

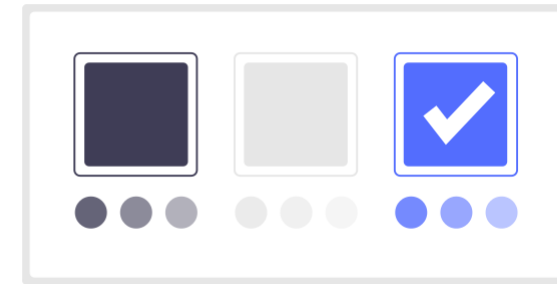


Get to Know **DATA SCIENCE**

Mohamad Irwan Afandi

Table of Content

1. What is data science
2. Why we need data science
3. Data science in daily life
4. The Essential skill needed by data scientist
5. Data science tools
6. Data science framework
7. Role in data world



What is **Data Science**

Data is a collection of facts, something that can be observations or just descriptions of things.

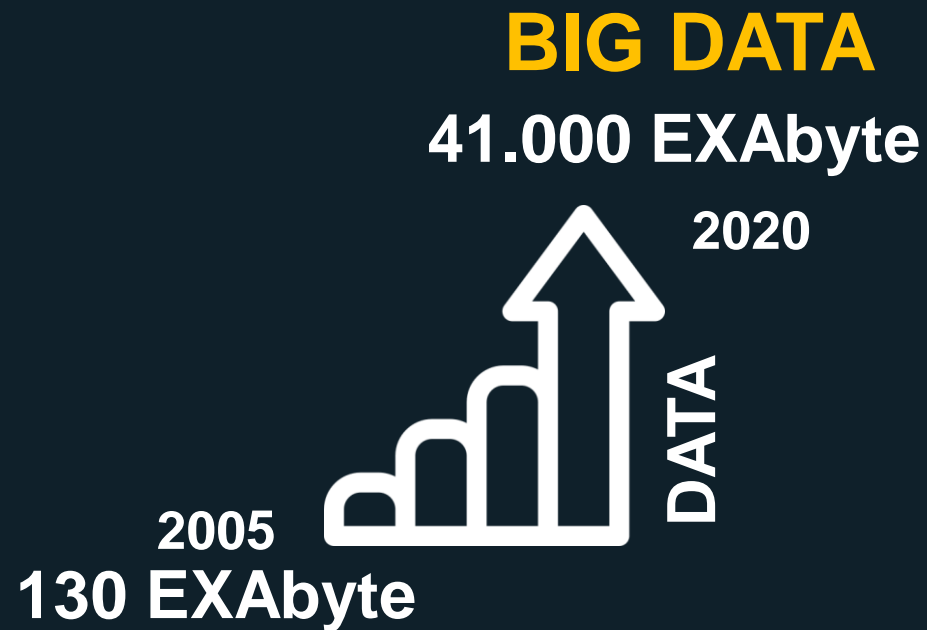
Science is knowledge to discover new things

Data Science is a process discover hidden patterns from the raw of data using science methodology (statistic, algorithm and machine learning principles).





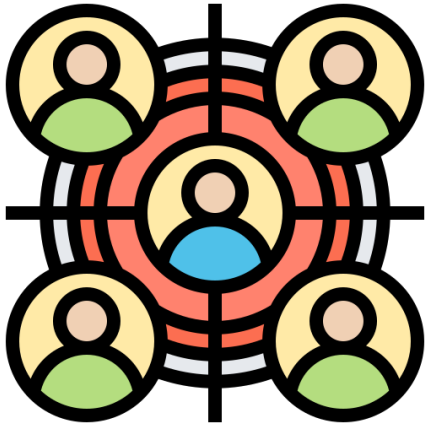
WHY WE NEED DATA SCIENCE



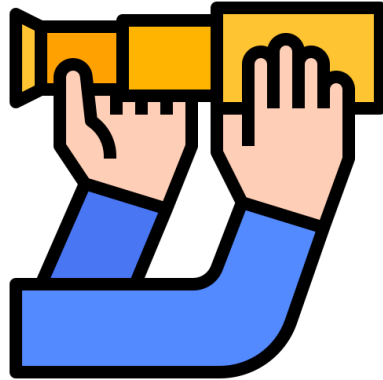
data that is not processed is called garbage

DATA SCIENCE

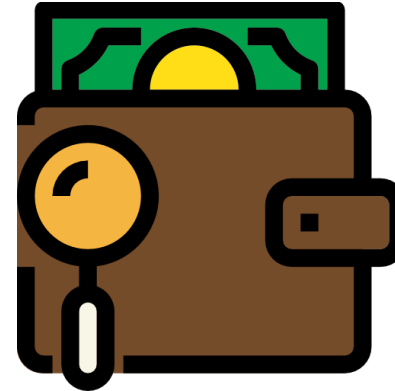
In Daily Life



Customer
Segmentation



Future
Prediction



Fraud
Detection

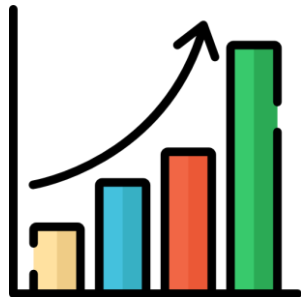


Market
Analysis

Safe driving car, customer churn, smart city, Etc.



PROGRAMMING



STATISTICS



BUSINESS

DATA SCIENCE Essential Skills

PREREQUISITES FOR DATA SCIENTIST

SOFTSKILL



Curiosity

High motivation to learn and never give up



Common sense

Know what we have to do with the problems



Communication Skill

How we present the results we get

HARDSKILL



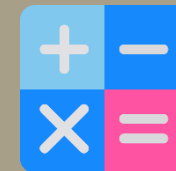
Machine Learning



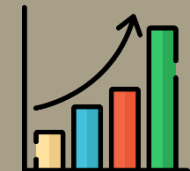
Programming



Database

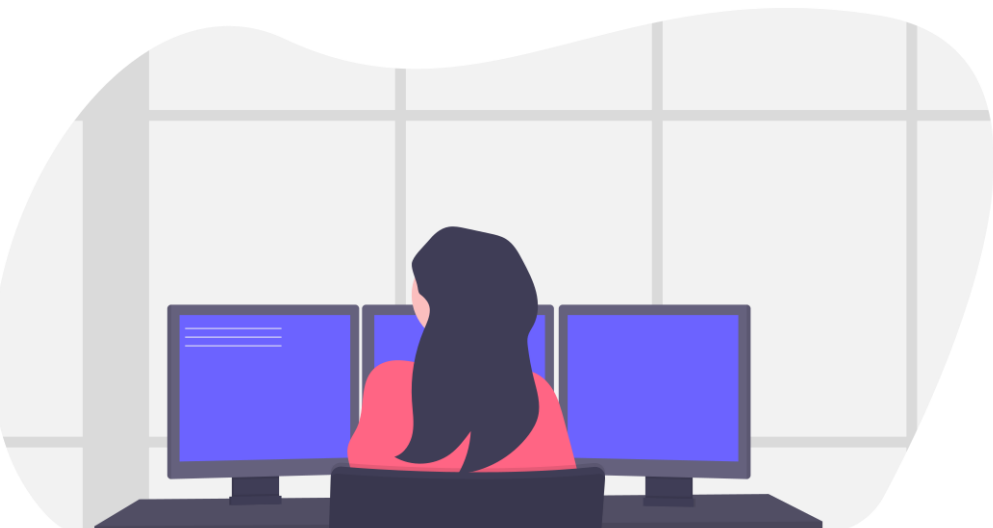


Math Modeling



Statistics

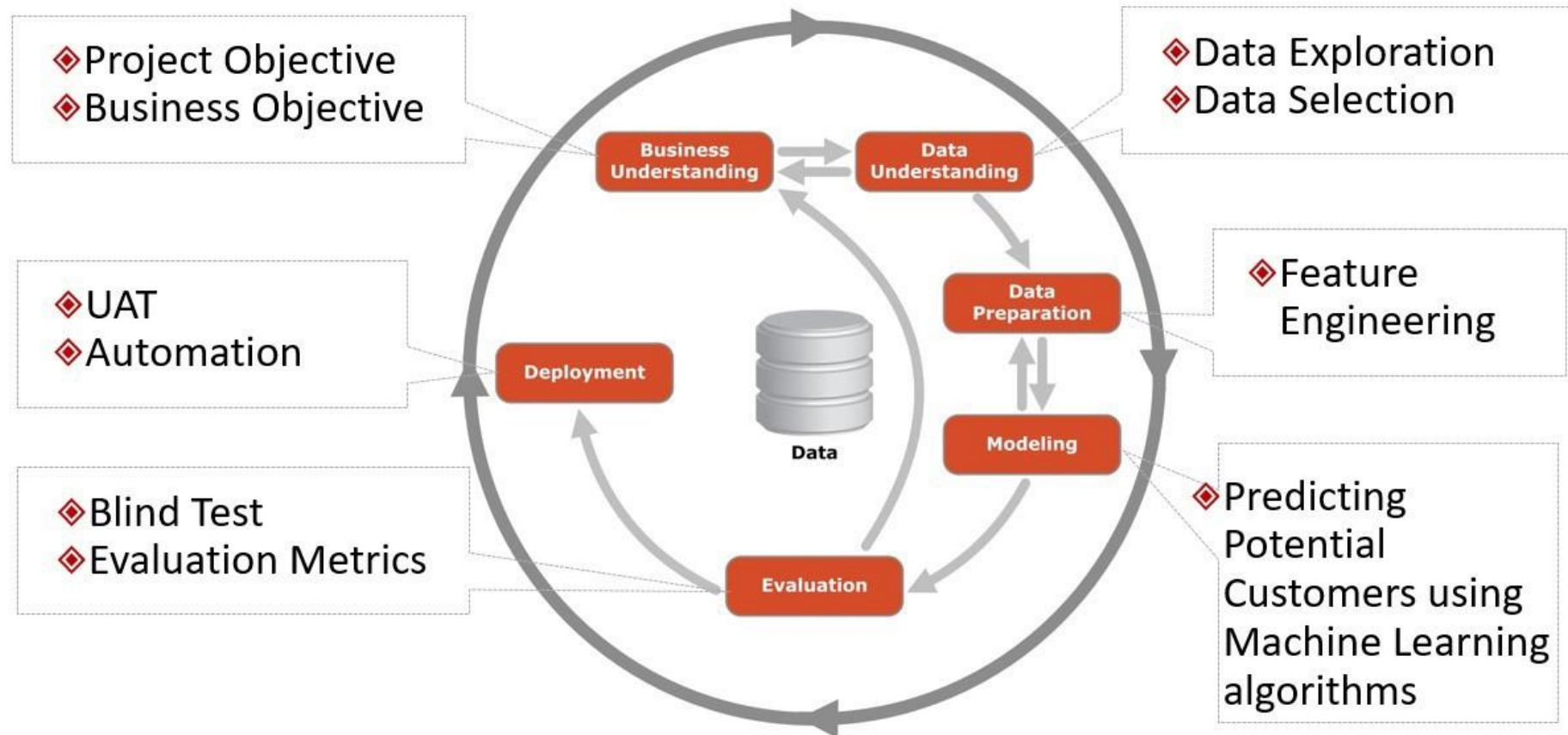
DATA SCIENTIST TOOLS



DATA SCIENCE

Framework CRISP-DM

Cross Industry Standard Process for data mining



BORROWER SCORING on Agriculture Sector

01



- Determine borrowers who deserve to be funded is very difficult
- Create recommended system to help credit scoring department

02



- Collect the data from previous data (RDBM), Weblog (NoSQL), or Third party like AFPI (JSon)
- Identify data problems
- Discover new insight

03



- Dealing with missing value, outlier, redundant data, different format, imbalance data, standardization, etc

04



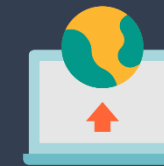
- Determine the methodology (machine learning or logic programming)
- Determine important factor
- Build the model

05



- Evaluate the model using some matrix accuracy to see the model performance.
- Determine whether go to the next step or return to previous

06



- Deploy the model in our software (website or mobile apps)

[2]

Data Engineer

Prepare the infrastructure for DS like database, server, create data processing pipeline, etc

[3]

Machine Learning Engineer

Focus on implementing the model made by DS

[4]

Data Scientist

Define problem, understanding the data, data exploration, modeling and evaluate

[1]

Data Analyst

Focus on Analyzing the data, get the insight, and implement on business.

[4]

Business Intelligence

Define business problem/goal, transfer data into business insight, and deliver to user





Thank you ...