



KMITL
FIGHT



คณะเทคโนโลยีสารสนเทศ
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

CUBE Analytics
CONSULTING

Master of Science Program in
Artificial Intelligence for Business Analytics

Data Science

Week 1: Introduction

Eakasit Pacharawongsakda, Ph.D.

eakasit@datacubeth.ai

Co-Founder and Data Science Team Lead

Cube Analytics Consulting Co., Ltd.

datacube ||| **CUBE Analytics**
CONSULTING

Copyright © 2025 Cube Analytics Consulting Co., Ltd.

All rights reserved. All course materials, slides and notes are protected by Cube Analytics Consulting Co., Ltd. Course materials are copyrighted according to the owner of the content or the principal body. You may take notes on your own, however, you are not allowed to reproduce, distribute, upload or display any of the course materials in any way – whether or not fees is charged – without express written consent from Cube Analytics Consulting Co., Ltd.

Introduction

สำหรับการบรรยายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL

ตัวอย่างการใช้ Big Data Analytics ในชีวิตประจำวัน

- ลักษณะของข้อมูล
- การจัดเก็บข้อมูล (Data Management)
 - ฐานข้อมูล (Database)
 - คลังข้อมูล (Data Warehouse)
- การจัดการข้อมูลขนาดใหญ่ (Big Data)
- ตัวอย่างการประยุกต์ใช้การวิเคราะห์ข้อมูลด้วยเทคโนโลยี Data Science

ในหนึ่งวันทำงาน

เวลา 07:00 น. ออกเดินทางไปทำงาน



source: <http://pad1.whstatic.com/images/thumb/a/aa/Reduce-Anxiety-About-Driving-if-You're-a-Teenager-Step-5-aid196018-728px-Reduce-Anxiety-About-Driving-if-You're-a-Teenager-Step-5-version-2.jpg>

 CUBE Analytics
CONSULTING

เวลา 07:45 น. ยังคงติดอยู่บนถนน



source: <http://www.elparkid.com/images/259/research-and-report-writing-9-22-129-11E2-q2r0wg-clipart.jpg>

 CUBE Analytics
CONSULTING

เวลา 08:00 น. เจ้านายโทรศัพท์เข้ามาดามงาน



source: https://d1ai9qtk9p41kl.cloudfront.net/assets/mc/psuderman/2011_07/text-drive.png

**CUBE Analytics
CONSULTING**

เวลา 08:05 น. ขับรถไปชนกับคันอื่น



**CUBE Analytics
CONSULTING**

เวลา 10:00 น. ถึงที่ทำงานและทำงานต่อไป



source: <http://stuffpoint.com/anime-and-manga/image/285181-anime-and-manga-girl-working-in-the-computer/>

**CUBE Analytics
CONSULTING**

เวลา 18:00 น. แวะซื้อของกลับบ้าน



**CUBE Analytics
CONSULTING**

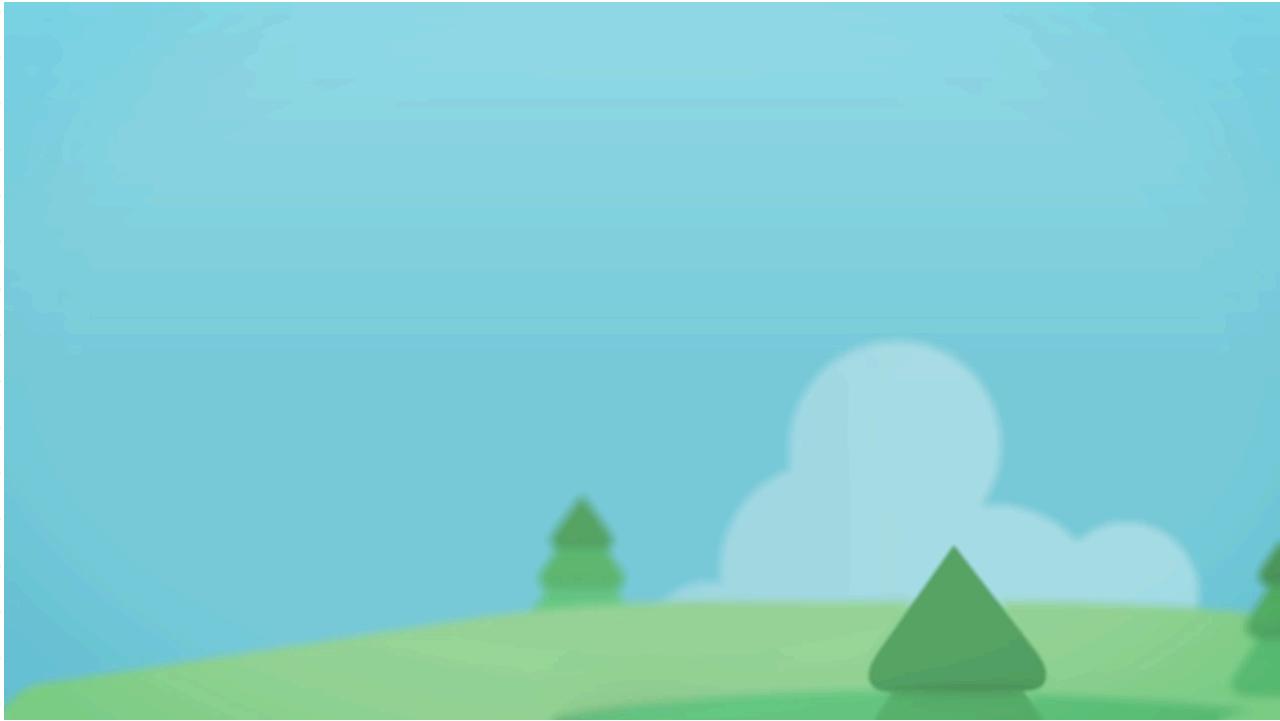
เวลา 20:00 น. กลับถึงบ้านและอยู่คนเดียว



**ในหนึ่งวันทำงานกับ
เทคโนโลยีข้อมูลขนาดใหญ่ (Big Data)**

ຮະບບນໍາທາງ

ສໍາหารັບກາຮຽນຮ່ວມມືດ ມະຫາວິທະຍາ ໃນຄະນະກູດ M: Sc. in Artificial Intelligence for Business Analytics, KMITL



Copyright © 2025 Cube Analytics Consulting Co., Ltd.

<http://facebook.com/datacube.th>

13

ຮະບບນໍາທາງ

ສໍາหารັບກາຮຽນຮ່ວມມືດ ມະຫາວິທະຍາ ໃນຄະນະກູດ M: Sc. in Artificial Intelligence for Business Analytics, KMITL

- ແອພພລິເຄຊັນ Waze

waze
Guided Tour

Copyright © 2025 Cube Analytics Consulting Co., Ltd.

<http://facebook.com/datacube.th>

14

รถที่ไม่ต้องมีคนขับ (self driving car)

สำหรับการบรรยายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL

- Waymo (Google self-driving car)



แพงไข่อัจฉริยะ

สำหรับการบรรยายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL

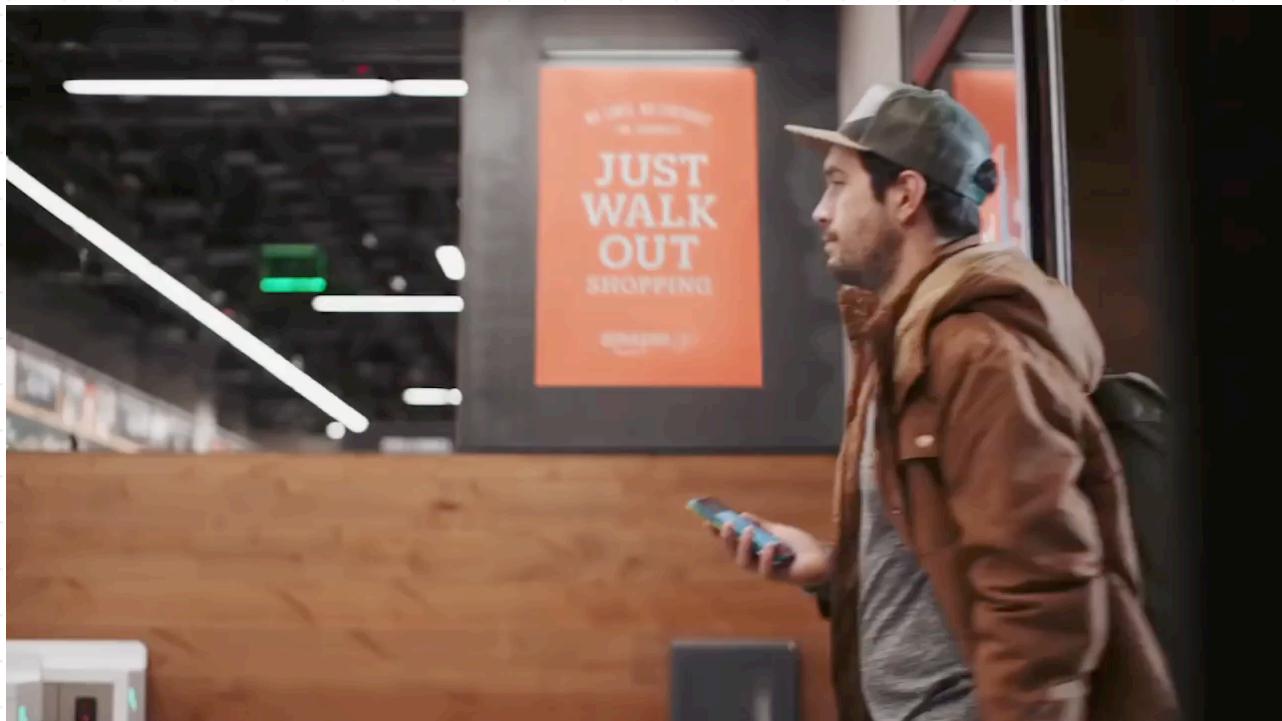
- Egg Minder



ร้านค้าที่ไม่ต้องรอคิว

สำหรับการบรรยายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL

- Amazon Go



เทคโนโลยีที่กำลังใช้เวตประจำวันสะดวกขึ้น

สำหรับการบรรยายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL



Business without analytics

สำหรับการบรรยายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL



Business without analytics

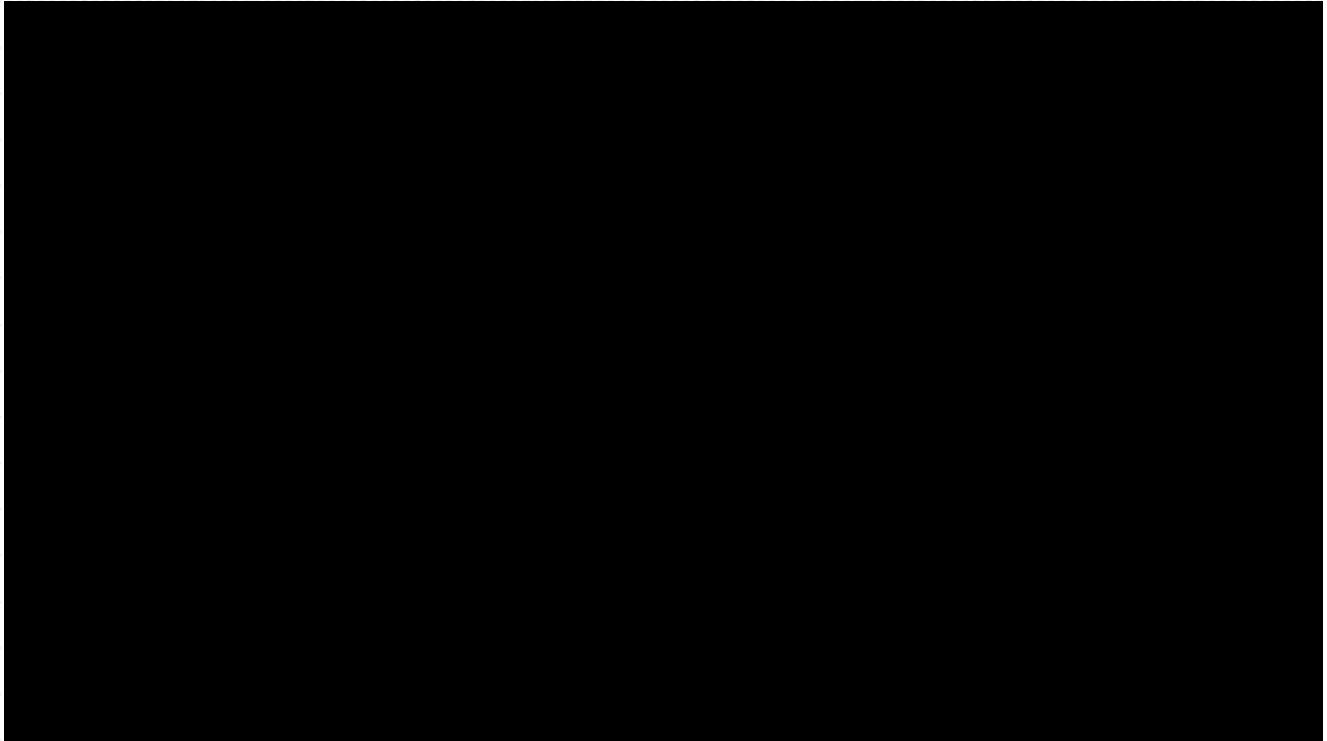
สำหรับการบรรยายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL



image source: <http://www.oknation.net/blog/print.php?id=434843>

Business with analytics

สำหรับการสูงสุด M: Sc. in Artificial Intelligence for Business Analytics, KMITL



source: <https://www.youtube.com/watch?v=7tAgbn9kpY>

<http://facebook.com/datacube.th>

21

Where does data come from?

สำหรับการสูงสุด M: Sc. in Artificial Intelligence for Business Analytics, KMITL



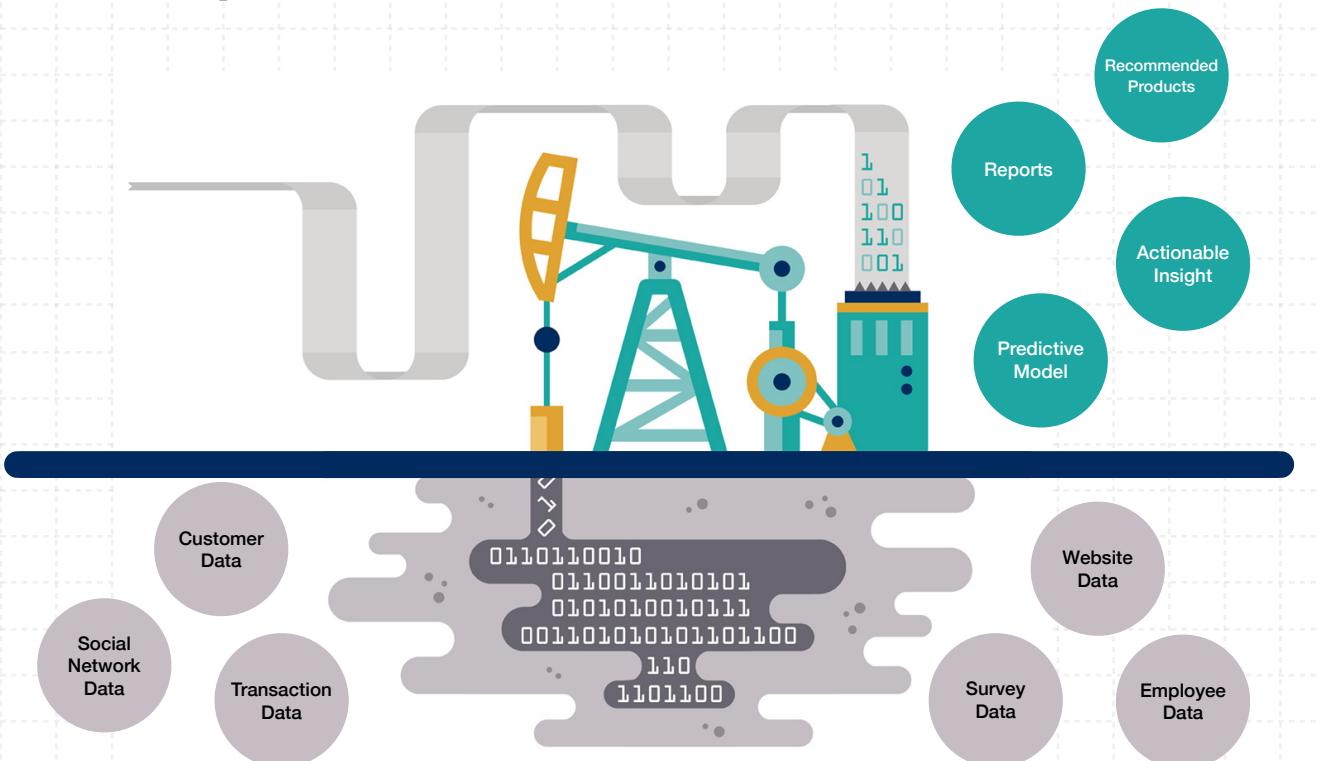
source: https://www.youtube.com/watch?v=Y_JlkzzhAgw

<http://facebook.com/datacube.th>

22

Data is a new oil

- สำหรับการบรรยายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL
- “เมื่อข้อมูลมีค่าดั่งน้ำมัน”



Source: <https://www.raconteur.net/technology/drilling-for-new-oil-of-big-data>

<http://facebook.com/datacube.th>

23

Introduction

- สำหรับการบรรยายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL
- ตัวอย่างการใช้ Big Data Analytics ในชีวิตประจำวัน

- ลักษณะของข้อมูล
- การจัดเก็บข้อมูล (Data Management)
 - ฐานข้อมูล (Database)
 - คลังข้อมูล (Data Warehouse)
- การจัดการข้อมูลขนาดใหญ่ (Big Data)
- ตัวอย่างการประยุกต์ใช้การวิเคราะห์ข้อมูลด้วยเทคโนโลยี Data Science

Data

- สำหรับการบรรยายในหลักสูตร M. Sc. in Artificial Intelligence for Business Analytics, KMITL
- แบ่งตาม **ลักษณะ** ของข้อมูล

- ข้อมูลเชิง **ประชากรศาสตร์** (Demographic Data)
- ข้อมูลเชิง **พฤติกรรม** (Behavioural Data)
- ข้อมูลเชิง **เครือข่าย** (Network Data)
- แบ่งตาม **รูปแบบ** ของข้อมูล
- แบ่งตาม **แหล่งที่อยู่** ของข้อมูล

Data

- สำหรับการบรรยายในหลักสูตร M. Sc. in Artificial Intelligence for Business Analytics, KMITL
- ข้อมูลเชิง **ประชากรศาสตร์**

(Demographic Data)

- อายุ
- เพศ
- ระดับการศึกษา
- อาชีพ
- รายได้
- สถานภาพการแต่งงาน
- จำนวนบุตร



| ชื่อ |
|------------------|
| ปิติ |
| นามสกุล |
| ยินดี |
| เพศ |
| ชาย |
| วัน เดือน ปีเกิด |
| 28 สิงหาคม 2524 |
| เบอร์โทรศัพท์ |
| 0891234567 |
| อาชีพ |
| วิศวกร |
| รายได้ (บาท) |
| 50,000 |

| ชื่อ |
|------------------|
| มานี |
| นามสกุล |
| มีความลุข |
| เพศ |
| หญิง |
| วัน เดือน ปีเกิด |
| 22 มิถุนายน 2524 |
| เบอร์โทรศัพท์ |
| 0898765432 |
| อาชีพ |
| ผู้จัดการฝ่ายขาย |
| รายได้ (บาท) |
| 70,000 |

Data

สำหรับการบรรยายในหลักสูตร M. Sc. in Artificial Intelligence for Business Analytics, KMITL

- ข้อมูลเชิงพฤติกรรม (Behavioural Data)

- ข้อมูลพฤติกรรมการใช้งานต่างๆ จากบุคคลหรือสิ่งของ (Internet of Things)



Data

สำหรับการบรรยายในหลักสูตร M. Sc. in Artificial Intelligence for Business Analytics, KMITL

TYPICAL VARIABLES HELD AT CUSTOMER LEVEL

| Retailer | Mobile Phone Operator | Retail Bank |
|--|--|---|
| Transactions (baskets) <ul style="list-style-type: none"> Date of last transaction Value of transaction Number of transactions <ul style="list-style-type: none"> - by day/time - by channel (store/online/telephone) Purchases <ul style="list-style-type: none"> Number of items purchased <ul style="list-style-type: none"> - by product groups Amount paid <ul style="list-style-type: none"> - by product groups - by method of payment Discount amount Loyalty point earned/redeemed | Calling behaviour <ul style="list-style-type: none"> Numbers and durations of call by type (voice/SMS/MMS) <ul style="list-style-type: none"> - by outbound/inbound - by destination type (mobile/land) - by source type (mobile/land) Usage by type by outbound/inbound <ul style="list-style-type: none"> - by day (weekday/weekend) - by time of day Revenue <ul style="list-style-type: none"> - by outbound/inbound Data <ul style="list-style-type: none"> Data usage, app usage, websites usage | Products held <ul style="list-style-type: none"> counts by product types balance by product types Transactions <ul style="list-style-type: none"> by product types numbers and value by types of transactions and split by channels (branch/ATM/online/telephone) numbers and values Credit <ul style="list-style-type: none"> Credit scores, days, overdrawn Delinquency indicators |

Data

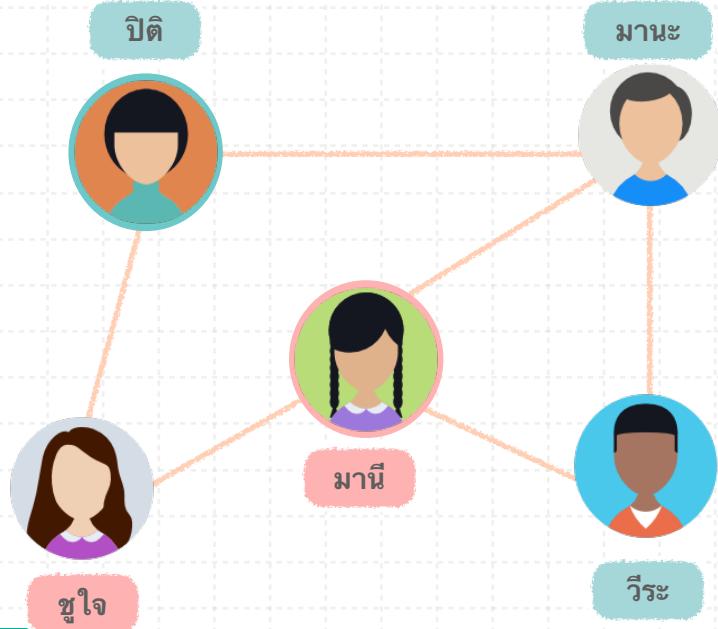
สำหรับการบรรยายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL

- ข้อมูลเชิงเครือข่าย (Network Data)

- ข้อมูลที่มีการเชื่อมโยงกันระหว่างบุคคล หรือสิ่งต่างๆ (เช่น sensor, โทรศัพท์มือถือ)

- การวิเคราะห์

- หา ความสัมพันธ์
- หา influencer
- หา community



Data

สำหรับการบรรยายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL

- แบ่งตามลักษณะของข้อมูล

- แบ่งตามรูปแบบของข้อมูล

- ข้อมูลแบบมีโครงสร้าง (Structured Data)
- ข้อมูลแบบไม่มีโครงสร้าง (Unstructured Data)
- ข้อมูลแบบกึ่งโครงสร้าง (Semi-structured Data)

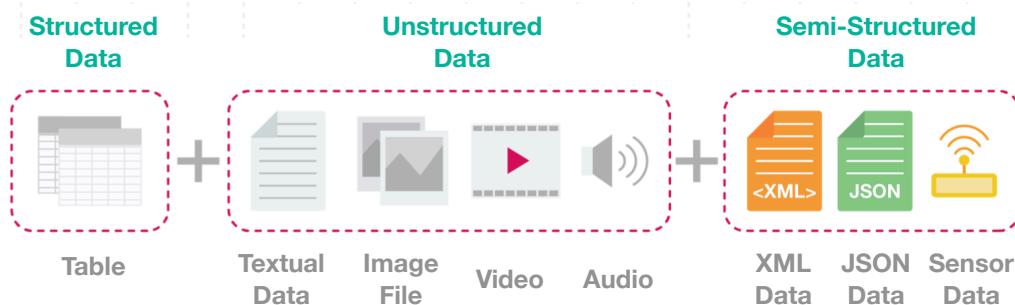
- แบ่งตามแหล่งที่อยู่ของข้อมูล

Data

สำหรับการบรรยายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL

- แบ่งตามรูปแบบของข้อมูล

- ข้อมูลแบบมีโครงสร้าง (Structured Data) เช่น ข้อมูลในตาราง
- ข้อมูลแบบไม่มีโครงสร้าง (Unstructured Data) เช่น ข้อความ รูปภาพ
- ข้อมูลแบบกึ่งโครงสร้าง (Semi-structured Data) เช่นไฟล์ JSON,ไฟล์ XML



Data

สำหรับการบรรยายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL

- แบ่งตามรูปแบบของข้อมูล

- ข้อมูลแบบมีโครงสร้าง (Structured Data) เช่น ข้อมูลที่เก็บในรูปแบบตาราง

| customer_id | age | gender | region | income | married | children | car | response |
|-------------|-----|--------|------------|---------|---------|----------|-----|----------|
| ID12101 | 48 | FEMALE | INNER_CITY | 17546.0 | NO | 1 | NO | NO |
| ID12102 | 40 | MALE | TOWN | 30085.1 | YES | 3 | YES | NO |
| ID12103 | 51 | FEMALE | INNER_CITY | 16575.4 | YES | 0 | YES | YES |
| ID12104 | 23 | FEMALE | TOWN | 20375.4 | YES | 3 | NO | NO |

- ข้อมูลแบบไม่มีโครงสร้าง (Unstructured Data) เช่น ข้อมูลที่เป็นข้อความ
ข้อมูลรูปภาพ



Data

- สำหรับการบรรยายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL
- แบ่งตามรูปแบบของข้อมูล

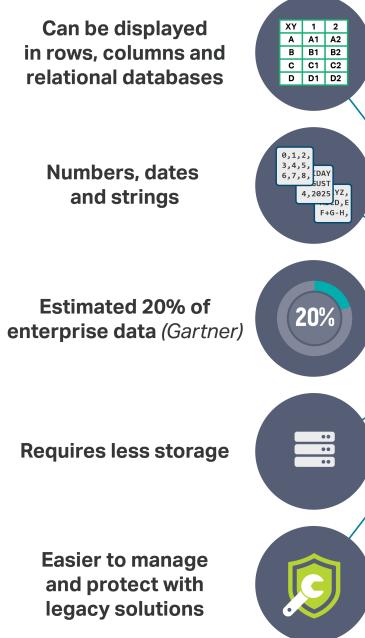
- ข้อมูลแบบกึ่งมีโครงสร้าง (semi-structure) เช่น ข้อมูล XML หรือ JSON

```
<?xml version="1.0" standalone="yes" ?>
<BankAccount>
    <Number>1234</Number>
    <Type>Checking</Type>
    <OpenDate>11/04/2020</OpenDate>
    <Balance>25999.00</Balance>
    <AccountHolder>
        <FirstName>Manee</FirstName>
        <LastName>Yindee</LastName>
    <AccountHolder>
</BankAccount>
```

Data

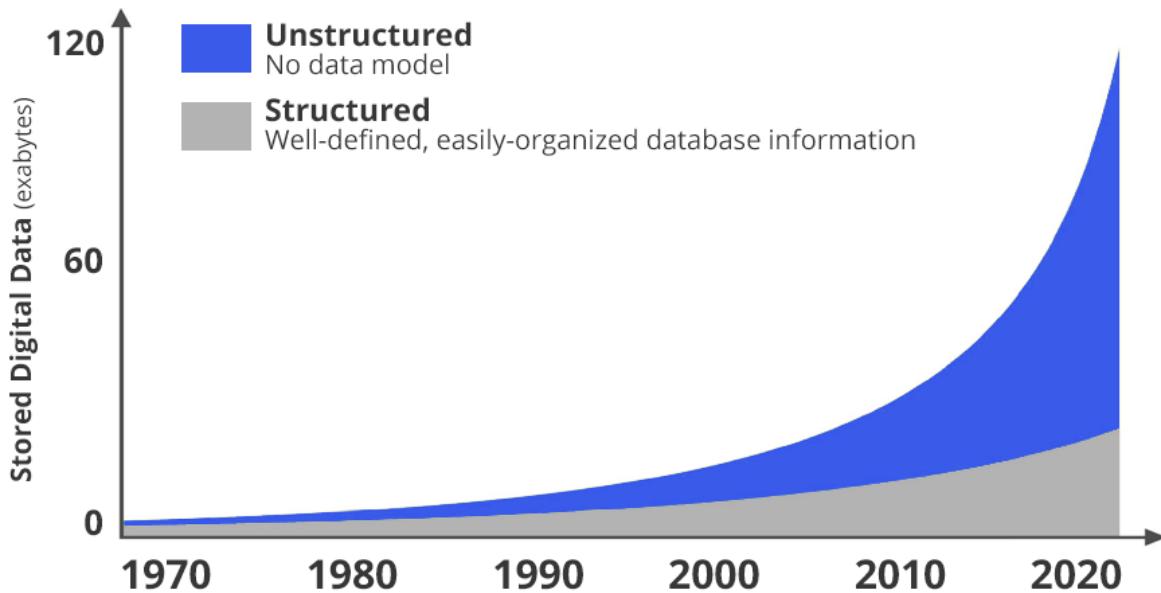
สำหรับการบรรยายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL

Structured Data vs Unstructured Data



Data

สำหรับการบรรยายในหลักสูตร M. Sc. in Artificial Intelligence for Business Analytics, KMITL



source: <https://www.m-files.com/blog/what-is-structured-data-vs-unstructured-data/>

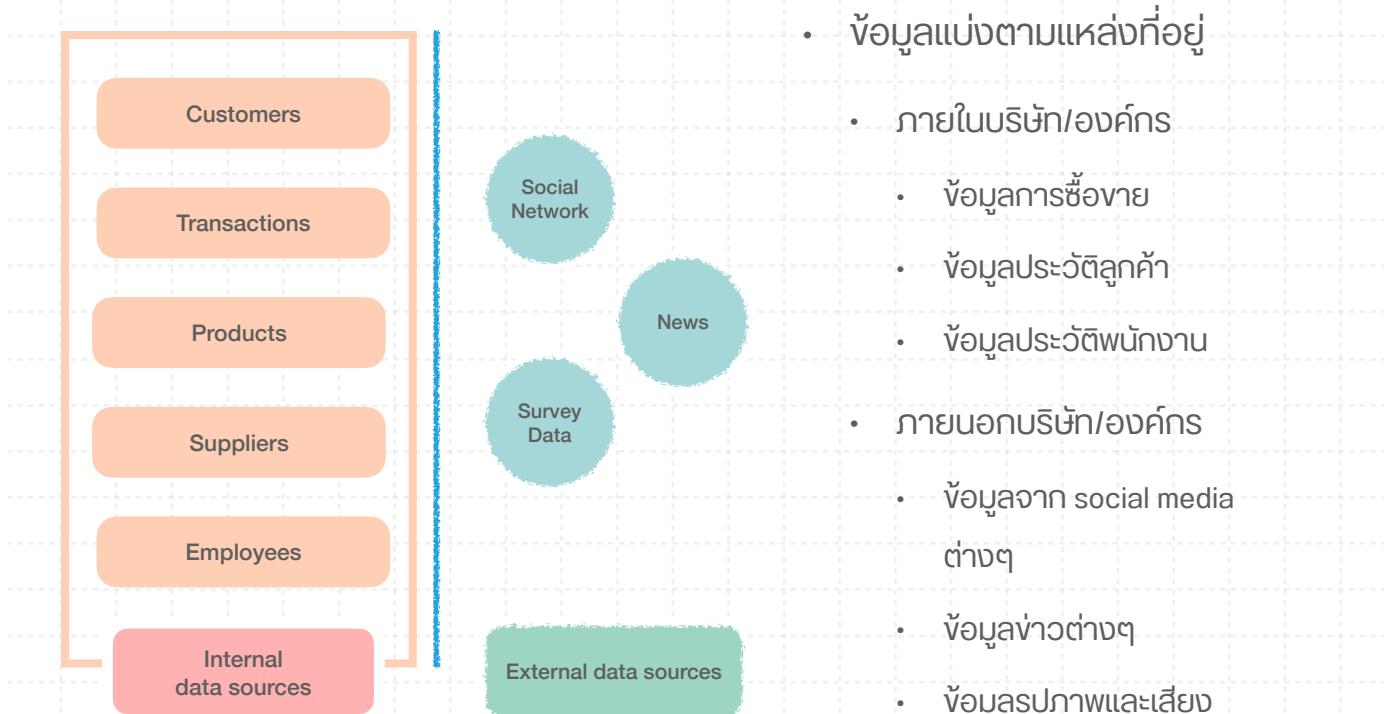
Data

สำหรับการบรรยายในหลักสูตร M. Sc. in Artificial Intelligence for Business Analytics, KMITL

- แบ่งตาม **ลักษณะ** ของข้อมูล
- แบ่งตาม **รูปแบบ** ของข้อมูล
- แบ่งตาม **แหล่งที่อยู่** ของข้อมูล
 - ข้อมูล **ภายใน** องค์กร (Internal Data Sources)
 - ข้อมูล **ภายนอก** องค์กร (External Data Sources)

Data

สำหรับการบรรยายในหลักสูตร M. Sc. in Artificial Intelligence for Business Analytics, KMITL



Data

สำหรับการบรรยายในหลักสูตร M. Sc. in Artificial Intelligence for Business Analytics, KMITL

• ข้อมูลที่ในองค์กรที่ใช้พิจารณาจะเป็น

- ข้อมูล**เชิงประชากรศาสตร์ (demographics)** และ ข้อมูล**เชิงพฤติกรรม (behavior)**
- ข้อมูล**ภายในองค์กร**
- ข้อมูลที่มี**โครงสร้าง**
- ตัวอย่างข้อมูล เช่น
 - ข้อมูลลูกค้า เป็นข้อมูล**เชิงประชากรศาสตร์ (demographics)**
 - ข้อมูลการซื้อสินค้า เป็นข้อมูล**เชิงพฤติกรรม (behavior)**

Introduction

- สำหรับการบรรยายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL
- ตัวอย่างการใช้ Big Data Analytics ในชีวิตประจำวัน
 - ลักษณะของข้อมูล
 - การจัดเก็บข้อมูล (Data Management)**
 - ฐานข้อมูล (Database)**
 - คลังข้อมูล (Data Warehouse)
 - การจัดการข้อมูลขนาดใหญ่ (Big Data)
 - ตัวอย่างการประยุกต์ใช้การวิเคราะห์ข้อมูลด้วยเทคโนโลยี Data Science

Data Storage

- สำหรับการบรรยายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL
- ตัวอย่างข้อมูลที่จัดเก็บในองค์กร/บริษัท

ตารางเก็บข้อมูลลูกค้า

| Customer ID | Name | Surname | Gender | Birthdate |
|-------------|------|---------|--------|------------|
| D2021-0001 | ปิติ | ยันดี | ชาย | 28/08/2545 |
| | | | | |
| | | | | |
| | | | | |

ชื่อ

ปิติ

นามสกุล

ยันดี

เพศ

ชาย

วันเดือนปีเกิด

28 สิงหาคม 2545

เบอร์โทรศัพท์

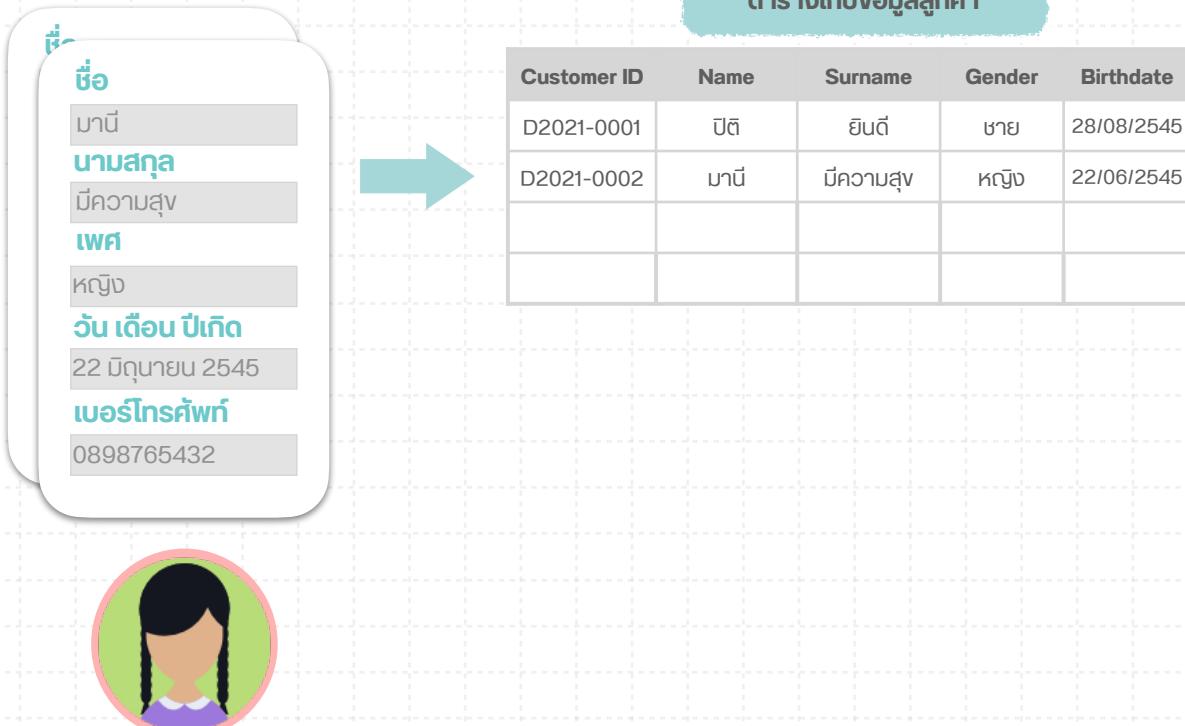
0891234567



Data Storage

- สำหรับการบรรยายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL
 • ตัวอย่างข้อมูลที่จัดเก็บในองค์กร/บริษัท

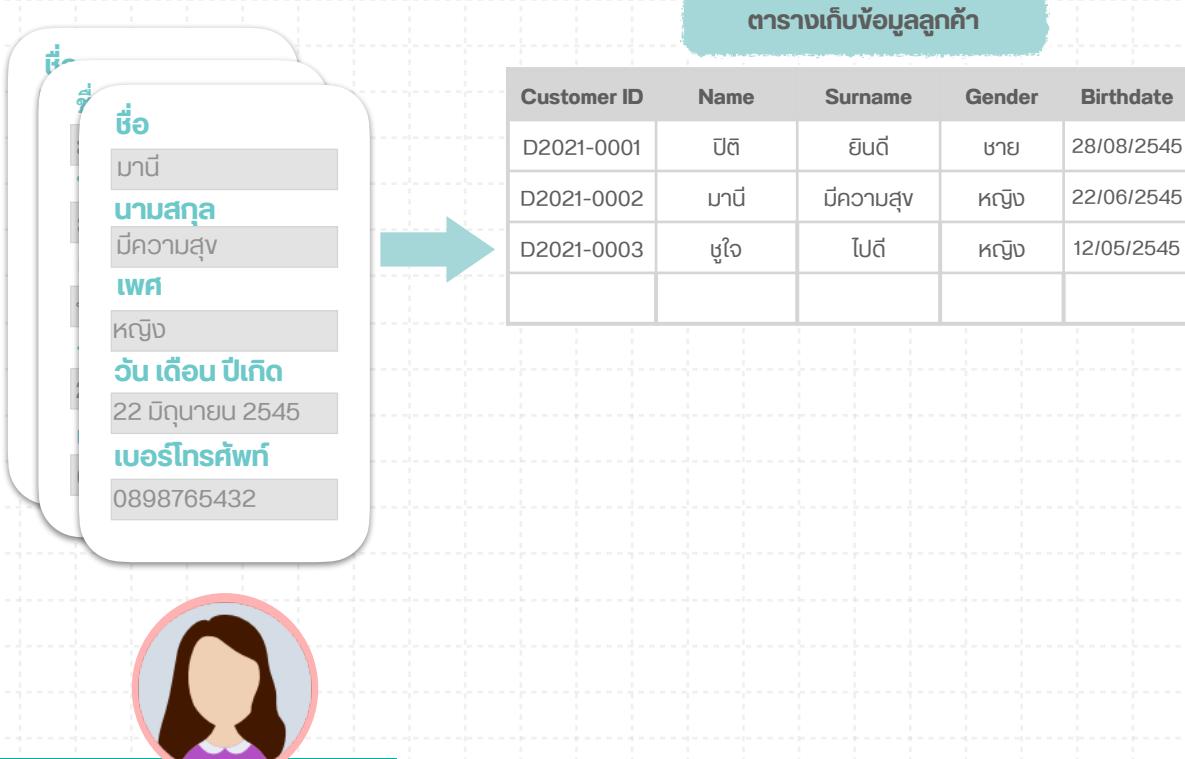
ตารางเก็บข้อมูลลูกค้า



Data Storage

- สำหรับการบรรยายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL
 • ตัวอย่างข้อมูลที่จัดเก็บในองค์กร/บริษัท

ตารางเก็บข้อมูลลูกค้า



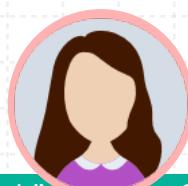
Data Storage

- สำหรับการบรรยายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL
 • ตัวอย่างข้อมูลที่จัดเก็บในองค์กร/บริษัท

ตารางเก็บข้อมูลลูกค้า



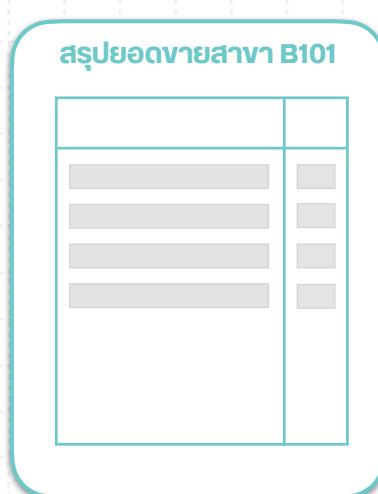
| Customer ID | Name | Surname | Gender | Birthdate |
|-------------|-------|-----------|--------|------------|
| D2021-0001 | ปีติ | ยันต์ | ชาย | 28/08/2545 |
| D2021-0002 | นาเน่ | มีความสุข | หญิง | 22/06/2545 |
| D2021-0003 | บูรพา | ໄปส์ | หญิง | 12/05/2545 |
| ... | ... | ... | ... | ... |



Data Storage

- สำหรับการบรรยายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL
 • ตัวอย่างข้อมูลที่จัดเก็บในองค์กร/บริษัท

ตารางเก็บข้อมูลยอดขาย



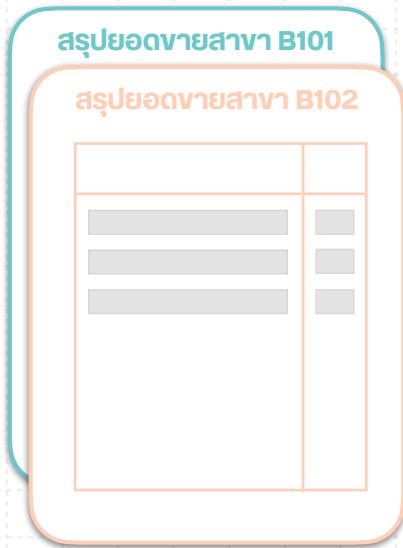
| Branch | Date | Product ID | QTY | Total |
|--------|------------|------------|-----|-------|
| B101 | 31-01-2021 | 00100 | 10 | 2,650 |
| B101 | 31-01-2021 | 00150 | 5 | 1,000 |
| B101 | 31-01-2021 | 00182 | 15 | 3,000 |
| B101 | 31-01-2021 | 00185 | 20 | 4,000 |

Data Storage

สำหรับการบรรยายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL

- ตัวอย่างข้อมูลที่จัดเก็บในองค์กร/บริษัท

ตารางเก็บข้อมูลยอดขาย



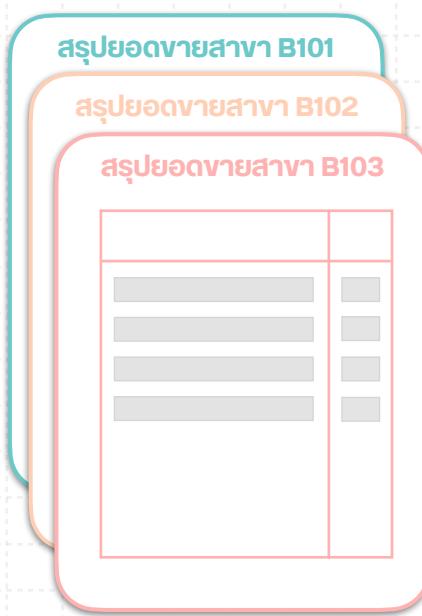
| Branch | Date | Product ID | QTY | Total |
|--------|------------|------------|-----|-------|
| B101 | 31-01-2021 | 00100 | 10 | 2,650 |
| B101 | 31-01-2021 | 00150 | 5 | 1,000 |
| B101 | 31-01-2021 | 00182 | 15 | 3,000 |
| B101 | 31-01-2021 | 00185 | 20 | 4,000 |
| B102 | 31-01-2021 | 00150 | 5 | 1,000 |
| B102 | 31-01-2021 | 00182 | 3 | 600 |
| B102 | 31-01-2021 | 00185 | 2 | 400 |

Data Storage

สำหรับการบรรยายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL

- ตัวอย่างข้อมูลที่จัดเก็บในองค์กร/บริษัท

ตารางเก็บข้อมูลยอดขาย



| Branch | Date | Product ID | QTY | Total |
|--------|------------|------------|-----|--------|
| B101 | 31-01-2021 | 00100 | 10 | 2,650 |
| B101 | 31-01-2021 | 00150 | 5 | 1,000 |
| B101 | 31-01-2021 | 00182 | 15 | 3,000 |
| B101 | 31-01-2021 | 00185 | 20 | 4,000 |
| B102 | 31-01-2021 | 00150 | 5 | 1,000 |
| B102 | 31-01-2021 | 00182 | 3 | 600 |
| B102 | 31-01-2021 | 00185 | 2 | 400 |
| B103 | 31-01-2021 | 00100 | 20 | 5,300 |
| B103 | 31-01-2021 | 00150 | 30 | 6,000 |
| B103 | 31-01-2021 | 00182 | 40 | 8,000 |
| B103 | 31-01-2021 | 00185 | 50 | 10,000 |

Data Storage

สำหรับการศึกษาในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL

- การเก็บข้อมูลแบบไฟล์ (file based system)
 - เก็บข้อมูลไว้ในไฟล์ เช่น Excel หรือ Word หรือ Text File
 - แต่ละไฟล์มีรูปแบบการจัดเก็บที่ต่างกัน
- ข้อจำกัดของการเก็บข้อมูลแบบไฟล์
 - ไม่มีรูปแบบมาตรฐาน (no standard)
 - ข้อมูลมีการเก็บซ้ำซ้อนกัน (duplicate) ระหว่างแต่ละคน
 - ข้อมูล update ไม่ตรงกัน
 - การดึงข้อมูลมาใช้งานจากหลายๆ ที่ทำได้ยาก

Data Storage: Database

สำหรับการศึกษาในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL

- ฐานข้อมูล (database)

- เป็นการเก็บข้อมูลไว้ในตารางซึ่งมีความสัมพันธ์กัน ทำให้สามารถ update, delete ได้อย่างถูกต้อง และช่วยให้การค้นหาข้อมูลง่ายขึ้น

ตาราง Customers

| Customer_ID | Name | Surname | Gender | Birthdate |
|-------------|-------|----------|--------|------------|
| D2021-0001 | ปัต | ยันต์ | ชาย | 28/08/2545 |
| D2021-0002 | นาเน่ | เมืองสุข | หญิง | 22/06/2545 |
| D2021-0003 | ชูใจ | ໄປดี | หญิง | 12/05/2545 |

ตาราง Transactions

| Customer_ID | Branch_ID | Product_ID | Qty | Date |
|-------------|-----------|------------|-----|------------|
| D2021-0001 | B101 | 00150 | 1 | 14/02/2021 |
| D2021-0001 | B102 | 00185 | 2 | 14/02/2021 |
| D2021-0002 | B102 | 182 | 1 | 12/05/2021 |
| D2021-0002 | B102 | 185 | 1 | 12/05/2021 |

Data Storage: Database

สำหรับการบรรยายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL

- ฐานข้อมูล (database) ประกอบด้วย

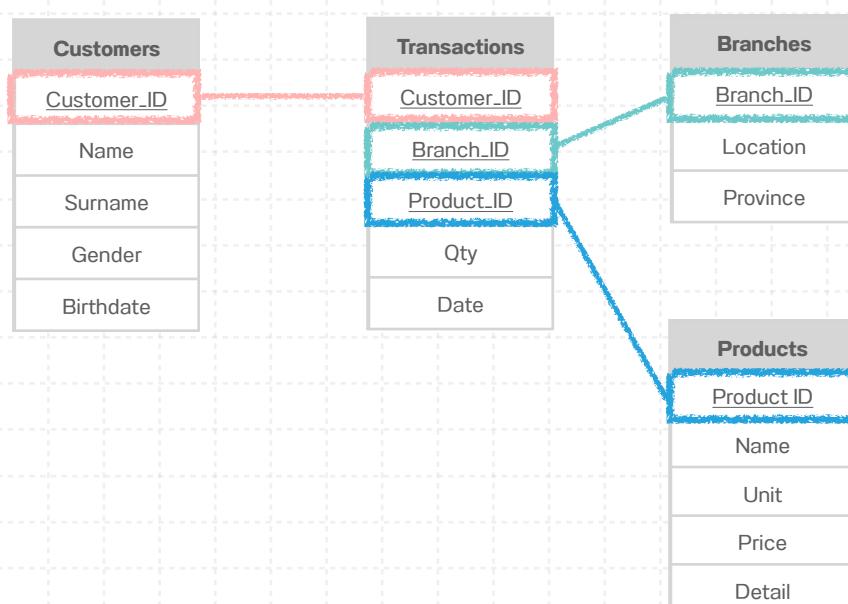
- ตาราง (table หรือ relation) เก็บข้อมูลส่วนต่างๆ ไว้ และประกอบด้วย
 - ฟิลด์** (field) แสดงคุณลักษณะ (attribute) ของข้อมูลแต่ละรายการ
 - เรคอร์ด** (record) แสดงข้อมูลข้อมูลแต่ละแถว (row) หรือตัวอย่าง (example)
- ความสัมพันธ์ (relation) และความสัมพันธ์ระหว่างตาราง โดยใช้คีย์ (key) เป็นพอล์ดกี่ซึ่งมีอยู่ในแต่ละตาราง



Data Storage: Database

สำหรับการบรรยายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL

- ฐานข้อมูลจะแสดงในรูปแบบของ schema diagram ได้ เพื่อให้เห็นฟิลด์และความสัมพันธ์ระหว่างตาราง



Big Data Careers

สำหรับการบรรยายให้นศ. โครงการ MBA IT-Smart



Data Analyst

(Business Analyst)

Skills

- Statistics
- Communication
- Business Knowledge

Tools

- Excel
- Tableau / Power BI
- SQL



Data Scientist

(Statisticians, Data Managers)

Skills

- Mathematics
- Programming
- Analytics

Tools

- SQL
- Python / R
- RapidMiner / SAS



Data Engineer

(Data Architects, Database Administrators)

Skills

- Programming
- Mathematics
- Big Data

Tools

- Hadoop / Spark
- NoSQL
- Python

Data Storage: Database

สำหรับการบรรยายในหลักสูตร M. Sc. in Artificial Intelligence for Business Analytics, KMITL

- ในการจัดการข้อมูลในฐานข้อมูลจะใช้คำสั่ง SQL (Structure Query Language)

• Data Definition Language (DDL)

- คำสั่ง **CREATE DATABASE** สำหรับสร้างฐานข้อมูล
- คำสั่ง **CREATE TABLE** สำหรับสร้างตารางต่างๆ ในฐานข้อมูล

• Data Manipulation Language (DML)

- คำสั่ง **INSERT** สำหรับเพิ่มข้อมูลเข้าไปในตาราง
- คำสั่ง **UPDATE** สำหรับแก้ไขข้อมูลเข้าในตาราง
- คำสั่ง **DELETE** สำหรับลบข้อมูลออกจากตาราง
- คำสั่ง **SELECT** สำหรับดึงข้อมูลในตารางมาแสดงผลในรูปแบบต่างๆ

Data Storage: Database

- สำหรับการบรรยายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL
- คำสั่ง **SELECT** ใช้ในการดึงข้อมูลจากตารางต่างๆ ในฐานข้อมูล

- มีโครงสร้างคำสั่ง

```
SELECT FIELD_1 , FIELD_2 , ... , FIELD_N  
FROM table_name  
WHERE FIELD_2 = CONDITION_1
```

- ตัวอย่างคำสั่ง

```
SELECT Name , Gender  
FROM Customers
```

```
WHERE Gender = "หญิง"
```

| | FIELD 1 | Name | Surname | Gender | Birthdate |
|-------------|---------|-------------|---------|------------|-----------|
| Customer_ID | | | | | |
| D2021-0001 | ปีเต | ยันดี | ชาย | 28/08/2545 | |
| D2021-0002 | มาเน | เมื่ความสุข | หญิง | 22/06/2545 | |
| D2021-0003 | ชูใจ | โปรดี | หญิง | 12/05/2545 | |

ตาราง Customers

Data Storage: Database

- สำหรับการบรรยายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL
- คำสั่ง **SELECT** ใช้ในการดึงข้อมูลจากตารางต่างๆ ในฐานข้อมูล

- มีโครงสร้างคำสั่ง

```
SELECT FIELD_1 , FIELD_2 , ... , FIELD_N  
FROM table_name  
WHERE FIELD_2 = CONDITION_1
```

- ตัวอย่างคำสั่ง

```
SELECT Name , Gender  
FROM Customers
```

```
WHERE Gender = "หญิง"
```

| | FIELD 1 | FIELD 2 | |
|-------------|---------|-------------|------|
| Customer_ID | | | |
| D2021-0001 | ปีเต | ยันดี | ชาย |
| D2021-0002 | มาเน | เมื่ความสุข | หญิง |
| D2021-0003 | ชูใจ | โปรดี | หญิง |

ตาราง Customers

Data Storage: Database

- สำหรับการบรรยายในหลักสูตร M.Sc. in Artificial Intelligence for Business Analytics, KMITL
- คำสั่ง **SELECT** ใช้ในการดึงข้อมูลจากตารางต่างๆ ในฐานข้อมูล

- มีโครงสร้างคำสั่ง

```
SELECT FIELD_1 , FIELD_2 , ... , FIELD_N
```

```
FROM table_name
```

```
WHERE FIELD_2 = CONDITION_1
```

- ตัวอย่างคำสั่ง

```
SELECT Name , Gender
```

```
FROM Customers
```

```
WHERE Gender = "หญิง"
```

| | FIELD 1 | FIELD 2 | | |
|-------------|----------------|----------------|--------|------------|
| Customer_ID | Name | Surname | Gender | Birthdate |
| D2021-0001 | ปีเต | ยันดี | ชาย | 28/08/2545 |
| D2021-0002 | นาเน | เมื่ความสุข | หญิง | 22/06/2545 |
| D2021-0003 | ชูใจ | โปรดี | หญิง | 12/05/2545 |

ตาราง Customers

Data Storage: Database

- สำหรับการบรรยายในหลักสูตร M.Sc. in Artificial Intelligence for Business Analytics, KMITL
- ตัวอย่างฐานข้อมูล

ตาราง Branches

| Branch_ID | Location | Province |
|-----------|----------|-----------|
| B101 | ลำลูกกา | ปทุมธานี |
| B102 | บางเขน | กรุงเทพฯ |
| B103 | บางกรวย | สหธรรมุรี |
| B104 | รามคำแหง | กรุงเทพฯ |

ตาราง Customers

| Customer_ID | Name | Surname | Gender | Birthdate |
|-------------|------|-------------|--------|------------|
| D2021-0001 | ปีเต | ยันดี | ชาย | 28/08/2545 |
| D2021-0002 | นาเน | เมื่ความสุข | หญิง | 22/06/2545 |
| D2021-0003 | ชูใจ | โปรดี | หญิง | 12/05/2545 |

ตาราง Transactions

| Customer_ID | Branch_ID | Product_ID | Qty | Date |
|-------------|-----------|------------|-----|------------|
| D2021-0001 | B101 | 00150 | 1 | 14/02/2021 |
| D2021-0001 | B102 | 00185 | 2 | 14/02/2021 |
| D2021-0002 | B102 | 00182 | 1 | 12/04/2021 |
| D2021-0002 | B102 | 00185 | 1 | 12/04/2021 |

ตาราง Products

| Product_ID | Name | Price |
|------------|--------------------------|-------|
| 00100 | Marketing 4.0 | 300 |
| 00150 | AI for Marketing | 450 |
| 00182 | Introduction to Database | 300 |
| 00185 | Predictive Marketing | 500 |

Data Storage: Database

สำหรับการบรรยายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL

- **SELECT * FROM Customers**

ตาราง Customers

| Customer.ID | Name | Surname | Gender | Birthdate |
|-------------|------|-----------|--------|------------|
| D2021-0001 | ปัต | ยันดี | ชาย | 28/08/2545 |
| D2021-0002 | มาเน | มีความสุข | หญิง | 22/06/2545 |
| D2021-0003 | ชูใจ | ໄປສີ | หญิง | 12/05/2545 |

Data Storage: Database

สำหรับการบรรยายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL

- **SELECT * FROM Customers**

ตาราง Customers

| Customer.ID | Name | Surname | Gender | Birthdate |
|-------------|------|-----------|--------|------------|
| D2021-0001 | ปัต | ยันดี | ชาย | 28/08/2545 |
| D2021-0002 | มาเน | มีความสุข | หญิง | 22/06/2545 |
| D2021-0003 | ชูใจ | ໄປສີ | หญิง | 12/05/2545 |



| Customer.ID | Name | Surname | Gender | Birthdate |
|-------------|------|-----------|--------|------------|
| D2021-0001 | ปัต | ยันดี | ชาย | 28/08/2545 |
| D2021-0002 | มาเน | มีความสุข | หญิง | 22/06/2545 |
| D2021-0003 | ชูใจ | ໄປສີ | หญิง | 12/05/2545 |

ผลลัพธ์ (output)

Data Storage: Database

สำหรับการบรรยายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL

- **SELECT Name FROM Customers**

ตาราง Customers

| Customer.ID | Name | Surname | Gender | Birthdate |
|-------------|-------|-----------|--------|------------|
| D2021-0001 | ปิติ | ยันดี | ชาย | 28/08/2545 |
| D2021-0002 | มาเน่ | มีความสุข | หญิง | 22/06/2545 |
| D2021-0003 | บุ๊จ | ໄປສີ | หญิง | 12/05/2545 |

Data Storage: Database

สำหรับการบรรยายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL

- **SELECT Name FROM Customers**

ตาราง Customers

| Customer.ID | Name | Surname | Gender | Birthdate |
|-------------|-------|-----------|--------|------------|
| D2021-0001 | ปิติ | ยันดี | ชาย | 28/08/2545 |
| D2021-0002 | มาเน่ | มีความสุข | หญิง | 22/06/2545 |
| D2021-0003 | บุ๊จ | ໄປສີ | หญิง | 12/05/2545 |



| Name |
|-------|
| ปิติ |
| มาเน่ |
| บุ๊จ |

ผลลัพธ์ (output)

Data Storage: Database

สำหรับการบรรยายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL

- **SELECT Name , Surname FROM Customers**

ตาราง Customers

| Customer.ID | Name | Surname | Gender | Birthdate |
|-------------|-------|-----------|--------|------------|
| D2021-0001 | ปิติ | ยันดี | ชาย | 28/08/2545 |
| D2021-0002 | มาเน่ | มีความสุข | หญิง | 22/06/2545 |
| D2021-0003 | บู๊จ | ໄປສີ | หญิง | 12/05/2545 |

Data Storage: Database

สำหรับการบรรยายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL

- **SELECT Name , Surname FROM Customers**

ตาราง Customers

| Customer.ID | Name | Surname | Gender | Birthdate |
|-------------|-------|-----------|--------|------------|
| D2021-0001 | ปิติ | ยันดี | ชาย | 28/08/2545 |
| D2021-0002 | มาเน่ | มีความสุข | หญิง | 22/06/2545 |
| D2021-0003 | บู๊จ | ໄປສີ | หญิง | 12/05/2545 |



| Name | Surname |
|-------|-----------|
| ปิติ | ยันดี |
| มาเน่ | มีความสุข |
| บู๊จ | ໄປສີ |

ผลลัพธ์ (output)

Data Storage: Database

สำหรับการสอนรายวิชาหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL

- **SELECT * FROM Products WHERE Price > 300**

ตาราง Products

| Product_ID | Name | Price |
|------------|--------------------------|-------|
| 00100 | Marketing 4.0 | 300 |
| 00150 | AI for Marketing | 450 |
| 00182 | Introduction to Database | 300 |
| 00185 | Predictive Marketing | 500 |

Data Storage: Database

สำหรับการสอนรายวิชาหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL

- **SELECT * FROM Products WHERE Price > 300**

ตาราง Products

| Product_ID | Name | Price |
|------------|--------------------------|-------|
| 00100 | Marketing 4.0 | 300 |
| 00150 | AI for Marketing | 450 |
| 00182 | Introduction to Database | 300 |
| 00185 | Predictive Marketing | 500 |



| Product_ID | Name | Price |
|------------|----------------------|-------|
| 00150 | AI for Marketing | 450 |
| 00185 | Predictive Marketing | 500 |

ผลลัพธ์ (output)

Data Storage: Database

สำหรับการสูงสุด M: Sc. in Artificial Intelligence for Business Analytics, KMITL

- **SELECT * FROM Products WHERE Price >= 300**

ตาราง Products

| Product_ID | Name | Price |
|------------|--------------------------|-------|
| 00100 | Marketing 4.0 | 300 |
| 00150 | AI for Marketing | 450 |
| 00182 | Introduction to Database | 300 |
| 00185 | Predictive Marketing | 500 |

Data Storage: Database

สำหรับการสูงสุด M: Sc. in Artificial Intelligence for Business Analytics, KMITL

- **SELECT * FROM Products WHERE Price >= 300**

ตาราง Products

| Product_ID | Name | Price |
|------------|--------------------------|-------|
| 00100 | Marketing 4.0 | 300 |
| 00150 | AI for Marketing | 450 |
| 00182 | Introduction to Database | 300 |
| 00185 | Predictive Marketing | 500 |



| Product_ID | Name | Price |
|------------|--------------------------|-------|
| 00100 | Marketing 4.0 | 300 |
| 00150 | AI for Marketing | 450 |
| 00182 | Introduction to Database | 300 |
| 00185 | Predictive Marketing | 500 |

ผลลัพธ์ (output)

Data Storage: Database

สำหรับการสูงสุด M: Sc. in Artificial Intelligence for Business Analytics, KMITL

- **SELECT * FROM Products WHERE Name LIKE "%Marketing%"**

ตาราง Products

| Product_ID | Name | Price |
|------------|--------------------------|-------|
| 00100 | Marketing 4.0 | 300 |
| 00150 | AI for Marketing | 450 |
| 00182 | Introduction to Database | 300 |
| 00185 | Predictive Marketing | 500 |

Data Storage: Database

สำหรับการสูงสุด M: Sc. in Artificial Intelligence for Business Analytics, KMITL

- **SELECT * FROM Products WHERE Name LIKE "%Marketing%"**

ตาราง Products

| Product_ID | Name | Price |
|------------|--------------------------|-------|
| 00100 | Marketing 4.0 | 300 |
| 00150 | AI for Marketing | 450 |
| 00182 | Introduction to Database | 300 |
| 00185 | Predictive Marketing | 500 |



| Product_ID | Name | Price |
|------------|----------------------|-------|
| 00100 | Marketing 4.0 | 300 |
| 00150 | AI for Marketing | 450 |
| 00185 | Predictive Marketing | 500 |

ผลลัพธ์ (output)

Data Storage: Database

สำหรับการเรียนรู้ในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL

- ถ้าต้องการเรียงลำดับข้อมูล ใช้คำสั่ง **ORDER BY** ต่อด้านหลัง และมีคำว่า

- **ASC** ในกรณีที่ต้องการเรียงข้อมูลจาก น้อยไปมาก
- **DESC** ในกรณีที่ต้องการเรียงข้อมูลจาก มากไปน้อย
- โครงสร้างคำสั่ง SQL ในการเรียงลำดับจาก น้อยไปมาก

```
SELECT field_1, field_2, ..., field_N
FROM table_name1
ORDER BY field_1 ASC
```

- โครงสร้างคำสั่ง SQL ในการเรียงลำดับจาก มากไปน้อย

```
SELECT field_1, field_2, ..., field_N
FROM table_name1
ORDER BY field_1 DESC
```

Data Storage: Database

สำหรับการเรียนรู้ในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL

- **SELECT * FROM Products ORDER BY Price ASC**

ตาราง Products

| Product_ID | Name | Price |
|------------|--------------------------|-------|
| 00100 | Marketing 4.0 | 300 |
| 00150 | AI for Marketing | 450 |
| 00182 | Introduction to Database | 300 |
| 00185 | Predictive Marketing | 500 |



| Product_ID | Name | Price |
|------------|--------------------------|-------|
| 00100 | Marketing 4.0 | 300 |
| 00182 | Introduction to Database | 300 |
| 00150 | AI for Marketing | 450 |
| 00185 | Predictive Marketing | 500 |

ผลลัพธ์ (output)

Data Storage: Database

สำหรับการบรรยายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL

- **SELECT * FROM Products ORDER BY Price DESC**



Data Storage: Database

สำหรับการบรรยายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL

- กรณีที่ต้องการแสดงข้อมูลแค่บางแท่ง เช่น ยอดซื้อสูงสุด 5 อันดับแรก จะต้อง

ใช้ **ORDER BY** คู่กับ **LIMIT**

- โครงสร้างคำสั่ง SQL ในการแสดงข้อมูลที่เรียงลำดับ **น้อยสุด 5 อันดับแรก**

```
SELECT field_1, field_2, ..., field_N
FROM table_name1
ORDER BY field_1 ASC
LIMIT 0,5
```

- โครงสร้างคำสั่ง SQL ในการแสดงข้อมูลที่เรียงลำดับ **มากสุด 5 อันดับแรก**

```
SELECT field_1, field_2, ..., field_N
FROM table_name1
ORDER BY field_1 DESC
LIMIT 0,5
```

Data Storage: Database

สำหรับการสูงสุด M: Sc. in Artificial Intelligence for Business Analytics, KMITL

- **SELECT * FROM Products ORDER BY Price DESC LIMIT 0,2**



Data Storage: Database

สำหรับการสูงสุด M: Sc. in Artificial Intelligence for Business Analytics, KMITL

- **SELECT * FROM Products ORDER BY Price DESC LIMIT 0,2**



Data Storage: Database

สำหรับการเรียนรู้ในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL

- ในหลายๆ ครั้งที่ข้อมูลกระจายไปอยู่ในหลายตาราง ต้องทำการเชื่อมโยงข้อมูล (join) เสียก่อน

- โครงสร้างคำสั่ง SQL ในการ join คือ

```
SELECT field_1, field_2, ..., field_N
FROM table_name1, table_name2
WHERE table_name1.field = table_name2.field
```

- ตัวอย่าง

```
SELECT Product_ID, Name, Qty, Price
FROM Transactions, Products
WHERE Transactions.Product_ID = Products.Product_ID
```

Data Storage: Database

สำหรับการเรียนรู้ในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL

- ```
SELECT Product_ID, Name, Qty, Price
FROM Transactions, Products
WHERE Transactions.Product_ID = Products.Product_ID
```

ตาราง Transactions

| Customer_ID | Branch_ID | Product_ID | Qty | Date       |
|-------------|-----------|------------|-----|------------|
| D2021-0001  | B101      | 00150      | 1   | 14/02/2021 |
| D2021-0001  | B102      | 00185      | 2   | 14/02/2021 |
| D2021-0002  | B102      | 00182      | 1   | 12/04/2021 |
| D2021-0002  | B102      | 00185      | 1   | 12/04/2021 |

ตาราง Products

| Product_ID | Name                     | Price |
|------------|--------------------------|-------|
| 00100      | Marketing 4.0            | 300   |
| 00150      | AI for Marketing         | 450   |
| 00182      | Introduction to Database | 300   |
| 00185      | Predictive Marketing     | 500   |

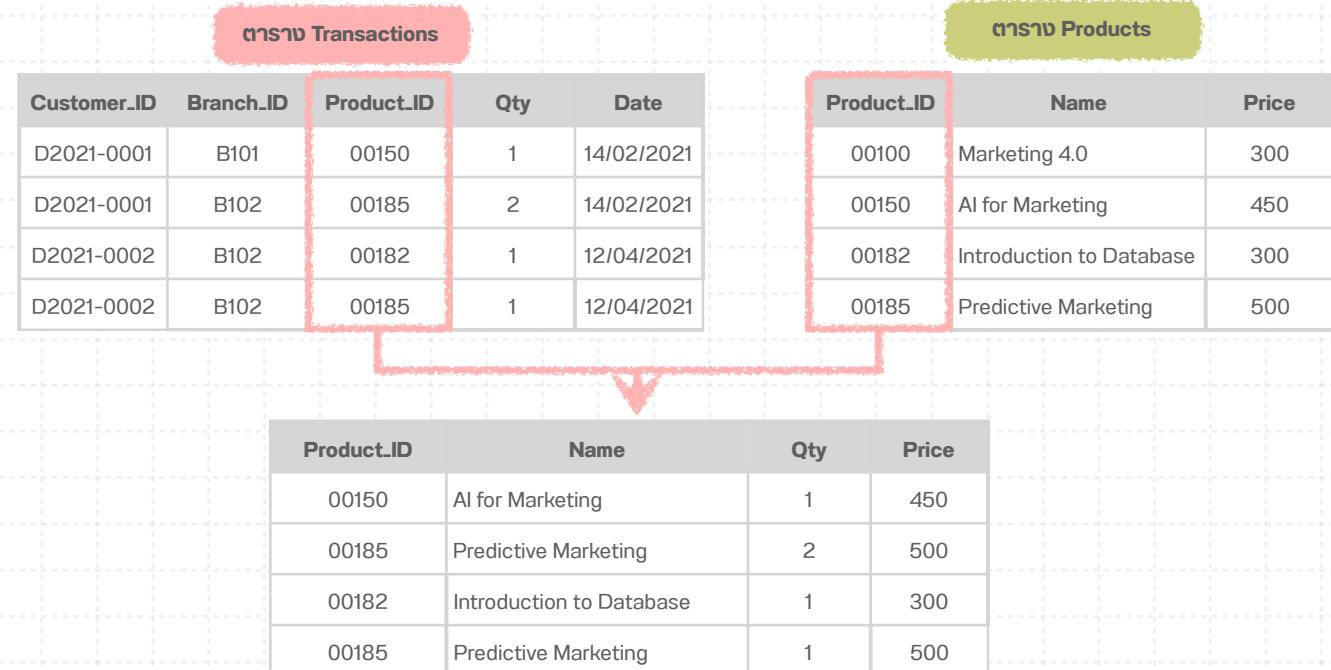
# Data Storage: Database

สำหรับการบรรยายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL

```

• SELECT Product_ID, Name, Qty, Price
 FROM Transactions, Products
 WHERE Transactions.Product_ID = Products.Product_ID

```



# Data Storage: Database

สำหรับการบรรยายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL

- ด้วยการค่าสรุป (aggregate) สามารถคำนวณได้โดยใช้ฟังก์ชัน

- COUNT** สำหรับนับจำนวนข้อมูล (row) ที่อยู่ในตาราง
- SUM** สำหรับรวมค่า (+) ของข้อมูลในตาราง
- AVG** สำหรับหาค่าเฉลี่ย (average) ของข้อมูลในตาราง
- MIN** หรือ **MAX** ใช้สำหรับหาค่าต่ำสุดหรือสูงสุดของข้อมูลในตาราง
- ส่วนใหญ่ใช้คู่กับ **group by** เพื่อกำหนดว่าจะจัดกลุ่มตามข้อมูลอะไร

# Data Storage: Database

สำหรับการอธิบายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL

- **SELECT SUM(Qty) FROM Transactions**

ตาราง Transactions

| Customer_ID | Branch_ID | Product_ID | Qty | Date       |
|-------------|-----------|------------|-----|------------|
| D2021-0001  | B101      | 00150      | 1   | 14/02/2021 |
| D2021-0001  | B102      | 00185      | 2   | 14/02/2021 |
| D2021-0002  | B102      | 00182      | 1   | 12/04/2021 |
| D2021-0002  | B102      | 00185      | 1   | 12/04/2021 |

# Data Storage: Database

สำหรับการอธิบายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL

- **SELECT SUM(Qty) as total FROM Transactions**

ตาราง Transactions

| Customer_ID | Branch_ID | Product_ID | Qty | Date       |
|-------------|-----------|------------|-----|------------|
| D2021-0001  | B101      | 00150      | 1   | 14/02/2021 |
| D2021-0001  | B102      | 00185      | 2   | 14/02/2021 |
| D2021-0002  | B102      | 00182      | 1   | 12/04/2021 |
| D2021-0002  | B102      | 00185      | 1   | 12/04/2021 |



| total |
|-------|
| 5     |

ผลลัพธ์ (output)

# Data Storage: Database

สำหรับการสูงสุด M: Sc. in Artificial Intelligence for Business Analytics, KMITL

- SELECT Product\_ID, SUM(Qty) FROM Transactions  
GROUP BY Product\_ID**

ตาราง Transactions

| Customer_ID | Branch_ID | Product_ID | Qty | Date       |
|-------------|-----------|------------|-----|------------|
| D2021-0001  | B101      | 00150      | 1   | 14/02/2021 |
| D2021-0001  | B102      | 00185      | 2   | 14/02/2021 |
| D2021-0002  | B102      | 00182      | 1   | 12/04/2021 |
| D2021-0002  | B102      | 00185      | 1   | 12/04/2021 |

# Data Storage: Database

สำหรับการสูงสุด M: Sc. in Artificial Intelligence for Business Analytics, KMITL

- SELECT Product\_ID, SUM(Qty) FROM Transactions  
GROUP BY Product\_ID**

ตาราง Transactions

| Customer_ID | Branch_ID | Product_ID | Qty | Date       |
|-------------|-----------|------------|-----|------------|
| D2021-0001  | B101      | 00150      | 1   | 14/02/2021 |
| D2021-0001  | B102      | 00185      | 2   | 14/02/2021 |
| D2021-0002  | B102      | 00182      | 1   | 12/04/2021 |
| D2021-0002  | B102      | 00185      | 1   | 12/04/2021 |

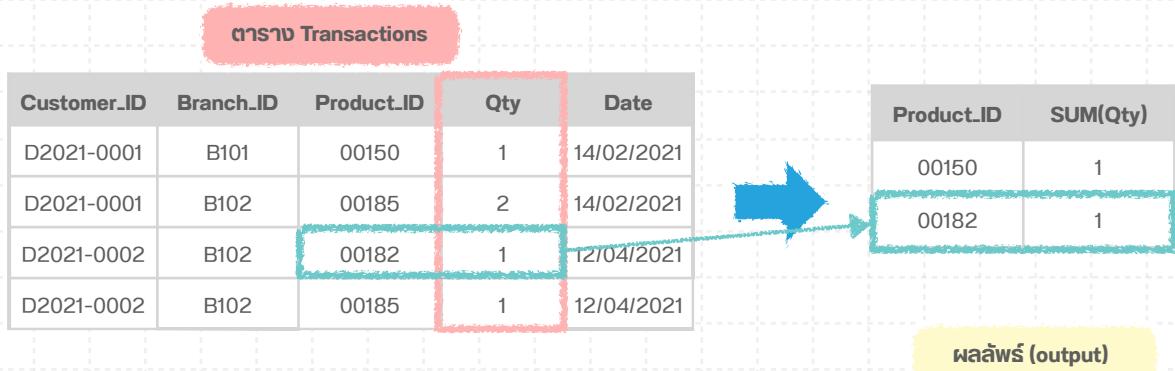
| Product_ID | SUM(Qty) |
|------------|----------|
| 00150      | 1        |

ผลลัพธ์ (output)

# Data Storage: Database

สำหรับการอธิบายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL

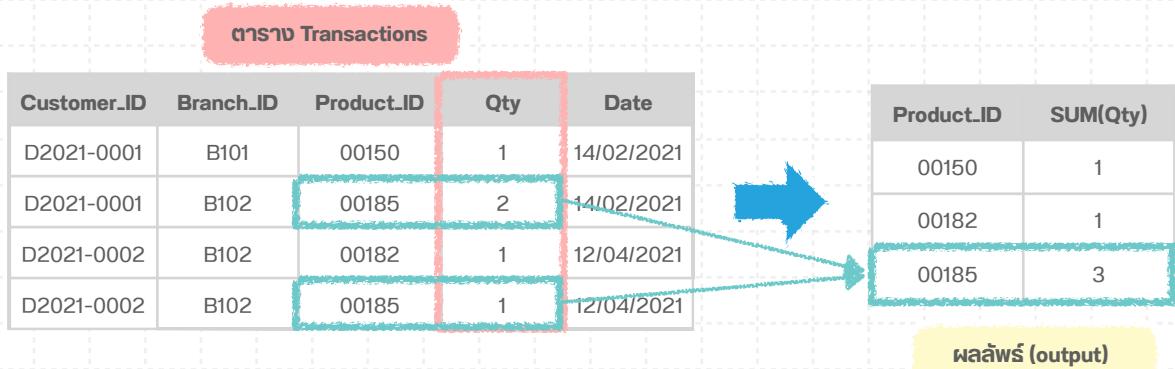
- `SELECT Product_ID, SUM(Qty) FROM Transactions  
GROUP BY Product_ID`



# Data Storage: Database

สำหรับการอธิบายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL

- `SELECT Product_ID, SUM(Qty) FROM Transactions  
GROUP BY Product_ID`



# Introduction

- สำหรับการบรรยายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL
- ตัวอย่างการใช้ Big Data Analytics ในชีวิตประจำวัน
- ลักษณะของข้อมูล
- การจัดเก็บข้อมูล (Data Management)**
  - ฐานข้อมูล (Database)
  - คลังข้อมูล (Data Warehouse)**
- การจัดการข้อมูลขนาดใหญ่ (Big Data)
- ตัวอย่างการประยุกต์ใช้การวิเคราะห์ข้อมูลด้วยเทคโนโลยี Data Science

## Data Storage: Data Warehouse

- สำหรับการบรรยายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL
- การเก็บข้อมูลในฐานข้อมูลมีข้อดีหลายอย่าง เช่น

- ทำให้แต่ละฝ่ายหรือแผนกมองเห็นข้อมูลที่ถูกต้องตรงกัน
- ฐานข้อมูลส่วนใหญ่จะ **เก็บไว้แยกจากกัน** เป็นแต่ละฝ่ายหรือแผนก เพื่อใช้ตอบโจทย์ของส่วนต้นเอง
- ถ้าผู้บริหาร **ต้องการดูภาพรวมของกิจกรรม** จำเป็นต้องดึงข้อมูลจากฐานข้อมูลต่างๆ เหล่านี้มาสรุปให้เห็นเป็น Dashboard หรือ รายงานต่างๆ
  - ค่อนข้างจะ **ใช้เวลานาน**
  - ข้อมูลแยกเป็นหลายๆ ตาราง การ **join** ข้อมูลข้างต้นอาจทำงานช้า
  - ข้อมูลที่มี **ความหลากหลาย** อาจจะเขียนไม่ตรงกัน
- มี concept ของ **คลังข้อมูล (data warehouse)** ที่ **รวบรวมข้อมูล** จากหลายๆ ฐานข้อมูลรวมกัน

# Data Storage: Data Warehouse

สำหรับการบรรยายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL



ฝ่ายขาย



ฝ่ายการตลาด



ฝ่ายบัญชี

- การเก็บข้อมูลในฐานข้อมูลมีข้อดีหลายอย่างทำให้แต่ละฝ่ายหรือแผนกมองเห็นข้อมูลที่ถูกต้องตรงกัน
- ฐานข้อมูลส่วนใหญ่จะเก็บไว้แยกจากกันเป็นแต่ละฝ่ายหรือแผนก เพื่อใช้ตอบโจทย์ของส่วนตนเอง
- ด้าผู้บริหารต้องการดูภาพรวมของกิจกรรมประจำเป็นต้องดึงข้อมูลจากฐานข้อมูลต่างๆ เหล่านี้มาสรุปให้เห็นเป็น Dashboard หรือรายงานต่างๆ
  - ค่อนข้างจะใช้เวลานาน
  - ข้อมูลแยกเป็นหลายๆ ตาราง การ join ข้อมูลข้ามตารางจะทำงานช้า
  - ข้อมูลที่มีความหลากหลาย

# Data Storage: Data Warehouse

สำหรับการบรรยายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL



ฝ่ายขาย



ฝ่ายการตลาด



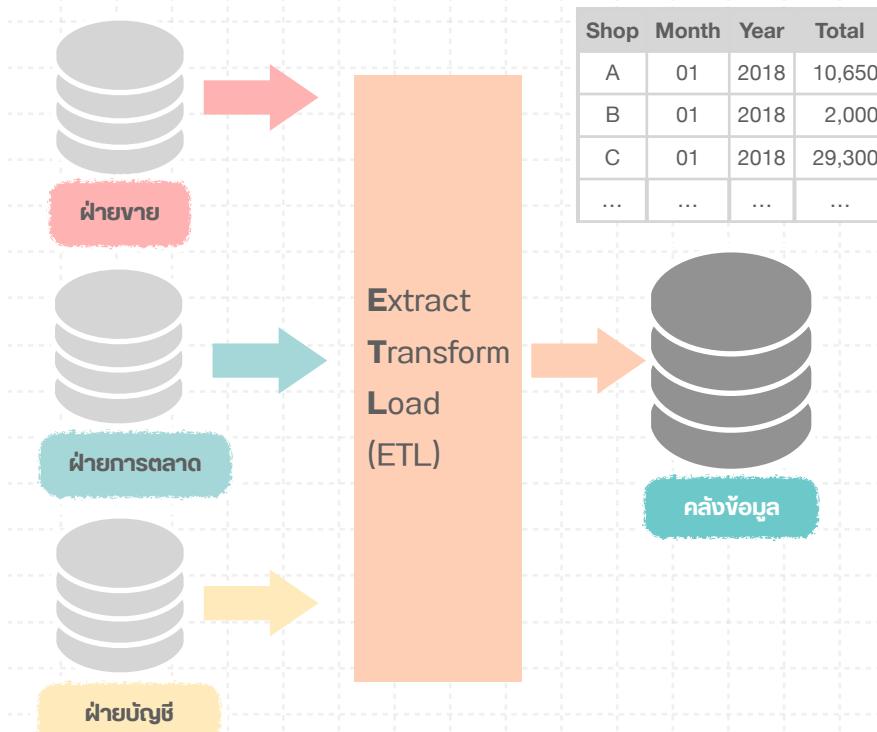
ฝ่ายบัญชี



- กระบวนการ ETL เป็น **การดึงข้อมูล (Extract)** **แปลงข้อมูล (Transform)** และ **บันทึกข้อมูล (Load)**
- เป็นการดึงข้อมูล (Extract) **เฉพาะด้านที่สนใจ** จากฐานข้อมูลต่างๆ มารวมกัน
- แปลงข้อมูลให้อยู่ในรูปแบบเดียวกัน** เช่น
  - แปลงข้อมูลเงินเดือนลาร์ให้เป็นเงินบาท
  - แปลงข้อมูลตัวย่อ เช่น F และ M ให้กลายเป็น Female และ Male
- บันทึกข้อมูล (Load) ลงไปในคลังข้อมูล (data warehouse)

# Data Storage: Data Warehouse

สำหรับการบรรยายในหลักสูตร M. Sc. in Artificial Intelligence for Business Analytics, KMITL



- ข้อมูลที่เก็บไว้ในคลังข้อมูล

มักจะอยู่ในรูปแบบของ

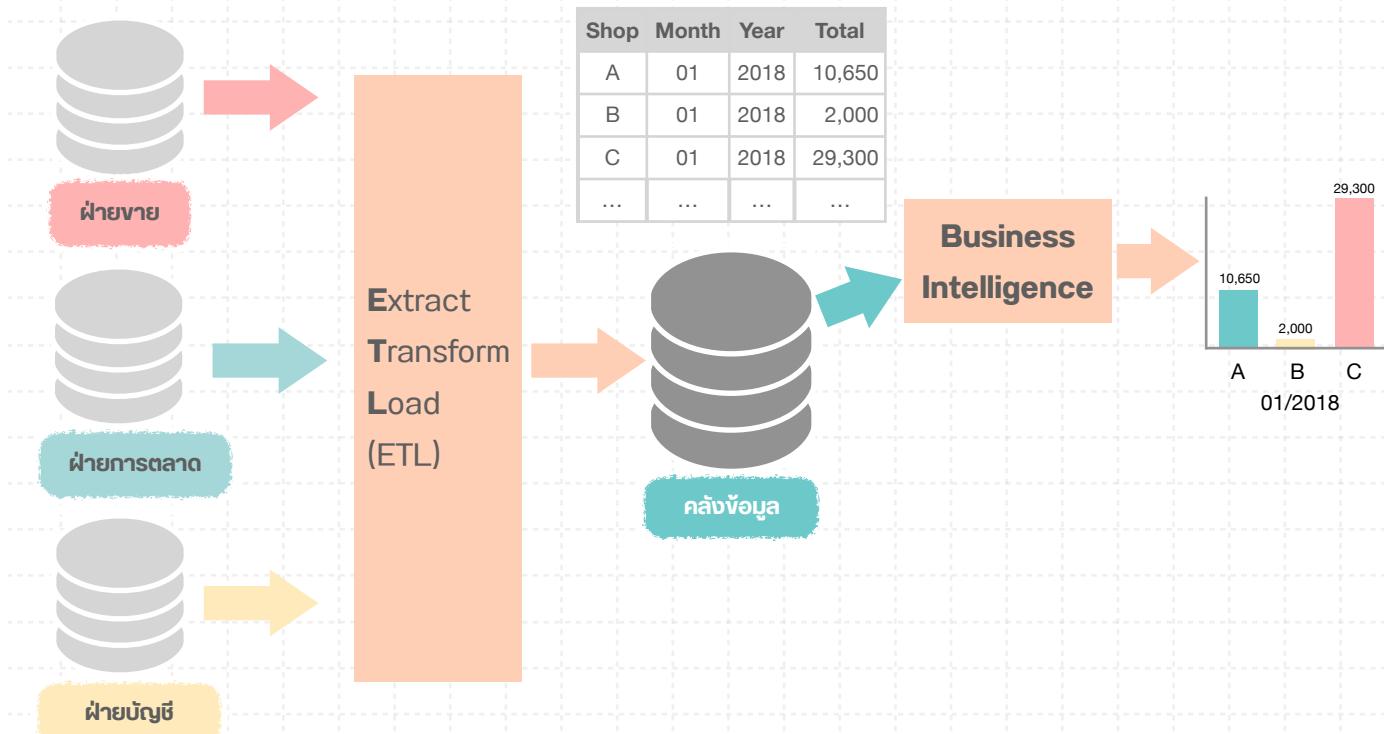
## การสรุป (summary)

ข้อมูลออกมา เช่น

- ยอดขายในแต่ละเดือน
- ยอดขายในแต่ละสาขา

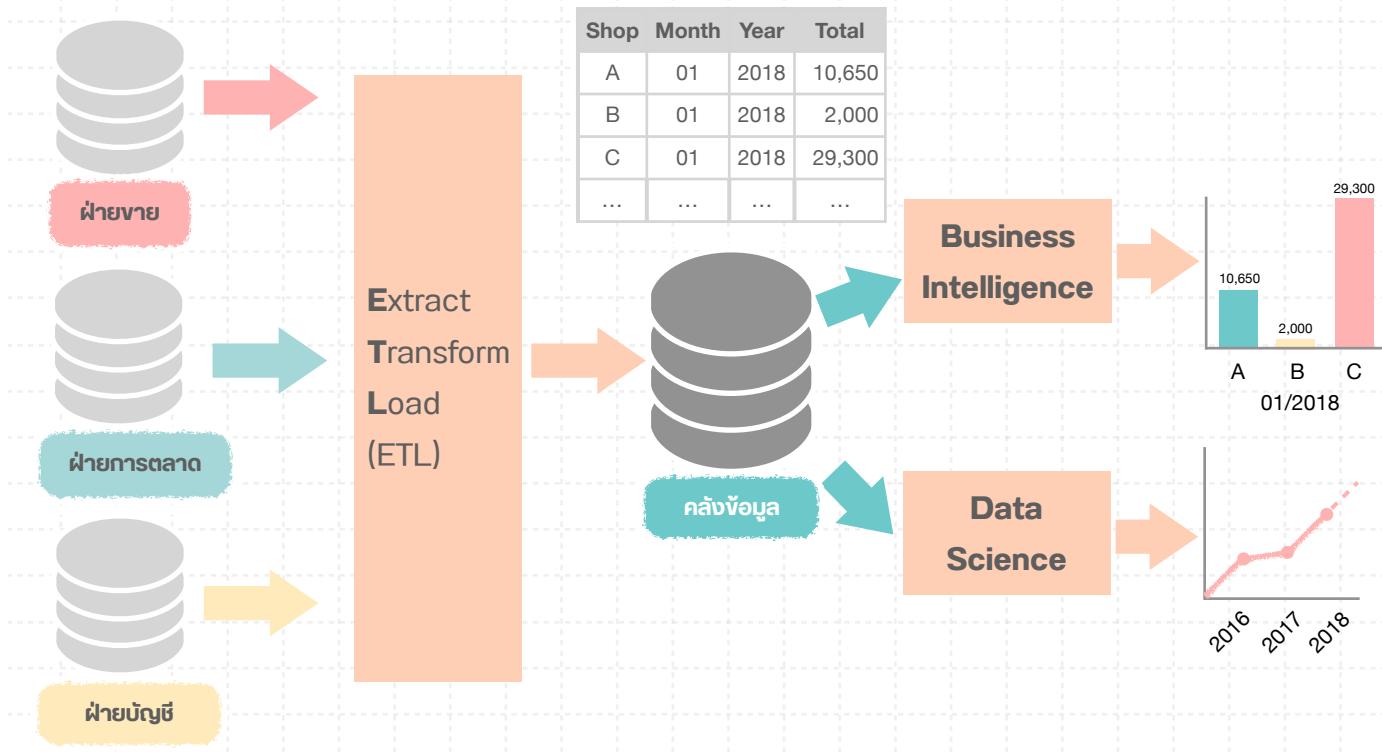
# Data Storage: Data Warehouse

สำหรับการบรรยายในหลักสูตร M. Sc. in Artificial Intelligence for Business Analytics, KMITL



# Data Storage: Data Warehouse

สำหรับการบรรยายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL



# Data Storage: Data Warehouse

สำหรับการบรรยายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL

- **Database**

- **ฐานข้อมูลใช้ในการจัดเก็บข้อมูล** ลดความซ้ำซ้อนของข้อมูล เป็นการจัดเก็บ เพื่อแก้ไข และลบข้อมูล

- **Data Warehouse**

- **คลังข้อมูลรวมข้อมูลจากหลายๆ ฐานข้อมูล** แปลงข้อมูลให้มีความเหมือนกัน หมายความว่า สำหรับการเรียกดู (view) เพื่อสร้างรายงานสรุป

- **Data Science/Machine Learning**

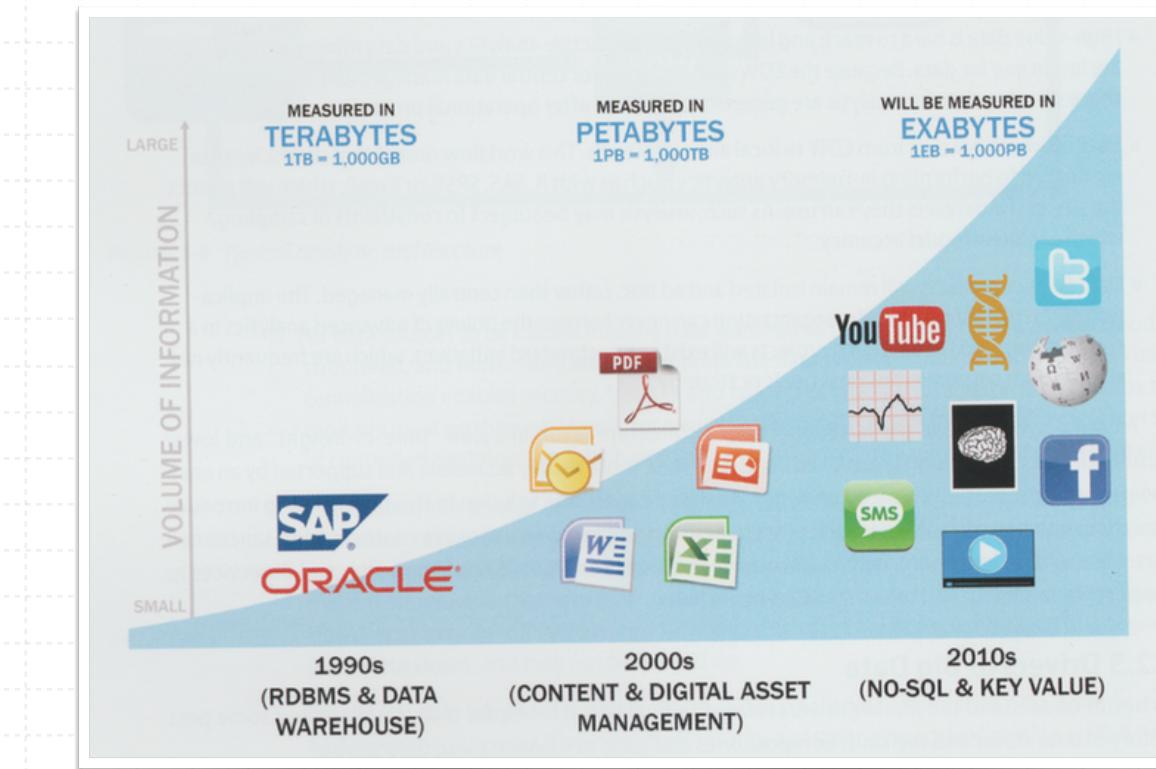
- เป็นการ **เรียนรู้จากข้อมูลในเดตเพื่อนำมาสร้างแบบจำลอง** สำหรับใช้ในการพยากรณ์เหตุการณ์ในอนาคต

# Introduction

- สำหรับการบรรยายในหลักสูตร M. Sc. in Artificial Intelligence for Business Analytics, KMITL
- ตัวอย่างการใช้ Big Data Analytics ในชีวิตประจำวัน
  - ลักษณะของข้อมูล
  - การจัดเก็บข้อมูล (Data Management)
    - ฐานข้อมูล (Database)
    - คลังข้อมูล (Data Warehouse)
  - **การจัดการข้อมูลขนาดใหญ่ (Big Data)**
  - ตัวอย่างการประยุกต์ใช้การวิเคราะห์ข้อมูลด้วยเทคโนโลยี Data Science

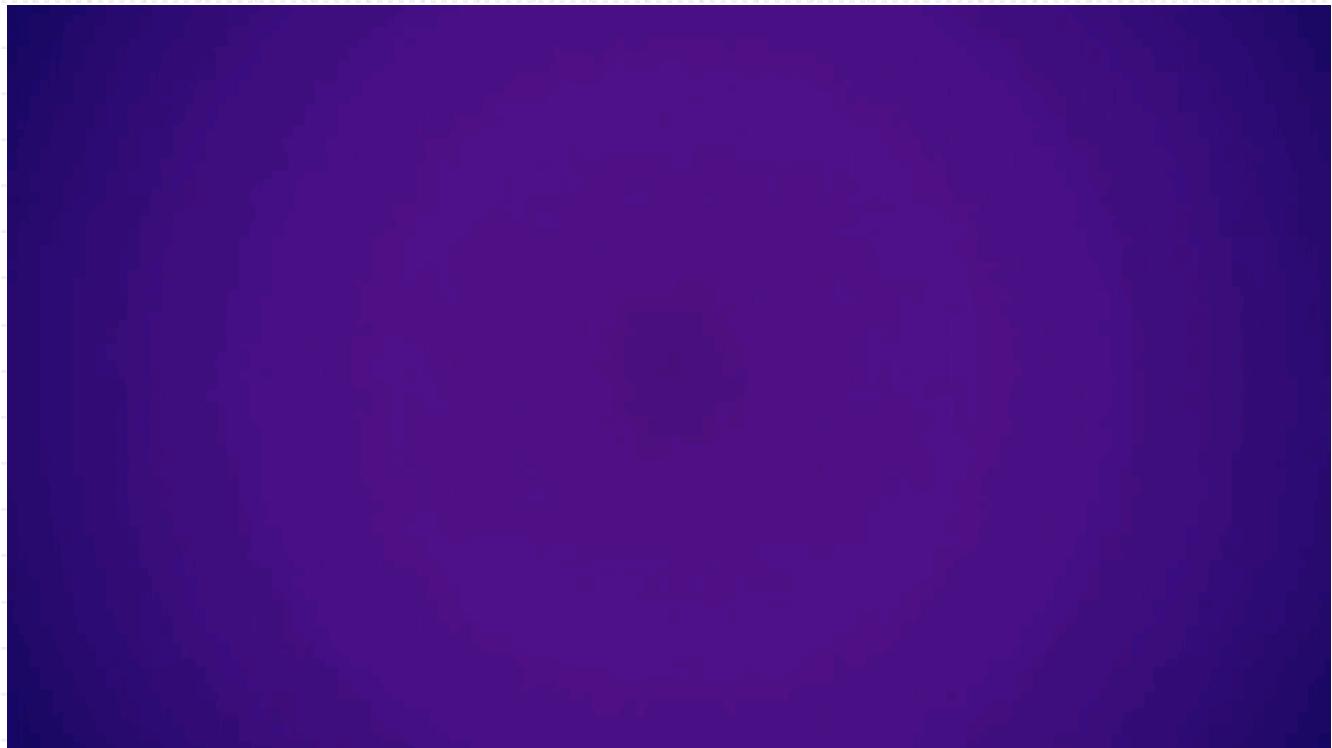
# Data Evolutions

สำหรับการบรรยายในหลักสูตร M. Sc. in Artificial Intelligence for Business Analytics, KMITL



# What is Big Data?

สำหรับการบรรยายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL



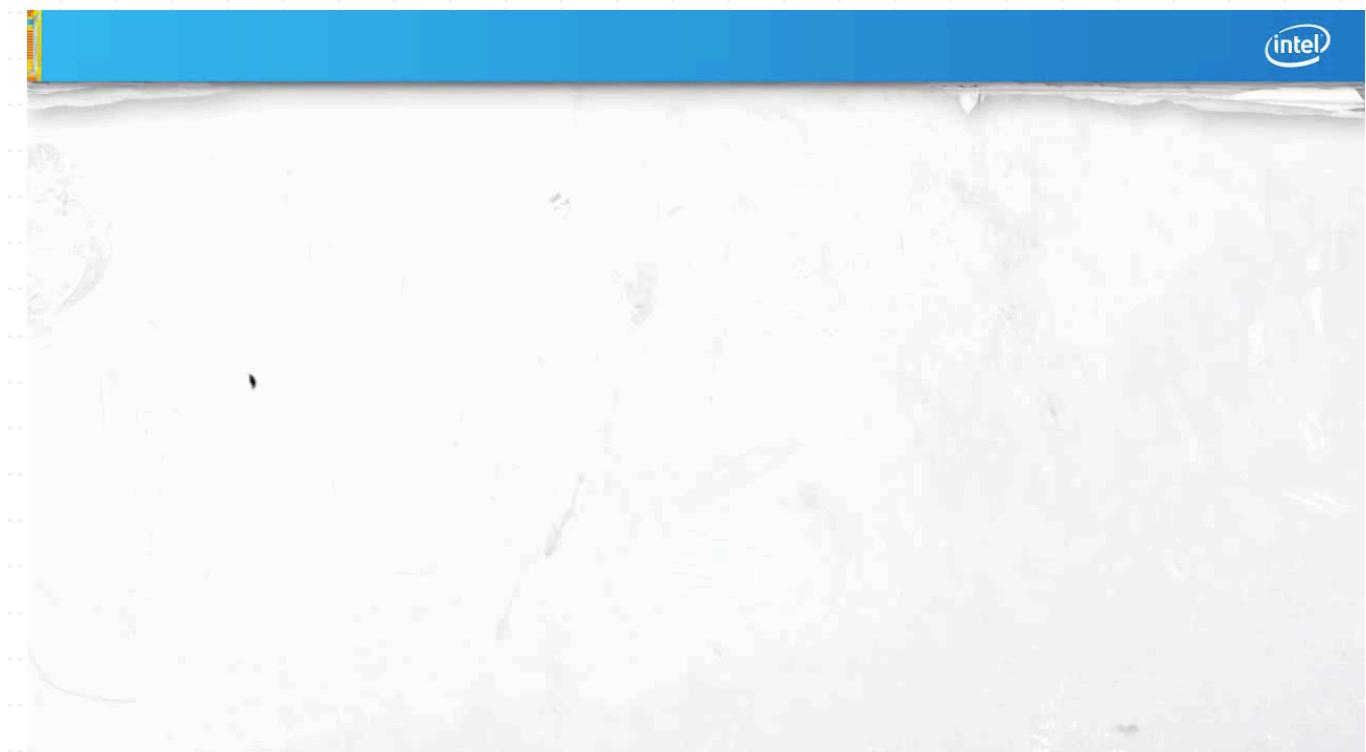
source: <https://www.youtube.com/watch?v=TzxmjbL-i4Y>

<http://facebook.com/datacube.th>

95

# What is Big Data?

สำหรับการบรรยายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL



source: <http://www.intel.com/content/www/us/en/big-data/big-data-101-animation.html#>

<http://facebook.com/datacube.th>

96

# What is Big Data?

- สำหรับการบรรยายในหลักสูตร M. Sc. in Artificial Intelligence for Business Analytics, KMITL
- Big Data ใช้กับด้วย 3 V

- Volume**

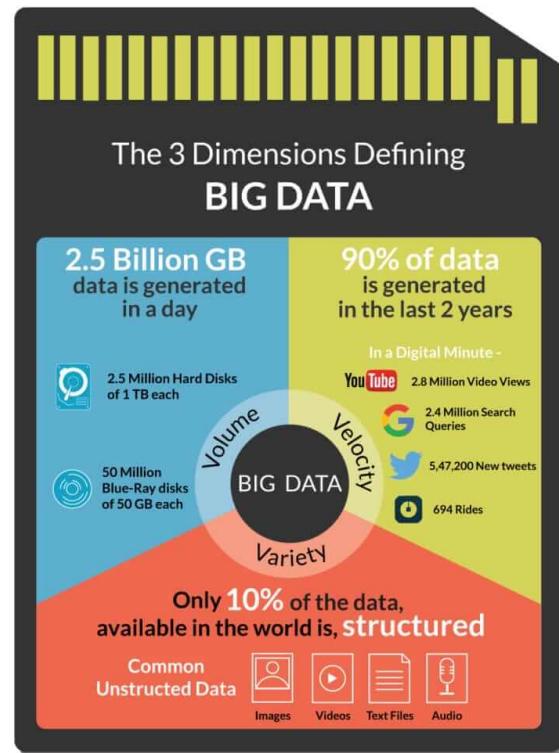
- ข้อมูลมีจำนวนเพิ่มขึ้นอย่างมหาศาล

- Velocity**

- ข้อมูลเพิ่มขึ้นอย่างรวดเร็ว

- Variety**

- ข้อมูลมีความหลากหลายมากขึ้น



source: <https://upxacademy.com/beginners-guide-to-big-data/>

<http://facebook.com/datacube.th>

97

# What is Big Data?

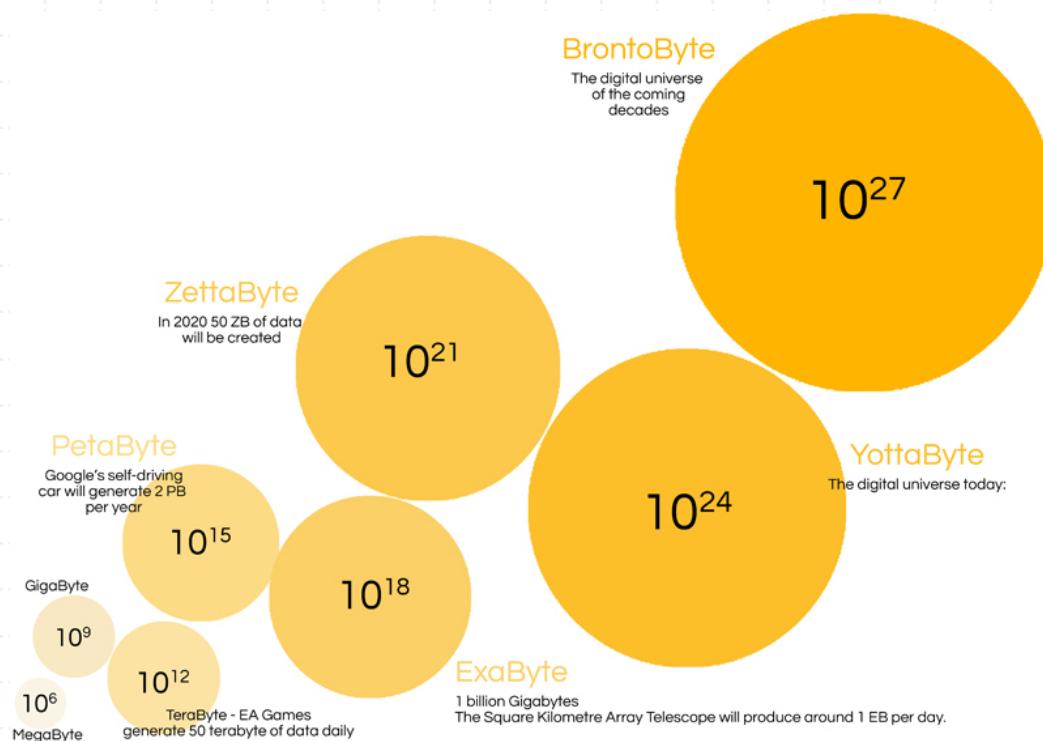
สำหรับการบรรยายในหลักสูตร M. Sc. in Artificial Intelligence for Business Analytics, KMITL

- Huge volume of data**

- ข้อมูลมีขนาดใหญ่มากๆ เช่น มีจำนวนเป็นพันล้านแถว (billion row) หรือ เป็นล้านคอลัมน์ (million columns)

# Big Data: Volume

สำหรับการสัมมนา M: Sc. in Artificial Intelligence for Business Analytics, KMITL



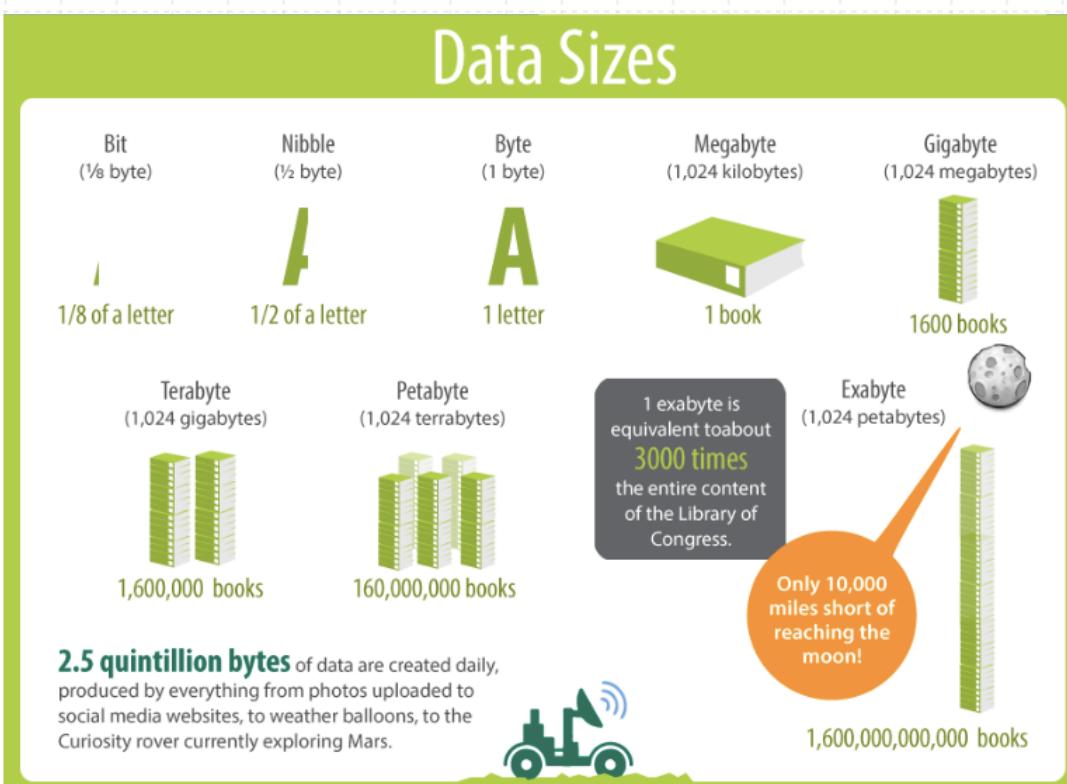
source: <https://datafloq.com/read/infographic/226>

<http://facebook.com/datacube.th>

99

# Big Data: Volume

สำหรับการสัมมนา M: Sc. in Artificial Intelligence for Business Analytics, KMITL



source: <https://www.adeptia.com>

<http://facebook.com/datacube.th>

100

# What is Big Data?

- Huge volume of data

- ข้อมูลมีขนาดใหญ่มากๆ เช่น มีจำนวนเป็นพันล้านแถว (billion row) หรือ เป็นล้านคอลัมน์ (million columns)
  - **Speed of new data creation and growth**
    - ข้อมูลเกิดขึ้นอย่างรวดเร็วมากๆ

Copyright © 2025 Cube Analytics Consulting Co., Ltd.

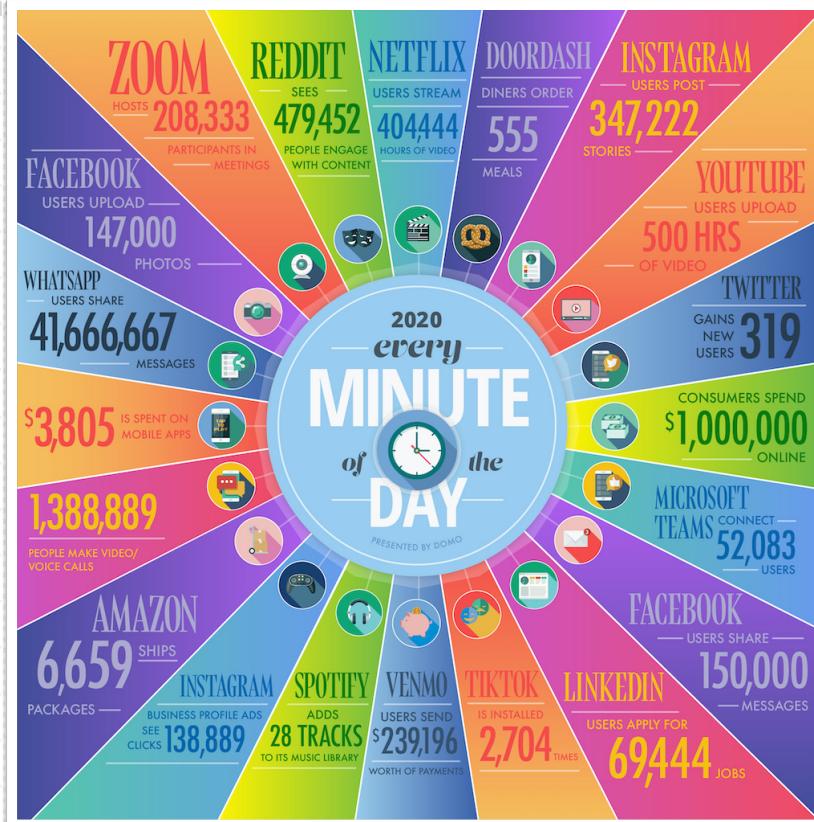
<http://facebook.com/datacube.th>

- 101 -

# Big Data: Velocity

ສຳຫຼັບກາ

CS: KMITL



**source:** <https://web-assets.domo.com/blog/wp-content/uploads/2020/08/20-data-never-sleeps-8-final-01-Resize.jpg>

1

- 102 -

# What is Big Data?

สำหรับการบรรยายในหลักสูตร M. Sc. in Artificial Intelligence for Business Analytics, KMITL

- Huge volume of data**

- ข้อมูลมีขนาดใหญ่มากๆ เช่น มีจำนวนเป็นพันล้านแถว (billion row) หรือ เป็นล้านคอลัมน์ (million columns)

- Speed of new data creation and growth**

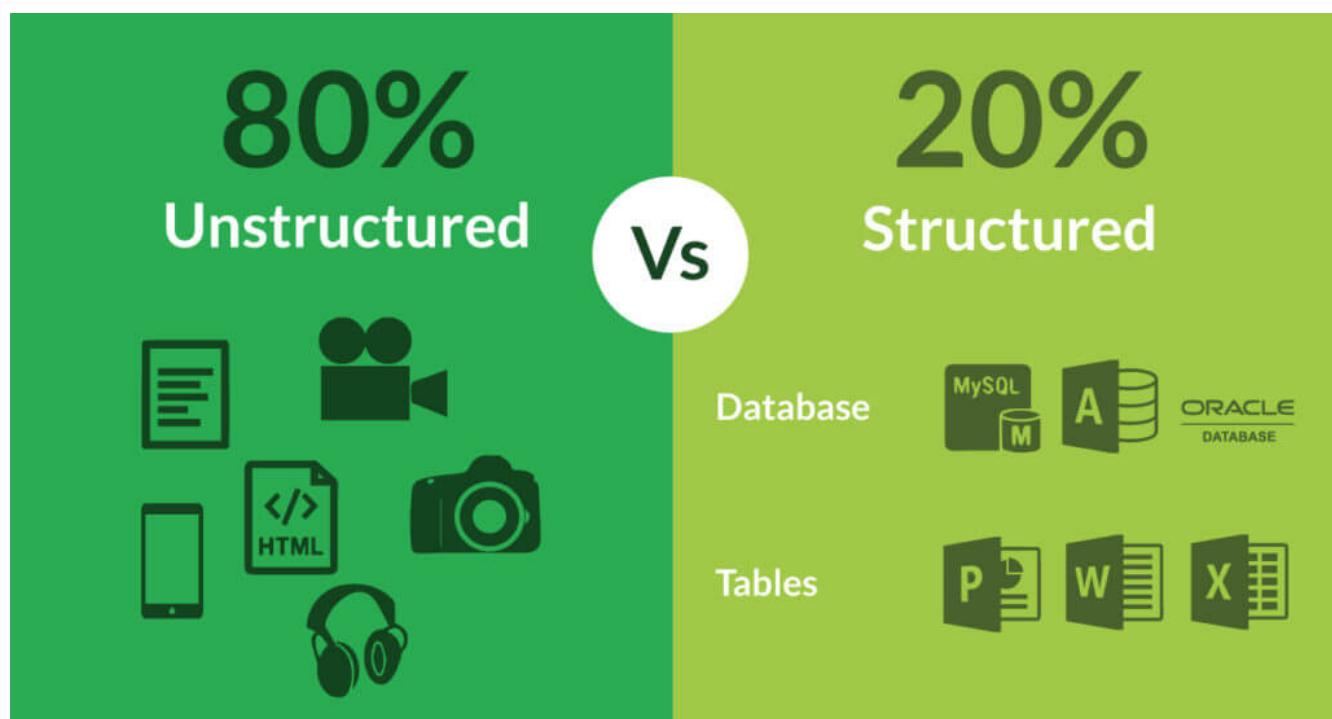
- ข้อมูลเกิดขึ้นอย่างรวดเร็วมากๆ

- Complexity of data types and structures**

- ข้อมูลมีความหลากหลาย ไม่ได้อยู่ในรูปแบบของตารางเท่านั้น อาจจะเป็นรูปแบบของข้อความ (text) รูปภาพ (images) หรือ วิดีโอ (video clip)

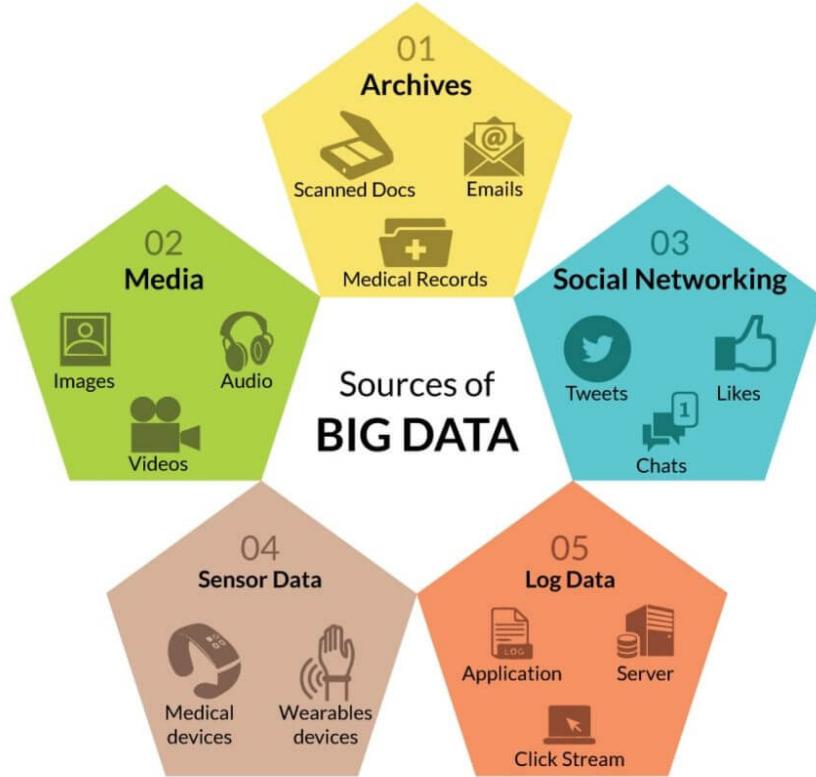
## Big Data: Variety

สำหรับการบรรยายในหลักสูตร M. Sc. in Artificial Intelligence for Business Analytics, KMITL



# Big Data: Variety

สำหรับการสูงสุด M: Sc. in Artificial Intelligence for Business Analytics, KMITL



source: <https://upxacademy.com/beginners-guide-to-big-data/>

<http://facebook.com/datacube.th>

105

## What is Big Data?

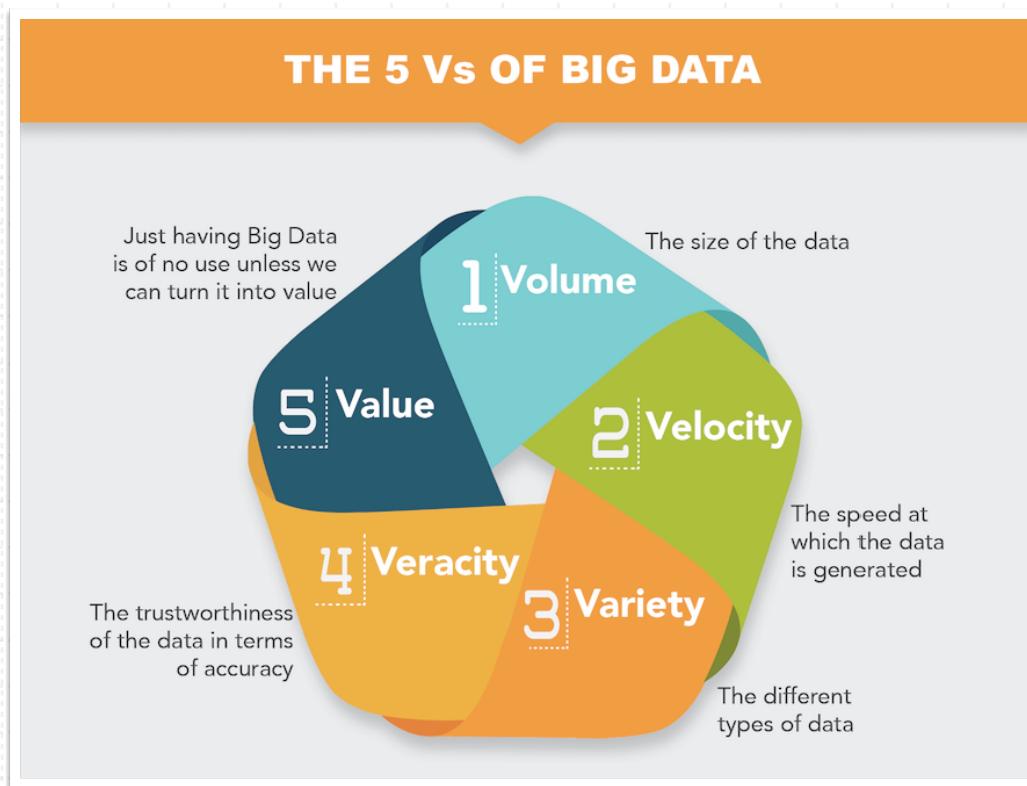
สำหรับการสูงสุด M: Sc. in Artificial Intelligence for Business Analytics, KMITL

ดังนั้น Big Data จึงหมายความว่า

“**ข้อมูลที่มี**หลากหลายรูปแบบ**  
ซึ่งก่อตัวขึ้นอย่าง**รวดเร็ว**จนกลายเป็น**  
**ข้อมูลที่มี**ขนาดมหาศาล****”

# What is Big Data?

สำหรับการบรรยายในหลักสูตร M. Sc. in Artificial Intelligence for Business Analytics, KMITL



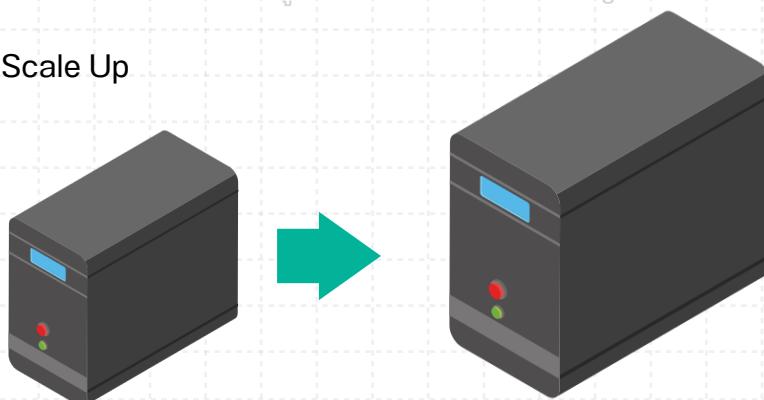
source: <http://dataeconomy.com/2014/08/infographic-how-to-explain-big-data-to-your-grandmother/> /facebook.com/datacube.th

107

## Scale Up v.s. Scale Out

สำหรับการบรรยายในหลักสูตร M. Sc. in Artificial Intelligence for Business Analytics, KMITL

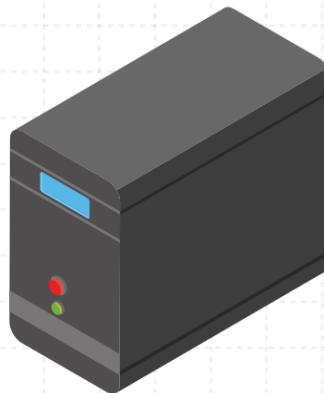
Scale Up



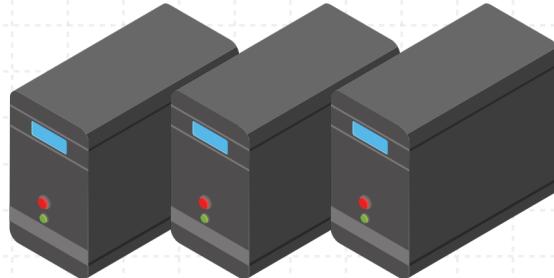
# Scale Up v.s. Scale Out

สำหรับการขยายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL

Scale Up



Scale Out



# Big Data Technology

สำหรับการขยายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL

- เนื่องจากข้อมูลมีขนาดใหญ่มากจนไม่สามารถประมวลผลด้วยเครื่อง

คอมพิวเตอร์เครื่องเดียวได้จึงมีเทคโนโลยีที่ช่วยประมวลผลแบบขนาน (parallel) เกิดขึ้น เช่น

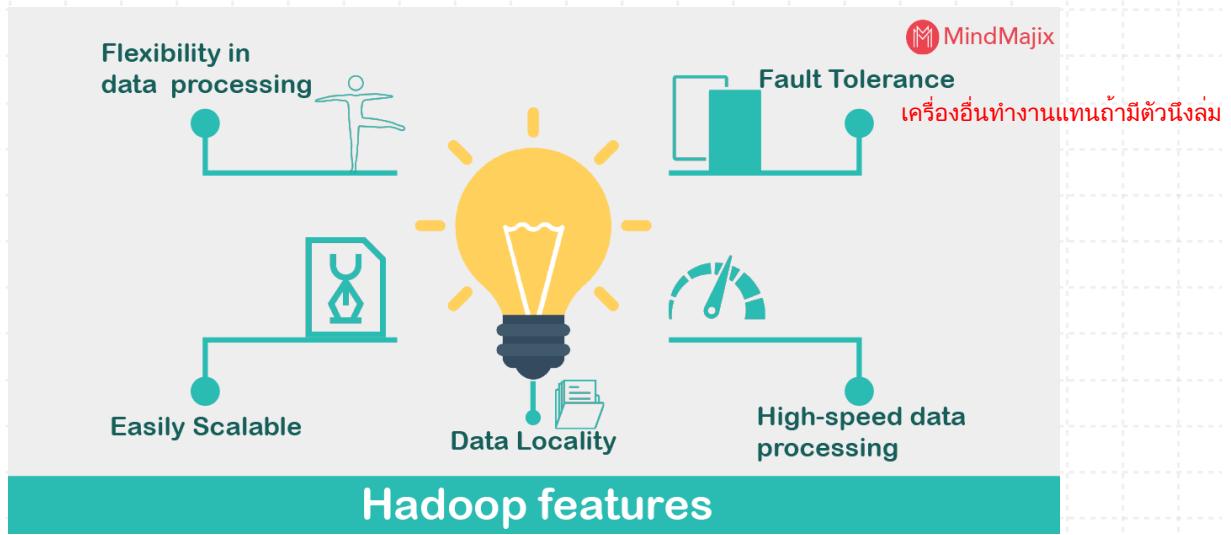
- MapReduce** เป็นเทคโนโลยีที่แบ่งข้อมูลและส่งไปประมวลผลในเครื่องคอมพิวเตอร์หลายๆ เครื่องๆ พร้อมกัน พัฒนาขึ้นจากวิศวกรของบริษัท Google
- Hadoop** เป็นระบบที่นำแนวคิดของ MapReduce มาพัฒนาต่อยอดร่วมกับ การจัดการไฟล์แบบขนาน (HDFS) และการทำงานแบบขนานต่างๆ พัฒนาขึ้นจากวิศวกรของบริษัท Yahoo!

# Big Data Technology

สำหรับการสูงสุด M: Sc. in Artificial Intelligence for Business Analytics, KMITL

- What is Hadoop?

- Hadoop is an open source software framework which is designed to **store** the enormous volumes of data sets in a distributed way on large clusters of the commodity.

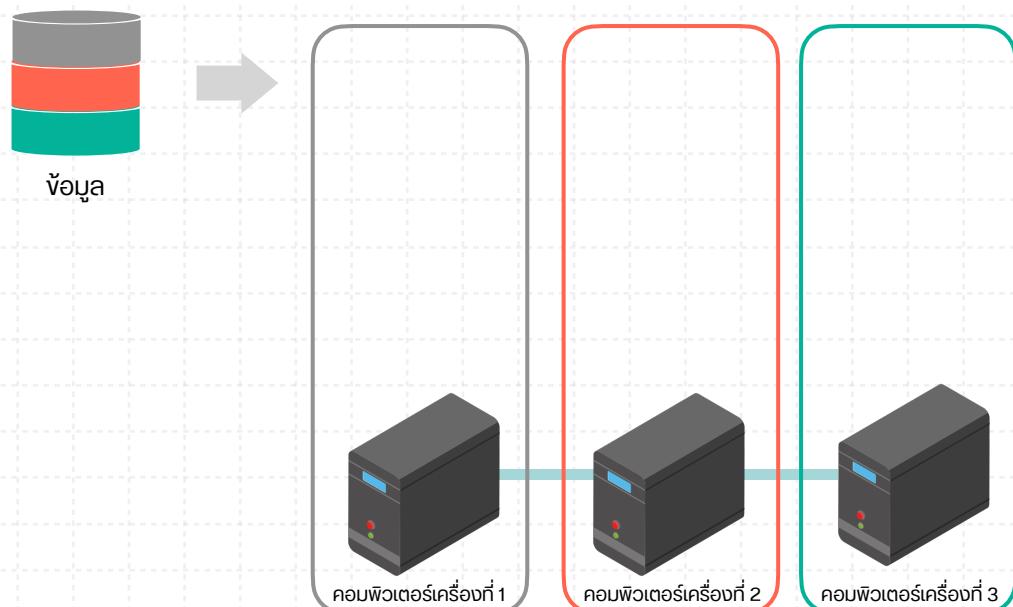


# Big Data Technology

สำหรับการสูงสุด M: Sc. in Artificial Intelligence for Business Analytics, KMITL

- How Does Hadoop Work?

ค่อนขึ้ป์ต์ของ Hadoop

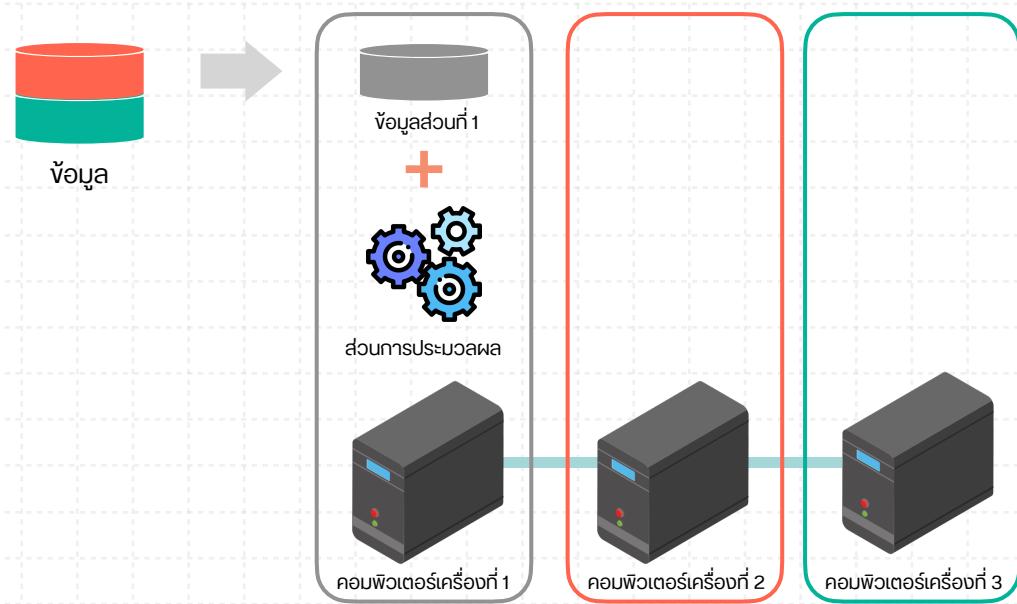


# Big Data Technology

สำหรับการบรรยายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL

- How Does Hadoop Work?

ค่อนซึป์ต์ของ Hadoop

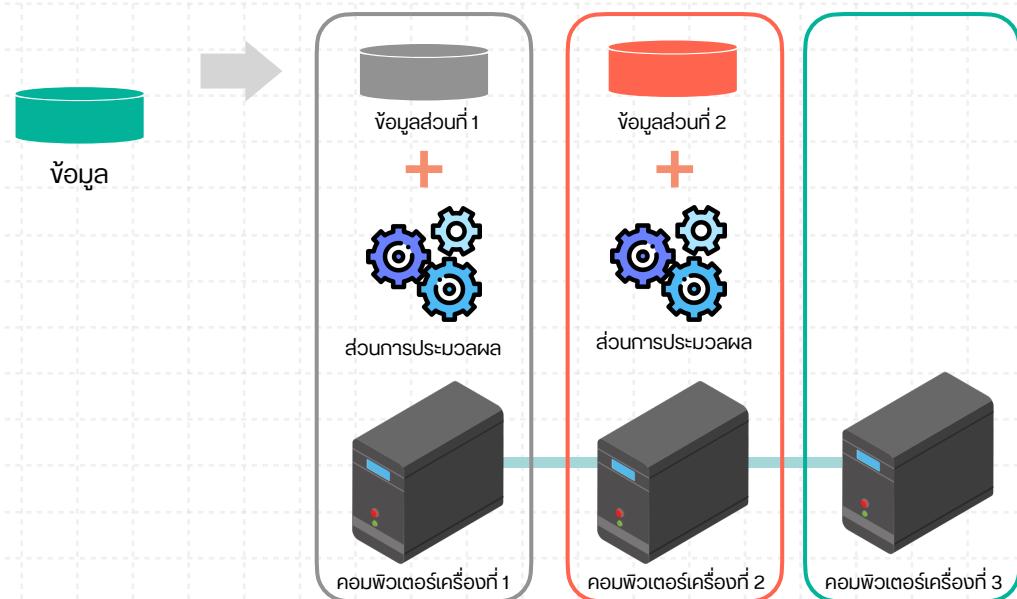


# Big Data Technology

สำหรับการบรรยายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL

- How Does Hadoop Work?

ค่อนซึป์ต์ของ Hadoop

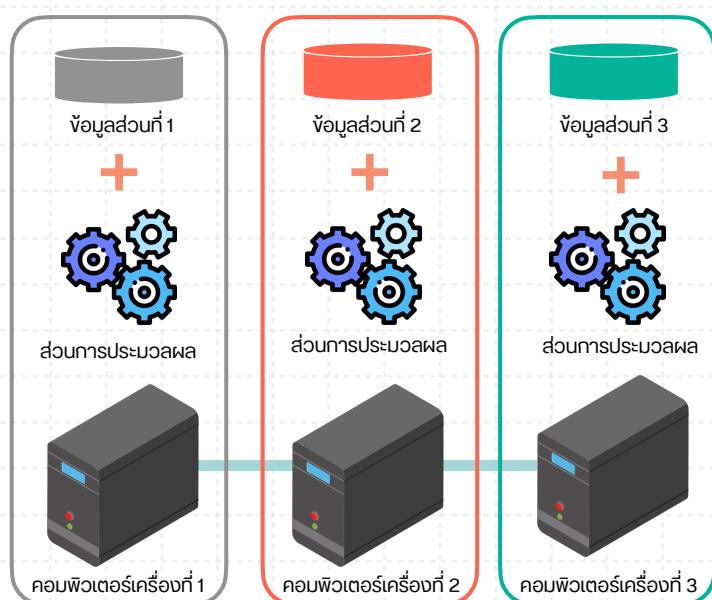


# Big Data Technology

สำหรับการบรรยายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL

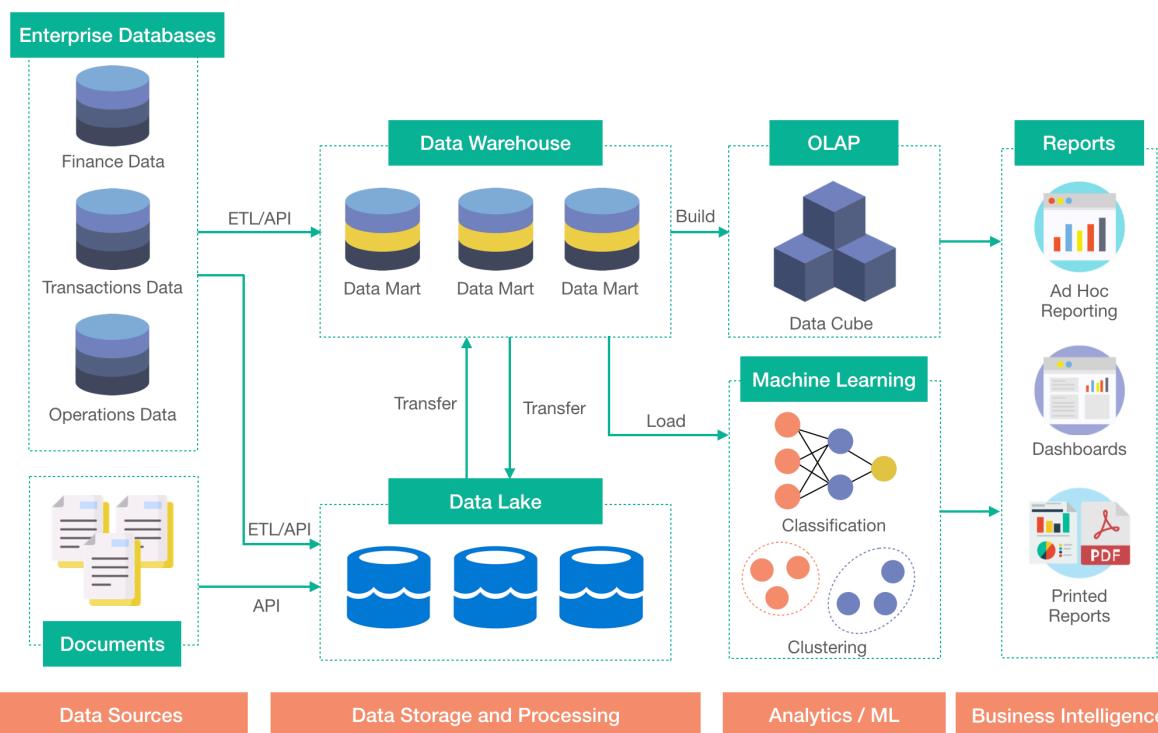
- How Does Hadoop Work?

ค่อนข้างเป็นท่อง Hadoop



# Enterprise Big Data Framework

สำหรับการบรรยายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL



# Introduction

- สำหรับการบรรยายในหลักสูตร M. Sc. in Artificial Intelligence for Business Analytics, KMITL
- ตัวอย่างการใช้ Big Data Analytics ในชีวิตประจำวัน
  - ลักษณะของข้อมูล
  - การจัดเก็บข้อมูล (Data Management)
    - ฐานข้อมูล (Database)
    - คลังข้อมูล (Data Warehouse)
  - การจัดการข้อมูลขนาดใหญ่ (Big Data)
  - **ตัวอย่างการประยุกต์ใช้การวิเคราะห์ข้อมูลด้วยเทคโนโลยี Data Science**

# Data Science Use Cases

สำหรับการบรรยายในหลักสูตร M. Sc. in Artificial Intelligence for Business Analytics, KMITL



## Customer Lifetime Value

แบ่งแยกลูกค้าตามมูลค่าที่สร้างให้กับบริษัท



## Next Best Action

นำเสนอสินค้า/บริการให้ตรงใจและเหมาะสมกับช่วงเวลาที่ต้องการ



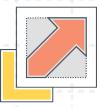
## Customer Segmentation

แบ่งกลุ่มลูกค้าตามพฤติกรรมต่างๆ เพื่อจะได้เข้าใจลูกค้าได้มากขึ้น



## Propensity to Buy

คาดการณ์ว่าลูกค้าคนใดบ้างที่มีโอกาสซื้อสินค้า/บริการ



## Up- and Cross-Selling

แนะนำสินค้า/บริการเพื่อให้ลูกค้าซื้อมากขึ้นหรือเพียงขึ้น



## Churn Prevention

คาดการณ์ว่าลูกค้าคนใดบ้างที่มีโอกาสจะยกเลิกการใช้บริการ

# Data Science Use Cases

สำหรับการบรรยายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL



## Fraud Detection

ค้นหาและคาดการณ์การใช้งานที่ผิดปกติ



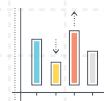
## Risk Management

จัดการกับความเสี่ยงต่างๆ



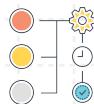
## Demand Forecast

พยากรณ์ยอดขายหรือยอดผลิตสินค้าต่างๆ



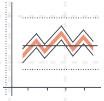
## Price Optimization

ปรับเปลี่ยนราคาของสินค้า/บริการให้สมดุล



## Quality Assurance

ค้นหาปัญหาที่เกี่ยวข้องกับคุณภาพของสินค้าก่อนการจำหน่าย



## Predictive Maintenance

คาดการณ์ความผิดปกติของเครื่องจักรก่อนปัญหาจะเกิดขึ้น

source: [www.rapidminer.com/](http://www.rapidminer.com/)

# Data Science Use Cases

สำหรับการบรรยายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL

| Use-cases                     | Retail | Banking | Telecommunication | Insurance | Manufacturing |
|-------------------------------|--------|---------|-------------------|-----------|---------------|
| Customer Lifetime Value (CLV) | ●      | ●       | ●                 | ●         |               |
| Customer Segmentation         | ●      | ●       | ●                 | ●         |               |
| Up- and Cross-Selling         | ●      | ●       | ●                 | ●         |               |
| Next Best Action              | ●      | ●       | ●                 | ●         |               |
| Propensity to Buy             | ●      | ●       | ●                 | ●         |               |
| Churn Prevention              | ●      | ●       | ●                 | ●         |               |
| Fraud Detection               |        | ●       |                   | ●         |               |
| Demand Forecast               | ●      |         | ●                 |           | ●             |
| Quality Assurance             |        |         |                   |           | ●             |
| Risk Management               |        | ●       |                   | ●         |               |
| Price Optimization            | ●      |         |                   |           |               |
| Predictive Maintenance        |        |         |                   |           | ●             |

# Data Science Applications

- สำหรับการบรรยายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL
- ตัวอย่างการนำ Data Science ไปใช้งาน



Copyright © 2025 Cube Analytics Consulting Co., Ltd.

source: <http://www.youtube.com/watch?v=f2Kji24833Y>  
<http://facebook.com/datacube.th>

121

# Data Science Applications

สำหรับการบรรยายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL

- คาดการณ์การตั้งครรภ์



- ห้าง Target ทำการวิเคราะห์พฤติกรรมการซื้อสินค้าของลูกค้าเพื่อถ่าย
- พบรูปแบบ (pattern) ว่า ถ้ามีการซื้อวัตถุนิยม ซื้ออาหารบำรุง หรือ ซื้อตุ๊กตาเด็กเพิ่ม ลูกค้าจะเริ่มตั้งครรภ์
- Target จะส่ง promotion ให้ลูกค้าเหล่านั้น

source: <http://www.youtube.com/watch?v=f2Kji24833Y>

Copyright © 2025 Cube Analytics Consulting Co., Ltd.

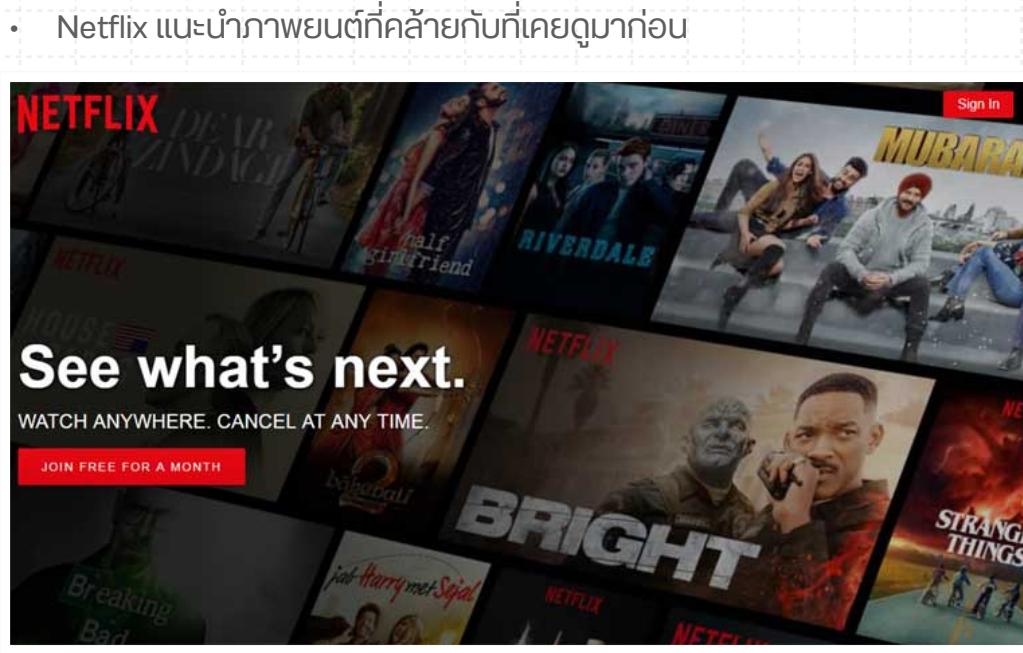
<http://facebook.com/datacube.th>

122

# Data Science Applications

สำหรับการบรรยายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL

- แนะนำสินค้าที่เกี่ยวข้อง



source: <https://indianexpress.com/article/technology/social/netflix-partnering-with-whatsapp-in-india-will-send-content-recommendations-5002335/>

Copyright © 2025 Cube Analytics Consulting Co., Ltd.

<http://facebook.com/datacube.th>

123

# Data Science Applications

สำหรับการบรรยายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL

- Google Self-Driving Car



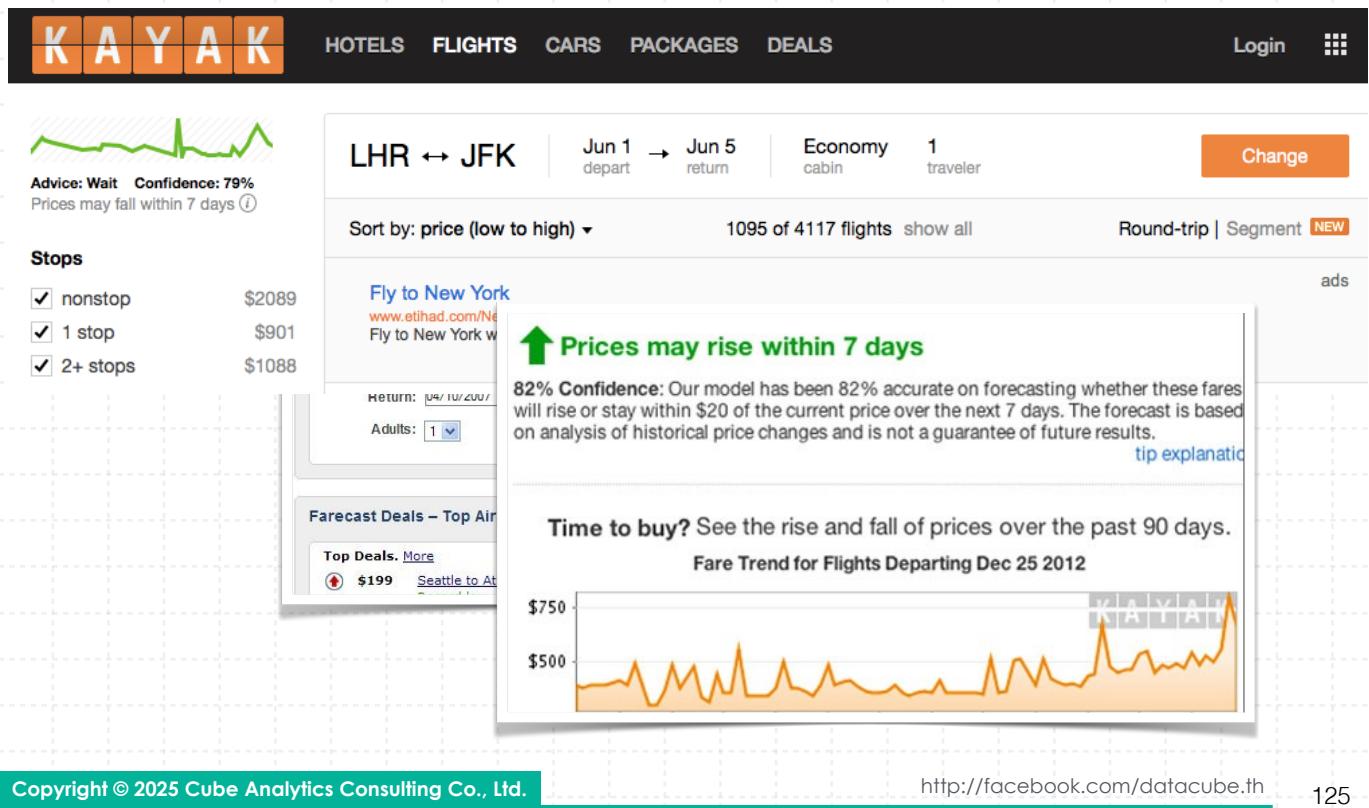
source: <https://www.youtube.com/watch?v=8fjNSUWX7nQ>

<http://facebook.com/datacube.th>

124

# Data Science Applications

- สำหรับการบรรยายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL
- แนวโน้มราคาตั๋วเครื่องบิน

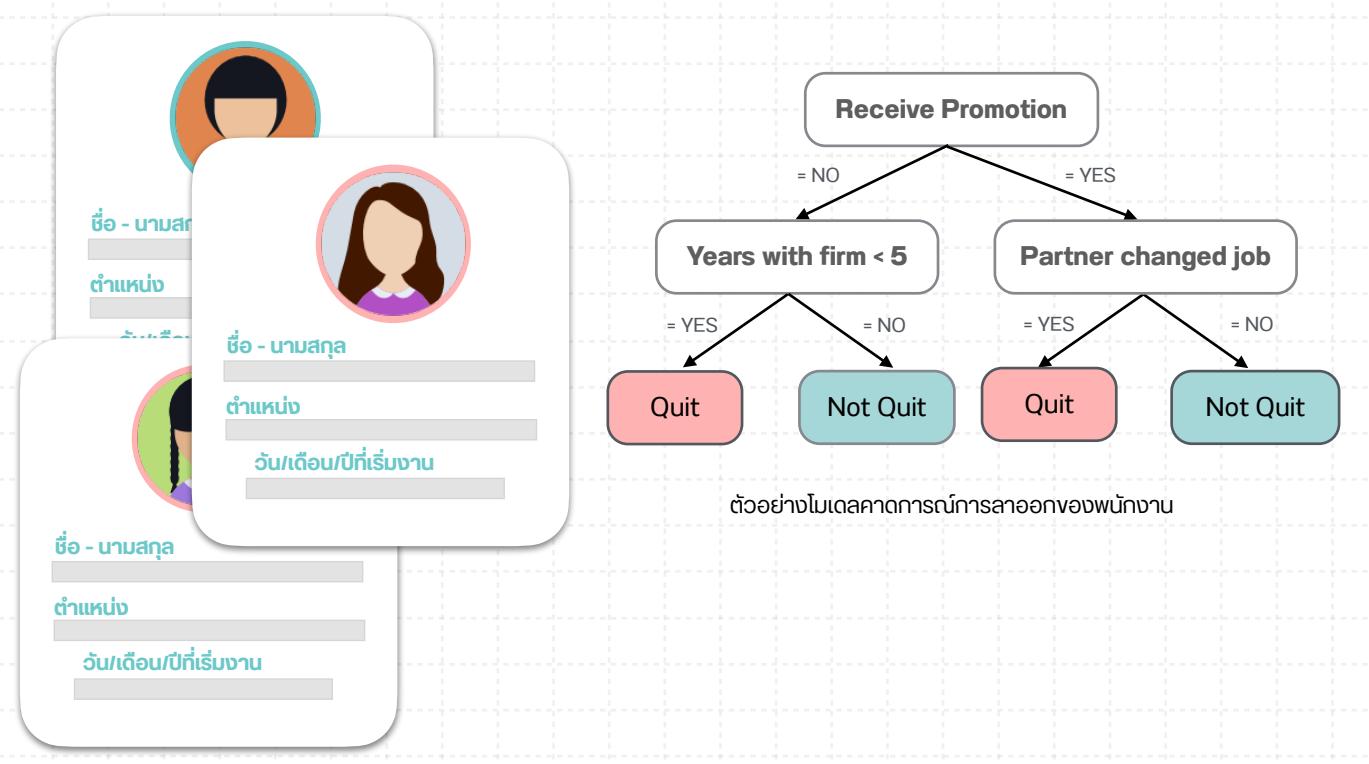


The screenshot shows the Kayak website interface for flight search. At the top, there's a navigation bar with 'HOTELS', 'FLIGHTS', 'CARS', 'PACKAGES', 'DEALS', 'Login', and a grid icon. The main search area displays a flight from London Heathrow (LHR) to New York (JFK) departing on June 1st and returning on June 5th, for one traveler in economy cabin. A green line graph at the top left indicates a price trend with the text 'Advice: Wait Confidence: 79% Prices may fall within 7 days'. Below the search form, there are filters for 'Stops' (nonstop, 1 stop, 2+ stops), a sort option ('Sort by: price (low to high)'), and a summary of 1095 of 4117 flights. A callout box for a 'Fly to New York' deal is visible. To the right, a green box highlights a price forecast: 'Prices may rise within 7 days' with an 82% confidence level. A chart titled 'Fare Trend for Flights Departing Dec 25 2012' shows price fluctuations over 90 days. The bottom of the page includes copyright information and a Facebook link.

Copyright © 2025 Cube Analytics Consulting Co., Ltd. <http://facebook.com/datacube.th> 125

# Data Science Applications

- สำหรับการบรรยายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL
- คาดการณ์การลาออกจากพนักงาน



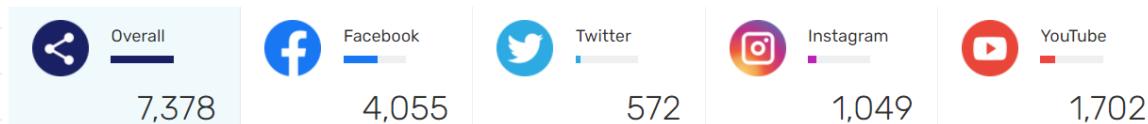
The diagram illustrates a decision tree for predicting employee turnover. The root node is 'Receive Promotion'. If 'NO', it leads to 'Years with firm < 5', which further branches into 'YES' (leading to 'Quit') and 'NO' (leading to 'Not Quit'). If 'YES' at the promotion node, it leads to 'Partner changed job', which further branches into 'YES' (leading to 'Quit') and 'NO' (leading to 'Not Quit'). A note at the bottom states: 'ตัวอย่างในการเดลคาดการณ์การลาออกจากพนักงาน' (Example for predicting employee turnover).

Copyright © 2025 Cube Analytics Consulting Co., Ltd. <http://facebook.com/datacube.th> 126

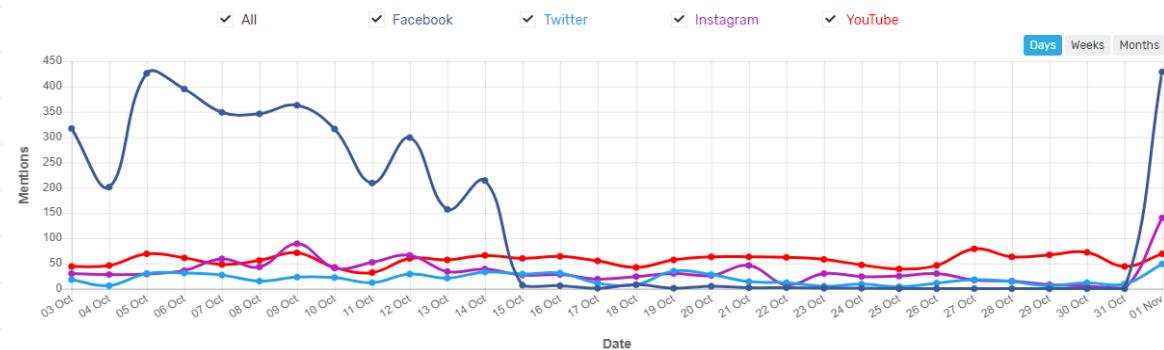
# Data Science Applications

สำหรับการบรรยายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL

- วิเคราะห์ศัพด์ในแง่ต่างๆ จากสังคมออนไลน์



## Data Timeline

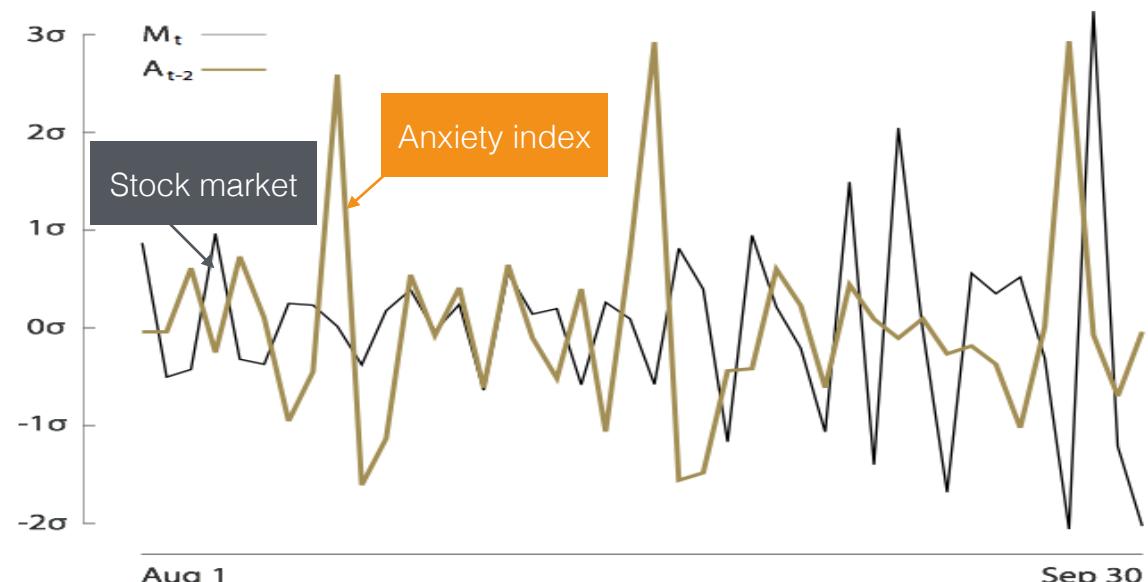


Reference: การวิเคราะห์การกล่าวถึงแนวโน้มการเมืองในประเทศไทย โดย Vivat Boonchun and Surasak Chantarach

# Data Science Applications

สำหรับการบรรยายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL

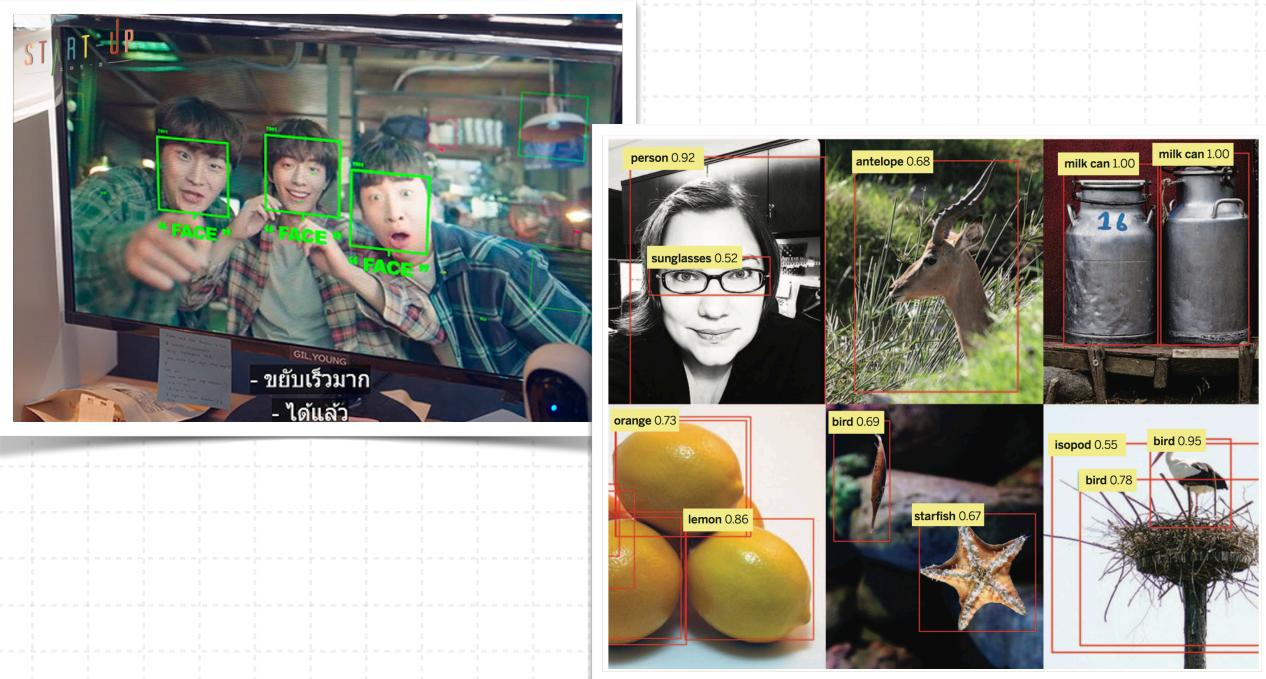
- คาดการณ์แนวโน้มของตลาดหุ้น



Reference: Eric Gilbert, Karrie Karahalios: Widespread Worry and the Stock Market. ICWSM 2010

# Data Science Applications

- สำหรับการบรรยายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL  
 • ระบบตรวจจับวัตถุ (Image Recognition)



ที่มา: [https://medium.com/@artisan\\_digital/เรื่องนวัตกรรม-ai-เส้นทางพัฒนาธุรกิจชั้นนำจากเชลล์-start-up-f287ddaa843](https://medium.com/@artisan_digital/เรื่องนวัตกรรม-ai-เส้นทางพัฒนาธุรกิจชั้นนำจากเชลล์-start-up-f287ddaa843)

<https://cs.uwaterloo.ca/~y328yu/mycourses/480-2018/readings/JordanMitchell.pdf>

datacube.th

129

# Data Science Applications

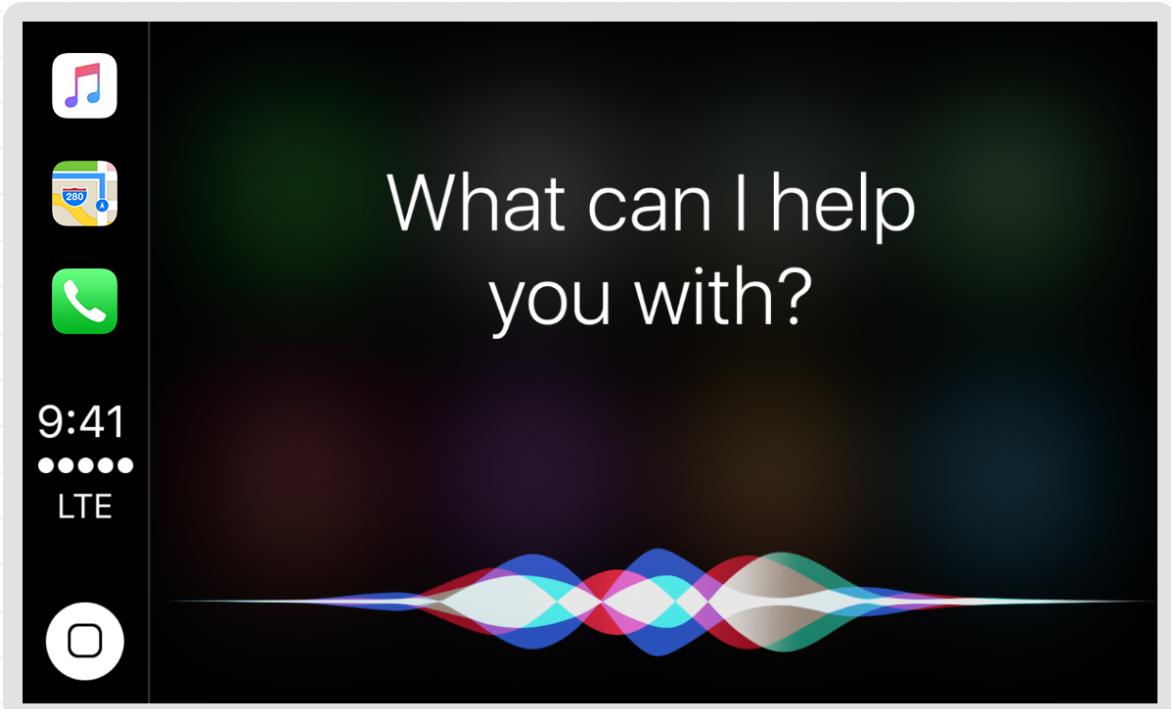
- สำหรับการบรรยายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL  
 • ระบบตรวจจับคนที่ไม่ใส่หน้ากากอนามัย



ที่มา: [https://www.youtube.com/watch?v=S4TUM\\_p6H3E&feature=emb\\_logo](https://www.youtube.com/watch?v=S4TUM_p6H3E&feature=emb_logo)

# Data Science Applications

- สำหรับการสูงสุด M: Sc. in Artificial Intelligence for Business Analytics, KMITL
- ระบบรู้จำเสียง (Speech Recognition)



ที่มา: <https://developer.apple.com/design/human-interface-guidelines/carplay/interaction/voice/>

# Data Science Applications

- สำหรับการสูงสุด M: Sc. in Artificial Intelligence for Business Analytics, KMITL
- ลำโพงอัจฉริยะ

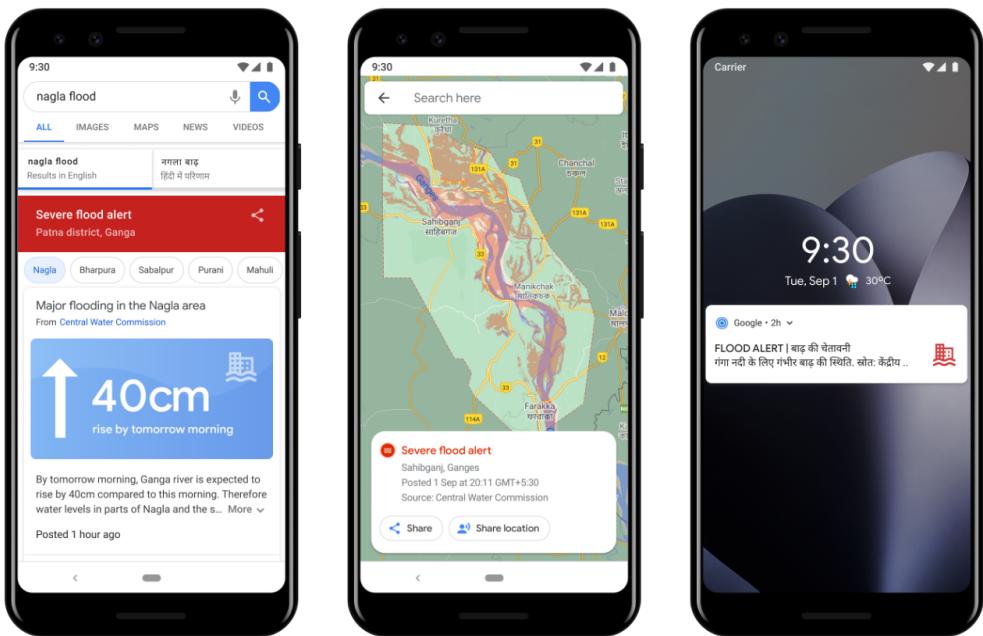


ที่มา: <https://www.zdnet.com/product/amazon-echo/>

<https://ninnygadget.com/review/mi-smart-speaker/>

# Data Science Applications

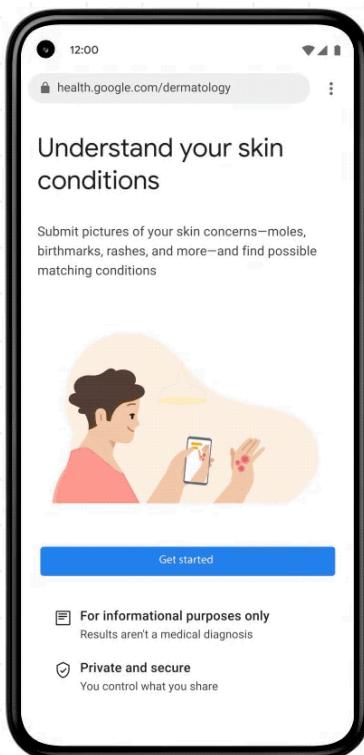
- สำหรับการสูงสุด M: Sc. in Artificial Intelligence for Business Analytics, KMITL
- ระบบคาดการณ์และแจ้งเตือนการเกิดอุทกภัย



ที่มา: <https://blog.google/technology/ai/flood-forecasts-india-bangladesh/>

# Data Science Applications

- สำหรับการสูงสุด M: Sc. in Artificial Intelligence for Business Analytics, KMITL
- ระบบคาดการณ์ความผิดปกติบนผิวหนัง



ที่มา: <https://blog.google/technology/health/ai-dermatology-preview-io-2021/>

# How to get dataset

สำหรับการอภิปรายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL

## UCI Machine Learning Repository

เป็นเว็บไซต์ที่รวบรวมข้อมูลที่ใช้ในการทำวิจัยทาง Machine Learning



Welcome to the UC Irvine Machine Learning Repository!

We currently maintain 588 data sets as a service to the machine learning community. You may [view all data sets](#) through our searchable interface. For a general overview of the Repository, please visit our [About](#) page. For information about citing data sets in publications, please read our [citation policy](#). If you wish to donate a data set, please consult our [donation policy](#). For any other questions, feel free to [contact the Repository librarians](#).

Supported By:  In Collaboration With: 

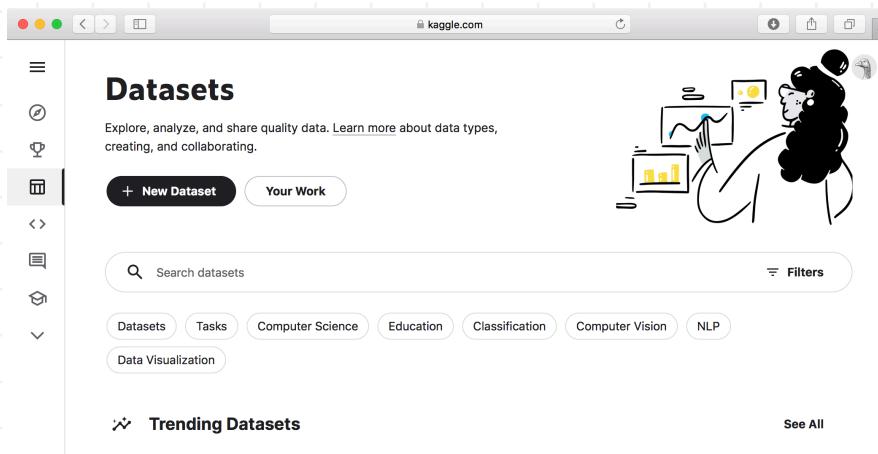
| Latest News:                                                                                                                                                                                                                                                                                                                                                                                                                                                       | Newest Data Sets:                                                                                                                                                                                                                                                                                                                                                                | Most Popular Data Sets (hits since 2007):                                                                                                                                                                                                                                                                    |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 09-24-2018: Welcome to the new Repository admins Dheeru Dua and Efi Karra Taniskidou!<br>04-04-2013: Welcome to the new Repository admins Kevin Bache and Moshe Lichman!<br>03-01-2010: Note from donor regarding Netflix data<br>10-16-2009: Two new data sets have been added.<br>09-14-2009: Several data sets have been added.<br>03-24-2008: New data sets have been added!<br>06-25-2007: Two new data sets have been added: UCI Donor Donations MAGIC Gamma | 04-21-2021:  Synchronous Machine Data Set<br>04-20-2021:  Wikipedia Math Essentials<br>04-20-2021:  Wikipedia Math Essentials | 4007794:  Iris<br>2161539:  Adult<br>1672204:  Wine |

# How to get dataset

สำหรับการอภิปรายในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL

## kaggle.com

เป็นเว็บไซต์ที่รวบรวมการแข่งขัน (competition) ทางด้าน Data Science / Machine Learning



**Datasets**

Explore, analyze, and share quality data. Learn more about data types, creating, and collaborating.

+ New Dataset   Your Work

Search datasets   Filters

Datasets   Tasks   Computer Science   Education   Classification   Computer Vision   NLP   Data Visualization

Trending Datasets   See All

# How to get dataset

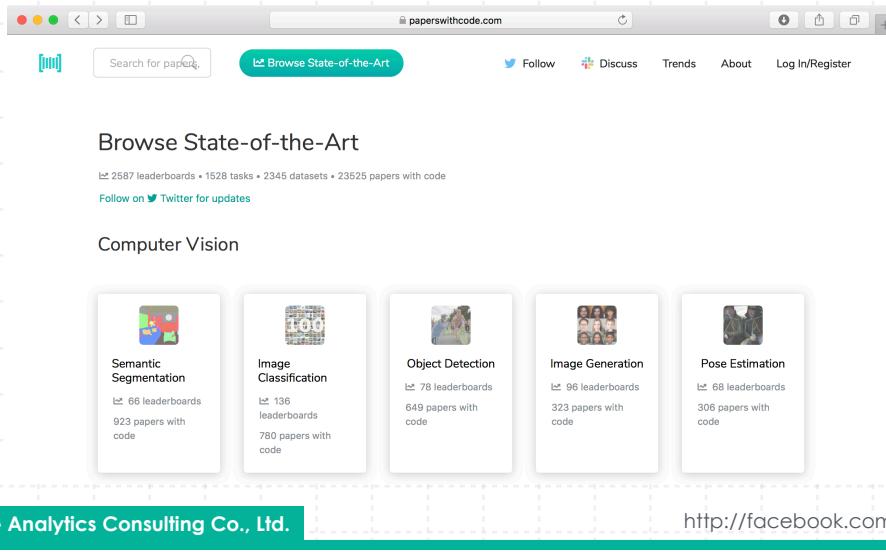
สำหรับการเรียนรู้ในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL

## [paperswithcode.com](https://paperswithcode.com)

เป็นเว็บไซต์ที่รวบรวมงานวิจัยทางด้าน

Machine Learning และมี Code

ที่ใช้ในการทำวิจัยให้ดาวน์โหลดได้โดยครับ



Copyright © 2025 Cube Analytics Consulting Co., Ltd.

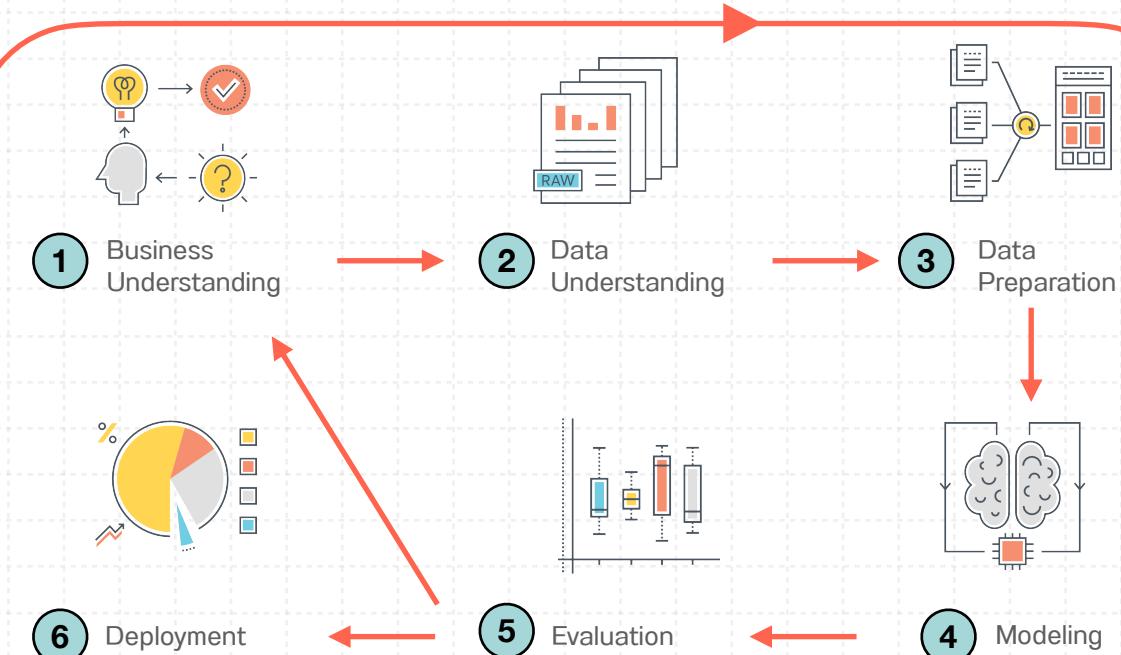
http://facebook.com/datacube.th

137

## CRISP-DM

สำหรับการเรียนรู้ในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL

• CRISP-DM ย่อมาจาก CRoss Industry Standard Process for Data Mining



# CRISP-DM

สำหรับการอัศวินในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL

| STEP                   | DESCRIPTION                                         |
|------------------------|-----------------------------------------------------|
| Business Understanding | Define the project.                                 |
| Data Understanding     | Examine the data; identify problems in the data.    |
| Data Preparation       | Fix problems in the data; create derived variables. |
| Modeling               | Build predictive or descriptive models.             |
| Evaluation             | Assess models; report on the expected effects of    |
| Deployment             | Plan for use of models.                             |

source: Applied Predictive Analytics: Principle and Techniques for the Professional Data Analyst

Copyright © 2025 Cube Analytics Consulting Co., Ltd.

<http://facebook.com/datacube.th>

139

# CRISP-DM

สำหรับการอัศวินในหลักสูตร M: Sc. in Artificial Intelligence for Business Analytics, KMITL

## • 1. Business Understanding

- ทำความเข้าใจกับปัญหา หรือ โอกาสเชิงธุรกิจ
- ระบุ output หรือเป้าหมายที่ต้องการได้จากการวิเคราะห์ข้อมูล
- ตัวอย่างเช่น
  - ทำอย่างไรเพิ่มยอดขายให้กับสินค้าชนิดต่างๆ ได้
  - ต้องการแบ่งกลุ่มนักศึกษาออกตามความสนใจ
  - ทำอย่างไรให้ลูกค้ากลับมาซื้อสินค้าได้อีก
  - อยากรู้ว่าลูกค้าคนใดบ้างมีโอกาสเป็นโภคภารโรง
  - อยากรู้ว่าลูกค้าคนใดบ้างมีโอกาสเป็นโภคภารโรง

# CRISP-DM

สำหรับการบรรยายในหลักสูตร M. Sc. in Artificial Intelligence for Business Analytics, KMITL

## • 2. Data Understanding

- ในขั้นตอนนี้เป็นการ
  - ตรวจสอบข้อมูลที่เกี่ยวข้อง
  - ข้อมูลถูกต้องบ่าเชื่อถือ
  - ข้อมูลที่ได้มีปริมาณมากพอหรือยัง
  - ข้อมูลที่ได้มีความเหมาะสม มีรายละเอียดเพียงพอต่อการนำไปใช้ในการวิเคราะห์
- ตัวอย่าง
  - ข้อมูลการซื้อสินค้าคนแต่ละบุคคล
  - ข้อมูลการลงทะเบียนและผลการศึกษาของนักศึกษา

# CRISP-DM

สำหรับการบรรยายในหลักสูตร M. Sc. in Artificial Intelligence for Business Analytics, KMITL

## • 3. Data Preparation

- ขั้นตอนการเตรียมข้อมูลเป็นขั้นตอนที่ใช้เวลานานที่สุด
- เนื่องจากไม่เดลก์ได้จากการนำตัวไม้บันจงให้ผลลัพธ์ที่ถูกต้องหรือไม่นั้น ขึ้นอยู่กับคุณภาพของข้อมูลที่ใช้ แบ่งออกได้เป็น 3 ขั้นตอนย่อยคือ

### • 3.1 ทำการคัดเลือกข้อมูล (Data Selection)

- กำหนดเป้าหมายก่อนว่าเราจะทำการวิเคราะห์อะไร
- เลือกใช้เฉพาะข้อมูลที่เกี่ยวข้องกับสิ่งที่เราจะทำการวิเคราะห์

# CRISP-DM

สำหรับการบรรยายในหลักสูตร M. Sc. in Artificial Intelligence for Business Analytics, KMITL

- 3. Data Preparation**

- 3.2 ทำการกรองข้อมูล (Data Cleaning)**

- ลบข้อมูลซ้ำซ้อน
- แก้ไขข้อมูลที่ผิดพลาด
  - ข้อมูลผิดรูปแบบ
  - ข้อมูลกี่หายไป
- ข้อมูล outlier
 

ที่แปลงแยกจากคนอื่น

ข้อมูลนักศึกษาชั้นปีที่ 1 ปีการศึกษา 2564

| รหัส  | เพศ  | อายุ | ความสูง | น้ำหนัก |
|-------|------|------|---------|---------|
| 64001 | ชาย  | 18   | 180     | 70      |
| 6402A | ณ    |      | 80      | 35      |
| 64123 | หญิง | 19   | 150     | 2500    |
| 65002 | ช    | 17   | 175     | 90      |



# CRISP-DM

สำหรับการบรรยายในหลักสูตร M. Sc. in Artificial Intelligence for Business Analytics, KMITL

- 3. Data Preparation**

- 3.3 แปลงรูปแบบของข้อมูล (data transformation)**

- เป็นขั้นตอนการเตรียมข้อมูลให้อยู่ในรูปแบบที่พร้อมนำไปใช้ในการวิเคราะห์ตามอัลกอริทึมของ Machine Learning ที่เลือกใช้

| ID | สินค้า | จำนวน |
|----|--------|-------|
| 1  | ปากกา  | 1     |
| 1  | ยางลบ  | 1     |
| 1  | คลิป   | 10    |
| 2  | สมุด   | 2     |
| 2  | ปากกา  | 2     |
| 3  | สมุด   | 1     |
| 3  | ปากกา  | 3     |
| 3  | ยางลบ  | 2     |



| ID | สมุด | ปากกา | ยางลบ | คลิป |
|----|------|-------|-------|------|
| 1  | -    | TRUE  | TRUE  | TRUE |
| 2  | TRUE | TRUE  | -     | -    |
| 3  | TRUE | TRUE  | TRUE  | -    |

ข้อมูลสำหรับการหากฎความสัมพันธ์ (Association Rules)

ข้อมูลใบฐานข้อมูล POS

# CRISP-DM

สำหรับการบรรยายในหลักสูตร M. Sc. in Artificial Intelligence for Business Analytics, KMITL

## • 3. Data Preparation

### • 3.3 แปลงรูปแบบของข้อมูล (data transformation)

- เป็นขั้นตอนการเตรียมข้อมูลให้อยู่ในรูปแบบที่พร้อมนำไปใช้ในการวิเคราะห์ ตามอัลกอริทึมของ Machine Learning ที่เลือกใช้

เมื่อวันที่ 4 มกราคม 2557 **เฟซบุ๊ก** เปิดตัวหน้าเพวใหม่ชื่อว่า Facebook A Look Back เมื่อผู้ใช้งานคลิกไปยังหน้านี้ก็จะแสดงคลิปวิดีโอที่บอกรเล่าเรื่องราวของผู้ใช้งานคนนั้นๆ เช่น เริ่มเล่น**เฟซบุ๊ก**ครั้งแรกปีไหน, โพสต์แรกบน**เฟซบุ๊ก**, รูปภาพที่ถูกกดไลค์มากที่สุด, รูปภาพที่ถูกแชร์มากที่สุด และ 20 อันดับเรื่องราวต่างๆ ที่เกิดขึ้นใน**เฟซบุ๊ก**ก็จะถูกแสดงและรวมไว้ในคลิปวิดีโอนี้



เอกสารทั่ว

| ID | เฟซบุ๊ก | รูปภาพ | ไลค์ | แชร์ | คลิปวิดีโอ |
|----|---------|--------|------|------|------------|
| 1  | 4       | 2      | 1    | 1    | 2          |
| 2  | ...     |        |      |      |            |

ตารางแสดงจำนวนความตื่นทองแต่ละคำ

# CRISP-DM

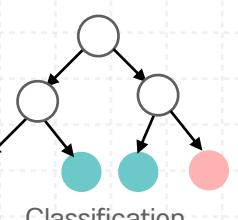
สำหรับการบรรยายในหลักสูตร M. Sc. in Artificial Intelligence for Business Analytics, KMITL

## • 4. Modeling

- เป็นขั้นตอนการวิเคราะห์ข้อมูลด้วยเทคนิคด้านตัวไม่นี่

### • Classification

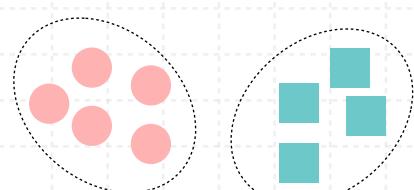
- สร้างโมเดลจากข้อมูลที่มีอยู่เพื่อกำหนดอนาคต
- เช่น กำหนดปริมาณน้ำฝนที่ตกลงวันต่อไป



Classification

### • Clustering

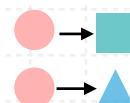
- แบ่งข้อมูลหลายกลุ่มตามความคล้ายคลึง
- เช่น แบ่งกลุ่มนักศึกษาตามคะแนนที่ได้



Clustering

### • Association Analysis

- หาความสัมพันธ์ของข้อมูลที่เกิดร่วมกัน
- เช่น ค้นหาสินค้าที่มีการซื้อร่วมกันบ่อยๆ



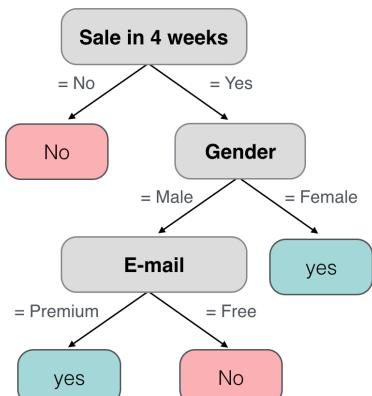
Association Analysis

# CRISP-DM

สำหรับการเรียนรู้ในหลักสูตร M. Sc. in Artificial Intelligence for Business Analytics, KMITL

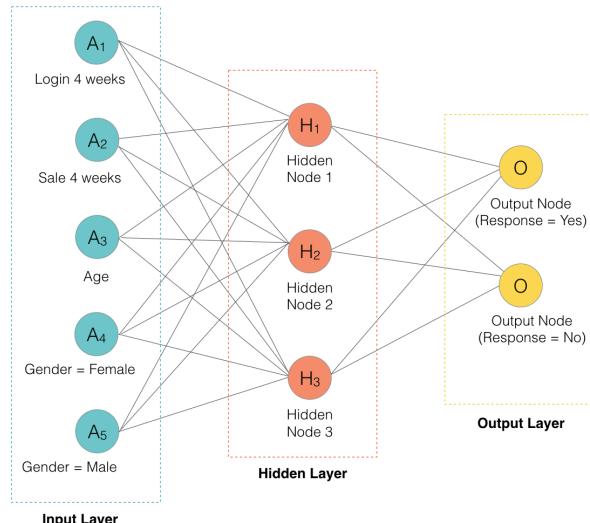
## • 5. Evaluation

- ประเมินหรือวัดประสิทธิภาพของโมเดลวิเคราะห์ข้อมูลในขั้นตอนก่อนหน้านั้น



โมเดล Decision Tree

VS



โมเดล Neural Network

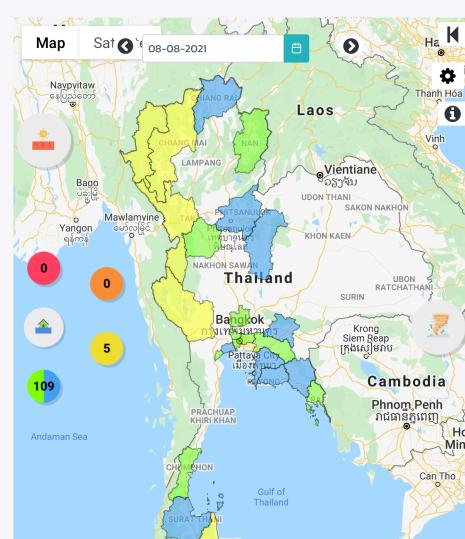
# CRISP-DM

สำหรับการเรียนรู้ในหลักสูตร M. Sc. in Artificial Intelligence for Business Analytics, KMITL

## • 6. Deployment

- นำโมเดลที่ได้ หรือ ผลการวิเคราะห์ที่ได้ไปใช้งานจริง

แผนที่คาดการณ์ภัยพิบัติ่วงหน้า 1 วัน (อุทกภัย/แมลง/ดินโคลนกลับ/วัตถุกัน)



### สรุปพื้นที่ได้รับผลกระทบ

24

จังหวัด

56

อำเภอ

119

ตำบล

35,830.97

พื้นที่ (ตร.กม.)

558,918

ครัวเรือน

1,378,409

ประชากร (คน)

### รายชื่อพื้นที่เสี่ยง

| # | จังหวัด ↑↓              | อำเภอ ↑↓    | ตำบล ↑↓ | พื้นที่ ↑↓    | การแจ้งเตือน ↑↓ | ครัวเรือน ↑↓ | ประชากร (คน) ↑↓ | ໂຄສະກ (%) ↑↓ |
|---|-------------------------|-------------|---------|---------------|-----------------|--------------|-----------------|--------------|
| 1 | กรุงเทพมหานคร เมตโรบุรี | ดาวคะนอง    | อุ久าดี  | สถานการณ์ปกติ | 11,881          | 17,392       | 59.00           |              |
| 2 | กรุงเทพมหานคร เมตโรบุรี | อุบลราชธานี | อุกุ้ม  | สถานการณ์ปกติ | 59,025          | 91,840       | 59.00           |              |
| 3 | สระบุรี ป่าตอง          | บ้านใหม่    | อุกุ้ม  | สถานการณ์ปกติ | 31,631          | 32,542       | 59.00           |              |
| 4 | พระนครศรีอยุธยาปะจิบัน  | หนองหลวง    | อุกุ้ม  | สถานการณ์ปกติ | 576             | 1,720        | 59.00           |              |
| 5 | พระนครศรีอยุธยาไกร      | โพแดง       | อุกุ้ม  | สถานการณ์ปกติ | 483             | 1,430        | 59.00           |              |
| 6 | ชลบุรี บางละมุง         | หัวหิน      | อุกุ้ม  | สถานการณ์ปกติ | 17,193          | 30,277       | 78.77           |              |
| 7 | ชลบุรี สังข์            | บางละมุง    | อุกุ้ม  | สถานการณ์ปกติ | 22,056          | 17,348       | 78.32           |              |

image source: ระบบสนับสนุนการตัดสินใจ กรมป้องกันและบรรเทาสาธารณภัย กระทรวงมหาดไทย