

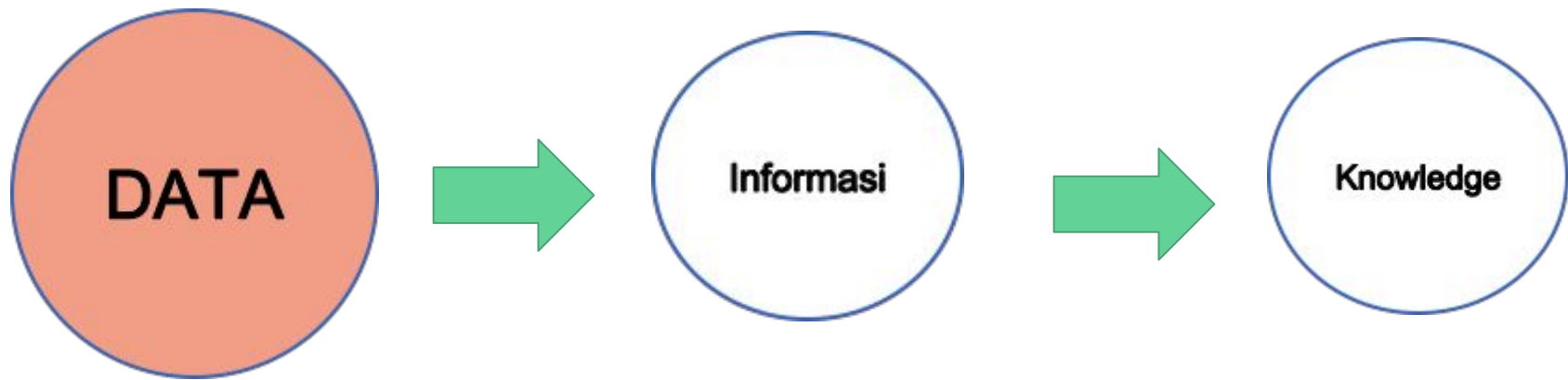
Internet Data Collection

Mutiara Auliya Khadija, S.Kom, M.Eng

mutiaraauliya@staff.uns.ac.id

Goals:

- Konsep data pada internet
- Internet Data Collection
- Data Gathering
- Mencari data yang sudah ada (public dataset)
- Memahami data yang anda miliki
- Praktik mencari dataset sesuai project



Internet Data Collection

- Semua data bisa anda dapatkan di internet baik itu gratis maupun berbayar
- Gratis: langsung cari di Google, data berita di detik.com, data Twitter, dll
- Berbayar: data jurnal berbayar, logo berbayar, tema berbayar

The image shows a promotional banner for Pngtree's 5th Anniversary. At the top, it says "Pngtree 5th Anniversary 87% OFF Lifetime Plan". Below this, there is a "Discount Countdown" section showing a timer with 59 minutes, 26 seconds, and 77 milliseconds remaining. To the right of the timer, the price for the "Pngtree Lifetime Plan" is shown as "us\$ 999" crossed out and "us\$ 129" in a large font, with a "BUY NOW" button below it. On the left, there is a "Privilege" section with three radio button options: "Unlimited downloads Forever", "Permanent Access to all products Forever", and "Valid commercial licenses Forever". The background of the banner features various cartoon illustrations of people and animals.

The image shows a Google search result for "sekolah vokasi". The search bar at the top contains the text "sekolah vokasi". Below the search bar, there are three search results. The first result is for "Sekolah Vokasi - Universitas Gadjah Mada" with a URL "https://ugm.ac.id" and a small image of a building. The second result is for "Sekolah Vokasi UGM - Universitas Gadjah Mada" with a URL "https://sv.ugm.ac.id" and a small image of a building. The third result is for "Sekolah Vokasi Fakultas dan Pascasarjana - Universitas ..." with a URL "https://uns.ac.id" and a small image of a building. At the bottom of the page, there is a small text "Open all personal permissions?".

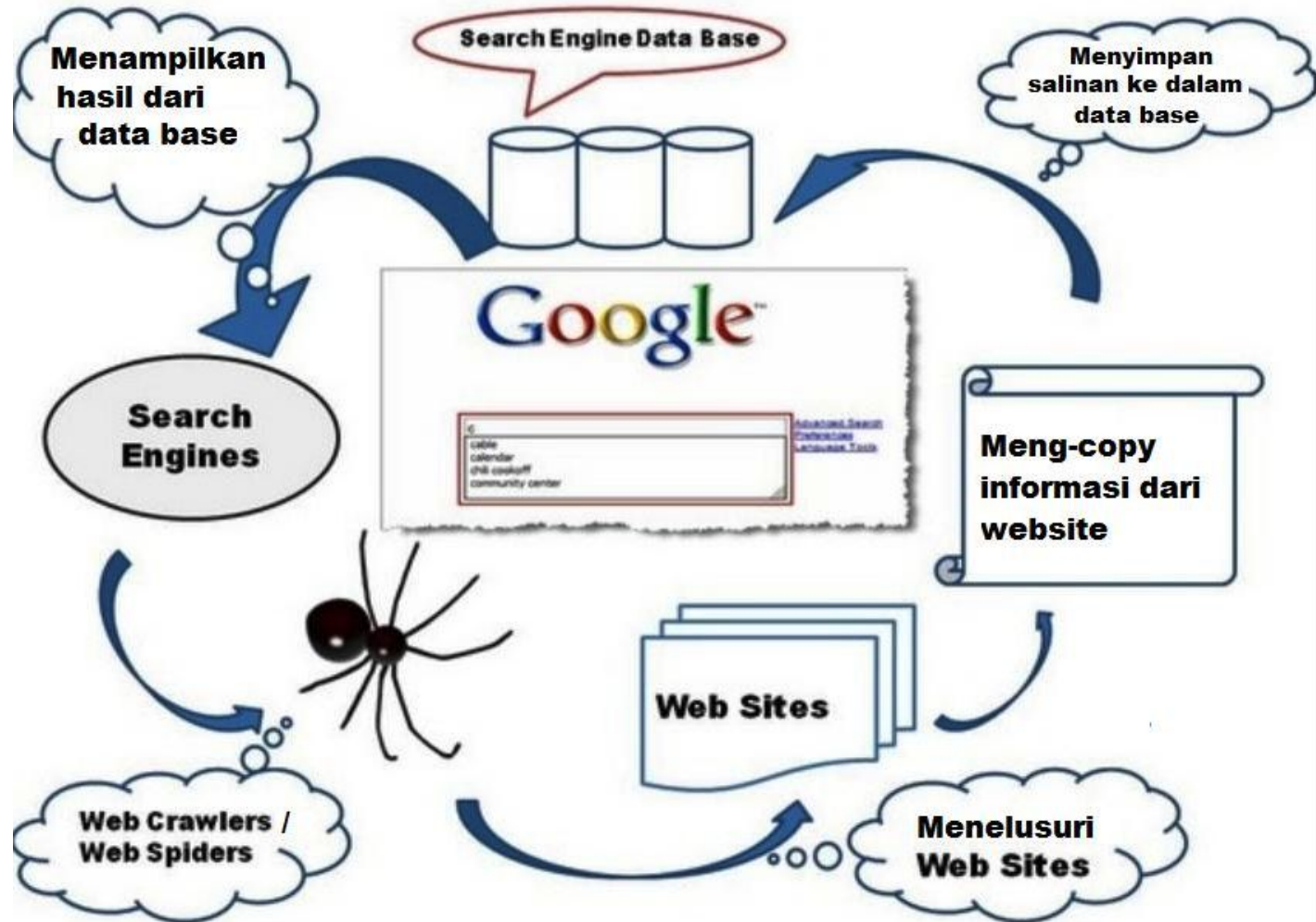
Apa sih Google
itu?



Google termasuk Search Engine!

- **Search engine** atau apabila diartikan dalam bahasa Indonesia adalah mesin pencari pada dasarnya merupakan program berbasis web yang diperuntukkan untuk mencari informasi di dalam World Wide Web (www).
- Pengguna internet memasukkan kata kunci informasi yang ingin dicari ke dalam perangkat search engine lalu perangkat tersebut akan mencari dan mengidentifikasi kata kunci tersebut ke dalam database informasi di internet.
- Contoh: Google, Microsoft Bing, Yahoo!, Baidu, Yandex

Cara Kerja Search Engine



Data Collection/ Data Gathering

- **Data collection/ Data gathering** adalah proses pengumpulan, pengukuran, dan analisis berbagai tipe informasi menggunakan teknik berstandar. Tujuan utama data collection adalah untuk mengumpulkan informasi dan data terpercaya sebanyak-banyaknya, yang kemudian dianalisis untuk membuat sebuah keputusan bisnis yang krusial.
- Proses data gathering bisa dilakukan dengan wawancara, kuesioner, internet, arsip, perpustakaan, dll

Dataset

- Dataset adalah istilah informal yang merujuk pada kumpulan data. Secara umum, dataset berisi lebih dari satu variabel dan menyangkut suatu topik tertentu.
- Pengertian dataset adalah sebuah kumpulan data yang berasal dari informasi-informasi pada masa lalu dan siap untuk dikelola menjadi sebuah informasi baru.

Jenis Dataset

1. Private Dataset

Private dataset adalah dataset yang dapat diambil dari sebuah organisasi yang akan dilakukan sebagai objek penelitian, seperti data bank, rumah sakit, sekolah, universitas, perusahaan, dan lain sebagainya

2. Public Dataset

Public dataset adalah dataset yang bisa diambil dari repository publik dan dibuka untuk umum.

Mencari Public Dataset

Mencari Public Dataset

data.go.id

data.go.id/home

Beranda Dataset Berita Tentang [Masuk](#)

Data Indonesia, Dalam Satu Portal

Temukan data-data Pemerintah dengan mudah!

Semua Data Cari Data Publik Indonesia

Topik

- Pertahanan dan Luar Negeri
- Ekonomi dan Industri
- Pembangunan Daerah
- Perlindungan Sosial dan Kesehatan
- Ketertiban Umum dan Keselamatan
- Pendidikan dan Tenaga Kerja
- Lingkungan dan Sumber Daya Alam
- Budaya dan Agama
- Pemerintahan Umum
- Pendukung Umum

Mencari Public Dataset

[kaggle.com](https://www.kaggle.com)

The screenshot shows the Kaggle website interface. The browser address bar displays 'kaggle.com/datasets?search=indonesia'. The left sidebar contains navigation options: Home, Competitions, Datasets (highlighted), Code, Discussions, Courses, and More. The main content area features a search bar with 'indonesia' entered, a 'Filters' button, and a category filter bar with options like Computer Science, Education, Classification, Computer Vision, NLP, Data Visualization, and Pre-Trained Model. Below this, there are 289 datasets listed. The top four datasets are:

Dataset Name	Author	Updated	Usability	Files	Size	Hotness
Indonesia-Coronavirus	Ardiansyah	2 years ago	9.4	9 Files (CSV)	10 kB	146
COVID-19 Indonesia Dataset	Hendratno	5 months ago	10.0	1 File (CSV)	924 kB	233
Lazada Indonesian Reviews	Griko Nibras	3 years ago	10.0	3 Files (CSV, other)	7 MB	59
Kue Indonesia	Ilham Firdausi Putra	2 years ago	8.8	1845 Files (other)	318 MB	16

Mencari Public Dataset

<https://www.bps.go.id/>

The screenshot shows the homepage of the Indonesian Central Bureau of Statistics (BPS). The header includes the BPS logo, the text "BADAN PUSAT STATISTIK", and navigation links for "Indonesia | English", "Cari", "Manual", "Tautan", "Peta Situs", and "S&K". A main navigation bar contains links for "Beranda", "Tentang Kami", "Berita", "Senarai Rencana Terbit", "Publikasi", "Berita Resmi Statistik", "Layanan", and "PPID".

On the left side, there are three main menu categories with expandable options:

- Sosial dan Kependudukan**: Gender, Geografi, Iklim. [Selengkapnya...](#)
- Ekonomi dan Perdagangan**: Ekspor-Import, Energi, Harga Eceran. [Selengkapnya...](#)
- Pertanian dan Pertambangan**: Hortikultura, Kehutanan, Perikanan. [Selengkapnya...](#)

At the top right, there is a search bar labeled "Cari Data.." with a "Cari" button. Below it is a "Website BPS Provinsi" dropdown menu.

The main content area features a banner for "TUJUAN PEMBANGUNAN BERKELANJUTAN | SUSTAINABLE DEVELOPMENT GOALS". Below this is a section titled "INFORMASI TERBARU" and "Berita Resmi Statistik".

The "Berita Resmi Statistik" section lists two news items:

- 17 Mei 2022 (Download Bahan Tayang)**
 - Ekspor April 2022 mencapai US\$27,32 miliar, naik 3,11 persen dibanding Maret 2022 & Import April 2022 senilai US\$19,76 miliar, turun 10,01 persen dibanding Maret 2022
 - Upah Nominal Harian Buruh Tani Nasional April 2022 Naik Sebesar 0,29 Persen
- 09 Mei 2022 (Download Bahan Tayang)**
 - Inflasi terjadi pada April 2022 sebesar 0,95 persen. Inflasi tertinggi terjadi di Tanjung Pandan sebesar 2,58 persen.
 - Jumlah kunjungan wisman ke Indonesia melalui pintu masuk utama pada Maret 2022 mencapai 40,79 ribu kunjungan dan Jumlah penumpang angkutan udara internasional pada Maret 2022 naik 84,40 persen.
 - Nilai Tukar Petani (NTP) April 2022 sebesar 108,46 atau turun 0,76 persen
 - Pada April 2022, Indeks Harga Perdagangan Besar (IHPB) Umum Nasional naik 0,97 persen
 - Harga Produsen Mengalami Inflasi 2,59 Persen di Triwulan I-2022

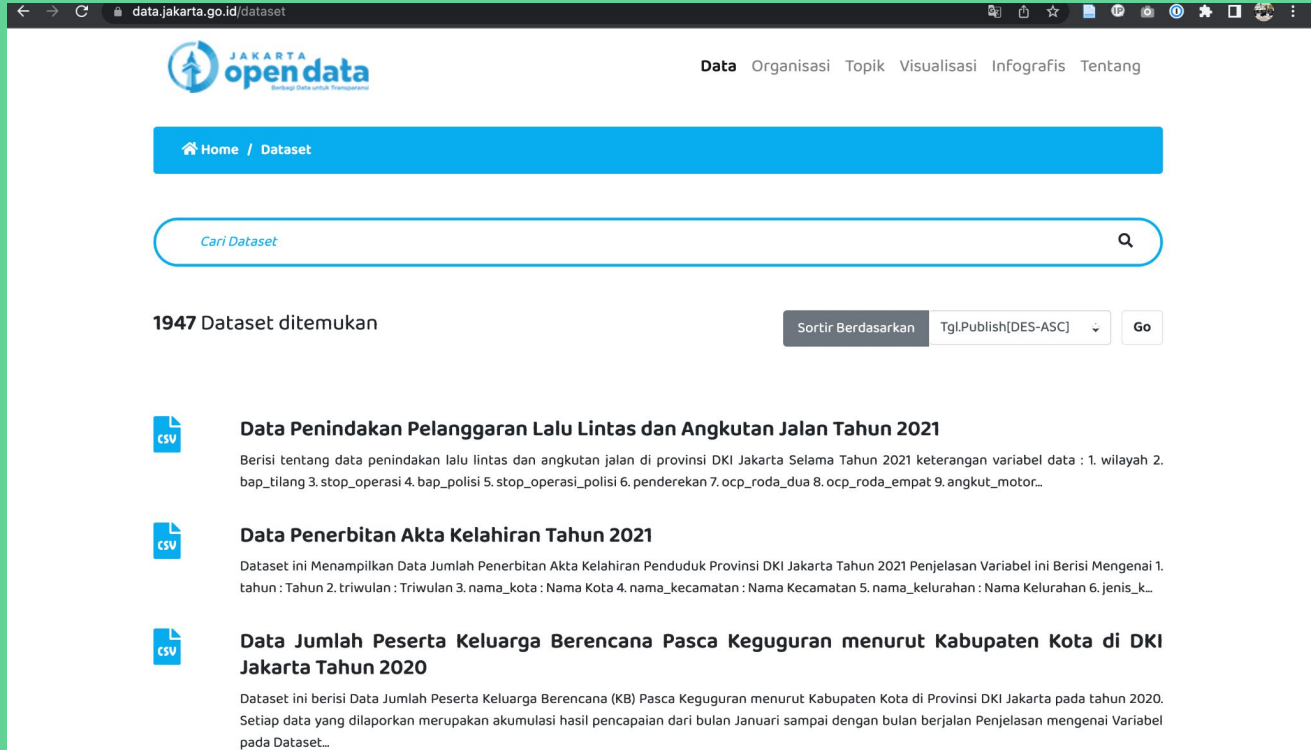
On the right side, there are several promotional tiles:

- Informasi Seputar SP2020** and **Sensus Penduduk 2020**
- Survei COVID-19 BPS** and **Publikasi COVID-19 BPS**
- DATA PERENCANAAN PEMBANGUNAN**
- Statistical Geographic Information System**
- Satu Data Indonesia**
- PELAYANAN STATISTIK TERPADU pst.bps.go.id**
- Aplikasi Allstats BPS**

At the bottom left, there is a "Galeri Infografis" section with a link to "Cara Mendapatkan Data BPS".

Mencari Public Dataset

<https://data.jakarta.go.id/dataset>

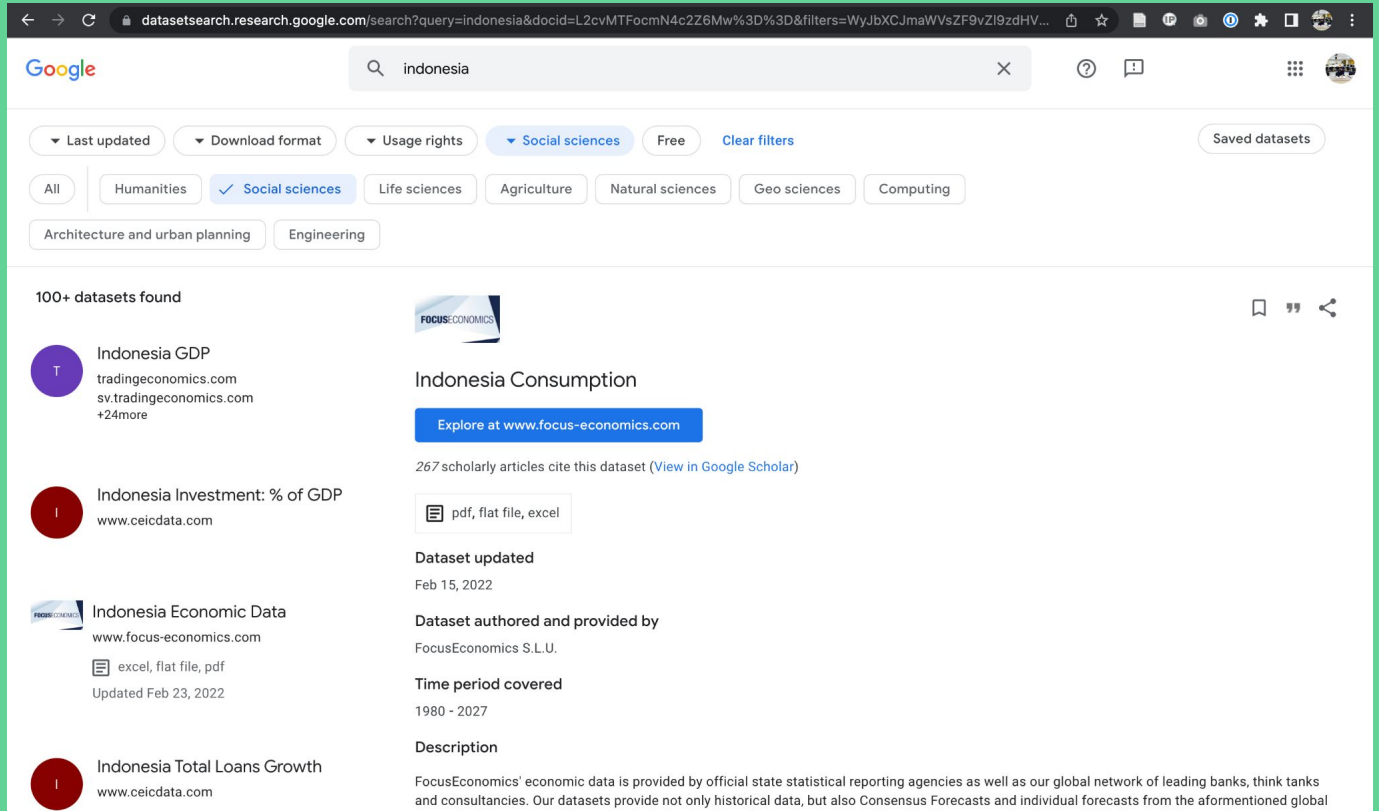


The screenshot shows the Jakarta Open Data website interface. At the top, there is a navigation menu with links for 'Data', 'Organisasi', 'Topik', 'Visualisasi', 'Infografis', and 'Tentang'. Below the navigation is a blue breadcrumb trail showing 'Home / Dataset'. A search bar contains the text 'Cari Dataset' and a magnifying glass icon. Below the search bar, it indicates '1947 Dataset ditemukan'. There are two sorting options: 'Sortir Berdasarkan' and 'Tgl.Publish[DES-ASC]'. Three dataset entries are visible, each with a CSV icon and a title:

- Data Penindakan Pelanggaran Lalu Lintas dan Angkutan Jalan Tahun 2021**
Berisi tentang data penindakan lalu lintas dan angkutan jalan di provinsi DKI Jakarta Selama Tahun 2021 keterangan variabel data : 1. wilayah 2. bap_tilang 3. stop_operasi 4. bap_polisi 5. stop_operasi_polisi 6. penderekan 7. ocp_roda_dua 8. ocp_roda_empat 9. angkut_motor...
- Data Penerbitan Akta Kelahiran Tahun 2021**
Dataset ini Menampilkan Data Jumlah Penerbitan Akta Kelahiran Penduduk Provinsi DKI Jakarta Tahun 2021 Penjelasan Variabel ini Berisi Mengenai 1. tahun : Tahun 2. triwulan : Triwulan 3. nama_kota : Nama Kota 4. nama_kecamatan : Nama Kecamatan 5. nama_kelurahan : Nama Kelurahan 6. jenis_k...
- Data Jumlah Peserta Keluarga Berencana Pasca Keguguran menurut Kabupaten Kota di DKI Jakarta Tahun 2020**
Dataset ini berisi Data Jumlah Peserta Keluarga Berencana (KB) Pasca Keguguran menurut Kabupaten Kota di Provinsi DKI Jakarta pada tahun 2020. Setiap data yang dilaporkan merupakan akumulasi hasil pencapaian dari bulan Januari sampai dengan bulan berjalan Penjelasan mengenai Variabel pada Dataset...

Mencari Public Dataset

<https://datasetsearch.research.google.com/>



The screenshot shows a search for 'indonesia' on the Google Dataset Search platform. The search results are filtered by 'Social sciences' and 'Free'. The top result is 'Indonesia Consumption' from FocusEconomics, which is cited by 267 scholarly articles. The dataset is updated as of February 15, 2022, and covers the period from 1980 to 2027. It is available in PDF, flat file, and Excel formats. Other results include 'Indonesia GDP' and 'Indonesia Investment: % of GDP'.

datasetsearch.research.google.com/search?query=indonesia&docid=L2cvMTFocmN4c2Z6Mw%3D%3D&filters=WyJlbXcJmaWVsZF9vZI9zdHV...

Google indonesia

Last updated Download format Usage rights Social sciences Free Clear filters Saved datasets

All Humanities Social sciences Life sciences Agriculture Natural sciences Geo sciences Computing

Architecture and urban planning Engineering

100+ datasets found

Indonesia Consumption
tradingeconomics.com
sv.tradingeconomics.com
+24more

Explore at www.focus-economics.com

267 scholarly articles cite this dataset (View in Google Scholar)

pdf, flat file, excel

Dataset updated
Feb 15, 2022

Dataset authored and provided by
FocusEconomics S.L.U.

Time period covered
1980 - 2027

Description
FocusEconomics' economic data is provided by official state statistical reporting agencies as well as our global network of leading banks, think tanks and consultancies. Our datasets provide not only historical data, but also Consensus Forecasts and individual forecasts from the aforementioned global

Indonesia GDP
tradingeconomics.com
sv.tradingeconomics.com
+24more

Indonesia Investment: % of GDP
www.ceicdata.com

Indonesia Economic Data
www.focus-economics.com
excel, flat file, pdf
Updated Feb 23, 2022

Indonesia Total Loans Growth
www.ceicdata.com

Praktek Analisis Dataset

1. [Data Destinasi Wisata Indonesia](#)
2. [Glassdoor Jobs Data Analysis](#)

Apa sih goal kita
ketika sudah
mencari
dataset?

1. Dibuat Infografis yang menarik
 2. Membuat dashboard interaktif dengan Google Data Studio
-

Contoh Dashboard Interaktif

[Dashboard Data Imunisasi BCG dan DPT Provinsi DKI Jakarta Tahun 2019](#)