

# DATA SCIENCE and BUSINESS INTELLIGENT

---

Author: Egi Safitri

Meeting 1



# Introduction to Data Science

1. Data Science dan Kepentingannya (Definisi dan Contoh)
2. Proses Data Science (Apa yang sebenarnya dilakukan oleh Data Science dan Data Scientist?)
3. Keterkaitan AI dan Data Science.
4. Prasyarat untuk Data Science (Apa yang perlu diketahui seseorang untuk menjadi Data Scientist?)

## What is Data Science?



Data Science adalah bidang interdisipliner yang menggunakan algoritma, prosedur, dan proses untuk menganalisis sejumlah besar data guna menemukan pola tersembunyi, menghasilkan wawasan, dan mendukung pengambilan keputusan.

## Importance of Data Science

1. Peluang Karir
2. Membantu merek memahami pelanggan mereka dengan lebih baik dan lebih mendalam
3. Memungkinkan merek untuk mengkomunikasikan cerita mereka dengan cara yang lebih menarik dan kuat
4. Big Data adalah bidang baru yang terus berkembang Temuan dan hasilnya dapat diterapkan pada hampir semua sektor seperti perjalanan, kesehatan, dan Pendidikan
5. Data Science dapat diakses oleh hampir semua sektor

## Application of Data Science?

1. Peluang Karir
2. Membantu merek memahami pelanggan mereka dengan lebih baik dan lebih mendalam
3. Memungkinkan merek untuk mengkomunikasikan cerita mereka dengan cara yang lebih menarik dan kuat
4. Big Data adalah bidang baru yang terus berkembang Temuan dan hasilnya dapat diterapkan pada hampir semua sektor seperti perjalanan, kesehatan, dan Pendidikan
5. Data Science dapat diakses oleh hampir semua sektor

## Importance of Data Science

1. Pencarian Internet
2. Iklan Digital Sistem Rekomendasi
3. Pemrosesan Gambar
4. Pengenalan Suara
5. Situs Perbandingan Harga
6. Perencanaan Rute Penerbangan
7. Deteksi Penipuan dan Risiko
8. Logistik Pengiriman
9. Internet of Things (IoT)
10. Kesehatan
11. Kecerdasan Buatan (AI)
12. Realitas Augmentasi (AR)
13. Analisis Grafik
14. Data Science Agile
15. Personalisasi Hyper
16. Pembelajaran Mesin (Machine Learning)
17. Analisis Perilaku
18. dan lain-lain.

## CONTOH

- Perusahaan pertanian Cargill mengembangkan aplikasi pelacakan data seluler yang membantu petani udang mengurangi tingkat kematian.
- Perusahaan logistik Pitt Ohio menggunakan data historis dan analitik prediktif untuk memperkirakan waktu pengiriman dengan akurasi 99%.
- Perusahaan minyak raksasa Shell menggunakan data science untuk mengantisipasi kegagalan mesin di fasilitas di seluruh dunia.
- Dan masih banyak lagi...

## Facts on Data Generation

1. Setiap hari, 2,5 kuintiliun byte data telah dibuat.
2. Dengan begitu banyak informasi di ujung jari kita, kita terus menambah tumpukan data setiap kali mencari jawaban melalui mesin pencari.
3. Statistik menunjukkan bahwa lebih dari 500 terabyte data baru dimasukkan ke dalam database Facebook setiap hari. Bagaimana dengan Instagram, TikTok, X, YouTube, dan lainnya?

# How Does Data Science Work?

1. **Mengumpulkan Data:** Data mentah dikumpulkan dari berbagai sumber yang menjelaskan permasalahan bisnis.
2. **Menganalisis Data:** Dengan menggunakan berbagai analisis statistik dan pendekatan machine learning, pemodelan data dilakukan untuk mendapatkan solusi optimal yang paling baik menjelaskan masalah bisnis.
3. **Wawasan:** Wawasan yang dapat ditindaklanjuti yang berfungsi sebagai solusi untuk masalah bisnis yang dikumpulkan melalui data science.

# Example



## Collect Data

Gather the previous data



## Analyze Data

Use statistical analysis to



## Insights

Use machine learning to



# AI AND DATA SCIENCE

*"Data Science dan AI bukanlah hal yang sama."*

- Data Science dan AI adalah dua teknologi yang mengubah dunia. Sementara AI mendukung operasi Data Science, Data Science tidak sepenuhnya bergantung pada AI.
- Data Science memimpin Revolusi Industri ke-4.
- Data Science juga memerlukan algoritma machine learning, yang menyebabkan ketergantungan pada AI.

# AI vs DS

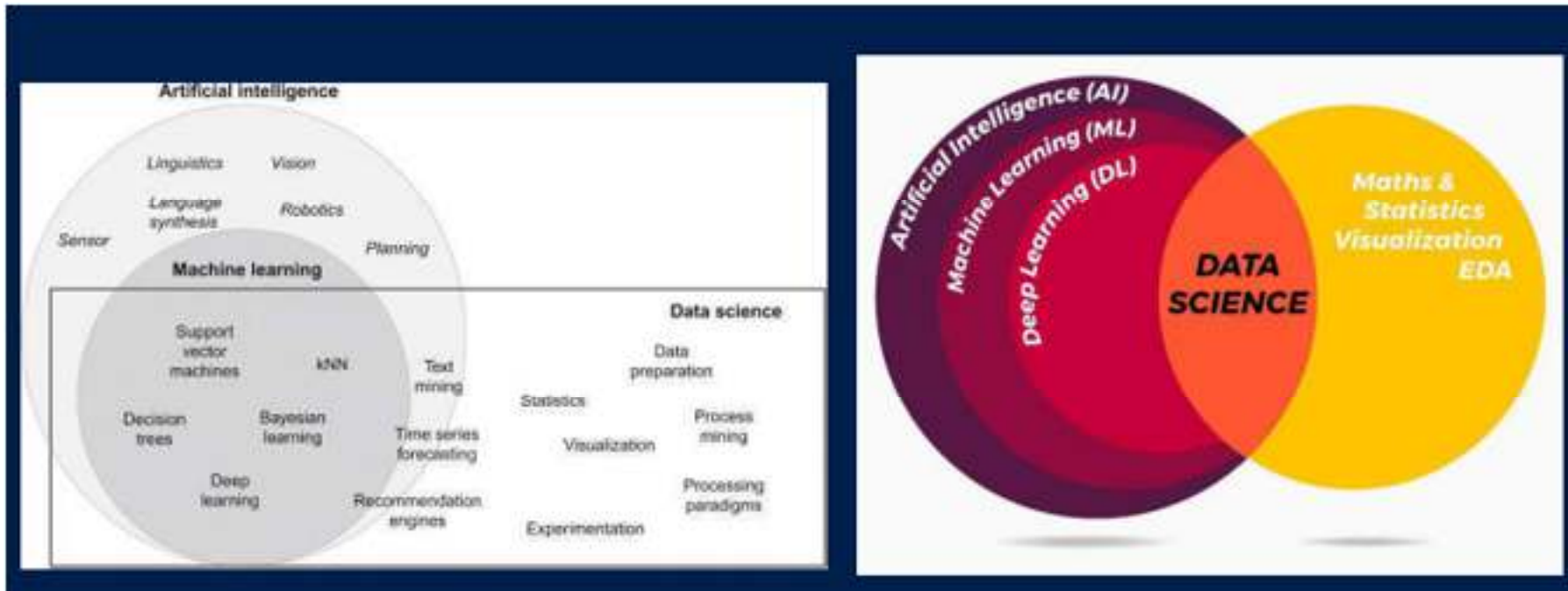
1. Pekerjaan di bidang Data Science membutuhkan pengetahuan bahasa pemrograman Machine Learning seperti R dan Python untuk melakukan berbagai operasi data serta keahlian ilmu komputer.
2. Data Science menggunakan lebih banyak alat selain AI, karena Data Science melibatkan beberapa langkah untuk menganalisis data dan menghasilkan wawasan.
3. Model Data Science dibangun untuk wawasan statistik, sedangkan AI digunakan untuk membangun model yang meniru kognisi dan pemahaman manusia.

# AI vs DS

Saat ini, industri membutuhkan keduanya, Data Science dan AI.

- **Data Science** membantu perusahaan mengambil keputusan berbasis data dan menilai kinerja mereka di pasar.
- **AI** membantu industri bekerja dengan perangkat dan perangkat lunak yang lebih cerdas, yang dapat meminimalkan beban kerja dan mengoptimalkan seluruh proses untuk meningkatkan inovasi.

# AI vs DS



# AI vs DS

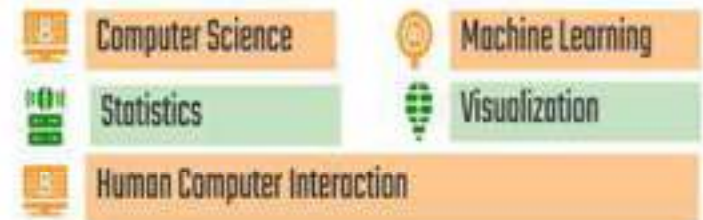
- **AI (Kecerdasan Buatan)** mencakup berbagai bidang seperti linguistik, visi komputer, robotika, dan perencanaan.
- **Machine Learning (ML)** adalah bagian dari AI yang mencakup algoritma seperti kNN, Support Vector Machines, Decision Trees, Deep Learning, dan Bayesian Learning.
- **Data Science** lebih fokus pada persiapan data, statistik, pemrosesan data, visualisasi, eksplorasi data, dan peramalan tren.

# AI vs DS

- AI mencakup berbagai teknologi yang meniru kecerdasan manusia.
- ML adalah subset AI yang memungkinkan sistem belajar dari data tanpa diprogram secara eksplisit.
- Data Science menggunakan AI dan ML untuk mengolah dan menganalisis data guna mendapatkan wawasan yang bermanfaat.

# PREREQUISITES FOR DATA SCIENCE

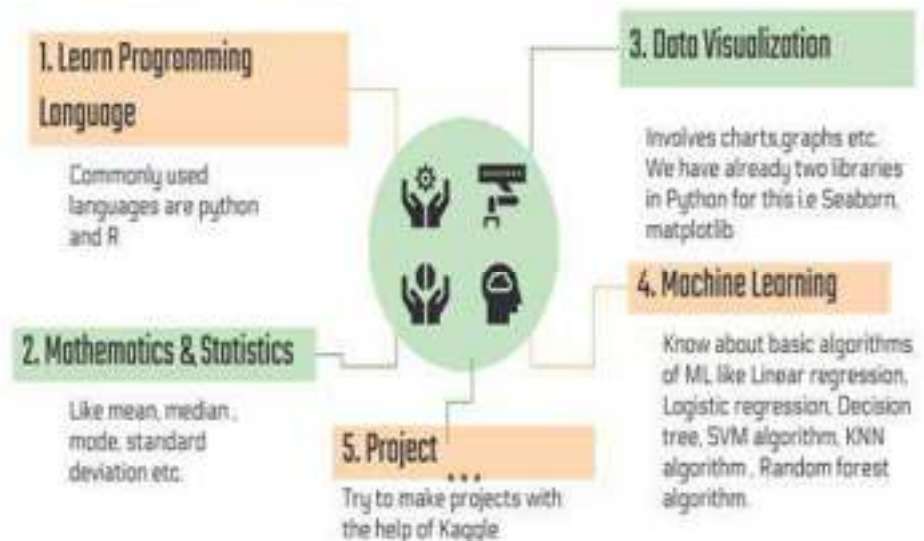
- **Ilmu Komputer (Computer Science)** – Pemrograman dan pengelolaan data.
- **Machine Learning** – Teknik pembelajaran mesin untuk analisis data.
- **Statistika (Statistics)** – Konsep dasar seperti mean, median, dan deviasi standar.
- **Visualisasi (Visualization)** – Kemampuan untuk menyajikan data secara visual.
- **Interaksi Manusia-Komputer (Human Computer Interaction)** – Kemampuan memahami bagaimana manusia berinteraksi dengan sistem berbasis data.




To collect, clean, integrate, analyze, visualize, interact with data to create data products.

# ROADMAP OF DATA SCIENCE

1. Belajar Bahasa Pemrograman – Python dan R adalah bahasa pemrograman yang umum digunakan dalam Data Science.
2. Matematika & Statistika – Memahami konsep dasar seperti mean, median, mode, dan deviasi standar.
3. Visualisasi Data – Menggunakan grafik dan diagram, dengan pustaka Python seperti Seaborn dan Matplotlib.
4. Pembelajaran Mesin (Machine Learning) – Memahami algoritma dasar seperti regresi linear, regresi logistik, pohon keputusan, algoritma SVM, algoritma KNN, dan algoritma random forest.
5. Proyek (Project) – Mempraktikkan Data Science dengan bantuan platform seperti Kaggle.





Start making small data science projects  
with the help of Kaggle

