a. Dimensi Bisnis (Business Dimension) dari Kebutuhan Data Warehouse

Dalam sebuah perusahaan yang fokus pada penjualan produk elektronik, beberapa dimensi bisnis penting yang perlu dipertimbangkan untuk Data Warehouse meliputi:

1. Produk: Berisi informasi tentang produk yang dijual, termasuk kode produk, nama, kategori, harga, dan deskripsi.
2. Waktu: Mengikuti kalender, dengan atribut seperti tahun, bulan, hari, dan kuartal.
3. Pelanggan: Informasi tentang pelanggan, seperti ID pelanggan, nama, alamat, dan segmentasi pelanggan.
4. Penjualan: Mencakup data penjualan seperti nomor faktur, tanggal penjualan, jumlah, diskon, dan metode pembayaran.
5. Lokasi: Meliputi lokasi penjualan atau distribusi, termasuk kota, negara bagian, dan wilayah.

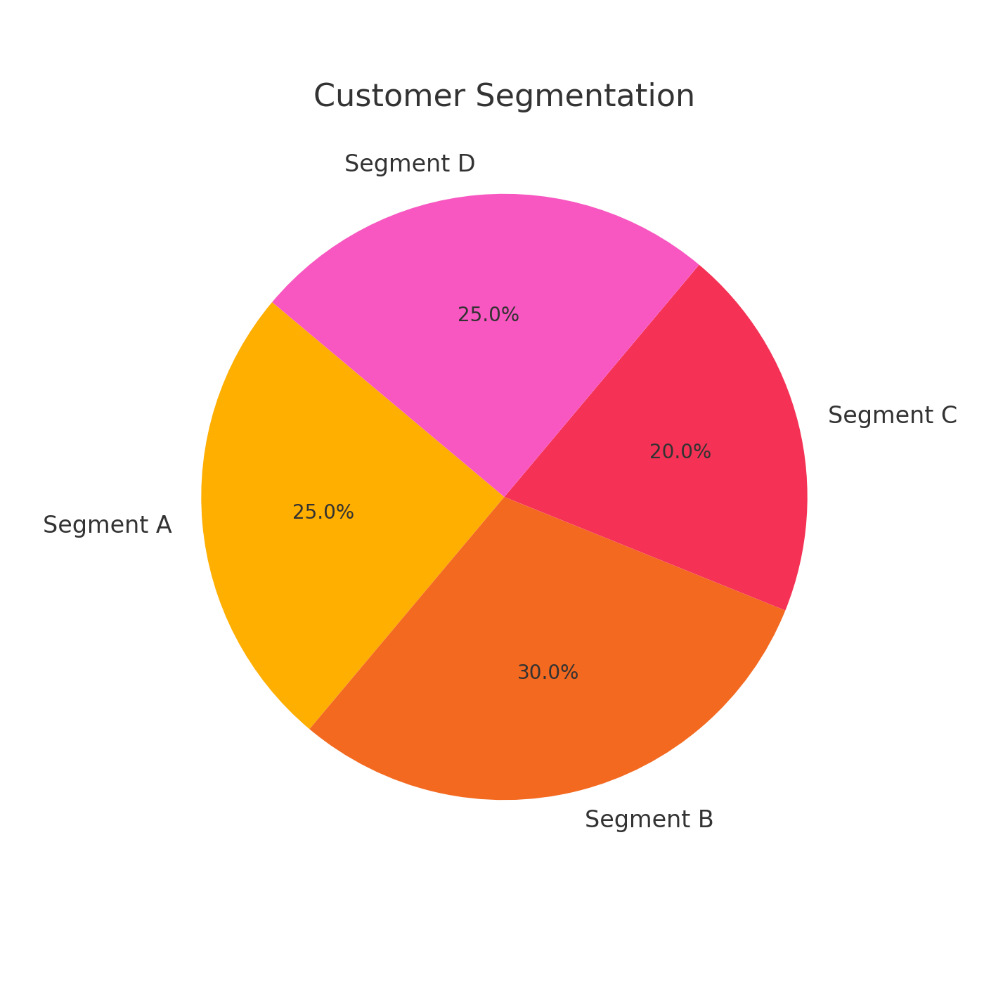
b. Common Subject Area dan Metric Bisnis Kunci

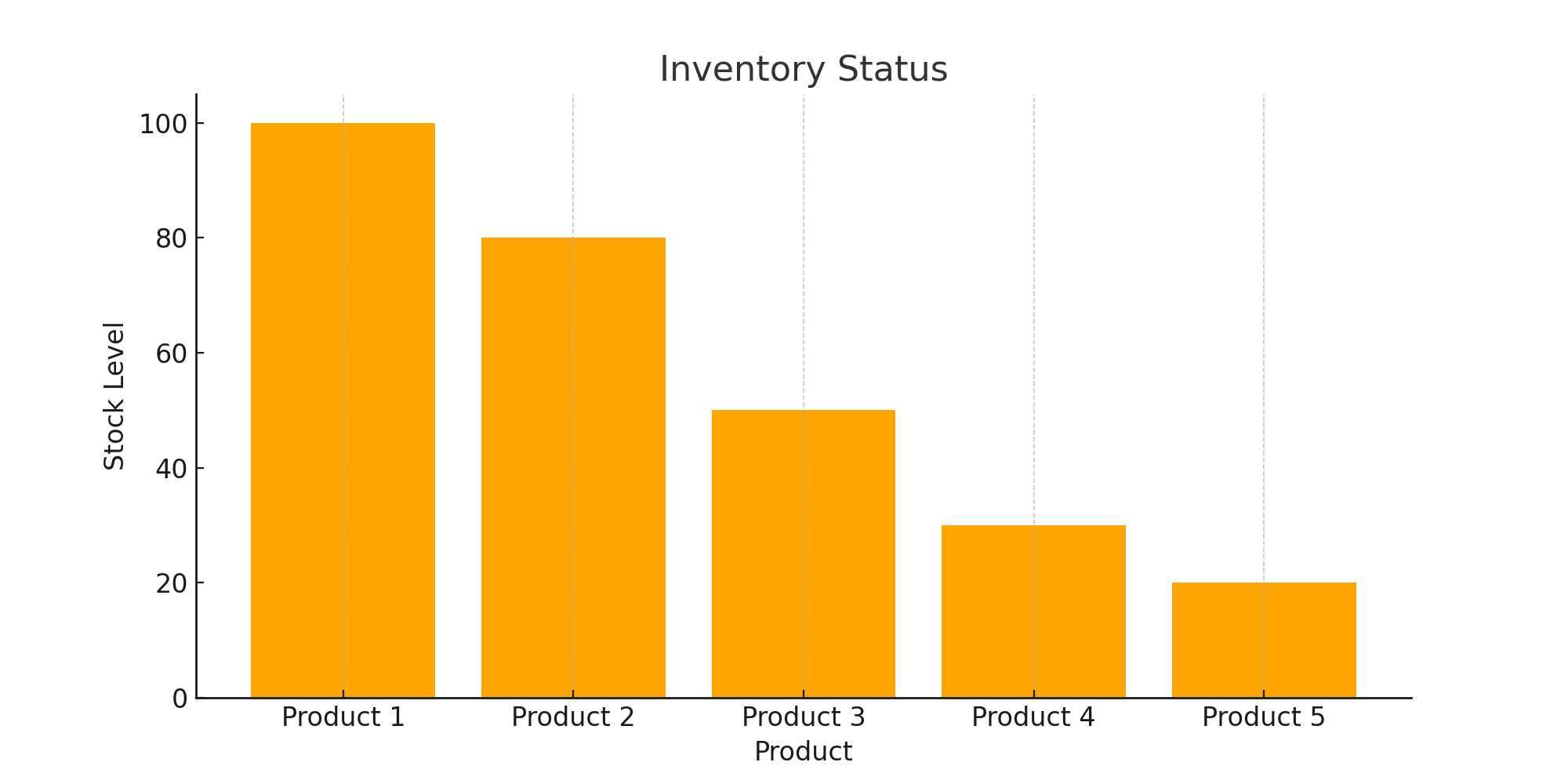
Common Subject Area:

1. Sales Data: Meliputi transaksi penjualan, produk yang terjual, pelanggan, dan lokasi penjualan.
2. Inventory Management: Informasi tentang stok produk, pengiriman, dan restocking.

Metric Bisnis Kunci:

1. Total Penjualan: Jumlah total penjualan yang dilakukan dalam periode waktu tertentu.
2. Revenue: Pendapatan yang dihasilkan dari penjualan.
3. Profit Margin: Perbandingan antara profit dan revenue.
4. Customer Retention Rate: Persentase pelanggan yang kembali melakukan pembelian dalam jangka waktu tertentu.
5. Inventory Turnover: Jumlah produk yang terjual dibandingkan dengan rata-rata persediaan produk.

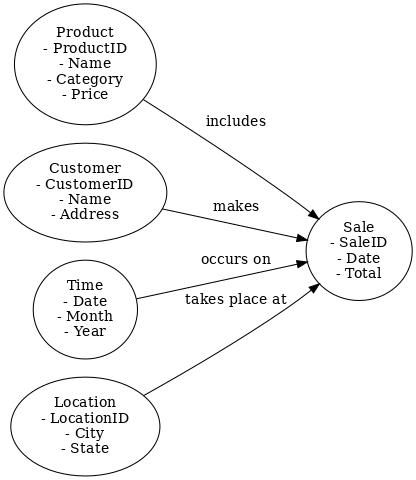




c. Arsitektur Data Warehouse

Arsitektur Data Warehouse biasanya terdiri dari beberapa komponen utama:

1. Source Systems: Sistem sumber data seperti ERP, CRM, dan aplikasi lainnya yang mengumpulkan data operasional.
2. Data Staging Area: Tempat di mana data dari berbagai sumber dikumpulkan dan diproses sebelum dimasukkan ke dalam Data Warehouse. Proses ETL (Extract, Transform, Load) dilakukan di sini.
3. Data Warehouse: Penyimpanan data yang diatur dalam skema yang sesuai, seperti skema bintang atau skema salju. Data di sini sudah dibersihkan dan diintegrasikan.
4. Data Marts: Subset dari Data Warehouse yang berfokus pada area spesifik, seperti penjualan atau inventaris.
5. OLAP Cubes: Digunakan untuk analisis multidimensi dan memungkinkan pengguna untuk melakukan drill-down atau roll-up data.
6. Data Mining/Analytics Tools: Alat-alat yang digunakan untuk analisis data lanjutan, seperti machine learning dan statistik.
7. BI Tools: Alat Business Intelligence yang digunakan untuk visualisasi dan pelaporan data.



d. Kebutuhan Perancangan Data, Spesifikasi Data Storage, dan Strategi Information Delivery

Kebutuhan Perancangan Data:

* Integrasi Data: Memastikan data dari berbagai sumber dapat digabungkan dengan benar.
* Konsistensi Data: Menjaga keselarasan data, baik secara temporal maupun dalam bentuk lainnya.
* Skalabilitas: Mampu menangani peningkatan volume data seiring dengan pertumbuhan bisnis.

Spesifikasi Data Storage:

* Data Storage: Menggunakan penyimpanan berbasis disk atau cloud dengan kemampuan kompresi dan de-duplikasi data.
* Database Management System (DBMS): Menggunakan DBMS relasional seperti MySQL, PostgreSQL, atau DBMS non-relasional seperti MongoDB untuk penyimpanan data.

Strategi Information Delivery:

* Dashboards: Menggunakan dashboard untuk memberikan visualisasi data kepada pengguna akhir.
* Scheduled Reports: Pembuatan laporan berkala yang dapat diakses oleh pengguna untuk analisis lebih lanjut.
* Ad-hoc Query: Memungkinkan pengguna untuk melakukan query ad-hoc untuk mendapatkan data yang diperlukan.

e. Mekanisme Data Cleansing dalam Proses Transformasi

Data cleansing adalah proses penting dalam ETL yang memastikan kualitas data dengan menghapus data yang tidak akurat, tidak lengkap, atau berlebihan. Berikut adalah contoh mekanisme data cleansing:

1. Validasi Data: Memeriksa apakah data memenuhi kriteria tertentu, seperti validasi format email atau nomor telepon.
2. Menghapus Duplikasi: Identifikasi dan penghapusan data duplikat berdasarkan kunci unik seperti ID pelanggan.
3. Normalisasi: Mengonversi data ke format standar, seperti mengubah tanggal ke format YYYY-MM-DD.
4. Penanganan Nilai Null: Mengisi nilai null dengan nilai default atau menjalankan logika bisnis untuk menentukan nilai yang tepat.

Contoh: Jika ada entri pelanggan dengan alamat yang hilang, mekanisme cleansing dapat mengisi alamat berdasarkan informasi pelanggan lain yang tersedia atau menandainya untuk perbaikan manual.

Soal 3: Contoh Meta Data di Data Warehouse

Meta data adalah data tentang data yang membantu dalam memahami dan mengelola Data Warehouse. Berikut adalah contoh meta data yang bisa ada dalam Data Warehouse:

1. Data Operasional:
   * Nama Tabel: Sales
   * Deskripsi Tabel: Menyimpan data transaksi penjualan
   * Kolom Utama: SaleID, CustomerID, ProductID, Date
2. History ETL:
   * Log ETL: Menyimpan informasi tentang proses ETL, termasuk waktu mulai dan selesai, jumlah record yang diproses, dan status keberhasilan/kegagalan.
3. Pengguna Data Warehouse/End User:
   * Nama Pengguna: JohnDoe
   * Peran: Analis Data
   * Akses Data: Diberikan akses ke tabel penjualan dan inventaris untuk analisis data.

Soal 4: Analisis Big Data dan Contoh Dashboard

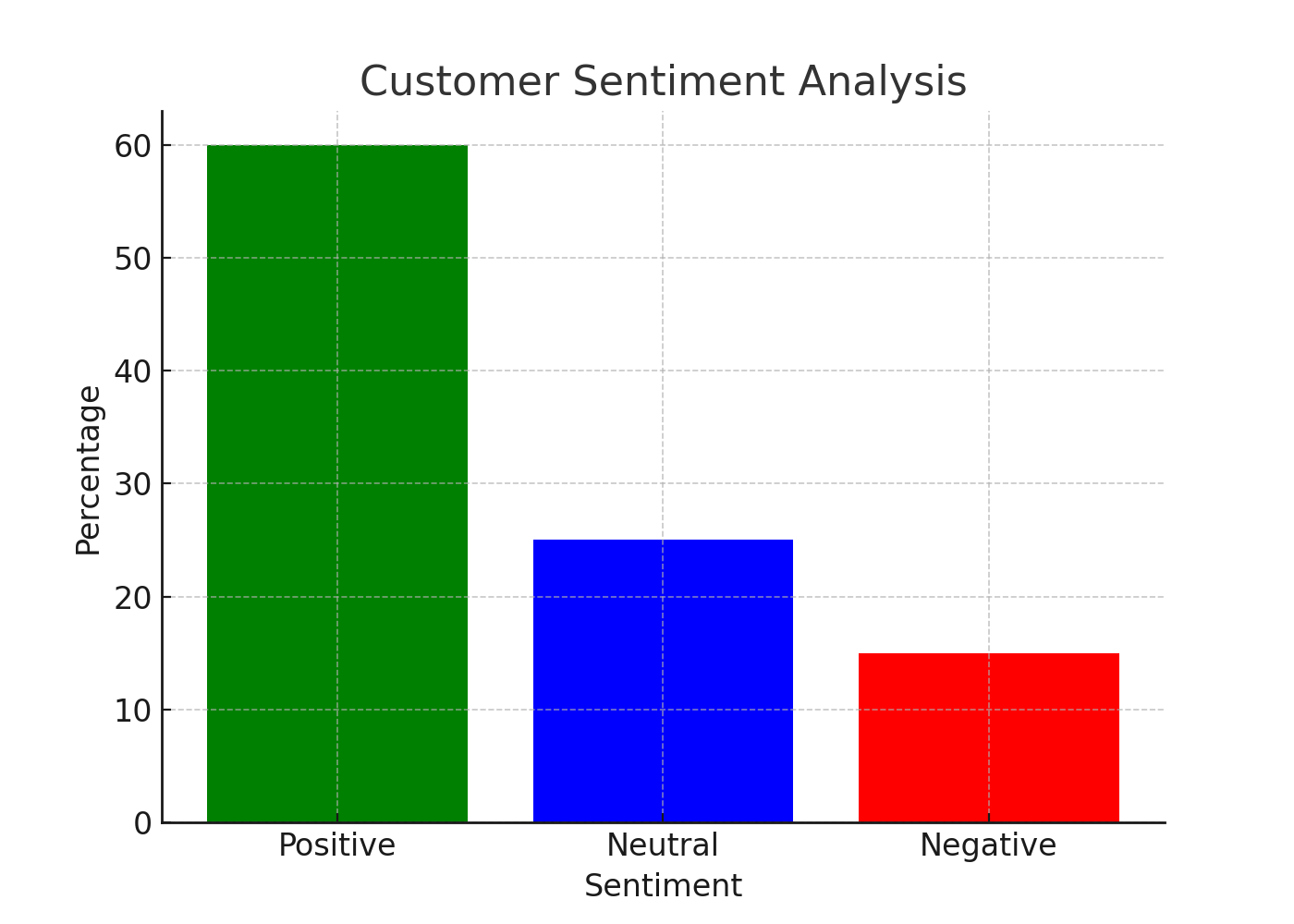
Arsitektur Big Data di Perusahaan

Perusahaan yang bergerak di bidang e-commerce menggunakan arsitektur Big Data yang terdiri dari beberapa komponen:

1. Data Sources: Termasuk log server web, data transaksi, data sosial media, dan data pelanggan.
2. Data Ingestion: Menggunakan Apache Kafka untuk ingest data real-time.
3. Data Storage: Hadoop HDFS untuk penyimpanan data yang tidak terstruktur dan semi-terstruktur.
4. Data Processing: Menggunakan Apache Spark untuk pemrosesan data paralel.
5. Data Analysis: Menggunakan alat analitik seperti Hadoop MapReduce dan SQL untuk analisis data.

Big Data Analytic yang Dilakukan

* Customer Segmentation: Menganalisis data pelanggan untuk mengelompokkan mereka berdasarkan perilaku pembelian, demografi, dan preferensi produk.
* Sentiment Analysis: Menggunakan data dari media sosial untuk memahami sentimen pelanggan terhadap produk atau merek.
* Predictive Analytics: Memprediksi tren penjualan dan permintaan produk menggunakan algoritma machine learning.



Contoh Dashboard

Dashboard untuk visualisasi data dapat mencakup beberapa widget seperti:

1. Sales Performance: Grafik batang atau garis yang menunjukkan penjualan harian, mingguan, atau bulanan.
2. Customer Segmentation: Diagram lingkaran yang menunjukkan distribusi pelanggan berdasarkan segmen.
3. Sentiment Analysis: Grafik kata atau sentimen yang menggambarkan perasaan umum pelanggan terhadap merek.
4. Inventory Status: Tabel atau grafik yang menunjukkan status stok produk secara real-time.

Dashboard ini membantu dalam pengambilan keputusan bisnis dengan menyajikan data yang relevan dan mudah dipahami.

