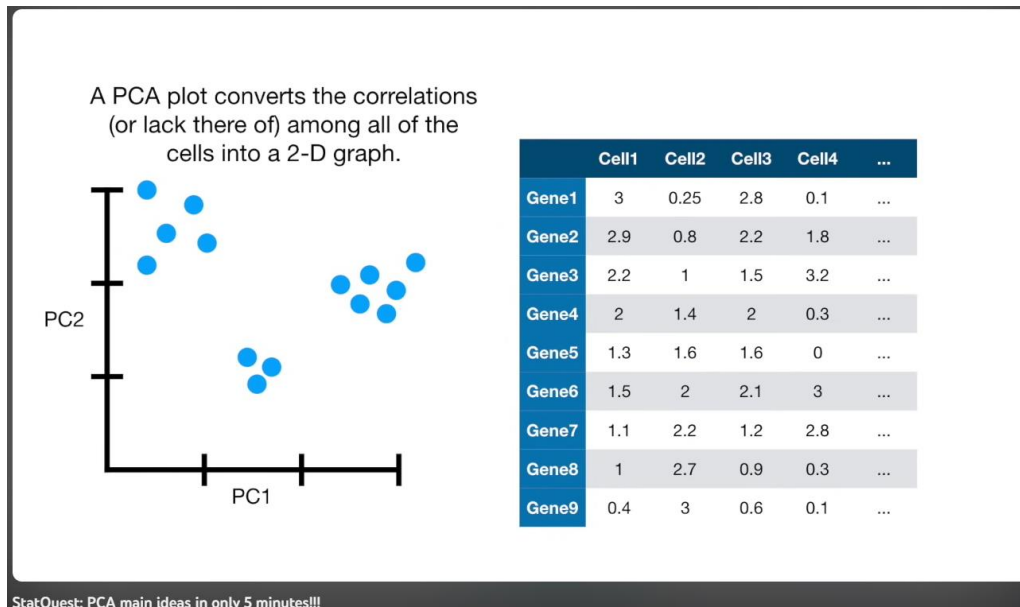


Muhammad Amirul Arsyad Arrayyan

1103204207

Understanding 3 link statquest (scan your note as PDF)

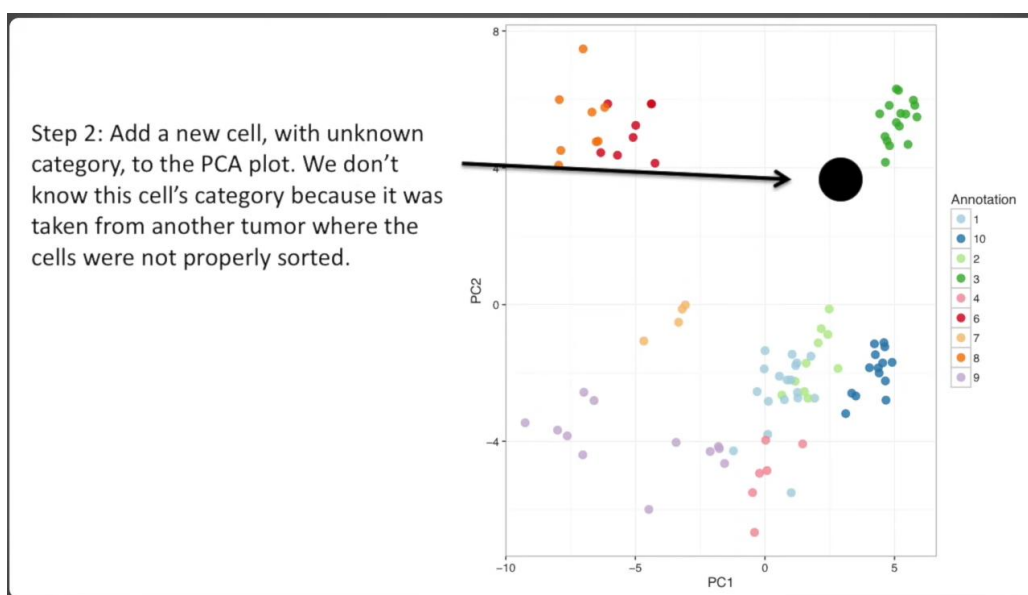
Vid1: StatQuest: PCA main ideas in only 5 minutes!!!



Menjelaskan bahwa dengan PCA dapat merencanakan analisis komponen utama dengan mengubah korelasi antara semua sel menjadi grafik 2D yang sebelumnya kemungkinan akan bertambah 1 dimensi setiap kolom data yang bertambah 1

Cara menafsirkan grafik PCA: Perbedaan PCA(PC 1) lebih penting daripada perbedaan Panjang PCA(PC 2), perbedaan antar 2 cluster berbeda di PC1 lebih berbeda dibanding 2 cluster berbeda di PC 2

Vid2: StatQuest: K-nearest neighbors, Clearly Explained

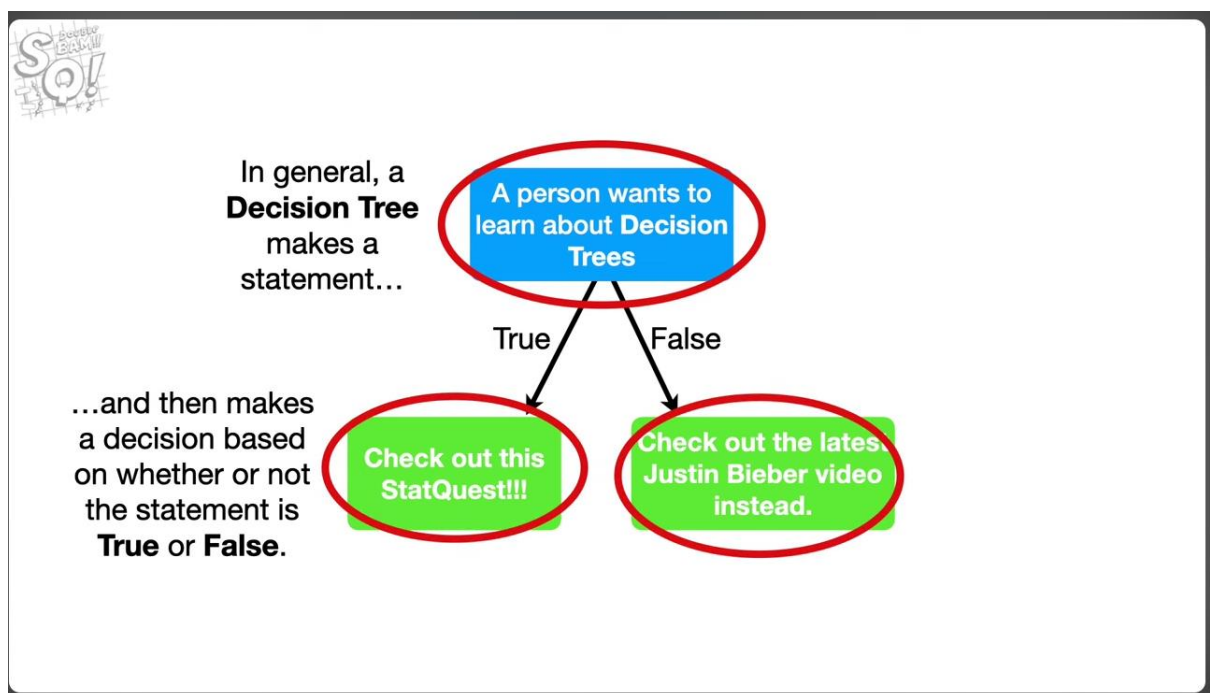


KNN mengklasifikasikan item baru yang tidak diketahui kategorinya dengan mencocokkan item baru tersebut dengan tetangga terdekatnya, K didefinisikan sebagai jumlah neighbour yang akan dibandingkan dengan item baru, apabila $K = 11$ maka item tersebut harus dibandingkan dengan 11 tetangganya.

Penentuan jumlah K tidak memiliki sebuah patokan, sehingga harus dilakukan uji coba terhadap value K untuk melihat mana yang paling cocok

Jika terlalu sedikit akan menimbulkan gangguan dan terkena efek outlier, sementara value yang terlalu besar memang dapat memperhalus berbagai hal tetapi jika salah satu kategori pada K memiliki jumlah yang kecil maka ia akan selalu kalah dengan kategori dengan jumlah besar.

Vid3: Decision and Classification Trees, Clearly Explained!!!



Decision Tree merupakan sebuah keputusan dari sebuah statement, umumnya keputusan tersebut terdiri dari "true" dan "false" dengan statement keputusan di setiap pilihannya.

Jika Decision Tree menunjukkan value numerik maka disebut sebagai Regression Tree

Statement yang bersifat True maka akan berada di sebelah kiri dan False di sebelah kanan