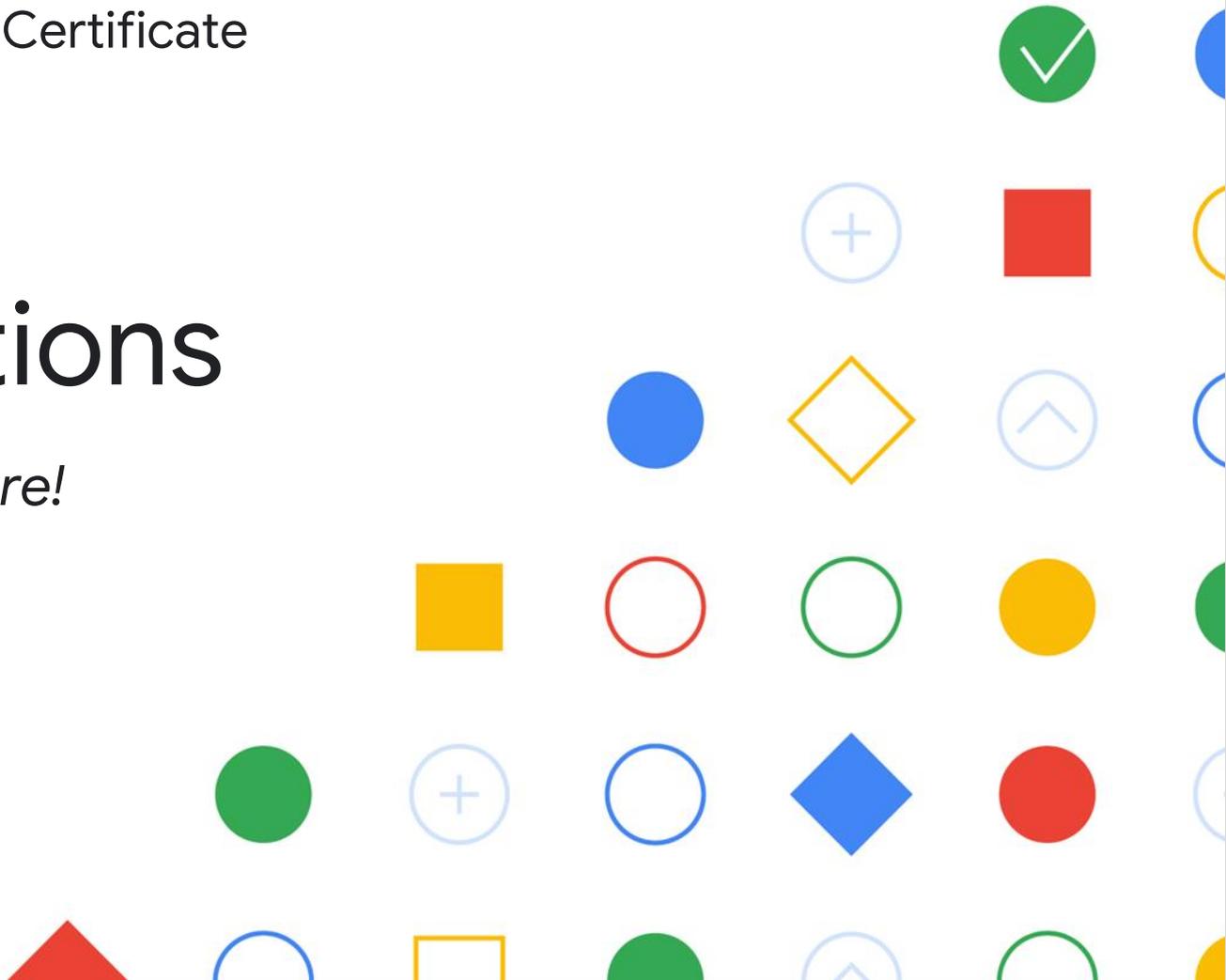


1. Foundations

Data, Data Everywhere!



Overview:

01

Introducing Data Analytics

02

All about Analytical Thinking

03

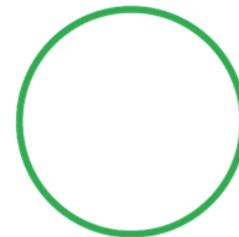
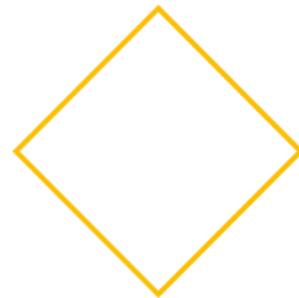
The Wonderful World of Data

04

Set Up Your Toolbox

05

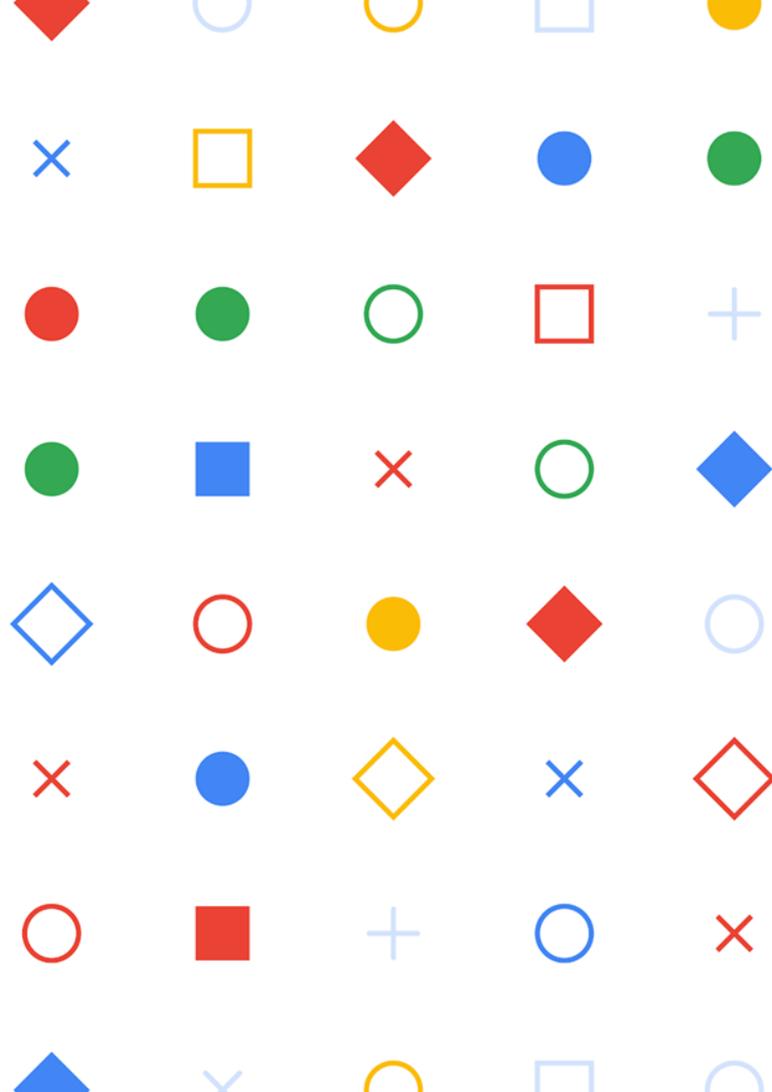
Endless Career Possibilities



Introducing Data Analytics

— *Part 1:*

1. Data Analysis Process vs. Tools
2. Technical Terms
3. Data Ecosystem
4. Data-Driven Decision Making



Data Analysis: Process vs. Tool

แยกแยะระหว่าง “กระบวนการ” กับ “เครื่องมือ”

- 1 **Ask**
ตั้งคำถาม ระบุปัญหา
- 2 **Prepare**
เตรียมการและจัดเก็บข้อมูล
- 3 **Process**
ดัดแปลงและทำความสะอาดข้อมูล
- 4 **Analyze**
ประมวลผลและวิเคราะห์ข้อมูล
- 5 **Share**
นำเสนอผลการวิเคราะห์ข้อมูล
- 6 **Act**
นำผลการวิเคราะห์ข้อมูลไปใช้จริง



Spreadsheets



SQL



Data Viz

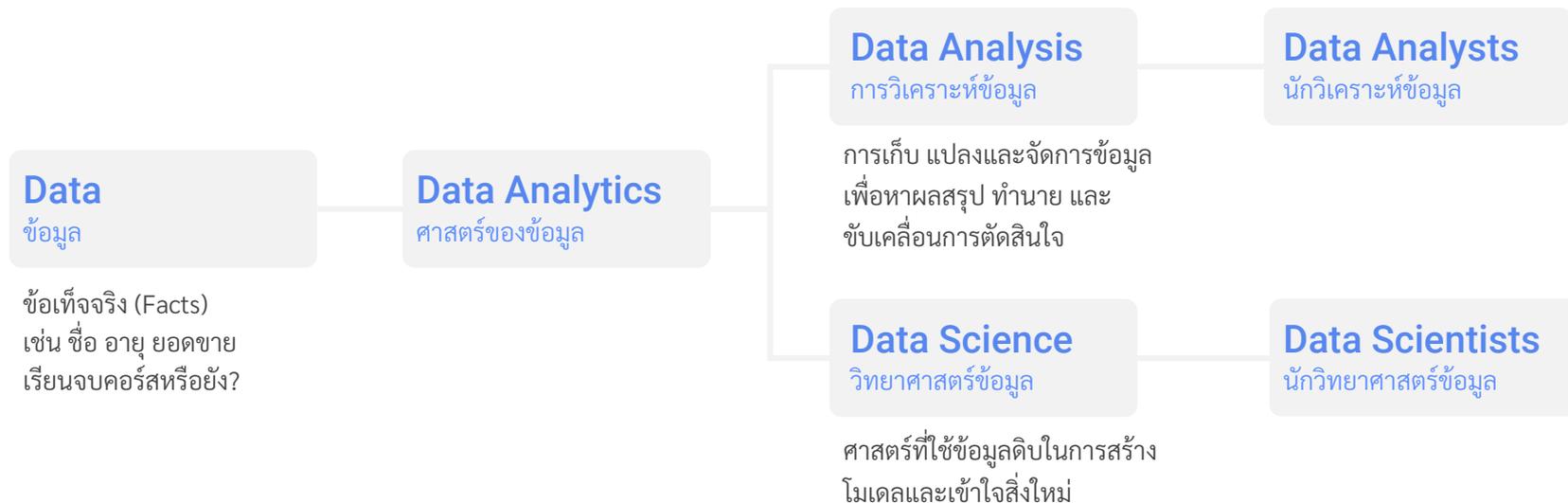


Programming



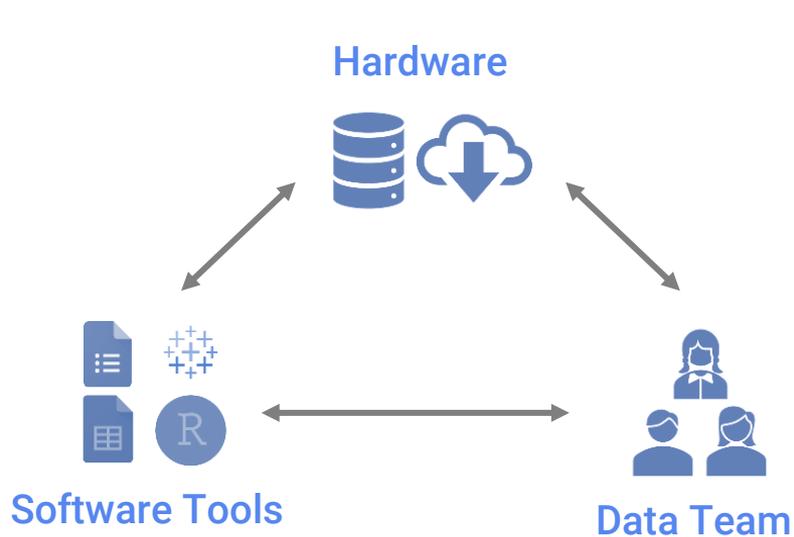
Technical Terms

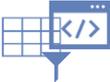
เข้าใจนิยามศัพท์ที่ใช้บ่อยเกี่ยวกับ Data



Data Ecosystem

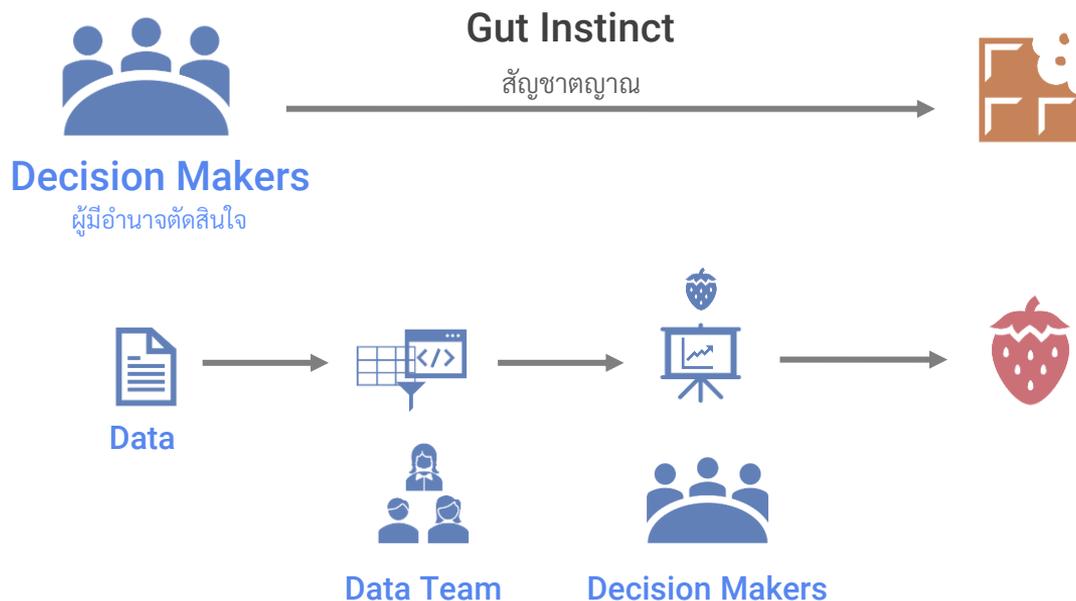
ระบบนิเวศน์ของข้อมูล



-  **Produce**
ผลิตข้อมูล
-  **Manage**
จัดการข้อมูล
-  **Store**
จัดเก็บข้อมูล
-  **Organize**
จัดระเบียบข้อมูล
-  **Analyze**
วิเคราะห์ข้อมูล
-  **Share**
นำเสนอข้อมูล

Data-Driven Decision Making

การใช้ข้อมูลในการกำหนดทิศทางกลยุทธ์ทาง “ธุรกิจ”



Data-Driven Decision Making

การใช้ข้อมูลในการกำหนดทิศทางกลยุทธ์ทาง “ธุรกิจ”

- 1 Ask**
ตั้งคำถาม ระบุปัญหา
- 2 Prepare**
เตรียมการและจัดเก็บข้อมูล
- 3 Process**
ดัดแปลงและทำความสะอาดข้อมูล
- 4 Analyze**
ประมวลผลและวิเคราะห์ข้อมูล
- 5 Share**
นำเสนอผลการวิเคราะห์ข้อมูล
- 6 Act**
นำผลการวิเคราะห์ข้อมูลไปใช้จริง



Business Needs

ความต้องการทางธุรกิจ



Collect Data

หาและเก็บข้อมูล



Trends and Patterns

แนวโน้มและรูปแบบ



Relationships

ความสัมพันธ์



Observations

ข้อสังเกต



Experience & Intuition

ประสบการณ์และสัญชาตญาณ



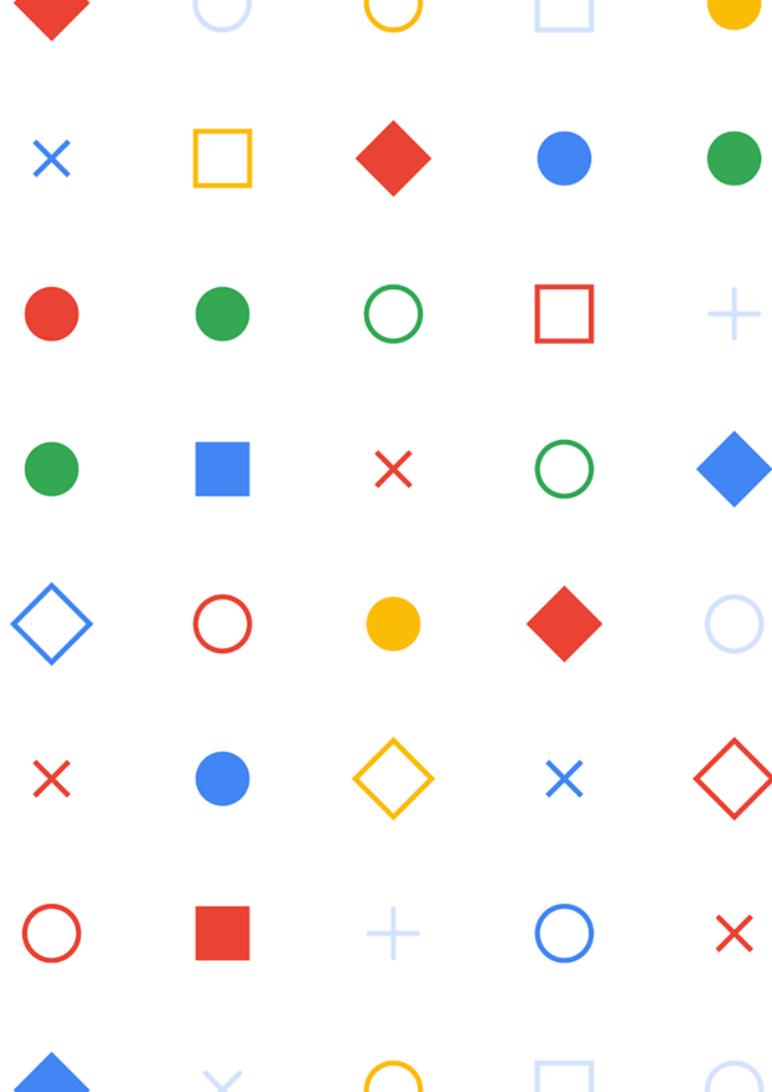
Subject Matter Experts

ผู้เชี่ยวชาญเฉพาะทาง

All about Analytical Thinking

— *Part2:*

1. Five Key Data Analytical Skills
2. Five Key Aspects of Analytical Thinking
3. Three Key Questions Data Analysts Ask



5 Key Data Analytical Skills

5 ทักษะสำคัญในการแก้ปัญหาด้วยการใช้ข้อเท็จจริง

1

Curiosity

มีความอยากรู้อยากเห็น

2

Context

เข้าใจบริบทของสิ่งที่กำลังวิเคราะห์ (Ex. motivation behind data collection)

3

Having a Technical Mindset

มีความสามารถในการแบ่งปัญหาออกเป็นขั้นตอนย่อย ๆ และจัดการอย่างเป็นลำดับขั้น (Orderly) ด้วยวิธีที่มีตรรกะ (Logical)

4

Data Design

มีวิธีจัดระเบียบข้อมูลให้ใช้งานได้ง่าย

5

Data Strategy

มีกลยุทธ์ในการบริหารผู้คน (People) กระบวนการ (Processes) และเครื่องมือ (Tools) สำหรับงานวิเคราะห์ข้อมูล

5 Key Aspects of Analytical Thinking

5 ลักษณะเด่นของการคิดเชิงวิเคราะห์

Analytical Thinking: Identifying and defining a problem, and then solving it by using data in an organized, step-by-step manner.

1

Visualization

การใช้ "ภาพ" แทนข้อมูลที่ซับซ้อน (Graphical representation of information)

2

Strategy

มีความคิดเชิงกลยุทธ์ในการจัดการกับข้อมูลเพื่อบรรลุเป้าหมาย

3

Problem-Orientation

โฟกัสที่ปัญหาที่จะแก้

4

Correlation

สามารถหา "ความสัมพันธ์" ระหว่างตัวแปรต่างๆ ของปัญหาที่กำลังแก้ได้

5

Big-Picture and Detail-Oriented Thinking

สามารถคิดได้ทั้งแบบภาพใหญ่ และเจาะรายละเอียด แล้วแต่สถานการณ์

3 Key Questions Data Analysts Ask

3 คำถามหลักที่นักวิเคราะห์ข้อมูลใช้ในการทำงาน

1

What is the root cause of the problem?

อะไรคือรากเหง้าของปัญหา?



5 Why's Technique

เมื่อเจอปัญหา ถามว่า "ทำไม" 5 ครั้ง

2

What are the gaps in our process?

กระบวนการของเราสามารถอุดช่องว่างระหว่างสภาพปัจจุบัน กับเป้าหมายที่อยากบรรลุได้แล้วหรือยัง?



Gap Analysis

วิเคราะห์หาช่องว่างระหว่างสภาพปัจจุบัน กับเป้าหมายที่อยากบรรลุ

3

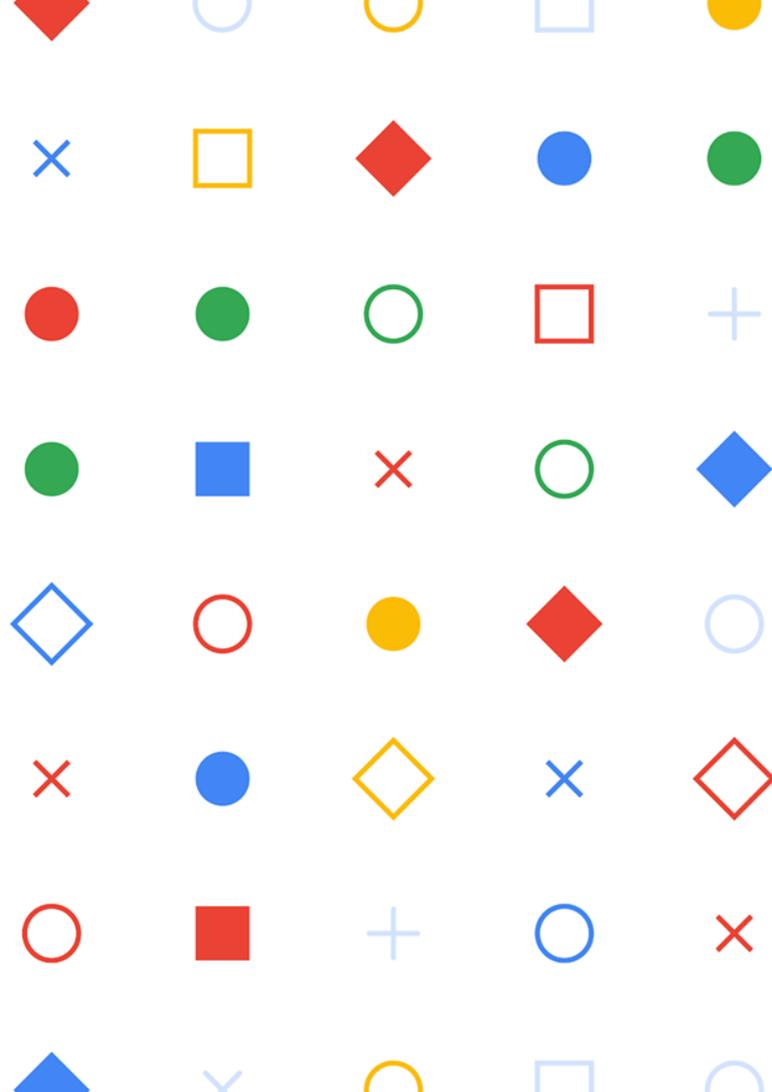
What did we not consider before?

ยังมีอะไรที่เรายังไม่ได้เอามาพิจารณาอีกบ้าง?

The Wonderful World of Data

— *Part3:*

1. Six Stages of Data Life Cycle
2. Six Phases of Data Analysis
3. Key Data Analyst Tools
4. Spreadsheets vs. Databases
5. Formula and Function



6 Stages of Data Life Cycle

6 ขั้นตอนของวงจรชีวิตของข้อมูล

1

Plan

วางแผนว่าต้องใช้ข้อมูลอะไรบ้าง จะจัดการอย่างไร และใครเป็นผู้รับผิดชอบ

2

Capture

เก็บข้อมูลจากแหล่งต่างๆ

3

Manage

ดูแลข้อมูล รวมถึงกำหนดด้วยว่าจะจัดเก็บที่ไหน อย่างไร และด้วยเครื่องมืออะไร

4

Analyze

ใช้ข้อมูลในการแก้ปัญหา ตัดสินใจ และสนับสนุนเป้าหมายทางธุรกิจ

5

Archive

เก็บข้อมูลเฉพาะส่วนที่จำเป็น (Relevant) สำหรับใช้อ้างอิงในระยะยาว (Long-term and future reference)

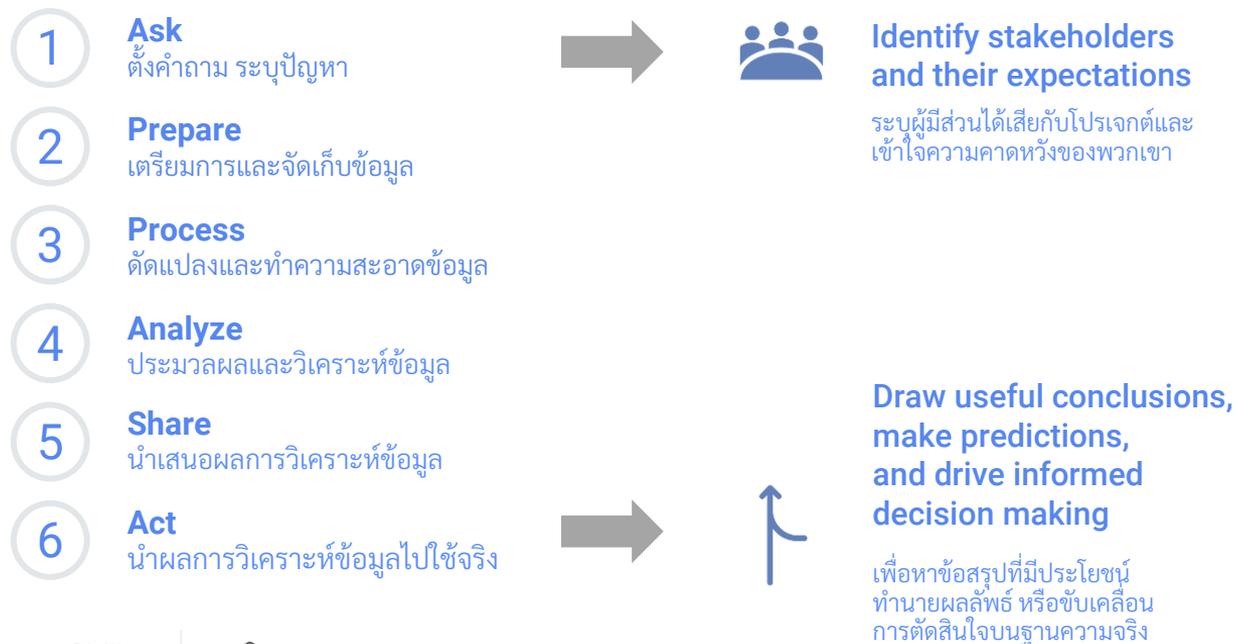
6

Destroy

ลบข้อมูลส่วนที่ไม่ใช้แล้วทิ้ง เพื่อประหยัดทรัพยากร และเพื่อความปลอดภัย

6 Phases of Data Analysis

6 ขั้นตอนของการวิเคราะห์ข้อมูล



Key Data Analyst Tools

เครื่องมือหลักของนักวิเคราะห์ข้อมูล



Spreadsheets

Collect / Store / Organize

เก็บ / บันทึก / จัดการ

Sort / Format / De-duplicate

เรียงลำดับ / เปลี่ยนประเภท / ลบแถวซ้ำ

Piece Data Together / Find Patterns / Perform Calculations

เชื่อมข้อมูล / หารูปแบบ / คำนวณ

Create Graphs and Charts

สร้างกราฟและแผนภาพ



SQL

Request and Retrieve Data from Databases

ขอและดึงข้อมูลจากฐานข้อมูล

Create, Update, Delete or Download Data

สร้าง อัปเดต ลบ หรือดาวน์โหลดข้อมูล



Data Viz

Create Graphs Maps and Charts

สร้างกราฟ แผนที่และแผนภาพ

Create Dashboards

สร้างแดชบอร์ด

Data Storytelling

เปลี่ยนข้อมูลเป็นเรื่องราว



Programming

Customized Analysis

การวิเคราะห์ที่เฉพาะเจาะจงมากๆ

Process Automation

ทำให้ขั้นตอนต่างๆ เป็นอัตโนมัติได้

Spreadsheets vs. Databases

โปรแกรมตารางคำนวณ vs. ฐานข้อมูล



Spreadsheets

เป็นซอฟต์แวร์แอปพลิเคชัน

เป็นแถว (Row) และคอลัมน์ (Column)

เก็บเป็นช่องๆ (Cells)

จำกัดและขนาดไม่ใหญ่มาก

แบบ manual

โดยทั่วไปคือทีละ 1 คน

คนใช้โปรแกรม (User)



Databases

เป็นคลังข้อมูล เข้าถึงได้ด้วยภาษา query เช่น SQL

กฎ (Rules) และความสัมพันธ์ (Relationships)

เก็บเป็นกลุ่มก้อนที่ซับซ้อน (Complex collections)

มหาศาล

แบบเข้มงวด (Strict) และคงเส้นคงวา (Consistent)

ได้ที่ละหลายๆ คน

ระบบดูแลฐานข้อมูล (Database management system)

ลักษณะ

โครงสร้างข้อมูล

การจัดเก็บข้อมูล

ปริมาณข้อมูลที่รองรับ

การนำเข้าข้อมูล

จำนวนผู้ใช้

ผู้ดูแล

Formula and Function

สูตรคำนวณและฟังก์ชัน

Formula

สูตรคำนวณ

ชุดคำสั่ง (A set of instructions) ที่นำข้อมูลในตารางมาคำนวณ เช่น + - * /

Function

ฟังก์ชัน

คำสั่งที่มีอยู่แล้วในโปรแกรมเช่น Google Sheets ที่นำเอาข้อมูลในตารางมาประมวลผลให้อัตโนมัติ เช่น หาผลรวม หาค่าเฉลี่ย

Ex. หาคะแนนรวม

C4	\downarrow	<i>fx</i>	=A2+2*B2+C2	
		A	B	C
1		Phys (100)	Chem (50)	Bio (100)
2		99	40	50
3				
4			Total	229

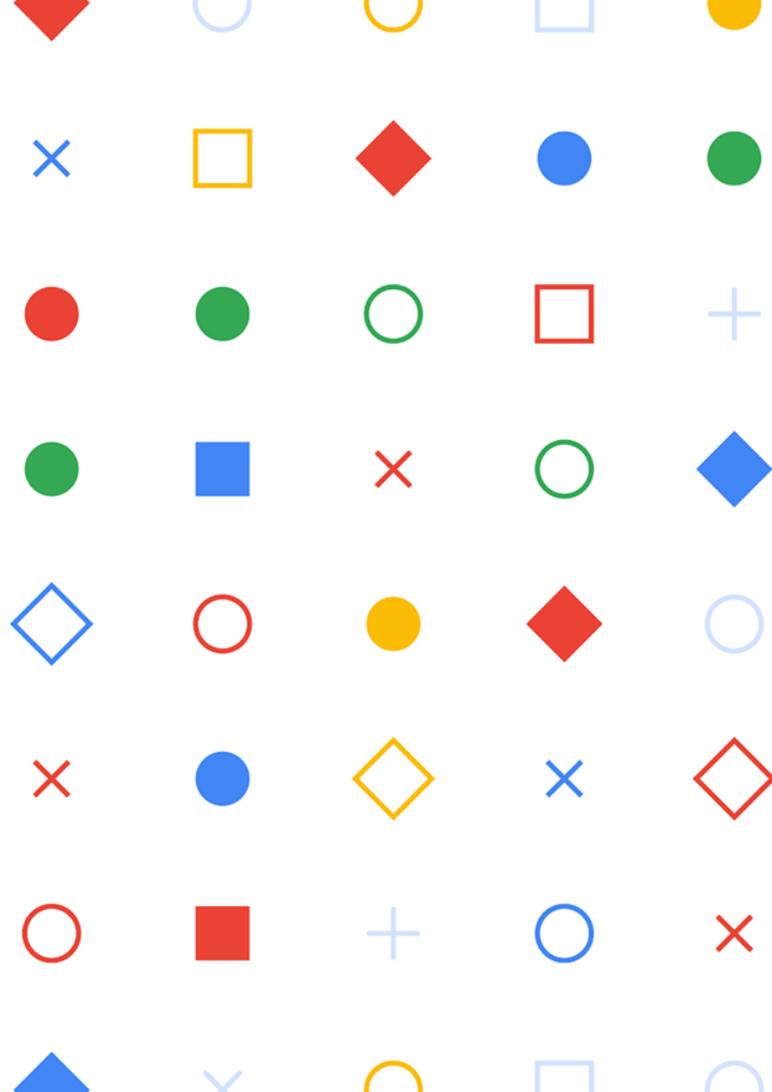
Ex. หาค่าเฉลี่ยส่วนสูง

B8	\downarrow	<i>fx</i>	=AVERAGE(B2:B6)	
		A	B	
1			Heights	
2		Dad	160	
3		Mom	148	
4		Great	172	
5		Best	155	
6		Fresh	154	
7				
8		Average	157.8	

Set Up Your Toolbox

— *Part4:*

1. Intro to Spreadsheets
2. Intro to Structured Query Language (SQL)
3. Choosing the Right Data Visualization

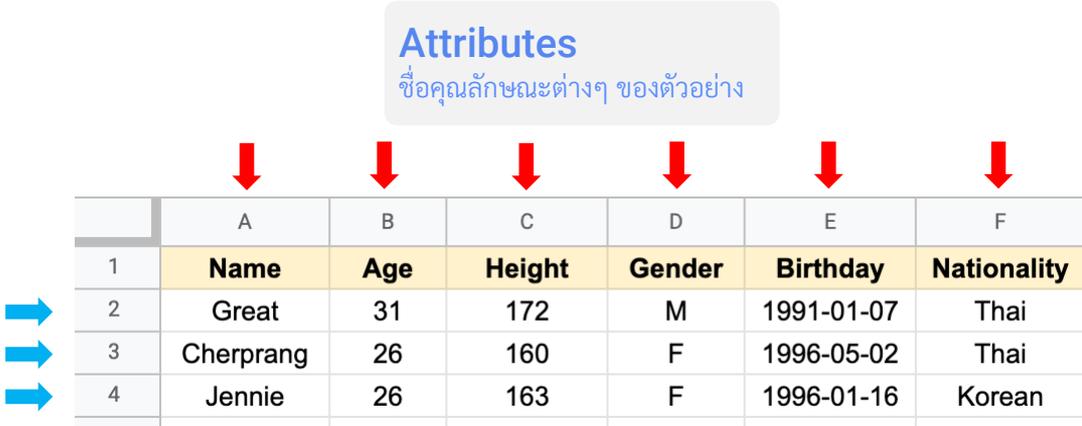


Intro to Spreadsheet

โปรแกรม Spreadsheet เช่น Google Sheets จะเก็บข้อมูลเป็นตารางและโดยทั่วไปจะมีลักษณะดังแสดง

Attributes

ชื่อคุณลักษณะต่างๆ ของตัวอย่าง



	A	B	C	D	E	F
1	Name	Age	Height	Gender	Birthday	Nationality
2	Great	31	172	M	1991-01-07	Thai
3	Cherprang	26	160	F	1996-05-02	Thai
4	Jennie	26	163	F	1996-01-16	Korean

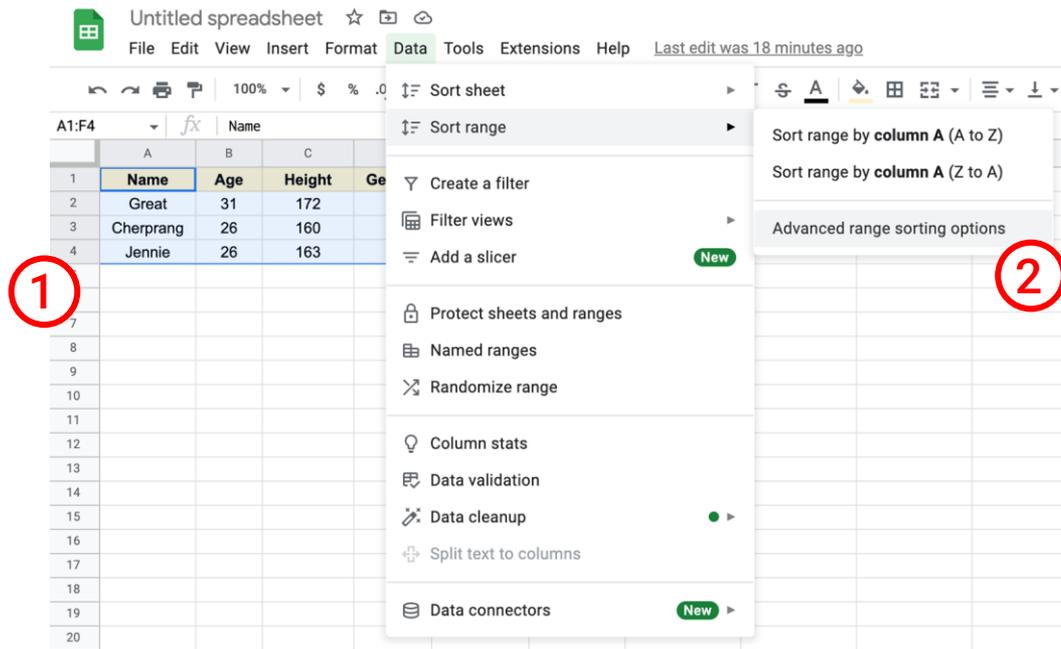
Observations

ค่าสังเกต

ก็คือค่าข้อมูลทั้งหมดของแต่ละตัวอย่าง (Instance)

Intro to Spreadsheet

เราใช้ประโยชน์จากเครื่องมือต่างๆ ในโปรแกรม Spreadsheet ได้ เช่น เรียงลำดับข้อมูล (Sort)



1

2

Sort sheet

Sort range

Sort range by column A (A to Z)

Sort range by column A (Z to A)

Create a filter

Filter views

Add a slicer

Protect sheets and ranges

Named ranges

Randomize range

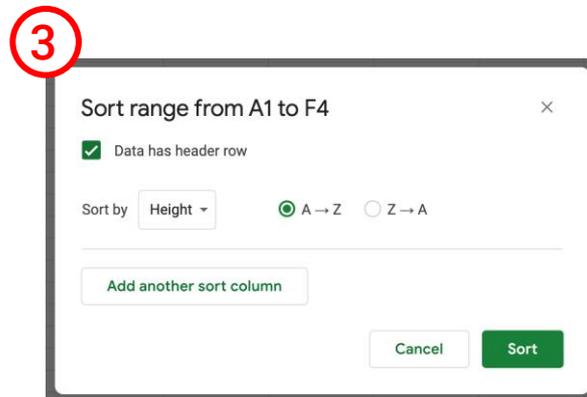
Column stats

Data validation

Data cleanup

Split text to columns

Data connectors



3

Sort range from A1 to F4

Data has header row

Sort by: Height

A → Z Z → A

Add another sort column

Cancel Sort

Result

	A	B	C	D	E	F
1	Name	Age	Height	Gender	Birthday	Nationality
2	Cherprang	26	160	F	1996-05-02	Thai
3	Jennie	26	163	F	1996-01-16	Korean
4	Great	31	172	M	1991-01-07	Thai

Intro to SQL

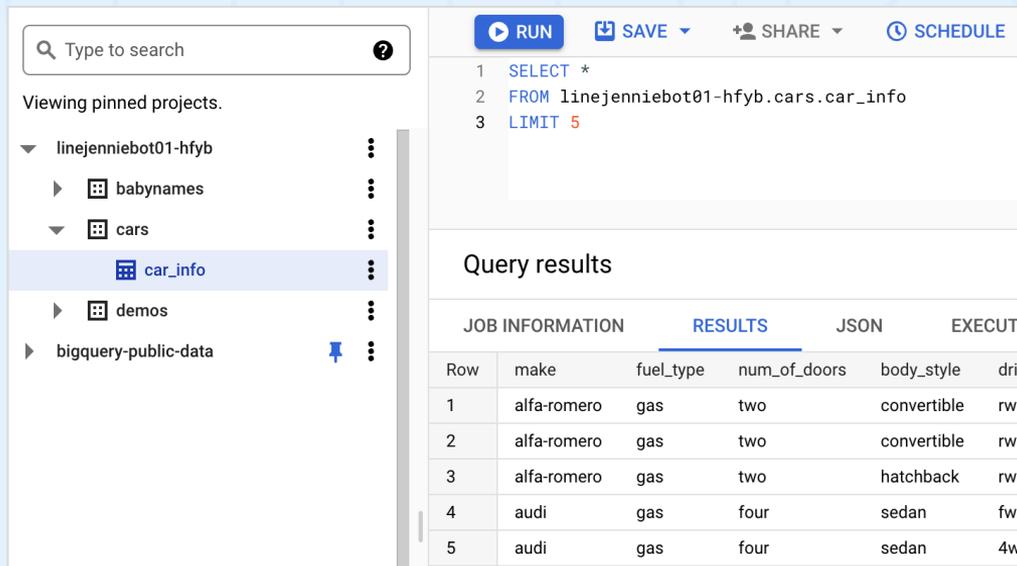
Query

การขอข้อมูล (request for data)
จากฐานข้อมูล (database)

Syntax:

SELECT	รายชื่อคอลัมน์ที่ต้องการ
FROM	ชื่อตาราง
WHERE	เงื่อนไข (Optional)
LIMIT	จำนวนแถวที่ต้องการ (Optional)

Ex. ดึงคอลัมน์ทั้งหมดของตาราง cars_info โดยเอามาแค่ 5 แถวแรกพอ



The screenshot shows a BigQuery interface. On the left, a project tree is visible with 'cars' selected and 'car_info' highlighted. The main area contains a SQL query editor with the following code:

```
1 SELECT *
2 FROM linejenniebot01-hfyb.cars.car_info
3 LIMIT 5
```

Below the query editor, the 'Query results' section is active, displaying a table with 5 rows of data:

Row	make	fuel_type	num_of_doors	body_style	drive
1	alfa-romero	gas	two	convertible	rwd
2	alfa-romero	gas	two	convertible	rwd
3	alfa-romero	gas	two	hatchback	rwd
4	audi	gas	four	sedan	fwd
5	audi	gas	four	sedan	4wd

Intro to SQL

Query

การขอข้อมูล (request for data)
จากฐานข้อมูล (database)

Syntax:

SELECT	รายชื่อคอลัมน์ที่ต้องการ
FROM	ชื่อตาราง
WHERE	เงื่อนไข (Optional)
LIMIT	จำนวนแถวที่ต้องการ (Optional)

Ex. โปรแกรมจะข้ามส่วนที่เรา "Comment"

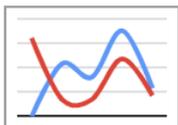
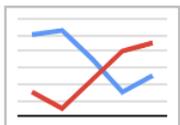
The screenshot shows a SQL query editor interface. On the left, there is a sidebar with a search bar and a list of pinned projects: linejennibot01-hfyb (containing babynames, cars, and car_info), demos, and bigquery-public-data. The 'car_info' table is selected. The main editor area contains a SQL query with comments: `1 SELECT make, engine_type, price`, `2 FROM linejennibot01-hfyb.cars.car_info`, `3 --WHERE make = 'audi' AND price > 17000`, `4 --LIMIT 5`, and `5 --Cherprang is very cute=`. The query is executed, and the results are displayed in a table. The table has columns 'Row', 'make', 'engine_type', and 'price'. The results are as follows:

Row	make	engine_type	price
1	alfa-romero	dohc	13495
2	alfa-romero	dohc	16500
3	alfa-romero	ohcv	16500
4	audi	ohc	13950
5	audi	ohc	17450
6	audi	ohc	15250
7	audi	ohc	17710
8	audi	ohc	18920
9	audi	ohc	23875

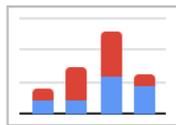
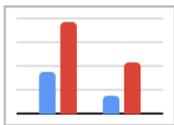
Choosing the Right Data Visualization

ใช้รูปภาพนำเสนอข้อมูลให้ตรงตามวัตถุประสงค์

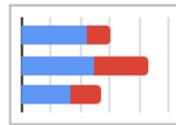
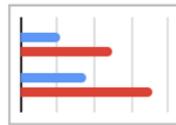
Line



Column



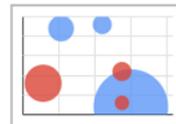
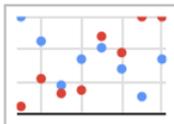
Bar



Pie



Scatter



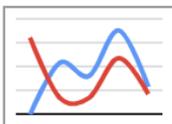
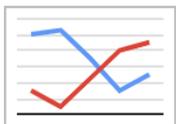
Map



Choosing the Right Data Visualization

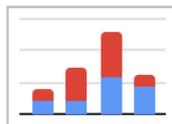
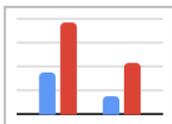
ใช้รูปภาพนำเสนอข้อมูลให้ตรงตามวัตถุประสงค์

Line



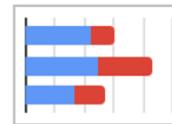
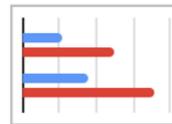
- ดูแนวโน้ม (Trend) รูปแบบ (Pattern) หรือ การเปลี่ยนแปลง (Change) ตามเวลา

Column



- เปรียบเทียบขนาดของ ปริมาณที่สนใจ (Compare and contrast using size)

Bar



Choosing the Right Data Visualization

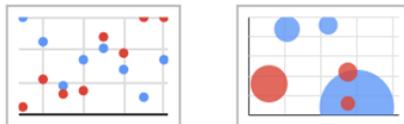
ใช้รูปภาพนำเสนอข้อมูลให้ตรงตามวัตถุประสงค์

Pie



- ดูสัดส่วนแบ่งตามกลุ่ม (Proportions)

Scatter



- ดูความสัมพันธ์ระหว่าง 2 ปริมาณขึ้นไป (Relationships between variables)

Map

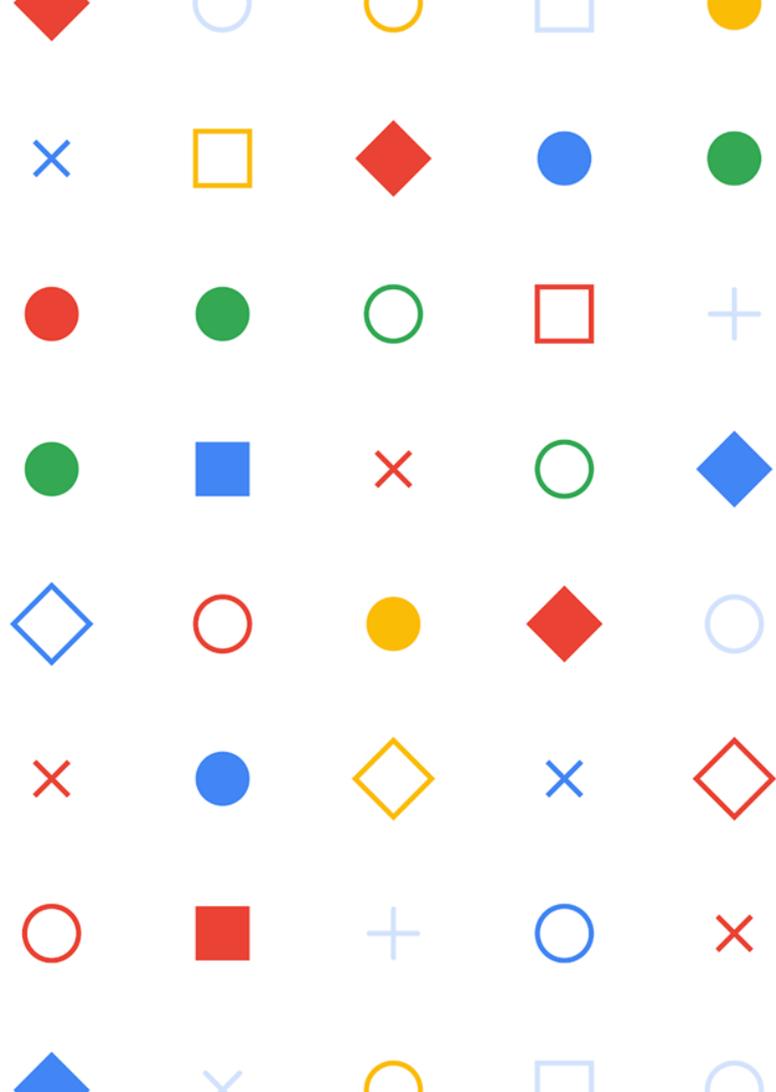


- เปรียบเทียบข้อมูลแบ่งตามตำแหน่งที่อยู่ (Locations)

Endless Career Possibilities

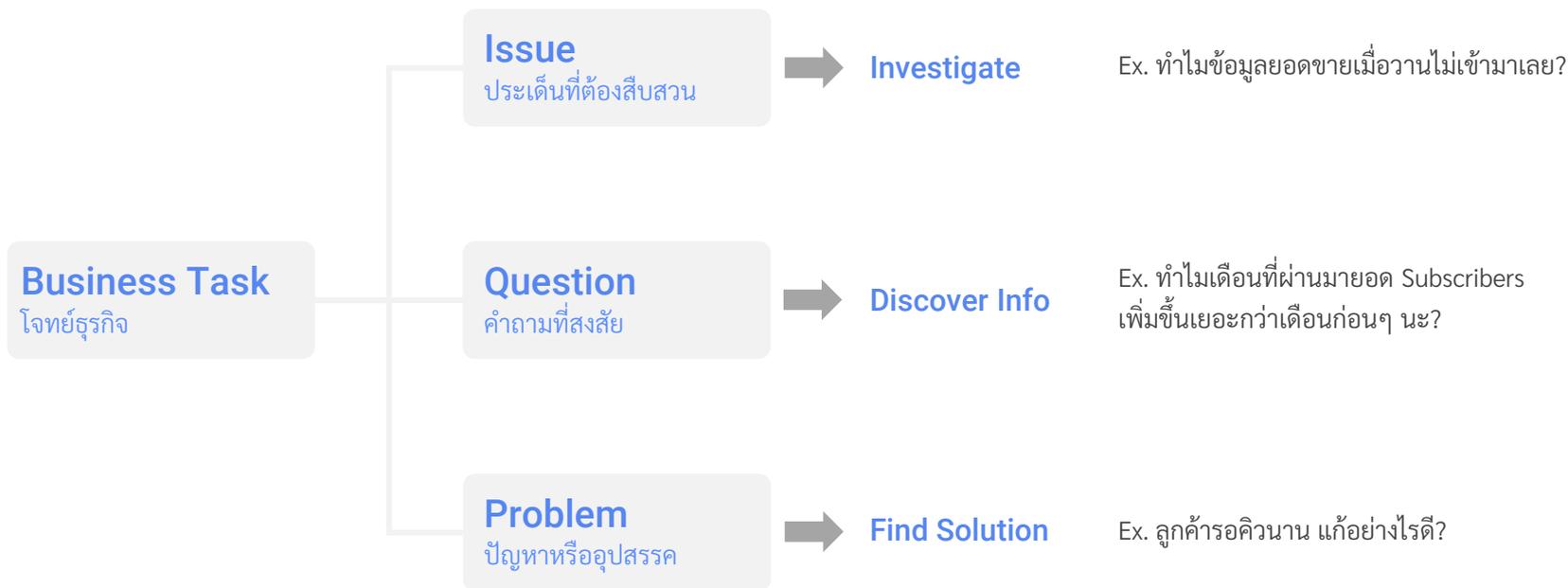
— *Part5:*

1. Data and Business Task
2. Data and Fairness



Data and Business Task

ใช้ข้อมูลตอบโจทย์ธุรกิจ



Data and Fairness

เก็บและใช้ข้อมูลอย่างเป็นธรรม ไม่มีความลำเอียง (Bias)

Fairness
ความเป็นธรรม

Data Collection

การเก็บข้อมูล

- กลุ่มตัวอย่างครอบคลุมคนทุกกลุ่ม (Inclusive samples)
- บริบททางสังคม (Social context)
- ใช้ข้อมูลรายงานด้วยตนเองของคนในกลุ่มตัวอย่างด้วย (Self-reported data)

Data Analysis

การวิเคราะห์ข้อมูล

- สรุปผลถูกต้อง (Correct) แต่ระวังว่าอาจจะไม่ยุติธรรม (Unfair)
- อาจขอความเห็นจากนักสังคมวิทยา (Social scientists) ในการป้องกันความลำเอียง (Bias)

