



Institut Informatika & Bisnis  
**DARMAJAYA**  
Yayasan Alfian Husin



**Kampus  
Merdeka**  
INDONESIA JAYA

**MERDEKA  
BELAJAR**

DATA SCIENCE DARMAJAYA  
“YOUR BEST FUTURE IN DATA”

**PERTEMUAN KE: 4**

# **Statistik Inferensial**

**KULIAH**

OLEH: Nurjoko



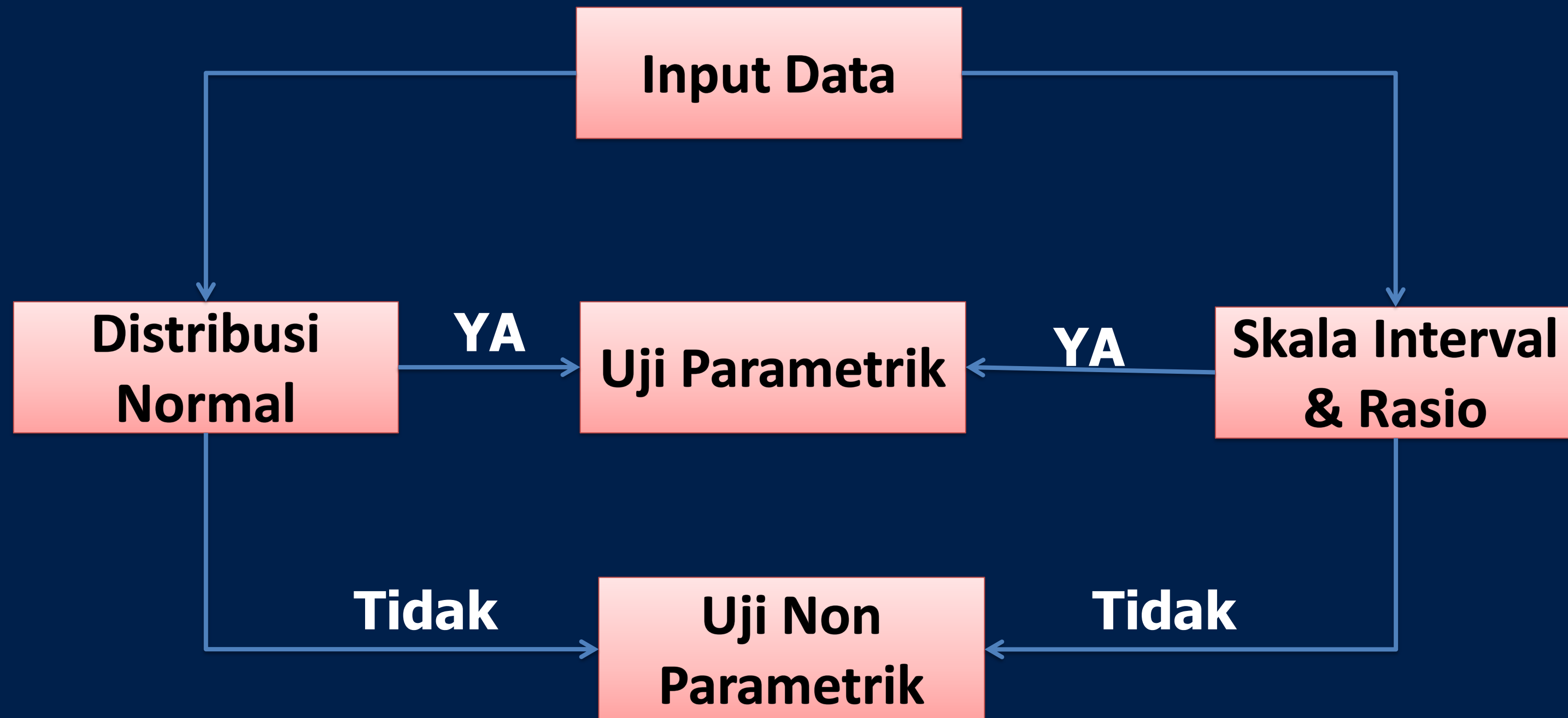
# Learning Objectives

- Memahami konsep statistik inferensial dan penerapannya
- Memahami sifat distribusi normal, termasuk rata-rata, deviasi standar, dan aturan empiris
- Mampu membedakan antara uji statistik parametrik dan non-parametrik
- Memahami perbedaan antara estimasi titik dan estimasi interval
- Memahami langkah-langkah yang terlibat dalam pengujian hipotesis, termasuk pemilihan tingkat signifikansi dan pelaksanaan uji hipotesis.



# Pengenalan Statistik Inferensia

- Yaitu statistika yang digunakan untuk menggeneralisasikan data sampel terhadap populasi.
- Untuk memberi peluang sejauh mana kebenaran data sampel terhadap populasi, maka peneliti menentukan nilai  $\alpha$  (tingkat signifikansi).
- fokus pada penggunaan data sampel untuk membuat inferensi atau kesimpulan tentang populasi yang lebih besar



## JENIS STATISTIK INFERENSIAL

- **Statistik parametrik**

Ilmu statistik yang digunakan untuk data-data yang memiliki sebaran normal dan memiliki skala interval atau rasio.

- **Statistik non parametrik**

Statistik nonparametrik disebut juga statistik bebas sebaran.

Statistik nonparametrik dapat digunakan pada data yang memiliki sebaran normal atau tidak dan memiliki skala nominal atau ordinal.

# Langkah-Langkah Pemilihan Metode Statistik

## 1. Apakah distribusi data diketahui?

- Jika distribusi data tidak diketahui maka statistik yang sesuai adalah statistik nonparametrik.
- Jika distribusi data diketahui, maka kita harus melihat jenis distribusi data tersebut.

## 2. Apakah data berdistribusi normal?

- Jika data tidak berdistribusi normal, maka statistik yang sesuai adalah statistik nonparametrik.
- Jika data berdistribusi normal, maka statistik yang sesuai adalah statistik parametrik.

### 3. Apakah sampel ditarik secara random?

- Jika sampel tidak ditarik secara random, maka statistik yang sesuai adalah statistik nonparametrik.
- Jika sampel ditarik secara random, maka statistik yang sesuai adalah statistik parametrik.

### 4. Apakah varians kelompok sama?

- Jika varians kelompok tidak sama, maka statistik yang sesuai adalah statistik nonparametrik.
- Jika varians kelompok sama, maka statistik yang sesuai adalah statistik parametrik.



## 5. Bagaimana jenis skala pengukuran data?

- Jika skala pengukuran data nominal dan ordinal, maka statistik yang sesuai adalah statistik nonparametrik.
- Jika skala pengukuran data interval dan rasio, maka statistik yang sesuai adalah statistik parametrik.

## Distribusi Probabilitas

Distribusi probabilitas adalah distribusi probabilitas yang digunakan untuk memodelkan atau menggambarkan variasi atau ketidakpastian dalam data dan hasil statistik.

Beberapa distribusi probabilitas dalam statistika inferensial:

- 1. Distribusi Normal (Gaussian):** Distribusi normal adalah salah satu distribusi probabilitas yang paling umum digunakan dalam statistika inferensial. Memiliki bentuk lonceng simetris dengan rata-rata (mean) dan deviasi standar (standard deviation) sebagai parameter utama.



- 2. Distribusi t-Student:** Distribusi t-Student digunakan dalam uji statistik seperti uji t. Ini berguna ketika kita bekerja dengan sampel kecil dan ingin mengestimasi rata-rata populasi atau membandingkan rata-rata dari dua grup.
- 3. Distribusi Chi-Kuadrat:** Distribusi chi-kuadrat umumnya digunakan dalam uji hipotesis yang melibatkan variansi, seperti uji chi-kuadrat untuk uji independensi dalam tabel kontingensi.
- 4. Distribusi F:** Distribusi F digunakan dalam uji statistik seperti analisis varians (ANOVA). Ini membandingkan varians antara beberapa kelompok atau perlakuan.

- 5. Distribusi Binomial:** Distribusi binomial digunakan ketika kita memiliki dua hasil yang mungkin (berhasil atau gagal) dalam serangkaian eksperimen independen. Ini digunakan dalam pengujian hipotesis dan estimasi probabilitas sukses dalam situasi seperti uji A/B.
- 6. Distribusi Poisson:** Distribusi Poisson digunakan untuk menggambarkan kejadian langka dalam interval waktu atau ruang tertentu. Ini sering digunakan dalam analisis data yang melibatkan jumlah kejadian yang jarang.
- 7. Distribusi Bernoulli:** Distribusi Bernoulli adalah kasus khusus dari distribusi binomial dengan satu percobaan. Ini digunakan untuk menggambarkan distribusi probabilitas dari variabel acak biner (dua hasil mungkin).

# STATISTIK PARAMETRIK

- **Statistik parametrik**: yaitu statistik yang digunakan untuk menguji hypotesis yang variabelnya terukur.
  - Contoh: “Berapa menit rata-rata tayangan iklan di TV?”
  - Variabel waktu tayangan iklan dapat terukur dalam menit (ada standar)
- **Statistik parametrik** adalah ilmu statistika yang mempertimbangkan jenis sebaran/distribusi data, yaitu apakah data menyebar normal atau tidak.

Pada umumnya, Jika data tidak menyebar normal, maka data harus dikerjakan dengan metode Statistika non-parametrik, atau setidaknya dilakukan transformasi agar data mengikuti sebaran normal, sehingga bisa dikerjakan dg statistika parametrik.

### Contoh metode statistika parametrik:

- uji-z (1 atau 2 sampel),
- uji-t (1 atau 2 sampel),
- korelasi pearson,
- Perancangan Percobaan (1 or 2-way ANOVA parametrik), dll.

# STATISTIK NON PARAMETRIK

## **Statistik Non parametrik adalah:**

statistik yang digunakan untuk menguji hypotesis yang variabelnya tidak memiliki kepastian (standar)

- **Contoh:** “Berapa besar kepuasan pasien terhadap pelayanan RS. X ?”
- Variabel kepuasan tidak memiliki standar pasti.

**Statistika Non Parametrik adalah** statistika bebas sebaran (tdk mensyaratkan bentuk sebaran parameter populasi, baik normal atau tidak).

Statistika non-parametrik biasanya digunakan untuk melakukan analisis pada data berjenis Nominal atau Ordinal. Data berjenis Nominal dan Ordinal tidak menyebar normal.

**Contoh** metode Statistika non-parametrik: Binomial test, Chi-square test, Median test, Friedman Test, dll.

Singkat kata **perbedaannya** kalau datanya memiliki sebaran atau **distribusi normal**, maka digunakan **statistika parametrik**. Kalo data **tidak memiliki sebaran normal**, maka digunakan **statistika nonparametrik**.



## Pengujian Hipotesis

Dalam statistik inferensial, pengujian hipotesis merupakan salah satu tujuan yang harus dilakukan untuk menjawab masalah penelitian secara rasional.

### **Tujuan pengujian hipotesis:**

untuk menentukan apakah jawaban teoritis yang terkandung dalam pernyataan hipotesis didukung oleh fakta yang dikumpulkan selama proses penelitian.

- Pengujian hipotesis (yang menggambarkan karakteristik populasi) dengan menggunakan data sampel (yang menggambarkan karakteristik sampel) pada dasarnya merupakan pembuatan keputusan melalui proses inferensi yang memerlukan akurasi peneliti dalam melakukan estimasi.
- Proses inferensi pada dasarnya dapat dilakukan melalui satu dari dua cara, yaitu:
  - estimasi nilai parameter populasi atau
  - membuat keputusan mengenai nilai parameter (proses pengujian hipotesis).



- Estimasi nilai parameter populasi akurasiya tergantung pada representasi sampel yang diambil dari populasi yang bersangkutan.
- Peneliti harus mempunyai kriteria atau standar yang digunakan untuk membuat keputusan terhadap hipotesis yang diuji berdasarkan sampel.
- Kriteria keputusan yang ditetapkan oleh peneliti dalam istilah statistik disebut tingkat signifikansi (*significance level*).

- **Tingkat signifikansi** adalah tingkat probabilitas yang ditentukan oleh peneliti untuk membuat keputusan menolak atau mendukung hipotesis.
- Kriteria keputusan berdasarkan tingkat signifikansi, misalnya **0,05** atau **0,01** menunjukkan bahwa keputusan yang dibuat oleh peneliti untuk menolak atau mendukung suatu hipotesis mempunyai probabilitas kesalahan sebesar lima persen atau sepuluh persen.



Dalam melakukan uji hipotesis, ada banyak faktor yang menentukan seperti:

- apakah sampel yang diambil berjumlah banyak atau hanya sedikit,
- apakah std deviasi populasi diketahui,
- apakah variansi dari populasi diketahui,
- apa metode parametrik yang digunakan, dst.

# 1. Prosedur Uji Hipotesis

a. Menentukan  $H_0$  dan  $H_1$

$H_0$  adalah NULL HYPOTHESIS

$H_1$  adalah ALTERNATIVE HYPHOTHESIS

b. Menentukan Uji (Prosedur) Statistik yang digunakan; apakah akan digunakan uji t, ANOVA, uji z, dan lain sebagainya.

c. Menentukan statistik tabel

1. Tingkat kepercayaan

2. Derajat kebebasan (df)

Derajat kebebasan atau *degree of freedom* sangat bervariasi tergantung dari metode yang dipakai dan jumlah sampel yang diperoleh.

3. Jumlah sampel yang didapat

#### d. Menentukan Statistik hitung

Nilai ini tergantung pada metode parametrik yang digunakan. Pada pengerjaan pada SPSS, nilai statistik hitung langsung ditampilkan nilai akhirnya; sedangkan proses perhitungannya sampai pada nilai akhir tersebut tidak diperlihatkan, termasuk angka statistik tabel.

#### e. Mengambil keputusan

Hal ini ditentukan dengan membandingkan nilai statistik hitung dengan nilai statistik tabel atau nilai kritisnya.

## 2. Berbagai Metode Parametrik

### a. Inferensi terhadap sebuah rata-rata populasi

Tujuan pengujian ini adalah ingin mengetahui apakah sebuah sampel berasal dari sebuah populasi yang mempunyai rata-rata (mean) yang sudah diketahui. Atau, bisa juga dikatakan ingin menguji apakah rata-rata sebuah sampel sudah bisa mewakili populasinya.

Pada inferensi ini, perlu diperhatikan ukuran sampel, apakah termasuk sampel besar ataukah sampel kecil.

- **Sampel Besar**

Dalam kasus dimana jumlah sampel yang diambil cukup besar atau varians populasi diketahui, maka bisa dipakai rumus (uji) Z.

- **Sampel Kecil**

Jika sampel kecil ( $<30$ ) dan varians populasi tidak diketahui, metode parametrik yang digunakan adalah uji t (student).

b.

## b. Inferensi terhadap dua rata-rata populasi

Dalam inferensi dua populasi, tujuan utama adalah ingin mengetahui apakah ada perbedaan antara dua rata-rata populasi.

- Sampel Besar

Metode yang digunakan adalah z test yang dimodifikasi.

- Sampel Kecil

Metode yang digunakan adalah:

-> t test yang dimodifikasi : sampel bisa saling berhubungan (dependen) maupun kedua sampel tidak ada hubungannya (independen).

-> F test

c. Inferensi terhadap lebih dari dua rata-rata populasi

Untuk lebih dari dua populasi, misal tiga jenis sampel, empat jenis sampel dan seterusnya, dipakai analisis ANOVA, yang terdiri atas:

- ANOVA satu faktor
- ANOVA dua faktor dengan *replacement*
- ANOVA dua faktor tanpa *replacement*

#### d. Inferensi untuk mengetahui apakah ada hubungan antara variabel

Inferensi ini bertujuan untuk mengetahui apakah ada hubungan yang signifikan antara dua variabel. Beberapa alat statistik untuk mengetahui hubungan antar variabel:

- Hubungan antar dua variabel, menggunakan korelasi dan regresi sederhana.
- Hubungan antara lebih dari dua variabel (tiga, empat, dan seterusnya), menggunakan metode korelasi dan regresi berganda.

## Contoh kasus penerapan statistika inferensial

**1. Uji Klinis Obat Baru:** Dalam pengembangan obat baru, uji klinis sering dilakukan untuk menguji efektivitas obat tersebut. Statistik inferensial digunakan untuk menentukan apakah obat tersebut lebih efektif daripada plasebo, dengan membandingkan hasil dari kelompok yang menerima obat dengan kelompok kontrol yang menerima plasebo.

**2. Pengendalian Kualitas Produksi:** Dalam industri manufaktur, statistika inferensial digunakan untuk menguji kualitas produk. Contohnya, pabrik mungkin mengambil sampel produk secara berkala dan menggunakan uji statistik untuk memastikan bahwa produk-produk tersebut memenuhi standar kualitas.

**3. Survei Opini Publik:** Dalam ilmu sosial, survei sering digunakan untuk mengumpulkan data opini publik. Statistik inferensial digunakan untuk mengambil kesimpulan tentang pendapat seluruh populasi berdasarkan sampel survei.



## Studi Mengenai Dukungan Pemilihan Umum:

Sebuah lembaga penelitian ingin mengukur tingkat dukungan masyarakat terhadap dua calon presiden dalam pemilihan umum. Mereka merancang survei opini publik dengan pertanyaan yang memungkinkan responden memilih satu dari dua calon presiden.

Langkah-langkah penerapan statistika inferensial dalam survei ini adalah sebagai berikut:

- 1. Pengambilan Sampel:** Lembaga penelitian akan mengambil sampel acak dari populasi yang mewakili pemilih yang akan berpartisipasi dalam pemilihan umum. Ini mungkin melibatkan rancangan sampel acak sederhana.
- 2. Desain Pertanyaan:** Pertanyaan survei dirancang dengan cermat untuk memungkinkan responden memilih calon presiden yang mereka dukung.
- 3. Pelaksanaan Survei:** Survei dilakukan dengan menghubungi responden yang terpilih secara acak dan mengumpulkan jawaban mereka.
- 4. Analisis Data:** Setelah mengumpulkan data, lembaga penelitian menggunakan statistika inferensial untuk mengambil kesimpulan tentang dukungan masyarakat. Mereka mungkin menghitung persentase pemilih yang mendukung masing-masing calon presiden.
- 5. Perhitungan Margin of Error:** Untuk memberikan gambaran tentang sejauh mana hasil survei mencerminkan seluruh populasi, lembaga penelitian dapat menghitung margin of error berdasarkan ukuran sampel dan level kepercayaan yang dipilih.
- 6. Presentasi Hasil:** Hasil survei, beserta margin of error, disajikan kepada publik dan pemangku kepentingan untuk memberikan informasi tentang tingkat dukungan masyarakat terhadap calon presiden.

Dalam contoh ini, statistika inferensial membantu dalam mengambil kesimpulan tentang tingkat dukungan masyarakat terhadap kedua calon presiden berdasarkan sampel survei. Hasil ini dapat digunakan oleh pemangku kepentingan, termasuk kampanye pemilu, untuk merencanakan strategi dan mengambil keputusan yang didasarkan pada data yang relevan.



# Learning Objective n

Fill in .....



# CONCLUSION

Fill in .....



# REFERENCES

Fill in IEEE Style



Institut Informatika & Bisnis  
**DARMAJAYA**  
Yayasan Alfian Husin



**Kampus  
Merdeka**  
INDONESIA JAYA

**MERDEKA  
BELAJAR**

# THANK YOU!!

DATA SCIENCE DARMAJAYA "YOUR BEST FUTURE IN DATA"