



\* Part of this slide is modified from a slide of Prof.Natawut

## Course Overview

*2110446 Data Science and Data Engineering*

Asst. Prof. Peerapon Vateekul, Ph.D.

Department of Computer Engineering,  
Faculty of Engineering, Chulalongkorn University

[Peerapon.v@chula.ac.th](mailto:Peerapon.v@chula.ac.th)

[www.cp.eng.chula.ac.th/~peerapon/](http://www.cp.eng.chula.ac.th/~peerapon/)

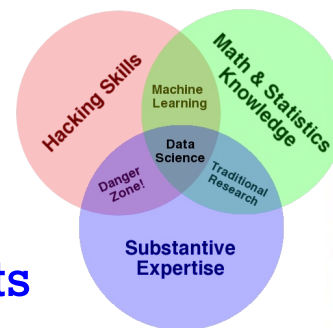
# What is this class about? (recap)

## Dr. Virote

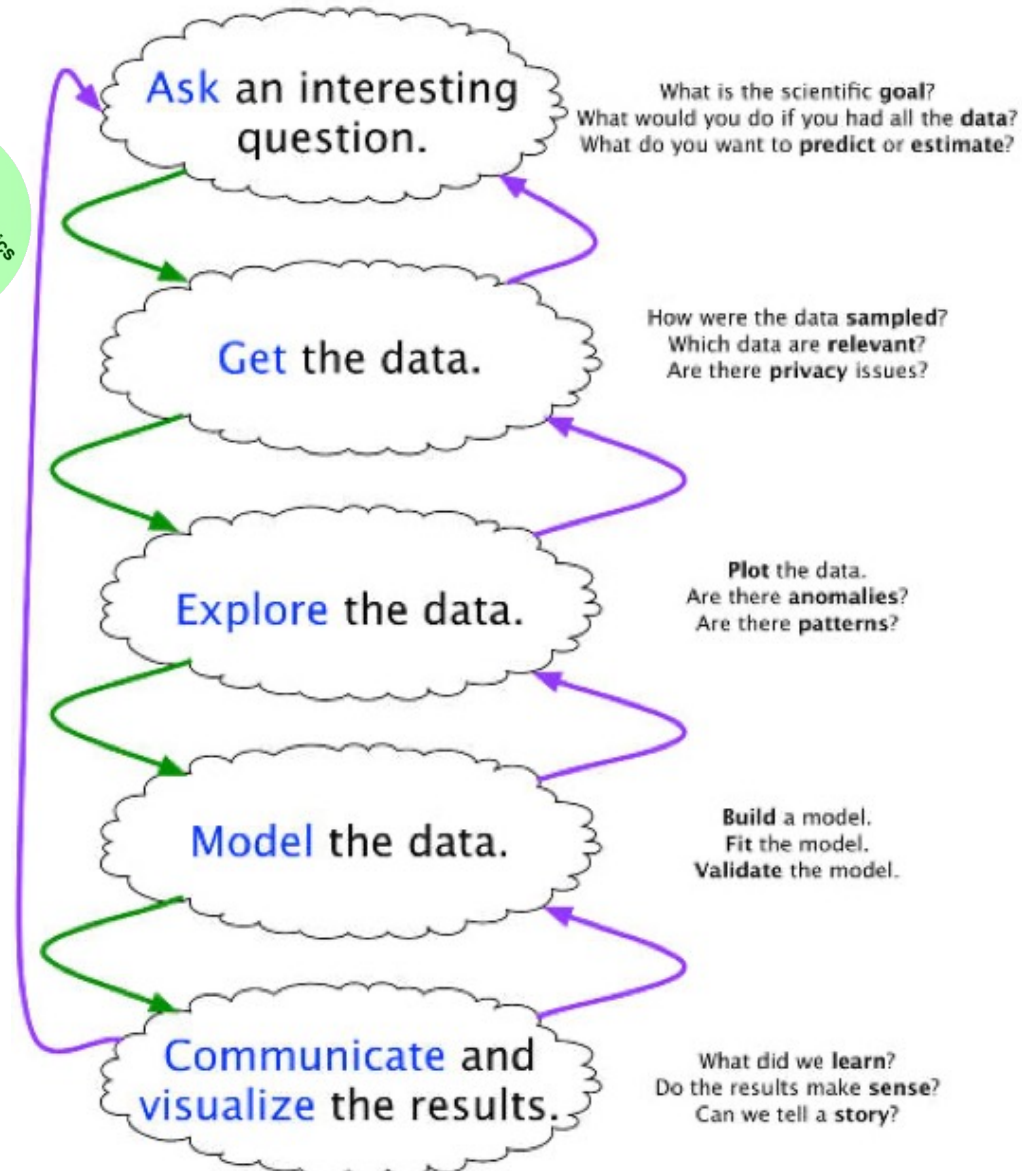
1. Transform data into **valuable insights**
2. Transform data into **data products**
3. Transform data into **interesting stories**

## Aj. Natawut

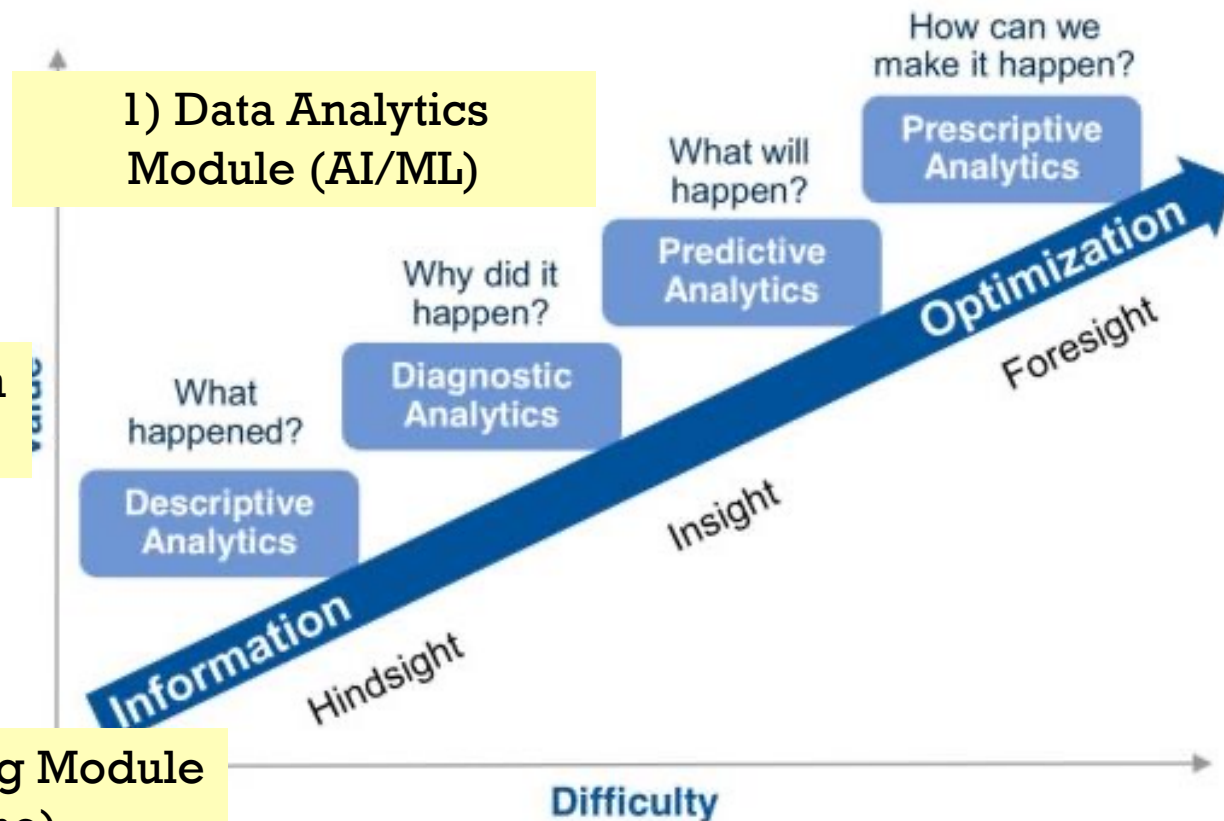
1. Measurement (**decision**)
2. Insights (**knowledge**)
3. Data Products (**Innovation, Intelligent**)



## The Data Science Process



# What is this class about? 3 Modules



1) Data Analytics  
Module (AI/ML)

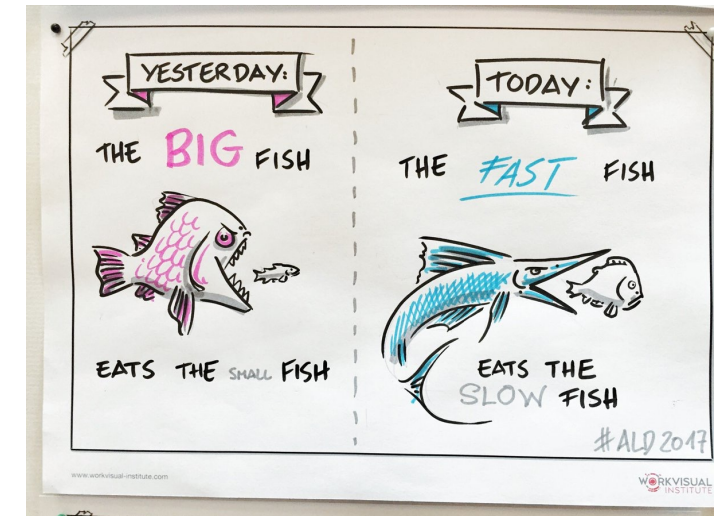
3) Data Visualization  
Module

2) Data Engineering Module  
(Data Pipeline)

Gartner's Top Strategic Technology Trends for 2023

Optimize	Scale	Pioneer
Digital Immune System	Industry Cloud Platforms	Superapps
Applied Observability	Platform Engineering	Adaptive AI
AI TRISM	Wireless-Value Realization	Metaverse
Sustainable Technology		

Gartner



#	Tue (1PM-4PM)	Contents	Instructor	Module
1	10-Jan-23	Introduction (Definition & Use cases) & Pandas	Aj.Peerapon	Data Science
2	17-Jan-23	Data preparation + EDA + Data Science Process + Statistical Analysis (Hypothesis Testing)	Aj.Peerapon	Data Science
3	24-Jan-23	Traditional ML (1)	Aj.Peerapon	Data Science
4	31-Jan-23	Traditional ML (2)	Aj.Peerapon	Data Science
5	7-Feb-23	DL (1)	Aj.Peerapon	Data Science
6	14-Feb-23	DL (2)	Aj.Peerapon	Data Science
7	21-Feb-23	Advanced topics & Guest speakers * Modern ML techniques (Transformer) * Sharing data scientist experience (in Thailand & Abroad) [2-3 guest speakers]	Aj.Peerapon	Data Science
8	28-Feb-23	Big data architecture + data storage	Aj.Natawut	Big Data Eng.
	7-Mar-23	Midterm Exam Week (7-10 Mar)		
9	14-Mar-23	Web scraping	Aj.Natawut	Big Data Eng.
10	21-Mar-23	Data ingestion	Aj.Natawut	Big Data Eng.
11	28-Mar-23	Big data processing (Spark)	Aj.Natawut	Big Data Eng.
12	4-Apr-23	Orchestration and serving (Airflow, MLflow, Seldon-core, FastAPI)	Aj.Natawut	Big Data Eng.
13	11-Apr-23	Data visualization	Aj.Veera	Data Viz.
14	18-Apr-23	Python visualization	Aj.Veera	Data Viz.
15	25-Apr-23	Graph analysis & spatial analysis	Aj.Veera	Data Viz.
16	2-May-23	Final Exam		
	9-May-23	Final Exam Week (8-19 May) Project presentation		



# Evaluation

■ Module1 Assignment (data analytics)	15%
■ Module2 Assignment (data engineering)	15%
■ Module3 Assignment (data visualization)	10%
■ Midterm Exam (Kaggle)	15%
■ Project	15%
■ Final Exam	30% (Lab Test)



# Class communication

- myCourseVille: “Avatar2”
- Discord: <https://discord.gg/eU93mtgYWf>



# Remarks

- This course assumes that:
- It is just a fundamental course to perform a data science task properly. Then, you can choose your own specialty in AI domains.
- It aims to utilize all tools and algorithms **without** theoretical proofs.
- All students should have a background in Python programming.

- There are many courses related to AI/ML/DS?
- AI
- Data Mining
- Machine Learning
- Neural Networks
- ASR
- NLP
- Computer Vision
- Etc.

ปีการศึกษา 2565/ภาคการศึกษาปลาย						
ทวิภาค						
2110446 DATA SCI/ENG						
วิทยาศาสตร์ข้อมูลและวิศวกรรมข้อมูล						
DATA SCIENCE AND DATA ENGINEERING						
คณะวิศวกรรมศาสตร์ ( ภาควิชาวิศวกรรมคอมพิวเตอร์ )						
3.0 CREDIT HOURS = LECT 3.0 CR						
(LECT 3.0 HR)						
เงื่อนไขรายวิชา : -						
วันสอบกลางภาค : TDF (รอประกาศ) วันสอบปลายภาค : TDF (รอประกาศ)						
ตอนเรียน	วิธีสอน	วัน-เวลาเรียน	อาคาร	ห้อง	ผู้สอน	หมายเหตุ
21	LECT	TU 13:00-16:00	EN100	201A	NNP,VMS,PVK	APPROVED CP
						จำนวนนิสิต Regis/Max
						100/100



Any questions? 😊