

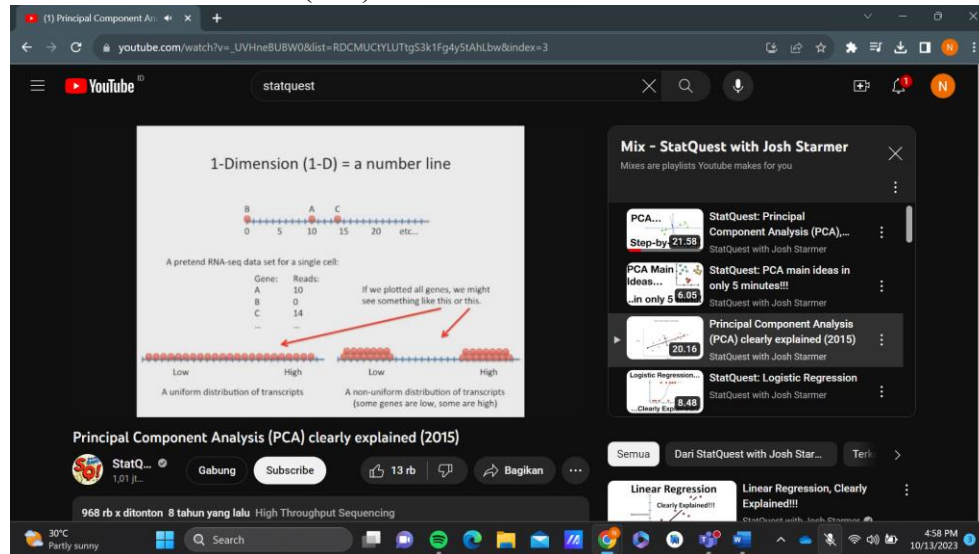
Nama : Nabilah Salwa

NIM : 1103204060

Understanding 3 Link StatQuest (Youtube: Josh Starmer)

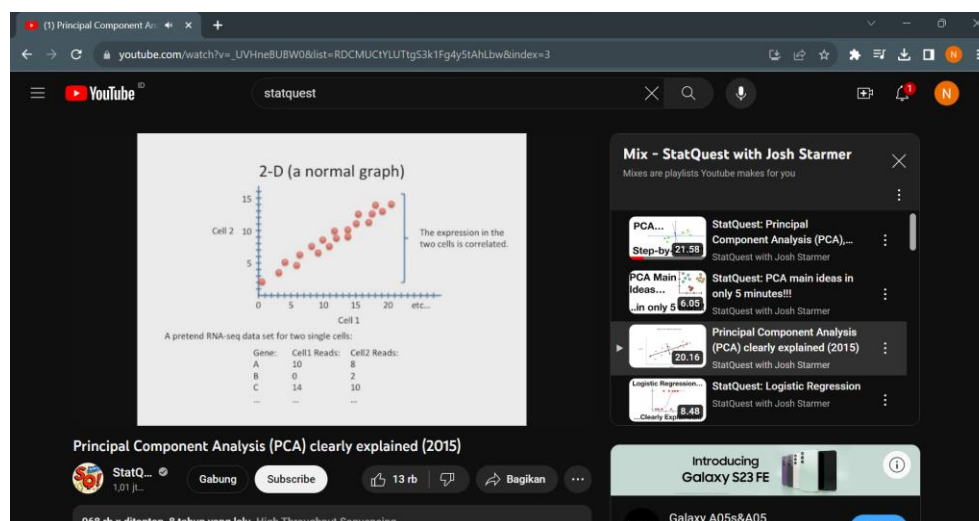
1. Principal Component Analysis (PCA) Clearly Explained (2015)

- PCA 1-Dimension (1-D) = a number line



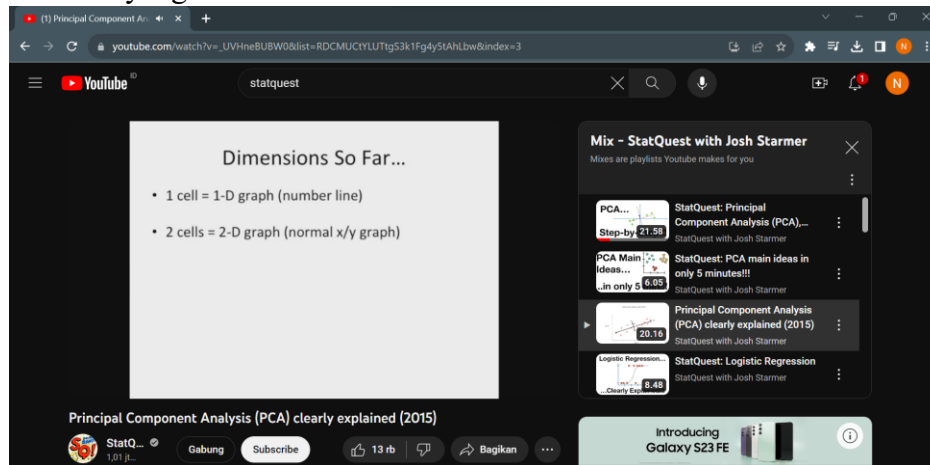
PCA 1-D adalah teknik yang mengurangi dimensi data dengan mencari arah perubahan yang paling besar dalam data dengan menggunakan contoh RNA-seq untuk single cell dengan melakukan plottingan genes nya.

- PCA 2-D



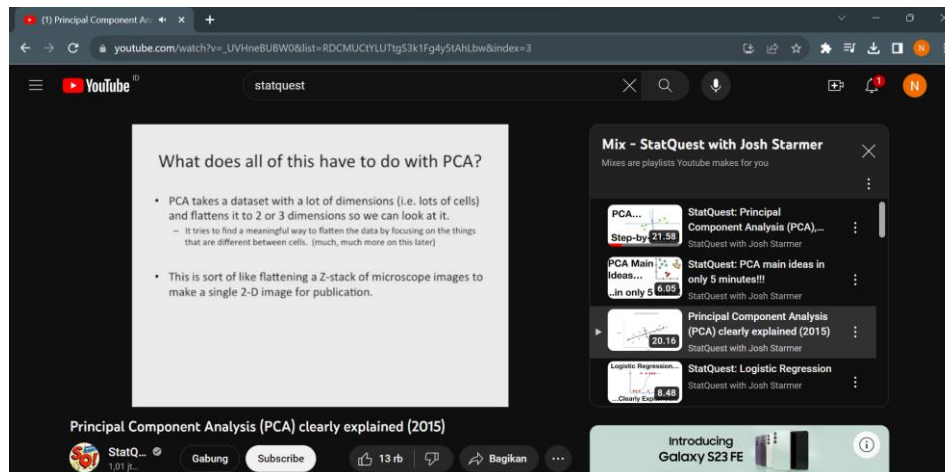
Masih menggunakan dataset yang sama dengan D-1, tapi disini dapat dilihat berbagai data yang memiliki korelasi dan data yang tidak memiliki korelasi pada genes dan cell reads nya

- Dimensi yang diketahui



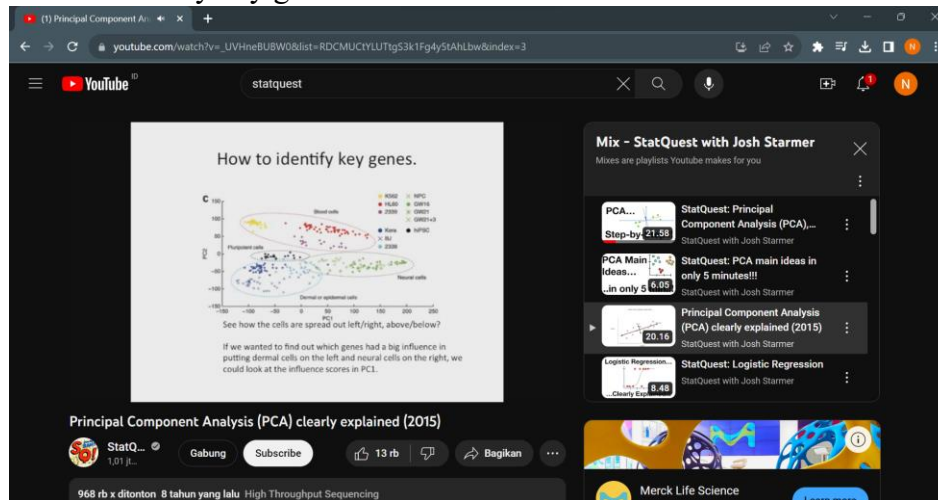
Disini dijelaskan bahwa setiap sell akan menambahkan dimensi dari graphnya.

- What does all of this have to do with PCA ?



PCA menggunakan dataset yang banyak dimensinya dan bisa diperkecil menjadi 2 atau 3 dimensi.

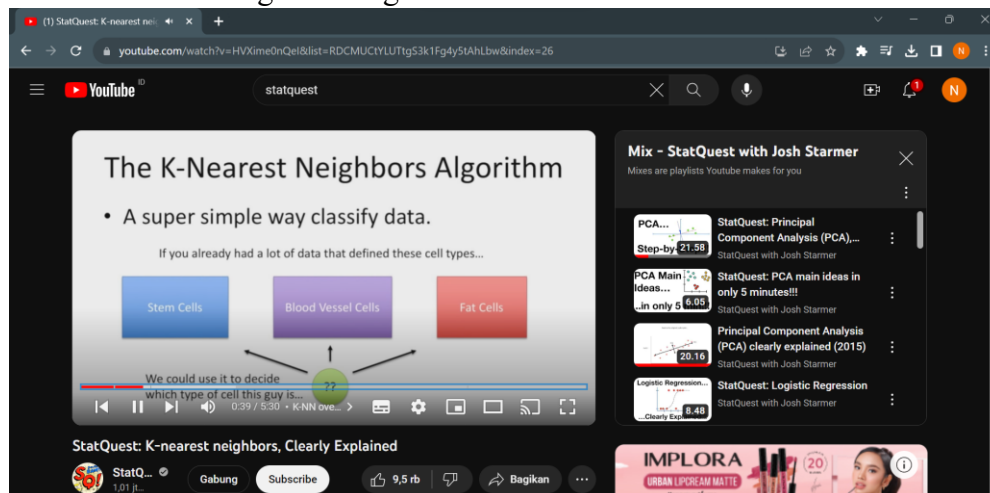
- How to identify key genes



Dengan cara melihat gen mana yang mempunyai pengaruh besar dalam menempatkan sel kulit disebelah kiri dan sel saraf disebelah kanan, kita dapat melihat skor pengaruhnya di PC1.

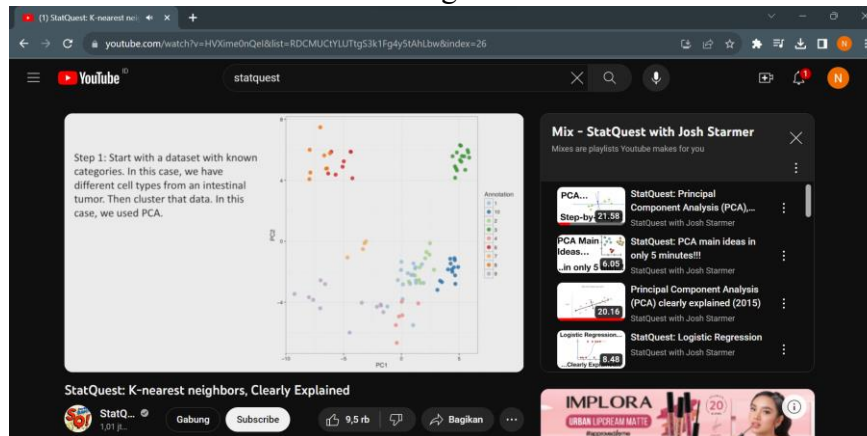
2. StatQuest: K-nearest neighbors, Clearly explained

- The K-nearest Neighbors Algorithm



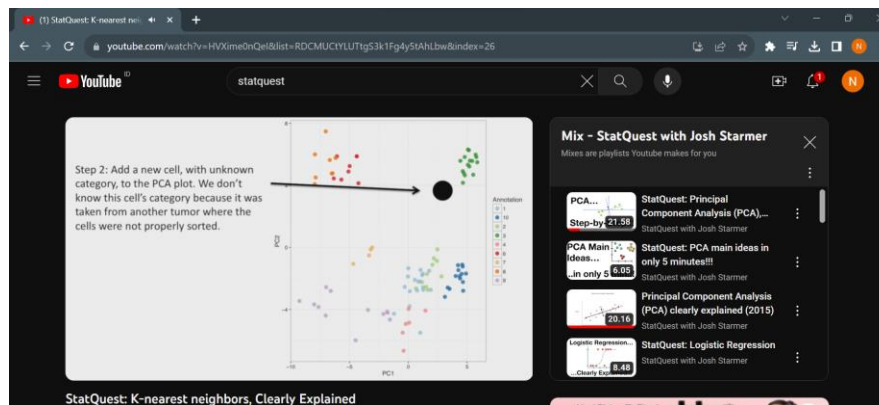
Ini digunakan untuk mengklasifikasikan data yang bentuknya seperti stem cells, blood vessel cell, dan fat cell untuk mengetahui cell mana yang masuk ke kategori mana

- Star with dataset with known categories

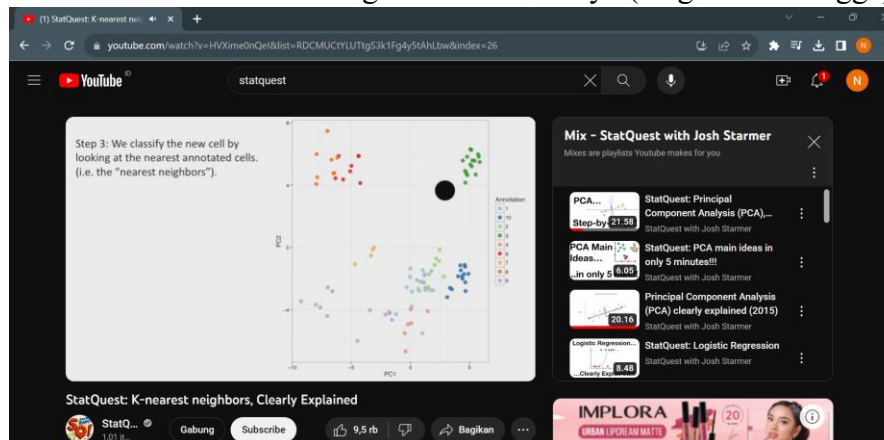


Disini ada beberapa tipe cell yang berbeda, datanya diclusterkan menggunakan PCA.

- Menambahkan cell baru

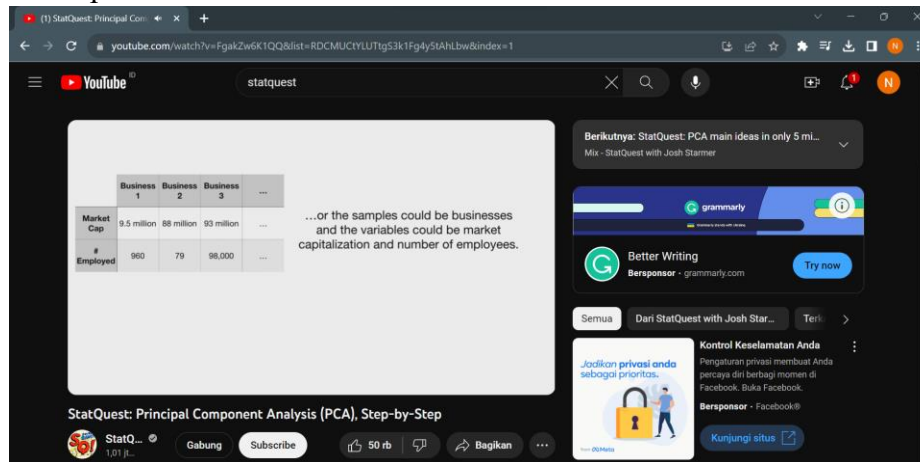


- Klasifikasi cell baru itu dengan cell terdekatnya (neighbors/tetangga)

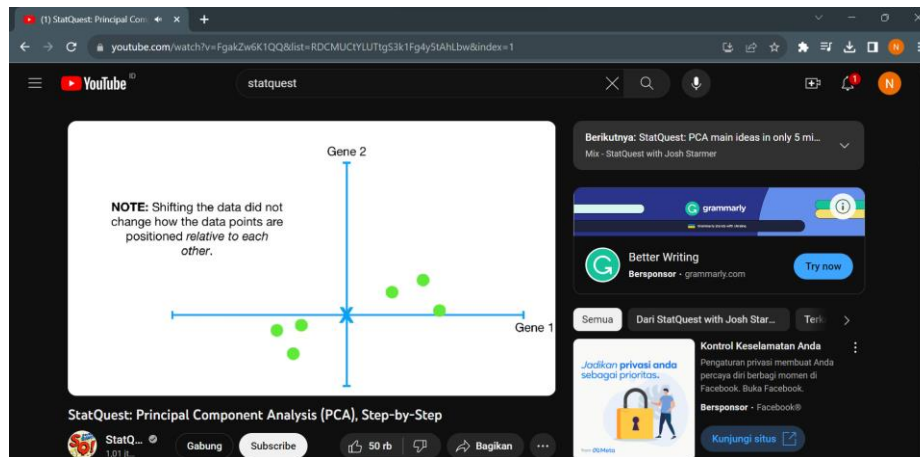


3. Principal Component Analysis (PCA) Step-by-step

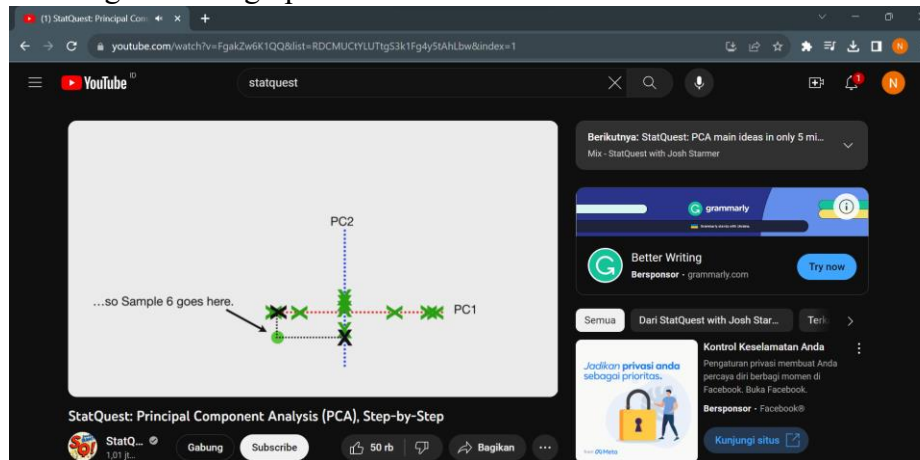
- Conceptual motivation for PCA



- Worked out for 2-dimension data



- Drawing the PCA graph



- PCA worked for 3 dimensional; data

