



**LEARNING PROGRESS REVIEW WEEK 1**

# **AI-POWERED DATA SCIENCE**

**FUNDAMENTALS OF DATA SCIENCE & METHODOLOGY  
| ESSENTIALS OF DATA AND DATABASES**

**TEAM 4  
DATA EXPLORERS**



# DATA EXPLORERS TEAM MEMBERS



Steven  
Rusly



Raihan  
Maulana  
Nilmada



Muhammad  
Syauqi  
Luvaniesky



Alfian  
Rosyid



Edvan  
Tazul  
Arifin




Pebry  
Zauhary




**SESI 1**

**INTRODUCE DIGITAL SKOLA**



# ABOUT US

Digital Skola adalah platform Edutech Indonesia yang didirikan sejak tahun 2020. Kami berkomitmen untuk membina talenta masa depan yang siap kerja dan sukses di dunia kerja. Melalui program Bootcamp, Mini Bootcamp, dan Corporate Training, Digital Skola membekali siswa dan profesional dengan pengetahuan dan keterampilan digital terkini yang memenuhi kebutuhan industri saat ini dan masa depan. Opsi pembelajaran kami dengan fleksibel mencakup kelas online dan offline serta real-experiences internship.





# OUR VISION

Digital Skola berupaya untuk membekali para profesional muda Indonesia dengan keterampilan yang dibutuhkan untuk memenuhi tuntutan industri saat ini dan masa depan, serta unggul dalam karir mereka.



**SESI 2**

**FUNDAMENTALS OF DATA SCIENCE &  
METHODOLOGY**

# DATA SCIENCE

## Pengertian Data Science

- Data science adalah sebuah keterampilan mengolah data-data untuk mendapatkan wawasan dan algoritma prediksi.

Data Science menggunakan tiga bidang ilmu:

1. Matematika atau Statistika
2. Computer Science
3. Keahlian dalam bidang tertentu



# HASIL KERJA

1. Decision Support: Membuat insight yang penting dari data-data untuk pengambilan keputusan yang lebih akurat dan specific
2. Scoring: Membangun algoritma scoring dari data scoring yang sudah dikumpulkan sebelumnya untuk memprediksi data scoring berikutnya
3. Recommendation: Membangun algoritma recommendation berdasarkan eksplorasi data untuk menemukan hubungan antarinformasi yang dapat menjadi rekomendasi untuk berbagai tujuan



# HASIL KERJA

1. Classification: Membangun algoritma classification berdasarkan data dari kategori sebelumnya sebagai formula untuk memprediksi kategori berikutnya
2. Forecasting: Membangun algoritma forecasting berdasarkan pola dan tren tiap periode waktu sebelumnya untuk memprediksi periode setelahnya
3. Detection: Membangun algoritma detection untuk membaca aktivitas yang tidak terdapat pada aktivitas sebelumnya sebagai bentuk tindakan antisipasi jika terjadi penyalahgunaan

# KETERAMPILAN DATA SCIENTIST

## HARDSKILLS

- SQL & Database

Berguna dalam pengumpulan, penyaringan, dan menjadi langkah awal pengolahan data.

- Python & R

Biasa digunakan untuk machine learning & visualisasi informasi, terutama untuk **Big Data**.

- Statistics

Penting untuk menganalisis berbagai jenis data mengetahui metode analisis yang paling efektif.

- Data Visualization & Business Intelligence (BI)

BI berguna untuk analisis informasi terkait bisnis. Analisis ini dibantu dengan visualisasi data.

- Machine Learning

Cabang **Artificial Intelligence (AI)** yang memungkinkan komputer mempelajari data dengan mandiri.

# KETERAMPILAN DATA SCIENTIST

## SOFTSKILLS

- Data Storytelling

Kemampuan menjelaskan data secara efektif dan mudah dipahami.

- Communication

Seorang Data Scientist harus mahir berkoordinasi dengan berbagai **stakeholders**.

- Critical Thinking

Berguna dalam **memahami konteks** dan masalah, mengolah data, dan menyajikan hasil kerja *data science*.

- Teamwork

*Data scientist* tidak hanya berkutat dengan data, tetapi juga **bekerja secara tim dan lintas disiplin**.

# DATA SCIENCE SUCCESS STORIES

## ● AirBnB

Sejak menerapkan **data analytics** pada 2008, valuasinya meningkat hingga **\$25,5 juta** dengan >25 juta orang pengguna di 192 negara.

### ● Implementasi:

- A/B Testing
- Image Recognition
- NLP
- Predictive Modelling
- Regression
- Collaborative Filtering

## ● Zara

Dengan **data science**, mampu menjual **11,000 jenis barang/tahun**. Pendapatan meningkat **220%** pada di periode 2004-2015.

### ● Implementasi:

- Prediction
- Decision Support

## ● Nike

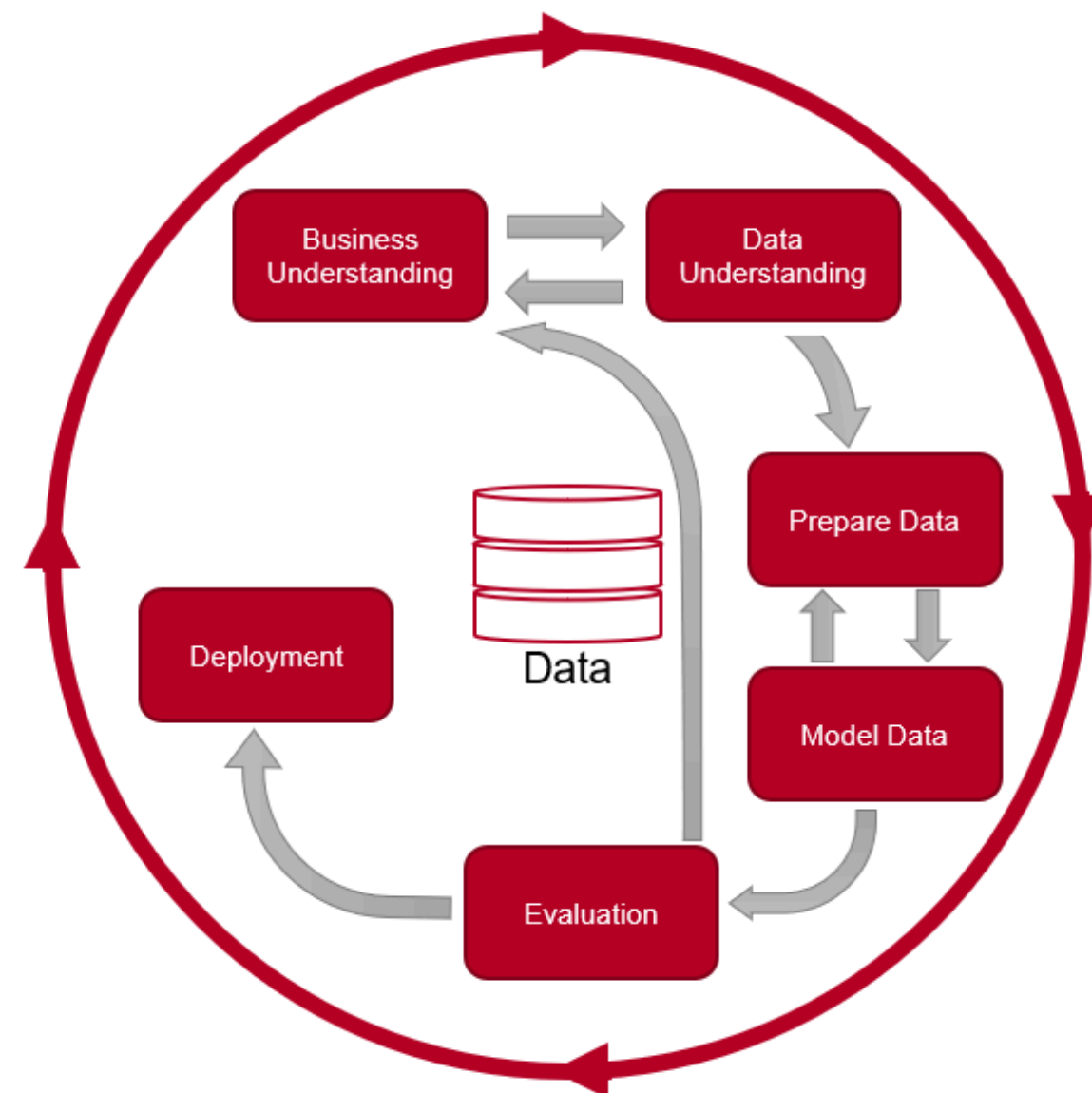
Mencatat peningkatan penjualan sebanyak **\$39,12 juta**. Nike berkontribusi **60%** dari bisnis digital global pada tahun 2018.

### ● Implementasi:

- AI for Customers
- Prediction

# FRAMEWORK DATA SCIENCE

**CRISP DM:** CRISP DM atau Cross Industry Standard Process for Data Mining ini adalah metodologi yang paling sering digunakan dalam dunia data science karena bisa diaplikasikan ke banyak Industri.





# FRAMEWORK DATA SCIENCE

- Business Understanding

Sebelum melakukan analisis, seorang data scientist harus memahami terlebih dahulu bisnis yang akan dianalisa agar bisa melakukan analisis secara efektif

- Data Understanding

Setelah proses business understanding telah dilaksanakan step selanjutnya adalah data understanding. Data apa saja yang kita punya? bagaimana korelasinya dll.

- Data Preparation

Setelah memahami data yang ada step selanjutnya adalah preparation yang meliputi cleaning, sampling, dan formatting



# FRAMEWORK DATA SCIENCE

- Data Modeling

setelah data sudah di pre process kita sudah bisa membuat dan melatih algoritma sesuai dengan kebutuhan masalah yang ingin dipecahkan dengan akurasi yang baik

- Deployment

Langkah terakhir adalah deployment. Yaitu algoritma yang tadi sudah di buat dan di latih algoritma ini sekarang bisa digunakan oleh bisnis yang bersangkutan






# **SESI 3 – ESSENTIALS OF DATA & DATABASES**



# ESSENTIALS OF DATA AND DATABASES


## PENGERTIAN DATA

Data merupakan kumpulan suatu fakta secara umum adalah kumpulan fakta, angka, simbol, atau informasi mentah yang belum diolah atau diinterpretasikan. Data dapat berbentuk angka, teks, gambar, suara, atau video, dan biasanya digunakan sebagai dasar untuk pengambilan keputusan atau analisis lebih lanjut.

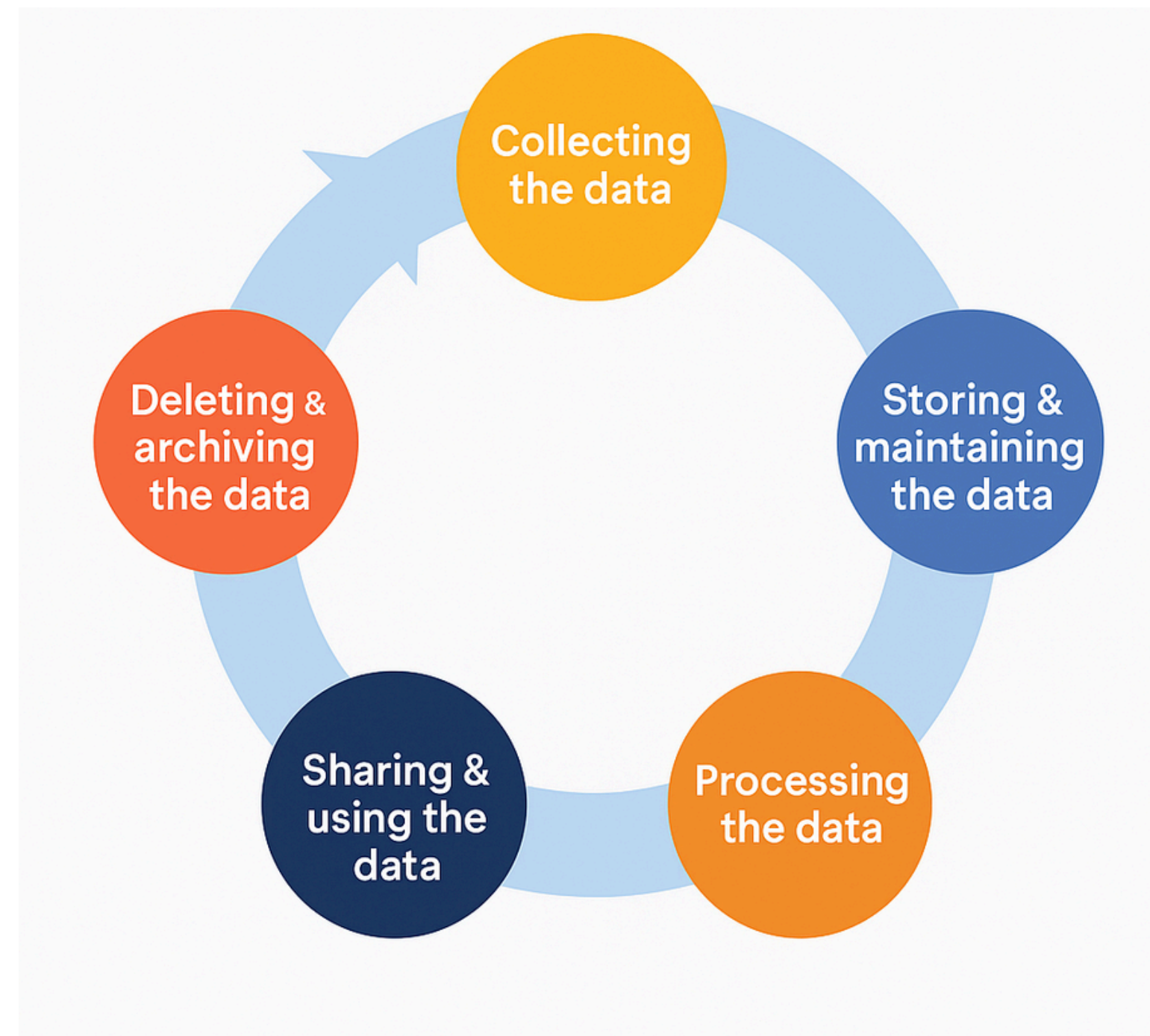




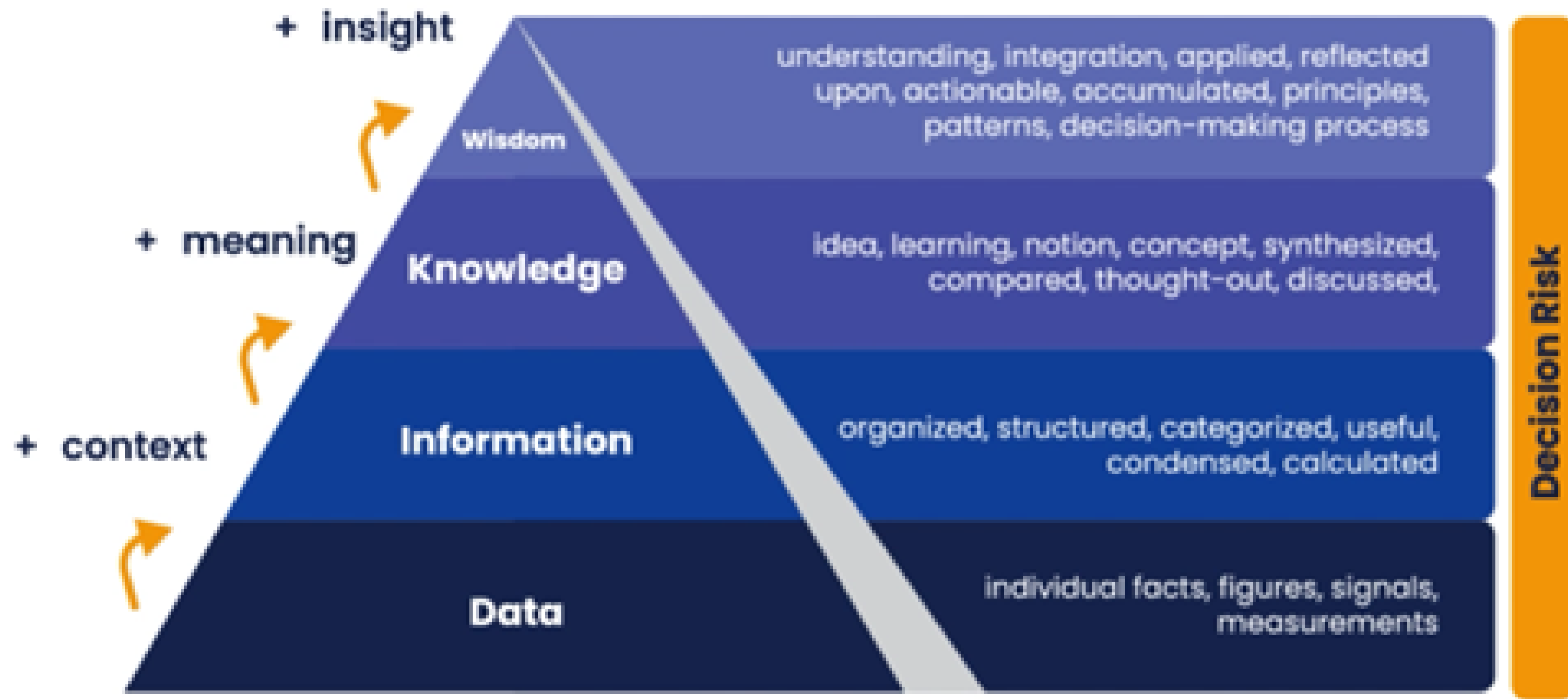
# STRUKTUR UMUM BASIS DATA :

- **Schema** : Struktur atau rancangan dari database; berisi definisi tabel, kolom, dan relasinya.
  - **Tabel** : Tempat menyimpan data dalam bentuk baris dan kolom.
  - **Column** : Bagian vertikal dalam tabel yang menunjukkan jenis data tertentu.
  - **Row** : Bagian horizontal dalam tabel, mewakili satu data lengkap.
  - **Value** ; Isi dari sel pada kolom dan baris tertentu.
- 

# DATA LIFECYCLE MANAGEMENT



# DIKW HIERARCHY



# KOMPONEN SQL

1. **Data Definition Language (DDL)** : CREATE, ALTER, DROP.
2. **Data Control Language (DCL)** : GRANT REVOKE.
3. **Data Manipulation Language (DML)** : SELECT, INSERT, UPDATE, DELETE



# ESSENTIALS OF DATA AND DATABASES

## JENIS JENIS DATA

1. **Data kuantitatif** – berbentuk angka dan dapat diukur (misal: tinggi badan, usia).
2. **Data kualitatif** – berbentuk deskripsi atau kategori (misal: warna favorit, jenis kelamin).





# DATA LIFECYCLE

1. **Collecting Data** : Proses pengumpulan data.
2. **Storing Data** : data yang telah terkumpul simpan ke Database.
3. **Processing data** : data diproses oleh data enginer agar data lebih rapi & terstruktur.
4. **Sharing & using the data** : Proses analisis data.
5. **Deleting & archiving data** : data yang sudah tidak relevan dapat diarsipkan.

The image features a white background with the text 'THANK YOU' centered. The corners are decorated with abstract geometric shapes in dark blue and light blue, connected by thin brown lines, resembling stylized buildings or flags.

**THANK YOU**