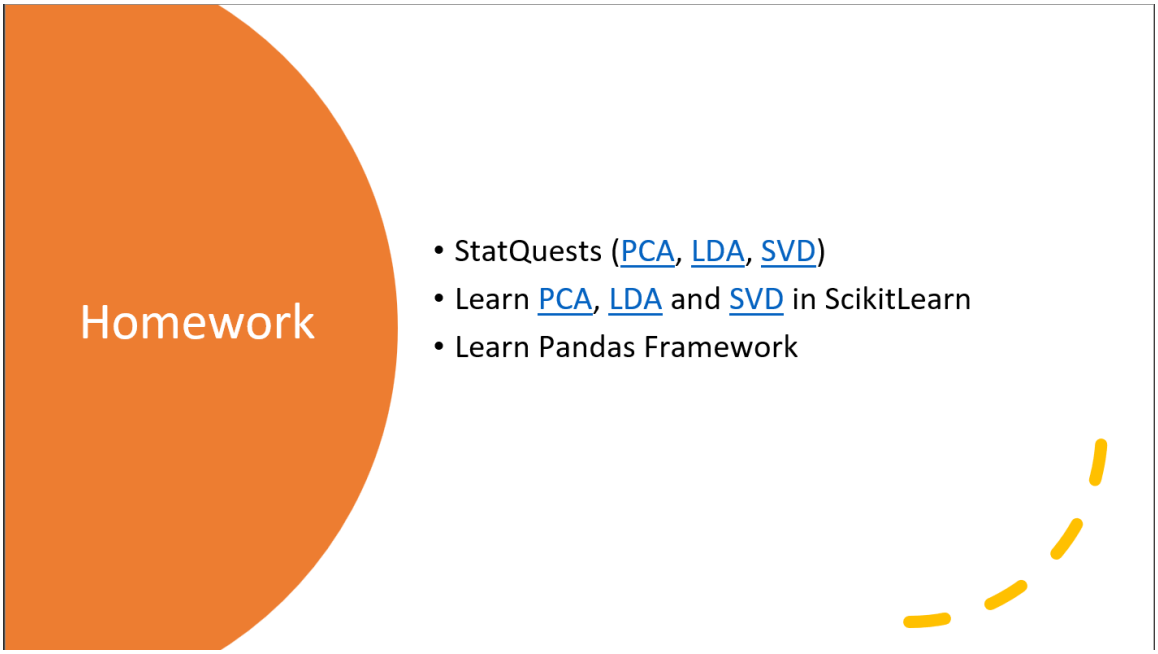
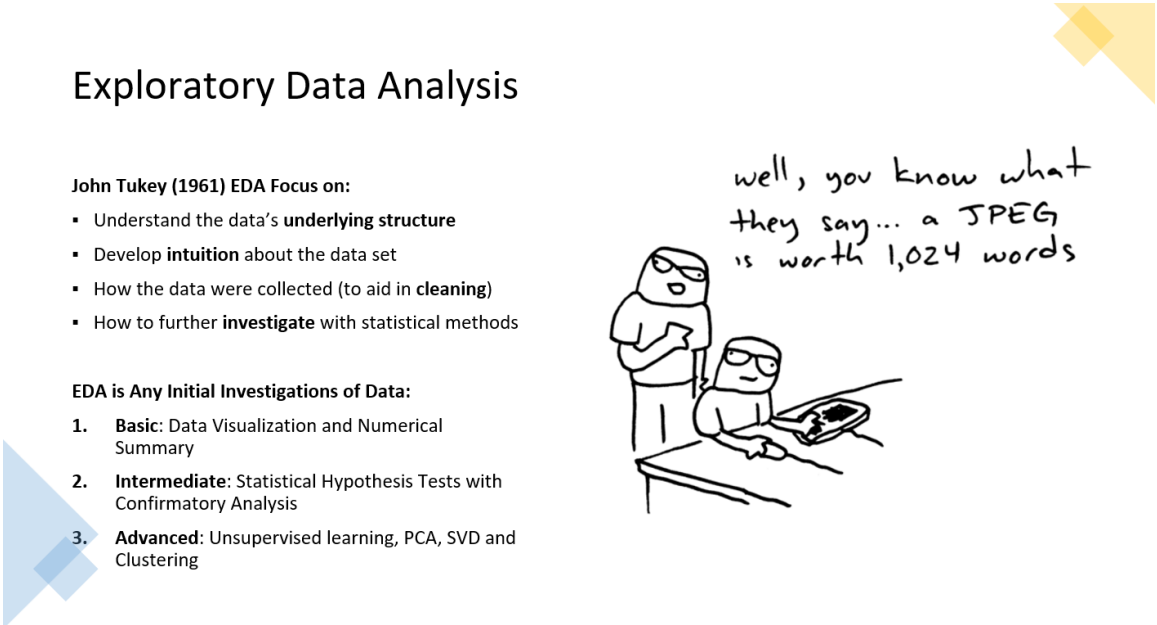
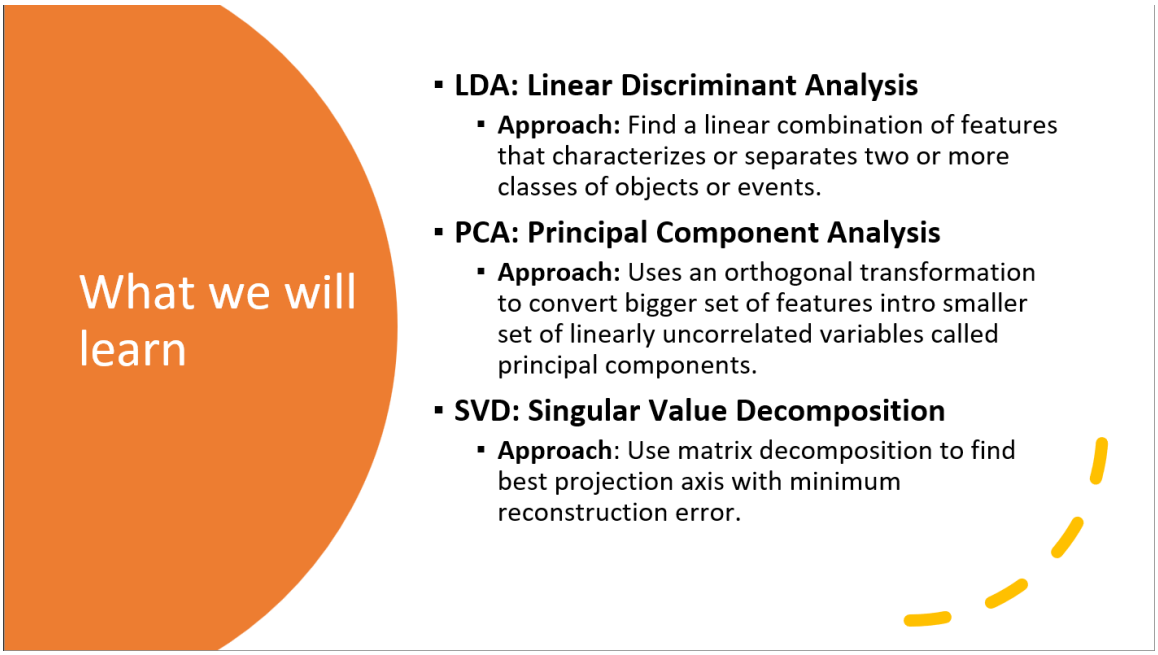
Muhammad Nur Maajid/1103228145



Eksplorasi Data atau EDA merupakan tahap yang sangat signifikan dalam proses analisis data. Ini memberikan cara untuk secara visual maupun kuantitatif memeriksa data, mengungkapkan pola, mengidentifikasi data yang tidak biasa atau keluar dari pola umum, menguji asumsi, dan membantu dalam pengembangan model awal. Fokus utama EDA adalah membantu para analis memahami karakteristik data dan memungkinkan mereka untuk membuat keputusan yang tepat selama tahapan analisis atau proses pemodelan selanjutnya.





1. PCA (Principal Component Analysis): PCA adalah teknik statistik yang digunakan untuk mengurangi dimensi data dengan cara mentransformasikan fitur-fitur asli ke fitur-fitur baru yang disebut sebagai komponen utama. Tujuan utama dari PCA adalah mengurangi dimensi data sambil mempertahankan sebanyak mungkin variasi yang terkandung dalam data. Komponen utama ini bersifat ortogonal satu sama lain, yang berarti mereka tidak berkorelasi. PCA berguna dalam mengatasi masalah multicollinearity dan memproyeksikan data ke ruang fitur yang lebih informatif.
2. LDA (Linear Discriminant Analysis): LDA adalah teknik yang digunakan untuk analisis statistik dan klasifikasi. Selain itu, LDA juga dapat digunakan untuk reduksi dimensi. Tujuannya adalah untuk memaksimalkan perbedaan antara kelas-kelas data, sehingga data dapat diproyeksikan ke ruang fitur yang memungkinkan pemisahan yang lebih baik antar kelas. Ini membuat LDA sering digunakan dalam masalah klasifikasi, terutama ketika Anda memiliki data yang terlabel.
3. SVD (Singular Value Decomposition): SVD adalah teknik faktorisasi matriks yang digunakan dalam berbagai aplikasi, termasuk sistem rekomendasi, analisis teks, dan reduksi dimensi. Dalam SVD, matriks data dibagi menjadi tiga matriks lainnya: matriks singular value, matriks left singular, dan matriks right singular. SVD berguna dalam mengungkapkan struktur intrinsik dalam data dan mengidentifikasi pola yang tersembunyi.