

# STUDI KASUS

PT **SmartMart Analytics** adalah perusahaan ritel nasional yang memiliki lebih dari 300 gerai fisik dan platform e-commerce yang melayani jutaan transaksi setiap bulan. Setiap hari perusahaan menghasilkan data dalam jumlah besar yang berasal dari:

- Transaksi penjualan (POS dan e-commerce)
- Data pelanggan (profil, riwayat belanja, loyalty program)
- Data media sosial (ulasan, komentar, sentimen)
- Data log sistem dan aplikasi mobile
- Data sensor gudang (stok dan distribusi)

Selama ini, PT SmartMart hanya menggunakan sistem database relasional konvensional untuk pelaporan bulanan. Namun, manajemen menghadapi berbagai permasalahan:

- Laporan lambat dan tidak real-time
- Kesulitan mengelola data tidak terstruktur
- Sulit menemukan pola perilaku pelanggan
- Infrastruktur lama tidak mampu menangani pertumbuhan data yang sangat cepat

Manajemen berencana mengadopsi **Big Data Architecture berbasis Hadoop Ecosystem** untuk mendukung pengambilan keputusan strategis dan keunggulan kompetitif perusahaan.

---

## Soal 1 – Asal Usul & Latar Belakang Big Data (CPMK-1 | 10 poin)

Jelaskan **asal usul dan faktor utama** berkembangnya Big Data yang relevan dengan kondisi PT SmartMart Analytics. Mengapa pendekatan database tradisional tidak lagi memadai?

- ⇒ Big Data berkembang karena digitalisasi bisnis, internet, media sosial, dan sistem sensor yang menghasilkan data sangat besar dan beragam.
- Pada PT SmartMart, data berasal dari transaksi, pelanggan, media sosial, log sistem, dan sensor gudang.
- Pendekatan database relasional tradisional tidak memadai karena sulit diskalakan, tidak optimal untuk data tidak terstruktur, dan umumnya hanya mendukung pemrosesan batch sehingga laporan tidak real-time

---

## Soal 2 – Jenis Implementasi Big Data (CPMK-2 | 10 poin)

Berdasarkan studi kasus, jelaskan **jenis implementasi Big Data** yang paling sesuai untuk PT SmartMart (operational analytics, customer analytics, real-time analytics, atau lainnya). Sertakan alasan pemilihan Anda.

Implementasi yang sesuai adalah:

- ⇒ **Customer analytics** untuk memahami perilaku dan meningkatkan loyalty
- ⇒ **Operational analytics** untuk optimasi stok dan distribusi
- ⇒ **Real-time analytics** mendukung monitoring stok, transaksi, dan promosi secara langsung di seluruh gerai dan e-commerce.
-

### Soal 3 – Sifat dan Karakteristik Big Data (CPMK-3 | 10 poin)

Identifikasikan dan jelaskan **sifat Big Data (5V atau lebih)** yang muncul pada data PT SmartMart Analytics, disertai contoh konkret dari kasus.

Karakteristik Big Data pada PT SmartMart meliputi:

- ⇒ **Volume:** Data transaksi dan pelanggan sangat besar
  - ⇒ **Velocity:** Data masuk terus-menerus dan real-time
  - ⇒ **Variety:** Data terstruktur, semi-terstruktur, dan tidak terstruktur
  - ⇒ **Veracity:** Data media sosial mengandung noise
  - ⇒ **Value:** Data bernilai untuk strategi bisnis
- 

### Soal 4 – Struktur Data & Analisis Big Data (CPMK-4 | 10 poin)

a. Klasifikasikan jenis data PT SmartMart ke dalam **structured, semi-structured, dan unstructured data**.

Jenis data:

- ⇒ **Structured:** Transaksi penjualan, data stok
- ⇒ **Semi-structured:** Log sistem, data JSON
- ⇒ **Unstructured:** Ulasan dan komentar pelanggan

b. Jelaskan tantangan analisis data berukuran besar berdasarkan struktur data tersebut.

- ⇒ Kesulitan integrasi data, pengolahan data tidak terstruktur, dan kebutuhan komputasi tinggi.
- 

### Soal 5 – Kompleksitas Big Data (CPMK-5 | 10 poin)

Analisis bentuk **kompleksitas Big Data** yang dihadapi PT SmartMart dari sisi teknis, organisasi, dan analisis data. Mengapa kompleksitas ini membutuhkan pendekatan Big Data?

Kompleksitas meliputi:

- ⇒ **Teknis:** Keterbatasan penyimpanan dan performa
  - ⇒ **Organisasi:** Kesiapan SDM dan budaya data
  - ⇒ **Analisis:** Data besar dan beragam sulit dianalisis
- Pendekatan Big Data diperlukan untuk mengatasi kompleksitas ini.
- 

### Soal 6 – Siklus Manajemen Data & Data Warehouse (CPMK-6 | 10 poin)

Jelaskan **siklus manajemen data** mulai dari akuisisi hingga visualisasi data pada PT SmartMart. Bandingkan peran **Data Warehouse** dan **Big Data Platform** dalam konteks studi kasus.

- ⇒ Siklus data:  
Akuisisi → Ingestion → Storage → processing → Analisis → Visualisasi
  - ⇒ Data Warehouse digunakan untuk laporan historis terstruktur, sedangkan Big Data Platform untuk data besar, tidak terstruktur, dan real-time.
-

### Soal 7 – Pemanfaatan Big Data di Berbagai Industri (CPMK-7 | 10 poin)

Jelaskan bagaimana penerapan Big Data pada PT SmartMart dapat disandingkan dengan pemanfaatan Big Data di **minimal dua industri lain** (misalnya perbankan, kesehatan, atau transportasi).

- ⇒ Di perbankan, Big Data digunakan untuk deteksi fraud.
  - Di kesehatan, untuk analisis rekam medis dan prediksi penyakit.
  - Di SmartMart, Big Data digunakan untuk analisis pelanggan dan optimasi operasional.
- 

### Soal 8 – Ekstraksi Insight & Pola Data (CPMK-8 | 10 poin)

Berikan contoh **insight atau pola (pattern)** yang dapat diekstraksi dari Big Data PT SmartMart. Jelaskan metode analisis yang dapat digunakan untuk memperoleh insight tersebut.

Insight yang diperoleh antara lain:

- ⇒ Pola pembelian pelanggan
  - ⇒ Rekomendasi produk
  - ⇒ Prediksi permintaan stok
- Metode analisis: data mining, machine learning, dan analisis sentimen.
- 

### Soal 9 – Framework Big Data (CPMK-9 | 10 poin)

Jelaskan **framework Big Data** yang sesuai untuk PT SmartMart Analytics. Uraikan peran masing-masing komponen utama dalam framework tersebut.

- ⇒ Framework Big Data yang sesuai adalah **Hadoop-based Big Data Framework**
    - **Data ingestion:** mengumpulkan data dari transaksi, media sosial, dan sensor.
    - **Distributed storage:** menyimpan data besar secara terdistribusi.
    - **Processing & resource management:** memproses data secara paralel dan efisien.
    - **Analytics layer:** menganalisis data untuk menemukan pola dan insight.
    - **Visualization:** menyajikan hasil analisis dalam bentuk laporan dan dashboard.
- 

### Soal 10 – Teknologi Hadoop Ecosystem (CPMK-10 | 10 poin)

Jelaskan peran teknologi dalam **ekosistem Hadoop** (misalnya HDFS, YARN, MapReduce, Hive, HBase, Spark) yang dapat digunakan untuk menyelesaikan permasalahan PT SmartMart.

- ⇒ **HDFS:** Penyimpanan data terdistribusi
- ⇒ **YARN:** Manajemen sumber daya
- ⇒ **MapReduce:** Pemrosesan batch
- ⇒ **Hive:** Query data berbasis SQL
- ⇒ **HBase:** Akses data real-time
- ⇒ **Spark:** Analisis cepat dan real-time