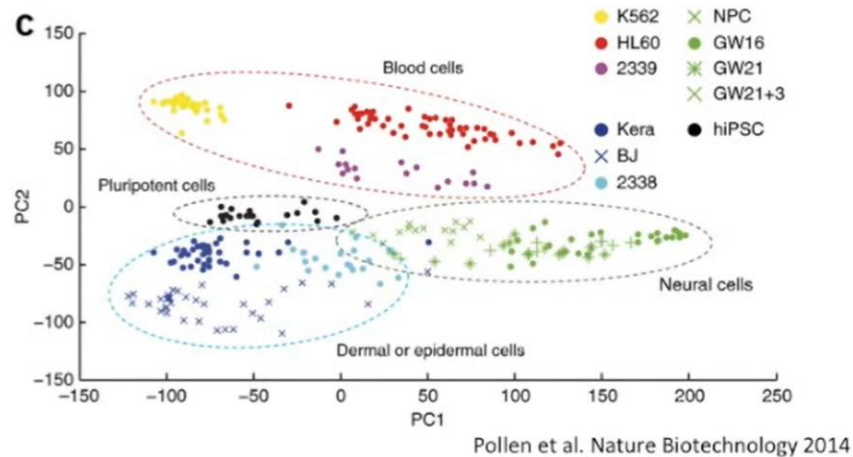


Nama : April Hamonangan Marbun

Kelas : TK4401

NIM : 1103202039

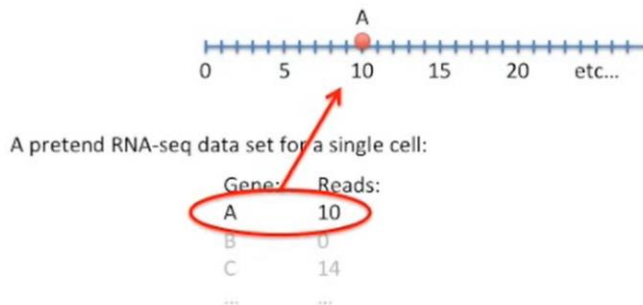
This PCA plot shows clusters of cell types.



PCA, atau Principal Component Analysis, adalah sebuah metode statistik yang digunakan untuk mengurangi dimensi data. Tujuan utama dari PCA adalah untuk mengidentifikasi pola dan struktur yang ada dalam data dengan mengurangi jumlah fitur (dimensi) dalam data, sambil mempertahankan sebanyak mungkin informasi yang relevan.

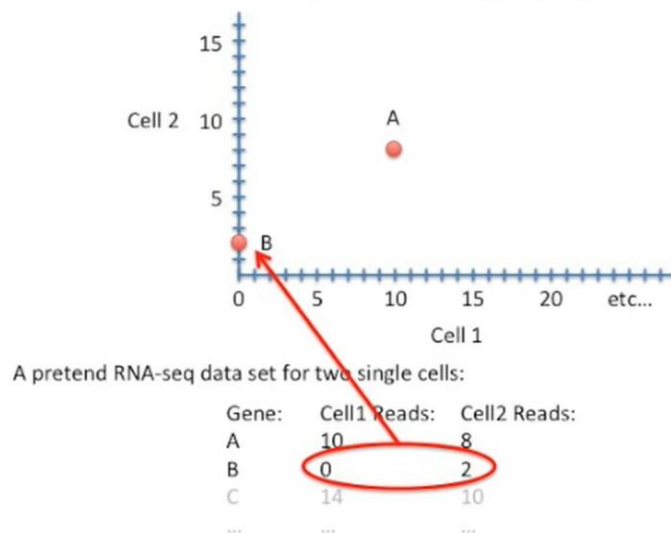
PCA bekerja dengan mentransformasikan data asli ke dalam sistem koordinat baru yang disebut "komponen utama" (principal components). Komponen utama ini adalah kombinasi linear dari fitur asli yang disusun sedemikian rupa sehingga komponen pertama mengandung sebagian besar variabilitas dalam data, diikuti oleh komponen kedua, dan seterusnya. Dengan demikian, PCA membantu mengidentifikasi pola yang mendasari data dan memungkinkan representasi data yang lebih sederhana dengan sedikit fitur yang tetap mempertahankan sebagian besar informasi penting. PCA sering digunakan dalam analisis data, pengolahan citra, kompresi data, dan berbagai aplikasi lainnya untuk mengatasi masalah high-dimensionality dan mengidentifikasi hubungan antara fitur-fitur dalam data.

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



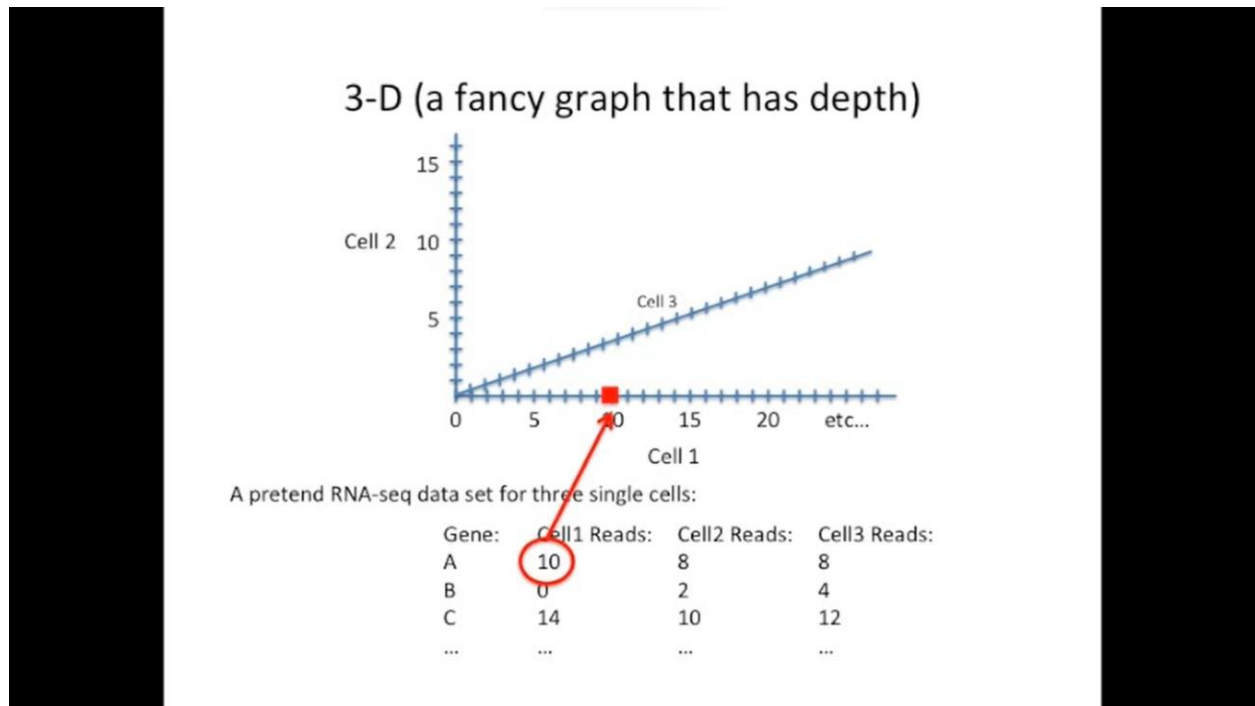
1-Dimension (1-D) atau dimensi satu adalah representasi matematis dari ruang yang hanya memiliki satu arah. Ini sering digambarkan sebagai "garis angka" atau "number line" yang terdiri dari titik-titik yang berjajar dalam satu baris lurus. Pada number line ini, setiap titik memiliki koordinat tunggal yang mewakili posisinya pada garis tersebut. Number line ini digunakan untuk merepresentasikan kuantitas satu dimensi seperti bilangan real, waktu, atau variabel tunggal. Contohnya, jika Anda ingin mengukur panjang atau jarak, Anda dapat menggunakan number line untuk menyatakan nilai dalam satu dimensi.

## 2-D (a normal graph)



Grafik distribusi normal 2-D (dua dimensi) mengacu pada grafik yang menunjukkan distribusi probabilitas untuk dua variabel acak yang secara bersamaan terdistribusi normal. Grafik distribusi normal 2-D ini umumnya menggambarkan kontur atau permukaan fungsi kepadatan probabilitas untuk distribusi normal bivariat ini. Biasanya digambarkan dalam bentuk elips dalam dua dimensi, dengan pusat elips mewakili rata-rata dari kedua variabel, dan bentuk serta orientasi elips ditentukan oleh varians dan kovariansi kedua variabel.

Grafik ini umumnya digunakan dalam statistik dan analisis data untuk memvisualisasikan hubungan antara dua variabel dan memahami bagaimana kedua variabel tersebut didistribusikan bersama. Hal ini membantu dalam menganalisis korelasi, kovarians, dan probabilitas yang terkait dengan distribusi bersama dari kedua variabel tersebut.



Grafik distribusi normal 3-D (tiga dimensi) adalah representasi grafis dari distribusi probabilitas untuk tiga variabel acak yang secara bersamaan terdistribusi normal. Grafik ini digunakan untuk memvisualisasikan bagaimana ketiga variabel tersebut berkorelasi dan didistribusikan bersama dalam ruang tiga dimensi.

Grafik distribusi normal 3-D umumnya digambarkan dalam bentuk permukaan yang memperlihatkan distribusi probabilitas dalam ruang tiga dimensi. Permukaan ini sering kali berbentuk seperti lonceng atau berbentuk bola terdistorsi, dengan pusatnya mewakili nilai rata-rata ketiga variabel tersebut. Bentuk dan orientasi permukaan ini ditentukan oleh varians dan kovariansi ketiga variabel.

Penggunaan grafik distribusi normal 3-D adalah untuk memahami sejauh mana ketiga variabel tersebut berkorelasi dan bagaimana mereka terdistribusi bersama dalam ruang tiga dimensi. Hal ini berguna dalam statistik, ilmu data, dan penelitian ilmiah di mana ada kebutuhan untuk memahami hubungan antara tiga variabel acak yang terdistribusi normal.

## Dimensions So Far...

- 1 cell = 1-D graph (number line)
- 2 cells = 2-D graph (normal x/y graph)
- 3 cells = 3-D graph (fancy graph with depth)
- 4 cells = 4-D graph (you can't draw it)
- 200 cells = 200-D graph (etc..)

Are all those dimensions super important? Or are some more important than others?

Grafik distribusi normal 4-D (empat dimensi) adalah representasi grafis dari distribusi probabilitas untuk empat variabel acak yang secara bersamaan terdistribusi normal. Ini adalah representasi yang lebih kompleks dibandingkan dengan grafik distribusi normal 2-D atau 3-D karena melibatkan ruang empat dimensi.

Dalam grafik distribusi normal 4-D, Anda akan memiliki tiga dimensi spasial (seperti panjang, lebar, dan tinggi) dan dimensi keempat mewakili nilai probabilitas atau tingkat kepadatan dalam ruang tersebut. Visualisasi 4-D bisa menjadi sangat sulit untuk dimengerti secara langsung oleh manusia, jadi sering kali diperlukan perangkat lunak khusus atau teknik visualisasi yang lebih canggih. Grafik distribusi normal 4-D digunakan dalam situasi di mana ada empat variabel acak yang ingin dianalisis bersama-sama, seperti dalam beberapa aplikasi ilmu data dan penelitian statistik yang lebih kompleks. Tujuannya adalah memahami hubungan antara keempat variabel tersebut dan bagaimana mereka didistribusikan bersama dalam ruang empat dimensi.