasi dan cenderung menurun

# Contoh Implementasi Pengumpulan Data Sederhana

## Data Storing :

- Penyimpanan data → sudah dilakukan oleh google melalui software Hadoop, dll
- Kita dapat download data ringkasan yang paling tidak berguna untuk analytics
- Jika ingin memperoleh data detail maka harus membayar → minimal U\$60/bulan



Excel File  
yang bisa  
didownload

Data ringkasan yang bersifat umum, karena tidak ada data :

Profiling :

Usia

Jenis kelamin

Pendidikan

Penghasilan

Jenis pekerjaan

Status keluarga

dll

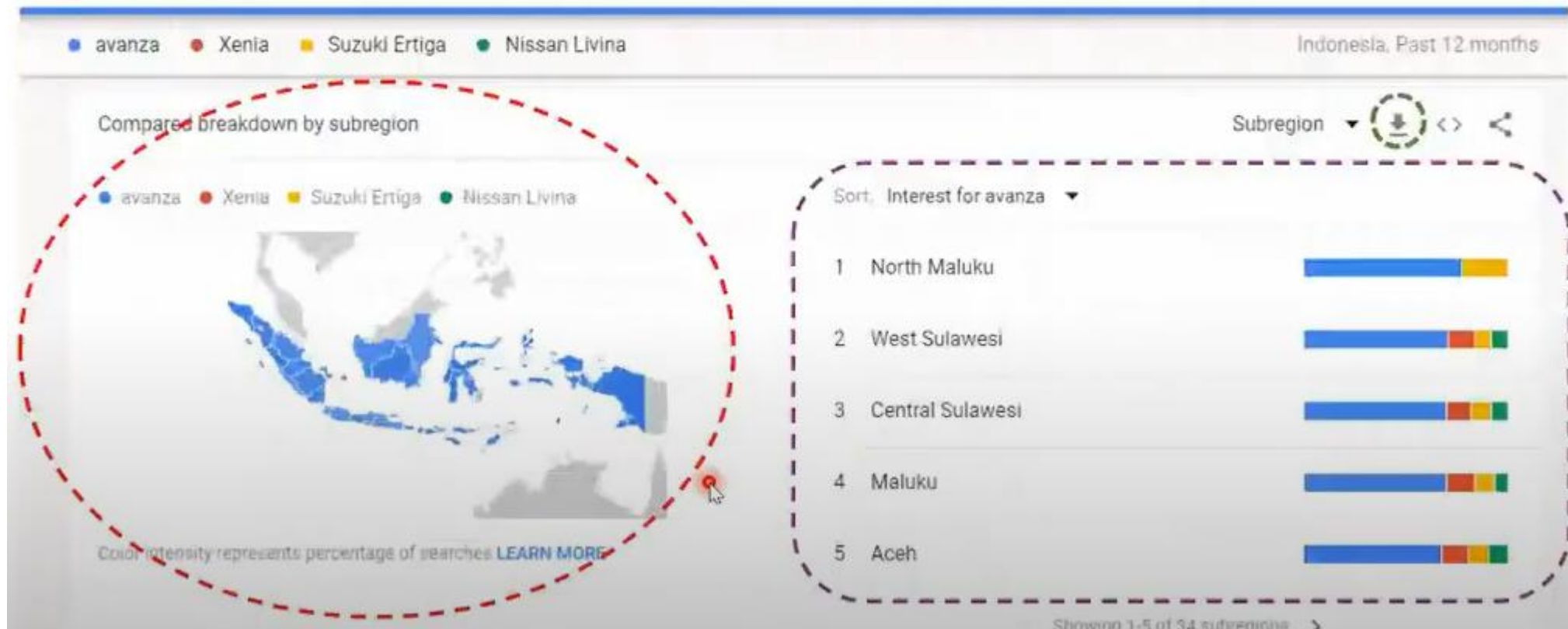




# Contoh Implementasi Pengumpulan Data Sederhana

**Data Cleaning :** Anda memilih apa saja data yang diperlukan untuk dilihat dengan data terbatas ini : compare seluruh mobil

1. Pemetaan keseluruhan informasi mobil yang dicari (garis lingkaran putus-putus warna merah)
2. Lokasi yang paling banyak mencari informasi (garis kotak putus-putus warna ungu)
3. Data excel yang bisa didownload (garis lingkaran putus-putus warna hijau)





# Contoh Implementasi Pengumpulan Data Sederhana

**Data Cleaning :** Anda memilih apa saja data yang diperlukan untuk dilihat dengan data terbatas ini : Avanza (lingkaran hitam putus-putus)

1. Pemetaan keseluruhan mobil (garis lingkaran putus-putus warna merah) → plot lokasi
2. Informasi excel yang bisa didownload → garis lingkaran hijau putus-putus → persentase lokasi yang paling banyak mencari mobil jenis ini
3. Jenis informasi yang dicari → garis kotak ungu putus-putus
4. Informasi excel yang bisa didownload → garis lingkaran orange putus-putus → persentase informasi yang paling banyak mencari mobil jenis ini

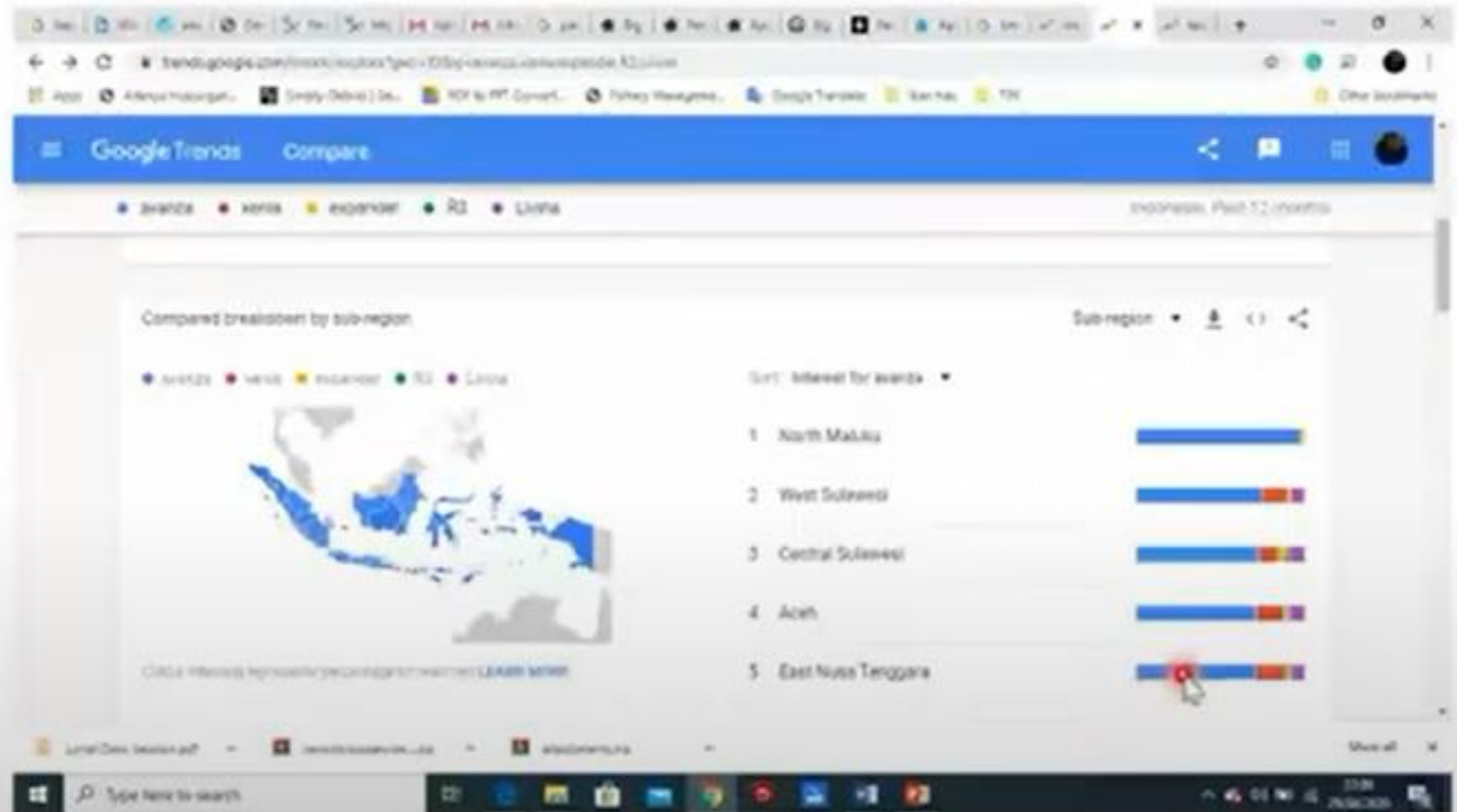




# Contoh Implementasi Pengumpulan Data Sederhana

## Data Analysis :

Perbandingan antar wilayah, yang paling tinggi ketertarikan terhadap avanza adalah Maluku utara







# Contoh Implementasi Pengumpulan Data Sederhana



## Data Consumption:

- Dari informasi tersebut dapat ditarik kesimpulan, bahwa pengguna internet di Indonesia memiliki ketertarikan terhadap avanza lebih tinggi dibandingkan jenis mobil lainnya
- Data yang paling dicari adalah data harga dan spesifikasi → strategi adalah promosi on the spot untuk menjelaskan harga dan spesifikasi
- Strategi yang dapat disarankan untuk Perusahaan mobil expander, Livina dan R3 diantaranya meningkatkan promosi sehingga masyarakat lebih mengenal dan tertarik dengan jenis mobil yang ditawarkan





# Contoh Implementasi Pengumpulan Data Sederhana (buka trends.google.com)



← → ↻ 🏠 🔒 trends.google.com/trends/explore?geo=ID ☆ ⬆️ 🌐 🏠 📱 📅 📌 📖

Apps detikcom - Informa... G Jurnal Bank Email & Com Kullah Math & Statistic Beasiswa Carbon IT and PC \* Other bookmarks Reading list

☰ Google Trends Explore 🔗 💬 ⋮ 👤

Add a search term

Indonesia ▼ Past 12 months ▼ All categories ▼ Web Search ▼

Search topics ⓘ Rising ▼ ⬇️ <> 🔗

1	2021 - Topic	Breakout
2	Pre-Employment Card - Topic	+1,050%
3	Coordinating Ministry for Economic Affairs - ...	+500%
4	TikTok - Topic	+450%

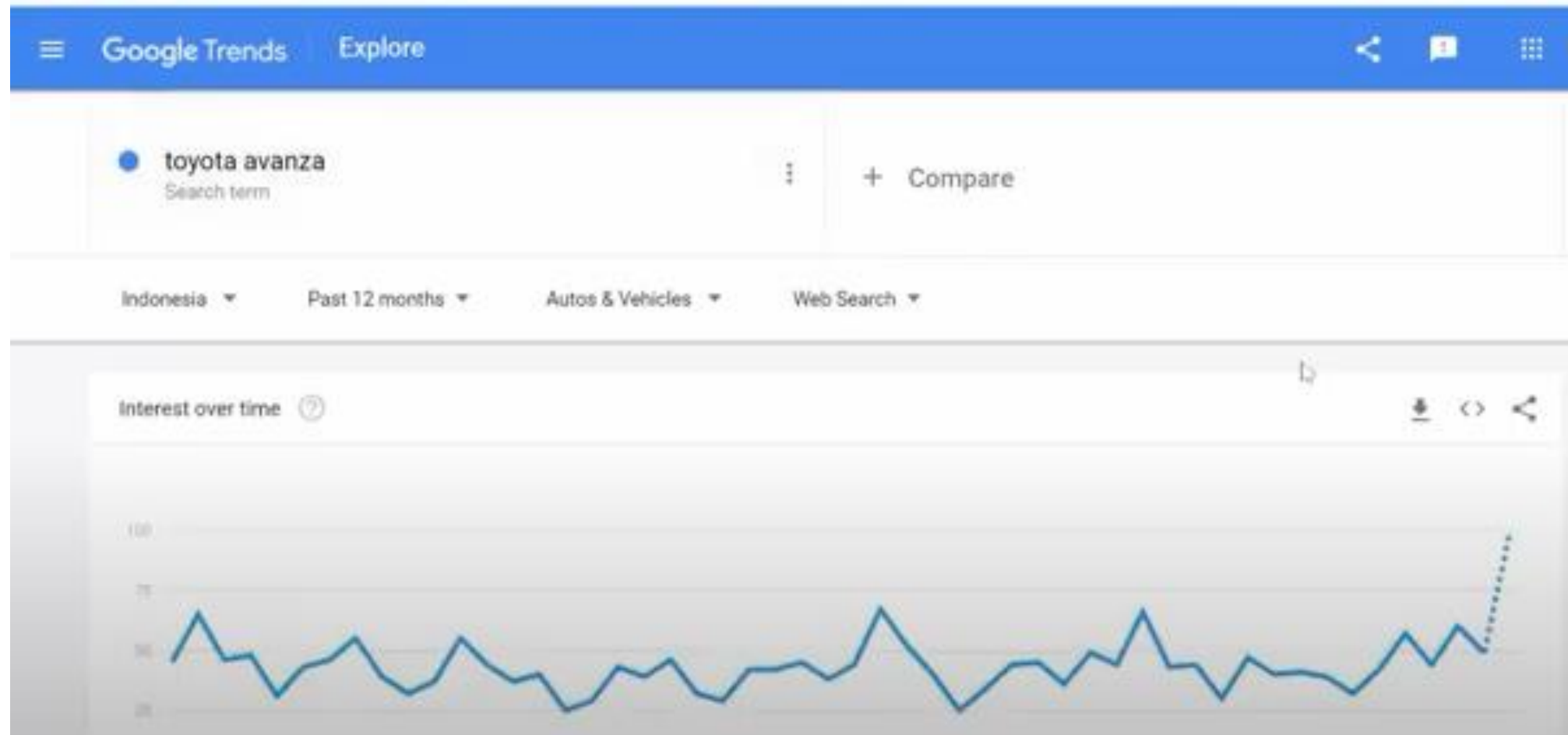
Search queries ⓘ Rising ▼ ⬇️ <> 🔗

1	snaptik	Breakout
2	download video tiktok	Breakout
3	ssstiktok	Breakout
4	tiktok watermark	+3,950%



# Contoh Implementasi Pengumpulan Data Sederhana

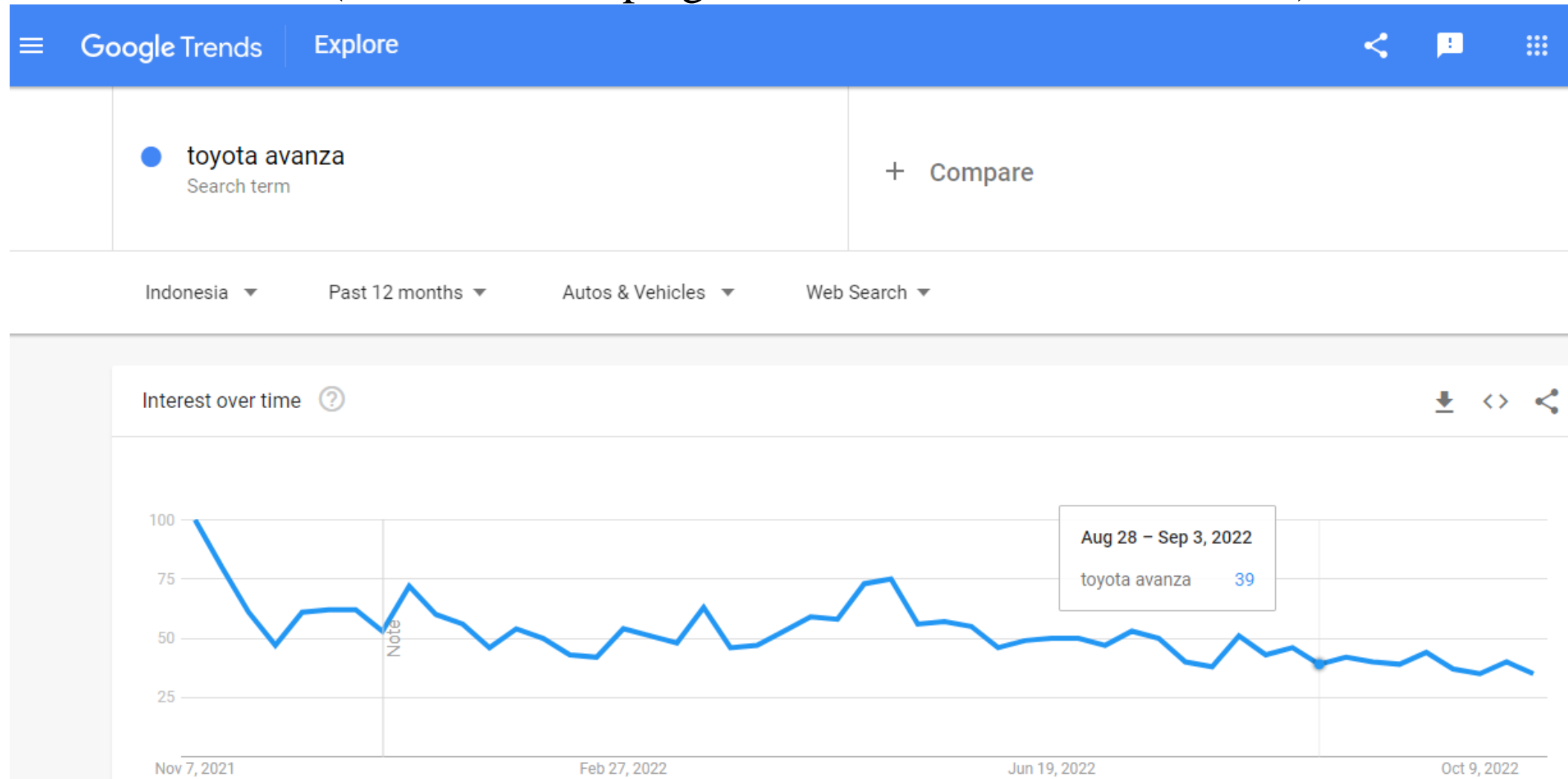
(ketikkan di pencarian : Toyota Avanza, di bagian kategori pilih Auto&Vehicles)





# Contoh Implementasi Pengumpulan Data Sederhana

(Terlihat kurva pergerakan dalam 12 bulan terakhir)



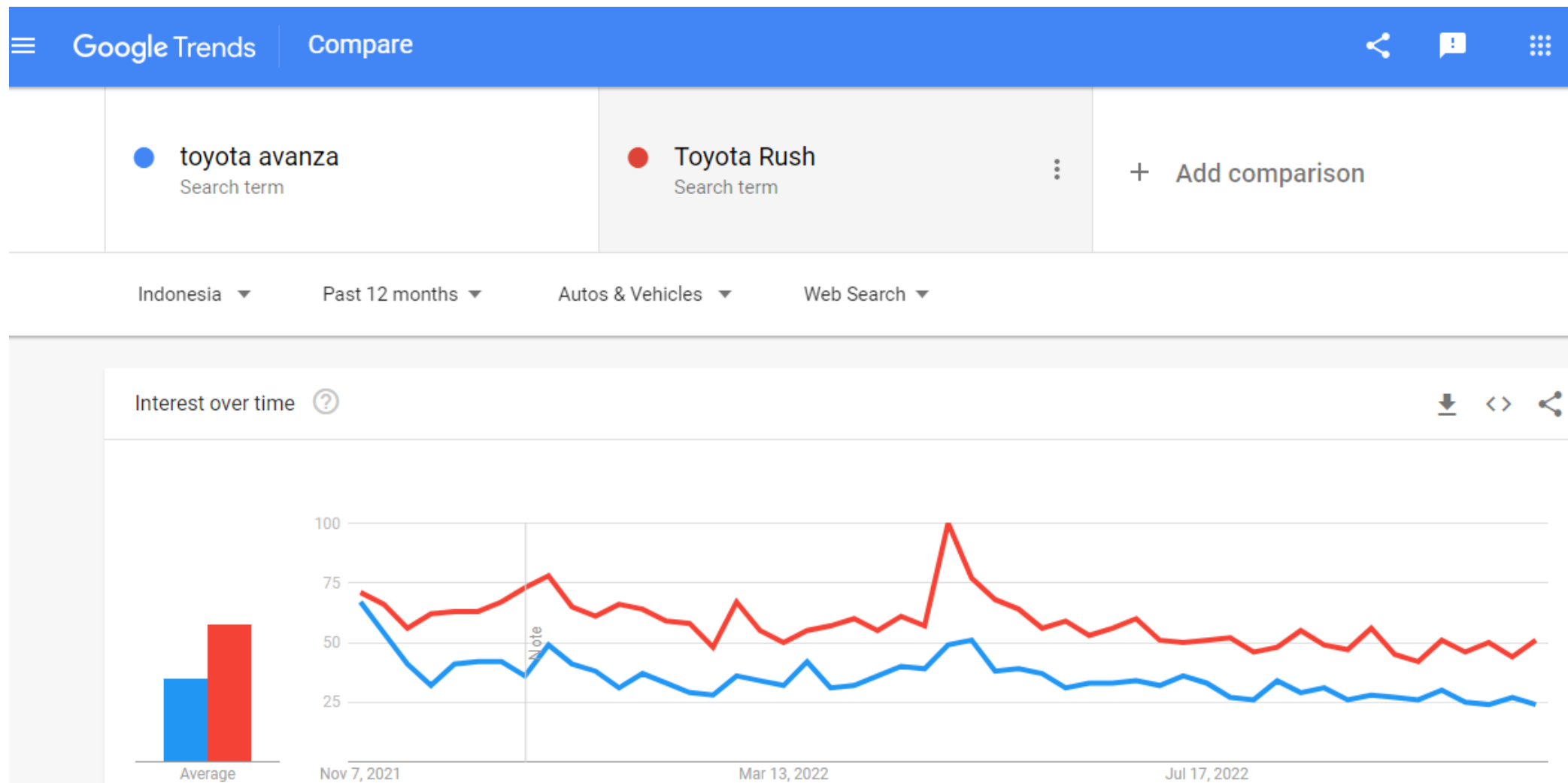




# Contoh Implementasi Pengumpulan Data Sederhana



(Klik bagian Compare, ketikkan Toyota Rush, terlihat perbandingannya)

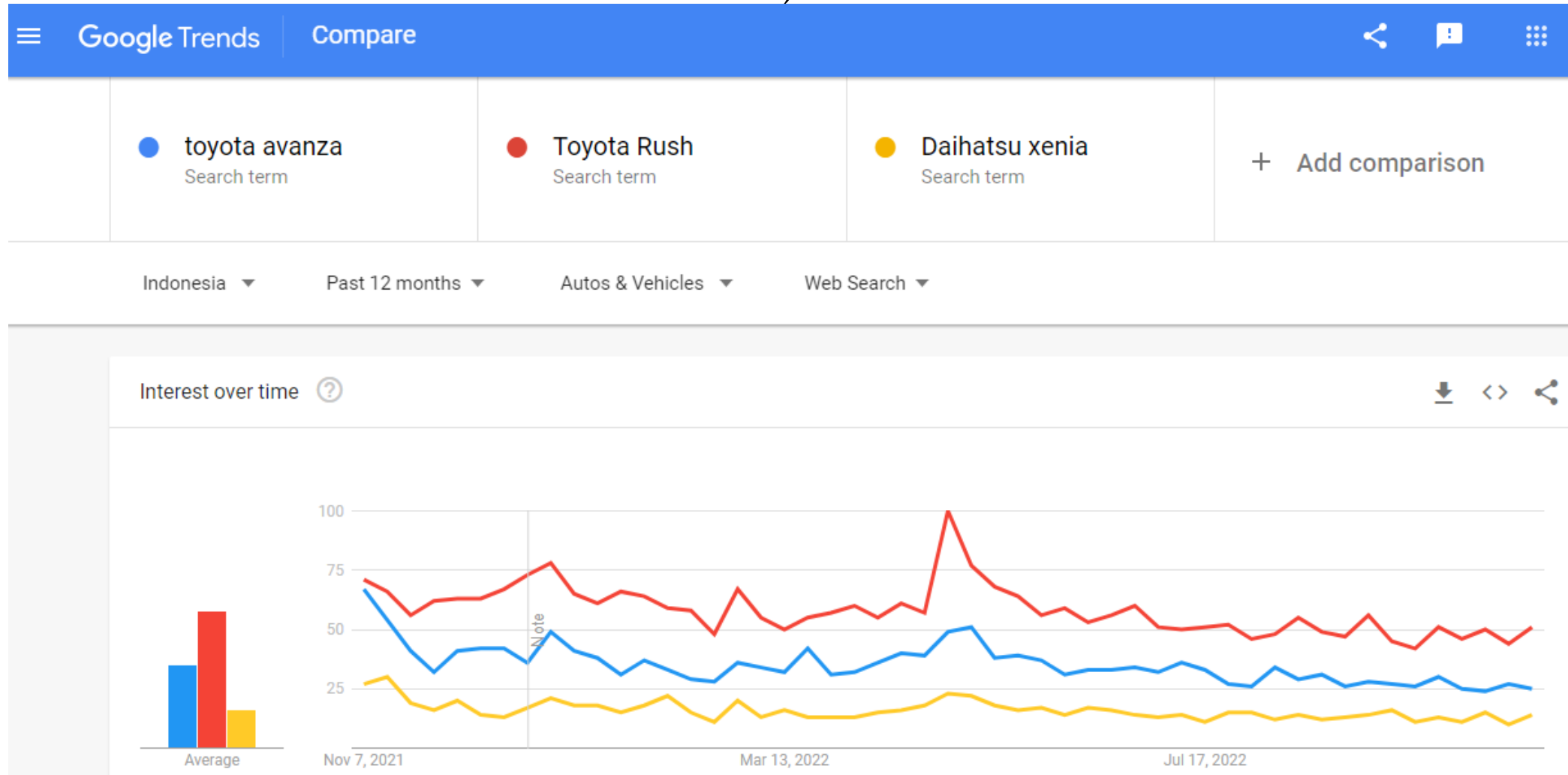




# Contoh Implementasi Pengumpulan Data Sederhana



(Klik lagi bagian Compare, ketikkan Daihatsu Xenia, ternyata penjualan xenia di bawah Avanza dan Rush)

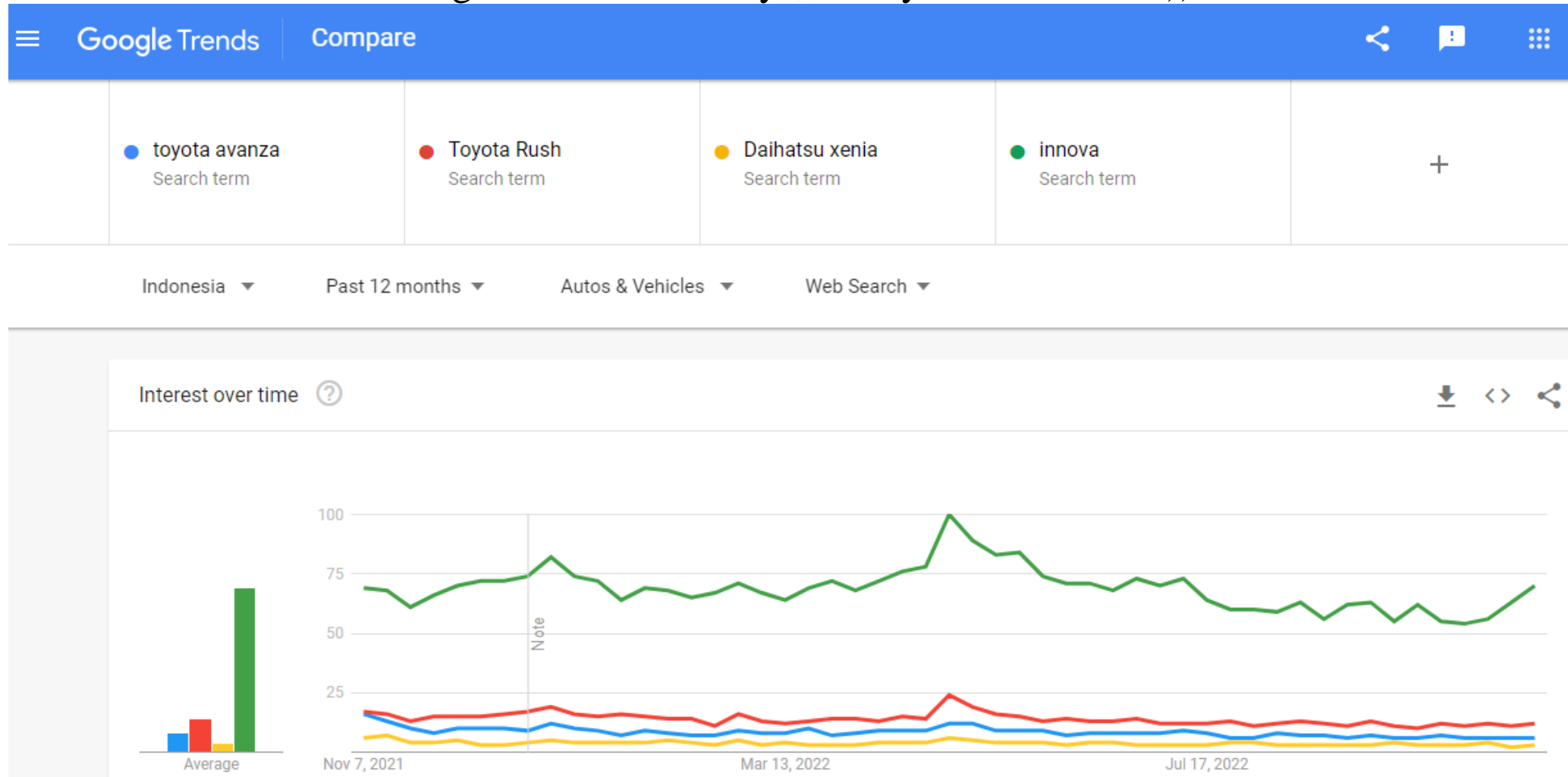




# Contoh Implementasi Pengumpulan Data Sederhana



(Klik lagi bagian Compare, ketikkan Toyota Innova, analisis penjualannya di atas ketiga mobil sebelumnya, trennya ada di innova))

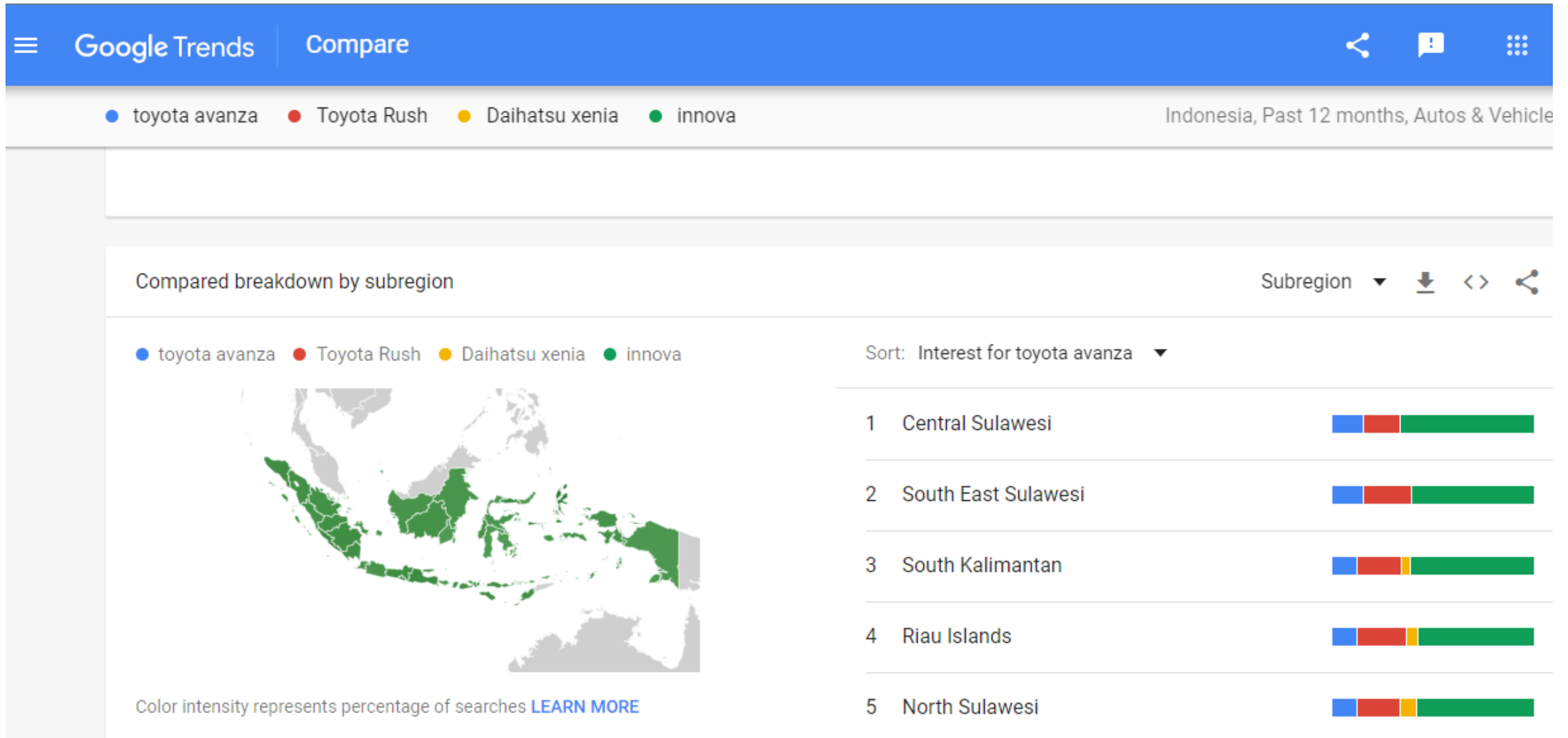






# Contoh Implementasi Pengumpulan Data Sederhana

(Rekapitulasi sebaran per wilayah)





# Contoh Implementasi Pengumpulan Data Sederhana

(Rekapitulasi sebaran per wilayah)

