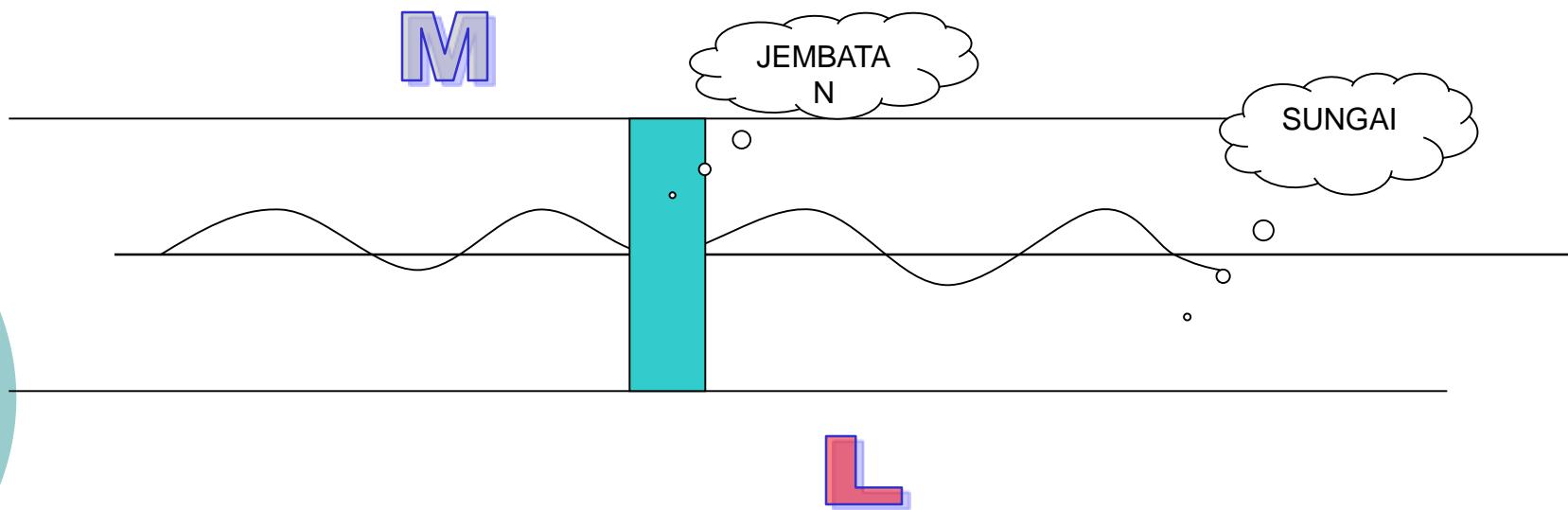




# Prioritas Anda Dalam Hidup

---

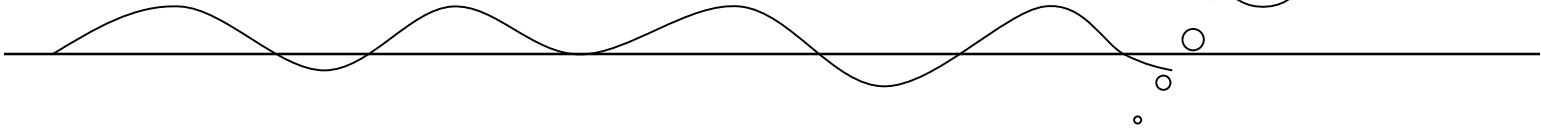
Sebuah test psikologi



Dua orang kekasih, **L** (perempuan) dan **M** (laki-laki) tinggal saling berseberangan. Mereka dipisahkan oleh sebuah sungai, yang di mana setiap kali ingin bertemu, mereka menggunakan satu-satunya jembatan penyeberangan yang ada di sana.



**M**



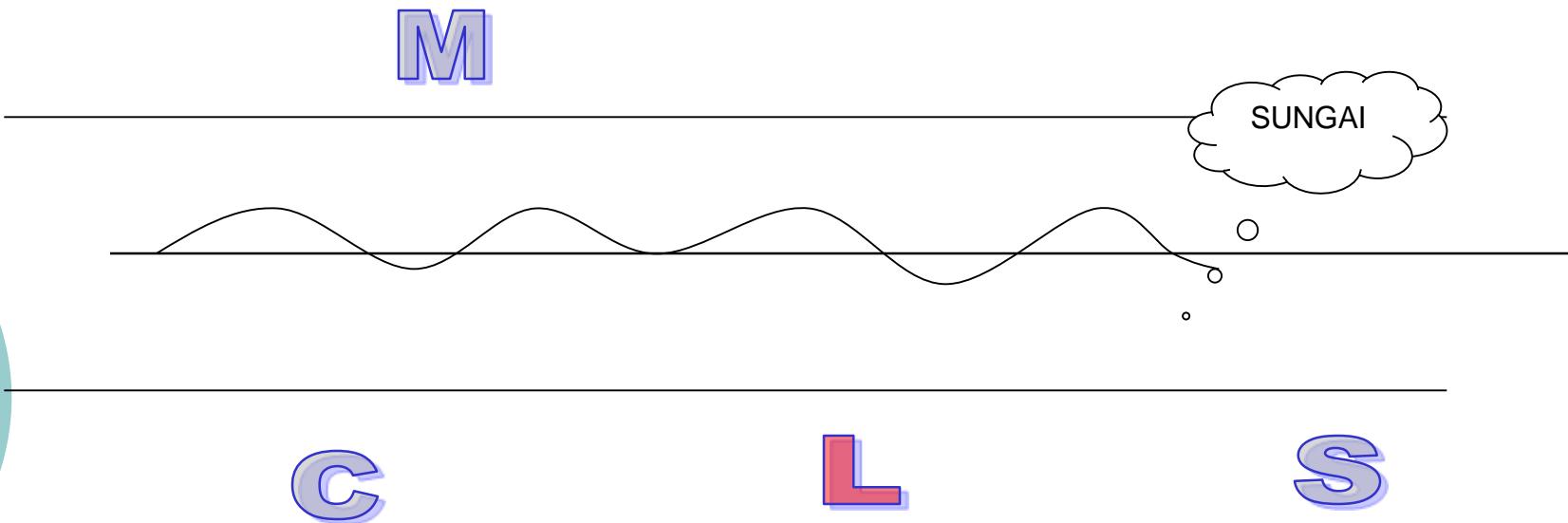
SUNGA  
I



**L**

Suatu hari, jembatan tersebut rusak dan hancur, sehingga kedua orang tersebut tidak dapat lagi saling bertemu.

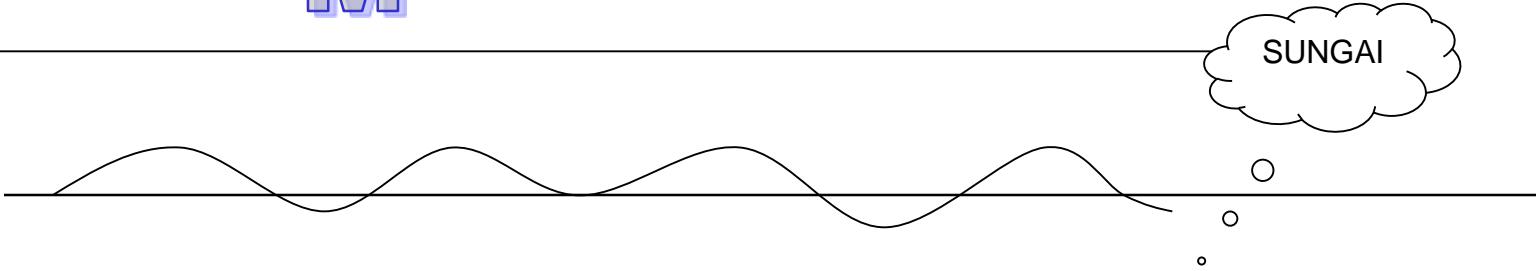
Waktu terus berjalan dengan kondisi seperti itu, sampai suatu saat **L** sangat rindu dan ingin bertemu dengan **M** ...



Dengan rusaknya jembatan, terbuka bisnis baru untuk penyeberangan, yaitu dengan menggunakan perahu penyeberangan. Ada 2 orang laki-laki yang menjalankan bisnis penyeberangan tersebut, masing-masing bernama **C** dan **S**.



M



SUNGAI

C L

S

L lalu mendatangi C

L: "Bisa Anda tolong seberangkan saya?"

C: "Bisa, asal Anda membayar saya sebesar 5 keping uang emas"

L: "Saya tidak punya uang sebanyak itu."

C: "Maaf, kalau begitu saya tidak bisa membantu"

L: "Tolonglah Pak, saya sangat butuh diseberangkan..."

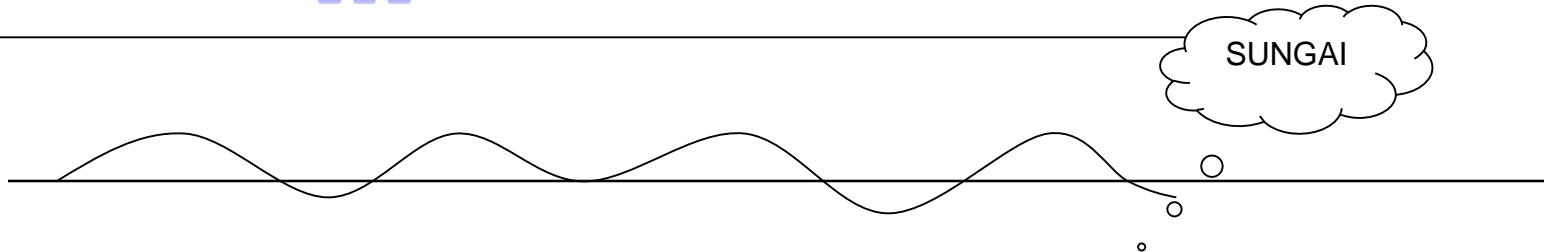
C: "Maaf, Nona. Saya hanya menjalankan bisnis saja.

Dan saya juga punya kebutuhan untuk menghidupi keluarga saya."

Dengan sedih kemudian L meninggalkan C ....



**M**



**C**

**L S**

**L** lalu mendatangi **S**

**L:** "Bisa Anda tolong seberangkan saya?"

**S:** "Bisa, asal Anda membayar saya sebesar 5 keping uang emas"

**L:** "Saya tidak punya uang sebanyak itu."

**S:** "Maaf, kalau begitu saya tidak bisa membantu"

**L:** "Tolonglah Pak, saya sangat butuh diseberangkan..."

**S:** "Hm... Mungkin saya bisa membantu .... Asalkan Nona mau menemani saya malam ini?"

**L:** (termenung agak lama) "Apakah tidak ada pilihan lain?"

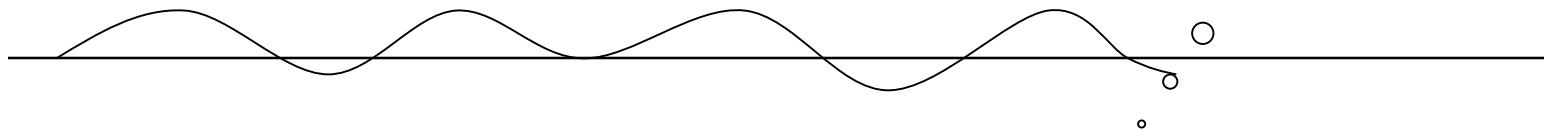
**S:** Maaf, Nona bayar saya dengan uang, atau dengan bermalam dengan saya, hanya itu pilihan Nona."

Karena rindu yang teramat dalam terhadap **M**, maka kemudian dengan berat hati **L** menyetujui persyaratan **S** untuk bermalam bersama.



**M** **L**

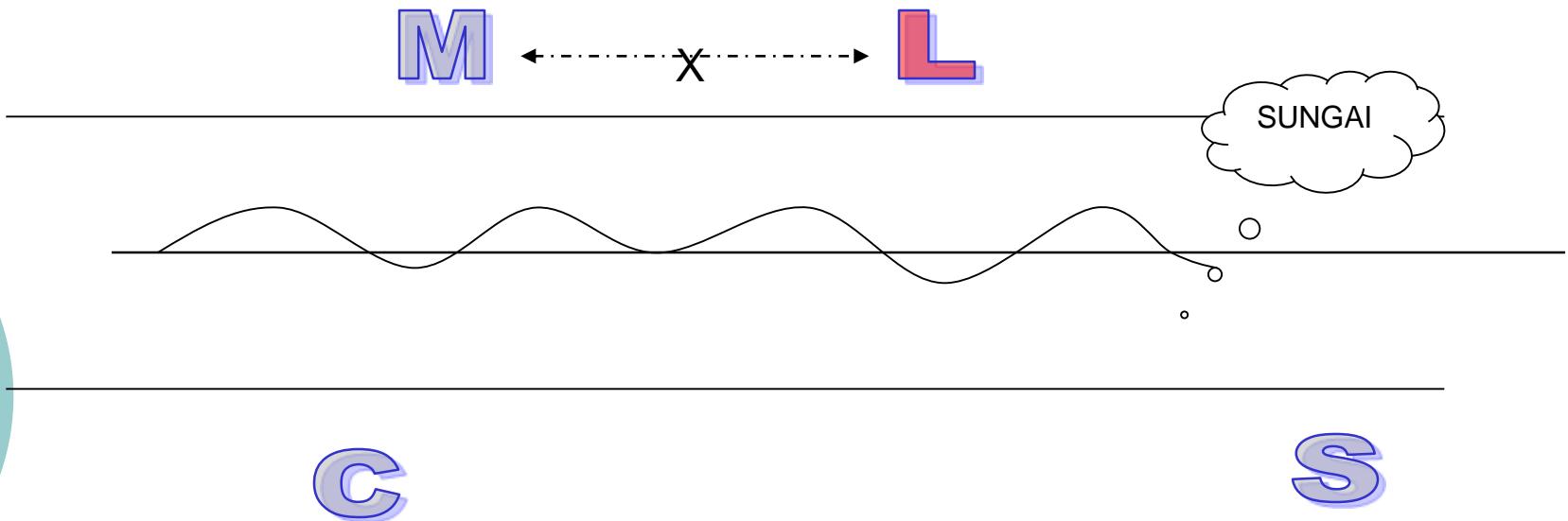
---



**C**

**S**

Akhirnya **L** bisa bertemu dengan **M**. Mereka berdua sangat bahagia, dan saling melepas rindu bersama. Namun setelah berhari-hari kemudian **M** akhirnya tidak tahan untuk menanyakan bagaimana cara **L** bisa menyeberangi sungai sementara jembatan masih belum diperbaiki dan harga sewa perahu sangatlah mahal. **L** menceritakan semuanya tanpa ada yang disembunyikan, bahwa akhirnya ia setuju untuk bermalam dengan **S** hanya demi usahanya untuk bertemu dengan **M** terlaksana, karena dia tidak punya uang sebanyak itu untuk menyeberang.

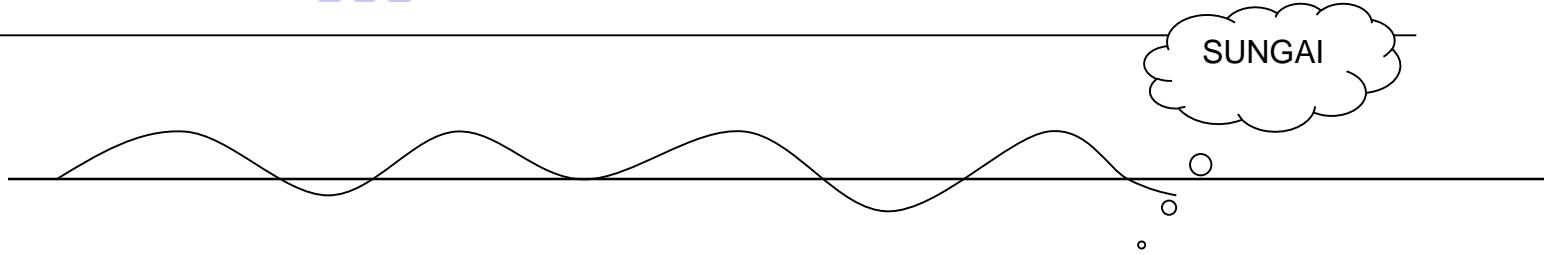


Mendengar pengakuan **L** yang apa adanya dan jujur tersebut, **M** sangat marah dan tersinggung. **M** tidak dapat menerima penjelasan dari **L** bahwa semua yang dilakukan **L** itu adalah semata-mata karena cintanya terhadap **M**.

**M** menganggap penjelasan **L** hanya alasan semata dan tidak dapat ditolerir, sehingga secara sepihak lalu **M** memutuskan **L** dan menyuruh **L** untuk melupakan hubungan mereka dan pergi meninggalkan **M**.



**M**



**C**

**LH**

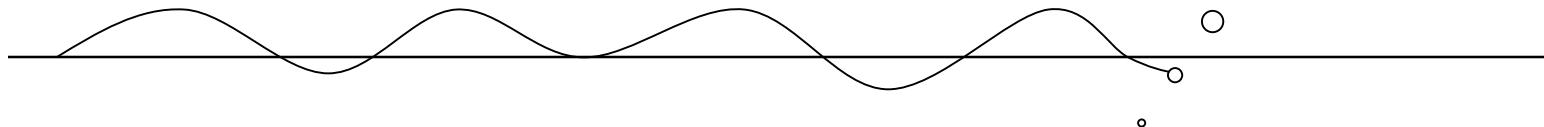
**S**

Dengan hati hancur, **L** pergi meninggalkan **M**. Tidak diceritakan bagaimana caranya, tapi **L** berhasil menyeberangi sungai dan kembali. Sekembalinya **L** menemui **H**, Seorang laki-laki yang menjadi sahabatnya sejak lama. **H** juga mengenal **M** dengan baik, sehingga **L** dapat mencerahkan seluruh isi hati dan kesedihannya kepada **H**. **H** dengan bijaksana kemudian menenangkan **L**, dan meyakinkan **L** bahwa apapun yang **L** buat sudah maksimal, dan bahwa **L** tidak dapat selalu memaksakan apa yang diinginkan menjadi kenyataan.

Hari demi hari berlalu, hubungan **L** dan **H** semakin akrab, dan akhirnya mereka Justru menjadi sepasang kekasih yang kemudian dalam selang waktu tidak lama keduanya melangsungkan pernikahan.



M



C

LH

S

### Tugas anda :

Dari ke-5 karakter diatas (M,C,H,L dan S), urutkanlah prioritas Anda terhadap karakter yang paling Anda sukai sampai pada karakter yang paling Anda tidak suka.

1. ... **(paling disukai)**

2. ...

3. ...

4. ...

5. ...**(paling tidak disukai)**

Tulislah hasilnya di kertas. Anda ingin tahu penjelasannya ....???

---

Anda benar-benar sudah mengurutkan karakter-karakter tersebut sesuai permintaan?

Berikut ini adalah interpretasi dari tiap-tiap karakter yang diwakilkan:

---

**M** Moral

**C** Career / Money

**L** Love

**H** Home

**S** Sex

Setelah Anda mengetahui arti masing-masing dari karakter tersebut, maka Anda bisa mengetahui prioritas Anda dalam hidup berdasarkan jawaban yang telah Anda susun sebelumnya ....



# **Ini sekedar permainan ...**

---

**Sesungguhnya ...  
ANDA ADALAH APA YANG ANDA  
PIKIR BAHWA ANDA ADALAH ANDA...  
You are what you think you are....**

# DATA SCIENCE



AI

# Data Science

## Pengertian

Bidang yang **mempelajari data** dan cara mengekstraksi **makna** darinya menggunakan serangkaian **metode, algoritma, system,** dan alat untuk mengekstrak wawasan dari data

## Objective

Pengetahuan itu kemudian diterapkan pada bisnis, pemerintah, dan badan lain untuk **membantu mendorong keuntungan, berinovasi produk dan layanan, membangun infrastruktur, dll**

# Data Science Applications



# Implementation of Data Science

Speech  
Recognition

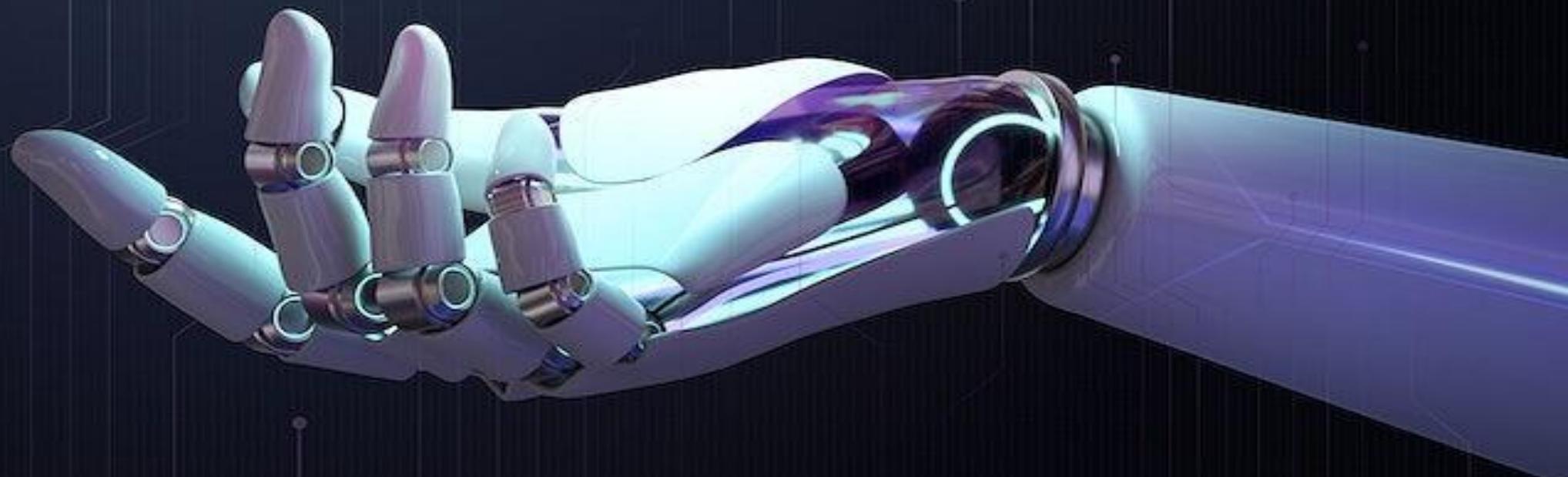
Recommender  
System

Hoax  
Detection

Fraud & Risk  
Prediction

Image  
Recognition

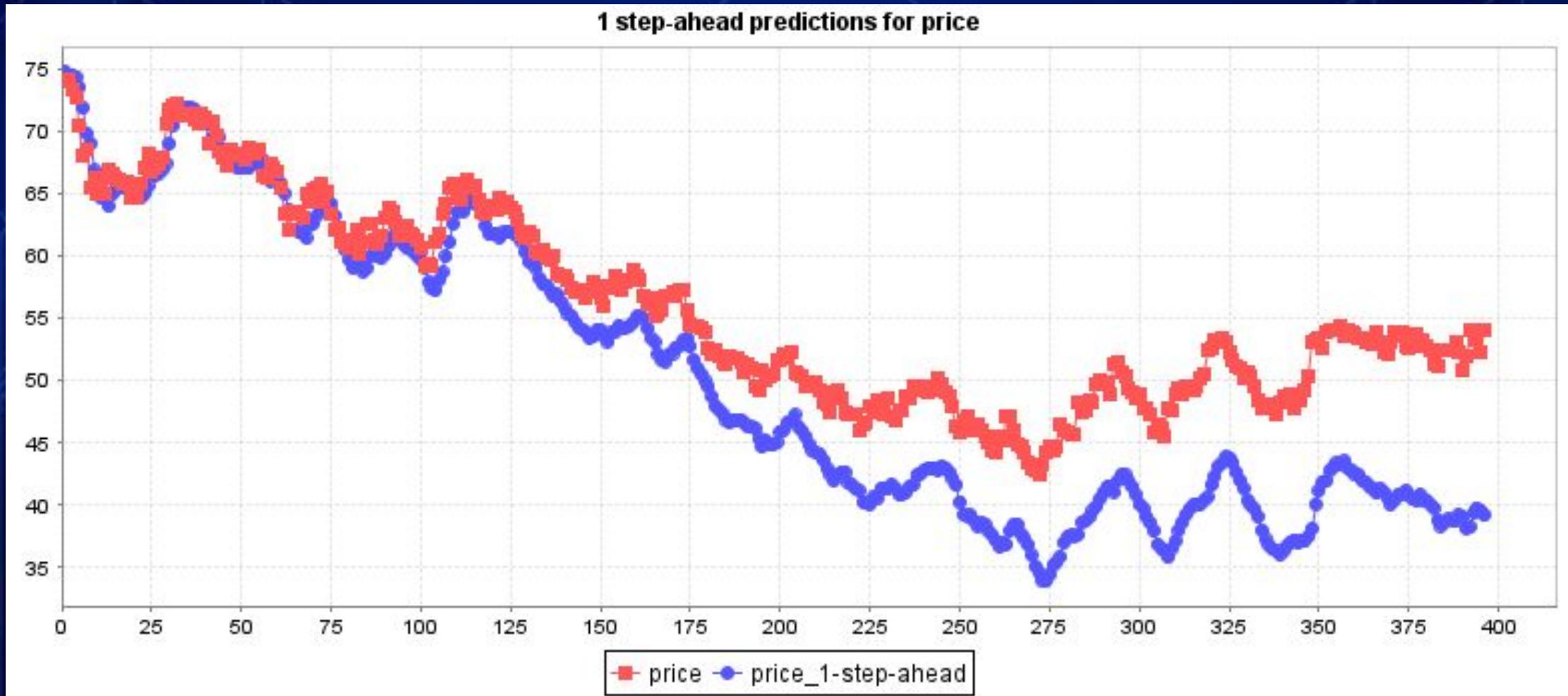
Disease  
Detection



# Exemple of Data Science Implementation

## Prediction

Prediksi Harga Minyak Indonesia



# Exemple of Data Science Implementation

## Image Recognition

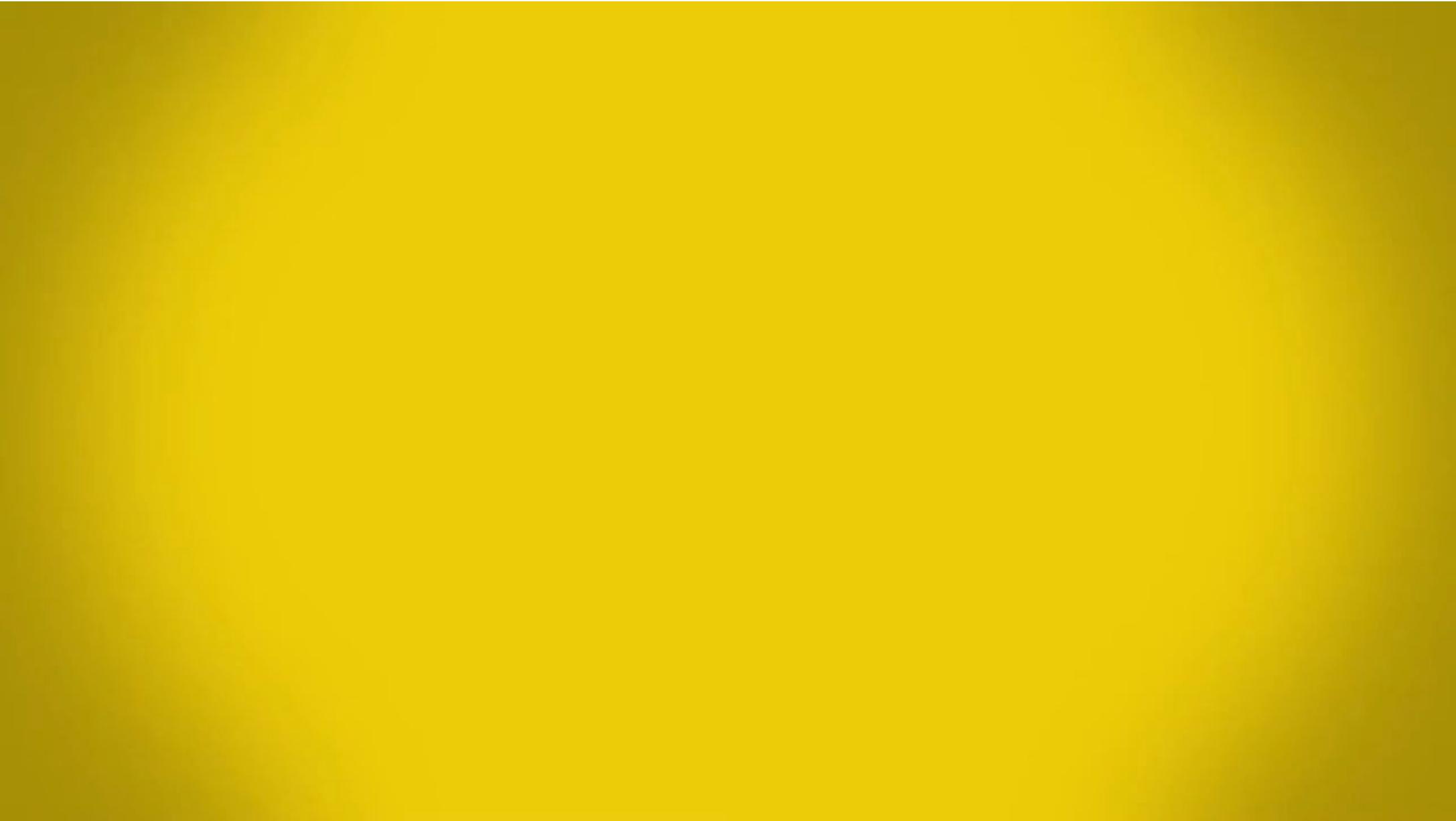


# Exemple of Data Science Implementation

## Sentiment Analysis

Headlines	SENTIMENT
gerak ihsg diprediksi lesu jelang akhir pekan	Neutral
inilah saham bidikan investor di salah satu ca...	Negative
bri raup laba bersih rp triliun sepanjang	Positive
nilai tukar rupiah di bca bni bri mandiri	Neutral

# **The Importance Role of Data Science**



# Data Science

## Advantages

A

It's in Demand

Abundance of  
Positions

B

C

Highly Paid Career

Highly Prestigious

D

E

Versatile

## Disadvantages

A

It is a Blurry Term

Mastering Data Science  
is near to impossible

B

C

Large amount of domain  
knowledge required

Arbitrary Data May Yield  
Unexpected Results

D

E

Problem of Data Privacy



# Data Scientist: Profesi Masa Depan



# Agenda:

1. Masa Depan

2. Profesi Masa Depan





Apa yang akan terjadi  
di masa depan?

# Menurut para futuris, kita akan mengendarai mobil self-driving pada 2030



Jalan tol khusus  
mobil self-driving



Armada Uber tanpa  
pengemudi



Pasar mobil self-driving  
senilai \$60 miliar



# Dan menerima belanjaan harian dari robot



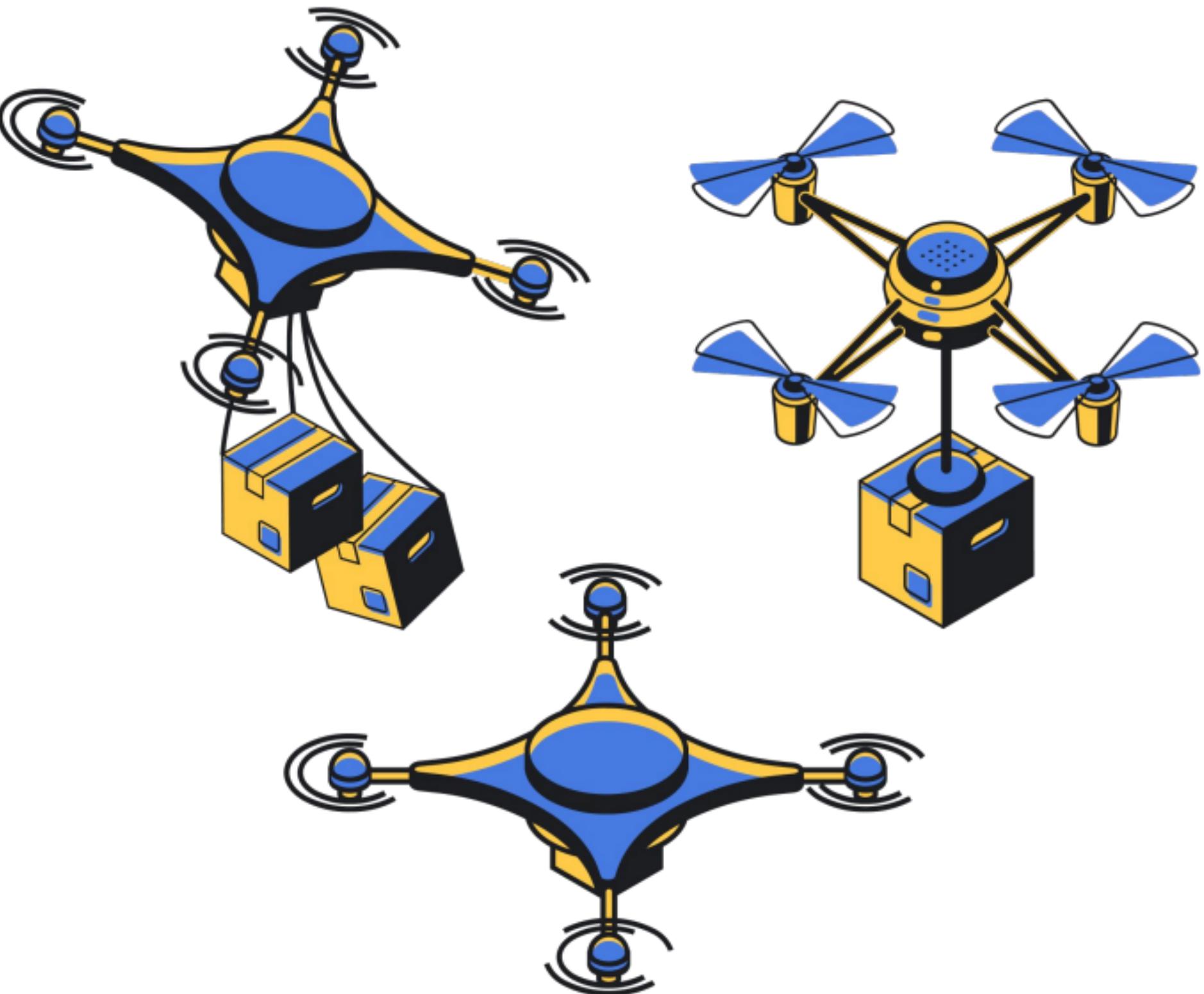
Drone mengantarkan berbagai produk dan sembako



Kurir digantikan robot



Pasar robot kurir senilai >2\$ miliar



# Dan membayar dengan mata



Gunakan gadget untuk  
main mata



Diagnosis yang  
lebih akurat



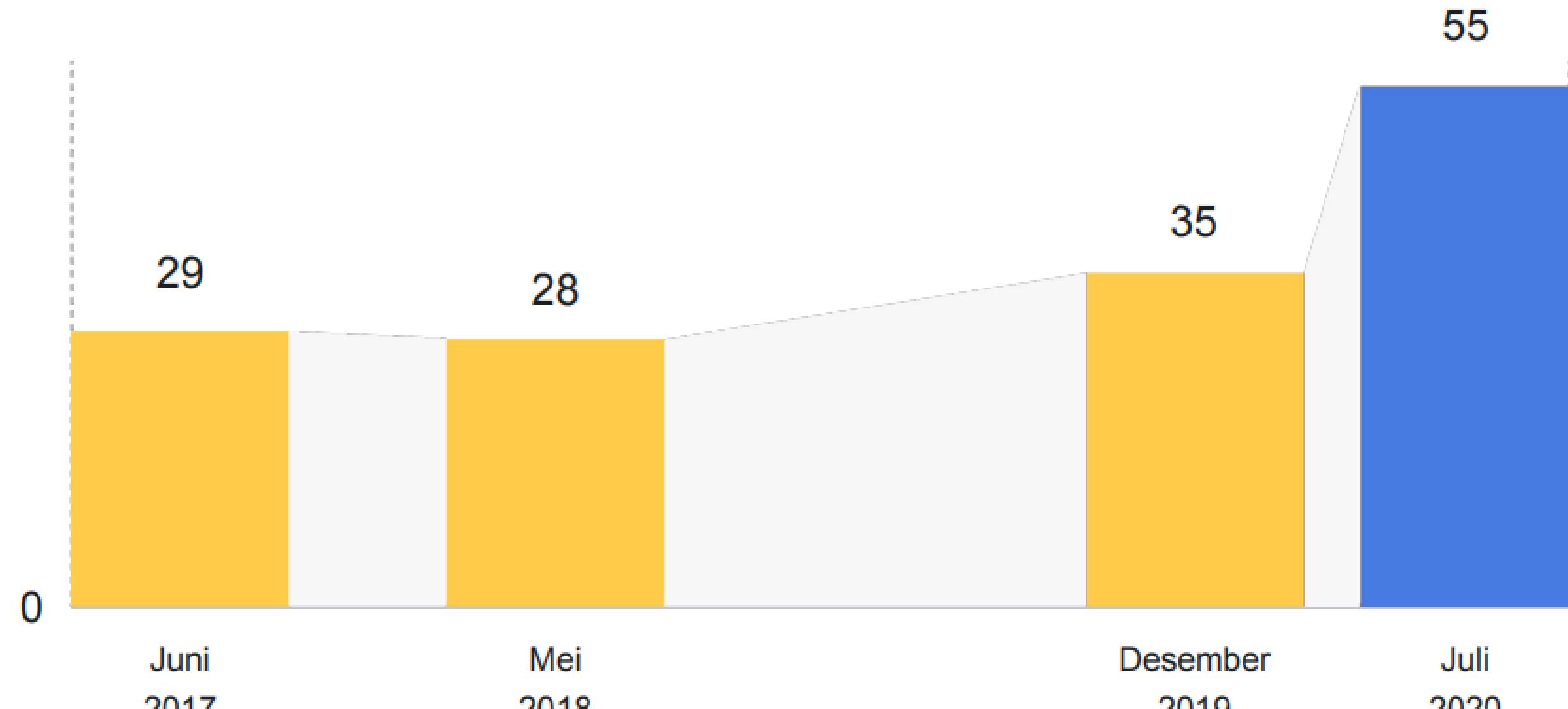
Pasar deteksi  
mata senilai \$5  
milyar



# Lockdown global mendekatkan kita ke masa depan ini

Covid-19 mempercepat penerapan teknologi digital hingga **7 tahun**

Dan **mendorong masif** masyarakat dan perusahaan menuju **digitalisasi**



Proporsi rerata produk dan layanan yang terdigitalisasi secara sebagian maupun sepenuhnya, %  
(McKinsey, 2021)



Siapa yang akan  
mengoperasikan masa  
depan?

Mereka berdiri di belakang layar dan mengembangkan teknologi ini

Artificial intelligence specialist

Robotics Engineer

Cloud Engineer

Data Specialist

Full Stack Engineer

Robotics Engineer

Full Stack Engineer

Cloud Engineer

Robotics Engineer

# Dan mereka masih kekurangan orang

Pada 2030, diperkirakan akan ada defisit global  
akan pekerja tech sebesar lebih dari  
**85 juta orang**

Data specialist adalah satu dari TOP-5  
pekerjaan paling dicari **di Indonesia**  
setelah pandemi

*termasuk validation analyst,  
machine learning engineer,  
dan data architect*

\* Sumber: Top 10 highest paying jobs in Indonesia for 2022, Michael Page - Jobs and recruitment agency in Indonesia

# 40%

Peningkatan gaji ahli data antara 2019  
dan 2021

# 62%

perusahaan Indonesia  
kekurangan data specialist  
(APEC)

# Mengapa Data Specialist begitu diminati?

# DATA SCIENTIST

Penambang Emas

Masa Depan



# Minyak adalah emas baru? Data adalah minyak baru

- Data specialists tidak menggali tanah, tapi menggali informasi
- Tidak dengan tangan sendiri, namun dengan algoritma
- Tidak mencari emas, namun mencari pola dan kesimpulan yang bermanfaat
- Namun keduanya mirip: data specialists dan penambang emas sama-sama menghasilkan uang

Jadi penambang emas  
di bidang apapun yang  
kamu sukai

Marketing

Kesehatan

Retail

Politik

Bisnis

Linguistik

HR

Keuangan

Pemerintahan

Media

# **TEMUKAN**

## **TAMBANG EMASMU!**

## References:

1. Folie Pengantar Sains Data UBM (DSC01)
2. Yandex Data School
3. Datanami.com

by: itsmecevi.github.io

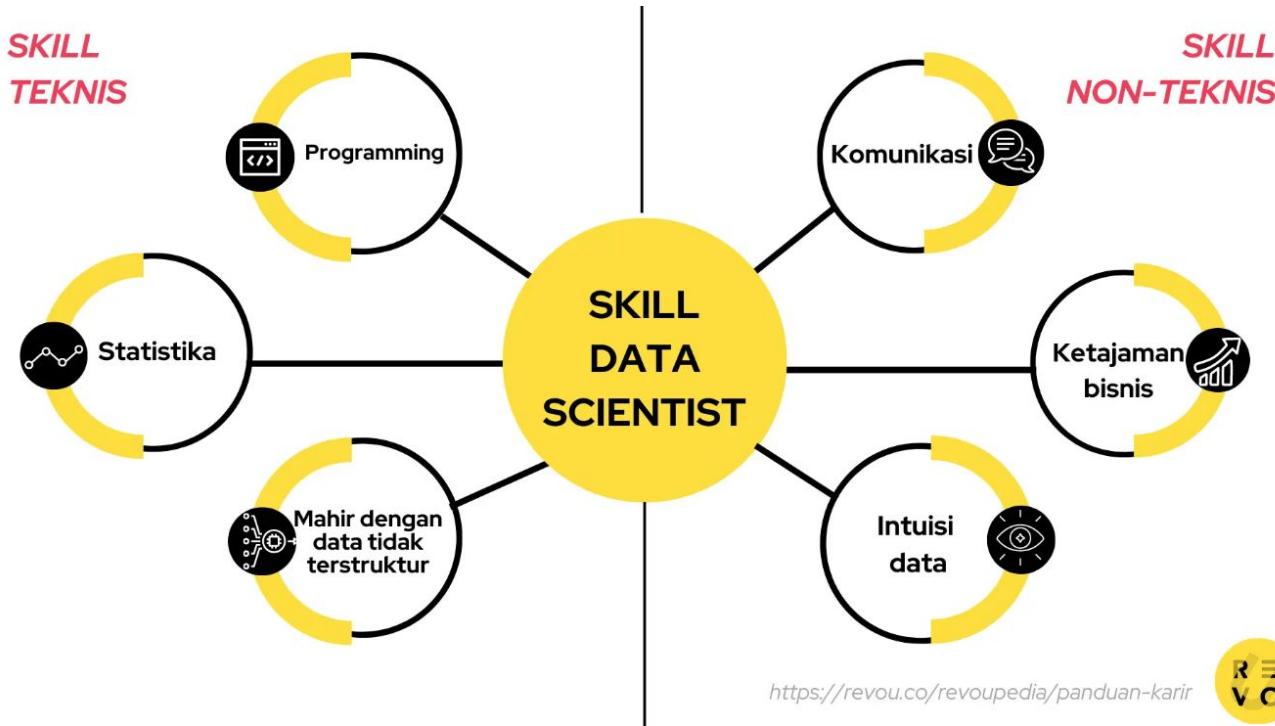
# PENGENALAN DATA SCIENCE

# Apa itu Data Science ?

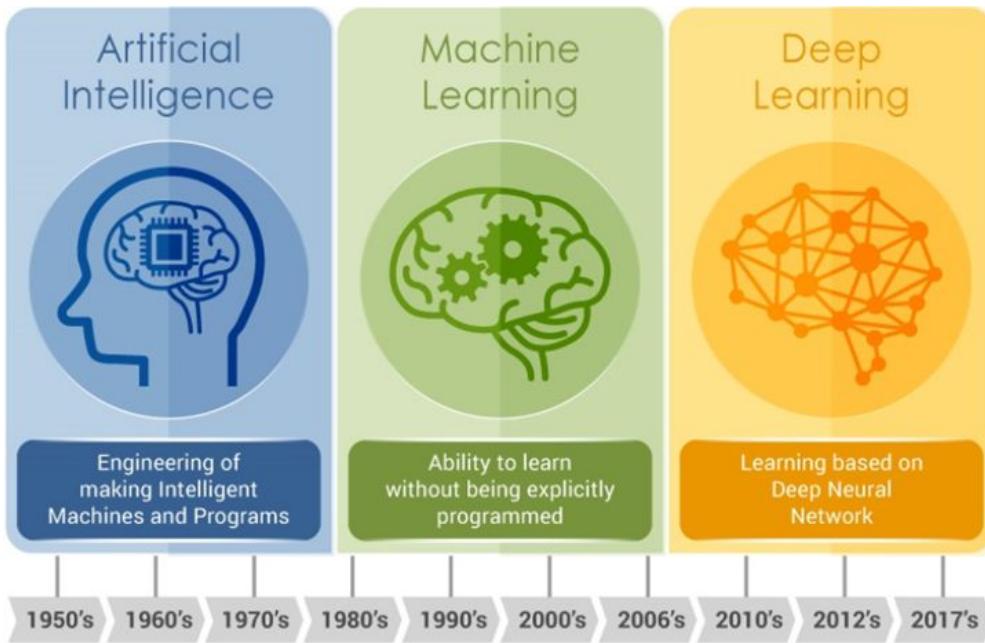
Data Science adalah ilmu yang menggabungkan matematika, statistika, dan ilmu komputer untuk menganalisa data skala kecil sampai dengan besar (big data) untuk menghasilkan suatu prediksi yang akurat untuk mengambil suatu keputusan dengan otomatis (AI) yang terus menerus belajar dan semakin cerdas

- Matematika
- Statistika
- Programming
- Big Data
- AI
- Machine Learning

# Kemampuan yang dibutuhkan seorang Data Scientist



# AI vs ML vs NN vs DL:



# Tujuan Data Science



- Menggali data dan memberikan informasi seakurat mungkin yang digunakan untuk :
- **Deskripsi** yaitu menampilkan pola data untuk dianalisa dan penemuan masalah.
- **Prediksi** yaitu melakukan prediksi berupa nilai, probabilitas maupun data dan kemudian merekomendasikan hasilnya untuk digunakan sebagai alat bantu pengambil keputusan maupun secara langsung digunakan secara otomatis oleh sistem.

# Penggunaan Data Science

Digunakan untuk banyak jenis seperti :

Prediksi populasi, cuaca, kondisi pasar, iklan, pendekripsi penipuan, klasifikasi gambar, pola kebiasaan pelanggan, jenis atau karakteristik pelanggan, marketing yang lebih terarah, rekomendasi produk, AI Game, pendekripsi anomali dan masih banyak kegunaan lainnya.

# Gambaran Cara Kerja Data Science

Study Hours	Rest Hours	Play Hours	Result
2	3	3	FAIL
4	2	1	PASS
3	2	1	PASS
1	1	1	PASS
1	2	3	FAIL

1	3	Testing	2
1	1	Set	1

← Predicted Value

- Agar mesin dapat belajar, dibutuhkan 2 jenis data yaitu
- **Data Training** berisi data fakta yang ada
- **Data Tes** berisi data yang ingin diprediksi

# MULAI DARI MANA?

- Menggali ide riset yang mau dikelola untuk menjadi data science
- Proses pengumpulan data (berdasarkan variable-variable yang relevan)
- Pengolahan data (mengolah data, menentukan type data, mengelompokan data)
- Menyajikan data (visualisasi data)



# Proses Pengumpulan data

## Data Primer

Data yang didapatkan langsung dari sumber aslinya, tidak melalui perantara.

Bisa berbentuk:

- Opini / persepsi secara individu dan kelompok
- Hasil observasi terhadap benda/kegiatan

## Data Sekunder

Data yang didapatkan secara tidak langsung melalui media perantara catatan/documenter penelitian arsip

Apa?

Siapa?

Kapan?

Dimana?

Bagaimana?

# Cara mengumpulkan Data

- Survey dan kuisioner
- Dataset statistik
- Wawancara
- Observasi
- Studi dokumen
- Teknik triangulasi



# DATA SCIENCE AT UBM

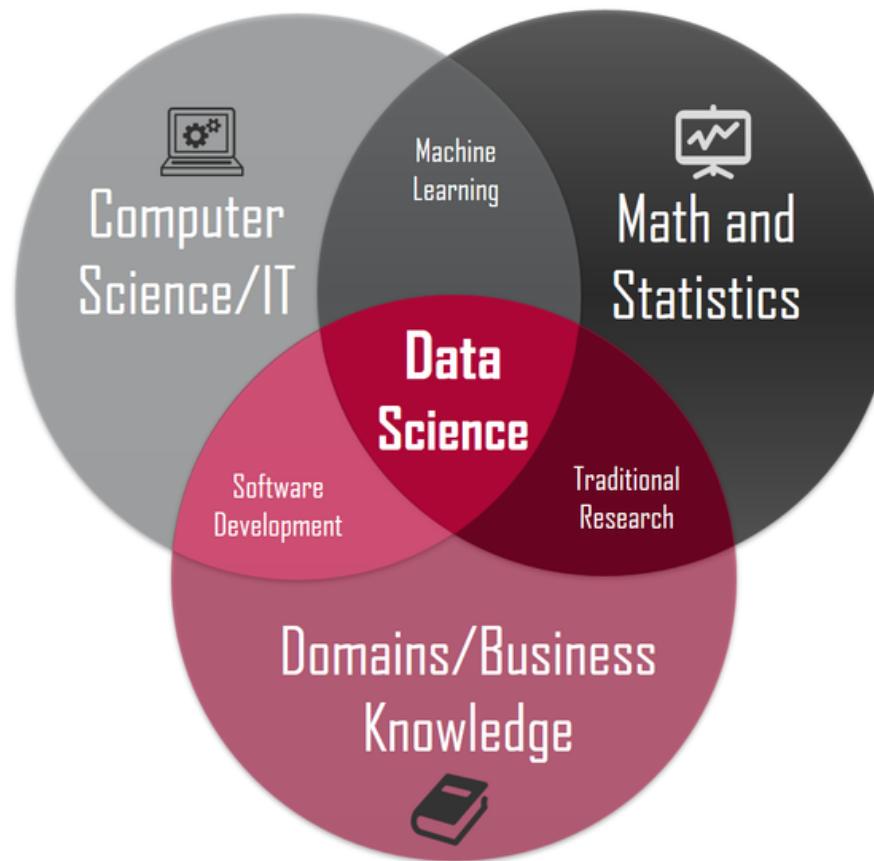
UNIVERSITAS BUNDA MULIA



# Data Visualization Options

# Overview

Secara garis besar, Data Science terdiri dari 3 domain utama



# What will you learn?

**Semester 1** -> Dasar-dasar Sains Data & Algoritma

**Semester 2** -> Struktur Data (Database) & Statistika

**Semester 3** -> R Programming

**Semester 4** -> Big Data, Data Visualisasi & Machine Learning

**Semester 5** -> Data Modeling

**Semester 6** -> Business Intelligence & Riset Sains Data

**Semester 7** -> Magang / Kampus Merdeka

**Semester 8** -> Skripsi



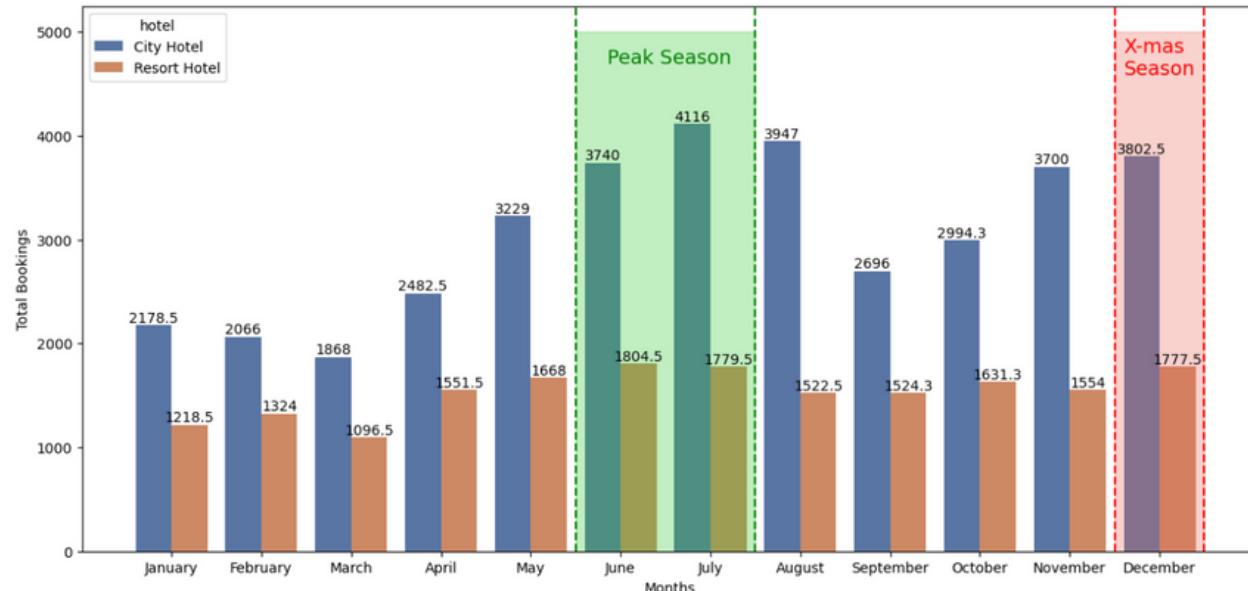
# Your Degree: **Sarjana Sains Data**

# A Glimpse of DS Project - Data Visualization

	hotel	arrival_date_month	hotel_count
4	City Hotel	January	2178.5
16	Resort Hotel	January	1218.5
3	City Hotel	February	2066.0
15	Resort Hotel	February	1324.0
7	City Hotel	March	1868.0

**Average Number of Bookings per Month for Every Hotel Category**

The figure shows that the number of hotel bookings started to rise gradually from the beginning of the year and reached its peaked in July at 4116 bookings for the City Hotel. In addition, the number of bookings for City Hotel always becomes a favourite among the customers every month.



```
plt.figure(figsize=(15,7))
fig_2 = sns.barplot(data=booking_trends_avg, x='arrival_date_month', y='bookings', hue='hotel')

plt.title('Average Number of Bookings per Month\nfor Every Hotel Category', fontsize=18, color='black', weight='bold', pad=10)
plt.text(x=-1, y=5450, s='The figure shows that the number of bookings for City Hotel started to rise gradually from the beginning of the year and reached its peaked in July at 4116 bookings. In addition, the number of bookings for City Hotel always becomes a favourite among the customers every month.', fontsize=10, style='italic')
```

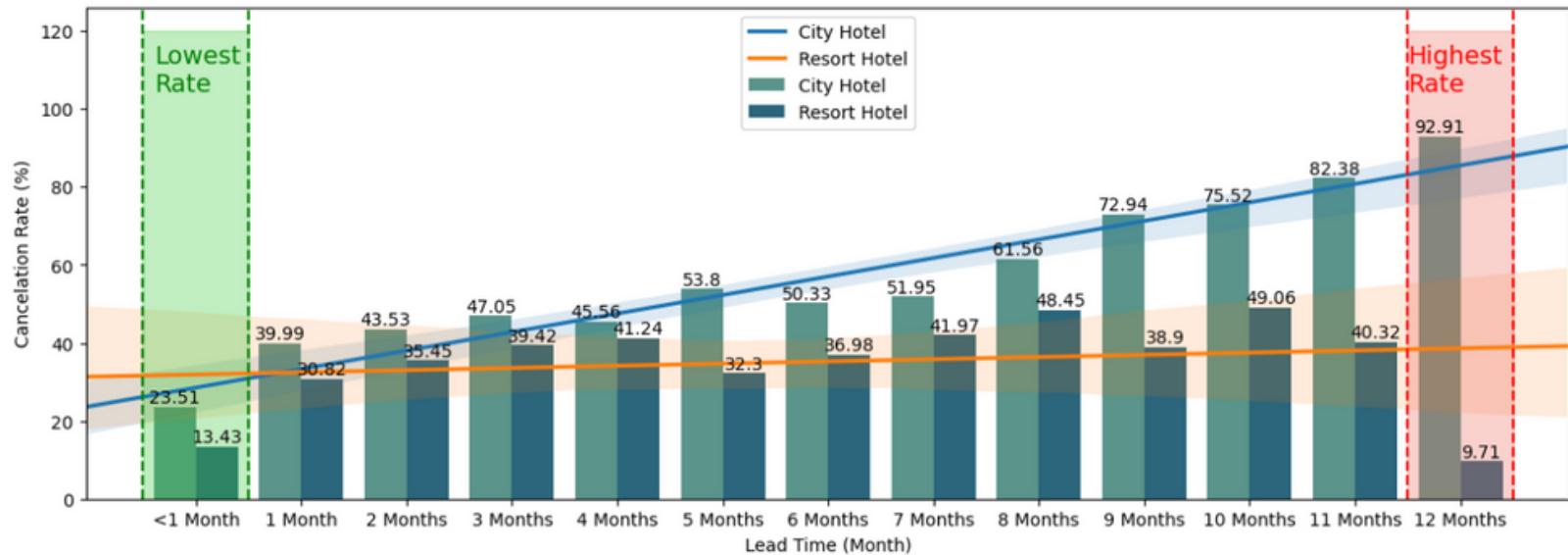
# A Glimpse of DS Project - Data Visualization

## Ratio of Cancelation Bookings With The Lead Time

The figure shows the correlation between the cancelation requests and the lead time.

It shows that, there is also a positive correlation between the two mentioned variables like the previous.

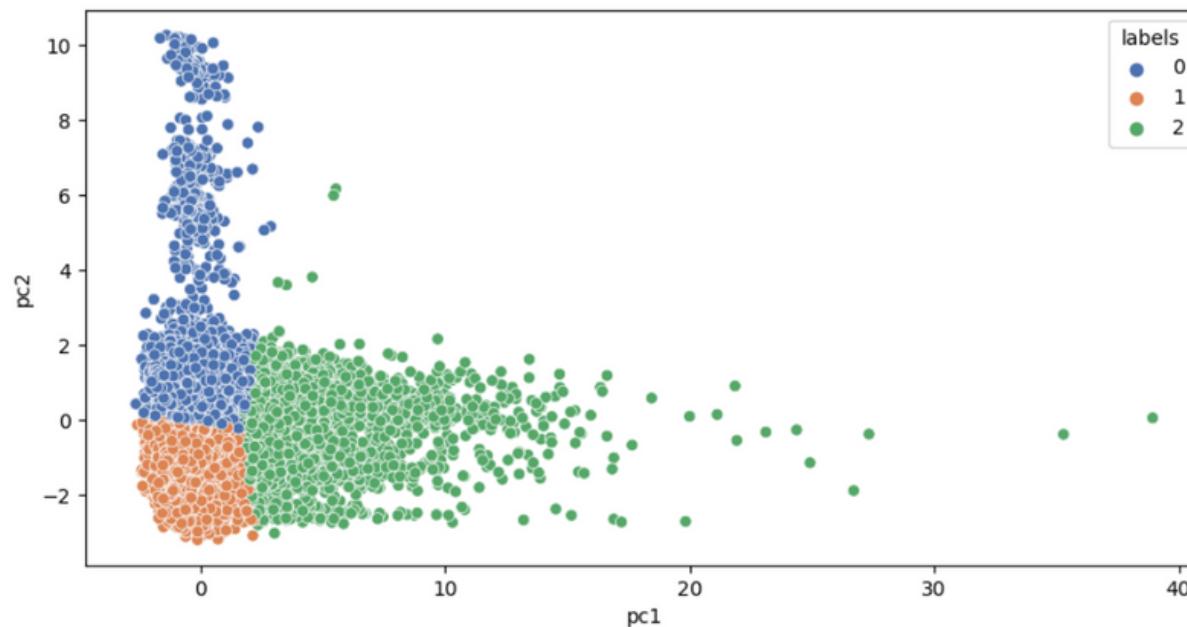
The longer the customers book the hotel reservation, the more probability of cancelation requests will also be made by the customers



# A Glimpse of DS Project - Data Modeling

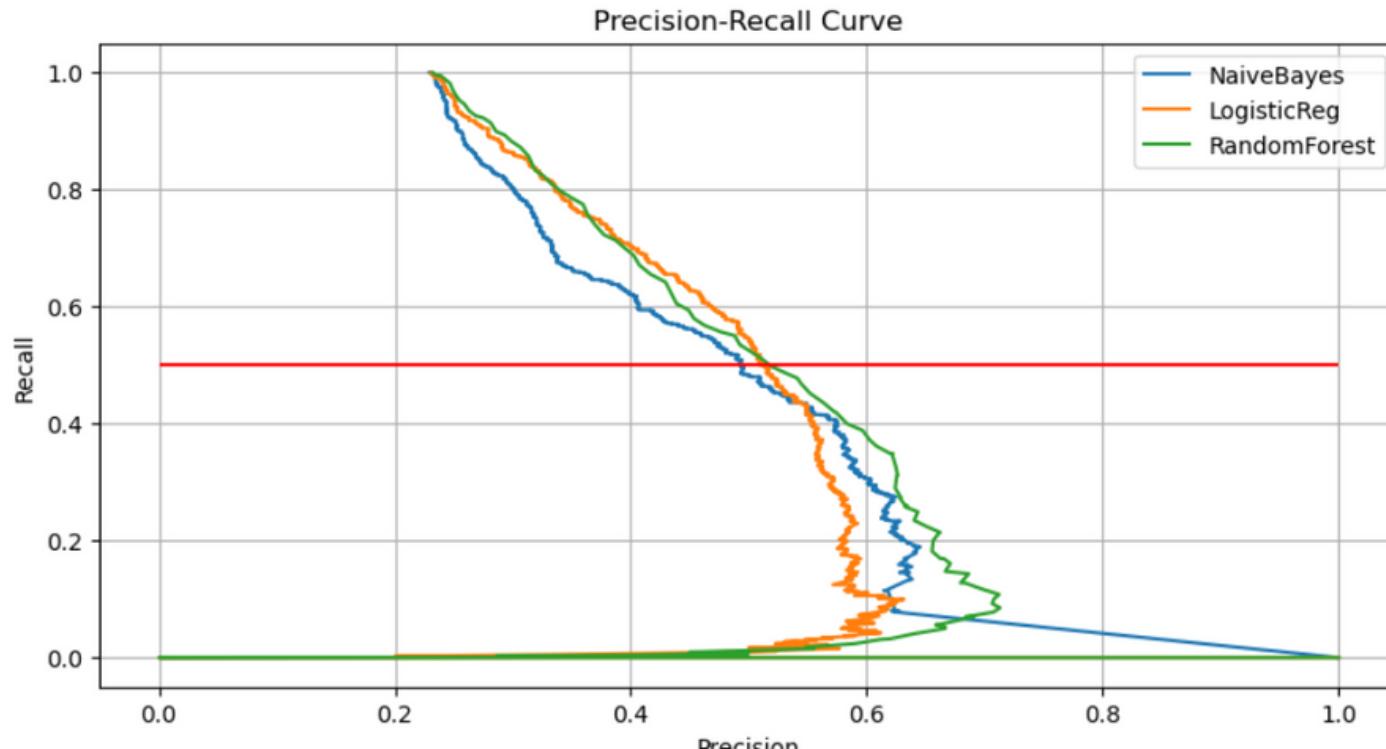
## Fitting Model(K-Means) ↗

The use of the following code resulting the clustering below.  
`KMeans(n_clusters=3, init='k-means++', max_iter=300, n_init=10, random_state=42)`



# A Glimpse of DS Project - Data Modeling

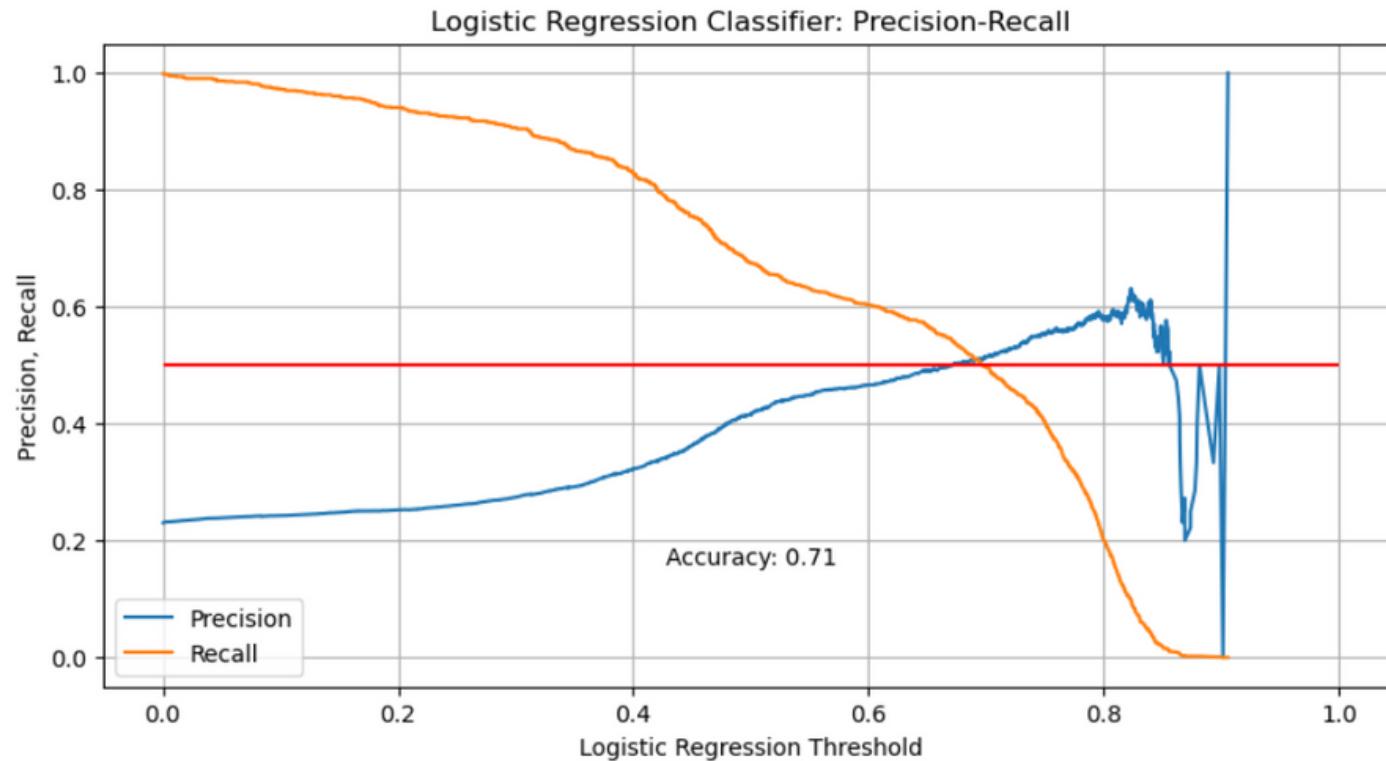
From DMA Charts above, it can be seen that even though Logistic Regression recall's score is below Naive Bayes, the accuracy of Logistic Regression is still better. Therefore, the Logistic Regression will be used in this case.



# A Glimpse of DS Project -

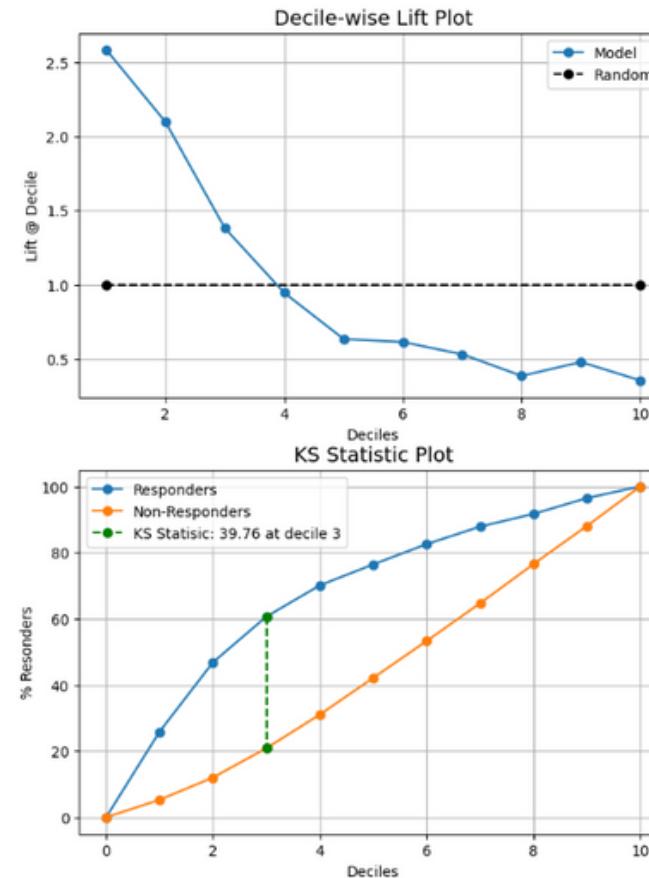
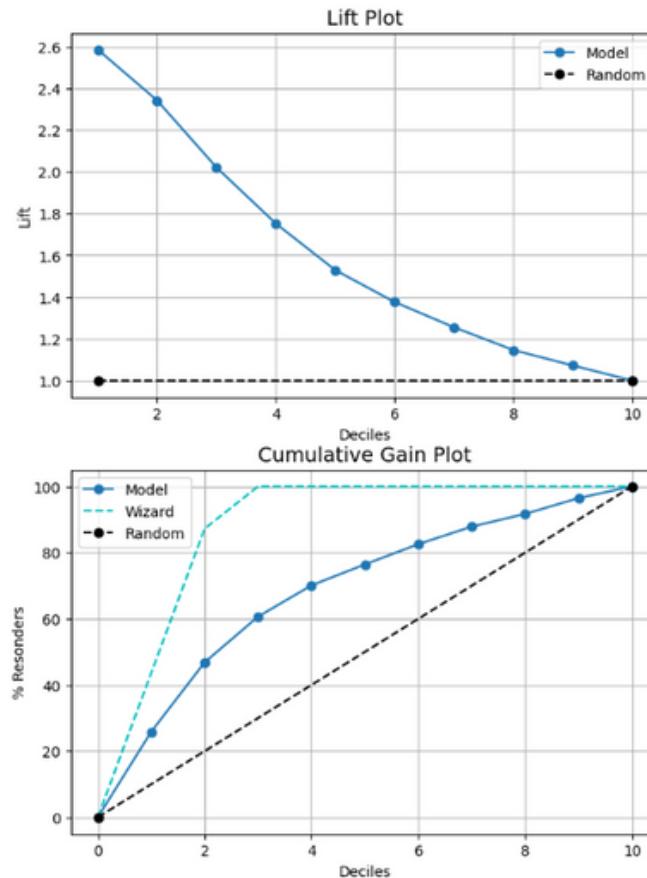
## Data Modeling

The precision-recall curve above tells that by giving the model threshold of 0.5, the Logistic Regression Model is still giving the best output of recall compared to the Naive Bayes and Random Forest model.



# A Glimpse of DS Project - Model Evaluation

By using the library from `kds`, we evaluate the model's performance by looking at the Gain and Lift charts. The code `kds.metrics.report(y_test, prob_glm[:,1])` is used to generate the desire report.





# Thank You