



Home

About

Content

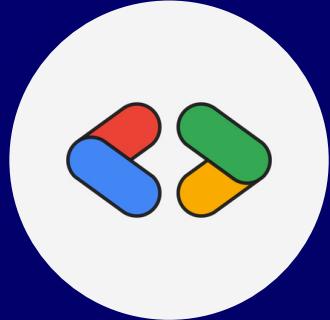
Others

Page 01

# INTRODUCTION TO DATA SCIENCE

---

By Jiryan Farokhi

[Home](#)[About](#)[Content](#)[Others](#)[Page 02](#)

# ABOUT ME

An Undergraduate Student from Mathematics ITS 2022 who has interest in Data Science and Artificial Intelligence

## Jiryan Farokhi

ML Curriculum Lead GDGoC ITS, Head of Data Science Avalon AI, Ex-Intern Data Scientists PT. LAPI DIVUSI, Freelancer Data and AI, Be finalist and participants in more of 5 Data Science Competitions

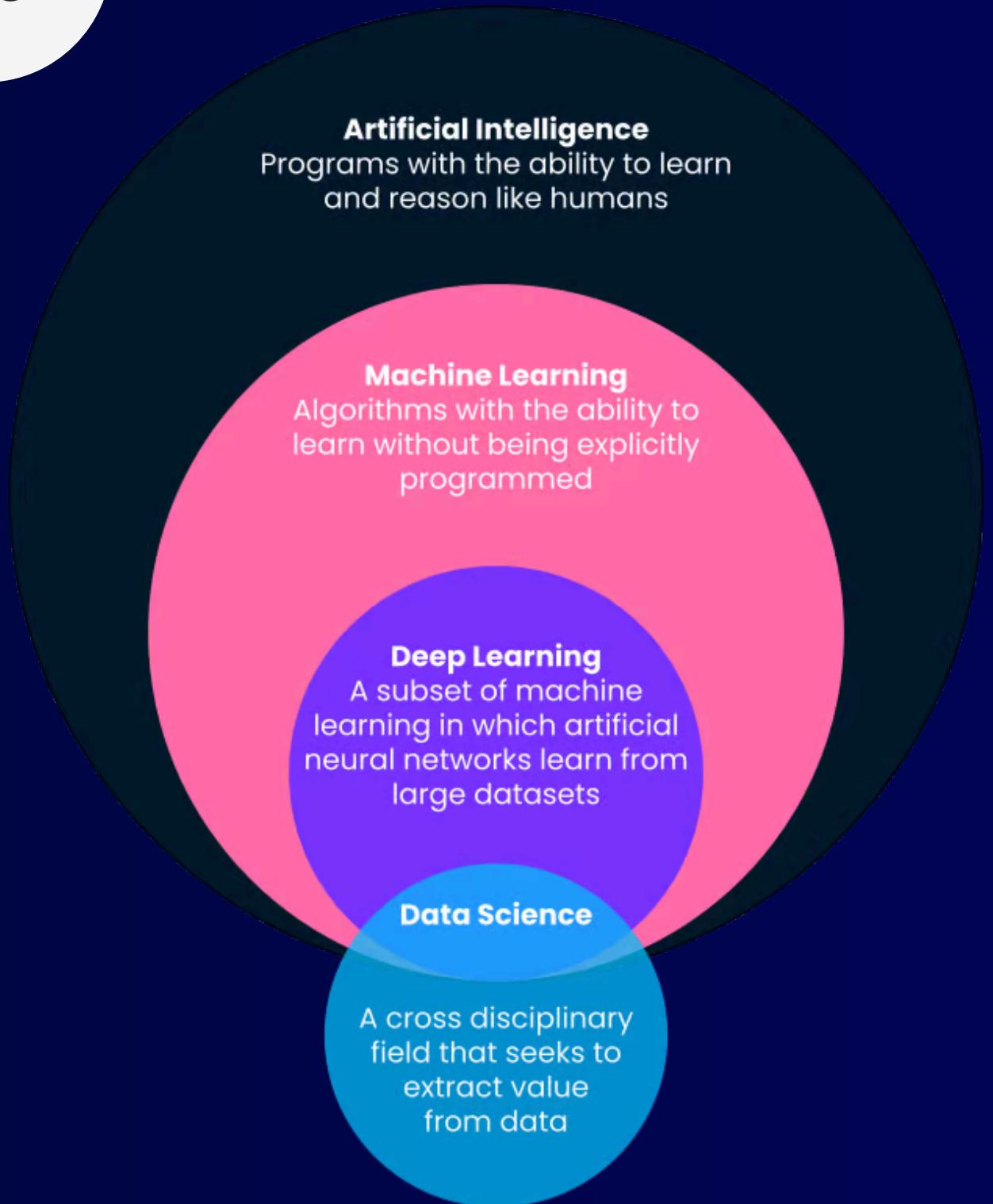
<https://github.com/JryFarr>

Jiryan Farokhi

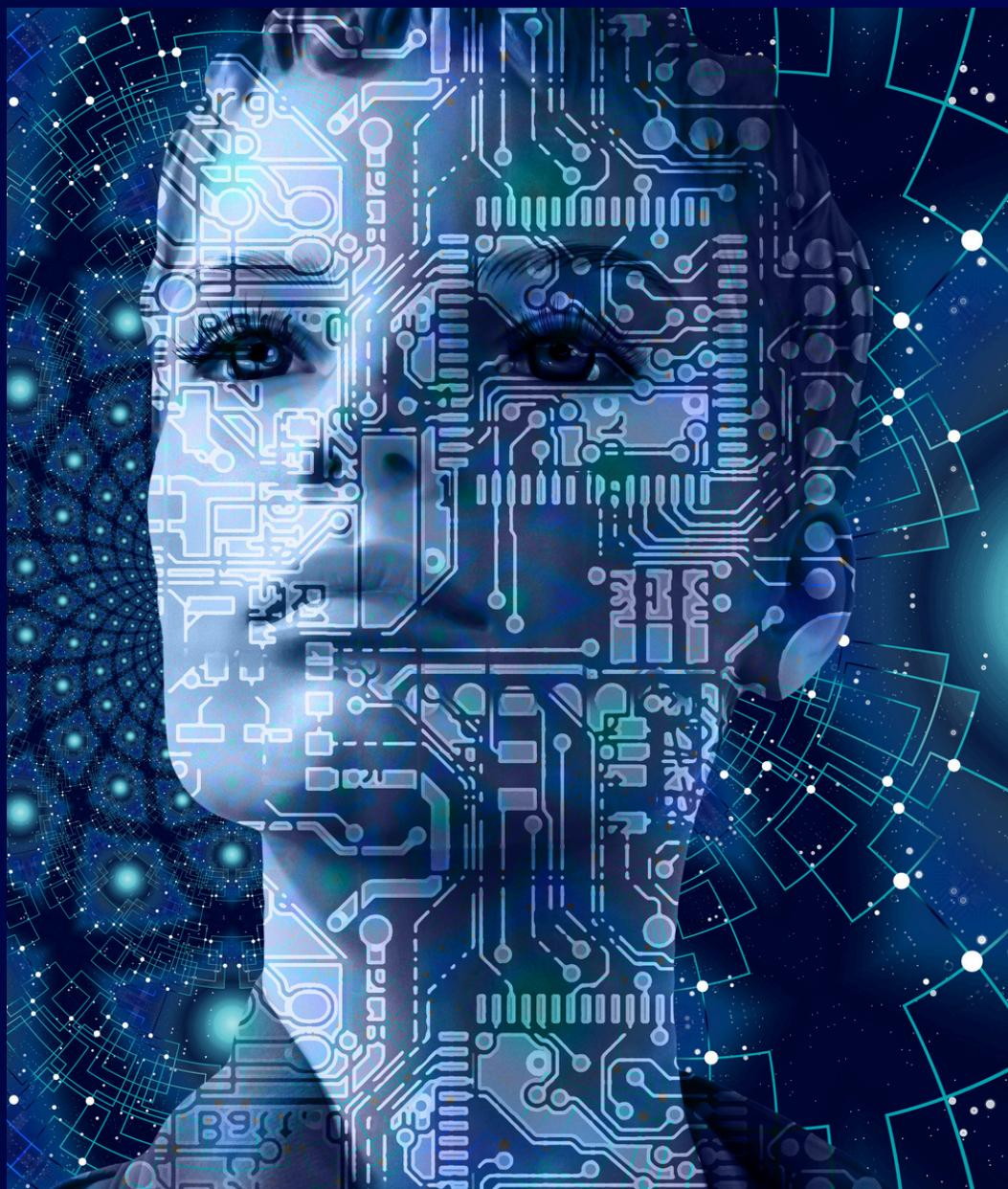




# KENAPA JUDULNYA DATA SCIENCE BUKAN AI ATAU MACHINE LEARNING ?



- Machine learning bagian dari data science
- Perlu waktu lebih lama
- Butuh resources lebih dalam tahap deployment (bisa berbayar)
- Butuh kemampuan matematika dan statistik yang kuat (bukan cuman menggunakan model tapi lebih dari itu)
- Alur path pembelajaran dari data ke AI disarankan
- Banyak algoritma yang perlu dipelajari
- **START FROM THE BIG PICTURE AND SMALL STEPS**



# DEFINITION

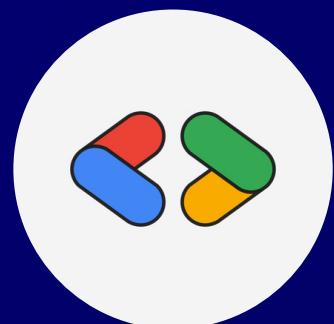
Data science is an interdisciplinary field that uses scientific methods, processes, algorithms, and systems to extract knowledge and insights from structured and unstructured data.

[Home](#)[About](#)[Content](#)[Others](#)

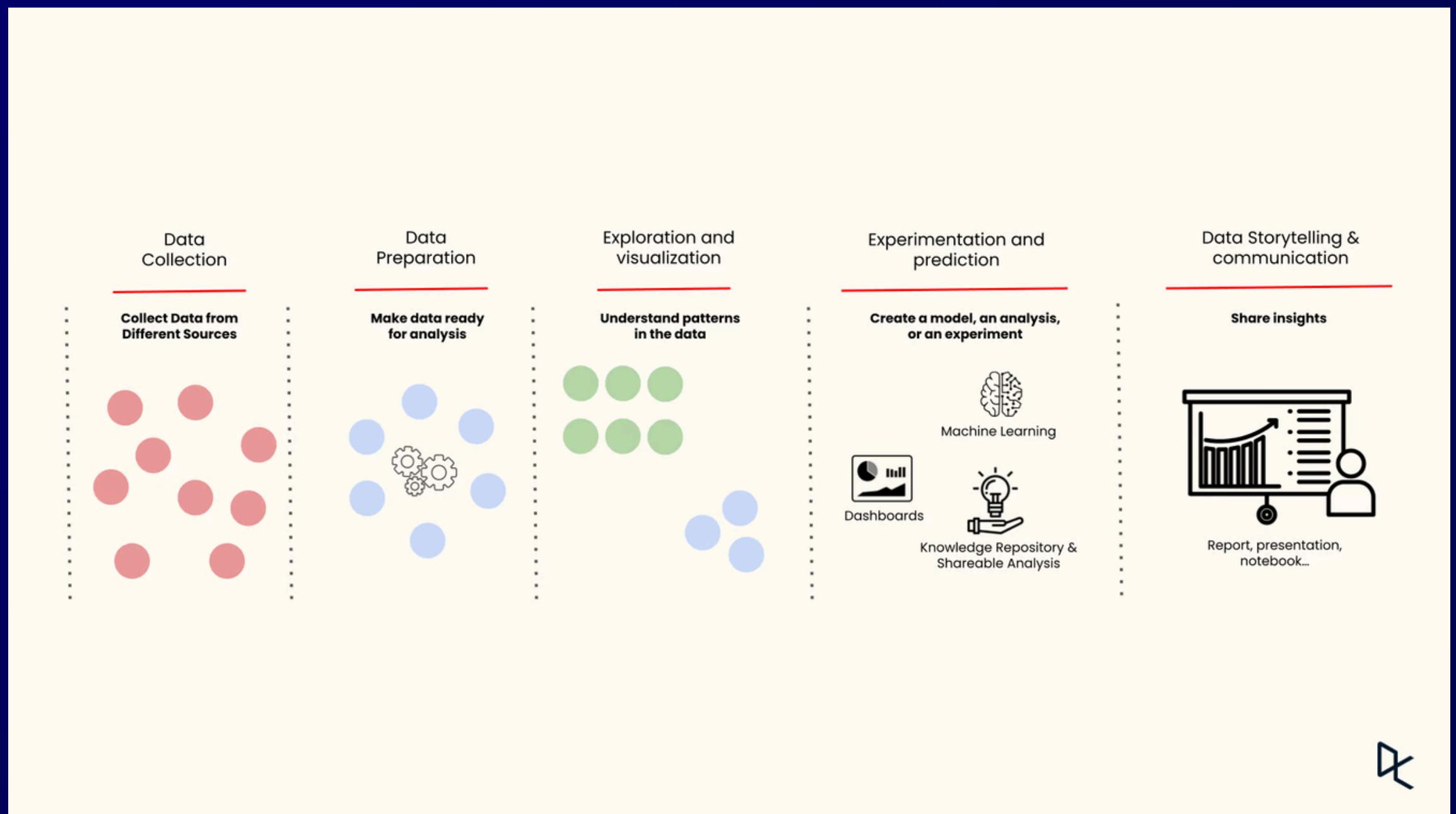
Page 06



# CØBA RENUNGIN GAMBARINI



# THE DATA SCIENCE LIFE CYCLE





# JOB IN DATA AND AI

## Data Engineers

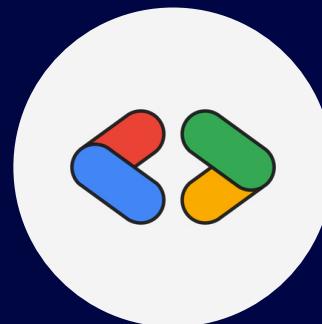
- Responsible for building and optimizing data pipelines, storing data in databases or cloud storage, and ensuring data is accessible to analysts and scientists.
- Skills :
  1. Programming : Python, Java, Scala, SQL
  2. Database : PostgreSQL, MySQL, MongoDB, Cassandra
  3. Teknologi Cloud & Big Data : AWS, Google Cloud, Azure, Hadoop, Spark, Kafka
  4. ETL (Extract, Transform, Load) : Apache NiFi, Airflow
  5. Pipeline Data Optimization
- Career path : Data Engineer → Senior Data Engineer → Engineering Manager

## Data Analyst

- Focuses on pulling, cleaning, analyzing, and visualizing data to derive actionable **business insights**. Uses tools like SQL, Excel, Tableau, and Power BI.
- Skills :
  1. Programming : SQL, Python (pandas, numpy)
  2. Data Analytics and Statistics: Excel, Google Sheets, statistic descriptive
  3. Data Visualization : Tableau, Power BI, Looker
  4. Business Understanding and Communication
- Career path : Data Analyst → Senior Data Analyst → Data Scientist/Business Intelligence Manager

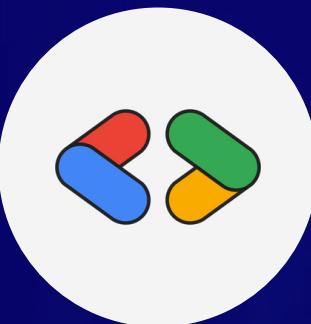
## Data Scientist

- Develops machine learning models, performs deep learning, and conducts A/B testing. Requires strong math, statistics, and programming skills.
- Skills
  1. Programming : Python (scikit-learn, TensorFlow, PyTorch), R
  2. Statistics and Mathematics : Kalkulus, probabilitas, regresi linier/logistik
  3. Machine Learning & AI: Supervised and unsupervised learning
  4. Big Data & Cloud Computing: Spark, AWS, Databricks
- Career path : Data Scientist → Senior Data Scientist → AI/ML Engineer / Head of Data Science

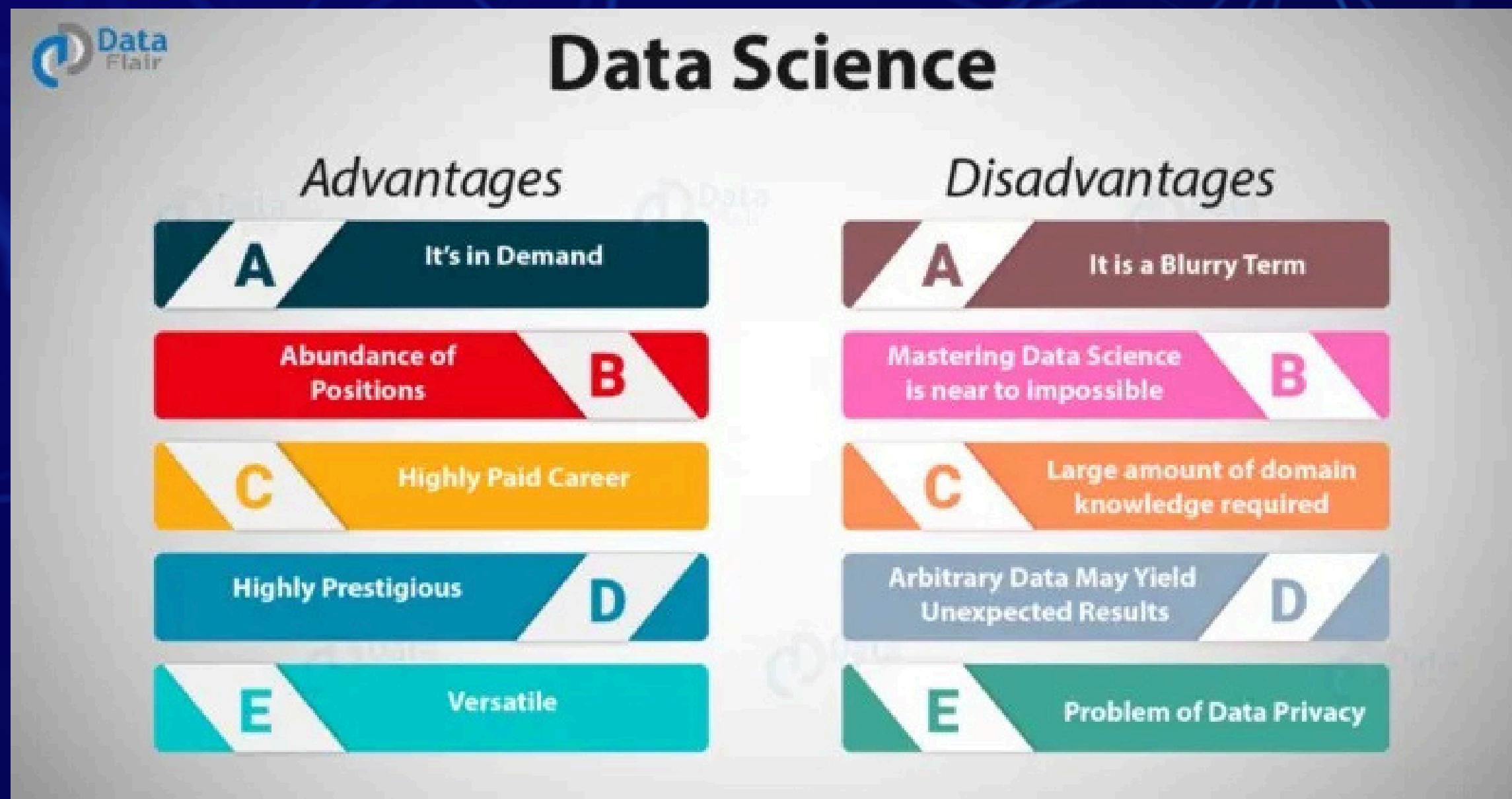


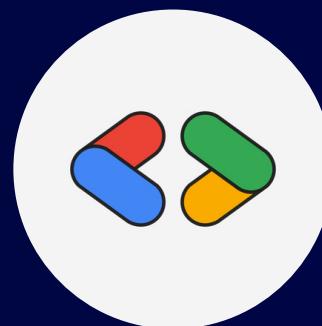
# AI ENGINEERS VS DATA SCIENTIST

- AI engineers and data scientists use data in similar ways
- AI modeling to teach machines (**software applications**) to make decisions faster than humans ever could.
- Mostly in every company, the tasks and data scientists are very similar.



# WHY IS DATA SCIENCE IS IMPORTANT





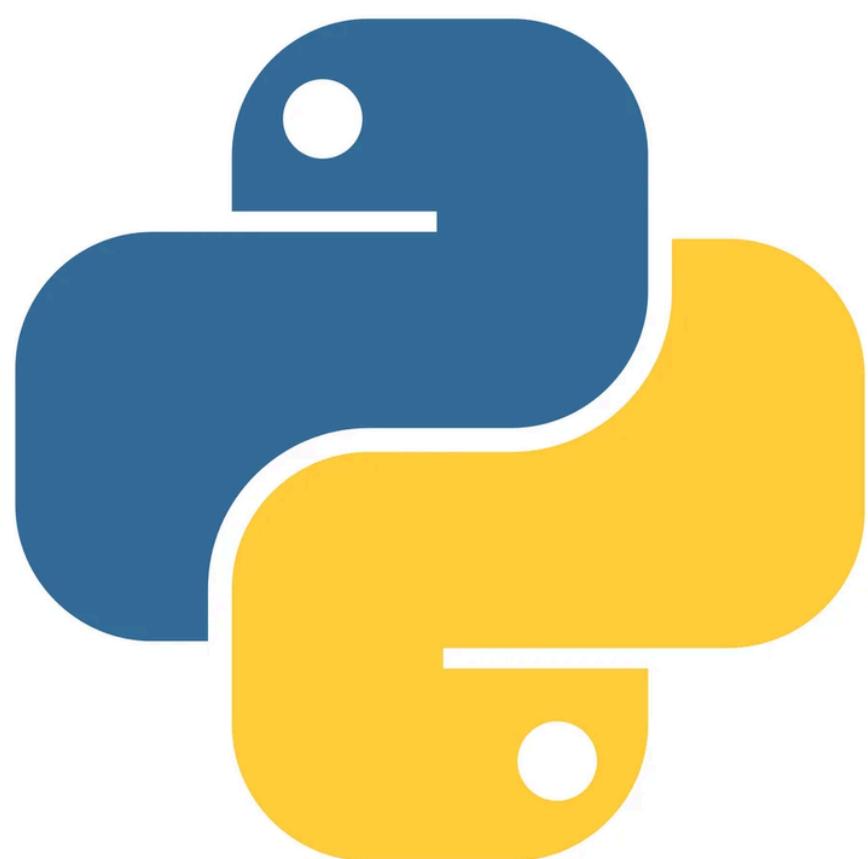
Home

About

**Content**

Others

**Page 11**



# FOCUS EN PYTHON



Home

About

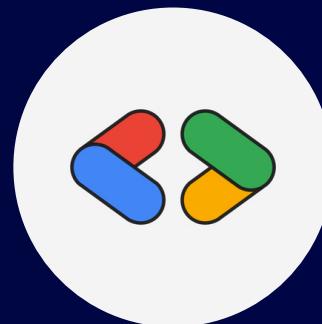
**Content**

Others

Page 12

# TOOLS DALAM PEMROGRAMAN





Home

About

Content

Others

Page 13

# POPULAR LIBRARY





Home

About

Content

Others

Page 14

# LET'S GET STARTED TO LEARN CØDE