

Data Visualization

- ***Plot***
Sebuah visualisasi yang menggambarkan/merepresentasikan suatu data, sering disebut sebagai **chart**.
- ***Data wrangling***
Serangkaian proses yang dilakukan untuk membersihkan/merapikan data mentah/awal, sering disebut sebagai **data preparation/data cleansing/data pre-processing**.
- ***Missing value***
Keadaan dimana data memiliki nilai yang hilang (tidak diketahui nilainya).
- ***Feature engineering***
Tahapan untuk menambah jumlah **variable**/kolom berdasarkan informasi dari **variable** lain yang sudah ada.
- ***Aggregation***
Rangkuman informasi suatu data yang sudah dikelompokkan berdasarkan suatu variabel/kolom tertentu.
- ***Base plot***
Plot sederhana yang merupakan fungsi/plot bawaan dari R.
- ***Exploratory***
Langkah awal dalam melakukan analisis data yang bertujuan untuk mengetahui karakteristik suatu data, biasanya dilakukan dengan membuat visualisasi sederhana menggunakan **base plot**.
- ***Explanatory***
Hal yang dilakukan untuk menyampaikan informasi yang terkandung dalam suatu data, biasanya dilakukan dengan membuat visualisasi yang lebih kompleks dan informatif.
- ***Box plot***
Plot yang digunakan untuk menggambarkan persebaran data yang bertipe numerik/angka.
- ***Scatter plot***
Plot yang digunakan untuk menggambarkan persebaran dan hubungan (**correlation**) antara dua variabel bertipe numerik/angka.
- ***Regression line***
Sebuah garis yang menggambarkan pola linier dan hubungan (**correlation**) antar variabel.
- ***Histogram***
Plot yang digunakan untuk menggambarkan persebaran data yang bertipe numerik/angka.
- ***Bins***
Rentang nilai setiap batang pada **histogram**.
- ***Density Plot***
Plot yang digunakan untuk menggambarkan persebaran data yang bertipe numerik/angka.
- ***Line chart***
Plot yang digunakan untuk menggambarkan pola trend suatu data bertipe numerik/angka.

- ***Pie chart***

Plot yang digunakan untuk menggambarkan jumlah/frekuensi data yang bertipe kategorik. **Pie chart** tidak disarankan untuk digunakan dalam **exploratory** maupun **explanatory** karena bentuknya yang menggunakan luas area untuk menggambarkan jumlah/frekuensi data.

- ***Bar plot***

Plot yang digunakan untuk menggambarkan jumlah/frekuensi data yang bertipe kategorik.

- ***ggplot2***

Kumpulan fungsi yang digunakan untuk membuat visualisasi.

- ***Mapping***

Memetakan/menentukan **axis** ataupun elemen lain (warna, ukuran, dll) pada **plot** berdasarkan variabel-variabel yang terdapat pada data.

- ***Axis***

Sumbu pada **plot**, terbagi menjadi sumbu horizontal(x) dan sumbu vertikal(y).

- ***Aesthetic***

“Sesuatu yang dapat ditampilkan” (**axis**, warna, ukuran, bentuk, dll).

- ***Layer***

Pada **ggplot2** dalam membuat **plot** menggunakan konsep lukisan, dimana **plot** dibuat dengan cara **layer-by-layer**.

- ***Geometry***

Jenis **plot** yang ingin digunakan/ditambahkan.

- ***Reshaping***

Mengubah bentuk data yang awalnya melebar menjadi memanjang, atau sebaliknya.

- ***Faceting***

Membuat jenis **plot** yang sama menjadi beberapa bagian berdasarkan variabel tertentu.

- ***Leaflet***

Salah satu **package** yang cukup populer untuk membuat visualisasi berupa peta interaktif.