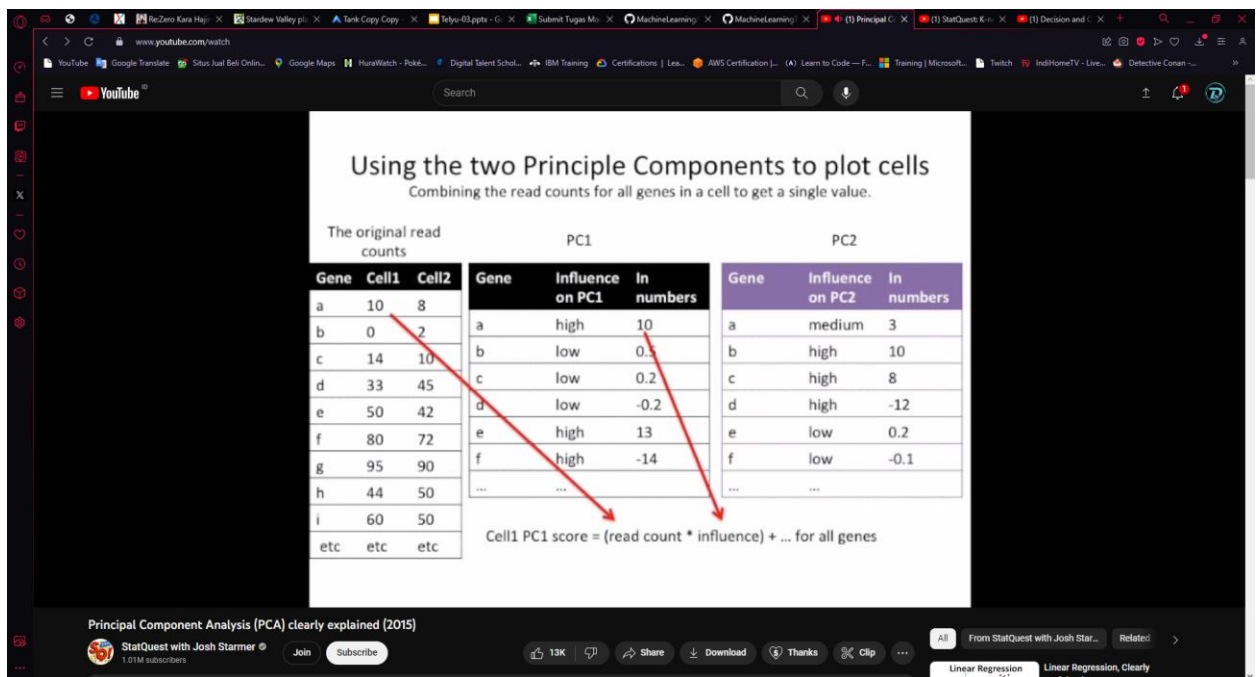
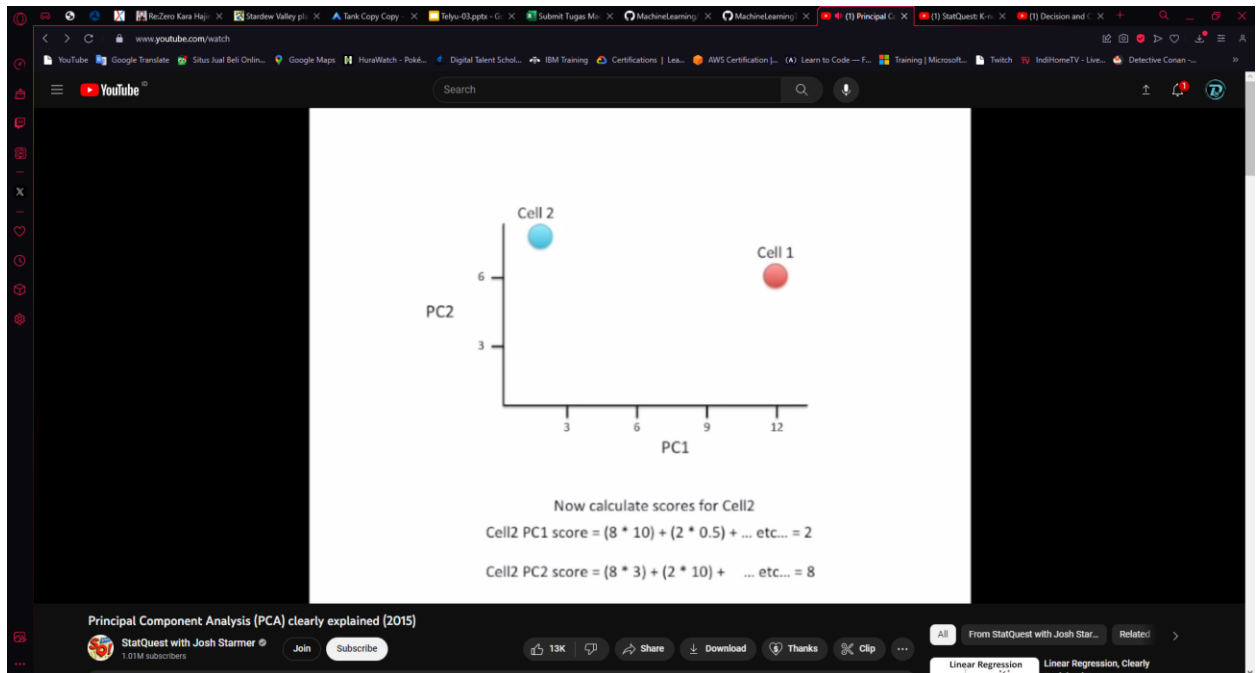


Muhammad Bustanul Arifin Nur Fathurrahman

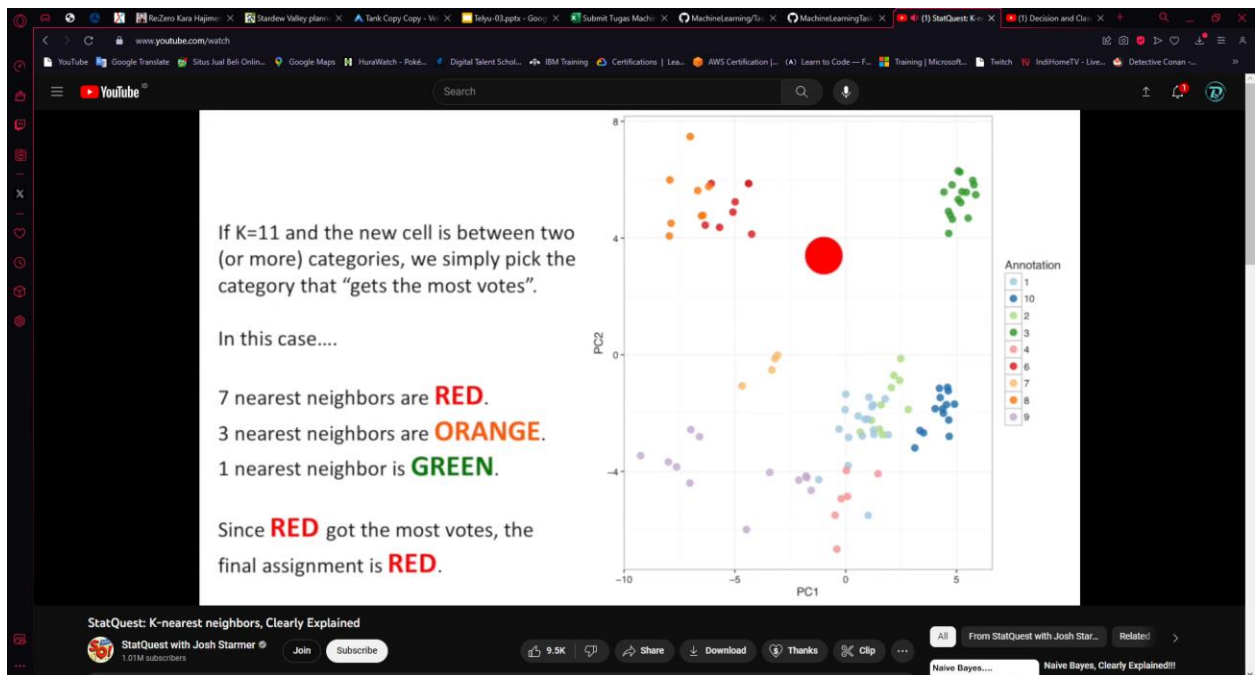
1103228224

## 1. PCA



PCA adalah metode penyampaian data dari database yang besar sehingga apabila ditampilkan dengan cara biasa maka tidak akan cukup. Tetapi karena data-data yang disampaikan itu penting, jadi saat ditampilkan data yang ditampilkan tidak boleh jauh berbeda dari data yang sebenarnya. Untuk itu cara yang dapat dilakukan adalah menarik garis lurus dari dua ujung data lalu ditentukan pengaruhnya, apakah besar atau tidak. Setelah itu data tadi dikali dengan pengaruh tersebut dan ditambah dengan semua kolom yang ada, hasil tersebut lah yang nantinya ditampilkan pada grafik dengan hanya 1 titik mewakili 1 kolom.

## 2. KNN



KNN atau K nearest neighbors merupakan metode untuk mengetahui cell yang tidak diketahui merupakan anggota cell mana, hal ini dilakukan dengan melihat cell terdekatnya. Apabila hanya dekat dengan 1 cell, maka cell tersebut menjadi bagian cell tersebut. Tetapi jika berada di Tengah-tengah, maka dilakukan “voting” dengan melihat jumlah “tetangga” terbanyak, seperti pada gambar di atas adalah merah.

### 3. Decision tree

The video shows a decision tree for the feature "Popcorn". The root node is "Popcorn" with a heart icon. It branches into "True" and "False".

- True branch:** Leaf node "Cool As Ice" with counts: Yes=1, No=3.
- False branch:** Leaf node "Cool As Ice" with counts: Yes=2, No=1.

Below the tree, the Gini Impurity for a Leaf is calculated:

$$\text{Gini Impurity for a Leaf} = 1 - (\text{the probability of "Yes"})^2 - (\text{the probability of "No"})^2$$
$$= 1 - \left(\frac{1}{1+3}\right)^2 - \left(\frac{3}{1+3}\right)^2$$
$$= 0.375$$

An arrow points from the text "And when we do the math, we get 0.375." to the result 0.375.

Video title: Decision and Classification Trees, Clearly Explained!!!  
Channel: StatQuest with Josh Starmer  
Chapters: Building a tree with Gini

Decision tree merupakan metode pengambilan keputusan berdasarkan hal-hal yang telah dipertimbangkan. Untuk pernyataan yang dijadikan acuan, dapat dipilih dengan cara menggunakan Gini Impurity dengan hasil dengan gini impurity terendah merupakan pernyataan yang dijadikan patokannya.

The video shows a decision tree for the feature "Loves Soda". The root node is "Loves Soda" with a heart icon. It branches into "True" and "False".

- True branch:** Leaf node "Cool As Ice" with counts: Yes=3, No=1.
- False branch:** Leaf node "Cool As Ice" with counts: Yes=0, No=3.

Below the tree, the Gini Impurity for Loves Soda is calculated:

$$\text{Gini Impurity for Loves Soda} = 0.214$$

An arrow points from the text "And because Loves Soda has the lowest the Gini Impurity overall..." to the result 0.214.

Below this, another decision tree is shown for the feature "Age < 15".

- True branch:** Leaf node "Cool As Ice" with counts: Yes=0, No=3.
- False branch:** Leaf node "Cool As Ice" with counts: Yes=3, No=2.

Below this tree, the Gini Impurity for Age < 15 is calculated:

$$\text{Gini Impurity for Age < 15} = 0.343$$

Video title: Decision and Classification Trees, Clearly Explained!!!  
Channel: StatQuest with Josh Starmer  
Chapters: Numeric and continuous