



VISUALISASI

Visualisasi adalah konversi data ke dalam format visual atau tabel sehingga karakteristik dari data dan relasi diantara item data atau atribut dapat di analisis atau dilaporkan

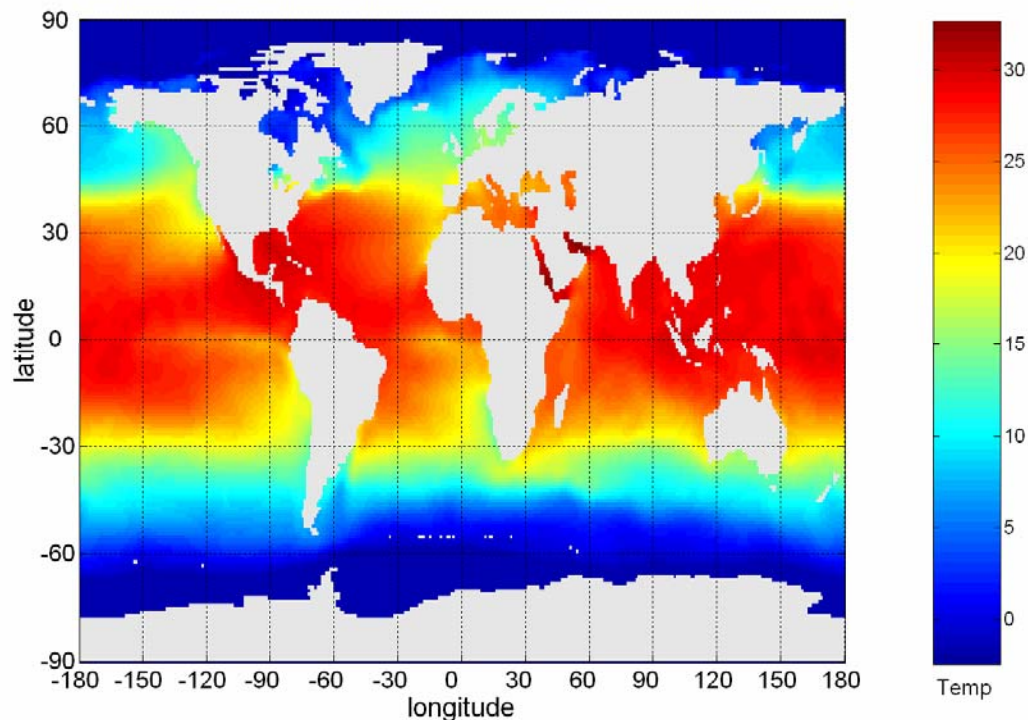
Visualisasi data adalah satu dari yang teknik paling baik dan menarik untuk eksplorasi data

Manusia memiliki kemampuan membangun yang baik untuk menganalisis sejumlah besar informasi yang dipresentasi secara visual

Ia dapat mendeteksi pola umum dan trend, pencilaan dan pola yang tidak umum

Gambar berikut ini menunjukkan suhu permukaan laut pada bulan Juli tahun 1982

Sepuluh dari ribuan titik data diringkas dalam satu gambar



Representasi



Representasi adalah pemetaan informasi menjadi format visual

Objek data, atribut-atributnya dan relasi diantara objek-objek data diterjemahkan ke dalam elemen grafis seperti titik, garis, bentuk-bentuk tertentu dan warna

Contoh :

Objek-objek sering direpresentasikan sebagai titik

Nilai atribut-atributnya dapat direpresentasikan sebagai posisi dari titik-titik atau karakteristik dari titik sebagai contoh warna, ukuran dan bentuk

Jika posisi di gunakan, maka relasi pada titik , apakah terbentuk dalam group atau sebuah titik pencilan, dapat dengan mudah dilihat

Penyusunan/Arrangement

Adalah penempatan elemen-elemen visual
diantara

Dapat membuat perbedaan besar mengenai
bagaimana mudahnya untuk memahami data

	1	2	3	4	5	6
1	0	1	0	1	1	0
2	1	0	1	0	0	1
3	0	1	0	1	1	0
4	1	0	1	0	0	1
5	0	1	0	1	1	0
6	1	0	1	0	0	1
7	0	1	0	1	1	0
8	1	0	1	0	0	1
9	0	1	0	1	1	0

	6	1	3	2	5	4
4	1	1	1	0	0	0
2	1	1	1	0	0	0
6	1	1	1	0	0	0
8	1	1	1	0	0	0
5	0	0	0	1	1	1
3	0	0	0	1	1	1
9	0	0	0	1	1	1
1	0	0	0	1	1	1
7	0	0	0	1	1	1

Seleksi



Seleksi adalah : eliminasi atau penekanan kembali dari beberapa objek tertentu dan atribut-atribut

Seleksi juga merupakan pemilihan subset dari atribut

- Reduksi secara dimensi sering digunakan untuk mengurangi jumlah dimensi menjadi 2 atau 3 dimensi
- Alternatifnya, sepasang atribut dapat dipertimbangkan

Seleksi dapat juga merupakan pemilihan subset dari objek

- Sebagian wilayah dari layar hanya dapat menunjukkan banyak titik

Teknik Visualisasi : Histogram

Histogram

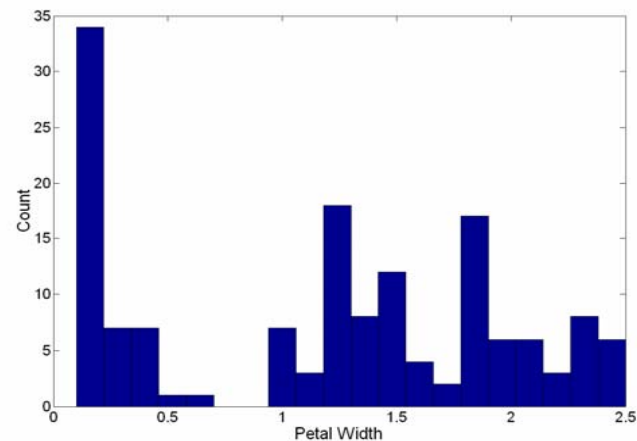
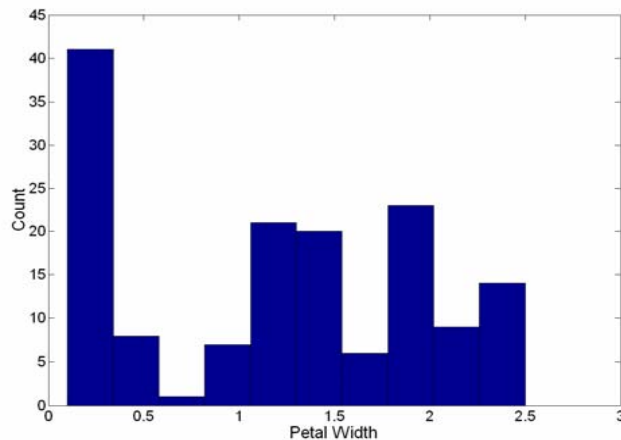
Biasanya menunjukkan distribusi dari nilai dari variabel tunggal

Histogram membagi nilai ke dalam *bin* dan menunjukkan batang plot dari sejumlah objek dalam setiap *bin*

Tinggi dari setiap batang menunjukkan jumlah dari objek

Bentuk histogram tergantung dari jumlah *bin*

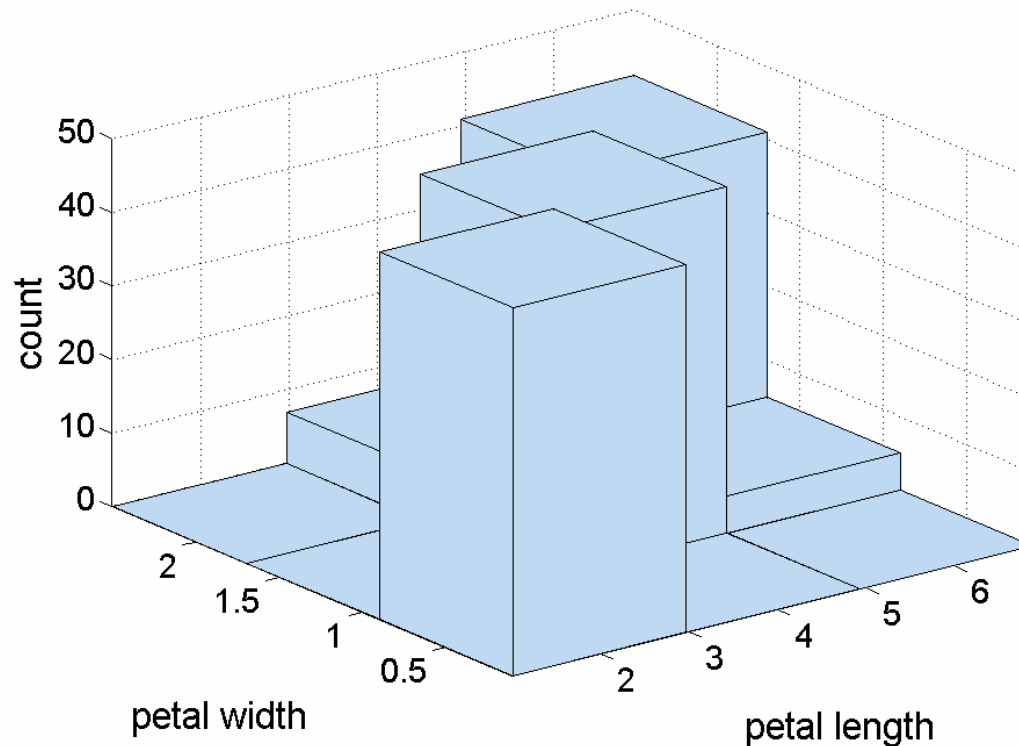
Contoh : lebar mahkota (10 dan 20 *bin*)



Histogram dua dimensi

Histogram jenis ini menunjukkan distribusi gabungan nilai-nilai dari atribut

Contoh : lebar mahkota dan panjang mahkota



Apa yang ingin disampaikan oleh gambar di samping ini ?

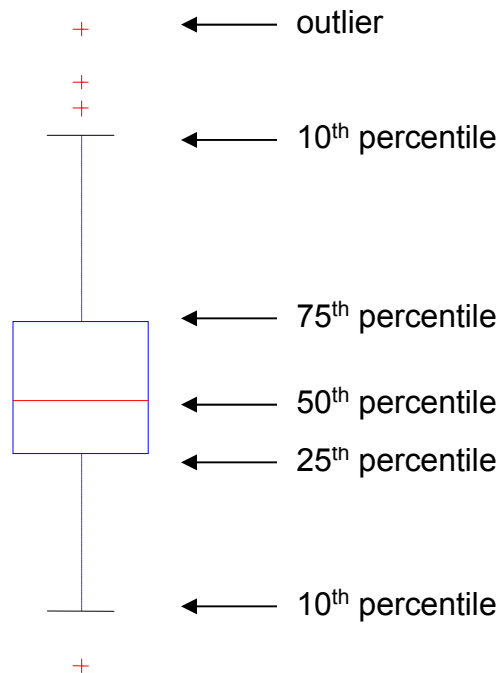
Teknik visualisasi : Box Plots

Box plots

Ditemukan oleh J. Tukey

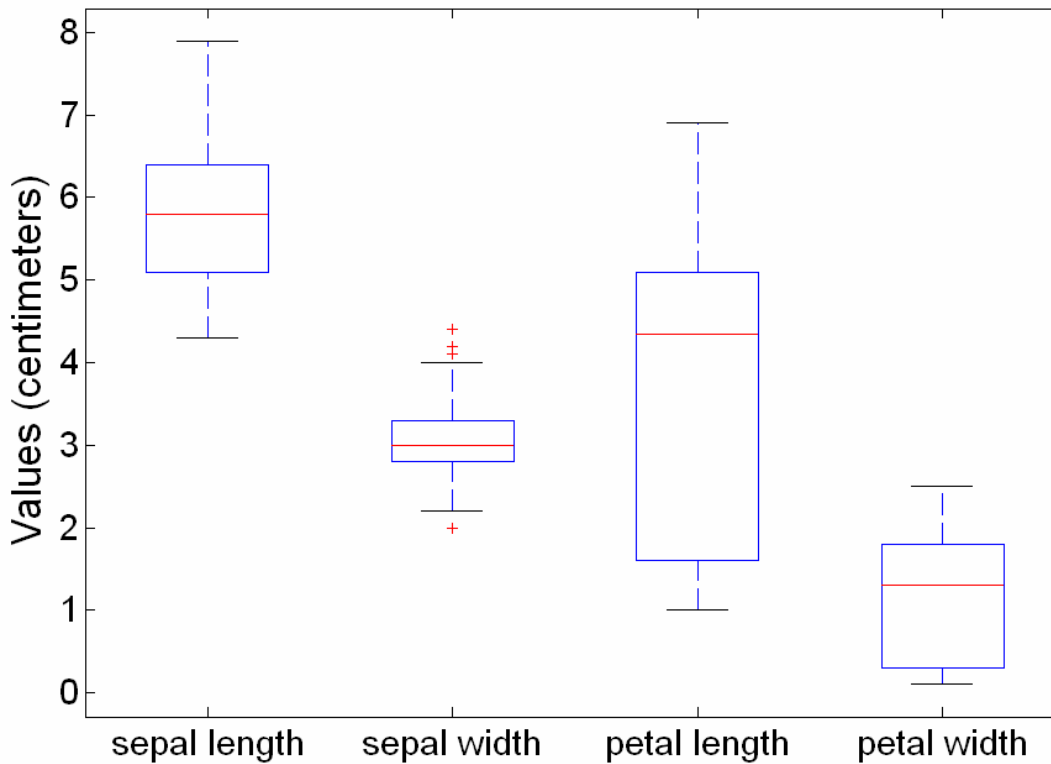
Box plot merupakan cara alternatif untuk menggambarkan distribusi data

Gambar berikut ini menunjukkan bagian dasar dari box plot



Contoh Box Plot

Box plot dapat digunakan untuk membandingkan atribut



Teknik Visualisasi : Scatter Plots

Scatter plot

Scatter plot merupakan nilai-nilai atribut yang menentukan posisi

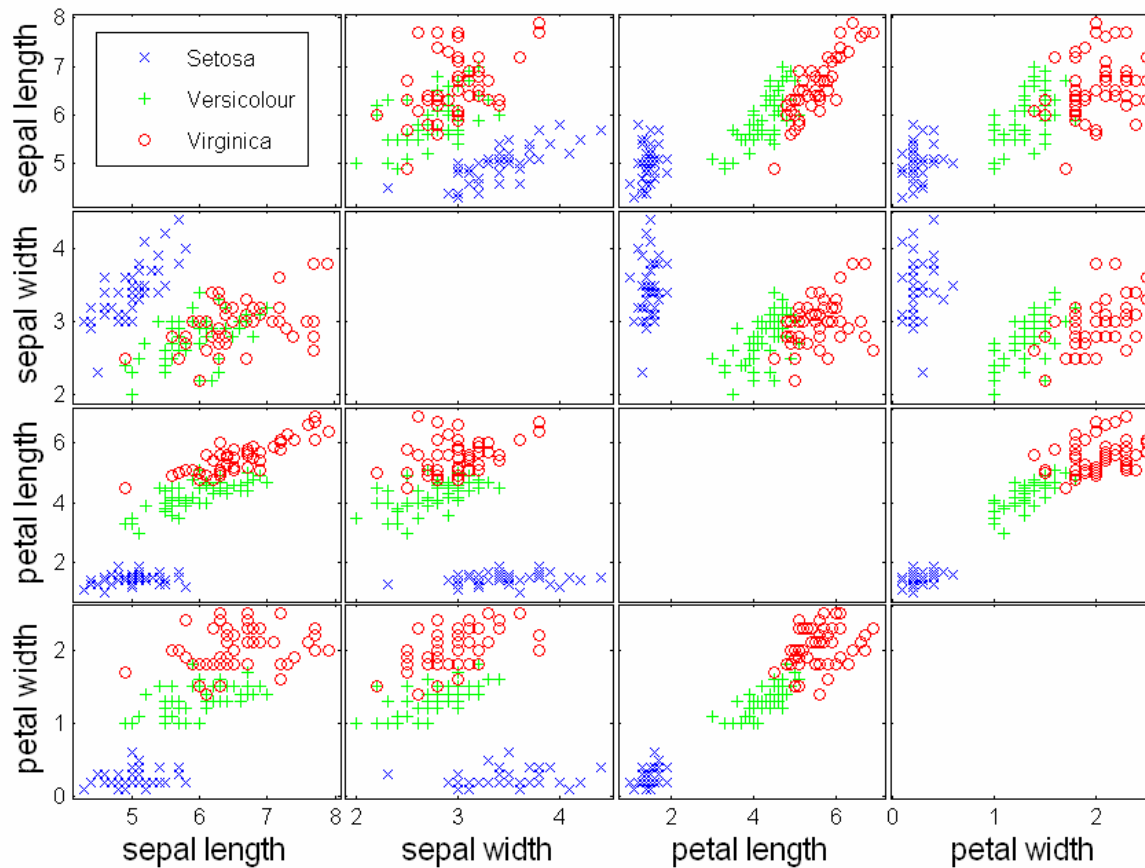
Scatter plot dua dimensi adalah bentuk yang paling umum, tapi dapat juga dibuat dalam scatter plot tiga dimensi

Kadangkala atribut tambahan dapat digambarkan dengan menggunakan ukuran, bentuk dan warna dari tanda yang mewakili objek

Akan sangat berguna jika kita memiliki array dari scatter plot yang secara terpadu merangkum hubungan dari beberapa pasang atribut

Lihat contoh pada slide berikut

Array Scatter Plot dari atribut Iris



Teknik visualisasi : contour plots

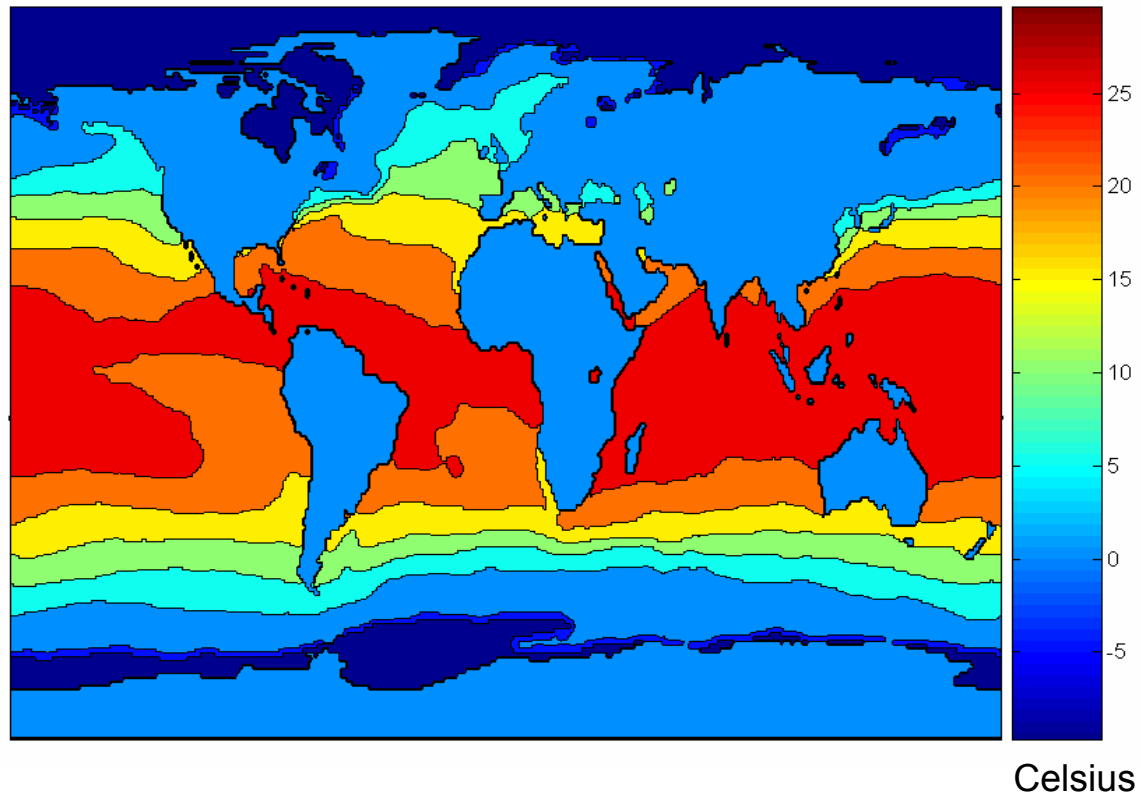


Contour plots

- Contour plots berguna jika atribut kontinu diukur dengan garis
- Contour plot membagi ke dalam daerah-daerah yang nilainya sama
- Garis contour yang membentuk batas dari daerah/wilayah menghubungkan titik-titik dengan nilai yang sama
- Contoh yang paling umum adalah *contour maps of elevation*
- Contour plots juga menggambarkan suhu, curah hujan, tekanan udara dsb.

Sbg contoh : adalah Sea Surface Temperature (SST) pada slide berikut

Contoh Contour Plot : Sea Surface Temperature December 1998



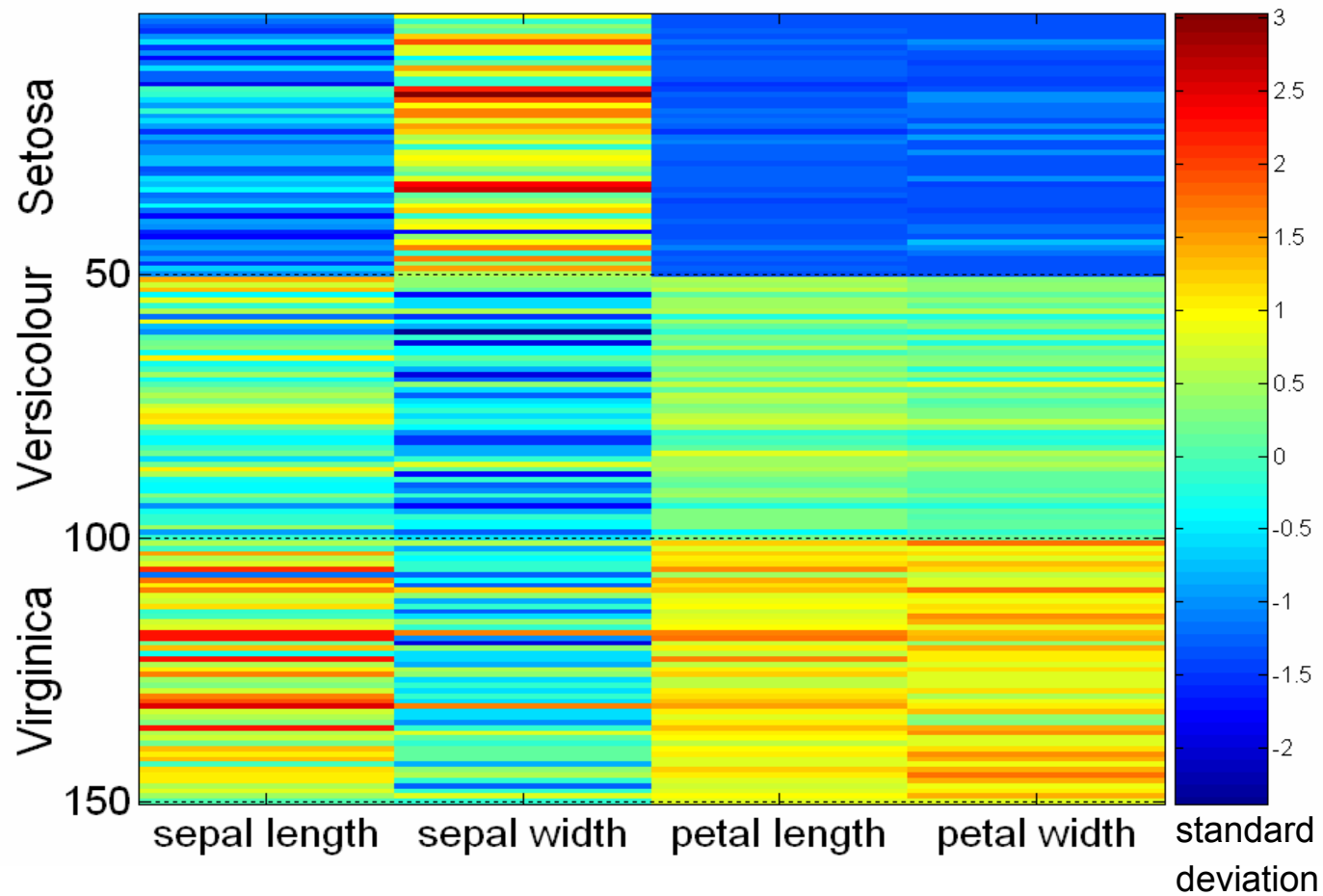
Teknik visualisasi : Matrix Plots



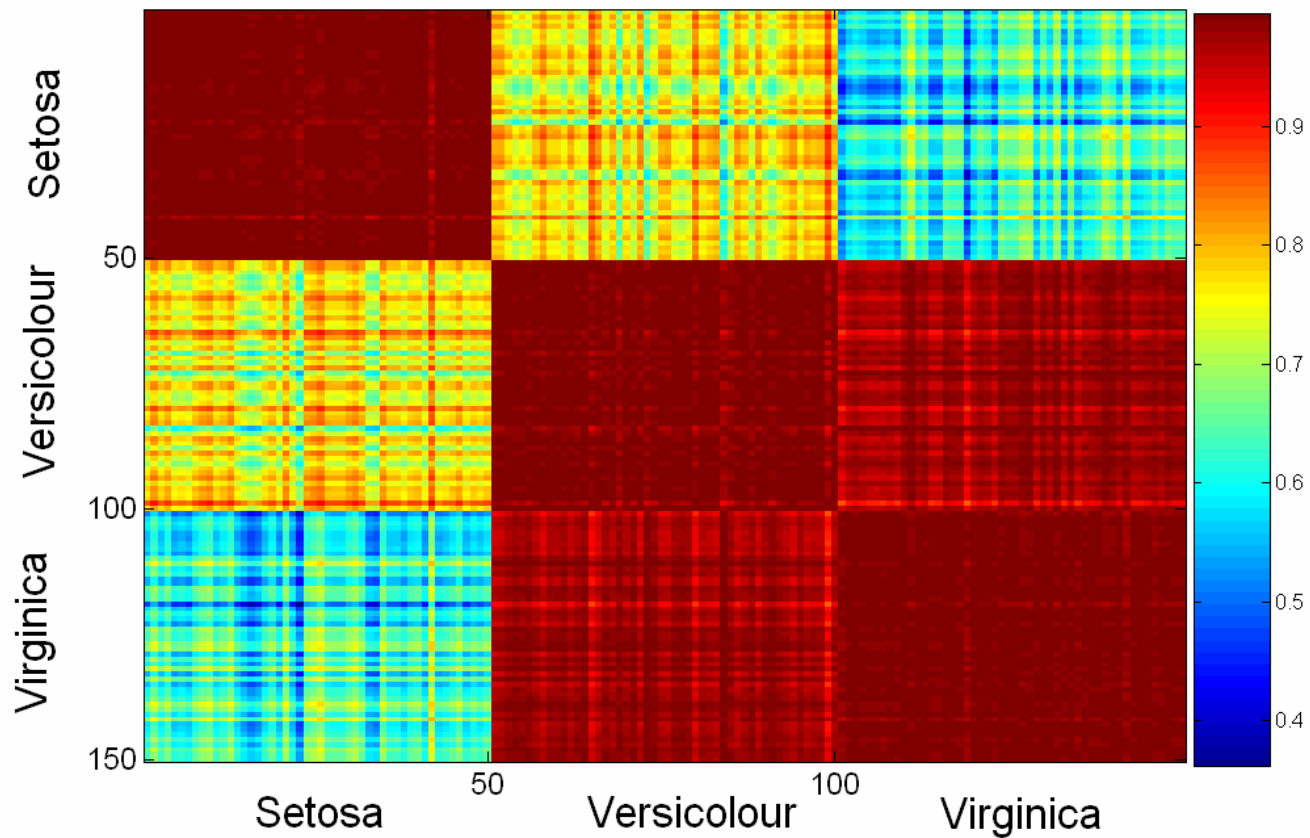
Matrix Plot :

- dapat menempatkan/plot data matriks
- Matriks plot ini dapat berguna pada saat objek diurut menurut kelas
- Biasanya, atribut-atribut dinormalisasikan untuk mencegah satu atribut dari dominasi plot
- Plot dari kesamaan atau jarak matriks dapat juga menjadi berguna untuk menggambarkan hubungan diantara objek-objek
- Contoh matriks plot dapat dilihat pada slide berikut

Visualisasi dari Matriks Data Iris



Visualisasi dari Matriks korelasi Iris



Teknik visualisasi : Koordinat paralel



Koordinat paralel

Berguna untuk menempatkan nilai atribut dari data dimensi tinggi

Daripada menggunakan *perpendicular axes*, maka digunakan *parallel axes*

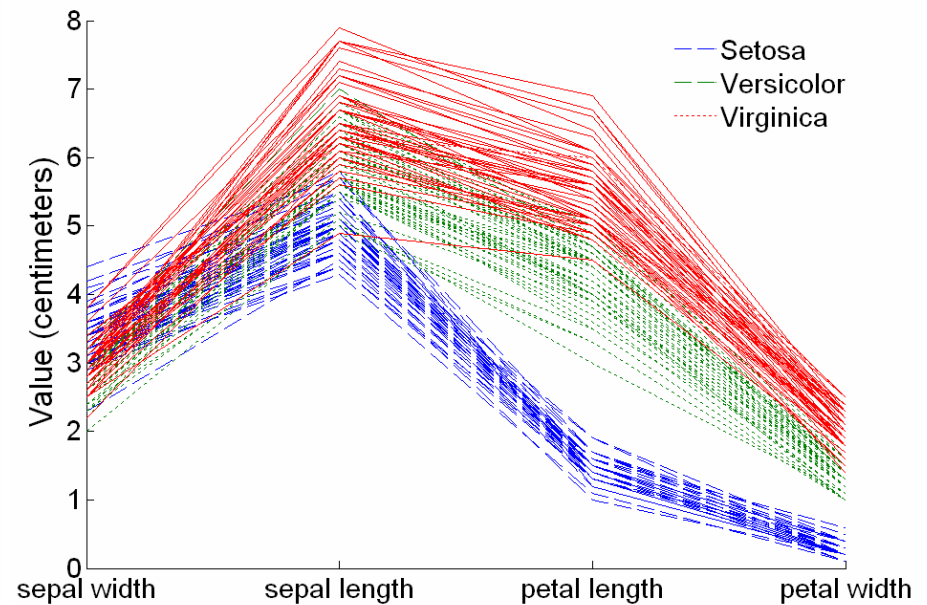
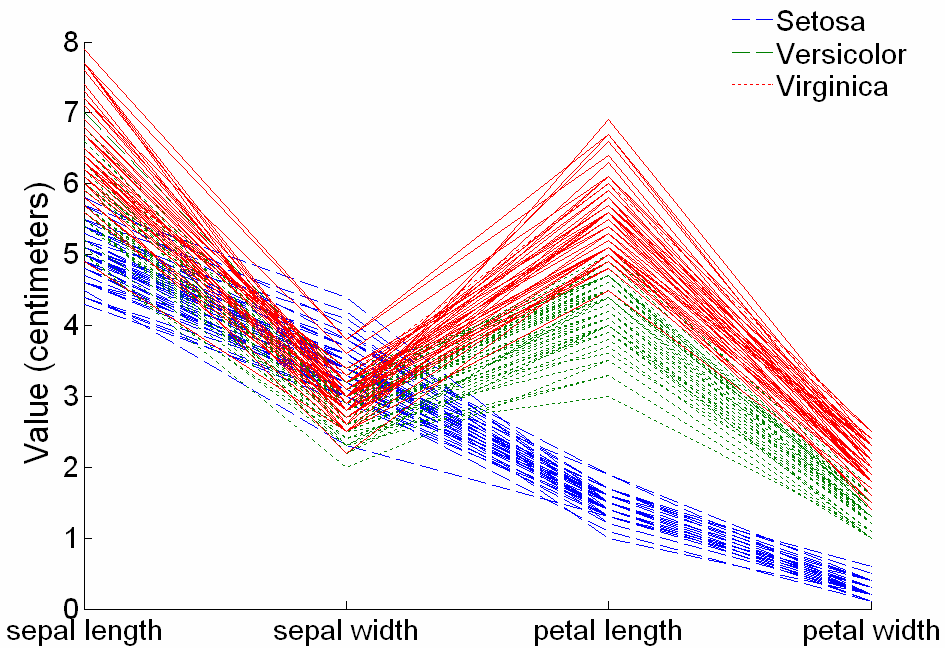
Nilai atribut untuk setiap objek di-plot sebagai titik pada setiap sumbu koordinat yang berhubungan dan titik-titik dihubungkan dengan garis

Kemudian, setiap objek direpresentasikan sebagai garis

Seringkali, garis merepresentasikan kelas yang berbeda dari group objek bersama, paling tidak untuk beberapa atribut

Pengurutan atribut adalah penting mengingat kondisi pengelompokan (group) seperti di atas

Plot Koordinat Paralel untuk data Iris



Teknik Visualisasi lainnya



❖ Star Plots

Teknik visualisasi dengan star plot pendekatannya sama dengan koordinat paralel, namun sumbu menyebar dari titik tengah

Garis yang menghubungkan nilai dari objek disebut dengan poligon

❖ Chernoff Faces

Pendekatan dengan chernoff faces diciptakan oleh Herman chernoff

Pendekatan ini berasosiasi dengan setiap atribut dengan karakteristik dari setiap wajah

Nilai dari setiap atribut menentukan penampakan dari hubungan karakteristik wajah

Setiap objek menjadi wajah yang terpisah

Chernoff faces ini tergantung pada kemampuan manusia dalam membedakan beberapa wajah

Star Plots untuk Data Iris



1



2



3

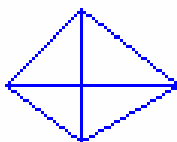


4

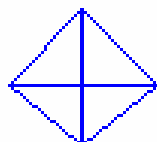


5

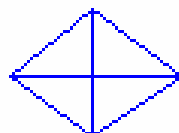
Setosa



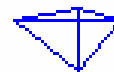
51



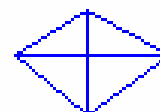
52



53

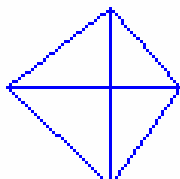


54

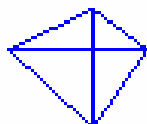


55

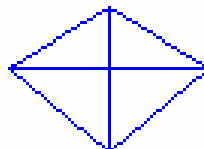
Versicolour



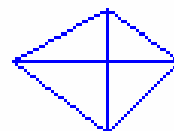
101



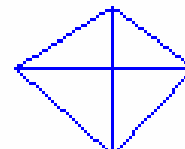
102



103



104



105

Virginica

Chernoff Faces untuk Data Iris



1



2



3



4



5

Setosa



51



52



53



54



55

Versicolour



101



102



103



104



105

Virginica