

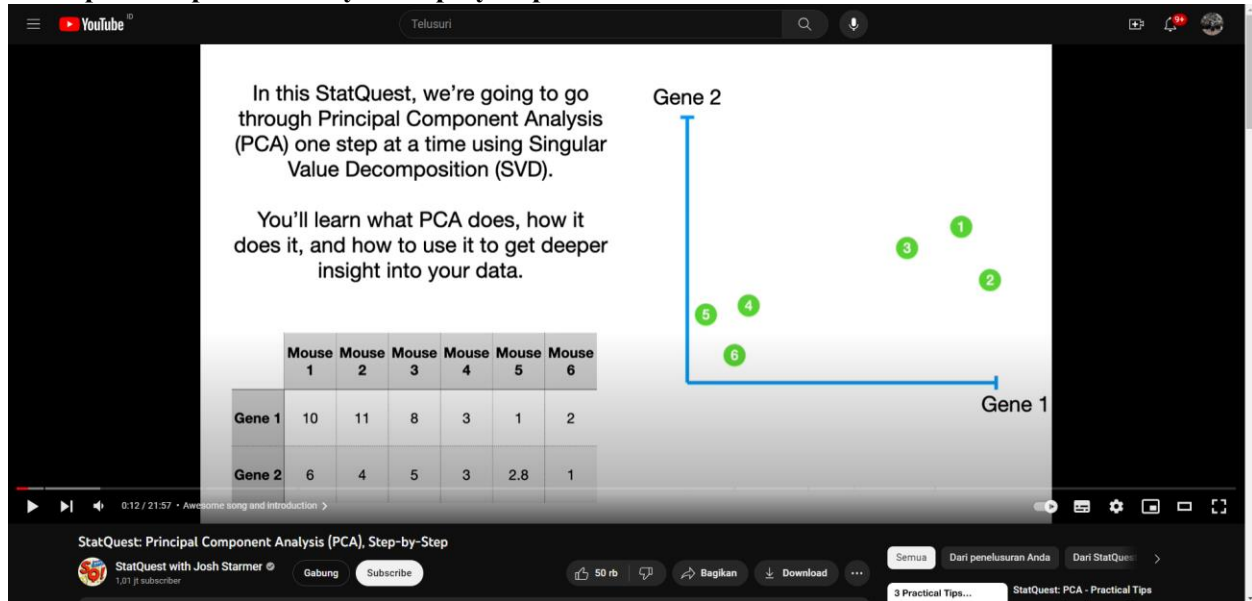
M Yusril Fauzan Hrp

1103204015

## Understanding 3 Link StatQuest

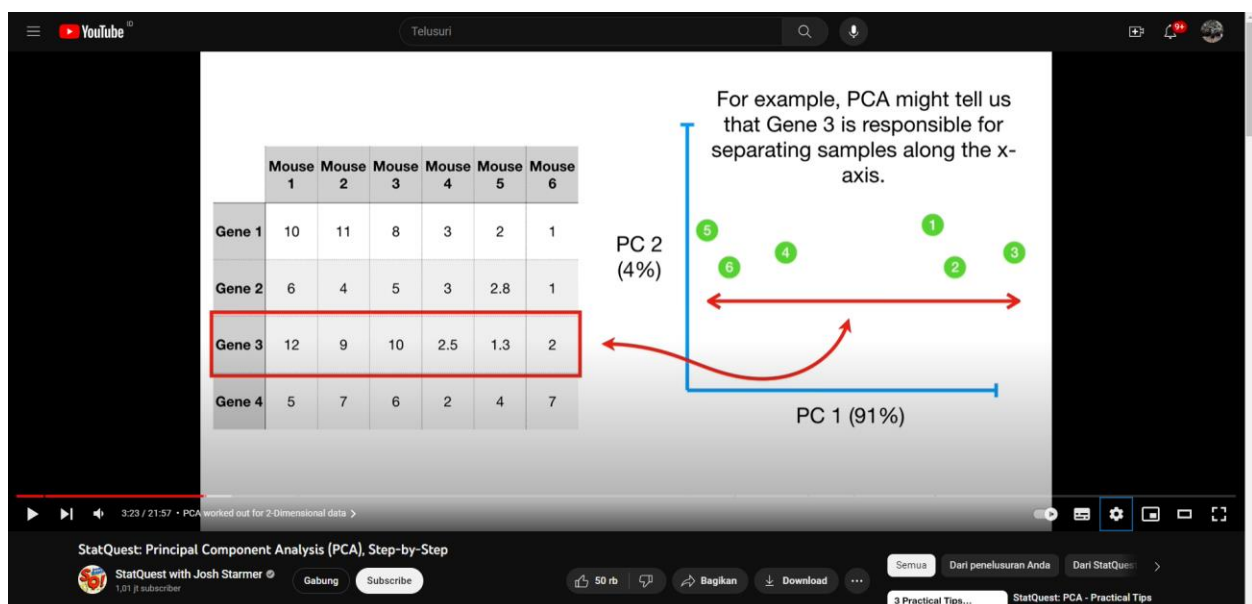
Based on Youtube : Josh starmer

### 1. Principal Component Analysis Step by step



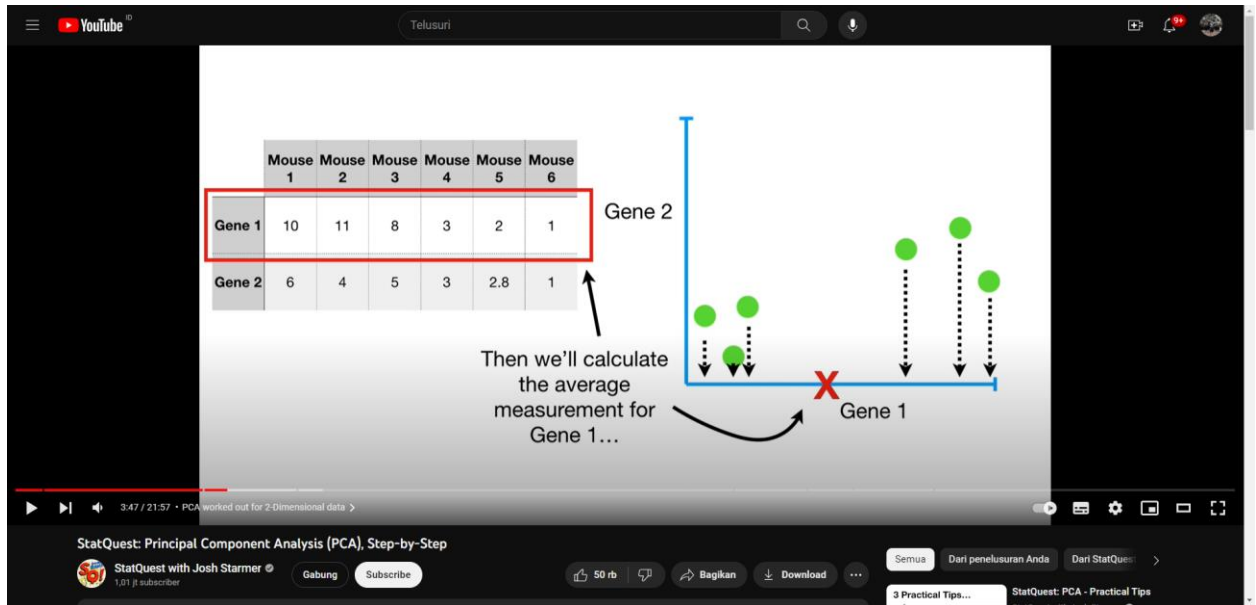
Principal Component Analysis merupakan suatu teknik statistik yang berguna dalam mengurangi dimensi data serta menjaga variabilitas data asli. Metode ini dapat diterapkan pada analisis data genetika "gene 1" dan "gene 2".

- Penerapan PCA pada sumbu x

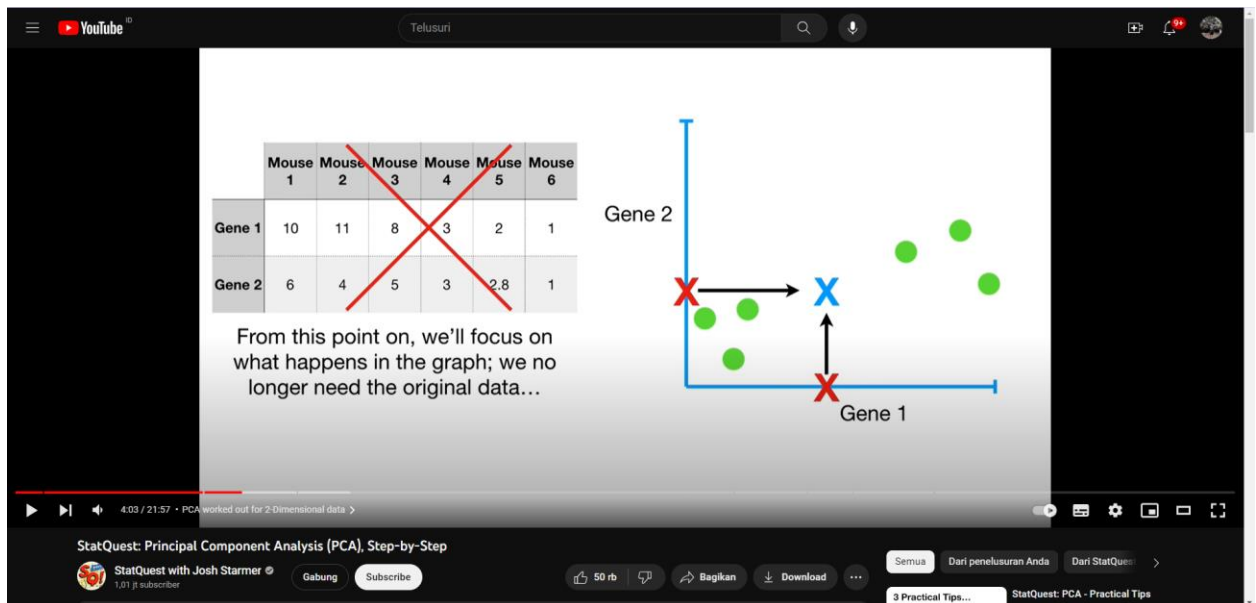


PC 1 akan menjadi arah dengan variabilitas tertinggi, dan PC 2 akan menjadi arah dengan variabilitas kedua tertinggi. PCA akan mengubah data dua dimensi menjadi data dua dimensi yang lebih kompak, tetapi mempertahankan sebagian besar variabilitas asli.

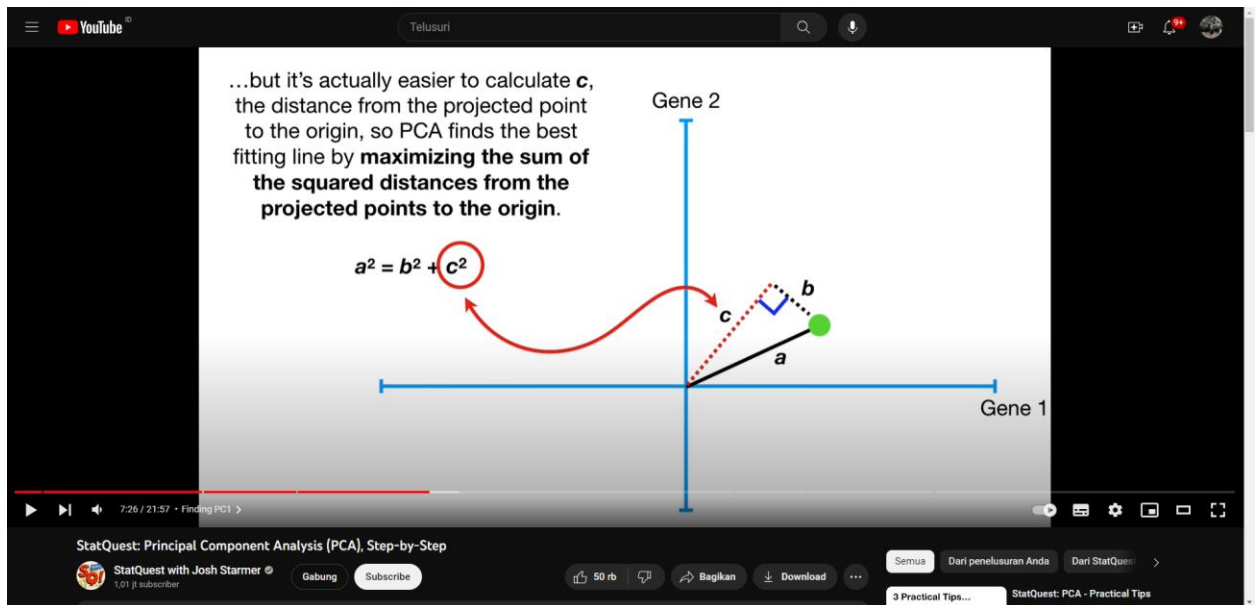
- PCA analytic for 2 dimensional



Standarisasi data untuk memastikan kedua variabel (x dan y) memiliki skala yang sama, biasanya dengan mengurangi rata rata dan membaginya dengan standar deviasi.

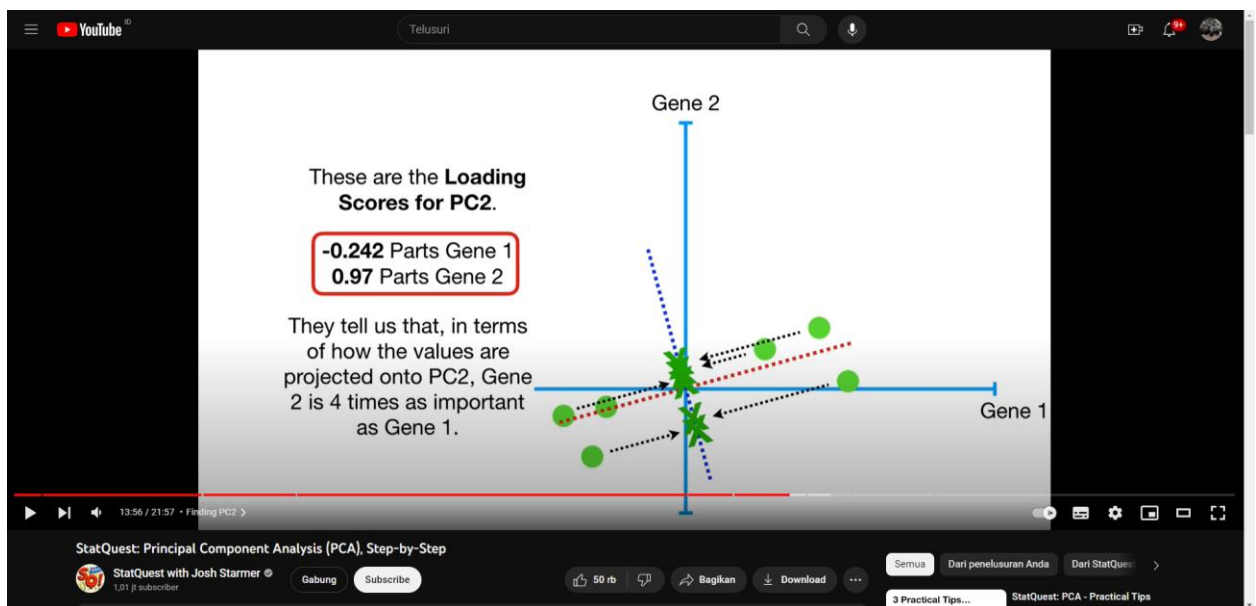


- Find PC 1



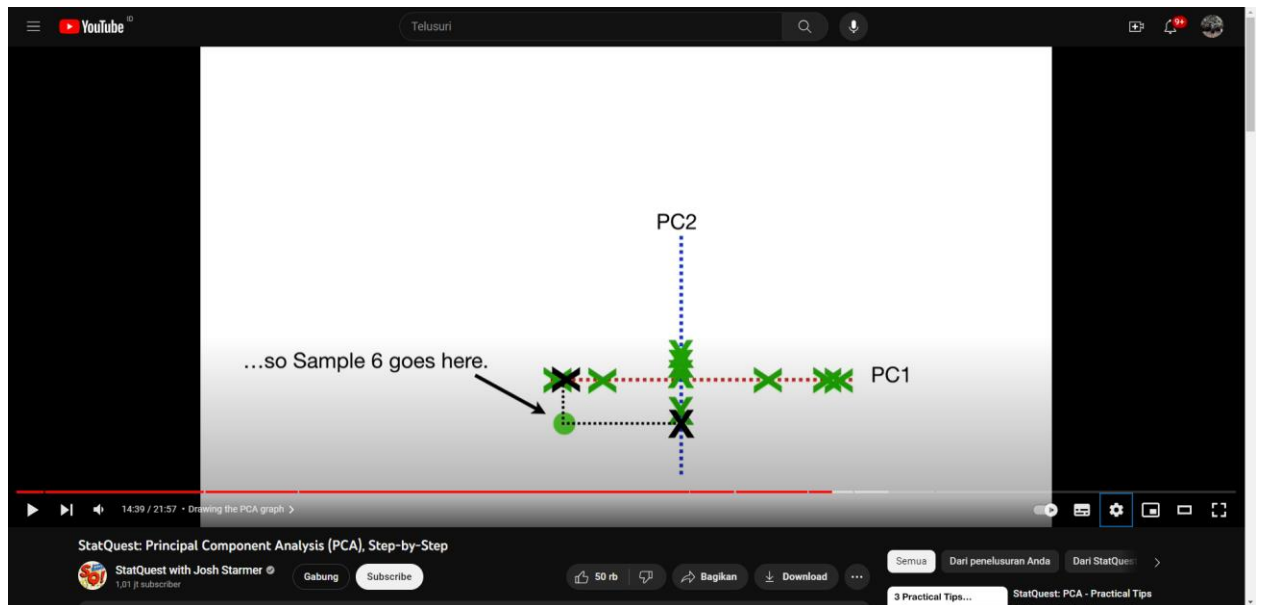
PC 1 diekstrak dari dataset dengan mengambil produk data standar yang asli dengan vektor yang sesuai dengan nilai terbesar.

- Find PC 2



Menghitung PC 2 dari data yang telah dipotong pada PC 1

- Draw PCA Graph



Jika poin poin data dalam grafik terpisah, maka semakin baik komponen utama dalam menjelaskan variasi dalam data