

Nama : Andi Cleopatra Maryam Jamila

Nim:1103213071

## Week 6

### Sebagai Kesimpulan:

Pemilihan model clustering yang terbaik sangat bergantung pada karakteristik data, tujuan analisis, dan sumber daya komputasi yang tersedia. **K-Means** cocok untuk dataset besar dan bentuk cluster yang relatif sederhana, namun sensitif terhadap inisialisasi centroid dan outliers. **Hierarchical Clustering** menawarkan fleksibilitas dalam bentuk cluster dan tidak memerlukan penentuan jumlah cluster sebelumnya, tetapi bisa menjadi kompleks untuk dataset besar. **Euclidean Clustering** adalah pilihan dasar, namun kurang efisien untuk dataset kompleks.

1. **Euclidean Clustering:** Model ini menggunakan jarak Euclidean untuk mengukur kemiripan antara data. Data yang jaraknya lebih dekat cenderung dikelompokkan bersama. Namun, Euclidean Clustering seringkali kurang efisien untuk dataset besar dan kompleks.
2. **Hierarchical Clustering:** Model ini membangun hierarki kelompok secara bertahap, mulai dari kelompok terkecil hingga terbesar. Ada dua pendekatan utama: agglomerative (menggabungkan kelompok) dan divisive (membagi kelompok). Fleksibilitas adalah keunggulannya, tetapi bisa menjadi kompleks untuk dataset besar.
3. **K-Means Clustering:** Model ini membagi data menjadi K kelompok yang tidak tumpang tindih. Algoritmanya relatif sederhana dan efisien, tetapi sangat sensitif terhadap inisialisasi centroid awal dan bentuk cluster.