

Kumpulan data dikategorikan sebagai Big Data ketika memiliki karakteristik tertentu, sering disebut dengan "5V". Berikut adalah penjelasan singkat dari masing-masing karakteristik tersebut:

1. Volume

Merujuk pada ukuran data yang sangat besar dan terus bertambah dengan cepat. Volume data ini bisa mencapai petabyte atau bahkan exabyte. Contohnya termasuk data dari media sosial, sensor IoT, transaksi online, dan lain-lain.

2. Velocity

Menunjukkan kecepatan tinggi di mana data dihasilkan dan perlu diproses. Data streaming dari media sosial, log transaksi online, dan data sensor adalah contoh di mana data perlu dianalisis secara real-time atau hampir real-time.

3. Variety

Menggambarkan berbagai jenis dan format data yang ditangani, mulai dari data terstruktur (seperti database), semi-terstruktur (seperti XML, JSON), hingga data tidak terstruktur (seperti teks, video, dan gambar). Varietas ini menambah kompleksitas dalam pengelolaan dan analisis data.

4. Veracity

Berkaitan dengan keakuratan dan keandalan data. Karena volume besar dan variasi sumber data, seringkali data tersebut mengandung ketidakpastian, ketidaklengkapan, atau bahkan kesalahan. Mengelola veracity berarti memastikan data yang dianalisis berkualitas tinggi dan dapat dipercaya.

5. Value

Menekankan pentingnya mengubah data besar menjadi nilai atau wawasan yang bermanfaat. Tidak cukup hanya memiliki akses ke data dalam jumlah besar; yang penting adalah kemampuan untuk mengolah data tersebut menjadi informasi yang dapat digunakan untuk pengambilan keputusan, inovasi, dan penciptaan keunggulan kompetitif.