

Nama : Benhard Simanullang

NIM : 24060120140136

### **Rangkuman sesi 1 : Big Data**

Big data adalah sebuah koleksi dataset yang tidak dapat di proses dengan teknik komputasi lawas, big data bergantung pada data yang di produksi pada berbagai macam device dan aplikasi.

Big data merupakan term yang digunakan untuk mendeskripsikan koleksi data yang sangat besar dan terus bertumbuh seiring waktu

Big data merupakan salah satu hal yang merubah cara kita melakukan bisnis dan memberi impact ke berbagai hal pada kehidupan sehari hari kita

Big data biasa di deskripsikan sebagai empat V, yaitu

1. Volume : mengarah kepada luasnya jumlah data baru yang dihasilkan setiap detik
2. Velocity ( kecepatan ) : merujuk pada kecepatan data baru yang dihasilkan dan kecepatan data bergerak.
3. Variety ( variasi ) : merujuk pada tipe data yang beragam
4. Veracity : merujuk pada kekacauan atau kepercayaan terhadap data

Berbagai macam data yang berada di dunia memberi memberi kan jumlah volume, velocity, variety, dan veracity yang sangat besar, dan sejalan dengan berkembangnya teknologi komputasi seperti komputasi awan maka kumpulan data ini memberikan kita nilai tambahan apabila dikelola dengan benar.

Proses pada data analitik

1. Koleksi data
2. Mengkategorikan data
3. Mengurutkan data
4. Melakukan ETL -> extract, transform, load
5. Menganalisis data
6. Mengshare data

## **Rangkuman sesi 2 : Visualisasi data**

Visualisasi data adalah praktik untuk menerjemahkan data ke dalam konteks visual seperti peta atau grafik, dengan tujuan mempermudah memahami data oleh manusia.

Visualisasi data juga merupakan disiplin seni yang bertujuan untuk mencari, memanipulasi, memformat, dan menyampaikan data dengan cara yang paling bagus.

Visualisasi data juga menyediakan cara yang cepat dan berkesan untuk menyampaikan informasi secara universal dengan konteks visual.

Keuntungan dari visualisasi data

1. Menyerap informasi dengan cepat dan memberikan keputusan yang tepat
2. Meningkatkan pemahaman tentang langkah selanjutnya yang harus diambil
3. Menyampaikan informasi yang mudah yang meningkatkan peluang pemahaman orang yang terlibat.
4. Tidak diperlukan nya data saintist karena data mudah di akses dan disajikan dengan bentuk yang lebih sederhana.

Gestalt Principles

1. Figure
2. Proximity
3. Similarity
4. parallelism
5. Common fate

Memilih grafik yang tepat

- a. Bagian garis (line chart) : bagian garis menampilkan bagaimana variable dapat berubah dari waktu ke waktu
- b. Area charts : metode visuali ini adalah variasi dari diagram garis yang menampilkan beberapa nilai dalam deret waktu, atau urutan data yang dikumpulkan