

Analisis Deskriptif

Tipe data

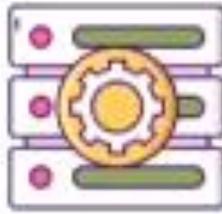
Tendensi pemusatan

Tendensi penyebaran

Outliers



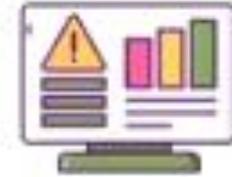
Penyimpanan



Transformasi



Analisis

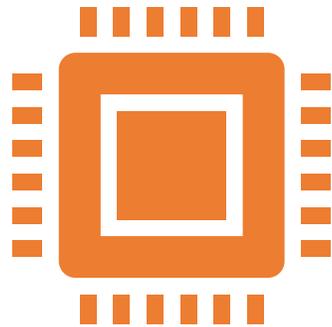


Visualisasi

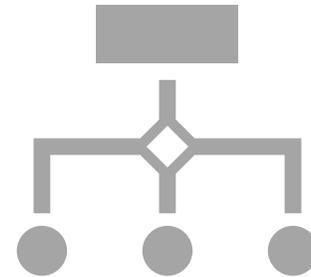
Tipe Data

- Kenapa kita perlu mengenal tipe data?

Tipe Data Kualitatif (Kategorik)



Nominal: data yang tidak dapat diurutkan.
Contoh: Jenis kelamin, Pekerjaan, Ras, dll

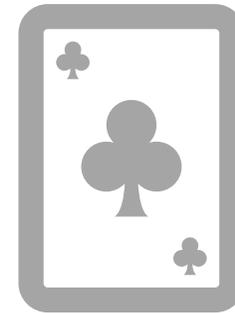


Ordinal: data yang dapat diurutkan. Contoh:
Tingkat kepuasan, ukuran benda

Tipe Data Kuantitatif (Numerik)



Interval: Data yang dapat diurutkan dan diukur perbedaan antara nilainya. Namun perbedaan tersebut tidak dapat diskalakan karena tidak memiliki nilai nol yang absolut. Contoh Nilai ujian, IPK, IQ, dll



Ordinal: data yang dapat diurutkan, diukur, dan dibandingkan nilai perbedaannya. Contoh berat bada, tinggi badan, suhu, dll

Aspek	Kualitatif		Kuantitatif	
	Nominal	Ordinal	Interval	Rasio
Dapat diurutkan	×	✓	✓	✓
Perbedaan nilai	×	×	✓	✓
Perbandingan	×	×	×	✓
Nilai 0 mutlak	×	×	×	✓

Varians

$$\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}$$

Standar Deviasi

$$\sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}}$$

Varians dan standar deviasi

- Merupakan sebuah ukuran penyebaran yang digunakan untuk melihat seberapa jauh jarak antara data dengan nilai tengahnya (rata-rata). Varians menggunakan rata-rata dari kuadrat jarak tersebut, sementara standar deviasi merupakan akar dari varians. Karena menggunakan rata-rata dan kuadrat jarak, kedua ukuran ini sensitive terhadap outliers.

Kuartil dan Persentil

- Kuartil (Quartile) merupakan nilai yang membagi data menjadi 4 bagian. Kuartil ke-1 (Q1) berarti nilai yang membagi 25% pertama, ke-2 (Q2) 25% kedua ($25\%+25\%=50\%$ =sama seperti median) dan kuartil ke-3 (Q3) membagi data menjadi 25% ketiga (75%)
- Persentil (Percentile) merupakan nilai yang membagi data menjadi x% bagian. Misalkan, kita ingin melihat 80% persentil, maka nilai 80% tersebut membagi data menjadi 80% dibawah nilai tersebut dan 20% di atas nilai tersebut

Jangkauan

Jangkauan (range) merupakan jarak antara nilai maksimum dan minimum. Jangkauan digunakan untuk mengestimasi seberapa lebar selang distribusi data yang kita miliki.

Jangkauan antarkuartil (interquartile range) merupakan jarak antara kuartil pertama dengan ketiga. Sementara, simpangan antarkuartil merupakan nilai Interquartile dibagi dua

