

VISUALISASI DATA

Exploratory Data Analytic (EDA)

Magister Teknik Informatika

Dr. Chairani, S.Kom., M.Eng

IIB DARMAJAYA, 2023/2024

Exploratory Data Analytic (EDA)

01

Lorem ipsum dolor sit amet, consectetur adipiscing. Donec risus dolor, porta venenatis neque sit amet, pharetra luctus felis.

02

Mengenal Tipe Data

Lorem ipsum dolor sit amet, consectetur adipiscing. Donec risus dolor, porta venenatis neque sit amet, pharetra luctus felis.

03

Analisis Ukuran Statistik

Lorem ipsum dolor sit amet, consectetur adipiscing. Donec risus dolor, porta venenatis neque sit amet, pharetra luctus felis.

04

Identifikasi & Penanganan Pencilan

Lorem ipsum dolor sit amet, consectetur adipiscing. Donec risus dolor, porta venenatis neque sit amet, pharetra luctus felis.

05

Analisis Distribusi Data

Lorem ipsum dolor sit amet, consectetur adipiscing. Donec risus dolor, porta venenatis neque sit amet, pharetra luctus felis.

06

Analisis Korelasi

Lorem ipsum dolor sit amet, consectetur adipiscing. Donec risus dolor, porta venenatis neque sit amet, pharetra luctus felis.

07


Penanganan Missing Value

Lorem ipsum dolor sit amet, consectetur adipiscing. Donec risus dolor, porta venenatis neque sit amet, pharetra luctus felis.




Apa itu EDA?

Proses eksplorasi dan analisis awal yang dilakukan pada dataset untuk memahami karakteristik, pola, dan hubungan antar variabel sebelum melakukan analisis lebih lanjut atau membangun model.

- 
- ❖ Memahami Data
 - ❖ Mengidentifikasi Pola
 - ❖ Menemukan Anomali
 - ❖ Mengeksplorasi Hubungan
 - ❖ Persiapan Data

Tujuan EDA

- 
- ❖ Statistik Deskriptif
 - ❖ Visualisasi
 - ❖ Analisis Korelasi
 - ❖ Analisis Outlier

Teknik EDA

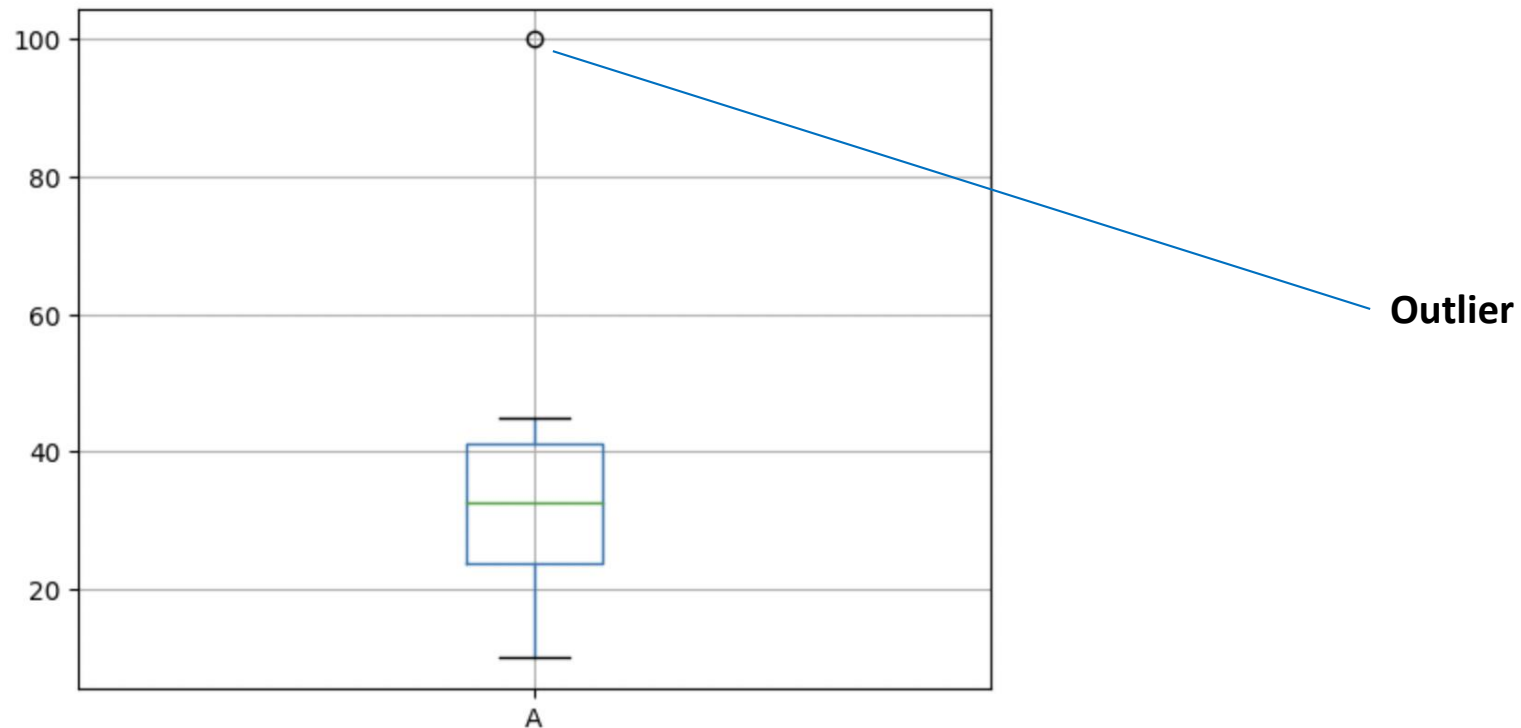
Tipe Data

- ❖ Data kategori/kualitatif
 - Binomial
 - Nominal
 - Ordinal
- ❖ Data Numerik/Kuantitatif
 - Data diskrit
 - Data kontinyu
- ❖ Data Interval
- ❖ Data Rasio

Analisis Ukuran Statistik

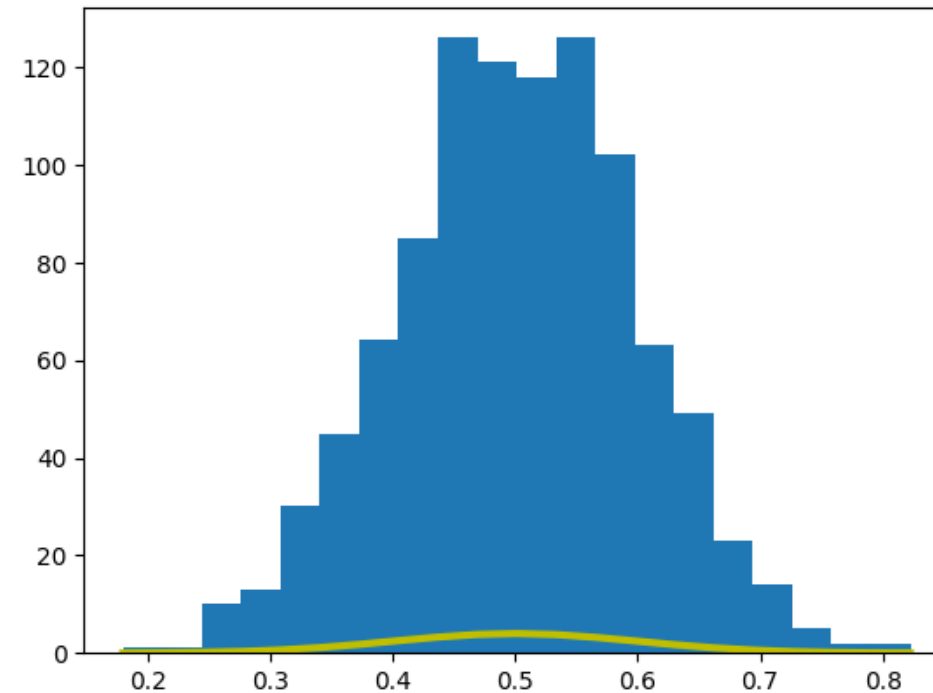
- ❖ Ukuran nilai tengah
 - Mean
 - Median
 - Modus
- ❖ Ukuran Persebaran
 - Deviasi Standar
 - Varian
 - Kuartil

Identifikasi & Penanganan Pencilan



Boxplot menunjukkan outlier

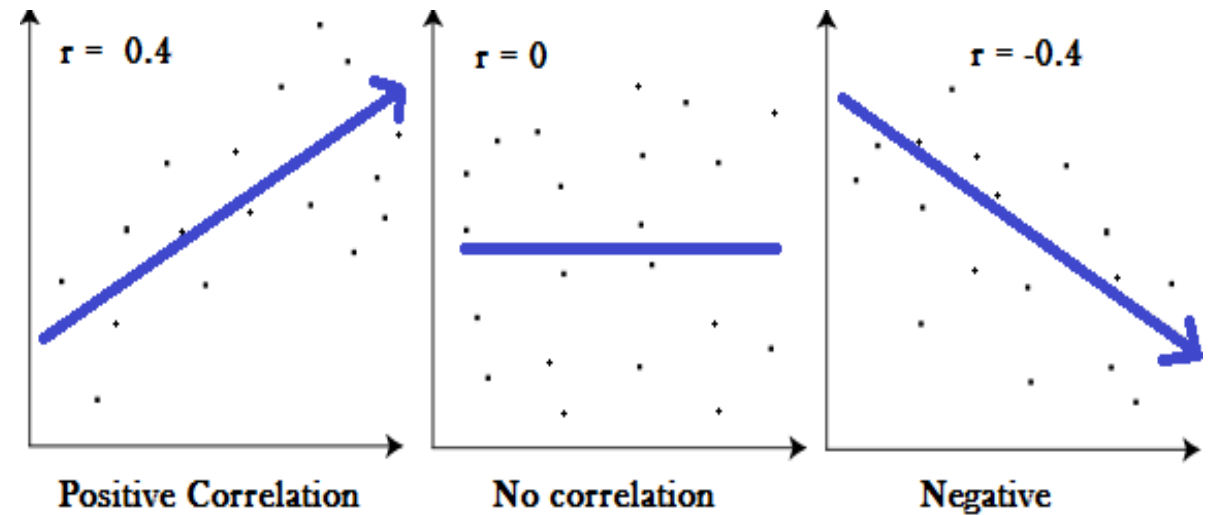
Analisis Distribusi Data



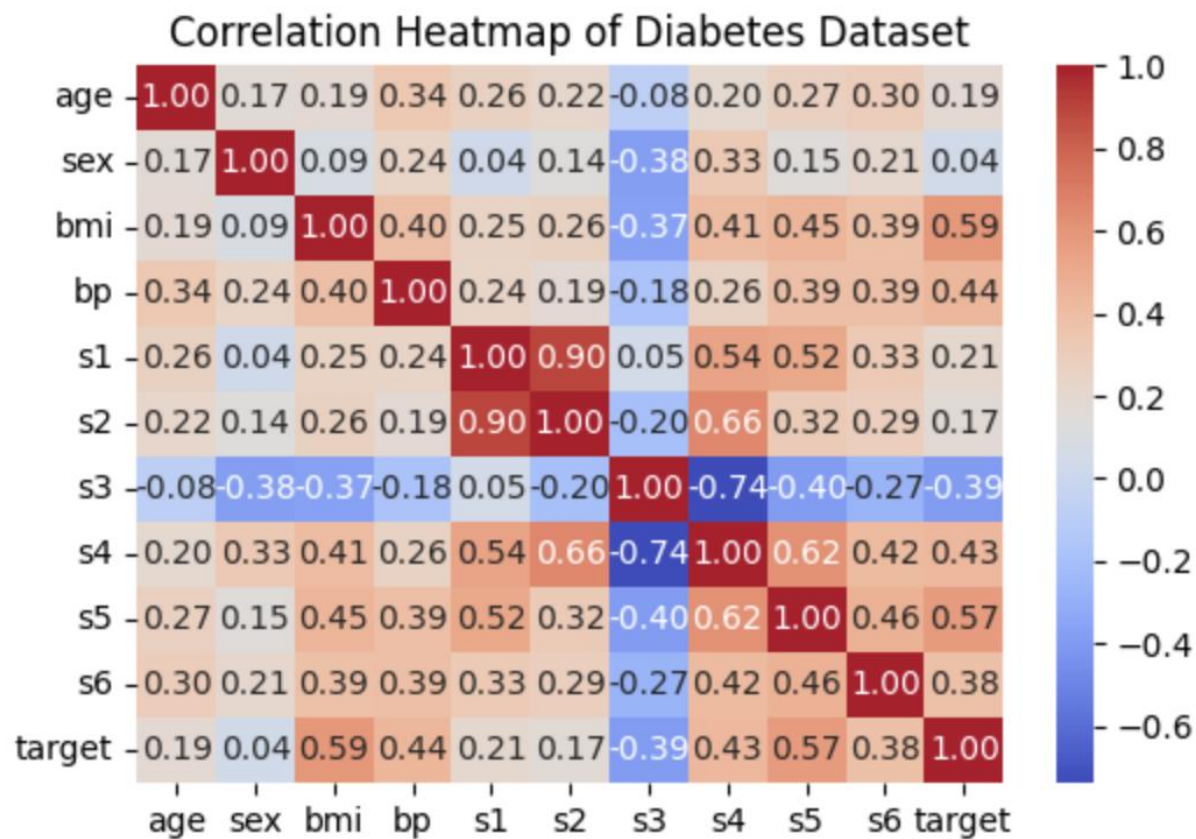
Histogram untuk melakukan pengecekan apakah distribusi data kita normal

Jenis Korelasi

- ❖ Korelasi Positif
Nilai koefisien korelasi > 0
- ❖ Korelasi Negatif
Nilai koefisien korelasi < 0
- ❖ Tidak ada korelasi
Nilai koefisien korelasi $= 0$



Menampilkan Korelasi antar Kolom



Apa itu Missing Values?

Missing value terjadi ketika tidak ada data yang terdeteksi untuk suatu fitur atau variabel pada sampel pengamatan tertentu di dataset. Missing value biasanya muncul sebagai NaN, ?, null atau tidak ada nilainya sama sekali alias blank cell.

Dapat diidentifikasi dengan fungsi Pandas **.isna()**

Penanganan Missing Values

- ❖ Menghapus baris di mana terdapat missing values
Fungsi Pandas: **.dropna()**
- ❖ Mengisi missing value dengan nilai mean, atau nilai 0
Fungsi Pandas: **.fillna()**

Tugas

- ❑ Melakukan EDA terhadap dataset [banking churn](#)

Terimakasih