



# DATA MINING

## PERTEMUAN 12

Memahami dan menjelaskan konsep Agregasi (*Aggregate*) dalam Data Mining





## Konsep Agregasi dalam Data Mining

**Agresasi** ada kaitannya dalam proses *Knowledge Discovery in Databases* (KDD) atau Data Mining

Data dunia nyata sering kali kotor (tidak lengkap, bising/noisy, tidak konsisten) dan kompleks (volume besar, dimensi tinggi). **Pre-processing** data adalah serangkaian teknik untuk membersihkan, mengintegrasikan, mentransformasi, dan mereduksi data agar siap untuk dianalisis oleh algoritma Data Mining.



## Definisi dan Konsep Agregasi

*Agregasi (Aggregation)* adalah salah satu teknik Data *Pre-processing* yang bertujuan untuk mengkombinasikan dua atau lebih objek (rekord/baris) menjadi sebuah objek tunggal atau merangkum data pada level yang lebih tinggi.

Agregasi sering kali melibatkan:

1. **Pengelompokan (*Grouping*)**: Data dikelompokkan berdasarkan satu atau lebih atribut kategori (misalnya, mengelompokkan transaksi berdasarkan ID Pelanggan atau Cabang Toko ).
2. **Perhitungan Fungsi Agregasi**: Menerapkan fungsi matematika atau statistik pada setiap kelompok untuk menghasilkan nilai ringkasan tunggal.



## Tujuan dan Manfaat Agregasi

**Agregasi** adalah langkah kunci untuk mengatasi beberapa masalah data, dengan manfaat utama:

Tujuan/Masalah Data	Manfaat Agregasi
<b>Volume Data Tinggi</b>	<b>Reduksi Data (Data Reduction):</b> Mengurangi jumlah baris (tuples) dalam dataset, membuat proses Data Mining berjalan lebih cepat dan efisien tanpa kehilangan informasi penting. Contoh: Dari jutaan transaksi harian, kita hanya perlu ringkasan penjualan bulanan per cabang.
<b>Kompleksitas Data</b>	<b>Memperoleh Pandangan Ringkas:</b> Data yang kompleks dan terperinci disederhanakan menjadi representasi yang lebih tinggi dan bermakna (higher-level view).
<b>Kualitas Data</b>	<b>Stabilitas Data:</b> Nilai rata-rata (misalnya) dari kelompok data lebih stabil dan kurang terpengaruh oleh noise (data bising) atau outliers dibandingkan nilai data mentah individual.



## Fungsi-fungsi Agregasi Umum

Berikut adalah beberapa **fungsi agregasi standar** yang paling sering digunakan:

Fungsi Agregasi	Deskripsi	Contoh Penggunaan
<b>SUM (Jumlah)</b>	Menghitung total nilai dari atribut numerik dalam satu kelompok.	Total Pendapatan yang dihasilkan per bulan.
<b>AVERAGE / MEAN (Rata-rata)</b>	Menghitung nilai rata-rata dari atribut numerik dalam satu kelompok.	Rata-rata umur pelanggan untuk setiap kategori produk.
<b>COUNT (Hitungan)</b>	Menghitung jumlah objek/baris dalam satu kelompok.	Jumlah total transaksi yang dilakukan oleh setiap pelanggan.
<b>MIN (Minimum)</b>	Menemukan nilai terkecil dari atribut numerik dalam satu kelompok.	Harga produk termurah yang dibeli dalam satu keranjang belanja.
<b>MAX (Maksimum)</b>	Menemukan nilai terbesar dari atribut numerik dalam satu kelompok.	Nilai transaksi tertinggi di setiap cabang.
<b>VARIANCE / STD DEV (Varian/Standar Deviasi)</b>	Mengukur sebaran data dalam satu kelompok.	Mengukur variabilitas penjualan harian dalam seminggu.
<b>MODE (Modus)</b>	Menemukan nilai yang paling sering muncul dalam atribut (biasanya kategori).	Kategori produk yang paling sering dibeli di setiap wilayah.



DATA MINING

## PRAKTIKUM AGREGATE

**Silahkan tonton video praktikum aggregate pada link dibawah ini**  
**Silahkan deskripsikan pemahaman anda terkait metode/algorithm datamining :**

**<https://youtu.be/oZNe7wQNkKI>**

**Praktikan dengan rapidminer dan buat laporan praktikumnya. Informasi detail ada pada elearning polsub**