Big Data, konsep pengelompokan atau pengumpulan data dalam skala besar.

3V Big Data

Big data adalah aset informasi bervolume tinggi (*high-volume*), berkecepatan tinggi (*high-velocity*), dan/atau memiliki banyak ragam (*high-variety*) yang menuntut bentuk pemrosesan informasi yang hemat biaya dan inovatif yang memungkinkan peningkatan wawasan, pengambilan keputusan, dan otomatisasi proses.

1. Volume

Jumlah data yang dihasilkan dari banyak transaksi serta volume data yang disimpan. Contohnya, seperti penggunaan *history browser*, pencatatan transaksi pada *e-commerce,*data ktp atau data penduduk Indonesia, data pelanggan pada perbankan dan masih banyak lagi. Ukuran big data biasanya menggunakan skala Terabytes (1000 Gigabytes) dan ukuran Petabytes (1.000.000 Gigabytes).

1. Variety

Variasi tipe dan variasi sifat dari data, apakah data tersebut bersifat terstruktur, semi terstruktur, ataupun tidak terstruktur. Mari kita bahas satu per satu.

1. Data Terstruktur

| **No** | **Kategori** | **Harga satuan** | **Paket** |
| --- | --- | --- | --- |
| 1 | Tiket Masuk Objek Wisata | Rp30.000 | Rp200.000 |
| 2 | Rafting 5 km/2 jam | Rp75.000 |
| 3 | 1x Makan Siang | Rp25.000 |
| 4 | Coffee Break | Rp15.000 |

Terlihat bahwa data pada tabel di atas tersusun rapi sesuai dengan struktur tabelnya. Jadi, orang dapat langsung memahami maksud yang ingin disampaikan dari tabel tersebut.

1. Data Tak Terstruktur

Terdapat beberapa format data, seperti gambar dan teks. Dari data tidak terstruktur, kita dapat mengetahui bahwa sebuah data tidak harus selalu berbentuk tabel dan angka, tetapi jauh lebih luas dari itu.

1. Data Semi Terstruktur

Data semi terstruktur adalah data dengan bentuk yang tidak dikenal. Pada dasarnya, jenis data ini memiliki atribut terstruktur yang membuatnya tetap dapat dianalisis. Namun, pada saat yang bersamaan, data semi terstruktur ini tidak mengikuti suatu susunan tertentu seperti data tidak terstruktur. Contoh data semi terstruktur di antaranya adalah berkas dalam bentuk .csv, email, halaman web, dan sejenisnya.

1. Velocity

Kecepatan dalam men-*generate* data, mengakses data serta memproses data. Big data *platform*dan big data*analytics software*tentu harus dapat memproses banyak data secepat mungkin ketika ada permintaan, contohnya adalah yang terdapat pada *search engine Google.*Ketika Anda ingin mencari suatu hal di Google, permintaan tersebut langsung diproses dan ditampilkan pada halaman Google.

Small data vs Big Data

| **Fitur** | **Small Data** | **Big Data** |
| --- | --- | --- |
| Variasi | Data biasanya terstruktur dan seragam. | Data seringkali tidak terstruktur dan heterogen. |
| Fakta | Data umumnya berkualitas tinggi dan dapat diandalkan. | Kualitas dan keandalan data dapat sangat bervariasi. |
| Teknologi | Tradisional | Modern |
| Volume | Data dalam kisaran puluhan atau ratusan Gigabyte. | Ukuran data lebih dari Terabyte. |
| Basis Data | SQL | NoSQL |
| Bahasa Pemrograman | SQL | Python, R, Java, SQL |
| Posisi (*Job*) | *Data Analyst*, *Database Administrators*, dan *Data Engineer* | *Data Scientist, Data Analyst, Database Administrators,*dan *Data Engineer.* |

Kenapa Big Data Penting?

Dengan menggunakan big data, sebuah perusahaan akan diberikan sebuah *insight*(wawasan) tentang social media yang dapat digunakan perusahaan untuk menyempurnakan pemasaran, periklanan, dan promosi mereka guna meningkatkan keterlibatan pelanggan serta tingkat konversi.

**Big Data dan Data Science**

urutan proyek big data di bawah ini.

1. *Identify the problem*
2. *Get the data*
3. *Prepare the data*
4. *Analyze the data*
5. *Generate reports and insights*
6. *Perform practical actions*

Dari urutan di atas, tahapan 1, 2, dan 3 adalah langkah-langkah yang ditangani dengan teknologi big data, sedangkan pada tahapan 4, 5, dan 6 menggunakan teknologi data science. Oleh karena itu, data science tidak akan luput dari big data.

**Peran Big Data pada Digital Transformation**

Transformasi digital membantu perusahaan merangkul perubahan dan tetap kompetitif di dunia yang semakin digital. Peran big data dalam transformasi digital berasal dari potensi sebuah organisasi untuk menggabungkan keduanya dalam usahanya. Oleh karena itu, memungkinkan digitalisasi dan otomatisasi operasi bisnis. Digitalisasi dan otomatisasi ini meningkatkan efisiensi, memacu inovasi, dan mengarah ke model bisnis baru.

Selain itu, big data analytics memungkinkan bisnis untuk mengamati informasi terperinci tentang kelompok pelanggan tertentu atau berbeda. Selain itu, bisa jadi untuk informasi lainnya berupa hal-hal yang dilakukan oleh *customer* saat berada di situs web mereka, barang yang dibeli, seberapa sering para *customer*membelinya, dan masih banyak lagi. Jadi, dengan menggunakan semua informasi ini, bisnis menerapkan perubahan untuk memenuhi kebutuhan pelanggan di masa depan. Oleh karena itu, untuk menyelesaikan transformasi digital, sektor bisnis perlu mengadopsi big data.