

Nama : Eka Milenia Ramadhani  
NIM : 1103228238

### **PCA StatQuest**

Teknik analisis data yang dikenal sebagai PCA (Principal Component Analysis) membantu mengatasi kompleksitas kumpulan data dengan cara mengurangi jumlah variabel menjadi beberapa faktor utama. Dengan demikian, PCA memfasilitasi pemahaman yang lebih baik terhadap struktur data dan penentuan variabel yang paling signifikan untuk memahami keseluruhan kumpulan data.

Proses PCA dimulai dengan proyeksi titik-titik data ke dalam garis, lalu mengukur jarak antara titik-titik tersebut dengan garis proyeksi. Komponen utama, disebut PC1, adalah garis yang memiliki jumlah kuadrat jarak terbesar antara titik proyeksi dan titik asal. Dengan mengurutkan komponen utama berdasarkan jumlah varians yang dijelaskan, PCA memungkinkan analisis distribusi variabilitas dalam data.

Selain itu, PCA juga memungkinkan reduksi dimensi dalam kumpulan data dengan menggabungkan berbagai bagian data, seperti sampel gen, dan menghitung kombinasi linier yang menghasilkan nilai tunggal terbesar, yang disebut nilai eigen. Vektor eigen ini digunakan untuk plot PCA, membantu mengidentifikasi variabel yang paling berkontribusi dalam menjelaskan data.

StatQuest menyajikan penjelasan yang jelas tentang perhitungan komponen utama dan cara memanfaatkannya untuk membuat plot data 2D yang lebih sederhana. Mereka juga memberikan tutorial tentang transformasi plot 2D menjadi grafik 3D untuk memudahkan pemahaman.

Dengan penerapan PCA, pemahaman struktur data menjadi lebih mudah dan memungkinkan pengambilan keputusan yang lebih baik berdasarkan pemahaman yang lebih mendalam tentang variabilitas dan keterkaitan dalam data.