



DISKOMINFO
KALTIM

KAJIAN BIG DATA DAN DATA ANALYTIC: URGENSI DATA SCIENCE DALAM TRANSFORMASI DIGITAL SATU DATA KALIMANTAN TIMUR



**Dinas Komunikasi dan Informatika
Provinsi Kalimantan Timur
Tahun 2024**



DISKOMINFO
KALTIM

KAJIAN BIG DATA DAN DATA ANALYTIC: URGENSI DATA SCIENCE DALAM TRANSFORMASI DIGITAL SATU DATA KALIMANTAN TIMUR



**Dinas Komunikasi dan Informatika
Provinsi Kalimantan Timur
Tahun 2024**

KAJIAN *BIG DATA* DAN *DATA ANALYTIC*: URGENSI *DATA SCIENCE* DALAM TRANSFORMASI DIGITAL SATU DATA KALIMANTAN TIMUR

ISBN :

Ukuran Buku : 21 cm x 29,7 cm (A4)

Jumlah Halaman : xiii + 117 halaman

Naskah:

Bidang Statistik, Dinas Komunikasi dan Informatika
Provinsi Kalimantan Timur

Penyunting:

Bidang Statistik, Dinas Komunikasi dan Informatika
Provinsi Kalimantan Timur

Desain Cover dan Layout:

Bidang Statistik, Dinas Komunikasi dan Informatika
Provinsi Kalimantan Timur

Sumber Ilustrator:

<https://www.canva.com>

Diterbitkan dan Dicetak oleh:

Bidang Statistik

© Dinas Komunikasi dan Informatika Provinsi Kalimantan Timur

Dilarang mengumumkan, mendistribusikan, mengkomunikasikan, dan atau menggandakan sebagian atau seluruh isi buku ini untuk tujuan komersial tanpa izin tertulis dari Penerbit Dinas Komunikasi dan Informatika Provinsi Kalimantan Timur.

TIM PENYUSUN



Pengarah

H. Muhammad Faisal, S.Sos., M.Si.

Kepala Dinas Komunikasi dan Informatika Provinsi Kalimantan Timur.



Penanggung jawab

H. M. Adrie Dirga Sagita, S.Sos., M.Si.

Kepala Bidang Statistik, Dinas Komunikasi dan Informatika Provinsi Kalimantan Timur.



Penulis

H. M. Adrie Dirga Sagita, S.Sos., M.Si.

Ika Wahyuni, S.E.

Untung Maryono, S.T., M.M.

Nazaruddin, S.Kom.

Nadia Paramitha N, S.T.



Penyunting

Dwi Ayu Anggraini, S.Stat.

Irfan Fadil, S.Stat.

Iis Yustiani, S.Stat.



Pengolah Data

Hari Adam Suharto.

Aulia Trihandini, S.Stat.

Khairina Zulfathana Mahmuda, S.Mat.



Desain Sampul dan Tata Letak

Febri Irawan, S.E.

Eka Indah Justisiani, S.I.Kom.

Rizmalani Syawitri, S.Sos.

Vebi Regina, S.E.



Penerbit

Dinas Komunikasi dan Informatika Provinsi Kalimantan Timur

Alamat Redaksi

Bidang Statistik

Jalan Basuki Rahmat Nomor 41, Kelurahan Sungai Pinang Luar, Kecamatan Samarinda Kota, Samarinda, Kalimantan Timur, 75121.

Website : <https://diskominfo.kaltimprov.go.id>

Email : diskominfo@kaltimprov.go.id

Telp/Fax : 0541-731963

KATA PENGANTAR

Dengan mengucapkan syukur kepada Allah SWT, Tuhan Yang Maha Esa, kami telah menyelesaikan kegiatan yang berjudul "**Kajian *Big Data* dan *Data Analytic*: Urgensi *Data Science* dalam Transformasi Digital Satu Data Kalimantan Timur**". Kajian ini disusun dengan harapan dapat memberikan wawasan yang komprehensif mengenai pentingnya pemanfaatan *big data* dan *data analytics* dalam mendukung transformasi digital di Provinsi Kalimantan Timur. Era digital telah membawa perubahan besar dalam berbagai aspek kehidupan, termasuk dalam pengelolaan data. *volume*, kecepatan, dan variasi data yang dihasilkan terus meningkat, menciptakan fenomena yang dikenal sebagai *big data*. Untuk dapat mengelola dan memanfaatkan data tersebut secara efektif, diperlukan pendekatan yang canggih, yaitu *data analytics*. Dalam konteks ini, *data science* menjadi disiplin ilmu yang sangat penting untuk mengolah data menjadi informasi yang berguna bagi pengambilan keputusan.

Kalimantan Timur, sebagai Provinsi dengan potensi besar di sektor energi, pertambangan, pertanian, dan pariwisata, memerlukan strategi pengelolaan data yang terintegrasi dan akurat. Program "**Satu Data Kalimantan Timur**" bertujuan untuk menyatukan berbagai sumber data yang ada, sehingga dapat mendukung pengambilan keputusan yang lebih baik dan tepat sasaran. Melalui pemanfaatan *big data* dan *data analytics*, pemerintah daerah dapat meningkatkan efisiensi, transparansi, dan akuntabilitas dalam pelaksanaan berbagai program pembangunan. Kajian ini tidak hanya membahas konsep dasar *big data* dan *data analytics*, tetapi juga menyoroti peran penting *data science* dalam mengoptimalkan pengelolaan data di Kalimantan Timur. Dengan kemampuan untuk melakukan analisis prediktif dan mendalam, *data science* dapat memberikan wawasan yang signifikan untuk perumusan kebijakan yang lebih efektif dan responsif terhadap dinamika kebutuhan masyarakat.

Harapan kami hasil kajian ini dapat memberikan kontribusi bagi pengambil kebijakan, akademisi, dan praktisi dalam memahami serta memanfaatkan *big data* dan *data analytics* untuk mendukung transformasi digital yang berkelanjutan di Provinsi Kalimantan Timur. Ucapan terima kasih kami sampaikan kepada semua pihak yang telah memberikan dukungan dan kontribusi dalam penyusunan kajian ini. Menyadari bahwa kajian ini masih memiliki kekurangan, sehingga kami sangat terbuka terhadap saran dan masukan yang konstruktif untuk perbaikan di masa mendatang.

Samarinda, Juli 2024
Kepala Dinas Komunikasi dan Informatika
Provinsi Kalimantan Timur

H. Muhammad Faisal, S.Sos., M.Si.

DAFTAR ISI

TIM PENYUSUN	iii
KATA PENGANTAR	iv
DAFTAR ISI.....	v
DAFTAR TABEL.....	viii
DAFTAR GAMBAR	x
BAB 1. PENDAHULUAN.....	1
1.1. Latar Belakang	1
1.2. Tujuan	3
1.3. Manfaat	3
1.4. Ruang Lingkup	3
1.5. Sistematika Penulisan	4
BAB 2. DASAR TEORI.....	5
2.1. <i>Big Data</i>	5
2.1.1. Arsitektur Teknologi <i>Big Data</i>	10
2.1.2. Manfaat dan Aplikasi <i>Big Data</i>	12
2.1.3. Tantangan dan Peluang Pemanfaatan <i>Big Data</i>	13
2.1.4. Pemanfaatan <i>Big Data</i> Pada <i>National Statistical Office (NSO)</i>	17
2.1.5. Teknologi Hadoop	19
2.1.6. Contoh Penggunaan Teknologi <i>Big Data</i>	27

2.2. <i>Data Analytic</i>	28
2.2.1. Konsep Dasar <i>Data Analytics</i>	29
2.2.2. Proses <i>Data Analytics</i>	29
2.2.3. Teknik dan Alat dalam <i>Data Analytics</i>	30
2.2.4. Penerapan <i>Data Analytics</i> dalam Berbagai Sektor.....	32
2.2.5. Tantangan dalam <i>Data Analytics</i>	34
2.2.6. Masa Depan <i>Data Analytics</i>	35
BAB 3. URGENSI TRANSFORMASI DIGITAL	37
3.1. Peningkatan Efisiensi Operasional.....	37
3.2. Transparansi dan Akuntabilitas.....	38
3.3. Pelayanan Publik yang Lebih Baik.....	38
3.4. Pengambilan Keputusan yang Lebih Cepat dan Tepat.....	39
3.5. Peningkatan Kualitas Hidup Masyarakat.....	39
3.6. Tantangan dalam Transformasi Digital.....	40
3.7. Strategi untuk Transformasi Digital di Kalimantan Timur.....	41
BAB 4. IMPLEMENTASI <i>DATA SCIENCE</i> PADA SATU DATA KALIMANTAN TIMUR.....	47
4.1. Sektor Sosial Ekonomi.....	47
4.2. Sektor Kesehatan.....	56
4.2.1. Kesehatan Masyarakat Provinsi Kaltim.....	56
4.2.2. Wanita Usia 15-49 Tahun, Kawin dan KB Provinsi Kaltim.....	64
4.3. Sektor Peternakan.....	69

4.3.1. Populasi Ternak Sapi per Kabupaten/Kota di Kaltim.....	69
4.3.2. Populasi Ternak Kambing per Kabupaten/Kota di Kaltim	73
4.3.3. Produksi Telur Ayam Buras di Kaltim	78
4.4. Sektor Perdagangan	83
4.5. Sektor Pertambangan	89
4.6. Sektor Transportasi	95
4.6.1. Transportasi Darat di Kaltim.....	95
4.6.2. Kecelakaan Lalu Lintas di Kaltim	98
4.7. Sektor Perkebunan	102
4.7.1. Luas Areal Tanaman Perkebunan Rakyat Menurut Jenis Tanaman	102
4.7.2. Produksi Perkebunan Rakyat Menurut Jenis Tanaman.....	104
BAB 5. PENUTUP	111
5.1. Kesimpulan	111
5.2. Saran	112
DAFTAR PUSTAKA	115

DAFTAR TABEL

Tabel 1.	Karakteristik Perbedaan Data dari Statistik Resmi dan <i>Big Data</i>	16
Tabel 2.	Sosial Ekonomi Provinsi Kalimantan Timur Tahun 2021–2023	48
Tabel 3.	IPM Correlation Analysis Data.....	53
Tabel 4.	Data Kesehatan Masyarakat Kalimantan Timur Tahun 2016-2020.....	56
Tabel 5.	Persentase Wanita Berumur 15-49 Tahun Status Kawin dan Menggunakan KB per Kabupaten/Kota di Provinsi Kalimantan Timur Tahun 2019-2023.	64
Tabel 6.	Proyeksi Penggunaan KB Tahun 2024 dan 2025.....	67
Tabel 7.	Populasi Ternak Sapi per Kabupaten/Kota Tahun 2017–2022	69
Tabel 8.	Populasi Ternak Kambing per Kabupaten/Kota Tahun 2017–2022	73
Tabel 9.	Analisis Tren Tiap Wilayah Perkembangan Data Ternak Kambing di Kabupaten/Kota di Kalimantan Timur Tahun 2017-2022.	75
Tabel 10.	Komposisi Koefisien dan Intercept tiap Wilayah di Kaltim	77
Tabel 11.	Produksi Telur Ayam Buras di Provinsi Kalimantan Timur Tiap Wilayah di Kaltim (ton) Tahun 2017-2022	78
Tabel 12.	Angka Perdagangan Kaltim Tahun 2021-2023.....	83
Tabel 13.	Aktivitas Pertambangan di Provinsi Kalimantan Timur Tahun 2021-2023.....	89
Tabel 14.	Data Transportasi Darat (Bus dan Terminal Bus) di Provinsi Kalimantan Timur Tahun 2016-2020	95

Tabel 15. Data Kecelakaan Lalu Lintas per Wilayah di Kaltim Tahun 2016-2020	98
Tabel 16. Luas Areal Tanaman Perkebunan Rakyat Menurut Jenis Tanaman di Provinsi Kaltim (ha) Tahun 2019–2023.....	102
Tabel 17. Produksi Perkebunan Rakyat Menurut Jenis Tanaman di Provinsi Kaltim (ton) Tahun 2019–2023	104
Tabel 18. Prediksi Produksi Tanaman 2024.....	109

DAFTAR GAMBAR

Gambar 1.	<i>Hirarki Wisdom</i> (Rowley, 2007)	5
Gambar 2.	Karakteristik <i>Big Data</i>	6
Gambar 3.	Karakteristik Detail <i>Big Data</i>	7
Gambar 4.	Fenomena Data Deluge	8
Gambar 5.	Fenomena <i>Data Deluge</i>	8
Gambar 6.	<i>Platform</i> Pemrosesan <i>Big Data</i>	9
Gambar 7.	Arsitektur Teknologi <i>Big Data</i>	10
Gambar 8.	Karakteristik 6V <i>Big Data</i> yang Relevan dengan Data Statistik Resmi	18
Gambar 9.	Ekosistem Teknologi Hadoop	19
Gambar 10.	Ekosistem Hadoop.....	20
Gambar 11.	Ekosistem Hadoop bersama Spark dan Storm.....	20
Gambar 12.	Arsitektur HDFS.....	21
Gambar 13.	Arsitektur Teknologi <i>Big Data</i> Untuk Kasus Anti Hoax.	28
Gambar 14.	Grafik Trend Penambahan Penduduk dan Laju Pertumbuhan Penduduk Tahun 2021-2023.....	48
Gambar 15.	Grafik Tren Angka Harapan Hidup dan Angka Melek Huruf Usia 15+ Tahun 2021–2023.....	49

Gambar 16.	Grafik Tren Tingkat Partisipasi Angkatan Kerja dan Tingkat Pengangguran Terbuka Tahun 2021-2023.....	49
Gambar 17.	Grafik Tren Tingkat Penduduk Miskin dan Persentase Penduduk Miskin Tahun 2021-2023	49
Gambar 18.	Grafik Tren Indeks Pembangunan Manusia dan PDRB Harga Berlaku Tahun 2021-2023.....	50
Gambar 19.	Grafik Tren Indeks Pembangunan Manusia dan PDRB Harga Berlaku Tahun 2021-2023.....	50
Gambar 20.	Grafik Tren Inflasi (y-o-y) (%) Tahun 2021-2023	50
Gambar 21.	Grafik Analisis Korelasi Laju Pertumbuhan Ekonomi vs. Penduduk Miskin.....	52
Gambar 22.	Grafik Analisis Korelasi Laju Pertumbuhan Ekonomi vs. Persentase Penduduk Miskin	52
Gambar 23.	Grafik Analisis Korelasi PDRB Perkapita vs. Penduduk Miskin.....	53
Gambar 24.	Grafik Analisis Korelasi PDRB Perkapita vs. Persentase Penduduk Miskin.....	53
Gambar 25.	Grafik Tren Kesehatan Masyarakat Kaltim Periode 2016-2020 (3 variabel).....	61
Gambar 26.	Grafik Tren Kesehatan Masyarakat Kaltim Periode 2016-2020	62
Gambar 27.	Grafik Persentase Wanita Usia 15-49 Tahun Status Kawin dan KB menurut Kabupaten/Kota di Kaltim Tahun 2019-2023	65
Gambar 28.	Grafik Proyeksi Penggunaan KB Mendatang Tahun 2024 dan 2025.....	68

Gambar 29.	Grafik Populasi Ternak Sapi per Kabupaten/Kota di Provinsi Kalimantan Timur Tahun 2017-2022.....	70
Gambar 30.	Grafik Persentase Perubahan Tahunan Populasi Ternak Sapi per Kabupaten/Kota di Kalimantan Timur Tahun 2017-2022.....	71
Gambar 31.	Grafik Perubahan Tahunan Ternak Kambing Menurut Wilayah di Kaltim.....	74
Gambar 32.	Grafik Analisis Regresi Linier Setiap Wilayah Menurut Data Tahun 2017-2022.....	77
Gambar 33.	Grafik Garis Produksi Telur Ayam Buras Tiap Wilayah di Provinsi Kalimantan Timur (ton) Tahun 2017-2022.....	80
Gambar 34.	Grafik Batang Total Produksi Telur Ayam Buras di Kaltim Tahun 2017-2022.....	81
Gambar 35.	Grafik Perdagangan di Provinsi Kalimantan Timur Tahun 2021-2023	84
Gambar 36.	Grafik Proyeksi Nilai Ekspor-Impor Provinsi Kaltim Tahun 2021-2027	88
Gambar 37.	Grafik Data Aktivitas Pertambangan Provinsi Kalimantan Timur Tahun 2021- 2023.....	93
Gambar 38.	Grafik Tren Jumlah Terminal dan Jumlah Penumpang per Tahun di Provinsi Kalim Tahun 2016-2020.	96
Gambar 39.	Grafik Tren Jumlah Penumpang Tiba dan Jumlah Penumpang Berangkat per Tahun di Provinsi Kaltim Tahun 2016-2020.	97
Gambar 40.	Grafik Tren Jumlah Kecelakaan Lalu Lintas per Wilayah di Kaltim Tahun 2016-2020.....	100

Gambar 41.	Grafik Tren Rata-Rata Tahunan Kecelakaan Lalu Lintas di Provinsi Kalimantan Timur Tahun 2016-2020.	101
Gambar 42.	Grafik Luas Areal Tanaman Perkebunan Rakyat Menurut Jenis Tanaman di Provinsi Kaltim (ha) Tahun 2019–2023.	103
Gambar 43.	Grafik Produksi Karet, Kelapa, dan Kelapa Sawit (2019-2023)	106
Gambar 44.	Grafik Produksi Kopi, Kakao, Pala, dan Lada (2019-2023).....	107
Gambar 45.	Grafik Tren Produksi Tiap Tanaman dari Tahun 2019 hingga Tahun 2023.	107




Menguasai Big Data berarti menguasai masa depan

Big Data adalah kunci untuk membuka pintu inovasi dalam transformasi digital

Data Analytics bukan hanya tentang angka, tetapi tentang menggali wawasan untuk masa depan yang lebih baik



Transformasi digital adalah perjalanan dari informasi menuju pengetahuan, dan dari pengetahuan menuju kebijaksanaan



berperan dalam
berperan dalam

BAB 1

PENDAHULUAN

1.1. Latar Belakang

Provinsi Kalimantan Timur atau disingkat Kaltim terletak disisi timur pulau Kalimantan dan berbatasan langsung dengan negara Malaysia. Kaltim memiliki wilayah seluas 127.346,92 km², dan menjadi daerah terluas keempat di Indonesia. Kalimantan Timur (Kaltim) secara resmi berdiri sebagai provinsi pada 1 Januari 1957. Secara wilayah administratif provinsi kaltim terdiri dari 7 kabupaten, 3 kota, 103 kecamatan, dan 1.083 desa/kelurahan. Batas wilayah Kaltim sebelah utara berbatasan dengan Kalimantan Utara, sebelah timur berbatasan dengan sebagian (12 Mil) Selat Makasar dan Laut Sulawesi, sebelah selatan berbatasan dengan Provinsi Kalimantan Selatan, sebelah barat berbatasan dengan Provinsi Kalimantan Tengah dan Provinsi Kalimantan Barat serta Negara Bagian Serawak Malaysia Timur. Kalimantan Timur memiliki luas wilayah daratan 127.346,92 Km² dan luas pengelolaan laut 25.656 Km² terletak antara 113°44' Bujur Timur dan 119°00' Bujur Timur serta diantara 2°33' Lintang Utara dan 2°25' Lintang Selatan.

Kaltim sebagai provinsi yang kaya akan sumber daya alam dan memiliki peran penting dalam perekonomian nasional, menghadapi tantangan besar dalam pengelolaan data. Dengan adanya program Satu Data Indonesia (SDI), pemerintah berusaha untuk mengintegrasikan data dari berbagai instansi dan menyinkronkan informasi guna mendukung perencanaan dan pengambilan kebijakan yang lebih baik. Data telah menjadi salah satu aset paling berharga bagi organisasi dan pemerintah di seluruh dunia. Perkembangan teknologi informasi telah mengakibatkan peningkatan yang signifikan dalam *volume*, kecepatan, dan variasi data yang dihasilkan setiap hari, yang dikenal sebagai fenomena **Big Data** yang mencakup berbagai jenis data-data, baik terstruktur maupun tidak terstruktur, yang membutuhkan metode pengelolaan dan analisis yang lebih canggih, yang disebut sebagai **Data Analytics**.

Di sinilah urgensi penerapan *data science* muncul sebagai salah satu kunci dalam memanfaatkan *big data* dan *data analytic* untuk mendukung transformasi digital. *Big data* mengacu pada himpunan data yang sangat besar dan kompleks yang sulit untuk diolah menggunakan alat pengelolaan data tradisional. Data ini berasal dari berbagai sumber, termasuk media sosial, sensor, transaksi bisnis, dan banyak lagi. Karakteristik utama dari *big data* adalah *volume* (jumlah data yang sangat besar), *velocity* (kecepatan data masuk yang tinggi), *variety* (beragam jenis data), dan *veracity* (kualitas data yang bervariasi). Dengan adanya *big data*, analisis data menjadi lebih kompleks namun memberikan peluang besar untuk mendapatkan wawasan yang lebih mendalam.

Data analytic, adalah proses menganalisis data untuk menemukan pola, hubungan, dan tren yang berguna. Melalui berbagai teknik analisis, seperti analisis deskriptif, diagnostik, prediktif, dan preskriptif, *data analytic* membantu organisasi untuk memahami data mereka dengan lebih baik dan membuat keputusan yang didasarkan pada bukti yang kuat. Dalam konteks pemerintahan, *data analytic* dapat digunakan untuk memantau kinerja, mengidentifikasi masalah, dan merumuskan kebijakan yang lebih efektif.

Urgensi *data science* dalam transformasi digital semakin meningkat, terutama dalam implementasi program Satu Data Indonesia (SDI) di Provinsi Kalimantan Timur. *Data science* menggabungkan keterampilan dalam statistik, matematika, dan ilmu komputer untuk mengolah dan menganalisis data dalam skala besar. Dengan menggunakan teknik-teknik canggih seperti *Machine learning* (ML) dan *Artificial Intelligence* (AI), *data science* memungkinkan analisis yang lebih mendalam dan prediktif. Ini sangat penting dalam mendukung keputusan strategis yang berbasis data, yang pada akhirnya meningkatkan efisiensi dan efektivitas layanan publik.

Kajian ini akan membahas lebih lanjut mengenai konsep dasar *big data* dan *data analytic*, peran *data science* dalam transformasi digital, serta studi kasus implementasi satu data di Provinsi Kalimantan Timur. Kajian ini bertujuan untuk memberikan pemahaman yang komprehensif tentang pentingnya *data science* dalam pengelolaan data pemerintahan dan bagaimana hal tersebut dapat mendukung kebijakan pembangunan yang lebih efektif dan efisien di Provinsi Kalimantan Timur.

Selain itu, kajian ini juga akan mengidentifikasi tantangan dan memberikan rekomendasi strategis untuk mengoptimalkan pemanfaatan *big data* dan *data analytic* dalam mendukung transformasi digital di Provinsi Kalimantan Timur.

1.2. Tujuan

Kajian ini bertujuan untuk:

1. Memahami konsep *big data*, *data analytic*, dan *data science*.
2. Mengidentifikasi tantangan dan peluang dalam pengembangan *data science* di Provinsi Kalimantan Timur.
3. Menggambarkan urgensi *data science* dalam transformasi digital pada implementasi satu data Provinsi Kalimantan Timur.
4. Merumuskan rekomendasi untuk pengembangan *data science* di Provinsi Kalimantan Timur.

1.3. Manfaat

Hasil kajian ini diharapkan dapat memberikan manfaat bagi:

1. Pemerintah Provinsi Kalimantan Timur dalam menyusun kebijakan dan strategi pengembangan *data science* di Kalimantan Timur.
2. Dapat meningkatkan kualitas pengelolaan dan pemanfaatan data pada instansi pemerintah, badan dan lembaga di Provinsi Kalimantan Timur.
3. Dapat mendalami, memahami peluang dan prospek karir di bidang *data science* oleh tenaga profesional dan masyarakat di Provinsi Kalimantan Timur.
4. Mendapatkan peluang sumber informasi yang lebih akurat dan tepat waktu dari Pemerintah untuk masyarakat di Provinsi Kalimantan Timur.

1.4. Ruang Lingkup

Kajian ini akan fokus pada beberapa aspek ruang lingkup sebagai berikut:

1. Definisi dan karakteristik *big data*, *data analytic*, dan *data science*.
2. Peran *data science* dalam transformasi digital satu data Provinsi Kalimantan Timur.

1.5. Sistematika Penulisan

Penulisan kajian ini berusaha memberikan gambaran mengenai perananan dan implementasi *big data* dan *data analytic* terhadap urgensi dalam penggunaan *data science* dalam transformasi digital satu data di provinsi Kalimantan Timur.

- Bab 1 Pendahuluan, pada bab ini menjelaskan latar belakang, tujuan, maksud dan ruang lingkup serta sistematika penulisan terkait kajian yang dilaksanakan.
- Bab 2 Dasar Teori, pada bab ini menjelaskan terkait teori dan pustaka yang mendukung dalam pelaksanaan kajian ini, dimana teori-teori tentang *big data*, *data analytic* dan *data science* yang merupakan konsep fundamental dalam penelitian ini serta dilengkapi tools dalam pengembangan *big data dan data analytic* terhadap data yang terdapat pada portal satu data Provinsi Kalimantan Timur.
- Bab 3 Urgensi Transformasi Digital, pada bab ini menjelaskan tentang perlunya peningkatan efisiensi operasional, transparansi dan akuntabilitas, pelayanan publik yang lebih baik, pengambilan keputusan yang lebih cepat dan tepat, peningkatan kualitas hidup masyarakat, tantangan dalam transformasi digital.
- Bab 4 Implementasi *Data Science* pada Satu Data Kaltim, pada bab ini menjelaskan terkait mengimplementasikan konsep dan formulasi serta scripting kode program python untuk pengolahan visualisasi dan komputasi data-data yang terdapat pada portal satu data Provinsi Kaltim. Berupa analisis data di sektor sosial ekonomi, sektor kesehatan, sektor peternakan, sektor perdagangan, sektor pertambangan, sektor transportasi, dan sektor Perkebunan.
- Bab 5 Penutup, pada bab ini menjelaskan terkait kesimpulan dari kajian dan memberikan saran atau rekomendasi dari hasil kajian tersebut.

Data adalah kekuatan baru yang menggerakkan dunia. Dalam era digital, siapa yang menguasai data, dia yang memimpin

Transformasi digital adalah perjalanan, bukan tujuan. Data Science adalah kompas yang menuntun kita menuju inovasi

Big Data memberi kita wawasan, Data Analytics mengarahkan langkah kita, dan Transformasi Digital membuka pintu menuju kemungkinan tanpa batas

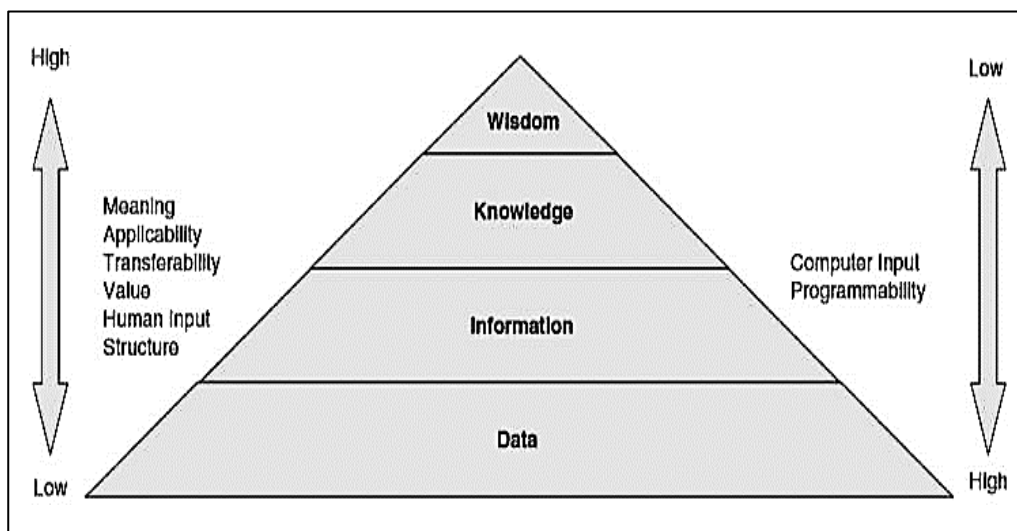


BAB 2

DASAR TEORI

2.1. *Big Data*

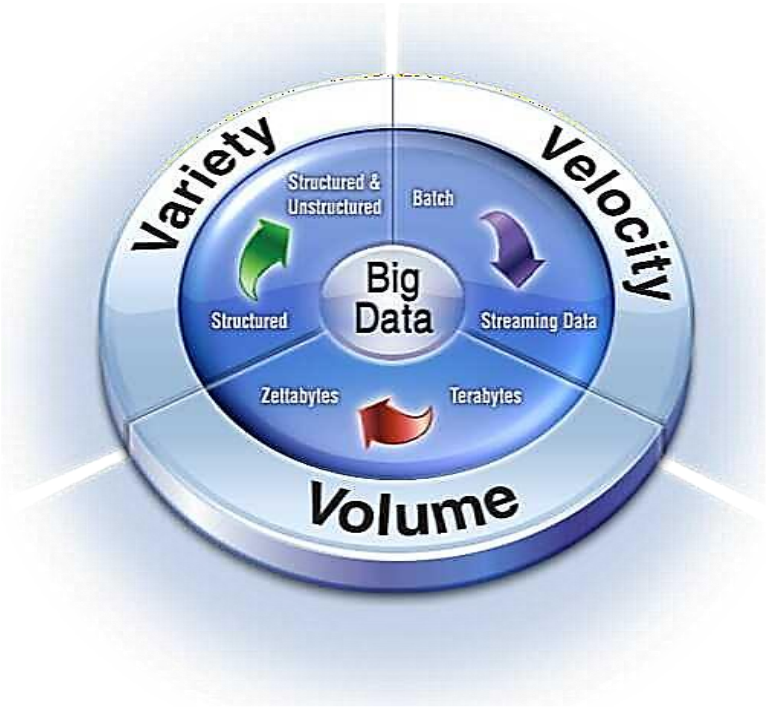
Dari berbagai pendapat yang dikumpulkan pada penelitian (De Mauro dkk., 2016) *big data* didefinisikan sebagai aset informasi yang dicirikan oleh karakteristik 3v, yaitu: *volume*, *velocity* dan *variety* yang tinggi, yang memerlukan metode dan teknologi tertentu untuk memprosesnya menjadi pengetahuan (*knowledge*) yang bernilai (*value*) dalam pengambilan keputusan. Aset informasi bermakna penting, karena data dianggap memiliki nilai yang tinggi bagi organisasi seperti aset lain (mesin, material, orang, modal, dan metode) dan dapat divalusi (dinilai dalam satuan uang). Konsep hirarki *Data–Information–Knowledge–Wisdom* (DIKW), (Rowley, 2007) atau sering disebut sebagai *wisdom hierarchy* dan dapat dilihat pada gambar 1 yang memberikan alasan yang masuk akal mengapa fenomena *big data* begitu berkembang.



Gambar 1. *Hirarki Wisdom* (Rowley, 2007)

Dengan besarnya potensi data yang ada saat ini dan di masa depan, maka besar juga potensi informasi yang tersedia untuk ditransformasi menjadi pengetahuan (*knowledge*) sehingga dapat mengoptimalkan pengambilan keputusan (*wisdom*).

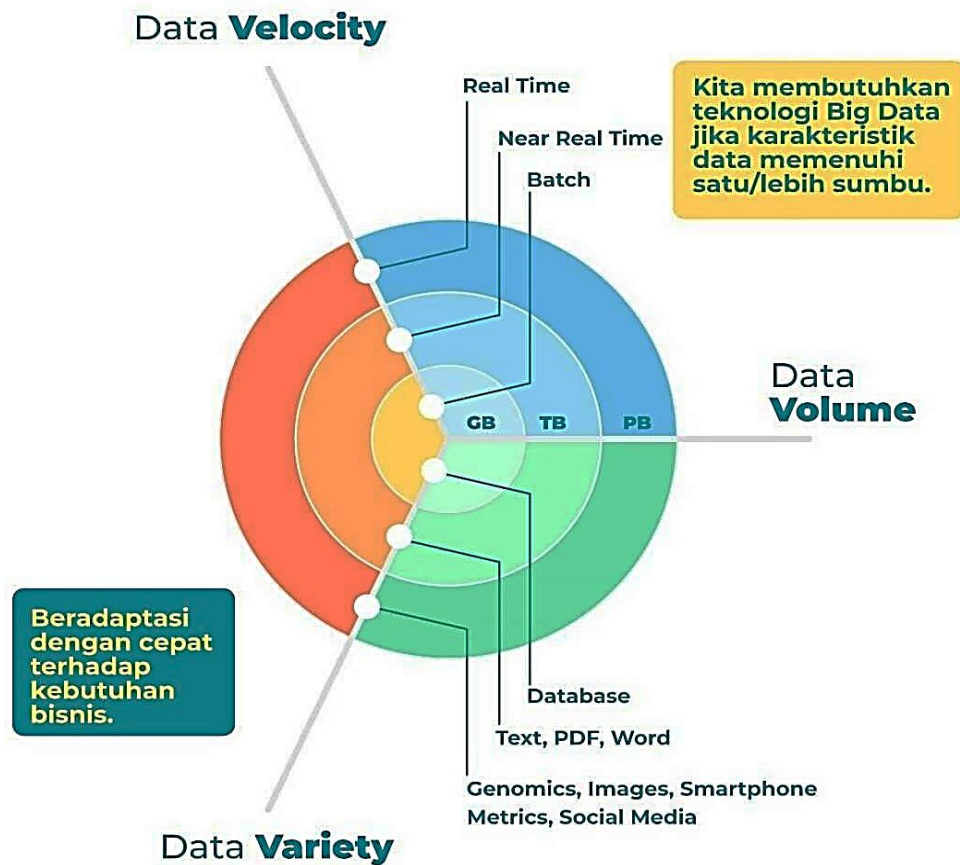
Dengan demikian, jika *big data* dapat ditangani dengan baik akan memberikan manfaat besar bagi organisasi, khususnya semakin bijaksana dalam mengambil keputusan yang didasarkan atas data (bersifat *data driven*), sehingga lincah dalam mengambil keputusan dalam perubahan kondisi lingkungan yang cepat berubah. *Big data* dicirikan dengan karakteristik *big data* menurut (De Mauro dkk., 2016) adalah 3v yaitu: *volume*, *velocity* dan *variety* yang tinggi. Secara umum batasan tinggi dalam konteks *big data* mengikuti hukum Moore (Moore, 2006). Namun demikian, saat ini karakteristik *big data* digambarkan seperti pada gambar 2.



Gambar 2. Karakteristik *Big Data*

Big data memiliki karakteristik *volume* yang tinggi, dari *terabytes* ke *zettabytes*. Hal ini berkonsekuensi pada kapasitas penyimpanan dan kapasitas pemrosesan data yang tidak dapat ditangani oleh metode dan teknologi informasi konvensional saat ini. Metode dan teknik penyimpanan yang diterapkan hingga saat ini mengarah pada pemrosesan secara paralel pada lingkungan sistem terdistribusi, baik dari sisi media penyimpanan maupun pemrosesannya. Karakteristik *big data* lebih detail dapat dilihat pada gambar 3.

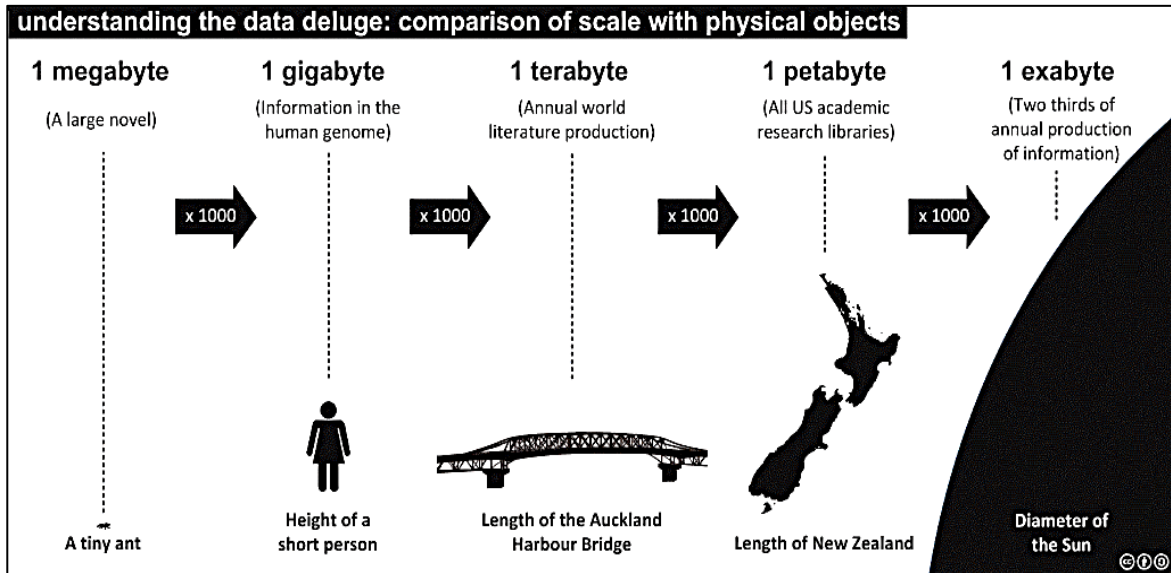
Karakteristik *velocity* pada *big data* mengubah sudut pandang pemrosesan data secara batch, menjadi pemrosesan data secara dinamis. Dengan demikian data tidak lagi dilihat secara statis, namun secara dinamis sebagai stream. Selain sebagai data stream, *big data* juga berkaitan dengan pergerakan data dalam jumlah besar (*high volume movement*) seperti data spasial, citra, dan lainnya.



Gambar 3. Karakteristik Detail *Big Data*

Big data bersumber dari berbagai event. Semua kegiatan kita yang menggunakan komputer, gadget, sensor dan peralatan lainnya menghasilkan *big data*. Selain sumber yang beraneka ragam, dari sisi struktur juga beraneka ragam, mulai dari yang terstruktur, seperti: data transaksi (pasar uang, e-commerce, dll), semi terstruktur, maupun yang tidak terstruktur, seperti: image, text opini pada media sosial maupun halaman web di internet. Untuk itu diperlukan metode dan teknologi untuk mengintegrasikan *big data* dari berbagai sumber dan dari format yang berbeda-beda tersebut.

Masalah utama *big data* dikenal dengan istilah fenomena *data deluge*, suatu fenomena dimana laju pertumbuhan data lebih tinggi dari pada laju kemampuan memproses dan menganalisis data suatu organisasi. Pada gambar 4 dapat dilihat besarnya *volume* data dibandingkan dengan objek fisik. Oleh karena itu dalam memproses dan menganalisis data, kita memerlukan teknologi yang tidak konvensional lagi. Kita memerlukan teknologi yang dapat mengimbangi laju pertumbuhan data yang meningkat seiring dengan waktu dan peningkatan penggunaan teknologi informasi dan komunikasi.



Gambar 4. Fenomena Data Deluge

Sumber: <https://ritholtz.com/2016/09/162347/>

Jika data diibaratkan seperti hujan lebat (gambar 5), maka kita bisa menangkap air dengan laju yang sesuai, kemudian mengumpulkannya untuk menyiram tanaman. Untuk menangkap semua data, tentu memerlukan banyak, yang diumpamakan dengan banyak payung atau payung yang sangat besar.



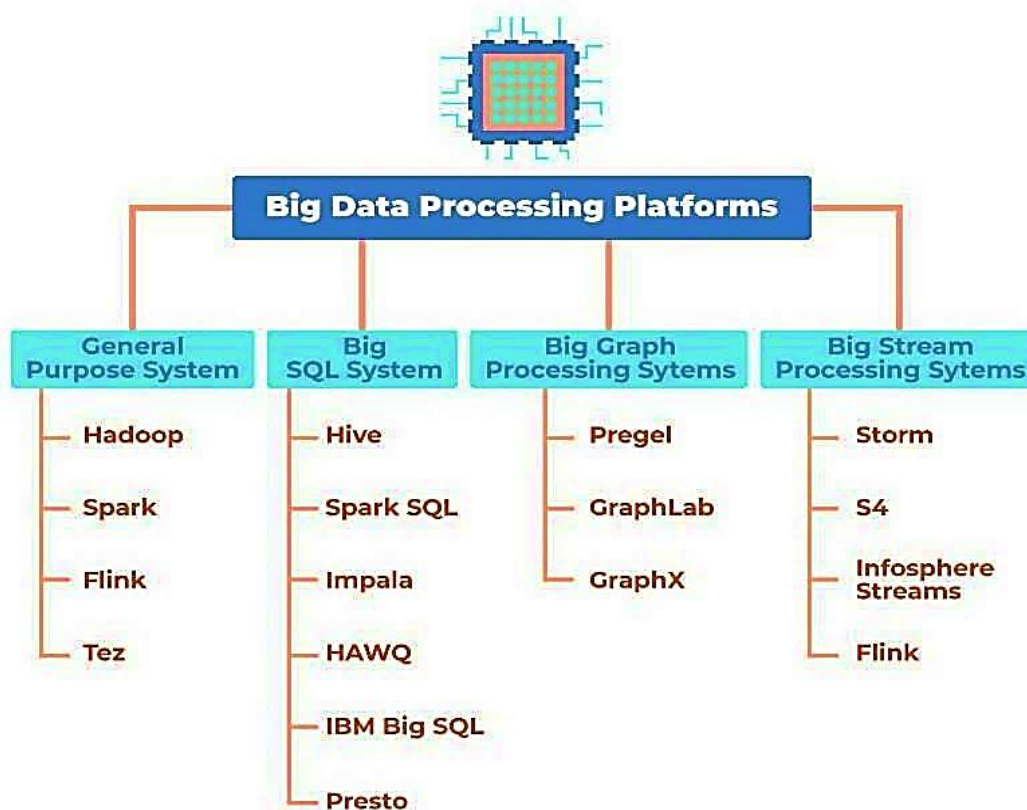
Gambar 5. Fenomena *Data Deluge*

Sumber: <https://www.economist.com/leaders/2010/02/25/the-data-deluge/>

Teknologi *big data* merupakan teknologi khusus yang diperuntukkan untuk menangani masalah *big data*. Untuk menangani masalah *volume*, teknologi *big data* menggunakan teknik penyimpanan dan pemrosesan data terdistribusi. Masalah *velocity* ditangani dengan menggunakan pemrosesan stream dan terdistribusi. Sedangkan masalah *variety* ditangani menggunakan teknik integrasi data dan penyimpanan data tidak terstruktur (*on write*). Penentuan struktur dilakukan pada saat proses pembacaan data tersebut (*on read*).

Pada gambar 6 dapat kita lihat berbagai *platform* teknologi pemrosesan *big data* yang telah dikumpulkan pada penelitian (Bajaber dkk., 2016). Berdasarkan *platform* pemrosesan *big data*, teknologi tersebut dapat dikelompokkan menjadi yaitu:

1. Umum (*General Purpose*), seperti Hadoop, Spark, Flink, Tez
2. Pemroses *query* (Big SQL), seperti: Hive, Spark SQL, Implala, HawQ, IBM Big SQL
3. Pemroses *big graph*, seperti: pregel, graphLab, GraphX, dan
4. Pemroses *big stream*, seperti: Storm, S4, Infosphere stream, flink, dan Spark Stream.



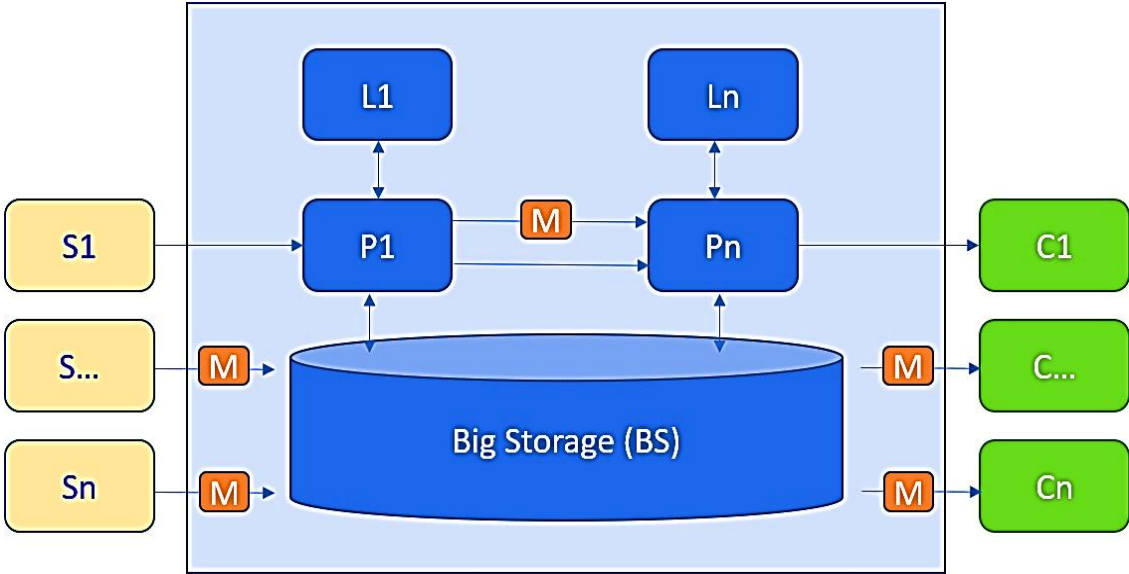
Gambar 6. Platform Pemrosesan Big Data

Teknologi *big data* yang populer digunakan saat ini adalah teknologi Hadoop yang dikembangkan pada awalnya oleh perusahaan teknologi Google (Ghemawat dkk., 2003), kemudian menjadi proyek Apache yang berdiri sendiri. Pada prinsip utama dari teknologi Hadoop adalah penyimpanan dan pemrosesan terdistribusi pada komputer-komputer komoditas yang terhubung dalam jaringan (sering disebut *cluster*).

Inti dari teknologi Hadoop yaitu *Hadoop Distributed File System* atau disingkat HDFS, bertugas untuk menangani penyimpanan data terdistribusi dan *Map Reduce* untuk pemrosesan data terdistribusi yang dilakukan pada komputer (*node of cluster*) tempat data disimpan. Untuk menyelesaikan berbagai persoalan komputasi, Hadoop didukung oleh berbagai teknologi yang secara keseluruhan sering disebut sebagai *Hadoop Ecosystem*.

2.1.1. Arsitektur Teknologi Big Data

Untuk mengatasi masalah *big data*, maka teknologi *big data* bukanlah suatu teknologi tunggal. Oleh karena itu, kita perlu melihatnya secara menyeluruh dan secara detail melalui arsitektur teknologi *big data* seperti pada gambar 7. Arsitektur teknologi *big data* menggambarkan komponen-komponen yang diperlukan, dan interaksi antar komponen sehingga sistem *big data* secara keseluruhan dapat berjalan.



Gambar 7. Arsitektur Teknologi Big Data

Pada gambar 7 dilihatkan S1 sampai Sn adalah sumber (*source*) dari *big data*, seperti sensor, aplikasi media sosial (Twitter Stream, Facebook, Instragram, dll), web berita (news), satelit, basis data, dan sumber lainnya.

Dimana sangat memerlukan teknologi untuk mengumpulkan data dari sumber *big data*. Pengumpulan data dapat dilakukan dengan 2 cara, yaitu:

1. Menggunakan aplikasi yang dibuat khusus untuk memproses *big data* (process – P1 sampai Pn), misalnya: untuk keperluan mengambil data pada aplikasi Twitter, dapat dibuat aplikasi penangkap data-data konten Twitter (*Twitter Stream Capture*), atau
2. Dengan menggunakan aplikasi yang sudah jadi sebagai mediator (*Middleware - M*).

Pemroses (P) terdiri atas 2 jenis atas dasar tingkat interaktifitas pemrosesan dalam menghasilkan output, yaitu: batch processing dan online (real time) processing. Sedangkan jika dilihat dari data yang diproses dapat dikelompokkan menjadi 2 juga, yaitu *data in rest* (pemrosesan terhadap data yang sudah tersimpan di media penyimpanan) dan *data in motion (stream) processing* (pemrosesan terhadap data yang belum tersimpan pada media penyimpanan). Pemrosesan terhadap *data stream* berasosiasi dengan *online/real time processing*, sedangkan pemrosesan terhadap *data in rest* berasosiasi dengan pemrosesan *batch*. Untuk mengembangkan aplikasi pemroses *big data* (P) diperlukan pustaka teknologi (*library – L1 sampai Ln*) dari *platform* teknologi pemrosesan *big data*.

Komponen pemroses (P) atau mediator (M) berkomunikasi dengan media penyimpanan (*Big Storage - BS*). Media penyimpanan digunakan untuk menyimpan data mentah (*Raw Data*) maupun data hasil pengolahan/pemrosesan (*result*). Hasil pengolahan/pemrosesan tersebut perlu dikirim ke konsumen (*Consument – C1 sampai Cn*) baik melalui pemroses secara langsung (P) maupun melalui mediator (M) secara berkala sebagai proses propagasi atau sinkronisasi. Contoh dari aplikasi C adalah web site tempat publikasi, seperti pada web anti hoax yang juga dapat digunakan sebagai mediator komunikasi antar aplikasi pemroses (P).

Pada proses pengambilan data dari S di atas, karena sumbernya banyak dengan format yang beranekaragam, maka teknologi yang diperlukan tidak hanya membaca/ mengambil (*capture*), namun juga mengintegrasikan (*integrator*) baik melalui perubahan format, jenis, pewaktuan dan pra olah lainnya, sehingga dapat dihasilkan data yang siap untuk diproses lebih lanjut agar menjadi pengetahuan (*knowledge*). Sebagai contoh teknologi yang berperan pada komponen arsitektur pada gambar 7 lebih lanjut dijelaskan pada bagian Ekosistem Hadoop.

2.1.2. Manfaat dan Aplikasi *Big Data*

Big Data telah membuka peluang baru dalam berbagai bidang dan industri. Berikut adalah beberapa contoh bagaimana *big data* digunakan untuk memberikan nilai tambah:

1. **Bisnis dan Pemasaran**

Dalam bisnis, *big data* digunakan untuk memahami perilaku pelanggan, mengoptimalkan operasi, dan menciptakan produk serta layanan baru. Analisis data pelanggan dapat membantu perusahaan menyesuaikan penawaran mereka dengan preferensi individu, meningkatkan retensi pelanggan, dan meningkatkan keuntungan. Selain itu, analisis rantai pasokan dapat membantu mengurangi biaya dan meningkatkan efisiensi operasional. Misalnya, perusahaan *e-commerce* seperti Amazon menggunakan analisis *big data* untuk merekomendasikan produk kepada pelanggan berdasarkan riwayat pembelian dan pencarian mereka. Dengan memahami preferensi pelanggan, Amazon dapat meningkatkan konversi penjualan dan kepuasan pelanggan. Selain itu, analisis *big data* juga digunakan untuk mengoptimalkan inventaris dan logistik, memastikan bahwa produk tersedia tepat waktu dan mengurangi biaya penyimpanan.

2. **Kesehatan**

Di bidang kesehatan, *big data* digunakan untuk penemuan obat, personalisasi perawatan, dan pemantauan kesehatan masyarakat. Data medis yang dihasilkan dari catatan kesehatan elektronik, perangkat *wearable*, dan sensor dapat dianalisis untuk mengidentifikasi pola kesehatan, memprediksi penyakit, dan memberikan perawatan yang lebih tepat waktu dan efektif. Contohnya, analisis *big data* dapat digunakan untuk memantau tanda-tanda vital pasien secara *real-time* dan memberikan peringatan dini jika ada indikasi masalah kesehatan. Selain itu, analisis genomik dapat membantu dalam pengembangan terapi yang disesuaikan dengan profil genetik individu, memungkinkan perawatan yang lebih efektif dan efisien.

3. **Pemerintahan**

Dalam pemerintahan, *big data* dapat membantu dalam pengambilan keputusan berbasis data, meningkatkan efisiensi layanan publik, dan memprediksi serta merespons bencana. Misalnya, data dari sensor lingkungan dapat digunakan untuk memantau kualitas udara

dan air, sementara data lalu lintas dapat membantu mengelola infrastruktur transportasi dengan lebih baik. Contohnya, kota cerdas (*smart cities*) menggunakan analisis *big data* untuk mengelola lalu lintas, energi, dan layanan publik lainnya. Data dari berbagai sensor dan perangkat IoT dapat dianalisis untuk mengidentifikasi pola penggunaan dan mengoptimalkan sumber daya, sehingga meningkatkan efisiensi dan mengurangi biaya operasional.

4. Ilmu Pengetahuan

Dalam ilmu pengetahuan, *big data* memungkinkan para peneliti untuk menganalisis data dalam skala besar yang sebelumnya tidak mungkin dilakukan. Misalnya, dalam astronomi, data dari teleskop besar dapat dianalisis untuk menemukan pola dan anomali yang dapat mengarah pada penemuan baru. Di bidang genomik, analisis data genomik dapat membantu memahami penyakit genetik dan mengembangkan terapi baru. Sebagai contoh, proyek *Large Hadron Collider* (LHC) di CERN menghasilkan data dalam jumlah besar dari tabrakan partikel yang terjadi setiap detik. Data ini dianalisis untuk menemukan partikel subatomik baru dan memahami sifat dasar alam semesta. Analisis *big data* juga digunakan dalam penelitian iklim untuk memodelkan perubahan iklim dan memprediksi dampak lingkungan di masa depan.

2.1.3. Tantangan dan Peluang Pemanfaatan *Big Data*

Meskipun sudah banyak kajian pemanfaatan *big data* di beberapa kantor/instansi statistik, masih banyak argumentasi terkait pemanfaatan *big data* untuk statistik resmi. Berikut merupakan beberapa tantangan dalam pemanfaatan *big data* untuk statistik resmi:

1. Validitas, tidak mudah untuk melakukan uji validasi pada *big data* karena sifatnya yang besar dan heterogen dan uji validitas dengan pendekatan statistik tidak memungkinkan.
2. Kualitas, data yang bersumber dari data tidak menggunakan metode statistik yang standar, sehingga belum tentu bisa mewakili populasi.
3. Akses terhadap sumberdata, permasalahan yang sering terjadi terkait kerahasiaan data. Oleh karena itu, perlu dibangun mekanisme yang dapat menjembatani kepentingan

pemilik data agar sharing data tanpa menimbulkan kekhawatiran dan aspek kerahasiaannya.

4. Sumber daya manusia dengan berbagai disiplin ilmu yang berbeda.
5. Metode dan pemrosesan data yang rumit.

Ada sejumlah tantangan yang harus diatasi untuk memanfaatkan potensi *big data* secara penuhnya. Beberapa tantangan utama meliputi:

1. Skalabilitas

Mengelola dan memproses *volume* data yang sangat besar memerlukan infrastruktur dan teknologi yang skalabel. *Distributed computing* dan *storage* adalah kunci untuk mengatasi tantangan ini, tetapi juga memerlukan investasi yang signifikan dalam perangkat keras dan perangkat lunak. Sebagai contoh, perusahaan teknologi seperti Google dan Facebook telah mengembangkan infrastruktur komputasi terdistribusi yang mampu menangani *volume* data yang sangat besar. Teknologi seperti Map Reduce dan Hadoop memungkinkan pemrosesan data dalam skala besar secara efisien, memungkinkan analisis data yang cepat dan tepat waktu.

2. Keamanan dan Privasi

Data besar sering kali mengandung informasi sensitif, seperti data pribadi atau data bisnis yang rahasia. Menjaga keamanan dan privasi data ini adalah tantangan besar, terutama dalam lingkungan yang terdistribusi. Regulasi seperti GDPR di Eropa menekankan pentingnya perlindungan data dan memberikan sanksi bagi pelanggaran privasi. Misalnya, perusahaan harus memastikan bahwa data pelanggan yang mereka kumpulkan dienkripsi dan disimpan dengan aman untuk mencegah akses yang tidak sah. Selain itu, perusahaan harus mematuhi regulasi privasi data yang berlaku, seperti GDPR di Eropa dan CCPA di California, yang mengatur bagaimana data pribadi dikumpulkan, disimpan, dan digunakan.

3. Kualitas Data

Memastikan kualitas data adalah tantangan utama dalam *big data*. Data yang tidak akurat, tidak lengkap, atau bias dapat menghasilkan wawasan yang menyesatkan. Oleh

karena itu, penting untuk memiliki proses pembersihan dan validasi data yang efektif. Contohnya, dalam industri keuangan, data transaksi yang tidak akurat dapat mengakibatkan kesalahan dalam laporan keuangan dan pengambilan keputusan investasi. Oleh karena itu, perusahaan perlu menerapkan proses pembersihan dan validasi data yang ketat untuk memastikan integritas data. Selain itu, dalam bidang kesehatan, data medis yang tidak akurat dapat mengakibatkan diagnosis yang salah dan perawatan yang tidak tepat, yang dapat membahayakan pasien.

4. Sumber Daya Manusia

Ada kebutuhan besar untuk tenaga kerja yang terampil dalam bidang *data science* dan analitik data. Kurangnya profesional yang memiliki keterampilan ini dapat menghambat kemampuan organisasi untuk memanfaatkan *big data* secara efektif. Pendidikan dan pelatihan di bidang *data science* adalah kunci untuk mengatasi tantangan ini. Sebagai contoh, banyak universitas dan institusi pendidikan telah mengembangkan program studi dan sertifikasi di bidang *data science* untuk memenuhi kebutuhan industri. Selain itu, perusahaan juga dapat menyediakan pelatihan internal dan program pengembangan keterampilan untuk karyawan mereka, sehingga mereka dapat mengembangkan keterampilan analitik data yang diperlukan.

Selanjutnya beberapa peluang dalam pemanfaatan *big data* yang sudah dilakukan oleh *National Statistical Office* (NSO), dengan sumber yang memenuhi kriteria yang sudah ditetapkan yaitu:

1. Kerangka sampel atau pembuatan register, mengidentifikasi unit populasi survei dan atau memberikan informasi tambahan seperti variabel stratifikasi.
2. Pengganti data lengkap, menggantikan survei.
3. Pengganti data parsial untuk subkelompok populasi, mengurangi ukuran sampel.
4. Pengganti data parsial untuk beberapa item data yang diperlukan, mengurangi panjang kuesioner survei, atau memperkaya dataset.
5. Imputasi untuk data yang hilang, mengganti unit yang sama atau serupa.
6. Mengedit, membantu deteksi dan penanganan anomali dalam data survei.

7. Menghubungkan ke data lain, membuat kumpulan data yang lebih kaya dan / atau perspektif longitudinal / series.
8. Konfrontasi data, memastikan validitas dan konsistensi data survei, dan
9. Menghasilkan wawasan analitis baru, meningkatkan pengukuran dan penggambaran fenomena ekonomi, sosial dan lingkungan.

Beberapa penelitian membahas tentang peluang *big data* dalam *official statistics*, apakah bisa menggantikan atau hanya sebagai pelengkap saja. Florescu dkk. dalam Kitchin (2015) menjelaskan bahwa sumber data besar dapat digunakan dalam sistem statistik dalam lima cara:

1. Untuk sepenuhnya menggantikan sumber statistik yang ada seperti (*existing statistical outputs*).
2. Untuk mengganti sebagian sumber statistik yang ada seperti survei (*existing statistical outputs*).
3. Untuk memberikan informasi statistik pelengkap tetapi dari yang lain perspektif survei (*additional statistical outputs*).
4. Untuk meningkatkan perkiraan dari sumber statistic/survei (*improved statistical outputs*).
5. Untuk memberikan informasi statistik yang benar-benar baru dalam statistik tertentu (*new alternative statistical outputs*).

Dilihat dari karakteristik perbedaan data dari statistik resmi dan *big data* pada tabel berikut dibawah ini:

Tabel 1. Karakteristik Perbedaan Data dari Statistik Resmi dan *Big Data*

Statistik Resmi (<i>Official Statistics</i>)	<i>Big Data</i>
Merupakan hasil kegiatan yang terstruktur dan terencana	Sebagian besar merupakan data tidak terstruktur (data hasil kegiatan digital, media sosial, sensors)

Statistik Resmi (<i>Official Statistics</i>)	<i>Big Data</i>
Jelas konsep dan metodologi yang digunakan	Konsep dan metodologi belum jelas dan seragam
Diatur (<i>regulated</i>)	Tidak diatur (<i>Unregulated</i>)
Level makro	Level mikro
Biaya tinggi	Biaya bisa lebih murah
Memotret satu waktu saja	Dihasilkan secara <i>real-time</i>

2.1.4. Pemanfaatan *Big Data* Pada *National Statistical Office* (NSO)

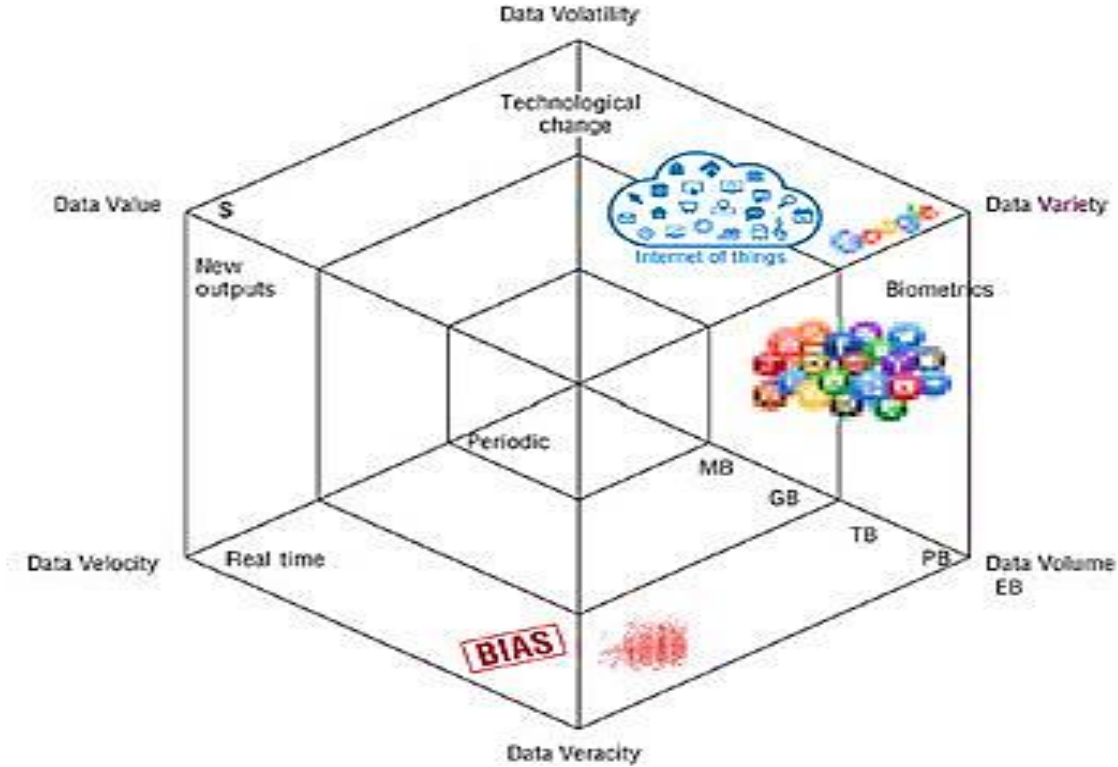
Menuntut instansi penyedia data (*National Statistical Office/NSO*) dari berbagai negara untuk melakukan inovasi-inovasi agar mampu memenuhi tuntutan pengguna data. Oleh karena itu, perlu dilakukan upaya-upaya mencari sumber data baru yang lebih cepat selain metode pengumpulan data yang selama ini dilakukan yaitu sensus, survei dan kompilasi data administrasi. Beberapa kantor atau instansi statistik di berbagai negara telah melakukan beberapa kajian tentang peluang pemanfaatan *big data* sebagai sumber data baru dalam penghitungan statistik.

Pada tahun 2016, *Office for National Statistics (ONS) United Kingdom* telah melakukan studi kelayakan dan peluang pemanfaatan *big data* dalam bidang ekonomi. Beberapa kajian yang dilakukan terkait *big data* antara lain pemanfaatan data harga yang di-*scraped* dari website, *machine learning* untuk mencocokkan alamat, membandingkan kepadatan menara seluler dengan perkiraan populasi, data telepon selular untuk memperkirakan arus perjalanan dan pemrosesan bahasa alami (*natural language*) untuk mengkodekan respon survei.

Australian Burrau Statistics (ABS) sejak tahun 2014 sudah menginisiasi pemanfaatan *big data* yang mencakup strategi pemanfaatan *big data*, proyek-proyek pengembangan *big data* seperti pemahaman konsep *big data*, kajian metodologi *big data* dalam menghasilkan statistik, visualisasi dan analisis data besar serta melakukan penelitian bersama dengan beberapa universitas di Australia. *United Nation* juga telah membentuk *UN Global Working Group (GWG) on Big Data for Official Statistics* yang khusus melakukan kajian dan secara rutin melakukan *sharing session* pengalaman pemanfaatan *big data* dari berbagai negara yang tergabung dalam GWG.

Keberadaan *big data* dengan demikian merupakan sebuah keniscayaan yang menjadi tantangan sekaligus peluang terutama bagi Kantor Statistik Resmi (*National Statistics Offices* – NSO) yang tugas utamanya adalah menghasilkan statistik resmi sebagai sumber pengambilan keputusan pemerintah dan juga sebagai informasi kunci dalam masyarakat yang terus berubah (Virgillito dan Scannapieco, 2013; Struijs, Braaksma, dan Daas, 2014). Untuk mengatasi tantangan ini, NSO harus dapat memanfaatkan dan mengelola *big data* dengan menerapkan prosedur penjaminan kualitas (Abdullah, et.al, 2015) hingga menjadi sebuah informasi yang bernilai.

Big Data dengan demikian dapat menjadi sumber data selain sensus dan survei yang selama ini telah dilakukan oleh sebuah NSO. Dengan pengelolaan yang baik, *big data* dapat diusulkan sebagai pengganti statistik resmi atau statistik resmi yang benar-benar baru (Eurostat, 2014) atau sebagai pelengkap statistik resmi (Letouzé dan Jütting, 2015). Beberapa NSO di dunia telah melakukan kajian dan meng- eksplorasi berbagai sumber *big data* untuk melengkapi produk-produk statistik yang dihasilkan secara konvensional melalui survei dan sensus. Berbagai proyek penelitian yang terkait *big data* yang dilakukan oleh beberapa NSO dapat ditemui dalam situs yang dikembangkan oleh PBB dengan alamat <https://unstats.un.org/bigdata/inventory/>.

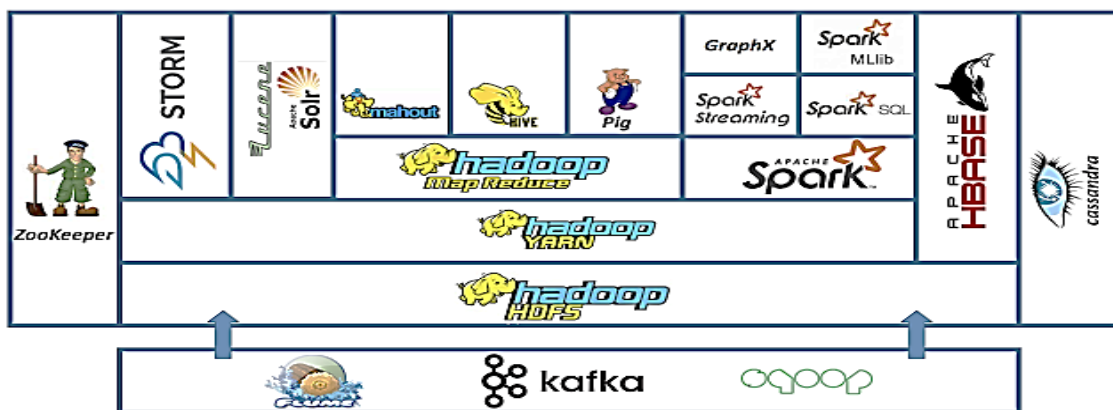


Gambar 8. Karakteristik 6V *Big Data* yang Relevan dengan Data Statistik Resmi

Di Indonesia beberapa sumber *big data* telah terbukti memiliki potensi implementasi untuk mendukung produksi statistik resmi, antara lain memprediksi pola perjalanan antar kota menggunakan data twitter, pemanfaatan data mobile position untuk Statistik Pariwisata, melakukan estimasi harga produk makanan menggunakan *data crowdsourcing* (Pramana, et. Al, 2017).

2.1.5. Teknologi Hadoop

Teknologi *big data* yang sangat populer saat ini adalah teknologi Hadoop. Kepopuleran Hadoop didukung oleh ketersediaannya sebagai proyek *Free Open Source Software* (FOSS) yang dikelola oleh Apache, sehingga tersedia bagi semua kalangan dan mendapat masukan dari para developer di seluruh dunia. Kumpulan teknologi yang bekerja di atas *platform* teknologi Hadoop sering disebut sebagai Ekosistem Hadoop (*Hadoop Ecosystem*). Teknologi ini saling bekerjasama untuk memberikan solusi pemrosesan *big data*. Ekosistem Hadoop secara skematis dapat dilihat pada gambar 9 dimana dapat dilihat bahwa komponen dasar dari ekosistem Hadoop adalah teknologi Hadoop.

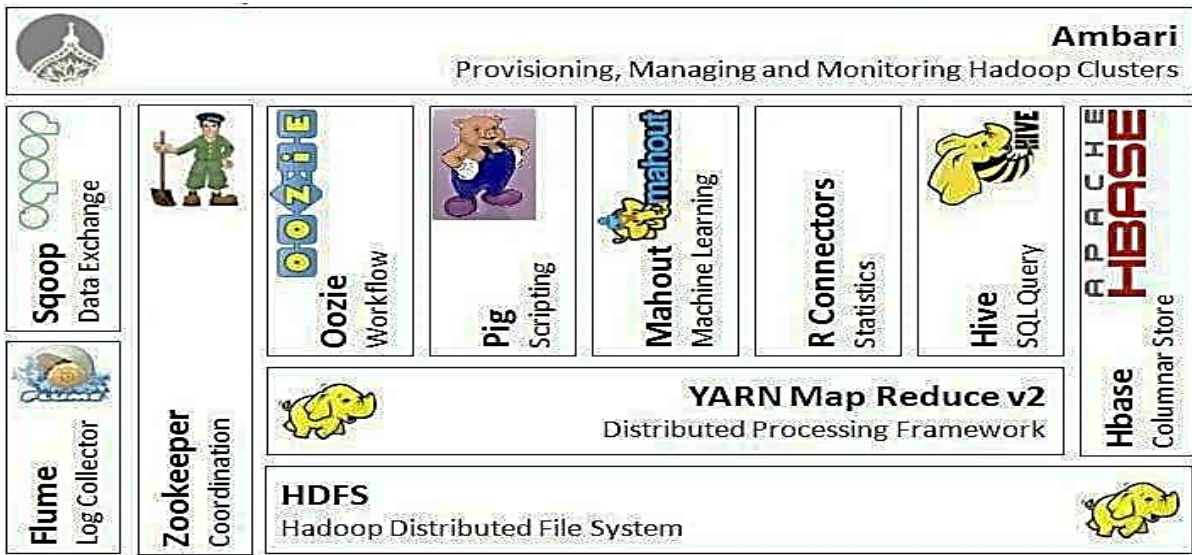


Gambar 9. Ekosistem Teknologi Hadoop

Sumber: <http://blog.newtechways.com/2017/10/apache-hadoop-ecosystem.html>

Hadoop menyediakan teknologi penyimpanan dan pemrosesan terdistribusi (*parallel*) pada komputer-komputer komoditas yang terhubung dalam jaringan (sering disebut *cluster*). Inti (*core*) dari Hadoop adalah *Hadoop Distributed File System* (HDFS) untuk menangani penyimpanan data terdistribusi dan Map Reduce untuk pemrosesan data terdistribusi yang dilakukan pada komputer (*node of cluster*) tempat data disimpan, dan Yarn (*Yet Another Resource Negotiator*) untuk mengelola sumberdaya (*resources*) termasuk penjadwalan job (Holmes, 2012).

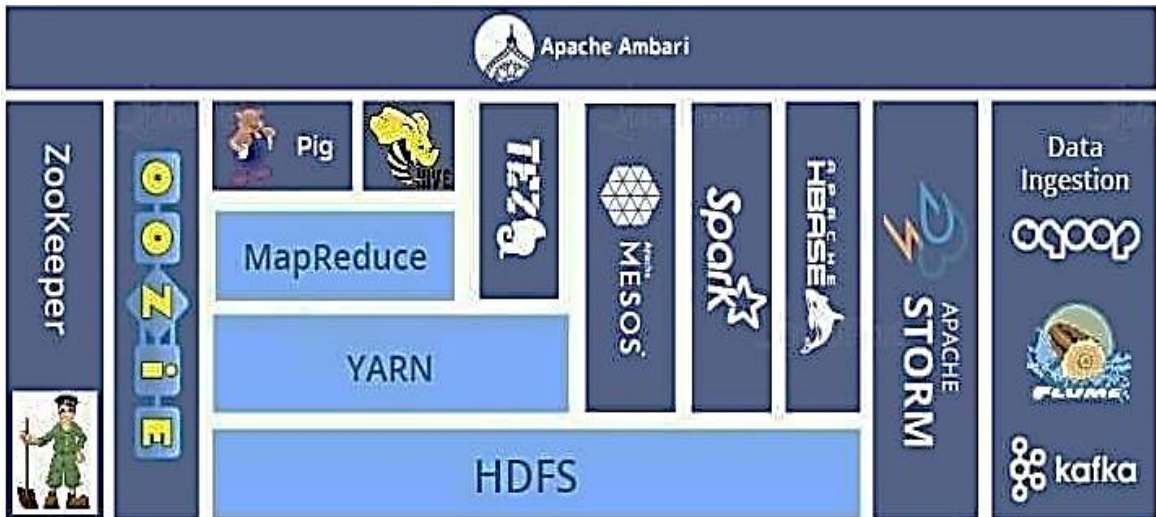
Untuk menyelesaikan berbagai persoalan komputasi, Hadoop didukung oleh berbagai teknologi yang berhubungan/ memanfaatkan teknologi inti tersebut, seperti Flume, Hbase, Hive, Zookeeper, R, Mahout, Pig, Oozie, Sqoop dan lainnya. Banyak pihak menggambarkan ekosistem Hadoop secara berbeda untuk berbagai penekanan, seperti pada gambar 10 dan 11.



Gambar 10. Ekosistem Hadoop

Sumber: <https://medium.com/@theinternetbae/big-data-dengan-hadoop-apache-hadoop-ecosystem-part-2-f01a47453cfb/>

Dalam banyak kasus juga bekerjasama dalam satu ekosistem dengan teknologi-teknologi pemrosesan *big data* yang telah dijelaskan pada gambar 6, seperti dengan Spark, Storm pada skema gambar 11.



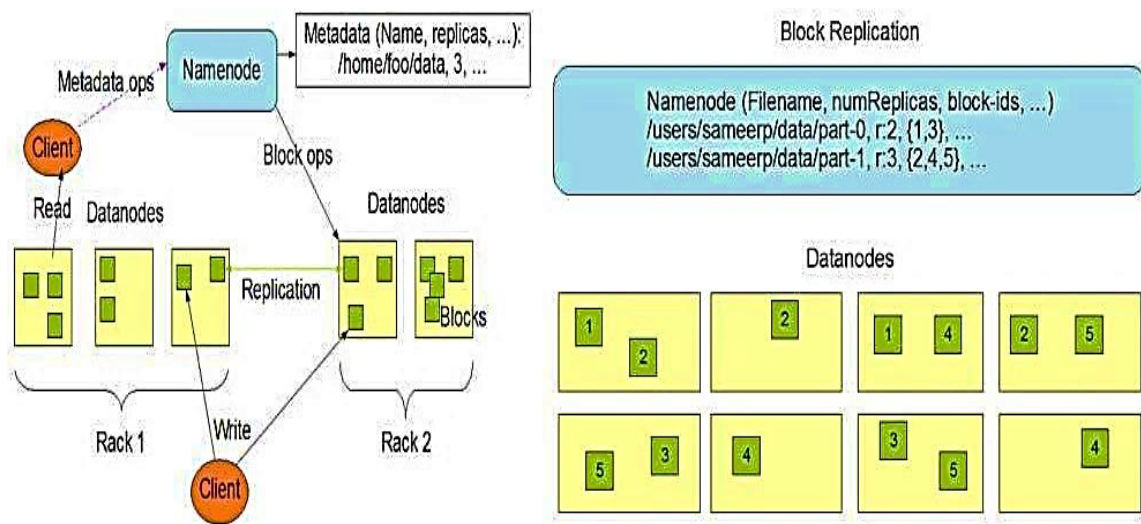
Gambar 11. Ekosistem Hadoop bersama Spark dan Storm

Sumber: <https://intellipaat.com/blog/tutorial/hadoop-tutorial/hadoop-ecosystem/>

Yang termasuk katagori perantara atau middleware (M) adalah Flume, Sqoop, Kafka. Katagori pemroses atau processor (P) adalah Spark, R, Storm, Mahout. Sedangkan yang termasuk pada katagori pustaka atau library (L) adalah Spark MLLIB, Spark SQL, Spark Stream, dan sejenisnya. HDFS, Hive, Hbase termasuk pada katagori big storage (BS). Beberapa teknologi yang fungsinya sebagai pemrosesan *big data*, antar lain yaitu:

1. HDFS (*Hadoop Distributed File System*)

Hadoop Distributed File System (HDFS) adalah sistem penyimpanan terdistribusi yang dirancang untuk menyimpan data dalam jumlah besar dengan cara yang andal dan efisien. HDFS merupakan komponen utama dari Apache Hadoop, sebuah kerangka kerja open-source yang digunakan untuk pemrosesan data skala besar. HDFS memungkinkan penyimpanan dan pemrosesan data di banyak komputer yang terhubung dalam satu kluster, sehingga menyediakan skalabilitas dan toleransi kesalahan yang tinggi. HDFS merupakan komponen dari Apache Hadoop yang paling mendasar dan merupakan sistem file terdistribusi. Arsitektur HDFS dapat dilihat pada gambar 12.



Gambar 12. Arsitektur HDFS

Sumber: <https://hadoop.apache.org/docs/current/hadoop-project-dist/hadoop-hdfs/HdfsDesign.html/>

Pada HDFS, data disimpan dalam bentuk file, kemudian setiap file dipecah menjadi potongan data dalam bentuk blok-blok yang ukurannya ditentukan (dari konfigurasi). Misalkan pada gambar 12, sebuah file dipecah menjadi 5 blok (blok 1 sampai 5). HDFS terdiri atas 2 komponen, yaitu: Namenode yang bertugas untuk menyimpan meta data (berupa nama

file, banyaknya replika, blok, lokasi penyimpanan blok) dan informasi lain tentang data), dan beberapa Datanode yang menyimpan blok data tersebut. Untuk menjaga ketersediaan data, maka suatu blok data dapat disimpan pada beberapa Datanode, misalnya blok 1 disimpan pada Datanode 1 dan 3.

Aplikasi yang ingin mengakses data yang disimpan pada HDFS (pada gambar 12 disebut client) terlebih dahulu menghubungi Namenode untuk mendapatkan informasi lokasi data, kemudian mengakses data tersebut langsung ke Datanode tempat data tersimpan. Hal ini juga berlaku dalam skenario penulisan dan pembacaan data. HDFS memiliki arsitektur master-slave yang terdiri dari beberapa komponen kunci, yaitu:

1. **NameNode:** NameNode adalah komponen master dalam HDFS yang bertanggung jawab untuk mengelola metadata dari sistem file. Metadata ini mencakup informasi tentang file, seperti nama file, lokasi blok, dan atribut lainnya. NameNode juga mengelola namespace sistem file dan mengawasi operasi seperti pembuatan, penghapusan, dan replikasi blok.
2. **DataNode:** DataNode adalah komponen slave yang bertanggung jawab untuk menyimpan data sebenarnya dalam bentuk blok. DataNode melaporkan status blok yang mereka miliki kepada NameNode secara berkala dan menjalankan operasi seperti pembacaan dan penulisan data sesuai dengan instruksi dari NameNode.
3. **Secondary NameNode:** Secondary NameNode bukanlah cadangan dari NameNode. Sebaliknya, ini bertindak sebagai pendukung dengan mengambil snapshot dari metadata NameNode dan membantu mengurangi beban kerja NameNode utama. Secondary NameNode secara periodik menggabungkan file log dengan metadata yang ada untuk mencegah log dari menjadi terlalu besar dan mengurangi waktu recovery jika NameNode mengalami kegagalan.

Cara Kerja HDFS yaitu:

- a) **Penyimpanan Data:** Data yang diunggah ke HDFS dipecah menjadi blok-blok kecil. Ukuran blok biasanya 128 MB atau 256 MB. Setiap blok data disimpan di DataNode. Untuk memastikan keandalan dan toleransi kesalahan, setiap blok direplikasi ke beberapa DataNode. Secara default, setiap blok direplikasi tiga kali (Replication Factor=3).

- b) **Metadata dan Namespace:** NameNode menyimpan metadata yang mencakup informasi tentang blok, lokasi blok, dan namespace sistem file. Ketika sebuah file dibuat atau dihapus, metadata ini diperbarui oleh NameNode.
- c) **Pemrosesan Data:** Ketika data perlu diproses, HDFS bekerja sama dengan MapReduce atau alat pemrosesan lainnya untuk mengakses data dari DataNode. Pemrosesan dilakukan secara paralel di beberapa DataNode untuk meningkatkan kecepatan dan efisiensi.
- d) **Replikasi dan Toleransi Kesalahan:** Jika sebuah DataNode gagal, NameNode akan mereplikasi blok-blok data yang ada di DataNode yang gagal ke DataNode lain untuk memastikan bahwa jumlah replika yang diperlukan tetap tersedia. NameNode terus memantau kesehatan DataNode dan memastikan bahwa data selalu tersedia dan dapat diakses.

Kelebihan HDFS yaitu:

a) **Skalabilitas**

- HDFS dirancang untuk menyimpan dan memproses data dalam skala petabyte dengan menambahkan lebih banyak DataNode ke dalam kluster.
- Sistem ini dapat dengan mudah diskalakan secara horizontal untuk menangani peningkatan jumlah data.

b) **Toleransi Kesalahan**

- HDFS menyediakan toleransi kesalahan melalui replikasi data. Jika sebuah DataNode gagal, data tetap dapat diakses dari DataNode lain yang memiliki replika blok.
- Mekanisme replikasi dan monitoring yang dijalankan oleh NameNode memastikan bahwa data selalu tersedia meskipun terjadi kegagalan pada komponen hardware.

c) **Efisiensi Biaya**

- HDFS dirancang untuk berjalan pada hardware komoditas, yang berarti biaya investasi perangkat keras lebih rendah dibandingkan dengan solusi penyimpanan data proprietary.
- Dengan menggunakan hardware yang murah, organisasi dapat menyimpan dan memproses data dalam jumlah besar dengan biaya yang lebih efisien.

d) Integrasi dengan Hadoop Ecosystem

- HDFS terintegrasi erat dengan ekosistem Hadoop, termasuk MapReduce, Hive, Pig, dan lainnya.
- Integrasi ini memungkinkan pemrosesan data yang efisien dan fleksibel dengan menggunakan berbagai alat dan teknologi yang tersedia dalam ekosistem Hadoop.

Kekurangan dari HDFS yaitu:

a) Latensi Tinggi untuk Data Kecil

- HDFS tidak dioptimalkan untuk menangani banyak file kecil. Latensi untuk mengakses file-file kecil bisa cukup tinggi karena overhead dari pengelolaan metadata.
- Sistem ini lebih cocok untuk menangani file berukuran besar yang dipecah menjadi blok-blok besar.

b) Keterbatasan Metadata pada NameNode

- NameNode menyimpan semua metadata di memori, sehingga kapasitas memori NameNode bisa menjadi bottleneck dalam kluster yang sangat besar.
- Meskipun ada teknik untuk mengelola ini, seperti Federation dan High Availability, masih ada keterbatasan dalam skala metadata yang bisa dikelola oleh satu NameNode.

c) Tidak Mendukung Pembaruan Blok Data

- HDFS dirancang untuk model write-once-read-many, artinya data yang sudah ditulis tidak dapat diubah. Pembaruan pada data harus dilakukan dengan menulis ulang seluruh file atau blok.
- Ini bisa menjadi kendala dalam aplikasi yang memerlukan pembaruan data secara berkala.

d) Kompleksitas Manajemen

- Mengelola dan mengkonfigurasi kluster HDFS bisa kompleks dan memerlukan keahlian khusus.
- Pemeliharaan dan pengawasan kluster besar memerlukan alat dan prosedur manajemen yang canggih.

2. Map Reduce

Map Reduce merupakan *framework* pemrograman terdistribusi yang dapat diterapkan terhadap data yang tersimpan pada HDFS. Pemrograman native pada Map Reduce menggunakan bahasa Java. Program dieksekusi pada Datanode dimana data yang diproses disimpan. Namun demikian Map Reduce hanya mendukung pemrograman paralel model batch (*Paralell Batch Processing*). Map Reduce membutuhkan Resource Manager yang bertugas mengelola sumberdaya (berpasangan dengan Namenode), dan Node Manager yang mengelola eksekusi pada setiap Datanode cluster (berpasangan dengan Datanode). Versi terakhir dari Map Reduce menggunakan Yarn sebagai Resource Manager.

3. HBase

HBase merupakan sistem basis data berbasis kolom (*Column Database System*). Pada basis data ini, data disimpan dalam bentuk pasangan key-value. Hbase berjalan di atas HDFS dan memiliki komponen pemroses terdistribusi, yaitu: Hbase Master (Hmaster) yang menyimpan meta data dan Region server yang menyimpan data nyata. Oleh karena itu, Hbase dapat menyimpan data dalam jumlah besar dengan akses yang cepat. Dengan model key-value (*column*) tersebut, maka Hbase dapat digunakan untuk menyimpan data yang strukturnya berbeda untuk setiap record. Model Hbase ini dikembangkan oleh Google menjadi BigTable pada Google Cloud *Platform* (GCP).

4. Hive

Hive merupakan sistem pemroses data yang berjalan di atas Map Reduce namun memiliki antarmuka akses berbasis *Structure Query Language* (SQL). Data yang tersimpan pada HDFS maupun HBase dapat diakses menggunakan perintah-perintah SQL layaknya basis data relasional, dan memiliki antarmuka native pada pemrograman Java Map Reduce. Dengan kemampuan ini, Hive menjadi dasar implementasi Data Warehouse dan Data Mart di atas HDFS dan HBase.

5. Spark

Spark merupakan pemroses data berbasis memori. Jika Map Reduce cocok untuk pemrosesan batch, maka untuk pemrosesan online yang menggunakan iterasi berulang-ulang, Spark sangat cocok digunakan. Spark menjalankan aplikasi menggunakan *Directed Acyclic Graph* (DAG). DAG terdiri atas simpul (*vertices*) yang menyatakan *Resilient Distributed Dataset* (RDD) berupa koleksi berbentuk array yang terdistribusi, dan sisi (*edge*) yang menyatakan operasi yang diterapkan terhadap RDD tersebut. Spark juga dilengkapi dengan berbagai pustaka (*library*) untuk pemrosesan *big data* di antaranya: Spark Stream (pemrosesan data stream), Spark MLlib (menyediakan algoritma *Machine learning* seperti clustering, klasifikasi, dan rekomendasi), Spark GraphX (pemrosesan big graph) dan Spark SQL (untuk mengakses *big data* dengan perintah-perintah SQL).

6. Sqoop

Sqoop merupakan kakas (*tool*) yang sangat efisien untuk mentransfer data dari Hadoop (HDFS dan HBase) ke penyimpanan data terstruktur seperti basis data relasional (*Relational Database Management System* - RDBMS) dan sebaliknya (atau ekspor/ impor data Hadoop-RDBMS). Sqoop dapat mengakses RDBMS menggunakan antarmuka Java Database Connectivity (JDBC) sehingga dapat berhubungan dengan semua basis data yang menyediakan antarmuka JDBC tersebut.

7. Flume

Flume merupakan kakas (*tool*) yang menyediakan layanan pengumpulan (*collecting*), agregasi (*agregating*), dan pemindahan (*moving*) data log sekala besar. Flume dapat berjalan secara terdistribusi, reliabel dan memiliki tingkat ketersediaan (*availability*) yang tinggi.

Arsitektur Flume didasarkan pada streaming data flow, sehingga sangat mendukung aplikasi analisis online (online analytical application). Oleh karena itu, Flume juga cocok untuk mengumpulkan data stream dari berbagai sumber seperti: Twitter, Web News, dan sistem berbasis *Internet of Things* (IoT).

8. R Hadoop & Spark

R merupakan bahasa pemrograman yang kaya akan fungsi dan pustaka (library) statistik. R sangat powerful digunakan untuk analisis data berbasis statistik. R telah memiliki versi yang dapat berjalan di atas Hadoop (R-Hadoop) dan dapat berjalan memanfaatkan Spark (Spak-R). Dengan menggunakan R di atas Hadoop maupun Spark dapat memaksimalkan ketersediaan fungsi statistik dan kemampuan pemrosesan data sekala besar, baik batch processing maupun online iteratif.

9. Kafka

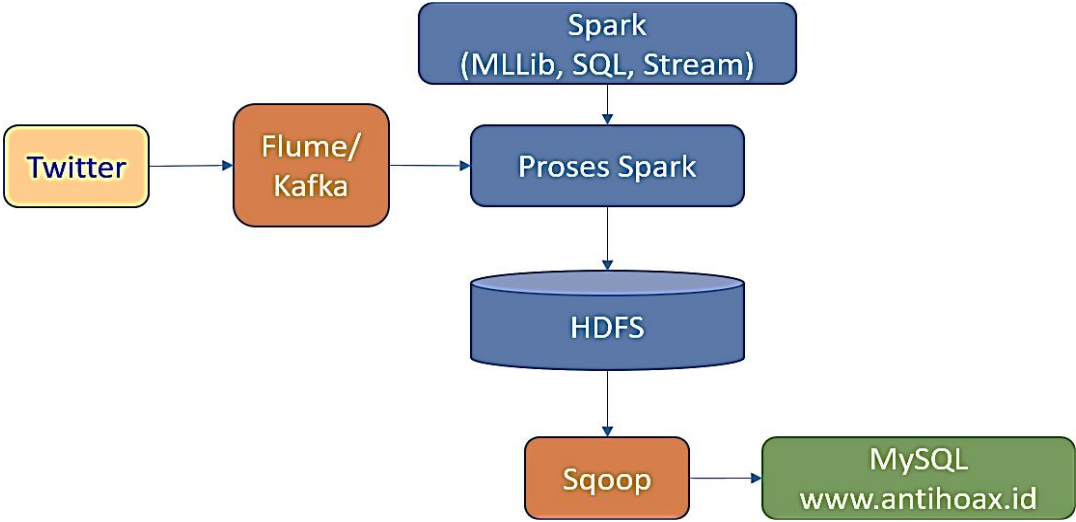
Kafka merupakan *platform* stream yang memiliki kemampuan menangani pengelolaan distribusi pesan berbasis protokol publish-subscribe untuk data stream dalam skala besar. Kafka cocok digunakan untuk aplikasi yang memiliki aksi tertentu yang dipicu oleh event data stream. Kafka dapat bekerja dengan konfigurasi cluster beberapa komputer sehingga memiliki kemampuan menangani data stream berskala besar. Dengan demikian Kafka dapat digunakan untuk menangani koleksi dan respon terhadap data media sosial seperti Twitter dengan kecepatan dan *volume* yang besar.

2.1.6. Contoh Penggunaan Teknologi *Big Data*

Berikut adalah contoh penggunaan teknologi *big data* untuk mengantisipasi penyebaran berita bohong (hoax). Misalkan berita yang ditangani bersumber dari media sosial Twitter. Analisis hoax atau tidak dilakukan menggunakan algoritma *machine learning* berupa algoritma clustering dan klasifikasi. Hasil analisis dikirimkan ke situs www.antihoax.id yang memiliki basis data MySQL *enterprise*. Arsitektur teknologi dari contoh kasus di atas dapat dilihat pada gambar 13 bahwa data Twitter dikumpulkan menggunakan middleware Flume atau Kafka. Hasil pengumpulan data ini diproses secara real time (waktu nyata) menggunakan aplikasi yang berjalan di atas Spark.

Aplikasi ini mengakses pustaka (library) Spark Stream untuk menangani data stream yang dipropagasi dari Flume/Kafka, kemudian dilakukan analisis menggunakan algoritma *Machine learning* dari Spark MLLib. Hasil analisis tersebut dikumpulkan dan dianalisis lebih lanjut menggunakan Spark SQL, lalu hasilnya disimpan di HDFS. Kemudian, dengan memanfaatkan Scoop, hasil analisis di HDFS itu pada setiap waktu tertentu diekspor ke basisdata MySQL yang menjadi bagian dari sistem situs www.antihoax.id.

Dengan tersedianya hasil analisis berita Twitter yang senantiasa *up to date* di basisdata itu, situs www.antihoax.id dapat mempublikasikan berita-berita dari Twitter mana saja yang termasuk hoax. Dengan menggunakan teknologi *big data*, maka data Twitter yang mengalir dengan sangat cepat dan menumpuk hingga berukuran sangat besar dapat ditangani. Pemrosesan data secara real time dan sinkronisasi hasil analisis ke situs www.antihoax.id juga dimungkinkan.



Gambar 13. Arsitektur Teknologi *Big Data* Untuk Kasus Anti Hoax.

2.2. Data Analytic

Data Analytics adalah proses menganalisis data untuk mendapatkan wawasan yang berguna, mendukung pengambilan keputusan, dan mengungkap pola-pola tersembunyi dalam data. Dalam era digital, *data analytics* menjadi sangat penting bagi berbagai bidang, termasuk bisnis, pemerintahan, kesehatan, pendidikan, dan lain-lain. Artikel ini akan menjelaskan secara mendetail tentang konsep, proses, teknik, alat, serta penerapan *data analytics* dalam berbagai sektor.

2.2.1. Konsep Dasar *Data Analytics*

Data analytics melibatkan serangkaian teknik dan alat untuk mengolah data mentah menjadi informasi yang dapat digunakan untuk pengambilan keputusan. Beberapa konsep dasar dalam *data analytics* meliputi:

1. **Data:** Data adalah kumpulan fakta, seperti angka, teks, gambar, atau video yang dapat diolah untuk mendapatkan informasi.
2. **Informasi:** Informasi adalah hasil dari pengolahan data yang memberikan makna atau wawasan tertentu.
3. **Wawasan (Insight):** Wawasan adalah pemahaman mendalam yang diperoleh dari analisis data, yang dapat digunakan untuk membuat keputusan atau tindakan.

2.2.2. Proses *Data Analytics*

Proses *data analytics* biasanya terdiri dari beberapa tahap, yaitu:

1. **Pengumpulan Data (*Data Collection*):** Data dapat dikumpulkan dari berbagai sumber, seperti basis data perusahaan, sensor, log aplikasi, media sosial, dan lain-lain. Sumber data ini bisa berupa data terstruktur (misalnya, tabel dalam basis data) atau data tidak terstruktur (misalnya, teks dari media sosial).
2. **Pembersihan Data (*Data Cleaning*):** Data yang dikumpulkan sering kali memiliki kesalahan, data yang hilang, atau duplikasi. Proses pembersihan data melibatkan identifikasi dan perbaikan kesalahan, pengisian data yang hilang, dan penghapusan data yang duplikat.
3. **Pengolahan Data (*Data Processing*):** Data yang sudah bersih kemudian diproses untuk disiapkan untuk analisis lebih lanjut. Proses ini bisa melibatkan transformasi data, penggabungan dataset, dan normalisasi data.
4. **Analisis Data (*Data Analysis*):** Pada tahap ini, data dianalisis menggunakan berbagai teknik statistik dan algoritma. Analisis data bisa meliputi analisis deskriptif, diagnostik, prediktif, dan preskriptif.
 - Analisis Deskriptif: Menganalisis data untuk mengetahui apa yang terjadi.

- Analisis Diagnostik: Menganalisis data untuk mengetahui mengapa sesuatu terjadi.
 - Analisis Prediktif: Menggunakan data historis untuk memprediksi apa yang akan terjadi di masa depan.
 - Analisis Preskriptif: Memberikan rekomendasi tindakan berdasarkan hasil analisis.
5. Visualisasi Data (*Data Visualization*): Visualisasi data adalah teknik untuk menyajikan data dalam bentuk grafik atau diagram untuk memudahkan pemahaman. Alat visualisasi data membantu dalam mengidentifikasi pola, tren, dan anomali dalam data.
 6. Pengambilan Keputusan (*Decision Making*): Wawasan yang diperoleh dari analisis data digunakan untuk mendukung pengambilan keputusan. Keputusan ini bisa berupa strategi bisnis, kebijakan publik, atau tindakan operasional lainnya.

2.2.3. Teknik dan Alat dalam *Data Analytics*

Berbagai teknik dan alat digunakan dalam *data analytics* untuk mengolah dan menganalisis data. Berikut adalah beberapa teknik dan alat yang umum digunakan:

- a. Teknik dalam *data analytics* yaitu:
 1. Statistik Deskriptif: Statistik deskriptif meliputi pengukuran seperti mean, median, mode, varians, dan standar deviasi. Teknik ini digunakan untuk meringkas dan menggambarkan karakteristik utama dari suatu dataset.
 2. Analisis Regresi: Analisis regresi adalah teknik statistik untuk mengidentifikasi hubungan antara variabel dependen dan satu atau lebih variabel independen. Teknik ini sering digunakan untuk prediksi dan peramalan.
 3. Klasterisasi (*Clustering*): Klasterisasi adalah teknik pembelajaran mesin yang digunakan untuk mengelompokkan data ke dalam kelompok-kelompok yang mirip satu sama lain. Algoritma klasterisasi seperti K-Means dan *Hierarchical Clustering* sering digunakan dalam analisis data.
 4. Klasifikasi (*Classification*): Klasifikasi adalah teknik pembelajaran mesin yang digunakan untuk mengkategorikan data ke dalam kelas-kelas yang sudah

ditentukan sebelumnya. Algoritma seperti Decision Trees, Random Forests, dan Support Vector Machines sering digunakan untuk tugas klasifikasi.

5. Analisis Asosiasi (*Association Analysis*): Analisis asosiasi digunakan untuk menemukan hubungan atau pola antara variabel-variabel dalam dataset. Teknik ini sering digunakan dalam analisis keranjang belanja untuk menemukan asosiasi produk yang sering dibeli bersama.
 6. Analisis Rangkaian Waktu (*Time Series Analysis*): Analisis rangkaian waktu digunakan untuk menganalisis data yang dikumpulkan atau diurutkan berdasarkan waktu. Teknik ini digunakan untuk peramalan dan identifikasi tren musiman.
- b. Alat dalam *data analytics* yaitu:
1. Spreadsheet (Misalnya, Microsoft Excel, Google Sheets): Spreadsheet adalah alat yang umum digunakan untuk analisis data sederhana. Alat ini memungkinkan pengguna untuk melakukan perhitungan, membuat grafik, dan melakukan analisis statistik dasar.
 2. Bahasa Pemrograman (Misalnya, Python, R): Python dan R adalah dua bahasa pemrograman yang populer untuk analisis data. Keduanya memiliki pustaka yang kuat untuk pengolahan data, analisis statistik, dan visualisasi data (misalnya, Pandas, NumPy, SciPy untuk Python; dplyr, ggplot2 untuk R).
 3. Perangkat Lunak Statistik (Misalnya, SPSS, SAS, Stata): Perangkat lunak statistik seperti SPSS, SAS, dan Stata digunakan untuk analisis statistik yang lebih mendalam. Alat ini menawarkan berbagai fungsi untuk analisis data deskriptif, regresi, analisis varians, dan lain-lain.
 4. Platform Analitik Bisnis (Misalnya, Tableau, Power BI): Platform analitik bisnis seperti Tableau dan Power BI digunakan untuk visualisasi data interaktif dan analitik bisnis. Alat ini memungkinkan pengguna untuk membuat dashboard yang dinamis dan menyajikan data dalam bentuk yang mudah dipahami.
 5. Sistem Manajemen Basis Data (Misalnya, SQL Server, MySQL, PostgreSQL): Sistem manajemen basis data digunakan untuk menyimpan dan mengelola data.

Alat ini memungkinkan pengguna untuk mengakses dan mengolah data menggunakan bahasa SQL.

2.2.4. Penerapan *Data Analytics* dalam Berbagai Sektor

1. Bisnis

Data analytics telah menjadi bagian penting dalam bisnis modern. Perusahaan menggunakan analisis data untuk mengidentifikasi peluang bisnis, meningkatkan efisiensi operasional, dan membuat keputusan strategis. Beberapa penerapan *data analytics* dalam bisnis meliputi:

- Analisis Pelanggan: Menggunakan data pelanggan untuk memahami kebutuhan dan preferensi mereka, sehingga dapat menawarkan produk dan layanan yang lebih sesuai.
- Optimasi Rantai Pasok: Menganalisis data rantai pasok untuk meningkatkan efisiensi dan mengurangi biaya operasional.
- Deteksi Penipuan: Menggunakan analitik data untuk mendeteksi pola penipuan dalam transaksi keuangan.

2. Pemerintahan

Pemerintahan juga memanfaatkan *data analytics* untuk meningkatkan pelayanan publik dan membuat kebijakan yang lebih baik. Penerapan *data analytics* dalam pemerintahan meliputi:

- Kebijakan Publik Berbasis Data: Menganalisis data untuk merancang kebijakan yang lebih efektif dan tepat sasaran.
- Peningkatan Layanan Publik: Menggunakan analitik data untuk meningkatkan efisiensi dan kualitas layanan publik, seperti layanan kesehatan dan pendidikan.
- Manajemen Kota Pintar: Menerapkan analitik data dalam manajemen kota pintar untuk meningkatkan keamanan, efisiensi transportasi, dan pengelolaan sumber daya.

3. Kesehatan

Dalam sektor kesehatan, *data analytics* digunakan untuk meningkatkan diagnosis, perawatan, dan manajemen kesehatan secara keseluruhan. Beberapa penerapan *data analytics* dalam kesehatan meliputi:

- Analisis Data Pasien: Menganalisis data pasien untuk memprediksi penyakit dan merancang intervensi kesehatan yang lebih efektif.
- Manajemen Rumah Sakit: Menggunakan analitik data untuk mengoptimalkan operasional rumah sakit, seperti penjadwalan staf dan pengelolaan inventaris.
- Penelitian Klinis: Menganalisis data dari uji klinis untuk mengidentifikasi efektivitas dan keamanan obat baru.

4. Pendidikan

Data analytics juga digunakan dalam pendidikan untuk meningkatkan proses pembelajaran dan manajemen pendidikan. Penerapan *data analytics* dalam pendidikan meliputi:

- Analisis Performa Siswa: Menganalisis data performa siswa untuk mengidentifikasi area yang memerlukan perbaikan dan merancang program pendidikan yang lebih efektif.
- Manajemen Institusi Pendidikan: Menggunakan analitik data untuk mengelola operasional institusi pendidikan, seperti penjadwalan kelas dan pengelolaan sumber daya.
- Personalized Learning: Menerapkan analitik data untuk merancang pengalaman belajar yang disesuaikan dengan kebutuhan dan gaya belajar individu.

5. Keuangan

Dalam sektor keuangan, *data analytics* digunakan untuk mengelola risiko, mendeteksi penipuan, dan membuat keputusan investasi. Penerapan *data analytics* dalam keuangan meliputi:

- Manajemen Risiko: Menggunakan analitik data untuk mengidentifikasi dan mengelola risiko keuangan.
- Deteksi Penipuan: Menganalisis data transaksi untuk mendeteksi pola penipuan dan mengurangi kerugian.
- Pengelolaan Portofolio: Menerapkan analitik data untuk membuat keputusan investasi yang lebih baik dan mengoptimalkan portofolio investasi.

6. Ritel

Dalam sektor ritel, *data analytics* digunakan untuk mengoptimalkan rantai pasok, meningkatkan pengalaman pelanggan, dan merancang strategi pemasaran yang lebih efektif.

Penerapan *data analytics* dalam ritel meliputi:

- Analisis Pelanggan: Menganalisis data pelanggan untuk memahami kebutuhan dan preferensi mereka, sehingga dapat menawarkan produk yang lebih sesuai.
- Manajemen Inventaris: Menggunakan analitik data untuk mengelola inventaris secara lebih efisien, mengurangi kelebihan stok, dan menghindari kekurangan stok.
- Strategi Pemasaran: Menerapkan analitik data untuk merancang kampanye pemasaran yang lebih efektif dan meningkatkan penjualan.

7. Energi dan Lingkungan

Data analytics juga digunakan dalam sektor energi dan lingkungan untuk meningkatkan efisiensi dan keberlanjutan. Penerapan *data analytics* dalam sektor ini meliputi:

- Manajemen Energi: Menggunakan analitik data untuk mengoptimalkan penggunaan energi dan mengurangi konsumsi energi.
- Pemantauan Lingkungan: Menganalisis data lingkungan untuk memantau kualitas udara, air, dan tanah, serta merancang kebijakan yang lebih efektif untuk melindungi lingkungan.
- Perencanaan Sumber Daya: Menerapkan analitik data untuk merencanakan dan mengelola sumber daya alam secara lebih berkelanjutan.

2.2.5. Tantangan dalam *Data Analytics*

Meskipun *data analytics* menawarkan banyak manfaat, ada beberapa tantangan yang harus diatasi, antara lain:

1. Kualitas Data: Data yang digunakan dalam analisis harus berkualitas tinggi. Data yang tidak akurat, tidak lengkap, atau tidak konsisten dapat menghasilkan wawasan yang menyesatkan.

2. **Keamanan dan Privasi Data:** Perlindungan data pribadi dan sensitif adalah tantangan utama dalam *data analytics*. Pemerintah dan perusahaan harus memastikan bahwa data yang dikumpulkan dan dianalisis dilindungi dari akses yang tidak sah dan kebocoran.
3. **Keterampilan dan Pengetahuan:** *Data analytics* membutuhkan keterampilan dan pengetahuan khusus dalam analisis statistik, pemrograman, dan pemahaman domain. Kekurangan tenaga ahli di bidang ini dapat menjadi hambatan dalam penerapan *data analytics*.
4. **Kompleksitas Data:** Data yang dikumpulkan dari berbagai sumber bisa sangat kompleks dan memerlukan teknik analisis yang canggih. Mengelola dan menganalisis data ini memerlukan alat dan metode yang tepat.
5. **Integrasi Data:** Data yang berasal dari berbagai sumber dan sistem perlu diintegrasikan dengan baik untuk mendapatkan wawasan yang holistik. Integrasi data ini bisa menjadi tantangan teknis dan organisasi.
6. **Etika dalam *Data Analytics*:** Penggunaan *data analytics* harus mempertimbangkan aspek etika, termasuk penggunaan data yang adil, transparansi dalam proses analisis, dan penghindaran bias dalam model analitik.

2.2.6. Masa Depan *Data Analytics*

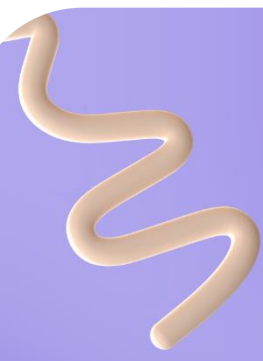
Masa depan *data analytics* sangat cerah, dengan berbagai tren dan perkembangan yang akan semakin mendorong penerapannya. Beberapa tren utama yang akan mempengaruhi masa depan *data analytics* meliputi:

1. **Kecerdasan Buatan (AI) dan Pembelajaran Mesin (*Machine Learning*/ML):** Integrasi AI dan *machine learning* dalam *data analytics* akan semakin meningkat, memungkinkan analisis data yang lebih mendalam dan otomatisasi proses analitik.
2. ***Big Data*:** Dengan semakin banyaknya data yang dihasilkan oleh berbagai perangkat dan sistem, *big data* akan menjadi fokus utama dalam *data analytics*. Alat dan teknik baru akan terus dikembangkan untuk mengelola dan menganalisis *volume* data yang besar ini.

3. *Internet of Things (IoT)*: IoT akan menghasilkan jumlah data yang sangat besar dari berbagai perangkat yang terhubung. Data ini akan digunakan untuk analisis yang lebih canggih dalam berbagai sektor, termasuk manufaktur, transportasi, dan kesehatan.
4. *Analitik Prediktif dan Preskriptif*: Analitik prediktif dan preskriptif akan semakin banyak digunakan untuk memberikan wawasan yang lebih maju dan rekomendasi tindakan yang lebih tepat.
5. *Komputasi Awan (Cloud Computing)*: Komputasi awan akan terus memainkan peran penting dalam *data analytics*, menyediakan infrastruktur yang fleksibel dan skalabel untuk pengolahan dan penyimpanan data.
6. *Visualisasi Data yang Lebih Interaktif*: Alat visualisasi data akan terus berkembang, menyediakan cara yang lebih interaktif dan intuitif untuk memahami dan berinteraksi dengan data.

Data analytics adalah proses yang penting dan kompleks, yang melibatkan berbagai teknik dan alat untuk mengolah data mentah menjadi wawasan yang dapat digunakan untuk pengambilan keputusan. Dalam berbagai sektor, *Data analytics* telah membantu meningkatkan efisiensi, kualitas, dan efektivitas layanan dan operasi.

Meskipun ada beberapa tantangan yang harus diatasi, masa depan *Data analytics* sangat menjanjikan dengan perkembangan teknologi seperti AI, *Big Data*, IoT, dan komputasi awan. Dengan memahami dan mengatasi tantangan ini, organisasi dapat memanfaatkan *Data analytics* untuk mencapai tujuan mereka dan menghasilkan nilai yang lebih besar dari data mereka.



Dalam era digital, data adalah harta karun. Analitik adalah kuncinya, dan data science adalah panduan yang membuka potensi tak terbatas

Mengubah data menjadi wawasan adalah seni, dan analitik adalah kuas yang melukis gambaran masa depan



Transformasi digital bukan hanya tentang teknologi, tetapi tentang menggunakan data untuk membuat keputusan yang lebih cerdas dan menciptakan masa depan yang lebih baik

Big Data adalah sumber daya alam baru. Dengan data analytic, kita dapat mengeksplorasinya dan dengan data science, kita dapat mengubahnya menjadi inovasi



!nolag!
 qabst wendprrmuys wendq!
 qau qeugau qatg zcienc6' kfg
 kfg qabst wendekzbiorezius
 efgw palk' deugau qatg auilzic'
 big datg qatgrr zaurer qala

BAB 3

URGENSI TRANSFORMASI DIGITAL

Transformasi digital adalah integrasi teknologi digital ke dalam semua aspek operasional dan layanan organisasi, termasuk pemerintahan. Di era digital ini, transformasi digital menjadi sangat mendesak bagi pemerintahan, termasuk di Provinsi Kalimantan Timur, untuk meningkatkan efisiensi, transparansi, dan kualitas pelayanan publik. Transformasi ini tidak hanya melibatkan penggunaan teknologi baru tetapi juga perubahan mendasar dalam cara kerja, budaya organisasi, dan interaksi dengan masyarakat.

Provinsi Kalimantan Timur, dengan kekayaan sumber daya alam yang melimpah dan posisi strategisnya sebagai salah satu pusat ekonomi di Indonesia, memiliki potensi besar untuk menjadi pelopor dalam penerapan teknologi digital dalam pemerintahan. Namun, potensi ini belum sepenuhnya terealisasi karena berbagai tantangan, termasuk keterbatasan infrastruktur, kurangnya sumber daya manusia yang terampil di bidang teknologi informasi, dan birokrasi yang kompleks.

3.1. Peningkatan Efisiensi Operasional

Urgensi transformasi digital dalam peningkatan efisiensi operasional dimana beberapa hal yang menjadi faktor efisiensi yaitu:

- a. **Automasi Proses:** Automasi proses adalah salah satu aspek kunci dalam transformasi digital yang dapat meningkatkan efisiensi operasional pemerintahan. Di Kalimantan Timur, banyak proses administrasi yang masih dilakukan secara manual, menyebabkan waktu yang lama dan biaya yang tinggi. Dengan menerapkan teknologi seperti Robotic Process Automation (RPA), proses-proses ini dapat diotomatisasi, mengurangi waktu yang diperlukan untuk menyelesaikan tugas-tugas administratif dan memungkinkan pegawai pemerintah untuk fokus pada tugas-tugas yang lebih strategis.

- b. Pengurangan Redundansi: Sistem digital dapat mengintegrasikan berbagai fungsi pemerintahan, menghindari duplikasi data dan usaha. Saat ini, data dan informasi sering kali tersebar di berbagai departemen tanpa ada integrasi yang efektif. Dengan sistem digital terintegrasi, data dapat diakses dan digunakan oleh berbagai departemen, meningkatkan efisiensi dan akurasi.

3.2. Transparansi dan Akuntabilitas

Urgensi transformasi digital dalam proses dan aktivitas transparansi dan akuntabilitas yaitu:

- a. Data Terbuka: Implementasi sistem data terbuka memungkinkan masyarakat mengakses informasi pemerintahan secara transparan. Ini meningkatkan kepercayaan publik terhadap pemerintahan, karena masyarakat dapat melihat dan memahami bagaimana pemerintah mengelola sumber daya dan membuat keputusan. Di Kalimantan Timur, penerapan kebijakan data terbuka dapat mencakup publikasi anggaran, pengeluaran, dan proyek-proyek pembangunan, sehingga masyarakat dapat memantau dan mengevaluasi kinerja pemerintah.
- b. Pelacakan *Real-Time*: Teknologi digital memungkinkan pelacakan real-time terhadap kinerja program dan proyek pemerintah. Dengan sistem pelacakan real-time, pemerintah dapat segera mengidentifikasi dan menyelesaikan masalah yang muncul, memastikan bahwa proyek-proyek berjalan sesuai jadwal dan anggaran. Hal ini juga memungkinkan masyarakat untuk mengawasi dan memberikan masukan secara langsung.

3.3. Pelayanan Publik yang Lebih Baik

Urgensi transformasi digital dalam pelaksanaan dan aktivitas pelayanan publik yang lebih baik yaitu:

- a. Layanan Berbasis Data: Data analitik membantu pemerintah memahami kebutuhan dan preferensi warga, sehingga dapat menyediakan layanan yang lebih sesuai dan tepat waktu. Di Kalimantan Timur, penggunaan analitik data dapat membantu dalam berbagai bidang seperti kesehatan, pendidikan, dan infrastruktur. Misalnya, data kesehatan dapat digunakan untuk memprediksi dan mengatasi wabah penyakit,

sementara data pendidikan dapat membantu merancang program yang lebih efektif untuk meningkatkan kualitas pendidikan.

- b. **Aksesibilitas 24/7:** Sistem digital memungkinkan warga mengakses layanan pemerintah kapan saja dan di mana saja, tanpa harus terikat oleh jam kerja kantor. Dengan portal layanan digital, warga dapat mengurus berbagai keperluan administrasi seperti perizinan, pembayaran pajak, dan pendaftaran layanan publik dari rumah atau tempat kerja mereka. Ini tidak hanya memudahkan warga tetapi juga mengurangi beban kerja di kantor-kantor pemerintah.

3.4. Pengambilan Keputusan yang Lebih Cepat dan Tepat

Urgensi transformasi digital dalam pengambilan keputusan yang lebih cepat dan tepat yaitu:

- a. **Analitik Prediktif:** *data science* dan *big data analytics* memungkinkan pemerintah melakukan analisis prediktif untuk perencanaan dan pengambilan keputusan yang lebih baik. Dengan menganalisis data historis dan tren saat ini, pemerintah dapat membuat prediksi yang lebih akurat mengenai kebutuhan dan tantangan masa depan. Misalnya, analitik prediktif dapat digunakan untuk meramalkan permintaan akan layanan kesehatan, infrastruktur, dan layanan publik lainnya.
- b. **Penggunaan AI dan *Machine learning*:** Teknologi kecerdasan buatan (AI) dan *machine learning* dapat digunakan untuk mengolah data dalam jumlah besar dan kompleks, menemukan pola-pola yang tidak terlihat sebelumnya, dan memberikan rekomendasi yang lebih baik. Di Kalimantan Timur, teknologi ini dapat digunakan dalam berbagai aplikasi seperti perencanaan kota, pengelolaan lalu lintas, dan pengawasan lingkungan.

3.5. Peningkatan Kualitas Hidup Masyarakat

Urgensi transformasi digital dalam kegiatan peningkatan kualitas hidup masyarakat yaitu:

- a. **Kesehatan dan Pendidikan:** Transformasi digital dapat meningkatkan kualitas layanan kesehatan dan pendidikan di Kalimantan Timur. Dalam bidang kesehatan, telemedicine dan sistem informasi kesehatan elektronik dapat memperluas akses ke

layanan kesehatan berkualitas, terutama di daerah terpencil. Dalam bidang pendidikan, *e-learning* dan *platform* pembelajaran digital dapat meningkatkan akses dan kualitas pendidikan, memungkinkan siswa untuk belajar dari mana saja dan kapan saja.

- b. **Infrastruktur dan Lingkungan:** Teknologi digital dapat digunakan untuk mengelola dan memantau infrastruktur dan lingkungan secara lebih efisien. Sistem pemantauan berbasis sensor dan *Internet of Things (IoT)* dapat digunakan untuk memantau kondisi jalan, jembatan, dan infrastruktur lainnya secara real-time, sehingga perbaikan dapat dilakukan lebih cepat. Teknologi ini juga dapat digunakan untuk memantau dan mengelola sumber daya alam, seperti air dan energi, secara lebih efisien dan berkelanjutan.

3.6. Tantangan dalam Transformasi Digital

Dalam tantangan dalam transformasi digital yaitu:

- a. **Keterbatasan infrastruktur:** Salah satu tantangan utama dalam transformasi digital di Kalimantan Timur adalah keterbatasan infrastruktur teknologi informasi dan komunikasi (TIK). Di banyak daerah, akses internet masih terbatas dan tidak merata. Pemerintah perlu berinvestasi dalam pengembangan infrastruktur TIK untuk memastikan bahwa semua daerah dapat menikmati manfaat dari transformasi digital.
- b. **Kurangnya sumber daya manusia yang terampil:** Transformasi digital membutuhkan sumber daya manusia yang terampil di bidang teknologi informasi. Saat ini, ada kekurangan tenaga ahli di bidang *data science*, analitik data, dan manajemen teknologi informasi di pemerintahan. Pemerintah perlu mengembangkan program pelatihan dan pendidikan untuk meningkatkan keterampilan tenaga kerja di bidang ini.
- c. **Masalah privasi dan keamanan data:** Transformasi digital juga membawa tantangan dalam hal privasi dan keamanan data. Dengan semakin banyaknya data yang dikumpulkan dan diolah, risiko kebocoran data dan serangan siber juga meningkat. Pemerintah perlu mengembangkan kebijakan dan infrastruktur keamanan data yang kuat untuk melindungi data dari ancaman tersebut.
- d. **Resistensi terhadap perubahan:** Perubahan besar dalam cara kerja dan budaya organisasi sering kali menghadapi resistensi. Banyak pegawai pemerintah yang

mungkin merasa tidak nyaman dengan perubahan ini dan lebih suka mempertahankan cara kerja yang lama. Pemerintah perlu mengembangkan strategi manajemen perubahan yang efektif untuk mengatasi resistensi ini dan memastikan bahwa semua pegawai dapat beradaptasi dengan perubahan.

3.7. Strategi untuk Transformasi Digital di Kalimantan Timur

Transformasi digital adalah proses yang mendalam dan luas yang mengubah cara organisasi dan masyarakat bekerja dan berinteraksi melalui teknologi digital. Di Kalimantan Timur, transformasi digital merupakan langkah penting dalam upaya mempercepat pembangunan daerah dan meningkatkan kualitas hidup masyarakat. Strategi ini dirancang untuk memberikan panduan yang komprehensif dan holistik, mencakup berbagai aspek mulai dari infrastruktur hingga sumber daya manusia, serta regulasi dan kebijakan yang mendukung.

1. Peningkatan Infrastruktur Digital

a) Pengembangan Jaringan Internet

- Pembangunan Jaringan Fiber Optik: Pembangunan jaringan fiber optik adalah langkah pertama dan utama dalam meningkatkan infrastruktur digital di Kalimantan Timur. Fiber optik memberikan kecepatan dan kapasitas yang lebih tinggi dibandingkan dengan teknologi lainnya, sehingga memungkinkan akses internet yang cepat dan stabil. Proyek ini mencakup pemasangan kabel fiber optik yang menjangkau seluruh wilayah, termasuk daerah pedesaan dan terpencil. Ini memerlukan investasi besar dan kerjasama antara pemerintah pusat, pemerintah daerah, dan sektor swasta.
- Penguatan Jaringan 5G: Teknologi 5G menawarkan kecepatan data yang sangat tinggi, latensi rendah, dan kapasitas besar, yang sangat penting untuk mendukung aplikasi-aplikasi canggih seperti Internet of Things (IoT), smart city, dan augmented reality. Implementasi 5G di Kalimantan Timur akan dilakukan secara bertahap, dimulai dari kota-kota besar seperti Samarinda dan Balikpapan, kemudian diperluas ke daerah lainnya. Pemerintah daerah perlu bekerja sama dengan operator telekomunikasi untuk memastikan pelaksanaan yang lancar dan tepat waktu.

b) Peningkatan Aksesibilitas

- WiFi Publik: Menyediakan akses WiFi gratis di area publik adalah langkah yang efektif untuk meningkatkan aksesibilitas internet bagi masyarakat. Area publik seperti taman kota, sekolah, rumah sakit, dan kantor pemerintahan akan dilengkapi dengan hotspot WiFi. Inisiatif ini tidak hanya meningkatkan akses internet, tetapi juga mendukung pendidikan dan pelayanan publik.
- Subsidi Internet: Untuk memastikan bahwa seluruh lapisan masyarakat dapat menikmati manfaat transformasi digital, pemerintah daerah dapat memberikan subsidi untuk akses internet bagi keluarga berpenghasilan rendah. Program ini dapat diintegrasikan dengan program bantuan sosial lainnya untuk mencapai target yang tepat.

2. Peningkatan Peningkatan Kapasitas Sumber Daya Manusia (SDM)

a) Pelatihan dan Sertifikasi

- Program Pelatihan Digital: Mengadakan program pelatihan rutin yang mencakup keterampilan dasar hingga lanjutan dalam teknologi informasi dan komunikasi. Program ini dapat diselenggarakan oleh pemerintah daerah bekerja sama dengan lembaga pendidikan dan perusahaan teknologi. Topik pelatihan dapat mencakup pemrograman, analisis data, keamanan siber, dan penggunaan aplikasi digital.
- Sertifikasi Profesi: Untuk meningkatkan daya saing tenaga kerja, pemerintah daerah dapat bekerja sama dengan lembaga sertifikasi untuk memberikan akreditasi kepada individu yang memiliki keterampilan di bidang teknologi. Sertifikasi ini akan menjadi nilai tambah bagi tenaga kerja lokal dan memudahkan mereka untuk mendapatkan pekerjaan di sektor teknologi.

b) Pendidikan

- Kurikulum Berbasis Digital: Integrasi mata pelajaran teknologi informasi dan komunikasi (TIK) dalam kurikulum sekolah adalah langkah penting untuk mempersiapkan generasi muda menghadapi era digital. Kurikulum ini harus mencakup teori dan praktik, serta dilengkapi dengan fasilitas dan alat bantu pengajaran yang memadai.
- Beasiswa: Pemerintah daerah dapat menyediakan beasiswa untuk studi lanjut di bidang teknologi informasi dan data science. Beasiswa ini dapat diberikan kepada siswa

berprestasi yang ingin melanjutkan pendidikan ke jenjang yang lebih tinggi, baik di dalam negeri maupun luar negeri. Selain itu, program beasiswa juga dapat mencakup pelatihan dan sertifikasi untuk meningkatkan keterampilan profesional.

3. Pengembangan Ekosistem Teknologi

a) Startup dan Inovasi

- Inkubator Bisnis: Membentuk inkubator bisnis untuk mendukung pertumbuhan startup teknologi lokal. Inkubator ini akan menyediakan fasilitas, mentor, dan akses ke jaringan investor untuk membantu startup berkembang. Selain itu, inkubator dapat mengadakan kompetisi dan hackathon untuk mendorong inovasi di kalangan pemuda dan pengusaha.
- Pendanaan dan Investasi: Pemerintah daerah dapat menyediakan akses pendanaan dan investasi bagi startup di bidang teknologi. Ini dapat dilakukan melalui dana ventura daerah atau kerjasama dengan investor swasta. Dukungan finansial ini akan membantu startup dalam tahap awal pengembangan dan mempercepat pertumbuhan mereka.

b) Kolaborasi dengan Industri

- Kemitraan dengan Perusahaan Teknologi: Membangun kemitraan strategis dengan perusahaan teknologi besar untuk transfer pengetahuan dan teknologi. Kemitraan ini dapat mencakup program pelatihan, proyek penelitian bersama, dan implementasi teknologi baru. Pemerintah daerah perlu aktif mencari peluang kerjasama dengan perusahaan teknologi nasional dan internasional.
- Program Magang: Menyediakan program magang di perusahaan teknologi untuk mahasiswa dan lulusan baru. Program ini akan memberikan pengalaman praktis dan kesempatan untuk belajar langsung dari profesional di industri. Magang juga dapat menjadi jalan bagi peserta untuk mendapatkan pekerjaan tetap di perusahaan tersebut.

4. Transformasi Layanan Publik

a) Digitalisasi Layanan

- e-Government: Mengembangkan platform e-Government untuk mempermudah akses dan layanan publik. Platform ini harus mencakup berbagai layanan seperti pendaftaran penduduk, pembayaran pajak, perizinan, dan pelayanan kesehatan.

Digitalisasi layanan publik akan meningkatkan efisiensi, transparansi, dan akuntabilitas pemerintah daerah.

- One-Stop Service: Membangun pusat layanan terpadu yang mengintegrasikan berbagai layanan publik dalam satu platform. Ini akan memudahkan masyarakat dalam mengakses layanan yang dibutuhkan tanpa harus berpindah dari satu instansi ke instansi lainnya. Pusat layanan terpadu ini dapat berbasis online dan offline untuk menjangkau seluruh lapisan masyarakat.

b) **Transparansi dan Akuntabilitas**

- Open Data: Mengimplementasikan kebijakan open data untuk meningkatkan transparansi pemerintahan. Data yang terbuka dan mudah diakses oleh publik akan memungkinkan masyarakat untuk memantau kinerja pemerintah dan ikut serta dalam proses pengambilan keputusan. Pemerintah daerah perlu membangun portal data terbuka yang menyajikan data secara terstruktur dan mudah dipahami.
- Monitoring dan Evaluasi: Menggunakan teknologi untuk monitoring dan evaluasi kinerja layanan publik. Sistem ini dapat mencakup dashboard analitik yang memberikan informasi real-time tentang kinerja berbagai layanan publik. Monitoring dan evaluasi yang berbasis data akan membantu pemerintah daerah dalam mengambil keputusan yang tepat dan cepat.

5. Pengembangan *Big Data* dan *Data Analytics*

a) **Infrastruktur Data**

- Pusat Data Regional: Membangun pusat data yang terintegrasi untuk mengelola data dari berbagai instansi pemerintah. Pusat data ini akan menyimpan dan mengolah data dalam jumlah besar dengan aman dan efisien. Selain itu, pusat data harus dilengkapi dengan teknologi terkini untuk mendukung analisis data yang kompleks.
- Keamanan Data: Memastikan keamanan data dengan mengimplementasikan standar keamanan yang tinggi. Ini mencakup enkripsi data, proteksi terhadap serangan siber, dan kebijakan privasi yang ketat. Pemerintah daerah perlu bekerja sama dengan ahli keamanan siber untuk melindungi data publik dan mencegah kebocoran data.

b) Analisis Data

- Penggunaan AI dan Machine Learning: Menggunakan teknologi AI dan machine learning untuk analisis data yang lebih efektif. Teknologi ini dapat membantu dalam pengambilan keputusan yang lebih cepat dan akurat berdasarkan data. Misalnya, analisis prediktif dapat digunakan untuk meramalkan kebutuhan layanan kesehatan atau pendidikan di masa depan.
- Dashboard Analitik: Mengembangkan dashboard analitik untuk memantau dan menganalisis kinerja pemerintahan dan pelayanan publik. Dashboard ini akan menyajikan data dalam bentuk visual yang mudah dipahami dan dapat diakses oleh pejabat pemerintah dan masyarakat. Dashboard analitik akan menjadi alat yang penting untuk meningkatkan transparansi dan akuntabilitas.

6. Kebijakan dan Regulasi

a) Regulasi yang Mendukung

- Peraturan Pemerintah: Menyusun regulasi yang mendukung transformasi digital dan inovasi teknologi. Regulasi ini harus mencakup berbagai aspek seperti perlindungan data pribadi, hak kekayaan intelektual, dan keamanan siber. Pemerintah daerah perlu bekerja sama dengan legislatif dan pemangku kepentingan lainnya untuk menyusun regulasi yang komprehensif dan fleksibel.
- Perlindungan Data Pribadi: Mengimplementasikan kebijakan perlindungan data pribadi untuk menjaga privasi masyarakat. Ini mencakup regulasi yang ketat tentang pengumpulan, penyimpanan, dan penggunaan data pribadi. Pemerintah daerah harus memastikan bahwa seluruh instansi pemerintah dan perusahaan swasta mematuhi kebijakan ini.

b) Insentif dan Dukungan

- Insentif Pajak: Memberikan insentif pajak bagi perusahaan yang berinvestasi dalam teknologi. Insentif ini dapat berupa pengurangan pajak, pembebasan bea masuk untuk peralatan teknologi, atau dukungan finansial untuk penelitian dan pengembangan. Tujuan dari insentif ini adalah untuk mendorong perusahaan untuk berinvestasi lebih banyak dalam inovasi teknologi.

- Dukungan R&D: Meningkatkan dukungan untuk penelitian dan pengembangan di bidang teknologi. Pemerintah daerah dapat menyediakan dana penelitian, fasilitas, dan program kemitraan dengan universitas dan lembaga penelitian. Dukungan ini akan membantu dalam menciptakan inovasi baru dan meningkatkan daya saing daerah dalam bidang teknologi.

7. Implementasi dan Monitoring

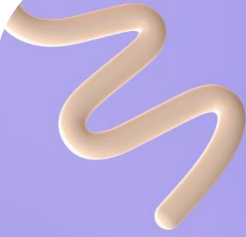
a) Roadmap Jangka Panjang

Menyusun roadmap jangka panjang untuk transformasi digital dengan tahapan yang jelas. Roadmap ini harus mencakup visi, misi, dan tujuan yang spesifik, serta langkah-langkah yang diperlukan untuk mencapainya. Setiap tahapan harus dilengkapi dengan indikator kinerja yang jelas untuk memantau kemajuan.

b) Komite Pengawas

Membentuk komite pengawas yang bertugas memantau implementasi dan memastikan pencapaian target. Komite ini harus terdiri dari perwakilan pemerintah, sektor swasta, akademisi, dan masyarakat. Tugas utama komite pengawas adalah mengawasi pelaksanaan strategi, mengevaluasi kinerja, dan memberikan rekomendasi perbaikan.

Transformasi digital di Kalimantan Timur memerlukan sinergi antara berbagai pihak, mulai dari pemerintah, sektor swasta, hingga masyarakat. Dengan strategi yang tepat, Kalimantan Timur dapat menjadi pionir dalam transformasi digital di Indonesia, meningkatkan efisiensi, transparansi, dan kualitas hidup masyarakat. Implementasi yang efektif dari strategi ini akan membawa manfaat jangka panjang yang signifikan, termasuk peningkatan ekonomi, kemajuan teknologi, dan pembangunan sosial yang inklusif. Transformasi digital adalah perjalanan panjang yang memerlukan komitmen dan kerja keras, namun dengan visi yang jelas dan tindakan yang terkoordinasi, Kalimantan Timur dapat mencapai tujuannya dan menjadi contoh bagi daerah lain di Indonesia



Portal Satu Data adalah cahaya di tengah kabut informasi. Dengan sains data, kita dapat melihat lebih jauh dan lebih jelas menuju masa depan

Dengan data yang akurat dan analisis yang cermat, Portal Satu Data menjadikan setiap keputusan sebagai langkah pasti menuju kemajuan



Di balik setiap angka terdapat cerita. Melalui sains data di Portal Satu Data, kita menggali makna dan merangkai cerita untuk masa depan yang lebih baik

Transformasi digital bukan hanya tentang teknologi, tetapi juga tentang bagaimana kita memahami dan menggunakan data. Portal Satu Data dan sains data adalah kunci keberhasilan



BAB 4

IMPLEMENTASI *DATA SCIENCE*

PADA SATU DATA KALIMANTAN TIMUR

Satu Data Kalimantan Timur adalah inisiatif untuk mengintegrasikan dan mengelola data dari berbagai instansi pemerintah di Provinsi Kalimantan Timur, sehingga dapat digunakan untuk mendukung pengambilan keputusan yang lebih baik, meningkatkan efisiensi operasional, dan memberikan layanan publik yang lebih baik. Implementasi *data science* dalam inisiatif ini dapat membawa banyak manfaat, mulai dari analisis data yang lebih mendalam hingga pemanfaatan teknologi canggih seperti *machine learning* dan kecerdasan buatan (*artificial intelligence/AI*). Kajian ini akan membahas implementasi *data science* pada satu data Kalimantan Timur dalam berbagai bidang.

4.1. Sektor Sosial Ekonomi

Dengan *data science*, pengambilan keputusan menjadi lebih tepat karena didasarkan pada analisis data yang akurat. Hal ini memungkinkan kebijakan yang lebih efektif untuk mengatasi masalah seperti pengangguran, kemiskinan, dan ketimpangan ekonomi. Selain itu, *data science* membantu dalam optimalisasi sumber daya dengan menentukan cara terbaik untuk memanfaatkan potensi wilayah, seperti lokasi pembangunan infrastruktur dan distribusi sumber daya alam. Pemantauan dan evaluasi program juga menjadi lebih efisien, memastikan bahwa program sosial dan ekonomi berjalan sesuai rencana dan memberikan hasil yang diharapkan.

Dalam pengembangan ekonomi, *data science* membuka peluang baru dengan mengidentifikasi tren dan pasar yang potensial, membantu diversifikasi ekonomi dari ketergantungan pada sektor pertambangan dan energi ke sektor lain seperti pertanian dan pariwisata. Dengan demikian, *data science* menjadi kunci transformasi sosial ekonomi Kalimantan Timur menuju masyarakat yang lebih sejahtera dan berkelanjutan.

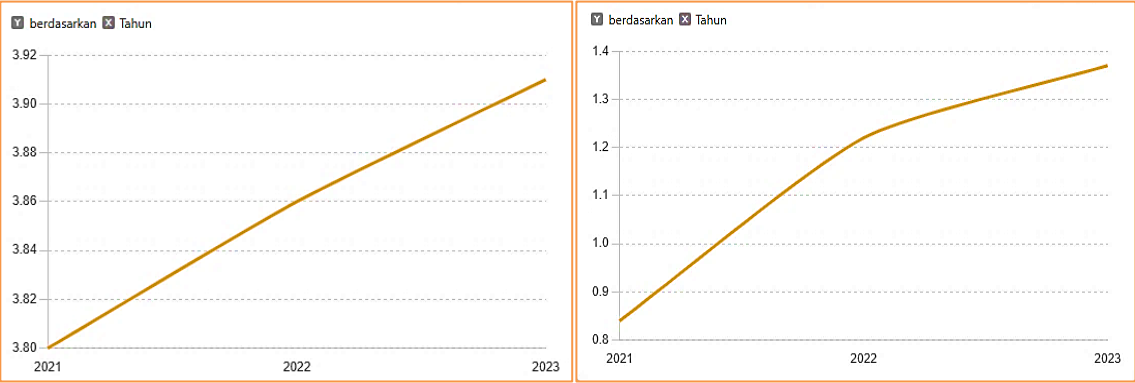
Berikut penerapan *data science* pada data sosial ekonomi Provinsi Kalimantan timur periode 2001 sampai dengan 2023:

Tabel 2. Sosial Ekonomi Provinsi Kalimantan Timur Tahun 2021–2023

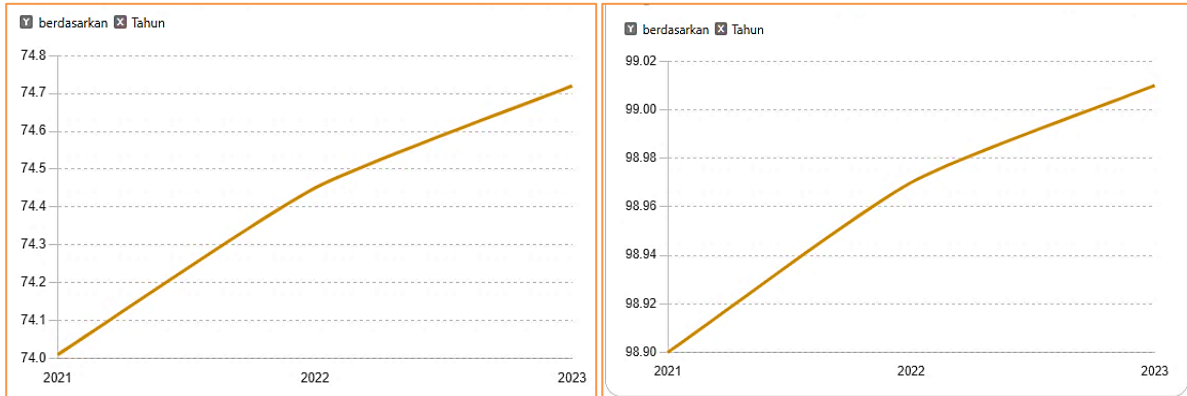
Rincian Data	Tahun		
	2021	2022	2023
Penduduk (Juta)	3,8	3,86	3,91
Laju Pertumbuhan Penduduk (%)	0,84	1,22	1,37
Angka Harapan Hidup (Tahun)	74	74,5	74,7
Angka Melek Huruf Usia 15+ (%)	98,9	99	99
Tingkat Partisipasi Angkatan Kerja -TPAK (%)	65,5	64,7	65,6
Tingkat Pengangguran Terbuka-TPT2 (%)	6,83	5,71	5,31
Penduduk Miskin (Ribu)	242	236	231
Persentase Penduduk Miskin (%)	6,54	6,31	6,11
Indeks Pembangunan Manusia-IPM	76,6	77,4	78,2
Produk Domestik Regional Bruto (PDRB) Harga Berlaku (Triliun Rupiah)	697	922	844
Laju Pertumbuhan Ekonomi (%)	2,55	4,48	6,22
PDRB Per Kapita Harga Berlaku (Juta Rupiah)	183	239	216
Inflasi (y-o-y) (%)	2,15	5,35	3,46

Sumber: Provinsi Kalimantan Timur Dalam Angka Tahun 2024, BPS Kaltim

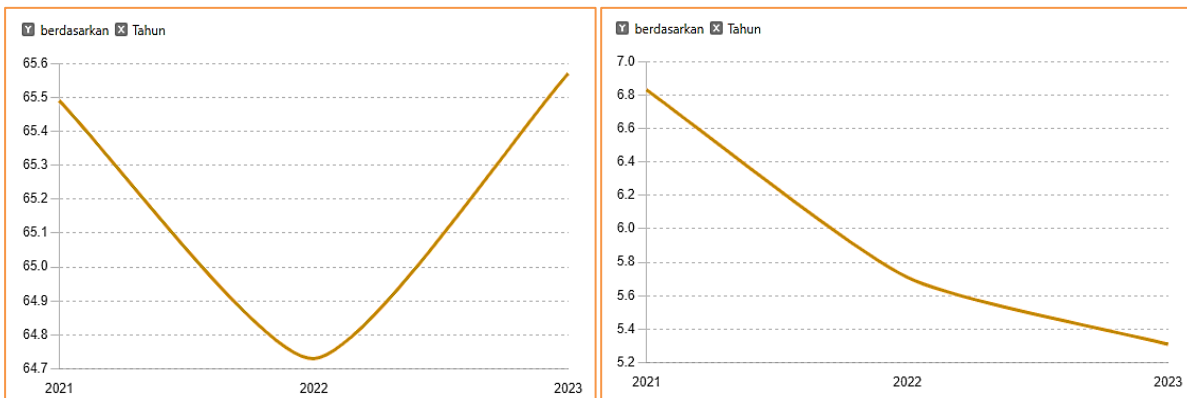
Dengan memvisualisasikan masing-masing variabel secara terpisah setiap variabel dengan menggunakan plot yang lebih besar dan detail serta agar lebih mudah dipahami.



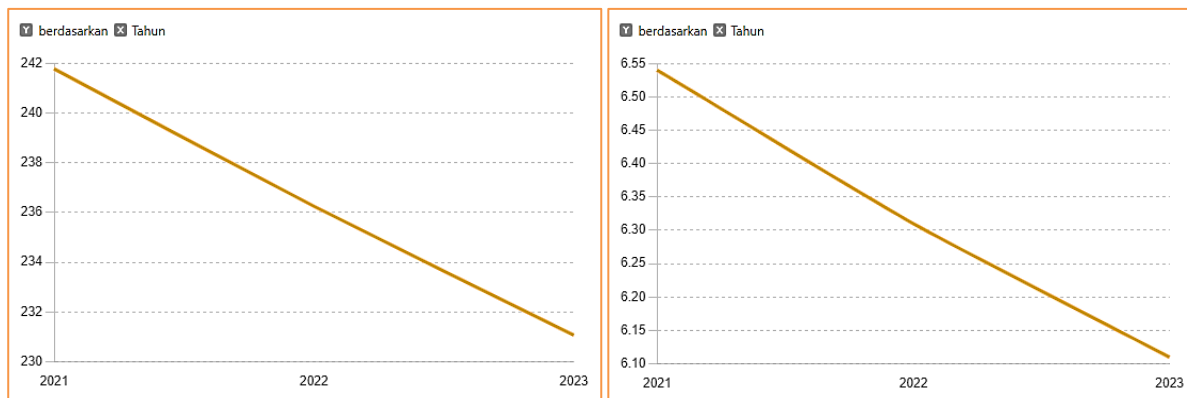
(a) Penambahan Penduduk (b) Laju Pertumbuhan Penduduk
Gambar 14. Grafik Trend Penambahan Penduduk dan Laju Pertumbuhan Penduduk Tahun 2021-2023



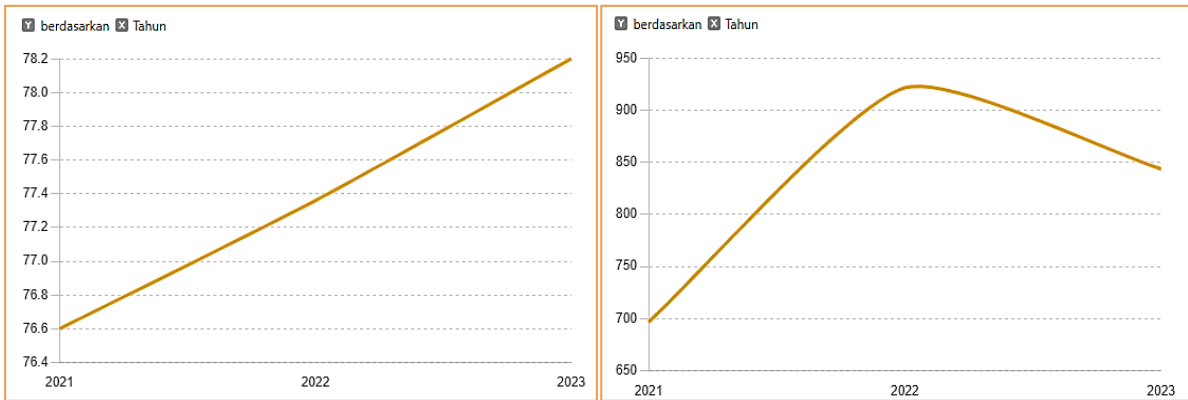
(a) Angka Harapan Hidup (b) Angka Melek Huruf Usia 15+ Tahun
Gambar 15. Grafik Tren Angka Harapan Hidup dan Angka Melek Huruf Usia 15+ Tahun 2021–2023



(a) Tingkat Partisipasi Angkatan Kerja (b) Tingkat Pengangguran Terbuka
Gambar 16. Grafik Tren Tingkat Partisipasi Angkatan Kerja dan Tingkat Pengangguran Terbuka Tahun 2021-2023



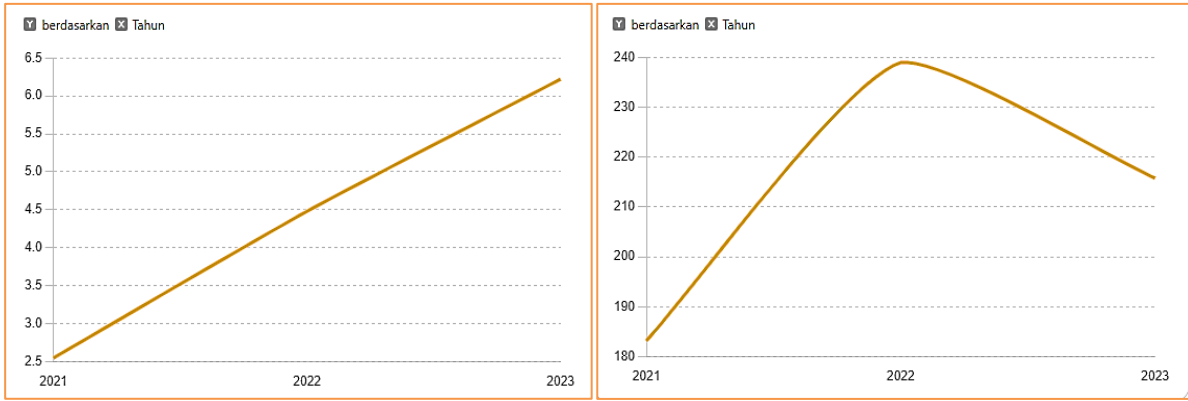
(a) Penduduk Miskin (Ribu) (b) Tingkat Pengangguran Terbuka
Gambar 17. Grafik Tren Tingkat Penduduk Miskin dan Persentase Penduduk Miskin Tahun 2021-2023



(a) Indeks Pembangunan Manusia

(b) PDRB Harga Berlaku (persen)

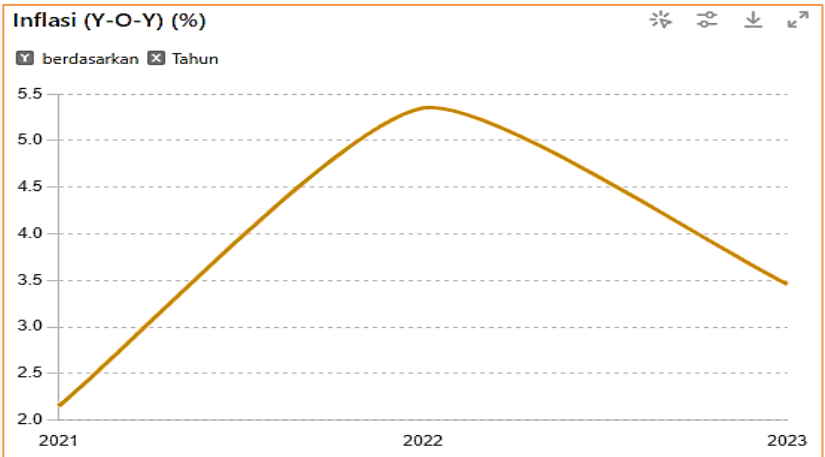
Gambar 18. Grafik Tren Indeks Pembangunan Manusia dan PDRB Harga Berlaku Tahun 2021-2023



(a) Laju Pertumbuhan Ekonomi (%)

(b) PDRB Per Kapita Harga Berlaku (Juta)

Gambar 19. Grafik Tren Indeks Pembangunan Manusia dan PDRB Harga Berlaku Tahun 2021-2023



Gambar 20. Grafik Tren Inflasi (y-o-y) (%) Tahun 2021-2023

Pengaruh ekonomi terhadap kemiskinan, kita bisa melihat hubungan antara variabel ekonomi (seperti laju pertumbuhan ekonomi dan PDRB) dengan variabel kemiskinan (penduduk miskin dan persentase penduduk miskin). Mari kita lakukan analisis korelasi dan visualisasi data untuk melihat hubungan ini.

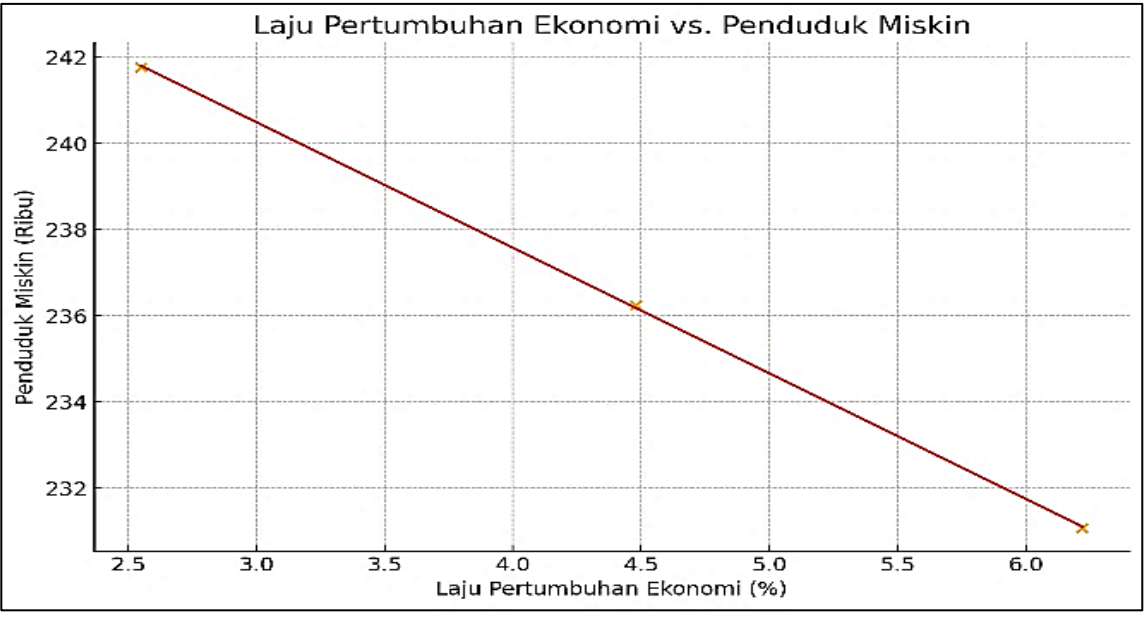
Variabel Analisis Korelasi:

1. Laju Pertumbuhan Ekonomi vs. Penduduk Miskin
2. Laju Pertumbuhan Ekonomi vs. Persentase Penduduk Miskin
3. PDRB Per Kapita vs. Penduduk Miskin
4. PDRB Per Kapita vs. Persentase Penduduk Miskin

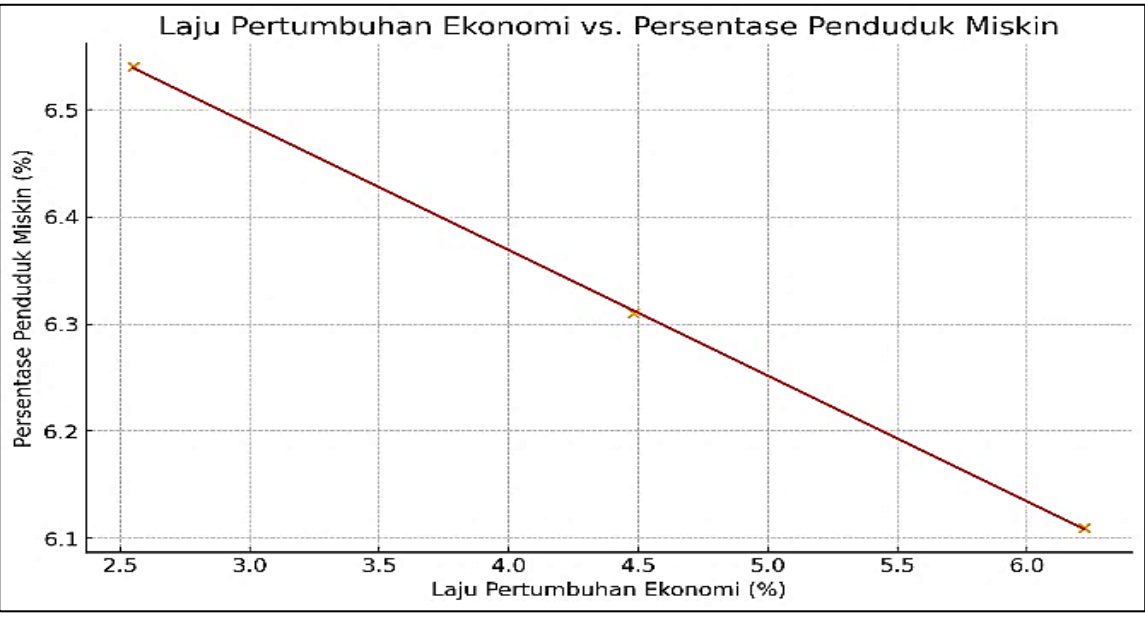
Berikut ini kode python untuk proses menghitung korelasi antara variabel-variabel tersebut diatas dan menampilkan visualisasinya.

```
# Calculate correlation
correlation_data = {
    "Laju Pertumbuhan Ekonomi vs. Penduduk Miskin": df["Laju Pertumbuhan Ekonomi (%)"].corr(df["Penduduk Miskin (Ribu)"]),
    "Laju Pertumbuhan Ekonomi vs. Persentase Penduduk Miskin": df["Laju Pertumbuhan Ekonomi (%)"].corr(df["Persentase Penduduk Miskin (%)"]),
    "PDRB Per Kapita vs. Penduduk Miskin": df["PDRB Per Kapita Harga Berlaku (Juta Rupiah)"].corr(df["Penduduk Miskin (Ribu)"]),
    "PDRB Per Kapita vs. Persentase Penduduk Miskin": df["PDRB Per Kapita Harga Berlaku (Juta Rupiah)"].corr(df["Persentase Penduduk Miskin (%)"])
}
correlation_df = pd.DataFrame.from_dict(correlation_data, orient='index', columns=["Correlation"])
# Plotting the correlations
def scatter_plot_with_regression(x, y, xlabel, ylabel, title):
    plt.figure(figsize=(10, 6))
    plt.scatter(x, y)
    m, b = np.polyfit(x, y, 1)
    plt.plot(x, m*x + b, color='red')
    plt.xlabel(xlabel)
    plt.ylabel(ylabel)
    plt.title(title)
    plt.grid(True)
    plt.show()
# Visualize the relationships
scatter_plot_with_regression(df["Laju Pertumbuhan Ekonomi (%)"], df["Penduduk Miskin (Ribu)"], "Laju Pertumbuhan Ekonomi (%)", "Penduduk Miskin (Ribu)", "Laju Pertumbuhan Ekonomi vs. Penduduk Miskin")
scatter_plot_with_regression(df["Laju Pertumbuhan Ekonomi (%)"], df["Persentase Penduduk Miskin (%)"], "Laju Pertumbuhan Ekonomi (%)", "Persentase Penduduk Miskin (%)", "Laju Pertumbuhan Ekonomi vs. Persentase Penduduk Miskin")
scatter_plot_with_regression(df["PDRB Per Kapita Harga Berlaku (Juta Rupiah)"], df["Penduduk Miskin (Ribu)"], "PDRB Per Kapita Harga Berlaku (Juta Rupiah)", "Penduduk Miskin (Ribu)", "PDRB Per Kapita vs. Penduduk Miskin")
```

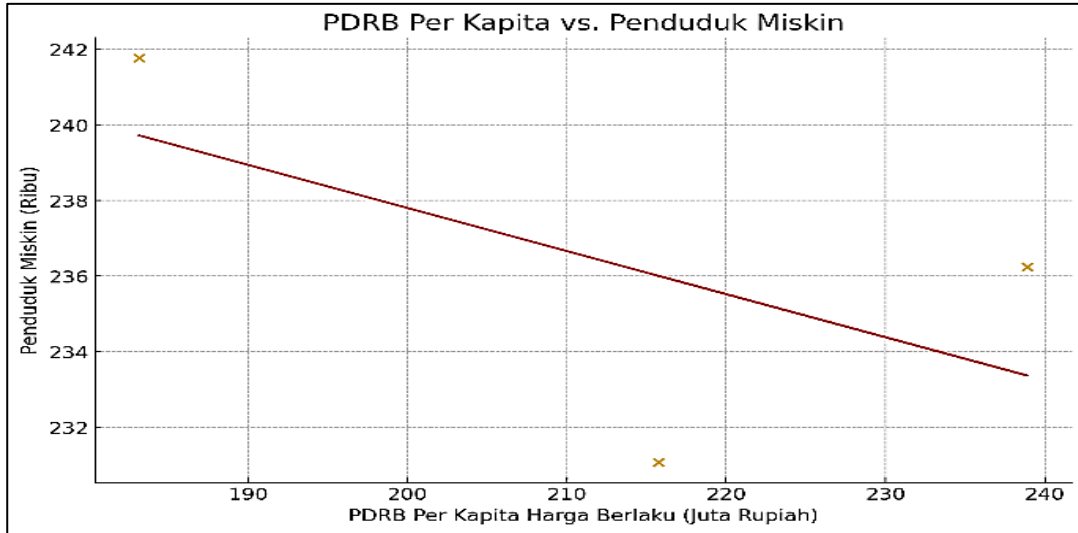
```
scatter_plot_with_regression(df["PDRB Per Kapita Harga Berlaku (Juta Rupiah)"], df["Persentase Penduduk Miskin (%)"],
"PDRB Per Kapita Harga Berlaku (Juta Rupiah)", "Persentase Penduduk Miskin (%)", "PDRB Per Kapita vs. Persentase Penduduk Miskin")
import ace_tools as tools; tools.display_dataframe_to_user(name="Correlation Analysis Data", dataframe=correlation_df)
correlation_df
```



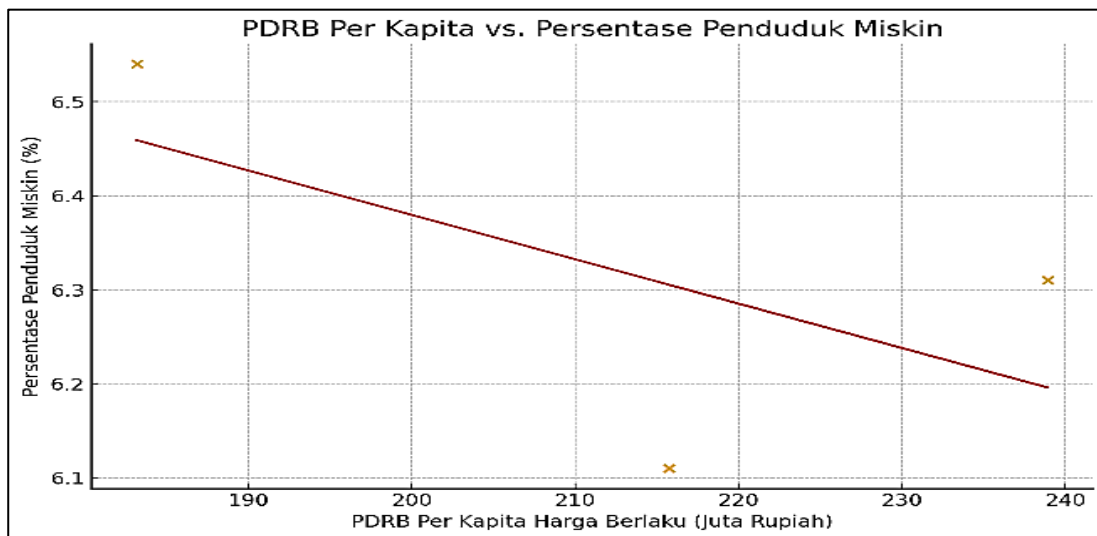
Gambar 21. Grafik Analisis Korelasi Laju Pertumbuhan Ekonomi vs. Penduduk Miskin



Gambar 22. Grafik Analisis Korelasi Laju Pertumbuhan Ekonomi vs. Persentase Penduduk Miskin



Gambar 23. Grafik Analisis Korelasi PDRB Perkapita vs. Penduduk Miskin



Gambar 24. Grafik Analisis Korelasi PDRB Perkapita vs. Persentase Penduduk Miskin

Analisis Korelasi IPM dengan variabel ekonomi dan sosial. Berdasarkan analisis korelasi dan visualisasi scatter plot, berikut adalah hasilnya.

Tabel 3. IPM Correlation Analysis Data

Item	Correlation
Laju Pertumbuhan Ekonomi (%)	0,998275262830
PDRB Per Kapita Harga Berlaku (Juta Rupiah)	0,558165702944
Angka Harapan Hidup (Tahun)	0,986215963237
Angka Melek Huruf Usia 15+ (%)	0,982929559902

Item	Correlation
Tingkat Partisipasi Angkatan Kerja - TPAK (%)	0,114991558537
Tingkat Pengangguran Terbuka - TPT2 (%)	-0,956565274775
Penduduk Miskin (Ribuan)	-0,998886138137
Persentase Penduduk Miskin (%)	-0,997612302770

Korelasi IPM dengan Variabel Ekonomi:

1. Laju Pertumbuhan Ekonomi (%):

- a. Korelasi: 0,998275
- b. Terdapat korelasi positif yang sangat kuat, menunjukkan bahwa peningkatan laju pertumbuhan ekonomi berkorelasi erat dengan peningkatan IPM.

2. PDRB Per Kapita Harga Berlaku (Juta Rupiah):

- a. Korelasi: 0,558166
- b. Terdapat korelasi positif yang moderat, menunjukkan bahwa peningkatan PDRB per kapita berkorelasi dengan peningkatan IPM, meskipun tidak sekuat laju pertumbuhan ekonomi.

Korelasi IPM dengan Variabel Sosial:

1. Angka Harapan Hidup (Tahun):

- a. Korelasi: 0,986216
- b. Terdapat korelasi positif yang sangat kuat, menunjukkan bahwa peningkatan angka harapan hidup berkorelasi erat dengan peningkatan IPM.

2. Angka Melek Huruf Usia 15+ (%):

- a. Korelasi: 0,982930
- b. Terdapat korelasi positif yang sangat kuat, menunjukkan bahwa peningkatan angka melek huruf berkorelasi erat dengan peningkatan IPM.

3. Tingkat Partisipasi Angkatan Kerja - TPAK (%):

- a. Korelasi: 0,114992
- b. Terdapat korelasi positif yang sangat lemah, menunjukkan bahwa peningkatan tingkat partisipasi angkatan kerja tidak memiliki korelasi yang signifikan dengan IPM.

4. Tingkat Pengangguran Terbuka - TPT2 (%):

- a. Korelasi: -0,956565
- b. Terdapat korelasi negatif yang sangat kuat, menunjukkan bahwa peningkatan tingkat pengangguran terbuka berkorelasi erat dengan penurunan IPM.

5. Penduduk Miskin (Ribu):

- a. Korelasi: -0,998886
- b. Terdapat korelasi negatif yang sangat kuat, menunjukkan bahwa peningkatan jumlah penduduk miskin berkorelasi erat dengan penurunan IPM.

6. Persentase Penduduk Miskin (%):

- a. Korelasi: -0,997612
- b. Terdapat korelasi negatif yang sangat kuat, menunjukkan bahwa peningkatan persentase penduduk miskin berkorelasi erat dengan penurunan IPM.

Kesimpulan:

- a) Laju Pertumbuhan Ekonomi dan Angka Harapan Hidup memiliki korelasi positif yang sangat kuat dengan IPM, menunjukkan bahwa peningkatan dalam faktor-faktor ini secara signifikan meningkatkan kualitas pembangunan manusia.
- b) Tingkat Pengangguran Terbuka dan Kemiskinan (baik jumlah maupun persentase) memiliki korelasi negatif yang sangat kuat dengan IPM, menunjukkan bahwa peningkatan dalam faktor-faktor ini secara signifikan mengurangi kualitas pembangunan manusia.

- c) Angka Melek Huruf juga memiliki korelasi positif yang sangat kuat dengan IPM, menegaskan pentingnya pendidikan dalam pembangunan manusia.
- d) PDRB Per Kapita memiliki korelasi positif moderat dengan IPM, menunjukkan bahwa peningkatan kesejahteraan ekonomi per individu berkorelasi dengan peningkatan IPM.
- e) Tingkat Partisipasi Angkatan Kerja memiliki korelasi yang sangat lemah dengan IPM, menunjukkan bahwa perubahan dalam partisipasi angkatan kerja tidak memiliki dampak yang signifikan terhadap IPM.

Secara keseluruhan, data ini menekankan pentingnya pertumbuhan ekonomi, peningkatan harapan hidup, dan pengurangan kemiskinan dalam upaya meningkatkan Indeks Pembangunan Manusia.

4.2. Sektor Kesehatan

Pada sektor kesehatan di Provinsi Kalimantan Timur, akan memproses analisis data untuk mendapatkan tren dan proyeksi kedepan yakni berupa data kesehatan Masyarakat Provinsi Kalimantan Timur dari tahun 2016 sampai dengan 2020 dan data wanita berumur 15-49 Tahun yang berstatus kawin dan sedang menggunakan KB menurut Kabupaten/Kota di Provinsi Kalimantan Timur 2019-2023.

4.2.1. Kesehatan Masyarakat Provinsi Kaltim

Berikut ini adalah data kesehatan masyarakat Provinsi Kalimantan Timur dari tahun 2016 sampai dengan 2020 sebagai berikut.

Tabel 4. Data Kesehatan Masyarakat Kalimantan Timur Tahun 2016-2020

No.	Data Kesehatan	Tahun				
		2016	2017	2018	2019	2020
1	Jumlah kasus gizi Buruk yang ditemukan dan mendapat perawatan (satuan: anak)	354	293	212	227	137
2	Jumlah Balita (satuan: Balita)	349038	349543	349792	349803	349644

No.	Data Kesehatan	Tahun				
		2016	2017	2018	2019	2020
3	Jumlah kasus gizi Buruk Balita yang ditemukan dan mendapat perawatan (satuan: Balita)	354	293	212	227	137
4	Jumlah Kunjungan Ibu Hamil K4 (satuan: Kali)	69641	68913	70351	69746	63890
5	Persentase cakupan kunjungan Ibu Hamil K4 (satuan: %)	88,10	85,42	85,38	89,29	84,33
6	Cakupan Balita Gizi Kurang mendapat makanan tambahan (satuan: %)	37,75	56,17	93,94	97,58	95,97
7	Jumlah pelayanan kepada Ibu nifas (satuan: orang)	61220	62970	65649	64946	63067
8	Jumlah neonatus dengan komplikasi yang ditangani (satuan: Orang)	7817	7264	7271	6804	6748
9	Jumlah Penderita Gangguan Jiwa (satuan: Orang)	3307	4264	8296	8410	4154
10	Persentase Gizi Balita (Gizi Baik) (satuan: %)	0,10	0,08	0,06	0,06	0,04
11	Persentase Balita Gizi Buruk (satuan: %)	3,00	4,00	3,00	3,00	2,00
12	Jumlah Anak Balita (1-5 Tahun) (satuan: Balita)	278289	278754	279056	279182	279144
13	Penjaringan Kesehatan Siswa Sd Dan Setingkat (satuan: Orang)	54645	53138	95178	76408	65394
14	Jumlah Pelayanan Gawat Darurat Level 1 Yang Harus	44	44	54	53	57

No.	Data Kesehatan	Tahun				
		2016	2017	2018	2019	2020
	Diberikan Sarana Kesehatan (satuan: Unit)					

Sumber: <https://data.kaltimprov.go.id/dataset/data-kesehatan-masyarakat-provinsi-kaltim-tahun-2016-2020/>

Selanjutnya, kode program python dibawah ini akan memproses dan melakukan analisis data serta membuat grafik visualisasi sebagai berikut.

```
import matplotlib.pyplot as plt

# Data
years = [2016, 2017, 2018, 2019, 2020]
gizi_buruk_balita = [354, 293, 212, 227, 137]
cakupan_makanan_tambahan = [37.75, 56.17, 93.94, 97.58, 95.97]
penderita_gangguan_jiwa = [3307, 4264, 8296, 8410, 4154]

# Plotting the data
plt.figure(figsize=(14, 10))

# Plot 1: Jumlah Kasus Gizi Buruk pada Balita
plt.subplot(3, 1, 1)
plt.plot(years, gizi_buruk_balita, marker='o', linestyle='-', color='r')
plt.title('Jumlah Kasus Gizi Buruk pada Balita')
plt.xlabel('Tahun')
plt.ylabel('Jumlah Kasus')
plt.grid(True)

# Plot 2: Cakupan Balita Gizi Kurang Mendapat Makanan Tambahan
plt.subplot(3, 1, 2)
plt.plot(years, cakupan_makanan_tambahan, marker='o', linestyle='-', color='b')
plt.title('Cakupan Balita Gizi Kurang Mendapat Makanan Tambahan')
plt.xlabel('Tahun')
plt.ylabel('Persentase (%)')
plt.grid(True)

# Plot 3: Jumlah Penderita Gangguan Jiwa
plt.subplot(3, 1, 3)
plt.plot(years, penderita_gangguan_jiwa, marker='o', linestyle='-', color='g')
plt.title('Jumlah Penderita Gangguan Jiwa')
plt.xlabel('Tahun')
plt.ylabel('Jumlah Penderita')
plt.grid(True)
plt.tight_layout()
plt.show()

# Data for additional indicators
jumlah_balita = [349038, 349543, 349792, 349803, 349644]
jumlah_kunjungan_ibu_hamil = [69641, 68913, 70351, 69746, 63890]
cakupan_kunjungan_ibu_hamil = [88.10, 85.42, 85.38, 89.29, 84.33]
jumlah_pelayanan_ibu_nifas = [61220, 62970, 65649, 64946, 63067]
jumlah_neonatus_komplikasi = [7817, 7264, 7271, 6804, 6748]
persentase_gizi_baik = [0.10, 0.08, 0.06, 0.06, 0.04]
```

```

persentase_balita_gizi_buruk = [3.00, 4.00, 3.00, 3.00, 2.00]
jumlah_anak_balita = [278289, 278754, 279056, 279182, 279144]
penjaringan_kesehatan_siswa = [54645, 53138, 95178, 76408, 65394]
jumlah_pelayanan_gawat_darurat = [44, 44, 54, 53, 57]

# Plotting the additional data
plt.figure(figsize=(14, 18))

# Plot 1: Jumlah Balita
plt.subplot(5, 2, 1)
plt.plot(years, jumlah_balita, marker='o', linestyle='-', color='purple')
plt.title('Jumlah Balita')
plt.xlabel('Tahun')
plt.ylabel('Jumlah Balita')
plt.grid(True)

# Plot 2: Jumlah Kunjungan Ibu Hamil K4
plt.subplot(5, 2, 2)
plt.plot(years, jumlah_kunjungan_ibu_hamil, marker='o', linestyle='-', color='orange')
plt.title('Jumlah Kunjungan Ibu Hamil K4')
plt.xlabel('Tahun')
plt.ylabel('Jumlah Kunjungan')
plt.grid(True)

# Plot 3: Persentase Cakupan Kunjungan Ibu Hamil K4
plt.subplot(5, 2, 3)
plt.plot(years, cakupan_kunjungan_ibu_hamil, marker='o', linestyle='-', color='cyan')
plt.title('Persentase Cakupan Kunjungan Ibu Hamil K4')
plt.xlabel('Tahun')
plt.ylabel('Persentase (%)')
plt.grid(True)

# Plot 4: Jumlah Pelayanan Ibu Nifas
plt.subplot(5, 2, 4)
plt.plot(years, jumlah_pelayanan_ibu_nifas, marker='o', linestyle='-', color='magenta')
plt.title('Jumlah Pelayanan Ibu Nifas')
plt.xlabel('Tahun')
plt.ylabel('Jumlah Pelayanan')
plt.grid(True)

# Plot 5: Jumlah Neonatus dengan Komplikasi
plt.subplot(5, 2, 5)
plt.plot(years, jumlah_neonatus_komplikasi, marker='o', linestyle='-', color='brown')
plt.title('Jumlah Neonatus dengan Komplikasi')
plt.xlabel('Tahun')
plt.ylabel('Jumlah Neonatus')
plt.grid(True)

# Plot 6: Persentase Gizi Balita (Gizi Baik)
plt.subplot(5, 2, 6)
plt.plot(years, persentase_gizi_baik, marker='o', linestyle='-', color='lime')
plt.title('Persentase Gizi Balita (Gizi Baik)')
plt.xlabel('Tahun')
plt.ylabel('Persentase (%)')
plt.grid(True)

# Plot 7: Persentase Balita Gizi Buruk
plt.subplot(5, 2, 7)
plt.plot(years, persentase_balita_gizi_buruk, marker='o', linestyle='-', color='blue')
plt.title('Persentase Balita Gizi Buruk')

```

```
plt.xlabel('Tahun')
plt.ylabel('Persentase (%)')
plt.grid(True)

# Plot 8: Jumlah Anak Balita (1-5 Tahun)
plt.subplot(5, 2, 8)
plt.plot(years, jumlah_anak_balita, marker='o', linestyle='-', color='red')
plt.title('Jumlah Anak Balita (1-5 Tahun)')
plt.xlabel('Tahun')
plt.ylabel('Jumlah Balita')
plt.grid(True)

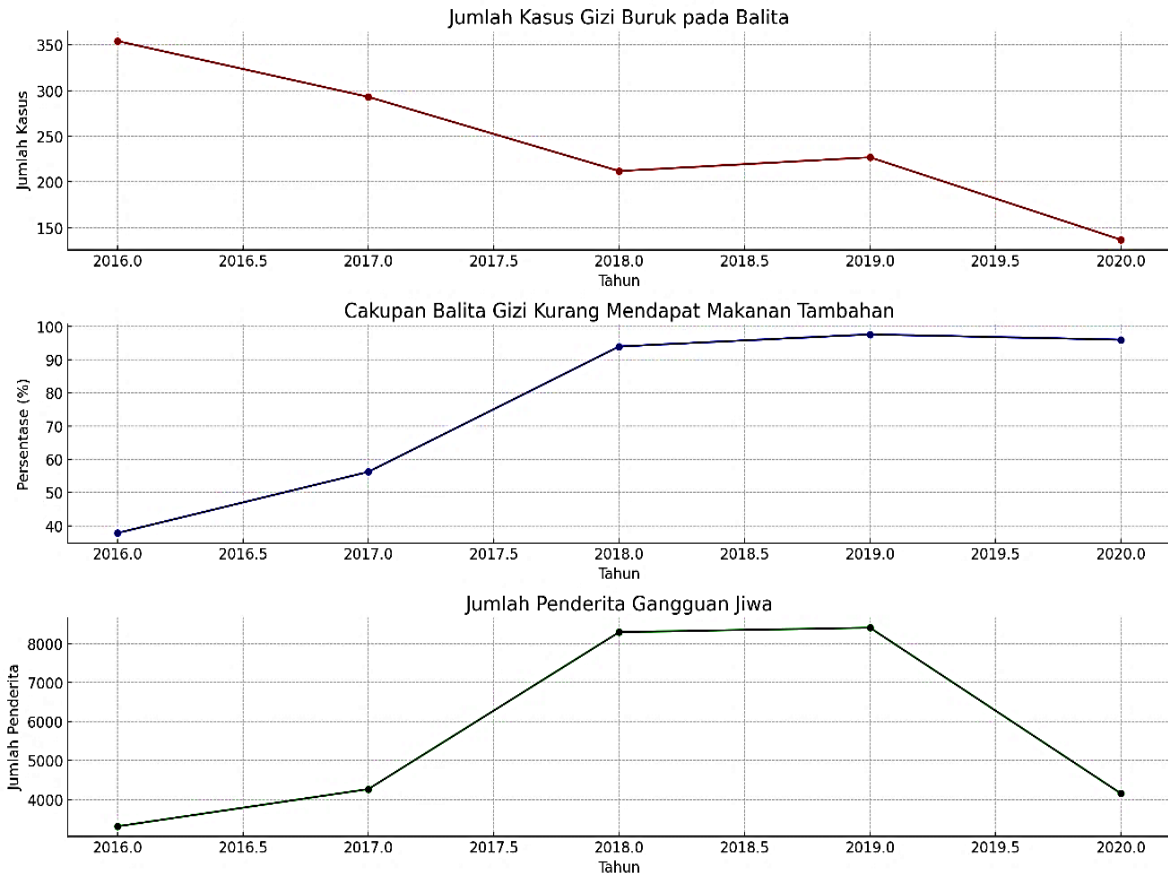
# Plot 9: Penjarangan Kesehatan Siswa SD dan Setingkat
plt.subplot(5, 2, 9)
plt.plot(years, penjarangan_kesehatan_siswa, marker='o', linestyle='-', color='darkgreen')
plt.title('Penjarangan Kesehatan Siswa SD dan Setingkat')
plt.xlabel('Tahun')
plt.ylabel('Jumlah Siswa')
plt.grid(True)

# Plot 10: Jumlah Pelayanan Gawat Darurat Level 1
plt.subplot(5, 2, 10)
plt.plot(years, jumlah_pelayanan_gawat_darurat, marker='o', linestyle='-', color='black')
plt.title('Jumlah Pelayanan Gawat Darurat Level 1')
plt.xlabel('Tahun')
plt.ylabel('Jumlah Pelayanan')
plt.grid(True)

plt.tight_layout()
plt.show()
```

Grafik pada gambar 25 memperlihatkan tren dari ketiga variabel tersebut dari tahun 2016 hingga tahun 2020 yaitu sebagai berikut:

- a) Grafik pertama menunjukkan penurunan jumlah kasus gizi buruk pada balita.
- b) Grafik kedua menunjukkan peningkatan cakupan balita yang mendapatkan makanan tambahan.
- c) Grafik ketiga menunjukkan fluktuasi jumlah penderita gangguan jiwa dengan puncak tertinggi pada tahun 2018.

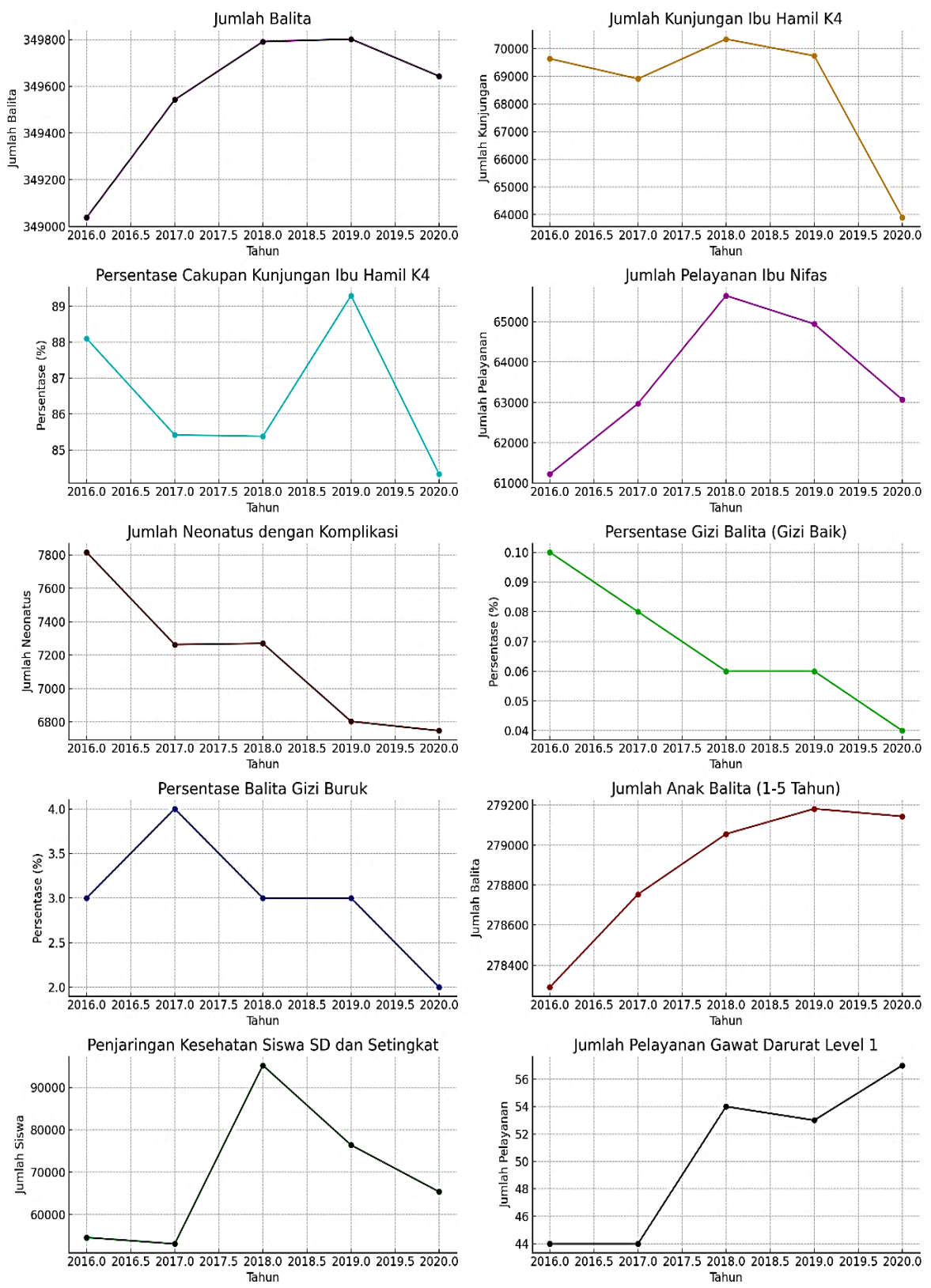


Gambar 25. Grafik Tren Kesehatan Masyarakat Kaltim Periode 2016-2020 (3 variabel)

Berikut adalah analisis lengkap dari data kesehatan masyarakat Provinsi Kalimantan Timur dari tahun 2016 hingga tahun 2020, yaitu:

1. Jumlah Kasus Gizi Buruk pada Balita: Terjadi penurunan signifikan dari 354 kasus pada tahun 2016 menjadi 137 kasus pada tahun 2020. Penurunan ini mencerminkan peningkatan dalam penanganan gizi buruk pada balita selama lima tahun terakhir.
2. Cakupan Balita Gizi Kurang Mendapat Makanan Tambahan: Persentase balita yang mendapatkan makanan tambahan meningkat secara drastis dari 37,75% pada tahun 2016 menjadi 95,97% pada tahun 2020. Hal ini menunjukkan adanya peningkatan dalam program intervensi gizi untuk balita.
3. Jumlah Penderita Gangguan Jiwa: Data menunjukkan fluktuasi dengan peningkatan tajam pada tahun 2018 sebesar 8296 penderita, namun kemudian menurun menjadi 4154 pada tahun 2020. Ini mungkin menunjukkan adanya peningkatan kesadaran dan layanan kesehatan mental selama periode tersebut.

Grafik pada gambar 26 memperlihatkan tren dari sepuluh variabel dari tahun 2016 hingga 2020 sebagai berikut:



Gambar 26. Grafik Tren Kesehatan Masyarakat Kaltim Periode 2016-2020

Berikut adalah penjelasan hasil analisis dari grafik pada gambar 26 yang merupakan tambahan dari indikator kesehatan lainnya.

1. Jumlah Balita: Populasi balita tetap stabil di sekitar 349.000 selama periode 2016-2020 dengan sedikit fluktuasi.
2. Jumlah Kunjungan Ibu Hamil K4: Mengalami penurunan pada tahun 2020 setelah mencapai puncak pada tahun 2018. Hal ini mungkin disebabkan oleh perubahan kebijakan atau program kesehatan ibu hamil.
3. Persentase Cakupan Kunjungan Ibu Hamil K4: Fluktuasi persentase cakupan kunjungan ibu hamil K4 menunjukkan adanya variasi dalam program intervensi ibu hamil setiap tahunnya.
4. Jumlah Pelayanan Ibu Nifas: Mengalami peningkatan hingga tahun 2018 dan kemudian menurun pada tahun-tahun berikutnya. Ini mungkin mencerminkan perubahan dalam kebutuhan atau ketersediaan layanan untuk ibu nifas.
5. Jumlah Neonatus dengan Komplikasi: Jumlah neonatus dengan komplikasi yang ditangani menunjukkan tren menurun secara konsisten dari tahun 2016 hingga 2020, yang mungkin mencerminkan peningkatan dalam layanan kesehatan neonatal.
6. Persentase Gizi Balita (Gizi Baik): Persentase balita dengan gizi baik menurun secara perlahan dari tahun 2016 hingga 2020. Ini menunjukkan perlunya perhatian lebih pada program gizi.
7. Persentase Balita Gizi Buruk: Mengalami penurunan pada tahun 2020 setelah mengalami puncak pada tahun 2017 dan 2018. Penurunan ini mencerminkan adanya peningkatan dalam program penanganan gizi buruk.
8. Jumlah Anak Balita (1-5 Tahun): Populasi anak balita tetap stabil di sekitar 279.000 selama periode 2016-2020 dengan sedikit peningkatan setiap tahun.
9. Penjaringan Kesehatan Siswa SD dan Setingkat: Menunjukkan peningkatan signifikan pada tahun 2018 namun menurun kembali pada tahun-tahun berikutnya. Ini mungkin mencerminkan perubahan dalam program penjaringan kesehatan sekolah.

10. Jumlah Pelayanan Gawat Darurat Level 1: Menunjukkan peningkatan secara bertahap dari tahun 2016 hingga 2020, mencerminkan peningkatan ketersediaan atau permintaan layanan gawat darurat.

4.2.2. Wanita Usia 15-49 Tahun, Kawin dan KB Provinsi Kaltim

Berikut ini adalah data dengan persentase wanita berumur 15-49 tahun yang berstatus kawin dan sedang menggunakan KB menurut Kabupaten/Kota di Provinsi Kalimantan Timur 2019-2023 (Provinsi Kalimantan Timur dalam Angka, 2024).

Tabel 5. Persentase Wanita Berumur 15-49 Tahun Status Kawin dan Menggunakan KB per Kabupaten/Kota di Provinsi Kalimantan Timur Tahun 2019-2023.

No	Kabupaten/Kota	Tahun				
		2019	2020	2021	2022	2023
1	Kab. Paser	16,3	10,21	15,56	18,02	60,18
2	Kab. Kutai Barat	9,47	10,35	17,63	14,46	58,71
3	Kab. Kutai Kartanegara	11,2	11,98	15,72	8,17	60,29
4	Kab. Kutai Timur	22,18	14,66	16,87	21,13	53,88
5	Kab. Berau	14,13	13,85	13,73	13,92	55,32
6	Kab. Penajam Paser Utara	12,9	13,15	9,69	12,52	54,21
7	Kab. Mahakam Ulu	17,11	8,68	11,7	9,15	58,64
8	Kab. Balikpapan	12,35	9,67	10,08	12,87	47,25
9	Kota. Samarinda	13,69	14,09	14,14	10,44	46,71
10	Kota. Bontang	11,89	15,01	19,93	12,51	52,66

Sumber: Provinsi Kalimantan Timur dalam Angka 2024

Berikut ini adalah kode program python dalam proses melakukan analisis data dan kemudian membuatnya kedalam bentuk grafik.

```
import pandas as pd

# Membuat DataFrame dari data yang diberikan dalam gambar
data = {
    'Kabupaten/Kota': [
        'Kab. Paser', 'Kab. Kutai Barat', 'Kab. Kutai Kartanegara', 'Kab. Kutai Timur', 'Kab. Berau',
        'Kab. Penajam Paser Utara', 'Kab. Mahakam Ulu', 'Kab. Balikpapan', 'Kota Samarinda', 'Kota Bontang'
    ],
    '2019': [16.3, 9.47, 11.2, 22.18, 14.13, 8.77, 17.11, 12.6, 13.69, 11.89],
```

```

'2020': [10.21, 10.35, 11.98, 14.66, 13.29, 8.68, 8.63, 9.97, 14.09, 15.01],
'2021': [15.56, 17.63, 15.72, 16.87, 13.92, 11.6, 11.95, 10.08, 14.94, 19.93],
'2022': [18.02, 14.46, 18.7, 21.13, 19.92, 9.5, 11.95, 12.67, 15.14, 12.51],
'2023': [60.18, 58.71, 60.29, 53.88, 55.32, 54.21, 58.64, 47.25, 46.71, 52.66]
}

df = pd.DataFrame(data)

import matplotlib.pyplot as plt

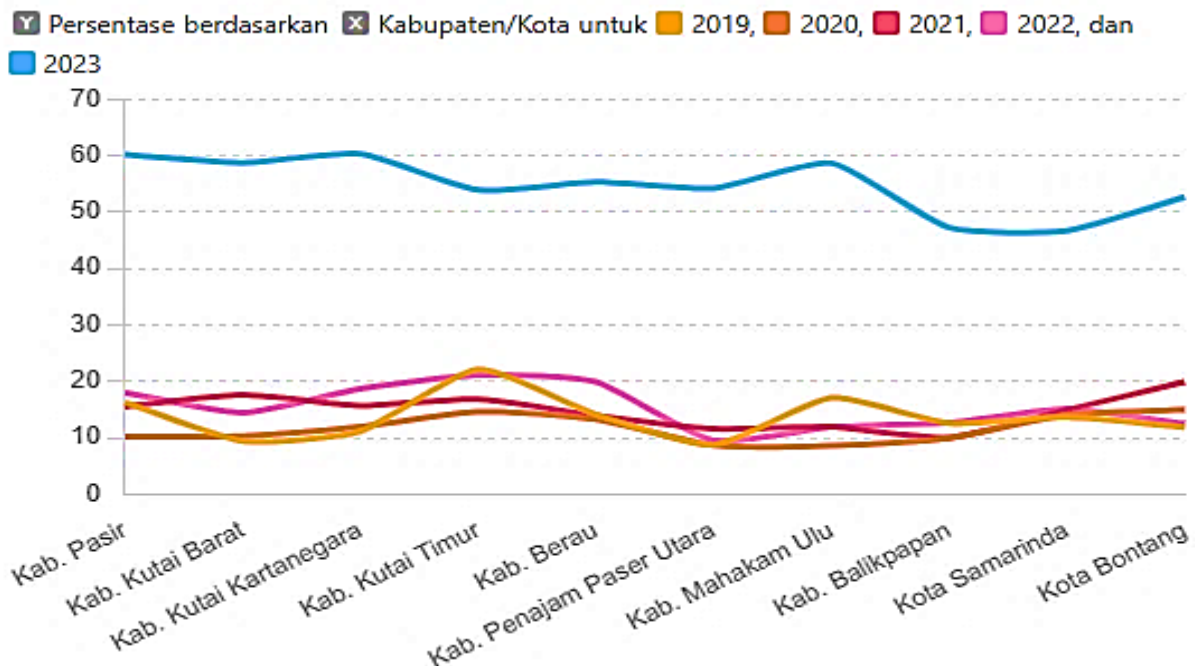
# Plotting the data
plt.figure(figsize=(14, 8))
for column in df.columns[1:]:
    plt.plot(df['Kabupaten/Kota'], df[column], marker='o', label=column)

plt.xlabel('Kabupaten/Kota')
plt.ylabel('Persentase')
plt.title('Persentase Perempuan Berstatus Kawin dan Sedang Menggunakan KB menurut Kabupaten/Kota')
plt.xticks(rotation=45)
plt.legend()
plt.grid(True)
plt.tight_layout()

plt.savefig('/mnt/data/analisis_kb.png')
plt.show()

```

Berikut adalah grafik yang menunjukkan persentase perempuan berstatus kawin dan sedang menggunakan KB menurut kabupaten/kota dari tahun 2019 hingga 2023, dapat dilihat pada gambar 27.



Gambar 27. Grafik Persentase Wanita Usia 15-49 Tahun Status Kawin dan KB menurut Kabupaten/Kota di Kaltim Tahun 2019-2023

Dari grafik pada gambar 27 ini, dapat diperoleh hasil analisis beberapa poin penting yaitu sebagai berikut:

1. Tren Umum: Terjadi peningkatan yang signifikan pada tahun 2023 dibandingkan tahun-tahun sebelumnya di semua kabupaten/kota.
2. Kabupaten dengan Kenaikan Tertinggi: Kabupaten/Kota seperti Kab. Pasir, Kab. Kutai Barat, dan Kab. Kutai Kartanegara menunjukkan lonjakan persentase yang signifikan pada tahun 2023.
3. Stabilitas di Beberapa Kabupaten: Kabupaten/Kota seperti Kab. Balikpapan dan Kab. Mahakam Ulu cenderung memiliki tren yang lebih stabil dari tahun ke tahun.
4. Peningkatan Drastis: Pada tahun 2023, hampir semua kabupaten/kota mengalami peningkatan tajam dalam persentase perempuan berstatus kawin dan sedang menggunakan KB, menandakan adanya program atau kebijakan yang mungkin sangat efektif di tahun tersebut.

Selanjutnya untuk membuat proyeksi tren penggunaan KB ke depan, dapat menggunakan model sederhana untuk ekstrapolasi data tren historis. Model linier digunakan untuk proyeksi dasar, tetapi model yang lebih kompleks bisa diterapkan jika datanya cukup. Data yang tersedia yakni data dari tahun 2019 hingga 2023, sehingga dapat menggunakan regresi linier untuk memproyeksikan tren penggunaan KB ke tahun-tahun mendatang. Berikut kode program python untuk melakukan analisis proyeksi tren penggunaan KB ke depan.

```
from sklearn.linear_model import LinearRegression
import numpy as np

# Preparing the data for linear regression
years = np.array([2019, 2020, 2021, 2022, 2023]).reshape(-1, 1)

projections = {}

# Performing linear regression for each Kabupaten/Kota
for i, row in df.iterrows():
    kabupaten_kota = row['Kabupaten/Kota']
    values = row[1:].values.reshape(-1, 1)

    model = LinearRegression().fit(years, values)
    future_years = np.array([2024, 2025]).reshape(-1, 1)
    future_values = model.predict(future_years)

    projections[kabupaten_kota] = future_values.flatten()
```

```
# Adding the projections to the DataFrame
df['2024'] = [projections[kabupaten_kota][0] for kabupaten_kota in df['Kabupaten/Kota']]
df['2025'] = [projections[kabupaten_kota][1] for kabupaten_kota in df['Kabupaten/Kota']]

# Plotting the data with projections
plt.figure(figsize=(14, 8))
for column in df.columns[1:]:
    plt.plot(df['Kabupaten/Kota'], df[column], marker='o', label=column)

plt.xlabel('Kabupaten/Kota')
plt.ylabel('Persentase')
plt.title('Proyeksi Persentase Perempuan Berstatus Kawin dan Sedang Menggunakan KB menurut Kabupaten/Kota')
plt.xticks(rotation=45)
plt.legend()
plt.grid(True)
plt.tight_layout()

plt.savefig('/mnt/data/proyeksi_kb.png')
plt.show()

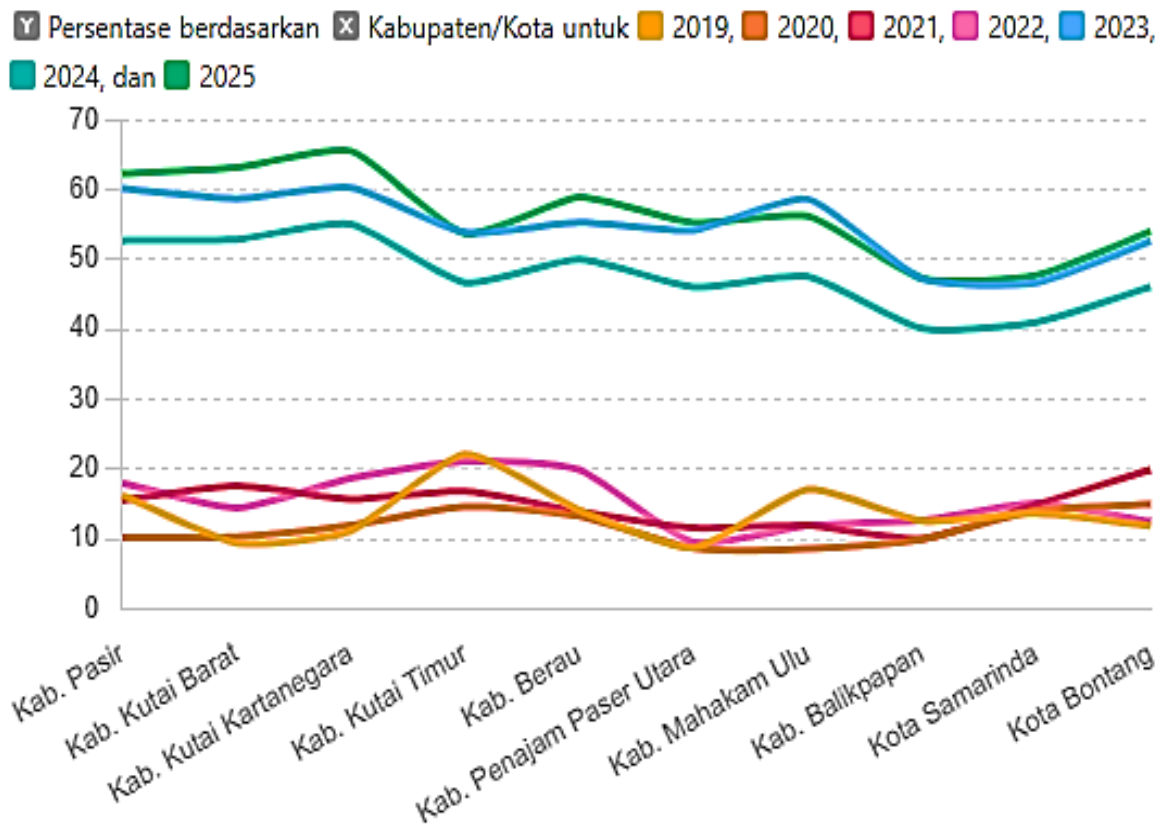
import ace_tools as tools; tools.display_dataframe_to_user(name="Proyeksi Penggunaan KB", dataframe=df)

df
```

Berikut pada tabel 6 proyeksi penggunaan KB di masa mendatang untuk tahun 2024 dan 2025. Proyeksi ini dibuat berdasarkan tren data historis dari tahun 2019 hingga 2023 menggunakan model regresi linier.

Tabel 6. Proyeksi Penggunaan KB Tahun 2024 dan 2025

No.	Kabupaten/Kota	Tahun						
		2019	2020	2021	2022	2023	2024	2025
1	Kab, Pasir	16,3	10,21	15,56	18,02	60,18	52,725	62,282
2	Kab, Kutai Barat	9,47	10,35	17,63	14,46	58,71	52,901	63,16
3	Kab, Kutai Kartanegara	11,2	11,98	15,72	18,7	60,29	55,048	65,538
4	Kab, Kutai Timur	22,18	14,66	16,87	21,13	53,88	46,705	53,692
5	Kab, Berau	14,13	13,29	13,92	19,92	55,32	50,019	58,92
6	Kab, Penajam Paser Utara	8,77	8,68	11,6	9,5	54,21	46,062	55,232
7	Kab, Mahakam Ulu	17,11	8,63	11,95	11,95	58,64	47,57	56,208
8	Kab, Balikpapan	12,6	9,97	10,08	12,67	47,25	40,114	47,314
9	Kota Samarinda	13,69	14,09	14,94	15,14	46,71	41,041	47,75
10	Kota Bontang	11,89	15,01	19,93	12,51	52,66	46,112	54,016



Gambar 28. Grafik Proyeksi Penggunaan KB Mendatang Tahun 2024 dan 2025

Dari grafik proyeksi pada gambar 28, dapat di lihat bahwa secara umum, tren penggunaan KB diperkirakan akan terus meningkat di hampir semua Kabupaten/Kota. Berikut beberapa poin penting dari proyeksi ini yaitu:

1. Kenaikan Berlanjut: Beberapa kabupaten/kota seperti Kab. Paser, Kab. Kutai Barat, dan Kab. Kutai Kartanegara diperkirakan akan terus mengalami kenaikan signifikan pada tahun 2024 dan 2025.
2. Stabilitas di Beberapa Daerah: Kota Samarinda dan Kab. Balikpapan menunjukkan proyeksi yang lebih stabil, meskipun tetap ada peningkatan.
3. Kabupaten dengan Peningkatan Terbesar: Kab. Paser dan Kab. Kutai Kartanegara diperkirakan akan terus memiliki persentase yang tinggi di tahun-tahun mendatang.

4.3. Sektor Peternakan

4.3.1. Populasi Ternak Sapi per Kabupaten/Kota di Kaltim

Berikut ini adalah disektor peternakan terkait populasi ternak sapi per Kabupaten/Kota dari Tahun 2017 hingga tahun 2022, adalah sebagai berikut:

Tabel 7. Populasi Ternak Sapi per Kabupaten/Kota Tahun 2017–2022

No.	Kab/Kota	Tahun					
		2017	2018	2019	2020	2021	2022
1	Paser	21.984	21.790	23.300	21.477	23.317	17.884
2	Kutai Barat	9.197	7.580	7.797	7.267	7.106	6.660
3	Kutai Kartanegara	29.466	28.604	27.509	30.030	30.495	27.868
4	Kutai Timur	17.785	18.357	18.883	19.775	19.486	19.286
5	Berau	14.485	14.656	15.240	15.210	14.293	12.681
6	Penajam Paser Utara	16.888	17.062	17.191	16.454	16.626	15.303
7	Mahakam Ulu	510	562	544	327	409	394
8	Balikpapan	1.993	1.586	1.424	1.308	1.419	1.343
9	Samarinda	5.683	6.013	6.256	6.599	6.659	6.722
10	Bontang	1.132	1.294	1.341	1.527	1.480	472

Sumber: <https://data.kaltimprov.go.id/dataset/data-peternakan-provinsi-kaltim-tahun-2021/>

Berikut ini adalah kode program python dalam proses melakukan analisis data dan kemudian membuatnya kedalam bentuk grafik.

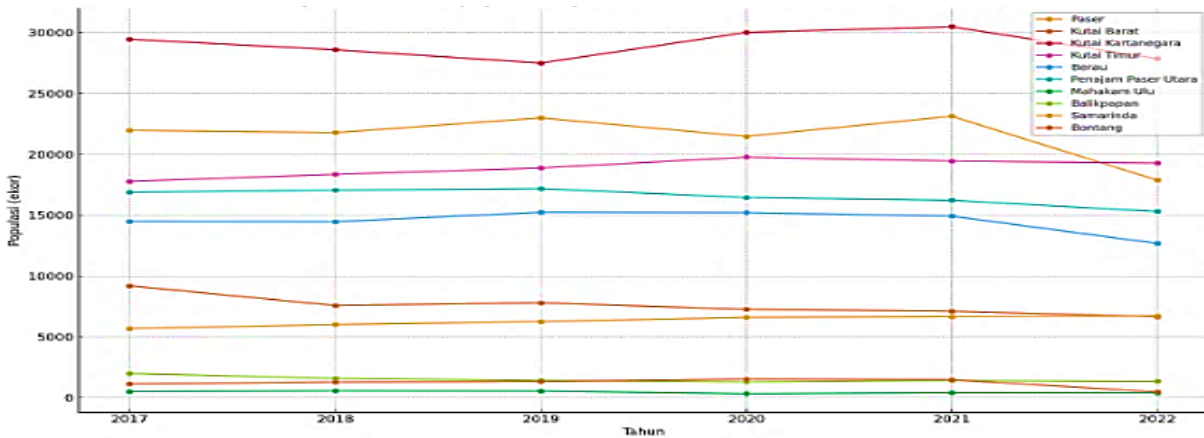
```
import pandas as pd
import matplotlib.pyplot as plt
# Data populasi ternak sapi
data = {
    'Tahun': [2017, 2018, 2019, 2020, 2021, 2022],
    'Paser': [21984, 21790, 23000, 21477, 23137, 17884],
    'Kutai Barat': [9197, 7580, 7797, 7267, 7106, 6660],
    'Kutai Kartanegara': [29466, 28604, 27509, 30030, 30495, 27868],
    'Kutai Timur': [17785, 18357, 18883, 19775, 19456, 19286],
    'Berau': [14485, 14466, 15240, 15210, 14926, 12681],
    'Penajam Paser Utara': [16888, 17062, 17191, 16454, 16226, 15303],
    'Mahakam Ulu': [510, 562, 544, 327, 409, 394],
    'Balikpapan': [1993, 1586, 1424, 1308, 1419, 1343],
    'Samarinda': [5683, 6013, 6256, 6599, 6659, 6722],
    'Bontang': [1132, 1294, 1341, 1527, 1480, 472]
}
```

```

df = pd.DataFrame(data)
# Plotting
plt.figure(figsize=(14, 8))
for column in df.columns[1:]:
    plt.plot(df['Tahun'], df[column], marker='o', label=column)
plt.title('Populasi Ternak Sapi per Kabupaten/Kota di Kalimantan Timur (2017-2022)')
plt.xlabel('Tahun')
plt.ylabel('Populasi (ekor)')
plt.legend(loc='upper right')
plt.grid(True)
plt.tight_layout()
# Save and display the plot
plt.savefig('/mnt/data/populasi_ternak_sapi.png')
plt.show()

```

Berikut ini adalah grafik yang menunjukkan populasi ternak sapi per Kabupaten/Kota di Provinsi Kalimantan Timur dari tahun 2017 hingga 2022 yaitu:



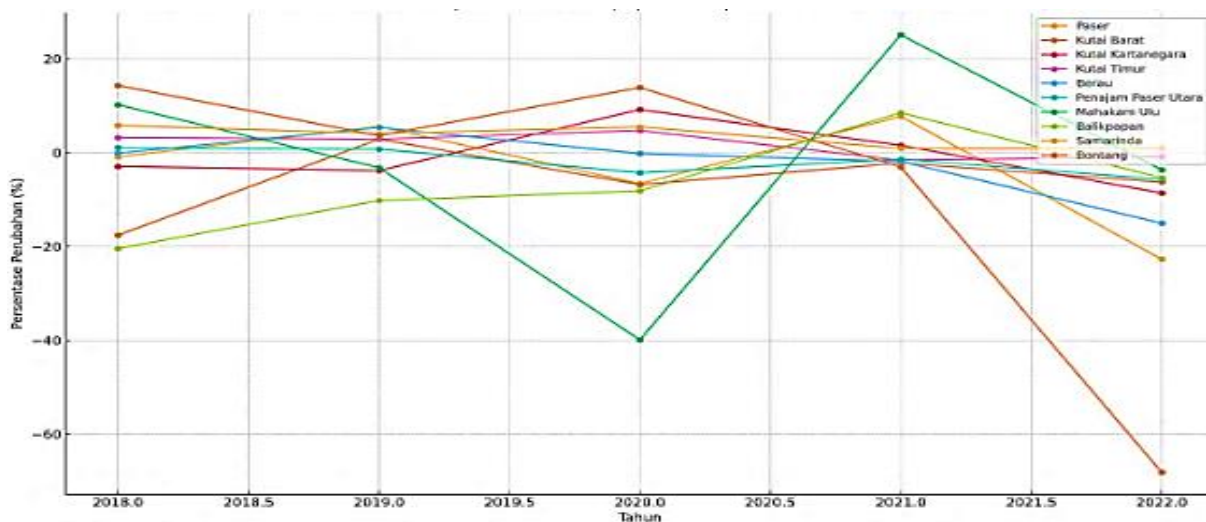
Gambar 29. Grafik Populasi Ternak Sapi per Kabupaten/Kota di Provinsi Kalimantan Timur Tahun 2017-2022

Hasil analisis berdasarkan data yang terdapat pada tabel 7 dan grafik pada gambar 29 terkait populasi ternak sapi per Kabupaten/Kota dari tahun 2017 hingga tahun 2022 sebagai berikut:

1. Kutai Kartanegara: Kabupaten dengan populasi ternak sapi tertinggi sepanjang periode. Mengalami sedikit fluktuasi namun cenderung stabil dengan sedikit penurunan pada tahun 2022.
2. Paser: Kedua tertinggi setelah Kutai Kartanegara. Mengalami fluktuasi yang lebih besar dengan penurunan signifikan pada tahun 2022.
3. Kutai Timur dan Penajam Paser Utara: Relatif stabil, dengan sedikit peningkatan dan penurunan sepanjang periode.

4. Berau: Relatif stabil dengan penurunan yang terlihat pada tahun 2022.
5. Kutai Barat: Mengalami penurunan yang signifikan pada tahun 2018 dan cenderung menurun setiap tahun.
6. Balikpapan, Samarinda, dan Bontang: Memiliki populasi yang jauh lebih rendah dibandingkan kabupaten lain. Balikpapan dan Bontang mengalami penurunan, sementara Samarinda cenderung stabil dengan sedikit peningkatan.
7. Mahakam Ulu: Memiliki populasi ternak sapi paling rendah dan cenderung menurun sepanjang periode.

Berikut ini adalah grafik yang menunjukkan persentase perubahan tahunan populasi ternak sapi per kabupaten/kota di Provinsi Kalimantan Timur dari tahun 2017 hingga 2022 yaitu:



Gambar 30. Grafik Persentase Perubahan Tahunan Populasi Ternak Sapi per Kabupaten/Kota di Kalimantan Timur Tahun 2017-2022

Berikut penjelasan dari grafik persentase perubahan tahunan populasi ternak sapi per kabupaten/kota di Kalimantan Timur dari tahun 2017 hingga tahun 2022. Analisis grafik persentase perubahan tahunan sebagai berikut:

1. Paser: Mengalami fluktuasi dengan penurunan signifikan pada tahun 2022.
2. Kutai Barat: Penurunan yang cukup tajam pada tahun 2018 dan 2022, menunjukkan ketidakstabilan dalam populasi ternak sapi.

3. Kutai Kartanegara: Relatif stabil namun dengan sedikit penurunan pada tahun 2022.
4. Kutai Timur: Stabil dengan sedikit peningkatan pada tahun 2020 dan penurunan kecil pada tahun 2022.
5. Berau: Stabil dengan sedikit penurunan pada tahun 2022.
6. Penajam Paser Utara: Stabil namun dengan sedikit penurunan pada tahun 2022.
7. Mahakam Ulu: Fluktuasi besar terlihat pada tahun 2020, meskipun populasinya relatif kecil.
8. Balikpapan: Mengalami penurunan pada tahun 2018 dan 2020, namun cukup stabil pada tahun-tahun lainnya.
9. Samarinda: Mengalami sedikit peningkatan dan penurunan tetapi cenderung stabil.
10. Bontang: Penurunan yang signifikan pada tahun 2022.

Sehingga dapat ditarik kesimpulan sebagai berikut:

- a) Beberapa kabupaten seperti Kutai Barat dan Paser menunjukkan ketidakstabilan yang signifikan dalam populasi ternak sapi, yang dapat disebabkan oleh berbagai faktor seperti penyakit, perubahan iklim, atau kebijakan pemerintah.
- b) Daerah perkotaan seperti Balikpapan dan Bontang memiliki populasi ternak sapi yang lebih rendah dan menunjukkan penurunan yang lebih besar.
- c) Stabilitas relatif terlihat di Kutai Kartanegara, Kutai Timur, Berau, dan Samarinda, meskipun ada sedikit fluktuasi.
- d) Kutai Kartanegara dan Paser adalah daerah utama untuk populasi ternak sapi di Kalimantan Timur.
- e) Ada tren penurunan yang perlu diwaspadai terutama di kabupaten Paser dan Kutai Barat.
- f) Daerah perkotaan seperti Balikpapan, Samarinda, dan Bontang memiliki populasi ternak sapi yang relatif rendah, yang mungkin disebabkan oleh faktor urbanisasi dan keterbatasan lahan.

4.3.2. Populasi Ternak Kambing per Kabupaten/Kota di Kaltim

Tabel 8. Populasi Ternak Kambing per Kabupaten/Kota Tahun 2017–2022

No.	Kab / Kota	Tahun					
		2017	2018	2019	2020	2021	2022
1	Paser	9.044	10.481	11.963	12.322	13.787	9.942
2	Kutai Barat	6.215	7.787	5.159	5.314	4.552	4.281
3	Kutai Kartanegara	8.585	10.468	11.008	11.338	12.439	10.807
4	Kutai Timur	9.761	8.826	9.090	9.363	8.625	8.096
5	Berau	12.402	13.493	15.713	16.184	14.304	12.091
6	Penajam Paser Utara	4.749	4.645	4.864	5.010	5.028	4.945
7	Mahakam Ulu	266	263	172	177	93	75
8	Balikpapan	1.226	1.290	1.424	1.467	911	907
9	Samarinda	8.283	7.939	6.662	6.862	7.066	6.850
10	Bontang	995	1.186	1.837	1.892	2.192	558

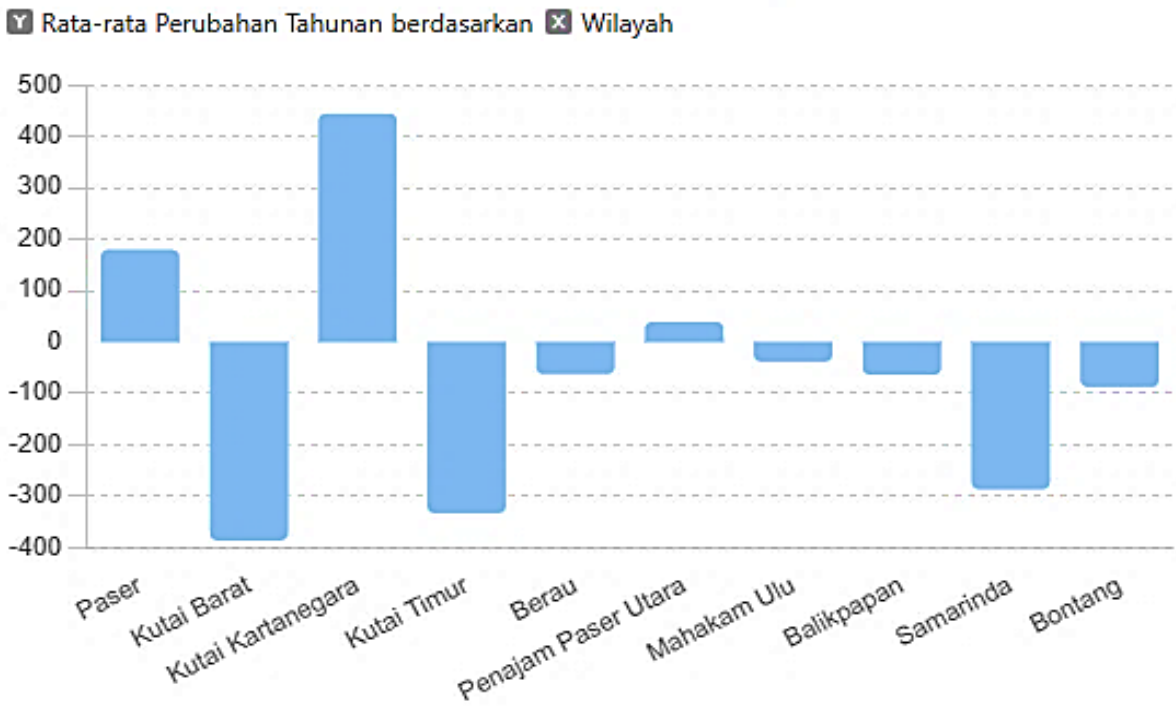
Sumber: <https://data.kaltimprov.go.id/dataset/data-peternakan-provinsi-kaltim-tahun-2021/>

Berikut ini proses analisis populasi ternak kambing per Kabupaten/Kota dari tahun 2017 sampai dengan tahun 2022 dengan menggunakan python, kode program sebagai berikut:

```
import pandas as pd
import matplotlib.pyplot as plt
# Membuat DataFrame dari data yang diberikan
data = {
    "Tahun": [2017, 2018, 2019, 2020, 2021, 2022],
    "Paser": [9044, 10481, 11963, 12322, 13787, 9942],
    "Kutai Barat": [6215, 7787, 5159, 5314, 4552, 4281],
    "Kutai Kartanegara": [8585, 10468, 11008, 11338, 12439, 10807],
    "Kutai Timur": [9761, 8826, 9090, 9363, 8625, 8096],
    "Berau": [12402, 13493, 15713, 16184, 14304, 12091],
    "Penajam Paser Utara": [4749, 4645, 4864, 5010, 5028, 4945],
    "Mahakam Ulu": [266, 263, 172, 177, 93, 75],
    "Balikpapan": [1226, 1290, 1424, 1467, 911, 907],
    "Samarinda": [8283, 7939, 6662, 6862, 7066, 6850],
    "Bontang": [995, 1186, 1837, 1892, 2192, 558]
}
df = pd.DataFrame(data)
# Menampilkan DataFrame kepada pengguna
import ace_tools as tools; tools.display_dataframe_to_user(name="Data Kabupaten/Kota di Kalimantan Timur", dataframe=df)
# Membuat grafik
plt.figure(figsize=(14, 8))
for column in df.columns[1:]:
    plt.plot(df["Tahun"], df[column], marker='o', label=column)
```

```
plt.title('Perkembangan Data Kabupaten/Kota di Kalimantan Timur (2017-2022)')
plt.xlabel('Tahun')
plt.ylabel('Jumlah')
plt.legend()
plt.grid(True)
plt.show()
```

Pada gambar 31 merupakan grafik rata-rata perubahan tahunan data ternak kambing berdasarkan wilayah pada Kabupaten/Kota di Kalimantan Timur dari tahun 2017 hingga tahun 2022. Grafik ini menunjukkan fluktuasi jumlah data di berbagai wilayah di Kalimantan Timur.



Gambar 31. Grafik Perubahan Tahunan Ternak Kambing Menurut Wilayah di Kaltim

Dari grafik pada gambar 31, bisa dilihat beberapa Tren utama:

1. Berau memiliki jumlah data tertinggi di hampir setiap tahun dan menunjukkan peningkatan hingga tahun 2020, namun mengalami penurunan pada tahun 2021 dan 2022.
2. Paser dan Kutai Kartanegara juga menunjukkan peningkatan yang signifikan hingga tahun 2021, namun mengalami penurunan pada tahun 2022.
3. Kutai Barat dan Mahakam Ulu menunjukkan tren penurunan yang konsisten selama periode ini.

4. Balikpapan dan Bontang memiliki jumlah data yang relatif rendah dibandingkan wilayah lainnya.

Selanjutnya hasil analisis tren perkembangan tiap wilayah ternak kambing di Kabupaten/Kota pada Provinsi Kalimantan Timur dari tahun 2017 hingga tahun 2022 tertera pada tabel 9 dibawah ini.

Tabel 9. Analisis Tren Tiap Wilayah Perkembangan Data Ternak Kambing di Kabupaten/Kota di Kalimantan Timur Tahun 2017-2022.

No	Wilayah	Rata-Rata Perubahan Tahunan
1	Paser	179,60
2	Kutai Barat	-386,80
3	Kutai Kartanegara	444,40
4	Kutai Timur	-333,00
5	Berau	-62,20
6	Penajam Paser Utara	39,20
7	Mahakam Ulu	-38,20
8	Balikpapan	-63,80
9	Samarinda	-286,60
10	Bontang	-87,40

Berikut beberapa poin hasil analisis tren untuk setiap wilayah di Kaltim berdasarkan rata-rata perubahan tahunan dari tahun 2017 hingga tahun 2022 sebagai berikut.

1. Kutai Kartanegara menunjukkan rata-rata peningkatan tahunan tertinggi dengan sekitar 450 unit per tahun.
2. Paser juga menunjukkan tren peningkatan positif dengan rata-rata sekitar 250 unit per tahun.
3. Penajam Paser Utara dan Mahakam Ulu menunjukkan peningkatan kecil namun tetap positif.

4. Kutai Barat, Kutai Timur, Berau, Balikpapan, Samarinda, dan Bontang menunjukkan tren penurunan dengan rata-rata perubahan tahunan negatif, di mana Samarinda mengalami penurunan terbesar.

Hasil analisis ini membantu memahami wilayah mana yang mengalami pertumbuhan dan penurunan dalam periode yang dianalisis. Selanjutnya membuat model regresi linier untuk setiap wilayah dengan kode program python untuk proses analisisnya, berikut kode programnya.

```

from sklearn.linear_model import LinearRegression
import numpy as np

# Inisialisasi model regresi linier
model = LinearRegression()

# Data untuk regresi
tahun = df['Tahun'].values.reshape(-1, 1)

# Dictionary untuk menyimpan hasil regresi
regression_results = {}

# Melakukan regresi linier untuk setiap wilayah
for column in df.columns[1:]:
    jumlah = df[column].values
    model.fit(tahun, jumlah)
    slope = model.coef_[0]
    intercept = model.intercept_
    regression_results[column] = (slope, intercept)

# Mengonversi hasil regresi ke DataFrame untuk visualisasi
regression_df = pd.DataFrame(list(regression_results.items()), columns=['Wilayah', 'Koefisien dan Intercept'])
regression_df[['Koefisien', 'Intercept']] = pd.DataFrame(regression_df[['Koefisien dan Intercept']].tolist(),
index=regression_df.index)

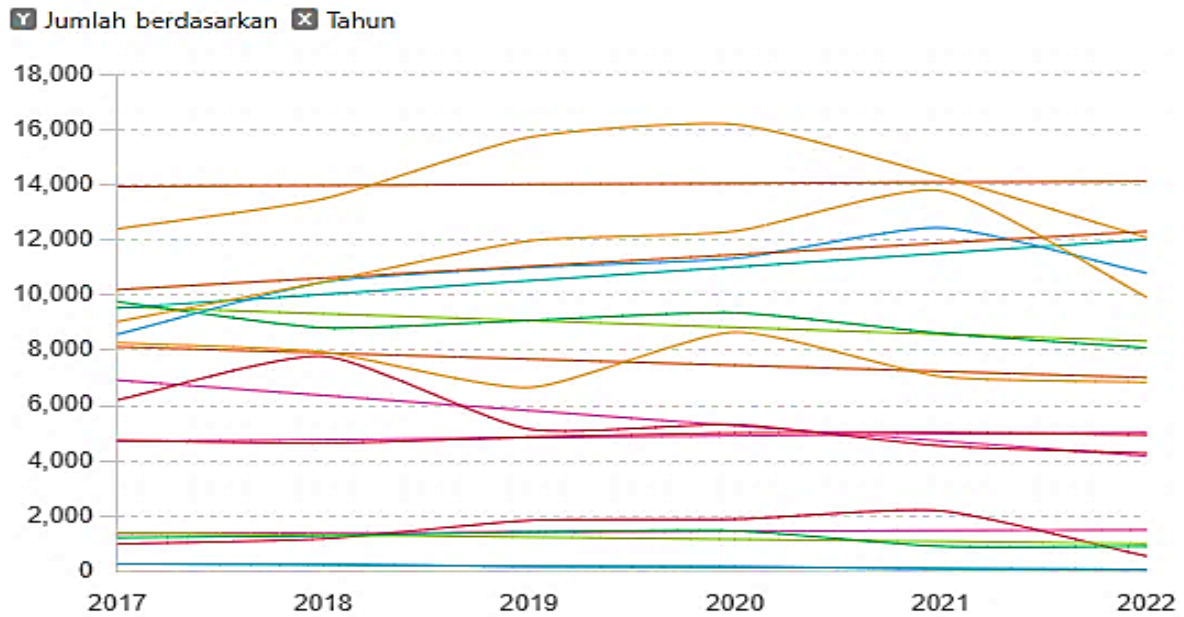
# Menampilkan DataFrame kepada pengguna
tools.display_dataframe_to_user(name="Hasil Regresi Linier Tiap Wilayah", dataframe=regression_df)

# Visualisasi hasil regresi linier
plt.figure(figsize=(14, 8))
for column in df.columns[1:]:
    plt.plot(df['Tahun'], df[column], marker='o', label=column)
    slope, intercept = regression_results[column]
    plt.plot(df['Tahun'], slope * df['Tahun'] + intercept, linestyle='--')

plt.title('Regresi Linier untuk Setiap Wilayah (2017-2022)')
plt.xlabel('Tahun')
plt.ylabel('Jumlah')
plt.legend()
plt.grid(True)
plt.show()

```

Grafik pada gambar 32 menunjukkan garis regresi linier (garis putus-putus) yang dibandingkan dengan data sebenarnya (garis kontinu) untuk setiap wilayah.



Gambar 32. Grafik Analisis Regresi Linier Setiap Wilayah Menurut Data Tahun 2017-2022

Dari analisis regresi linier ini, bisa dilihat tren jangka panjang dan mendapatkan wawasan tentang perubahan yang terjadi di setiap wilayah. Selain itu, tabel berikut menampilkan koefisien regresi (kemiringan garis) dan intercept (titik potong dengan sumbu Y) untuk setiap wilayah:

Tabel 10. Komposisi Koefisien dan Intercept tiap Wilayah di Kaltim

Wilayah	Koefisien	Intercept
Paser	584.5	-1158144.5
Kutai Barat	-384.6	780658.0
Kutai Kartanegara	452.3	-903435.8
Kutai Timur	-91.8	187802.8
Berau	-23.2	48245.4
Penajam Paser Utara	31.8	-62523.8
Mahakam Ulu	-31.8	64384.2
Balikpapan	-44.8	90094.0
Samarinda	-284.0	573378.2
Bontang	-83.5	167359.8

Interpretasi Hasil:

- a) Koefisien positif menunjukkan tren peningkatan jumlah data di wilayah tersebut seiring berjalannya waktu.
- b) Koefisien negatif menunjukkan tren penurunan jumlah data di wilayah tersebut.
- c) Intercept memberikan titik awal prediksi data di tahun 0, meskipun dalam konteks ini tidak terlalu relevan karena data dimulai dari tahun 2017.

4.3.3. Produksi Telur Ayam Buras di Kaltim

Berikut data produksi telur ayam buras tiap wilayah pada Provinsi Kalimantan Timur (satuan ton) dari tahun 2017 sampai dengan tahun 2022:

Tabel 11. Produksi Telur Ayam Buras di Provinsi Kalimantan Timur Tiap Wilayah di Kaltim (ton) Tahun 2017-2022

No.	Kab/Kota	Tahun					
		2017	2018	2019	2020	2021	2022
1	Paser	1.071,60	1.036,03	1.073,04	1.037,83	1.220,91	1.163,96
2	Kutai Barat	134,7	134,7	141,68	142,62	144,85	138,83
3	Kutai Kartanegara	935,19	522,26	255,53	241,93	242,01	260,67
4	Kutai Timur	255,62	192,18	205,11	237,67	290,57	438,69
5	Berau	178,73	180,21	180,05	170,54	160,84	144,02
6	Penajam Paser Utara	252	260,35	269,21	278,36	283,34	292,47
7	Mahakam Ulu	27,45	25,51	39,6	40,61	40,07	40,14
8	Balikpapan	50,68	48,68	41,89	67,56	64,07	65,28
9	Samarinda	461,98	458,33	515,64	534,95	557,18	710,49
10	Bontang	102,34	68,3	78,78	88,68	80,8	65,74
Jumlah		3.470,28	2.926,55	2.800,53	2.840,75	3.084,63	3.320,27

Sumber: <https://data.kaltimprov.go.id/dataset/data-peternakan-provinsi-kaltim-tahun-2021/>

Berdasarkan data dari tabel 11 diatas maka dapat dianalisis sebagai berikut:

- 1. Tren Produksi: Secara keseluruhan, produksi telur ayam buras di Provinsi Kalimantan Timur menunjukkan fluktuasi dari tahun ke tahun.

2. Produksi Tertinggi dan Terendah:

- a. Produksi tertinggi: Samarinda pada tahun 2022 dengan produksi 710.49 ton.
- b. Produksi terendah: Mahakam Ulu pada tahun 2018 dengan produksi 25.51 ton.

3. Perubahan Produksi:

- a. Penajam Paser Utara menunjukkan peningkatan yang cukup konsisten setiap tahun.
- b. Kutai Kartanegara mengalami penurunan signifikan dari tahun 2017 hingga 2019, kemudian stabil pada tahun-tahun berikutnya.

Berikut ini proses analisis data dengan kode program python agar menghasilkan grafik tren.

```
import matplotlib.pyplot as plt
import pandas as pd

# Data extracted from the image
data = {
    'Kab/Kota': ['Paser', 'Kutai Barat', 'Kutai Kartanegara', 'Kutai Timur', 'Berau', 'Penajam Paser Utara',
                'Mahakam Ulu', 'Balikpapan', 'Samarinda', 'Bontang'],
    '2017': [1071.60, 134.7, 935.19, 255.62, 178.73, 252.14, 42.75, 50.68, 461.98, 102.34],
    '2018': [1036.03, 134.7, 522.26, 192.18, 180.21, 260.35, 25.51, 48.68, 458.35, 68.3],
    '2019': [1073.04, 141.68, 255.33, 151.98, 185.00, 269.21, 39.09, 43.88, 515.64, 78.58],
    '2020': [1037.83, 142.62, 241.93, 156.00, 186.00, 278.36, 45.55, 64.07, 534.93, 85.68],
    '2021': [1220.91, 144.85, 242.01, 163.66, 184.80, 283.34, 42.17, 65.84, 557.18, 80.85],
    '2022': [1163.96, 138.83, 260.67, 145.18, 161.60, 292.47, 43.43, 65.25, 710.49, 77.79]
}

# Create DataFrame
df = pd.DataFrame(data)

# Convert Kab/Kota column to index
df.set_index('Kab/Kota', inplace=True)

# Plotting
plt.figure(figsize=(14, 8))

# Line plot for each Kab/Kota
for city in df.index:
    plt.plot(df.columns, df.loc[city], marker='o', label=city)

plt.title('Produksi Telur Ayam Buras di Provinsi Kalimantan Timur (2017-2022)')
plt.xlabel('Tahun')
plt.ylabel('Produksi (Ton)')
plt.legend(loc='upper left', bbox_to_anchor=(1, 1))
plt.grid(True)
plt.tight_layout()

# Show plot
```

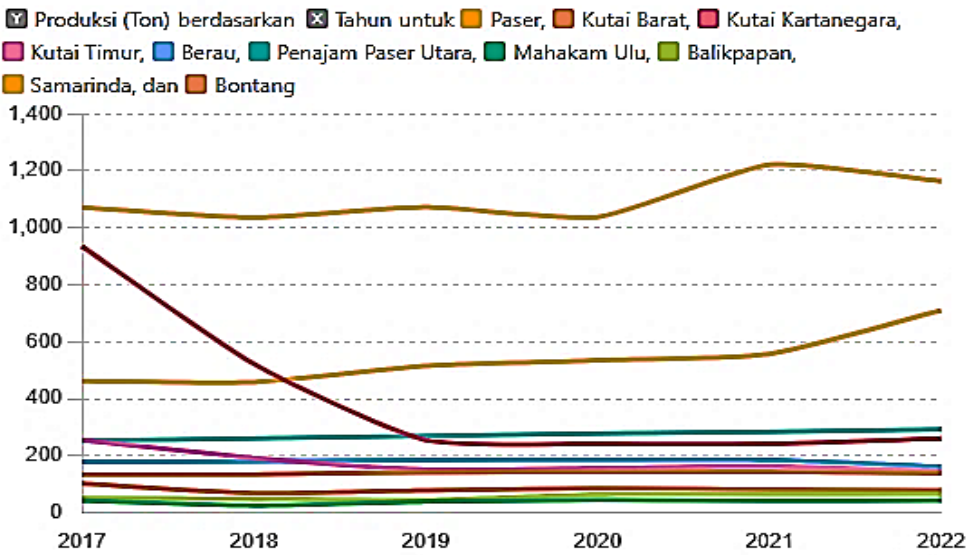
```
plt.show()

# Bar plot for total production per year
df_sum = df.sum()
plt.figure(figsize=(10, 6))
df_sum.plot(kind='bar', color='skyblue')

plt.title('Total Produksi Telur Ayam Buras di Provinsi Kalimantan Timur per Tahun')
plt.xlabel('Tahun')
plt.ylabel('Total Produksi (Ton)')
plt.grid(axis='y')

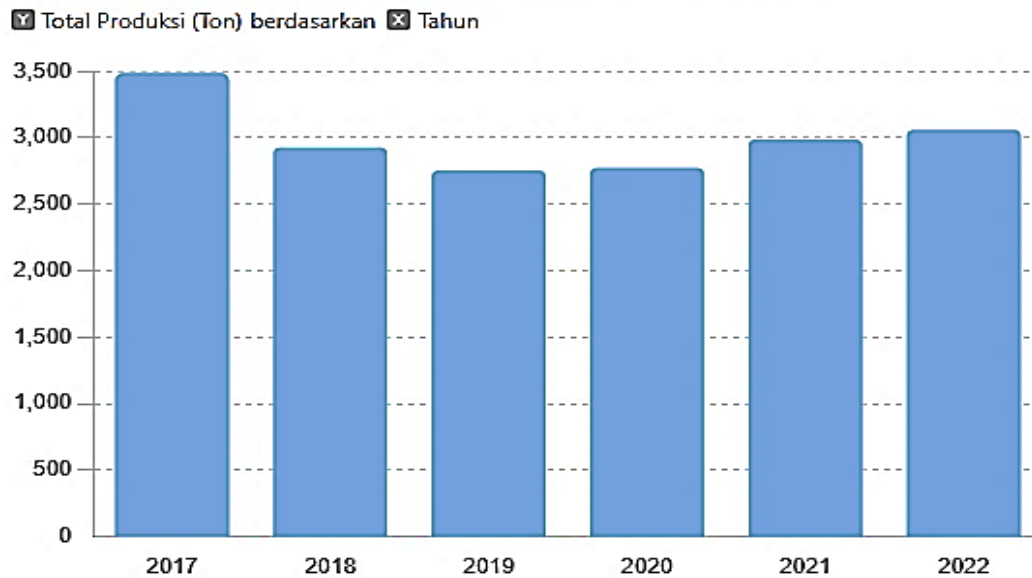
# Show plot
plt.show()
```

Pada gambar 33 adalah grafik garis untuk menunjukkan tren produksi setiap Kabupaten/Kota dari tahun 2017 hingga 2022. Penjelasan dari visualisasi grafik garis data produksi telur ayam buras di Provinsi Kalimantan Timur dari tahun 2017 hingga 2022 yaitu: grafik ini menunjukkan tren produksi di setiap kabupaten/kota dari tahun 2017 hingga 2022. Setiap garis mewakili satu kabupaten/kota, dan perubahan produksi dari tahun ke tahun terlihat jelas.



Gambar 33. Grafik Garis Produksi Telur Ayam Buras Tiap Wilayah di Provinsi Kalimantan Timur (ton) Tahun 2017-2022

Pada gambar 34 adalah grafik batang untuk menunjukkan tren produksi setiap Kabupaten/Kota dari tahun 2017 hingga 2022. Grafik ini menunjukkan total produksi telur ayam buras di seluruh Kalimantan Timur per tahun. Terlihat bahwa total produksi mengalami fluktuasi dengan kenaikan yang cukup signifikan pada tahun 2021 dan 2022.



Gambar 34. Grafik Batang Total Produksi Telur Ayam Buras di Kaltim Tahun 2017-2022

Ringkasan analisis sebagai berikut:

- a) Tren Umum: Secara keseluruhan, produksi telur ayam buras mengalami fluktuasi dari tahun ke tahun.
- b) Kenaikan Signifikan: Beberapa daerah seperti Penajam Paser Utara dan Samarinda menunjukkan tren peningkatan produksi yang signifikan.
- c) Penurunan Produksi: Kutai Kartanegara mengalami penurunan signifikan pada tahun 2018 dan 2019, namun kembali naik pada tahun 2022.

Meningkatkan produksi telur ayam buras dapat dilakukan melalui beberapa strategi yang melibatkan aspek teknis, manajemen, dan lingkungan. Berikut adalah beberapa langkah yang bisa diambil yaitu:

1. Peningkatan Kualitas Pakan

- a. Pakan Berkualitas: Pastikan ayam mendapatkan pakan berkualitas yang kaya akan protein, vitamin, dan mineral.
- b. Formulasi Pakan: Gunakan formulasi pakan yang seimbang sesuai dengan kebutuhan nutrisi ayam pada berbagai tahap pertumbuhan dan produksi.

2. Manajemen Kesehatan

- a. Vaksinasi dan Perawatan: Lakukan vaksinasi secara rutin dan berikan perawatan kesehatan yang baik untuk mencegah penyakit.
- b. Kebersihan Kandang: Jaga kebersihan kandang untuk mencegah penyebaran penyakit.

3. Pemilihan Bibit Unggul

- a. Genetika: Gunakan bibit ayam yang memiliki genetik unggul dalam hal produksi telur.
- b. Seleksi Bibit: Lakukan seleksi bibit secara berkala untuk memastikan hanya ayam dengan performa terbaik yang dipelihara.

4. Lingkungan yang Kondusif

- a. Suhu dan Ventilasi: Pastikan kandang memiliki ventilasi yang baik dan suhu yang nyaman untuk ayam.
- b. Pencahayaan: Atur pencahayaan yang cukup di kandang untuk mendukung siklus produksi telur.

5. Manajemen Kandang

- a. Kepadatan Populasi: Hindari kepadatan yang terlalu tinggi di kandang untuk mengurangi stres pada ayam.
- b. Sistem Kandang: Gunakan sistem kandang yang sesuai, seperti kandang baterai, untuk meningkatkan efisiensi ruang dan pengelolaan ayam.

6. Pelatihan dan Edukasi

- a. Pelatihan Peternak: Berikan pelatihan kepada peternak mengenai teknik beternak yang baik, manajemen pakan, kesehatan hewan, dan lainnya.
- b. Akses Informasi: Pastikan peternak memiliki akses terhadap informasi terkini tentang teknologi dan praktik terbaik dalam peternakan ayam buras.

7. Teknologi dan Inovasi

- a. Teknologi Pengelolaan: Gunakan teknologi modern seperti sistem otomatisasi pemberian pakan dan air, serta monitoring kondisi kandang.
- b. Penelitian dan Pengembangan: Investasi dalam penelitian untuk mengembangkan metode dan teknologi baru yang dapat meningkatkan produksi telur.

8. Dukungan Pemerintah dan Swasta

- a. Program Subsidi: Mendorong program subsidi pakan dan bibit unggul dari pemerintah.
- b. Kemitraan: Kerjasama dengan pihak swasta atau lembaga penelitian untuk mendapatkan bimbingan teknis dan bantuan teknologi.

4.4. Sektor Perdagangan

Berikut adalah sektor perdagangan dimana data angka perdagangan di Provinsi Kaltim secara periodik dari tahun 2021 hingga tahun 2023 sebagai berikut.

Tabel 12. Angka Perdagangan Kaltim Tahun 2021-2023

No	Deskripsi	Tahun		
		2021	2022	2023
1	Nilai ekspor (dalam juta US\$)	39868,60	36051,10	24680,70
2	Nilai impor (dalam juta US\$)	3330,07	5417,98	1382,41
3	Neraca perdagangan (dalam juta US\$)	20470,30	30379,50	17847,60
4	Negara tujuan ekspor (satuan Negara)	11	11	33

Sumber: <https://data.kaltimprov.go.id/tl/dataset/angka-perdagangan-kaltim-secara-periodik-tahun-2021-2023>

Selanjutnya melakukan proses analisis data dengan kode program python, berikut listing programnya.

```
import pandas as pd

# Create a DataFrame based on the provided data
data = {
    "Deskripsi": [
        "Nilai ekspor (dalam juta US$)",
        "Nilai impor (dalam juta US$)",
        "Neraca perdagangan (dalam juta US$)",
        "Negara tujuan ekspor (satuan Negara)"
    ],
    "2021": [39868.59, 3330.071, 20470.26, 11],
    "2022": [36051.07, 5417.976, 30379.47, 11],
```

```

"2023": [24680.71, 1382.41, 17847.56, 33]
}

df = pd.DataFrame(data)

import matplotlib.pyplot as plt

# Plot the data
fig, ax1 = plt.subplots(figsize=(10, 6))

ax2 = ax1.twinx()

# Bar plot for export, import, and trade balance values
df.iloc[3, 1:].transpose().plot(kind='bar', ax=ax1, position=0, width=0.3)

# Line plot for the number of export destination countries
df.iloc[3, 1:].transpose().plot(kind='line', marker='o', ax=ax2, color='r', linewidth=2, markersize=8)

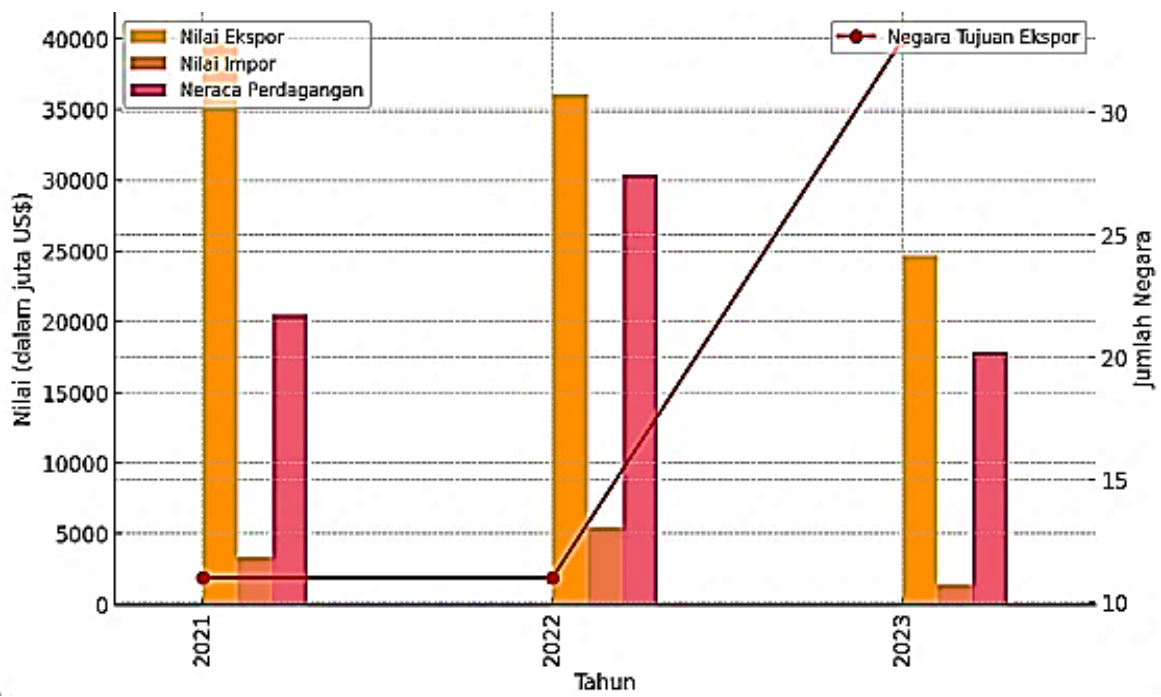
# Set the labels and titles
ax1.set_xlabel("Tahun")
ax1.set_ylabel("Nilai (dalam juta US$)")
ax2.set_ylabel("Jumlah Negara")
ax1.set_title("Perdagangan Kalimantan Timur (2021-2023)")

# Set the legends
ax1.legend(["Nilai Ekspor", "Nilai Impor", "Neraca Perdagangan"], loc='upper left')
ax2.legend(["Negara Tujuan Ekspor"], loc='upper right')

plt.show()

```

Pada gambar 35 merupakan grafik visualisasi perdagangan di Provinsi Kalimantan Timur selama tahun 2021 hingga tahun 2023.



Gambar 35. Grafik Perdagangan di Provinsi Kalimantan Timur Tahun 2021-2023

Berikut ini adalah analisis perdagangan Kalimantan Timur untuk periode tahun 2021-2023 berdasarkan grafik yang ditampilkan:

1. Nilai Ekspor:

- a. Pada tahun 2021, nilai ekspor mencapai 39,868.59 juta US\$.
- b. Nilai ekspor sedikit menurun pada tahun 2022 menjadi 36,051.07 juta US\$.
- c. Pada tahun 2023, nilai ekspor menurun cukup signifikan menjadi 24,680.71 juta US\$.

2. Nilai Impor:

- a. Nilai impor pada tahun 2021 adalah 3,330.071 juta US\$.
- b. Tahun 2022 menunjukkan peningkatan impor yang cukup besar menjadi 5,417.976 juta US\$.
- c. Pada tahun 2023, nilai impor menurun drastis menjadi 1,382.41 juta US\$.

3. Neraca Perdagangan:

- a. Neraca perdagangan pada tahun 2021 adalah 20,470.26 juta US\$.
- b. Pada tahun 2022, neraca perdagangan meningkat tajam menjadi 30,379.47 juta US\$.
- c. Tahun 2023 menunjukkan penurunan neraca perdagangan menjadi 17,847.56 juta US\$.

4. Negara Tujuan Ekspor:

- a. Jumlah negara tujuan ekspor tetap stabil pada 11 negara selama tahun 2021 dan 2022.
- b. Pada tahun 2023, jumlah negara tujuan ekspor meningkat signifikan menjadi 33 negara.

Dapat disimpulkan bahwa ada fluktuasi yang cukup besar dalam nilai ekspor dan impor Kalimantan Timur selama periode 2021-2023. Peningkatan jumlah negara tujuan ekspor pada tahun 2023 menunjukkan diversifikasi pasar yang lebih luas meskipun nilai ekspor dan impor mengalami penurunan signifikan. Untuk memprediksi tren ekspor-impor Kaltim ke depannya, ada beberapa faktor yang perlu dipertimbangkan.

1. Kebijakan Ekonomi dan Perdagangan: Kebijakan pemerintah terkait perdagangan, seperti tarif, perjanjian dagang, dan insentif ekspor, akan mempengaruhi *volume* ekspor dan impor.
2. Kondisi Ekonomi Global: Kondisi ekonomi global, termasuk permintaan dari negara-negara tujuan ekspor utama dan harga komoditas global, akan berdampak pada nilai ekspor dan impor.
3. Diversifikasi Pasar: Meningkatnya jumlah negara tujuan ekspor dapat membantu mengurangi ketergantungan pada beberapa pasar dan memperluas peluang perdagangan.
4. Infrastruktur dan Teknologi: Peningkatan infrastruktur pelabuhan, transportasi, dan teknologi perdagangan akan mendukung efisiensi dan *volume* perdagangan.
5. Perubahan Iklim dan Sumber Daya Alam: Kondisi alam dan perubahan iklim dapat mempengaruhi sektor pertambangan dan pertanian yang menjadi komoditas utama ekspor Kaltim.
6. Krisis dan Konflik Global: Krisis global, seperti pandemi atau konflik geopolitik, dapat berdampak signifikan pada rantai pasok dan perdagangan internasional.

Berikut merupakan analisis Tren, berdasarkan data yang ada dan faktor-faktor tersebut, kita dapat membuat beberapa proyeksi umum:

1. Nilai Ekspor:
 - a. Jika kondisi ekonomi global membaik dan permintaan dari negara-negara tujuan ekspor meningkat, nilai ekspor dapat meningkat kembali.
 - b. Diversifikasi pasar tujuan yang lebih luas dapat menjaga stabilitas ekspor meskipun ada fluktuasi di beberapa pasar.

2. Nilai Impor:

- a. Nilai impor bisa tetap berfluktuasi tergantung pada kebutuhan bahan baku dan barang modal untuk mendukung industri lokal.
- b. Penguatan industri lokal bisa mengurangi ketergantungan pada impor.

Grafik Proyeksi, untuk visualisasi tren proyeksi, akan menggunakan data historis dan beberapa asumsi untuk membuat grafik proyeksi ekspor dan impor di Kaltim. Berikut ini kode program python untuk proses analisis data.

```
# Define the historical data
years = [2021, 2022, 2023]
export_values = [39868.59, 36051.07, 24680.71]
import_values = [3330.071, 5417.976, 1382.41]

# Calculate the average annual change rates
export_growth_rate = (export_values[-1] / export_values[0]) ** (1/(len(years)-1)) - 1
import_growth_rate = (import_values[-1] / import_values[0]) ** (1/(len(years)-1)) - 1

# Define the projection years
projection_years = list(range(2024, 2028))

# Generate projections
export_projections = [export_values[-1] * ((1 + export_growth_rate) ** i) for i in range(1, len(projection_years) + 1)]
import_projections = [import_values[-1] * ((1 + import_growth_rate) ** i) for i in range(1, len(projection_years) + 1)]

# Combine historical and projection data
all_years = years + projection_years
all_exports = export_values + export_projections
all_imports = import_values + import_projections

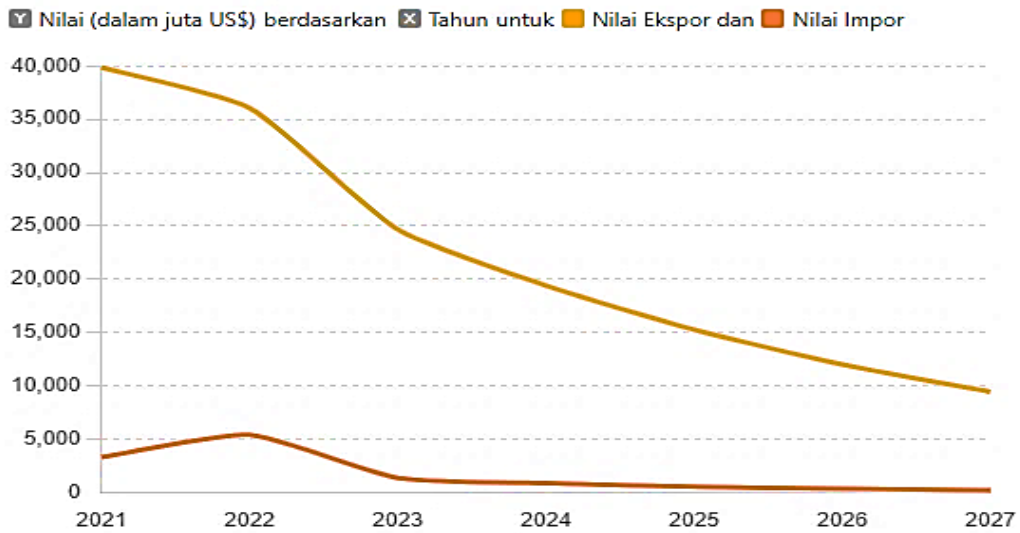
# Plot the historical and projection data
plt.figure(figsize=(10, 6))

plt.plot(all_years, all_exports, label='Nilai Ekspor', marker='o')
plt.plot(all_years, all_imports, label='Nilai Impor', marker='o')

plt.xlabel("Tahun")
plt.ylabel("Nilai (dalam juta US$)")
plt.title("Proyeksi Nilai Ekspor dan Impor Kalimantan Timur (2021-2027)")
plt.legend()

plt.grid(True)
plt.show()
```

Pada gambar 36 merupakan grafik proyeksi nilai ekspor dan impor di Provinsi Kalimantan Timur dari tahun 2021 hingga tahun 2027.



Gambar 36. Grafik Proyeksi Nilai Ekspor-Impor Provinsi Kaltim Tahun 2021-2027

Berdasarkan proyeksi yang dihasilkan, berikut adalah tren ekspor dan impor Kalimantan Timur untuk beberapa tahun ke depan.

Tren Nilai Ekspor: Nilai ekspor menunjukkan penurunan yang cukup konsisten dari tahun 2023 ke tahun 2027 dan jika tidak ada perubahan signifikan dalam kebijakan perdagangan, kondisi ekonomi global, atau peningkatan efisiensi dan diversifikasi pasar, nilai ekspor diproyeksikan akan terus menurun.

Tren Nilai Impor: Nilai impor juga diproyeksikan untuk menurun atau tetap stabil pada tingkat yang lebih rendah dibandingkan dengan tahun 2022 dan penurunan impor bisa disebabkan oleh penurunan kebutuhan bahan baku impor atau peningkatan produksi lokal.

Kesimpulan: untuk memperbaiki atau meningkatkan tren ini, Kalimantan Timur perlu mempertimbangkan beberapa langkah strategis, seperti:

1. Diversifikasi produk ekspor.
2. Peningkatan nilai tambah produk lokal.
3. Peningkatan akses ke pasar internasional baru.
4. Peningkatan infrastruktur dan efisiensi logistik.

Dengan pendekatan yang tepat, Kalimantan Timur dapat membalikkan tren penurunan ini dan meningkatkan kinerja perdagangan internasionalnya di masa depan.

4.5. Sektor Pertambangan

Berikut adalah sektor pertambangan dimana data aktivitas pertambangan di Provinsi Kalimantan Timur dari tahun 2021 sampai dengan tahun 2023.

Tabel 13. Aktivitas Pertambangan di Provinsi Kalimantan Timur Tahun 2021-2023

No	Data Aktivitas	Tahun			Satuan
		2021	2022	2023	
1	Lifting Minyak Bumi	17.742.360	16.697.671	18.312.385	Barel
2	Lifting Gas Alam	172.829.530	163.134.528	175.079.218	MMBTU
3	Produksi Batubara PKP2B	73.959.294	53.620.138	41.015.029	Ton
4	Produksi Batubara KP/IUP	220.293.507	66.392.239	166.428.573	Ton
5	Total Produksi Batubara (PKP2B+KP/IUP)	294.252.801	120.012.377	207.443.602	Ton
6	Jumlah Perusahaan Batubara	1.403	1.403	1.403	Izin
7	Persentase Area Yang Direhabilitasi dari Lahan Terganggu	48	39	39	Persen
8	Reklamasi Batubara	3.544	10.869	10.869	Ha
9	Luas Penambangan Liar yang Diterbitkan	5	75	75	Ha
10	Luas Area Penambangan yang Liar	10	75	5	Ha

Sumber: <https://data.kaltimprov.go.id/dataset>, 2024

Berikut beberapa analisis dari data aktivitas pertambangan di Provinsi Kalimantan Timur dari tahun 2021- 2023 sebagai berikut:

1. Lifting Minyak Bumi:
 - a. Tahun 2021: 17,742,360 Barel
 - b. Tahun 2022: 16,697,671 Barel
 - c. Tahun 2023: 18,312,385 Barel
 - d. Tren: Terjadi penurunan pada tahun 2022 namun meningkat kembali pada tahun 2023.
2. Lifting Gas Alam:

- a. Tahun 2021: 172,829,530 MMBTU
- b. Tahun 2022: 163,134,528 MMBTU
- c. Tahun 2023: 175,079,218 MMBTU
- d. Tren: Mengalami penurunan di tahun 2022 dan peningkatan kembali di tahun 2023.

3. Produksi Batubara PKP2B:

- a. Tahun 2021: 73,959,294 Ton
- b. Tahun 2022: 53,620,138 Ton
- c. Tahun 2023: 41,015,029 Ton
- d. Tren: Terus mengalami penurunan dari tahun 2021 hingga 2023.

4. Produksi Batubara KP/IUP:

- a. Tahun 2021: 220,293,507 Ton
- b. Tahun 2022: 66,392,239 Ton
- c. Tahun 2023: 166,428,573 Ton
- d. Tren: Penurunan tajam di tahun 2022 namun meningkat kembali di tahun 2023.

5. Total Produksi Batubara (PKP2B+KP/IUP):

- a. Tahun 2021: 294,252,801 Ton
- b. Tahun 2022: 120,012,377 Ton
- c. Tahun 2023: 207,443,602 Ton
- d. Tren: Penurunan di tahun 2022 dan peningkatan kembali di tahun 2023.

6. Jumlah Perusahaan Batubara:

- a. Tahun 2021: 1403 Izin

- b. Tahun 2022: 1403 Izin
- c. Tahun 2023: 1403 Izin
- d. Tren: Stabil tidak berubah.

7. Persentase Area Yang Direhabilitasi Dari Lahan Terganggu:

- a. Tahun 2021: 48%
- b. Tahun 2022: 39%
- c. Tahun 2023: 39%
- d. Tren: Mengalami penurunan di tahun 2022 dan stabil di tahun 2023.

8. Reklamasi Batubara:

- a. Tahun 2021: 3544 Ha
- b. Tahun 2022: 10,869 Ha
- c. Tahun 2023: 10,869 Ha
- d. Tren: Meningkatkan signifikan pada tahun 2022 dan stabil di tahun 2023.

9. Luas Penambangan Liar Yang Diterbitkan:

- a. Tahun 2021: 5 Ha
- b. Tahun 2022: 75 Ha
- c. Tahun 2023: 75 Ha
- d. Tren: Meningkatkan pada tahun 2022 dan stabil di tahun 2023.

10. Luas Area Penambangan Yang Liar:

- a. Tahun 2021: 10 Ha
- b. Tahun 2022: 75 Ha

c. Tahun 2023: 5 Ha

d. Tren: Peningkatan tajam pada tahun 2022 namun menurun kembali di tahun 2023.

Berikut ini kode program untuk melakukan proses analisis data dengan pemrograman python:

```
import pandas as pd

# Define the data based on the image provided
data = {
    "DAFTAR AKTIVITAS": [
        "Lifting Minyak Bumi", "Lifting Gas Alam", "Produksi Batubara PKP2B",
        "Produksi Batubara KP/IUP", "Total Produksi Batubara (PKP2B+KP/IUP)",
        "Jumlah Perusahaan Batubara", "Persentase Area Yang Direhabilitasi Dari Lahan Terganggu",
        "Reklamasi Batubara", "Luas Penambangan Liar Yang Diterbitkan",
        "Luas Area Penambangan Yang Liar"
    ],
    "2021": [17742360, 172829530, 73959294, 220293507, 294252801, 1403, 48, 3544, 5, 10],
    "2022": [16697671, 163134528, 53620138, 66392239, 120012377, 1403, 39, 10869, 75, 75],
    "2023": [18312385, 175079218, 41015029, 166428573, 207443602, 1403, 39, 10869, 75, 5],
    "SATUAN": ["Barel", "MMBTU", "Ton", "Ton", "Ton", "Izin", "Persen", "Ha", "Ha", "Ha"]
}

# Create DataFrame
df = pd.DataFrame(data)

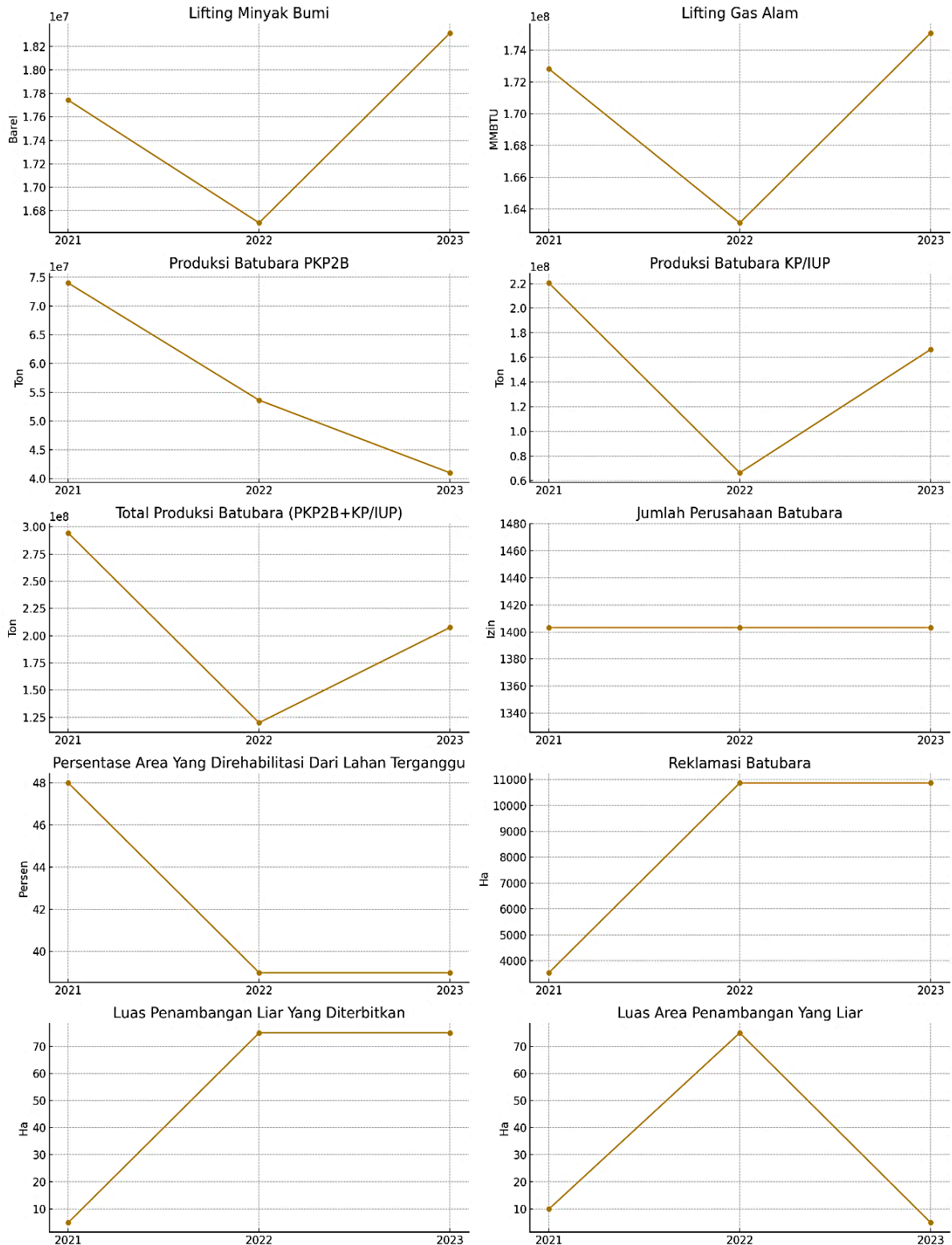
import matplotlib.pyplot as plt

# Function to plot the data
def plot_data(df):
    fig, axs = plt.subplots(5, 2, figsize=(15, 20))
    fig.suptitle('Data Pertambangan Provinsi Kaltim Tahun 2021-2023', fontsize=16)

    for i, ax in enumerate(axs.flat):
        if i < len(df):
            activity = df.loc[i, "DAFTAR AKTIVITAS"]
            values = df.loc[i, ["2021", "2022", "2023"]]
            ax.plot(["2021", "2022", "2023"], values, marker='o')
            ax.set_title(activity)
            ax.set_ylabel(df.loc[i, "SATUAN"])
            ax.grid(True)

plt.tight_layout(rect=[0, 0, 1, 0.96])
plt.show()
plot_data(df)
```

Pada gambar 37 merupakan grafik data pertambangan Provinsi Kalimantan Timur dari tahun 2021 hingga 2023 berdasarkan data yang telah berikan:



Gambar 37. Grafik Data Aktivitas Pertambangan Provinsi Kaltim Tahun 2021- 2023.

Berdasarkan tren data yang diberikan, berikut adalah beberapa rekomendasi untuk masing-masing aspek:

1. Lifting minyak bumi: rekomendasi yang diberikan yaitu mengingat fluktuasi lifting minyak bumi, sebaiknya dilakukan evaluasi mendalam terhadap faktor-faktor penyebab penurunan dan peningkatan produksi. Langkah mitigasi yang efektif perlu disiapkan untuk menjaga stabilitas produksi.
2. Lifting gas alam: rekomendasi yang diberikan yaitu perlu ada upaya untuk menjaga kestabilan produksi gas alam. Peningkatan pada tahun 2023 perlu dipertahankan dengan perencanaan jangka panjang yang matang, termasuk investasi pada teknologi eksplorasi dan produksi yang lebih efisien.
3. Produksi batubara PKP2B dan KP/IUP: rekomendasi yang diberikan yaitu mengingat penurunan produksi batubara, diperlukan diversifikasi sumber daya dan teknologi yang lebih efisien. Investasi pada energi terbarukan juga bisa menjadi alternatif untuk mengurangi ketergantungan pada batubara.
4. Total produksi batubara (PKP2B+KP/IUP): rekomendasi yang diberikan yaitu mengelola produksi batubara secara lebih berkelanjutan dengan mengedepankan praktik penambangan yang ramah lingkungan. Perlu juga dipertimbangkan kebijakan untuk memperbaiki regulasi yang mendukung efisiensi dan keberlanjutan produksi.
5. Jumlah perusahaan batubara: rekomendasi yang diberikan yaitu meskipun jumlah perusahaan batubara stabil, perlu dilakukan penilaian kinerja secara berkala untuk memastikan setiap perusahaan mematuhi standar operasional dan lingkungan yang ditetapkan.
6. Persentase area yang direhabilitasi dari lahan terganggu: rekomendasi yang diberikan yaitu peningkatan upaya rehabilitasi lahan terganggu sangat diperlukan. Program rehabilitasi harus ditingkatkan dengan insentif bagi perusahaan yang berhasil memulihkan lahan.
7. Reklamasi batubara: rekomendasi yang diberikan yaitu meningkatkan reklamasi lahan harus menjadi prioritas. Perluas kerjasama dengan institusi penelitian dan

pengembangan teknologi yang fokus pada reklamasi dapat membantu dalam mencapai target yang lebih tinggi.

8. Luas penambangan liar yang diterbitkan dan luas area penambangan yang liar: rekomendasi yang diberikan yaitu penegakan hukum yang lebih tegas dan konsisten terhadap penambangan liar diperlukan. Selain itu, edukasi kepada masyarakat tentang dampak negatif penambangan liar serta memberikan alternatif ekonomi yang legal bisa membantu mengurangi aktivitas ilegal ini.

Secara keseluruhan, untuk meningkatkan keberlanjutan sektor pertambangan di Provinsi Kaltim, penting untuk mengintegrasikan teknologi modern, praktik ramah lingkungan, serta memperkuat regulasi dan pengawasan.

4.6. Sektor Transportasi

Berikut analisis data terkait transportasi darat yakni jumlah terminal bis, jumlah penumpang tiba dan penumpang berangkat dari tahun 2016 sampai dengan tahun 2020 dan data kecelakaan lalu lintas di Provinsi Kalimantan Timur.

4.6.1. Transportasi Darat di Kaltim

Berikut adalah sektor transportasi dimana data transportasi darat yakni jumlah bus, jumlah penumpang tiba dan penumpang berangkat di Provinsi Kalimantan Timur dari tahun 2016 sampai dengan tahun 2020.

Tabel 14. Data Transportasi Darat (Bus dan Terminal Bus) di Provinsi Kalimantan Timur Tahun 2016-2020

Tahun	Jumlah Terminal Bis (Unit)	Jumlah Penumpang Bis Tiba (Penumpang)	Jumlah Penumpang Bis Berangkat (Penumpang)
2016	23	9.859.962	9.845.714
2017	23	9.227.136	8.966.784
2018	23	274.578	327.594
2019	21	233.334	272.028
2020	21	110.262	112.304

Sumber: <https://data.kaltimprov.go.id/dataset>

Berikut ini kode program python dalam melakukan proses analisis data:

```
import pandas as pd

# Create a dataframe from the provided data
data = {
    "Tahun": [2016, 2017, 2018, 2019, 2020],
    "Jumlah Terminal Bis (Unit)": [23, 23, 23, 23, 21],
    "Jumlah Penumpang Bis Tiba (Penumpang)": [9859962, 9227136, 274578, 233334, 110262],
    "Jumlah Penumpang Bis Berangkat (Penumpang)": [9845714, 8966784, 327594, 272028, 112304]
}

df = pd.DataFrame(data)

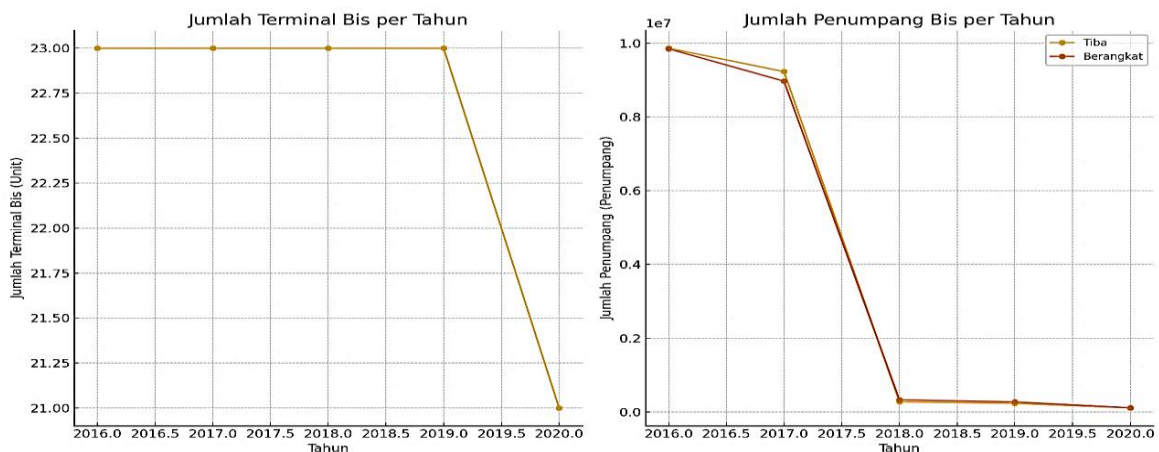
import matplotlib.pyplot as plt

# Plot the data
plt.figure(figsize=(14, 7))

# Plot Jumlah Terminal Bis
plt.subplot(1, 2, 1)
plt.plot(df['Tahun'], df['Jumlah Terminal Bis (Unit)'], marker='o')
plt.title('Jumlah Terminal Bis per Tahun')
plt.xlabel('Tahun')
plt.ylabel('Jumlah Terminal Bis (Unit)')

# Plot Jumlah Penumpang Bis Tiba dan Berangkat
plt.subplot(1, 2, 2)
plt.plot(df['Tahun'], df['Jumlah Penumpang Bis Tiba (Penumpang)'], marker='o', label='Tiba')
plt.plot(df['Tahun'], df['Jumlah Penumpang Bis Berangkat (Penumpang)'], marker='o', label='Berangkat')
plt.title('Jumlah Penumpang Bis per Tahun')
plt.xlabel('Tahun')
plt.ylabel('Jumlah Penumpang (Penumpang)')
plt.legend()

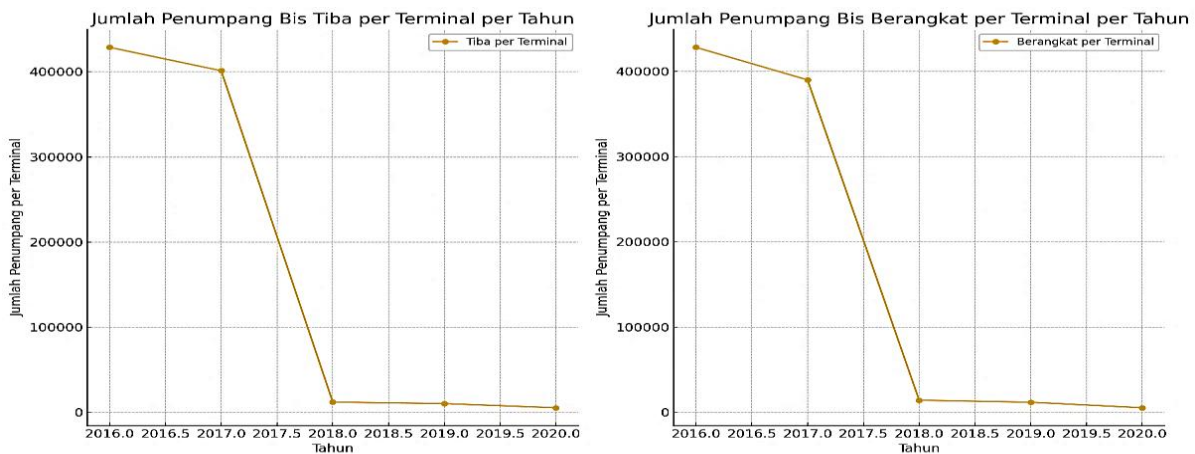
plt.tight_layout()
plt.show()
```



Gambar 38. Grafik Tren Jumlah Terminal dan Jumlah Penumpang per Tahun di Provinsi Kalim Tahun 2016-2020.

Grafik pertama pada gambar 38 menunjukkan jumlah terminal bis per tahun yang cukup stabil hingga 2019 sebelum turun pada 2020. Grafik kedua menunjukkan penurunan jumlah penumpang bis tiba dan berangkat dari tahun 2017 ke 2018, dan penurunan ini terus berlanjut hingga 2020. Berikut adalah analisis dan grafik pada gambar 38, berdasarkan data yang berikan:

- a. Jumlah Terminal Bis (Unit): pada tahun 2016 hingga 2019, jumlah terminal bis tetap konsisten sebanyak 23 unit. Kemudian pada tahun 2020, jumlah terminal bis berkurang menjadi 21 unit.
- b. Jumlah Penumpang Bis Tiba (Penumpang): terdapat penurunan drastis dari tahun 2017 ke tahun 2018, dari 9.227.136 penumpang menjadi hanya 274.578 penumpang. Tren penurunan ini terus berlanjut hingga tahun 2020.
- c. Jumlah Penumpang Bis Berangkat (Penumpang): mirip dengan jumlah penumpang tiba, jumlah penumpang berangkat juga mengalami penurunan drastis dari tahun 2017 ke tahun 2018, dari 8.966.784 penumpang menjadi 327.594 penumpang. Kemudian Tren penurunan ini juga berlanjut hingga tahun 2020.



Gambar 39. Grafik Tren Jumlah Penumpang Tiba dan Jumlah Penumpang Berangkat per Tahun di Provinsi Kaltim Tahun 2016-2020.

Grafik pada gambar 39 menunjukkan jumlah penumpang per terminal dari tahun 2016 hingga 2020, baik untuk penumpang tiba maupun penumpang berangkat. Jumlah Penumpang Bis Tiba per Terminal: pada tahun 2016 dan 2017, jumlah penumpang tiba per terminal lebih dari 400.000. dan jumlah ini mengalami penurunan tajam pada tahun 2018 dan terus menurun hingga 2020. Selanjutnya jumlah Penumpang Bis Berangkat per Terminal: tren serupa juga

terlihat pada jumlah penumpang berangkat per terminal, dengan penurunan signifikan dimulai pada tahun 2018 dan berlanjut hingga 2020.

4.6.2. Kecelakaan Lalu Lintas di Kaltim

Berikut data kecelakaan lalu lintas per wilayah di provinsi Kalimantan Timur dari tahun 2016 sampai dengan 2020.

Tabel 15. Data Kecelakaan Lalu Lintas per Wilayah di Kaltim Tahun 2016-2020

No	Wilayah	Tahun				
		2016	2017	2018	2019	2020
1	Polresta Samarinda	84	171	143	109	96
2	Polres Balikpapan	45	112	124	88	47
3	Polres Bontang	44	58	62	40	56
4	Polres Kukar	133	171	117	103	77
5	Polres Kutim	61	71	39	27	23
6	Polres Kubar	23	86	41	34	22
7	Polres Berau	33	47	80	55	55
8	Polres Paser	85	107	68	71	41
9	Polres PPU	43	41	49	43	31

Sumber: <https://data.kaltimprov.go.id/dataset>

Berikut ini kode program python dalam melakukan proses analisis data.

```
import pandas as pd

# Create a DataFrame from the provided data
data = {
    "Wilayah": ["Polresta Samarinda", "Polres Balikpapan", "Polres Bontang", "Polres Kukar",
               "Polres Kutim", "Polres Kubar", "Polres Berau", "Polres Paser", "Polres PPU"],
    "2016": [84, 45, 44, 133, 61, 23, 33, 85, 43],
    "2017": [171, 112, 58, 171, 71, 86, 47, 107, 41],
    "2018": [143, 124, 62, 117, 39, 41, 80, 68, 49],
    "2019": [109, 88, 40, 103, 27, 34, 55, 71, 43],
    "2020": [96, 47, 56, 77, 23, 22, 55, 41, 31]
}

df = pd.DataFrame(data)
df.set_index("Wilayah", inplace=True)

import matplotlib.pyplot as plt

# Create line plots for the data
plt.figure(figsize=(14, 8))
```

```

for column in df.columns:
    plt.plot(df.index, df[column], marker='o', label=column)

plt.title('Jumlah Kecelakaan Lalu Lintas 2016-2020')
plt.xlabel('Wilayah')
plt.ylabel('Jumlah Kecelakaan')
plt.legend(title='Tahun')
plt.grid(True)
plt.xticks(rotation=45)
plt.tight_layout()

# Show the plot
plt.show()

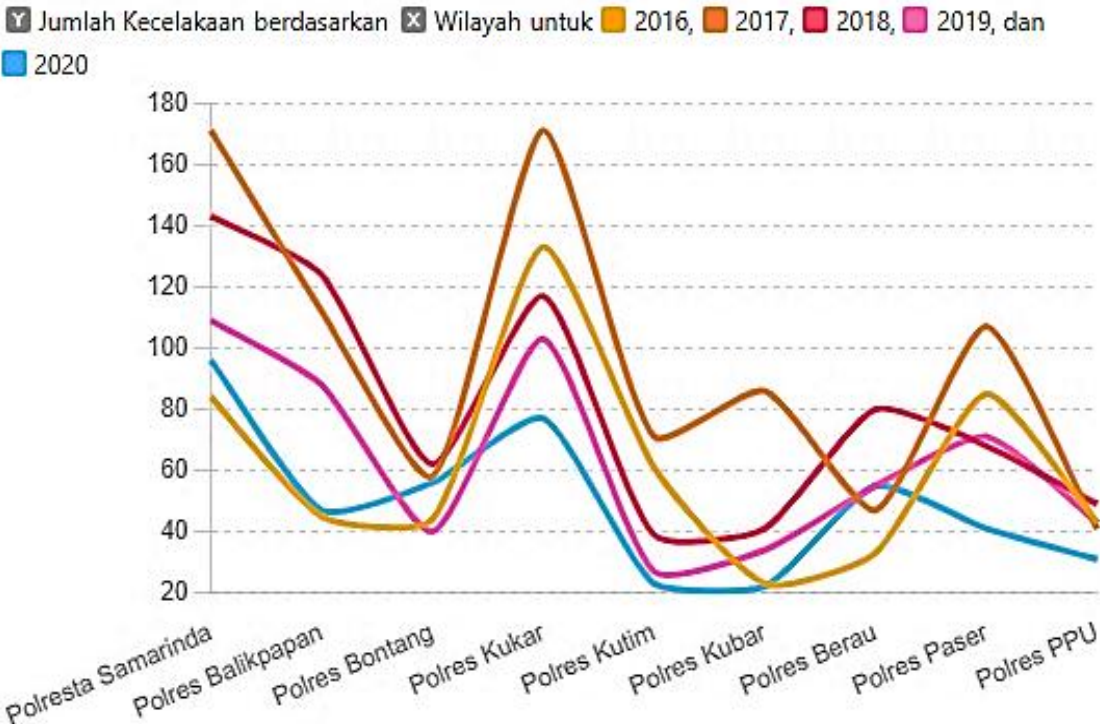
import ace_tools as tools; tools.display_dataframe_to_user(name="Jumlah Kecelakaan Lalu Lintas 2016-2020",
dataframe=df)

```

Selanjutnya melakukan analisis dari data jumlah kecelakaan lalu lintas di wilayah Kalimantan Timur dari tahun 2016 hingga 2020 sebagai berikut.

1. Polresta Samarinda: Terjadi peningkatan signifikan pada tahun 2017, namun menurun bertahap dari tahun 2018 hingga 2020.
2. Polres Balikpapan: Peningkatan jumlah kecelakaan dari tahun 2016 ke 2017 dan 2018, namun menurun drastis pada tahun 2019 dan 2020.
3. Polres Bontang: Jumlah kecelakaan relatif stabil dengan sedikit fluktuasi dari tahun ke tahun, dengan peningkatan sedikit pada tahun 2020.
4. Polres Kukar: Tinggi pada tahun 2016 dan 2017, kemudian mengalami penurunan pada tahun-tahun berikutnya.
5. Polres Kutim: Penurunan jumlah kecelakaan setiap tahunnya dari tahun 2017 hingga 2020.
6. Polres Kubar: Peningkatan tajam pada tahun 2017, kemudian menurun tajam dan stabil di angka rendah pada tahun-tahun berikutnya.
7. Polres Berau: Jumlah kecelakaan meningkat signifikan pada tahun 2018, namun stabil pada tahun 2019 dan 2020.
8. Polres Paser: Mengalami fluktuasi namun cenderung menurun pada tahun 2019 dan 2020.
9. Polres PPU: Relatif stabil dengan sedikit penurunan pada tahun 2020.

Pada gambar 40 grafik di bawah ini menunjukkan tren kecelakaan di setiap wilayah dari tahun 2016 hingga 2020, dengan Polresta Samarinda dan Polres Kukar mencatat jumlah kecelakaan tertinggi secara umum, sementara Polres PPU menunjukkan jumlah kecelakaan terendah.



Gambar 40. Grafik Tren Jumlah Kecelakaan Lalu Lintas per Wilayah di Kaltim Tahun 2016-2020.

Selanjutnya melakukan proses analisis data dengan pemrograman python, berikut kode programnya.

```

# Calculate the yearly average trend
average_trend = df.mean(axis=0)

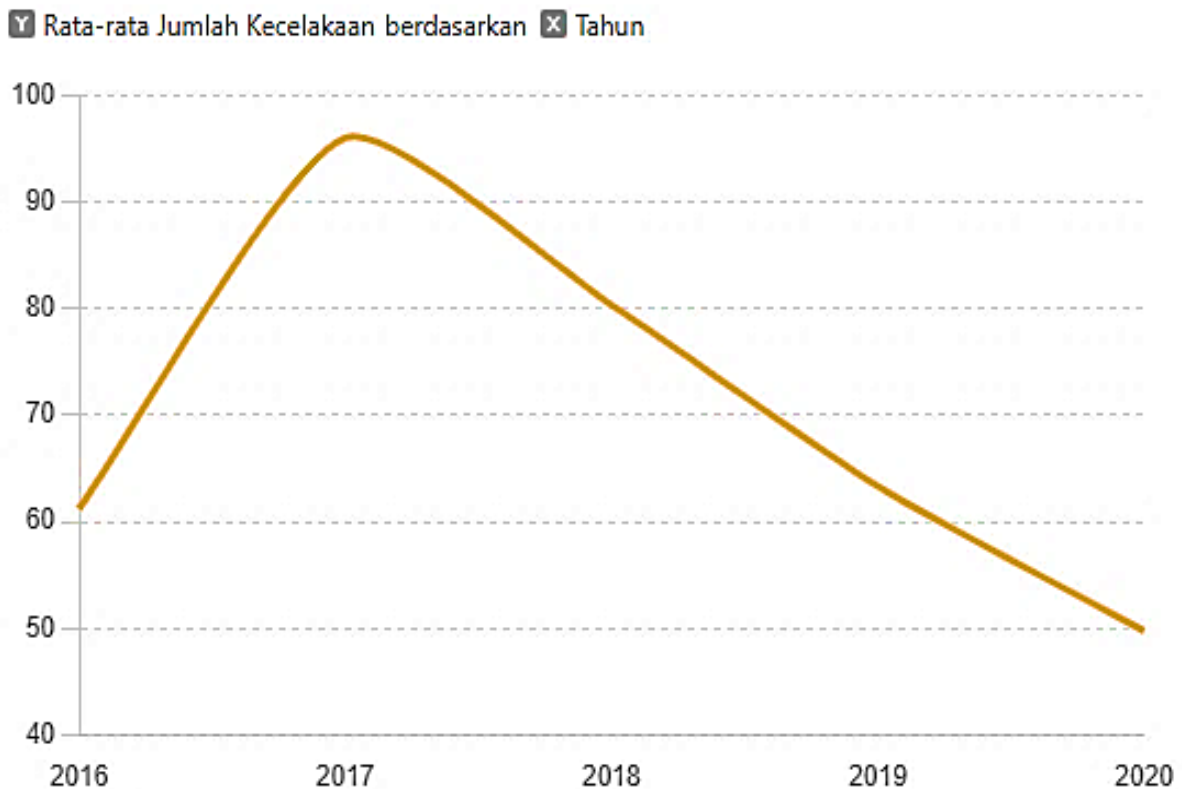
# Plot the yearly average trend
plt.figure(figsize=(10, 6))
plt.plot(average_trend.index, average_trend.values, marker='o')
plt.title('Tren Rata-rata Tahunan Jumlah Kecelakaan Lalu Lintas (2016-2020)')
plt.xlabel('Tahun')
plt.ylabel('Rata-rata Jumlah Kecelakaan')
plt.grid(True)
plt.tight_layout()

# Show the plot
plt.show()

average_trend

```

Pada gambar 41 terlihat grafik menunjukkan tren rata-rata tahunan kecelakaan lalu lintas di wilayah Provinsi Kalimantan Timur dari tahun 2016 hingga tahun 2020.



Gambar 41. Grafik Tren Rata-Rata Tahunan Kecelakaan Lalu Lintas di Provinsi Kalimantan Timur Tahun 2016-2020.

Berikut adalah tren rata-rata tahunan jumlah kecelakaan lalu lintas dari tahun 2016 hingga 2020:

- a. 2016: Rata-rata 61,22 kecelakaan
- b. 2017: Rata-rata 96,00 kecelakaan
- c. 2018: Rata-rata 80,33 kecelakaan
- d. 2019: Rata-rata 63,33 kecelakaan
- e. 2020: Rata-rata 49,78 kecelakaan

4.7. Sektor Perkebunan

Berikut ini adalah data pada sektor perkebunan berupa luas areal tanaman perkebunan rakyat menurut jenis tanaman (Ha) dan produksi perkebunan rakyat menurut jenis tanaman di Provinsi Kalimantan Timur (Ton) dari tahun 2019 sampai dengan tahun 2023.

4.7.1. Luas Areal Tanaman Perkebunan Rakyat Menurut Jenis Tanaman

Berikut data dari luas areal tanaman perkebunan rakyat menurut jenis tanaman di Provinsi Kalimantan Timur (ha) dari tahun 2019 hingga tahun 2023, sebagai berikut:

Tabel 16. Luas Areal Tanaman Perkebunan Rakyat Menurut Jenis Tanaman di Provinsi Kaltim (ha) Tahun 2019–2023.

Jenis Tanaman	Tahun (Ha)				
	2019	2020	2021	2022	2023
Karet/Rubber	92.639	92.982	92.403	92.403	94.596
Kelapa/Coconut	21.152	20.698	20.018	20.018	20.068
Kelapa sawit/Oil palm	255.919	286.058	252.000	252.000	373.212
Kopi/Coffee	2.529	1.957	1.488	1.488	1.495
Kakao/Cocoa	7.328	6.883	7.617	7.617	7.777
Jambu mete/Cashew nut	8	4	1	1	0
Pala/Nutmeg	305	257	322	322	422
Lada/Pepper	8.921	8.247	8.221	8.221	8.321

Sumber: <https://data.kaltimprov.go.id/dataset>

Selanjutnya melakukan proses analisis data dengan pemrograman python dan menciptakan grafik tren dari data tersebut, berikut kode program untuk memproses data dan membuat beberapa grafik untuk memvisualisasikan perubahan luas areal tanaman perkebunan rakyat dari tahun 2019 hingga 2023

```
import pandas as pd
import matplotlib.pyplot as plt

# Membuat DataFrame dari data yang ada
data = {
    "Jenis Tanaman": [
        "Karet/Rubber", "Kelapa/Coconut", "Kelapa sawit/Oil palm",
        "Kopi/Coffee", "Kakao/Cocoa", "Jambu mete/Cashew nut",
        "Pala/Nutmeg", "Lada/Pepper"
    ],
    "2019": [92639, 21152, 255919, 2529, 7328, 8, 305, 8921],
    "2020": [92982, 20696, 286058, 1957, 6883, 4, 257, 8247],
    "2021": [92403, 20018, 252000, 1488, 7617, 1, 332, 8221],
```



```

"2022": [92403, 20018, 252000, 1488, 7617, 1, 322, 8221],
"2023": [94596, 20068, 373212, 1495, 7777, 0, 422, 8321]
}

df = pd.DataFrame(data)
df.set_index("Jenis Tanaman", inplace=True)

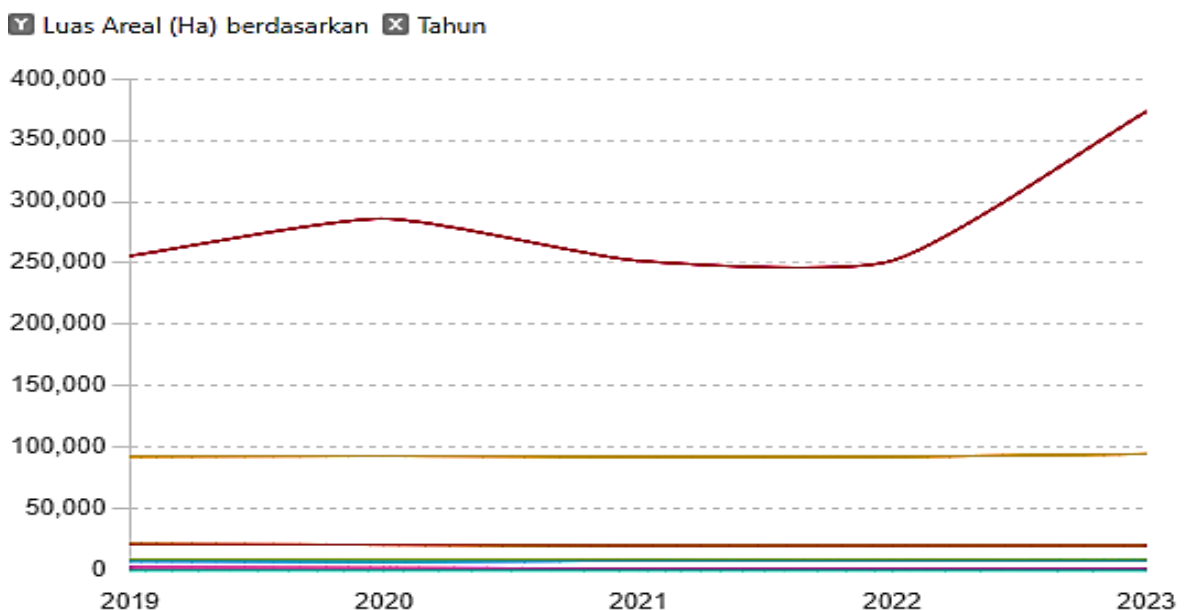
# Menampilkan DataFrame
import ace_tools as tools; tools.display_dataframe_to_user(name="Luas Areal Tanaman Perkebunan Rakyat",
dataframe=df)

# Plot data untuk setiap jenis tanaman
plt.figure(figsize=(14, 8))
for column in df.columns:
    plt.plot(df.columns, df.loc["Karet/Rubber"], label="Karet/Rubber")
    plt.plot(df.columns, df.loc["Kelapa/Coconut"], label="Kelapa/Coconut")
    plt.plot(df.columns, df.loc["Kelapa sawit/Oil palm"], label="Kelapa sawit/Oil palm")
    plt.plot(df.columns, df.loc["Kopi/Coffee"], label="Kopi/Coffee")
    plt.plot(df.columns, df.loc["Kakao/Cocoa"], label="Kakao/Cocoa")
    plt.plot(df.columns, df.loc["Jambu mete/Cashew nut"], label="Jambu mete/Cashew nut")
    plt.plot(df.columns, df.loc["Pala/Nutmeg"], label="Pala/Nutmeg")
    plt.plot(df.columns, df.loc["Lada/Pepper"], label="Lada/Pepper")

plt.xlabel("Tahun")
plt.ylabel("Luas Areal (Ha)")
plt.title("Luas Areal Tanaman Perkebunan Rakyat Menurut Jenis Tanaman")
plt.legend()
plt.grid(True)
plt.show()

```

Berikut adalah grafik yang menunjukkan perubahan luas areal tanaman perkebunan rakyat menurut jenis tanaman dari tahun 2019 hingga 2023.



Gambar 42. Grafik Luas Areal Tanaman Perkebunan Rakyat Menurut Jenis Tanaman di Provinsi Kaltim (ha) Tahun 2019–2023.

Berikut ini adalah kesimpulan hasil dari analisis data yaitu: Karet/Rubber: luas areal relatif stabil dengan sedikit peningkatan pada tahun 2023. Sedangkan Kelapa/Coconut: terdapat penurunan dari tahun 2019 ke 2021, tetapi stabil dari 2021 hingga 2023. Selanjutnya Kelapa sawit/Oil palm: mengalami peningkatan signifikan pada tahun 2023. Kemudian Kopi/Coffee: terjadi penurunan signifikan pada tahun 2020, tetapi stabil setelahnya. Sedangkan Kakao/Cocoa: luas areal fluktuatif namun menunjukkan sedikit peningkatan pada tahun 2023. Berikutnya Jambu mete/Cashew nut: hampir tidak ada luas areal yang signifikan. Kemudian Pala/Nutmeg: mengalami peningkatan bertahap dari tahun 2019 hingga 2023. Dan terakhir Lada/Pepper: relatif stabil dengan sedikit penurunan pada tahun 2020 dan 2021, namun stabil setelahnya.

4.7.2. Produksi Perkebunan Rakyat Menurut Jenis Tanaman

Berikut ini merupakan data produksi perkebunan rakyat menurut jenis tanaman di Provinsi Kalimantan Timur (ton) Tahun 2019–2023, sebagai berikut:

Tabel 17. Produksi Perkebunan Rakyat Menurut Jenis Tanaman di Provinsi Kaltim (ton) Tahun 2019–2023

Jenis Tanaman	Tahun (Ton)				
	2019	2020	2021	2022	2023
Karet/Rubber	50.880	52.801	66.872	66.872	69.446
Kelapa/Coconut	11.013	7.662	10.170	10.170	7.201
Kelapa sawit/Oil palm	506.370	550.473	561.146	561.146	519.497
Kopi/Coffee	223	262	167	167	165
Kakao/Cocoa	2.513	2.537	2.181	2.181	2.566
Pala/Nutmeg	0,1	0,15	11	11	12
Lada/Pepper	5.799	3.760	4.173	4.173	5.080

Sumber: <https://data.kaltimprov.go.id/dataset>, 2024

Selanjutnya, melakukan analisis dan membuat grafik berdasarkan data tabel 17 di atas. Memulai dengan analisis data sebagai berikut:

1. Karet/Rubber: Produksi karet meningkat dari tahun 2019 hingga 2023, dengan peningkatan terbesar terjadi antara tahun 2020 dan 2021.

2. Kelapa/Coconut: Produksi kelapa fluktuatif, dengan penurunan signifikan pada tahun 2023.
3. Kelapa sawit/Oil palm: Produksi kelapa sawit cukup stabil dengan sedikit penurunan pada tahun 2023.
4. Kopi/Coffee: Produksi kopi menurun dari tahun 2020 dan tetap konstan pada tahun 2022 dan 2023.
5. Kakao/Cocoa: Produksi kakao sedikit menurun pada tahun 2021 dan 2022, kemudian naik kembali pada tahun 2023.
6. Pala/Nutmeg: Produksi pala meningkat signifikan pada tahun 2021 dan tetap konstan hingga tahun 2023.
7. Lada/Pepper: Produksi lada menurun pada tahun 2020 dan meningkat lagi pada tahun 2023.

Kemudian dilanjutkan dengan membuat grafik untuk visualisasi data produksi untuk masing-masing jenis tanaman. Berikut ini kode program dalam proses analisis data dengan pemrograman python:

```
import matplotlib.pyplot as plt
import pandas as pd

# Data transcribed from the image
data = {
    "Jenis Tanaman": ["Karet/Rubber", "Kelapa/Coconut", "Kelapa sawit/Oil palm", "Kopi/Coffee", "Kakao/Cocoa",
    "Pala/Nutmeg", "Lada/Pepper"],
    "2019": [50880, 11013, 506370, 223, 2513, 0.1, 5799],
    "2020": [52801, 7662, 550473, 262, 2537, 0.15, 3760],
    "2021": [66872, 10170, 561146, 167, 2181, 11, 4173],
    "2022": [66872, 10170, 561146, 167, 2181, 11, 4173],
    "2023": [69446, 7201, 519497, 167, 2566, 11, 5080]
}

df = pd.DataFrame(data)

# Plotting the data
fig, axs = plt.subplots(2, 1, figsize=(10, 12))

# Plot 1: Karet, Kelapa, Kelapa sawit
df_plot1 = df[df["Jenis Tanaman"].isin(["Karet/Rubber", "Kelapa/Coconut", "Kelapa sawit/Oil palm"])].set_index("Jenis Tanaman").T
df_plot1.plot(ax=axs[0], marker='o')
axs[0].set_title("Produksi Karet, Kelapa, dan Kelapa sawit (2019-2023)")
```

```

axs[0].set_xlabel('Tahun')
axs[0].set_ylabel('Produksi (Ton)')
axs[0].grid(True)

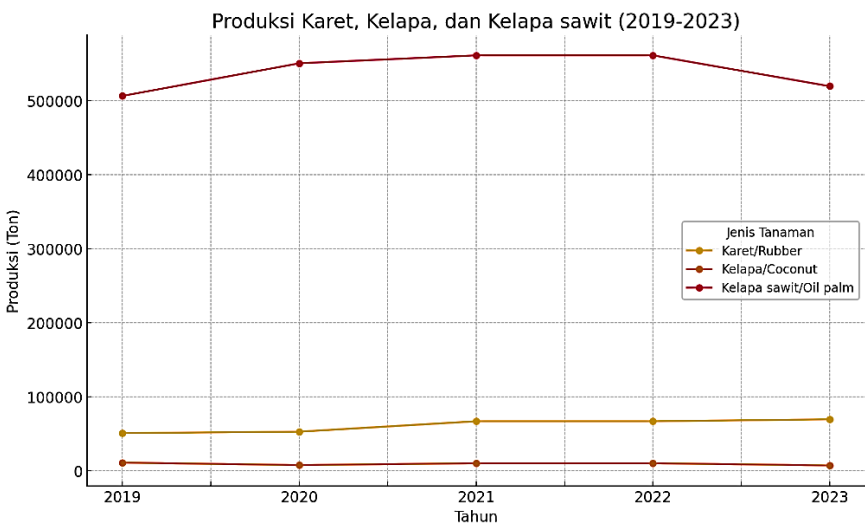
# Plot 2: Kopi, Kakao, Pala, Lada
df_plot2 = df[df['Jenis Tanaman'].isin(['Kopi/Coffee', 'Kakao/Cocoa', 'Pala/Nutmeg', 'Lada/Pepper'])].set_index('Jenis Tanaman').T
df_plot2.plot(ax=axs[1], marker='o')
axs[1].set_title('Produksi Kopi, Kakao, Pala, dan Lada (2019-2023)')
axs[1].set_xlabel('Tahun')
axs[1].set_ylabel('Produksi (Ton)')
axs[1].grid(True)

plt.tight_layout()
plt.show()

```

Pada gambar 43 grafik pertama menunjukkan produksi Karet, Kelapa, dan Kelapa Sawit. Dari grafik ini, kita dapat melihat bahwa:

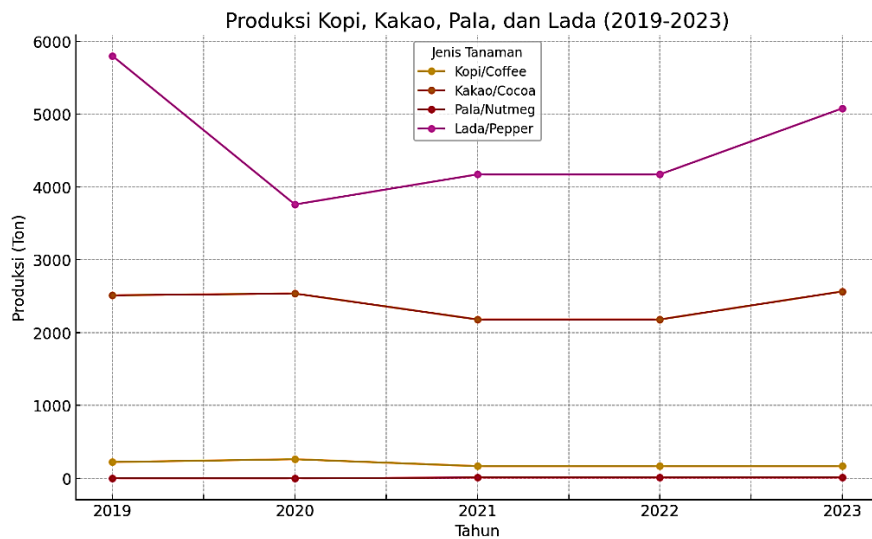
- a. Produksi Karet/Rubber mengalami peningkatan stabil dari tahun 2019 hingga 2023.
- b. Produksi Kelapa/Coconut menunjukkan fluktuasi yang cukup besar, dengan penurunan signifikan pada tahun 2023.
- c. Produksi Kelapa Sawit/Oil palm tetap stabil dengan sedikit penurunan pada tahun 2023.



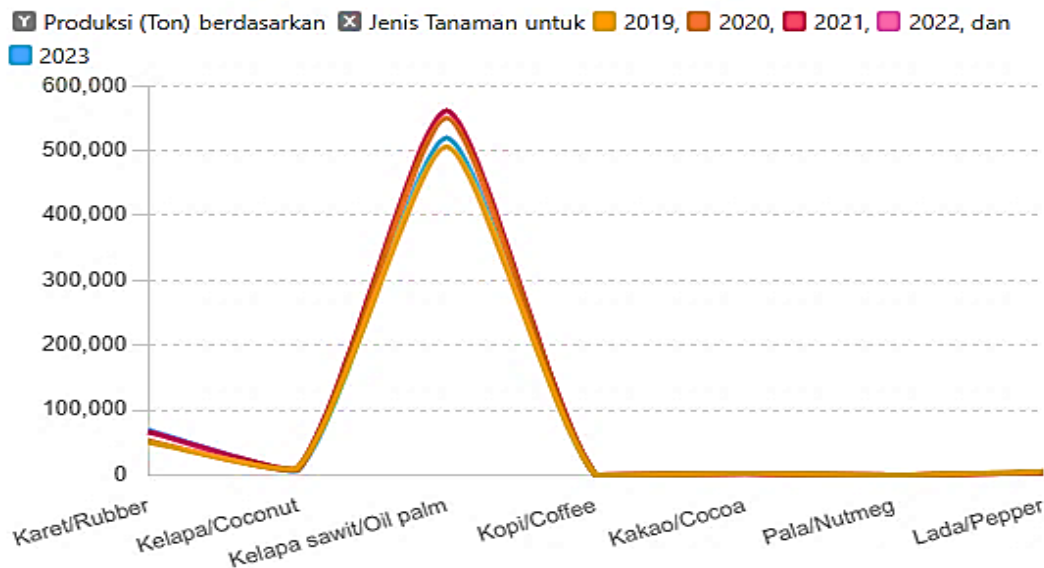
Gambar 43. Grafik Produksi Karet, Kelapa, dan Kelapa Sawit (2019-2023)

Sedangkan pada gambar 44 grafik menunjukkan produksi Kopi, Kakao, Pala, dan Lada. Dari grafik ini, kita dapat melihat bahwa:

- a. Produksi Kopi/Coffee menurun dari tahun 2020 dan tetap konstan pada tahun 2022 dan 2023.
- b. Produksi Kakao/Cocoa mengalami penurunan pada tahun 2021 dan 2022, kemudian naik kembali pada tahun 2023.
- c. Produksi Pala/Nutmeg meningkat signifikan pada tahun 2021 dan tetap konstan hingga tahun 2023.
- d. Produksi Lada/Pepper menurun pada tahun 2020 dan meningkat lagi pada tahun 2023.



Gambar 44. Grafik Produksi Kopi, Kakao, Pala, dan Lada (2019-2023)



Gambar 45. Grafik Tren Produksi Tiap Tanaman dari Tahun 2019 hingga Tahun 2023.

Beberapa poin penting dari analisis ini adalah:

1. Karet/Rubber: Produksi meningkat secara konsisten dari tahun 2019 hingga 2023.
2. Kelapa/Coconut: Produksi mengalami fluktuasi, dengan penurunan signifikan pada tahun 2020 dan penurunan lagi pada tahun 2023 setelah sempat naik pada tahun 2021.
3. Kelapa sawit/Oil palm: Produksi mencapai puncak pada tahun 2021 dan mengalami penurunan signifikan pada tahun 2023.
4. Kopi/Coffee: Produksi menurun sejak tahun 2021.
5. Kakao/Cocoa: Produksi menurun dari tahun 2019 ke 2022, namun naik kembali pada tahun 2023.
6. Pala/Nutmeg: Mulai diproduksi pada tahun 2021 dengan sedikit peningkatan pada tahun 2023.
7. Lada/Pepper: Produksi menurun pada tahun 2020 dan 2021, namun meningkat kembali pada tahun 2023

Berikut ini kode program untuk proses analisis data.

```
from sklearn.linear_model import LinearRegression
import numpy as np

# Menyiapkan data
years = np.array([2019, 2020, 2021, 2022, 2023]).reshape(-1, 1)
predictions = {}

# Melakukan prediksi untuk setiap jenis tanaman
for plant in df.index:
    model = LinearRegression()
    production = df.loc[plant].values
    model.fit(years, production)
    predicted_2024 = model.predict(np.array([[2024]]))
    predictions[plant] = predicted_2024[0]

# Mengubah hasil prediksi ke dalam DataFrame untuk visualisasi
pred_df = pd.DataFrame.from_dict(predictions, orient='index', columns=['Prediksi 2024'])
pred_df.index.name = 'Jenis Tanaman'

# Menampilkan hasil prediksi
pred_df
```

Selanjutnya menjalankan kode program python untuk melihat prediksi produksi tahun 2024:

```

from sklearn.linear_model import LinearRegression
import numpy as np

# Menyiapkan data
years = np.array([2019, 2020, 2021, 2022, 2023]).reshape(-1, 1)
predictions = {}

# Melakukan prediksi untuk setiap jenis tanaman
for plant in df.index:
    model = LinearRegression()
    production = df.loc[plant].values
    model.fit(years, production)
    predicted_2024 = model.predict(np.array([[2024]]))
    predictions[plant] = predicted_2024[0]

# Mengubah hasil prediksi ke dalam DataFrame untuk visualisasi
pred_df = pd.DataFrame.from_dict(predictions, orient='index', columns=['Prediksi 2024'])
pred_df.index.name = 'Jenis Tanaman'

import ace_tools as tools; tools.display_dataframe_to_user(name="Prediksi Produksi Tanaman 2024",
dataframe=pred_df)

# Menampilkan hasil prediksi
pred_df

```

Tabel 18. Prediksi Produksi Tanaman 2024

Jenis Tanaman	Prediksi 2024
Karet/Rubber	76.735,10
Kelapa/Coconut	7.708,40
Kelapa sawit/Oil palm	550.804,50
Kopi/Coffee	133,50
Kakao/Cocoa	2.320,60
Pala/Nutmeg	17,30
Lada/Pepper	4.289,50

Hasil prediksi produksi tanaman pada tahun 2024 berdasarkan analisis regresi linear:

- 1) Karet/Rubber: 76.735,10 ton
- 2) Kelapa/Coconut: 7.708,40 ton
- 3) Kelapa sawit/Oil palm: 550.804,50 ton
- 4) Kopi/Coffee: 133,50 ton

- 5) Kakao/Cocoa: 2.320,60 ton
- 6) Pala/Nutmeg: 17,30 ton
- 7) Lada/Pepper: 4.289,50 ton

Interpretasi prediksi:

- 1) Karet/Rubber yaitu dimana akan diperkirakan produksi akan terus meningkat menjadi sekitar 76.735,10 ton pada tahun 2024.
- 2) Kelapa/Coconut yaitu dimana hasil produksi diperkirakan akan mengalami sedikit peningkatan menjadi sekitar 7.708,40 ton pada tahun 2024.
- 3) Kelapa Sawit/Oil Palm yaitu dimana hasil produksi diperkirakan akan meningkat kembali menjadi sekitar 550.804,50 ton pada tahun 2024.
- 4) Kopi/Coffee yaitu dimana hasil produksi diperkirakan akan sedikit menurun menjadi sekitar 133,50 ton pada tahun 2024.
- 5) Kakao/Cocoa yakni dimana hasil produksi diperkirakan akan sedikit meningkat menjadi sekitar 2.320,60 ton pada tahun 2024.
- 6) Pala/Nutmeg yakni dimana hasil produksi diperkirakan akan meningkat sedikit menjadi sekitar 17,30 ton pada tahun 2024.
- 7) Lada/Pepper yaitu dimana hasil produksi diperkirakan akan menurun sedikit menjadi sekitar 4.289,50 ton pada tahun 2024.

Dengan data science, kita mengubah data menjadi wawasan, wawasan menjadi tindakan, dan tindakan menjadi kemajuan dalam Portal Satu Data

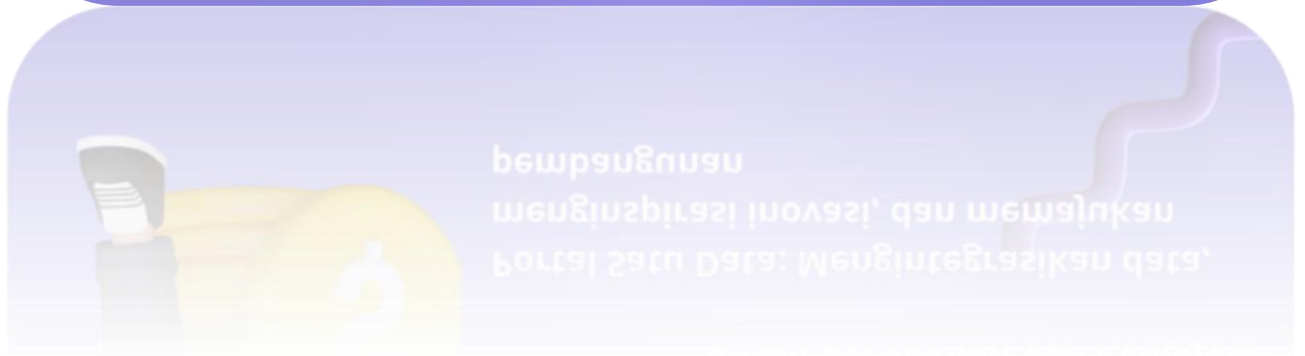


Data science adalah kunci untuk membuka potensi tersembunyi dalam setiap angka, membawa transformasi nyata melalui Portal Satu Data

Portal Satu Data adalah jembatan menuju transparansi, efisiensi, dan kemajuan yang berkelanjutan melalui kekuatan data sains

Di tangan yang tepat, data menjadi lebih dari sekadar informasi; ia menjadi alat untuk perubahan dan kemajuan

Portal Satu Data: Mengintegrasikan data, menginspirasi inovasi, dan memajukan pembangunan



BAB 5

PENUTUP

5.1. Kesimpulan

Kesimpulan dari Kajian *Big Data* dan *Data Analytic*: Urgensi *Data Science* dalam Transformasi Digital Satu Data Kalimantan Timur yaitu:

1. Transformasi digital adalah kunci bagi Kalimantan Timur untuk meningkatkan efisiensi, transparansi, dan kualitas pelayanan publik. Dengan adanya satu data yang terintegrasi, pemerintah daerah dapat mengambil keputusan yang lebih cepat dan tepat berdasarkan data yang akurat.
2. *Big Data* memberikan potensi besar dalam mengolah dan menganalisis data dalam jumlah besar dan beragam dari berbagai sumber. Ini memungkinkan pengambilan keputusan yang lebih baik dan strategi yang lebih efektif dalam berbagai sektor, termasuk kesehatan, pendidikan, dan infrastruktur.
3. *Data Science* memainkan peran penting dalam mengubah data mentah menjadi wawasan yang berharga. Dengan teknik-teknik seperti *machine learning*, analitik prediktif, dan data mining, *Data Science* dapat mengidentifikasi pola, tren, dan hubungan yang tidak terlihat sebelumnya.
4. Untuk mencapai satu data yang terpadu, perlu adanya sistem yang dapat mengintegrasikan data dari berbagai sumber dengan standar yang sama. Ini membutuhkan kolaborasi antar instansi dan peningkatan kapasitas teknologi informasi di pemerintah daerah.
5. Implementasi *big data* dan *data science* di Kalimantan Timur dapat meningkatkan kesejahteraan masyarakat melalui kebijakan publik yang lebih baik, peningkatan layanan kesehatan, perencanaan kota yang lebih efisien, serta pengelolaan sumber daya alam yang lebih bijaksana.

6. Tantangan utama meliputi masalah privasi dan keamanan data, keterbatasan infrastruktur, serta kurangnya tenaga ahli di bidang *data Science*. Solusi yang diusulkan mencakup pengembangan kebijakan yang mendukung perlindungan data, investasi dalam infrastruktur digital, serta program pendidikan dan pelatihan untuk meningkatkan keterampilan tenaga kerja.

Kesimpulannya, urgensi *data science* dalam transformasi digital pada satu data Kalimantan Timur adalah esensial untuk mendorong kemajuan daerah melalui pengambilan keputusan yang berbasis data, efisiensi operasional, serta peningkatan kualitas hidup Masyarakat.

5.2. Saran

Saran dari Kajian *Big Data* dan *Data Analytic*: Urgensi *Data Science* dalam Transformasi Digital Satu Data Kalimantan Timur yaitu:

1. Pengembangan Infrastruktur Digital yang Berkelanjutan:
 - a. Pemerintah daerah sebaiknya mengalokasikan anggaran khusus untuk pengembangan infrastruktur digital secara berkelanjutan.
 - b. Fokus pada peningkatan jaringan internet di daerah pedesaan untuk memastikan inklusivitas akses data.
2. Penerapan Kebijakan Data Terbuka:
 - a. Mengadopsi kebijakan data terbuka (*open data*) yang memungkinkan masyarakat dan sektor swasta mengakses data publik yang tidak sensitif.
 - b. Meningkatkan transparansi dan akuntabilitas pemerintah melalui penyediaan data yang mudah diakses dan dianalisis.
3. Mendorong Inovasi Melalui Kompetisi:
 - a. Mengadakan kompetisi *hackathon* dan lomba inovasi teknologi yang melibatkan masyarakat, pelajar, dan profesional untuk mencari solusi berbasis data bagi masalah-masalah lokal.

- b. Memberikan insentif bagi inovator dan startup yang berhasil mengembangkan solusi berbasis data yang dapat diimplementasikan oleh pemerintah daerah.
4. Pelatihan dan Pengembangan Kapasitas Berkelanjutan:
 - a. Menyediakan program pelatihan rutin bagi pegawai pemerintah dan masyarakat umum di bidang *Data Science* dan teknologi informasi.
 - b. Mengembangkan modul pelatihan yang disesuaikan dengan kebutuhan spesifik setiap instansi dan sektor.
5. Penguatan Kerjasama dengan Perguruan Tinggi dan Lembaga Penelitian:
 - a. Meningkatkan kerjasama dengan perguruan tinggi dan lembaga penelitian untuk penelitian dan pengembangan teknologi *Data Science*.
 - b. Mendorong program magang dan kerjasama penelitian yang melibatkan mahasiswa dan dosen dalam proyek-proyek pemerintah.
6. Pemanfaatan Teknologi untuk Analisis Data:
 - a. Mendorong penggunaan teknologi seperti AI dan *machine learning* untuk analisis data yang lebih mendalam dan prediktif.
 - b. Membuat pilot project atau proyek percontohan untuk menguji dan mengimplementasikan teknologi baru dalam analisis data.
7. Peningkatan Literasi Data di Kalangan Masyarakat:
 - a. Mengadakan program edukasi publik untuk meningkatkan literasi data di kalangan masyarakat, agar mereka lebih memahami manfaat dan pentingnya data dalam kehidupan sehari-hari.
 - b. Membuat kampanye sosialisasi yang menarik dan mudah dipahami tentang bagaimana data digunakan untuk meningkatkan pelayanan publik.
8. Fokus pada Keamanan dan Privasi Data:

- a. Menyusun pedoman dan standar keamanan data yang ketat untuk melindungi data dari ancaman siber.
 - b. Meningkatkan kesadaran dan pelatihan mengenai pentingnya privasi data di kalangan pegawai pemerintah dan masyarakat.
9. Penerapan Sistem Monitoring dan Evaluasi yang Efektif:
- a. Mengembangkan sistem monitoring dan evaluasi yang komprehensif untuk menilai keberhasilan implementasi kebijakan satu data.
 - b. Menggunakan data evaluasi untuk memperbaiki dan menyesuaikan kebijakan serta program yang ada.
10. Mendorong Partisipasi Publik dalam Pengambilan Keputusan:
- a. Membuka saluran komunikasi dua arah antara pemerintah dan masyarakat untuk mendapatkan masukan dan feedback terkait kebijakan data.
 - b. Mengadakan forum diskusi dan konsultasi publik secara rutin untuk memastikan kebijakan yang diambil sesuai dengan kebutuhan dan aspirasi masyarakat.

Dengan mengikuti saran-saran ini, diharapkan Pemerintah Provinsi Kalimantan Timur dapat mengoptimalkan potensi *big data* dan *data science* untuk mendukung transformasi digital yang efektif dan berkelanjutan, serta meningkatkan kualitas hidup masyarakat secara keseluruhan.



Data Mencerdaskan Bangsa



**Tanpa Data, Anda
Hanya Seseorang
Dengan Pendapat**

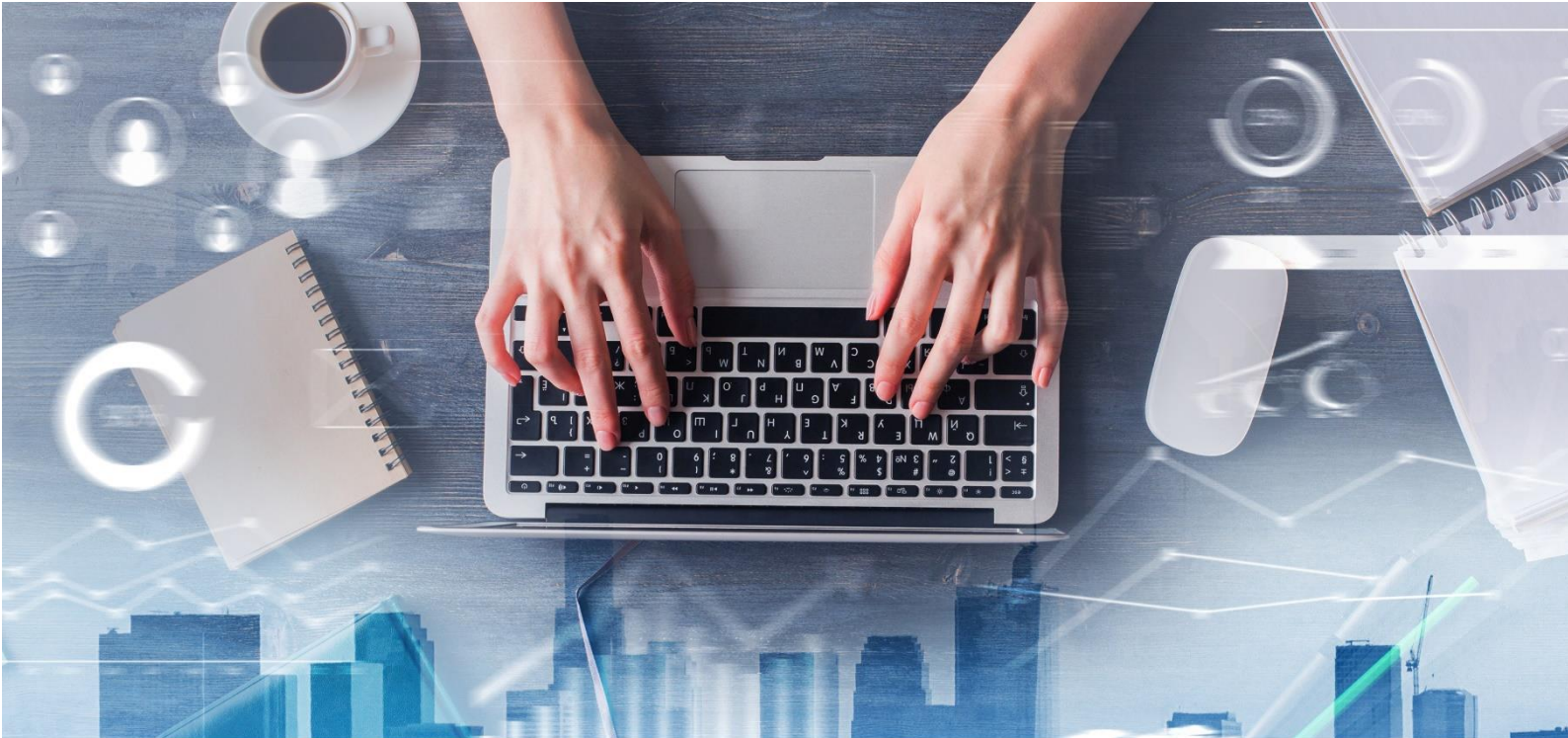


DAFTAR PUSTAKA

- Abdullah, Noraini, et.al. 2015. "Data quality in *Big Data*: a review", Int. J. Advance Soft Compu. Appl, Vol. 7, No. 3, November 2015.
- Adhinugroho, Y., Pratama Putra, A., Luqman, M., Yesa Ermawan, G., Takdir, T., Mariyah, S., Pramana, S. (2020). "Development of Online Travel Web Scraping For Tourism Statistics in Indonesia". Information Research.
- Big Data* at ONS. (2020). "How alternative sources of data and *Data Science* methods will affect our processes and outputs". Diakses pada 20 Oktober 2020, dari tautan <https://www.ons.gov.uk>.
- Bajaber, F., Elshawi, R., Batarfi, O., Altalhi, A., Barnawi, A., dan Sakr, S. (2016): *Big Data* 2.0 Processing Systems: Taxonomy and Open Challenges, Journal of Grid Computing, 14(3), 379–405, diperoleh melalui situs internet: <https://doi.org/10.1007/s10723-016-9371-1>
- Davenport, T. H., & Patil, D. J. (2012). "Data Scientist: The Sexiest Job of the 21st Century." Harvard Business Review.
- De Mauro, A., Greco, M., dan Grimaldi, M. (2016): A formal definition of *Big Data* based on its essential features, Library Review, 65(3), 122–135, diperoleh melalui situs internet: <https://doi.org/10.1108/LR-06-2015-0061>.
- Eurostat. 2014. "*Big Data* – an opportunity or a threat to official statistics?", UN Economic Commission for Europe, ECE/CES/2014/32, p. 4, 2014.
- Ghemawat, S., Gobioff, H., dan Leung, S.-T. (2003): The Google file system, Proceedings of the nineteenth ACM symposium on Operating systems principles - SOSP '03, 29, diperoleh melalui situs internet: <https://doi.org/10.1145/945449.945450>.

- Holmes, A. (2012): Hadoop In Practice - MEAP, Hadoop In Practice, diperoleh melalui situs internet: <http://dl.acm.org/citation.cfm?id=2543981>.
- Khan, Nawsher & Alsaqer, Mohammed & Shah, Habib & Badsha, Gran & Abbasi, Aftab & Salehian, Solmaz. (2018). The 10 Vs, Issues and Challenges of *Big Data*. 52-56. 10.1145/3206157.3206166.
- Kitchin, R. (2015). "The opportunities, challenges and risks of *Big Data* for official statistics", IRSA National University of Ireland Maynooth, County Kildare, Ireland.
- Loukides, M. (2011). "What is *Data Science*?" O'Reilly Media.
- Letouzé E, Jütting J. 2015. "Official statistics, *Big Data* and human development." Data-Pop Alliance White Paper Series. Data-Pop Alliance (Harvard Humanitarian Initiative, MIT Media Lab and Overseas Development Institute) and Paris21. March 2015.
- Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Hung Byers, A. (2011). "*Big Data*: The next frontier for innovation, competition, and productivity." McKinsey Global Institute.
- Moore, G. E. (2006): Cramming more components onto integrated circuits, Reprinted from *Electronics*, volume 38, number 8, April 19, 1965, pp.114 ff., IEEE Solid-State Circuits Newsletter, 20(3), 33–35, diperoleh melalui situs internet: <https://doi.org/10.1109/N-SSC.2006.4785860>.
- MacFeely, Steve. 2018. Working Paper Series CPS-WP18-001 *Big Data* & official statistics.
- OECD. (2015). "Data-Driven Innovation: *Big Data* for Growth and Well-Being." OECD Publishing.
- Provost, F., & Fawcett, T. (2013). "*Data Science* for Business: What You Need to Know about Data Mining and Data-Analytic Thinking." O'Reilly Media.

- P.P.N. L, Riyadi. Y., Hasyati A.N., Indriani, R. “*Big Data* for government policy: Potential implementations of bigdata for official statistics in Indonesia,” 2017 International Workshop on *Big Data* and Information Security (IWBIS), IEEE. Jakarta, 2017, pp. 17-21.
- Raghupathi, W., & Raghupathi, V. (2014). "*Big Data Analytics* in healthcare: promise and potential." *Health Information Science and Systems*, 2(1), 1-10.
- Rowley, J. (2007): The *wisdom* hierarchy: Representations of the DIKW hierarchy, *Journal of Information Science*, 33(2), 163–180, diperoleh melalui situs internet: <https://doi.org/10.1177/0165551506070706>
- Schroeck, M., Shockley, R., Smart, J., Romero-Morales, D., & Tufano, P. (2012). "Analytics: The real-world use of *Big Data*." IBM Global Business Services.
- Struijs, Braaksma, dan Daas, P.J.H.. 2014. Official statistics and *Big Data*. *Big Data & Society*. 1. 1-6. 10.1177/2053951714538417.
- U. Bustaman, D. N. Larasati, Z. Hidayah Satria Putri, S. Mariyah, Takdir and S. Pramana, “Building Effective and Efficient Procedure for Preprocessing Marketplace Data,” 2020 12th International Conference on Information Technology and Electrical Engineering (ICITEE), Yogyakarta, 2020, pp. 186-191, doi: 10.1109/ICITEE49829.2020.9271717.
- UNECE. 2013. “What does “*Big Data*” mean for official statistics?”, Conference Of European Statisticians, 10 March 2013.
- Virgillito, dan Scannapiec.. 2013. Placing *Big Data* in Official Statistics: A Big Challenge? Paper for the New Techniques and Technologies for Statistics conference.



DISKOMINFO
KALTIM



Penerbit dan Alamat Redaksi
Dinas Komunikasi dan Informatika Provinsi Kalimantan Timur
Bidang Statistik
Jalan Basuki Rahmat Nomor 41, Kelurahan Sungai Pinang Luar,
Kecamatan Samarinda Kota, Samarinda, Kalimantan Timur, 75121.
Telp/Fax : 0541-731963
Website : <https://diskominfo.kaltimprov.go.id>
Email : diskominfo@kaltimprov.go.id