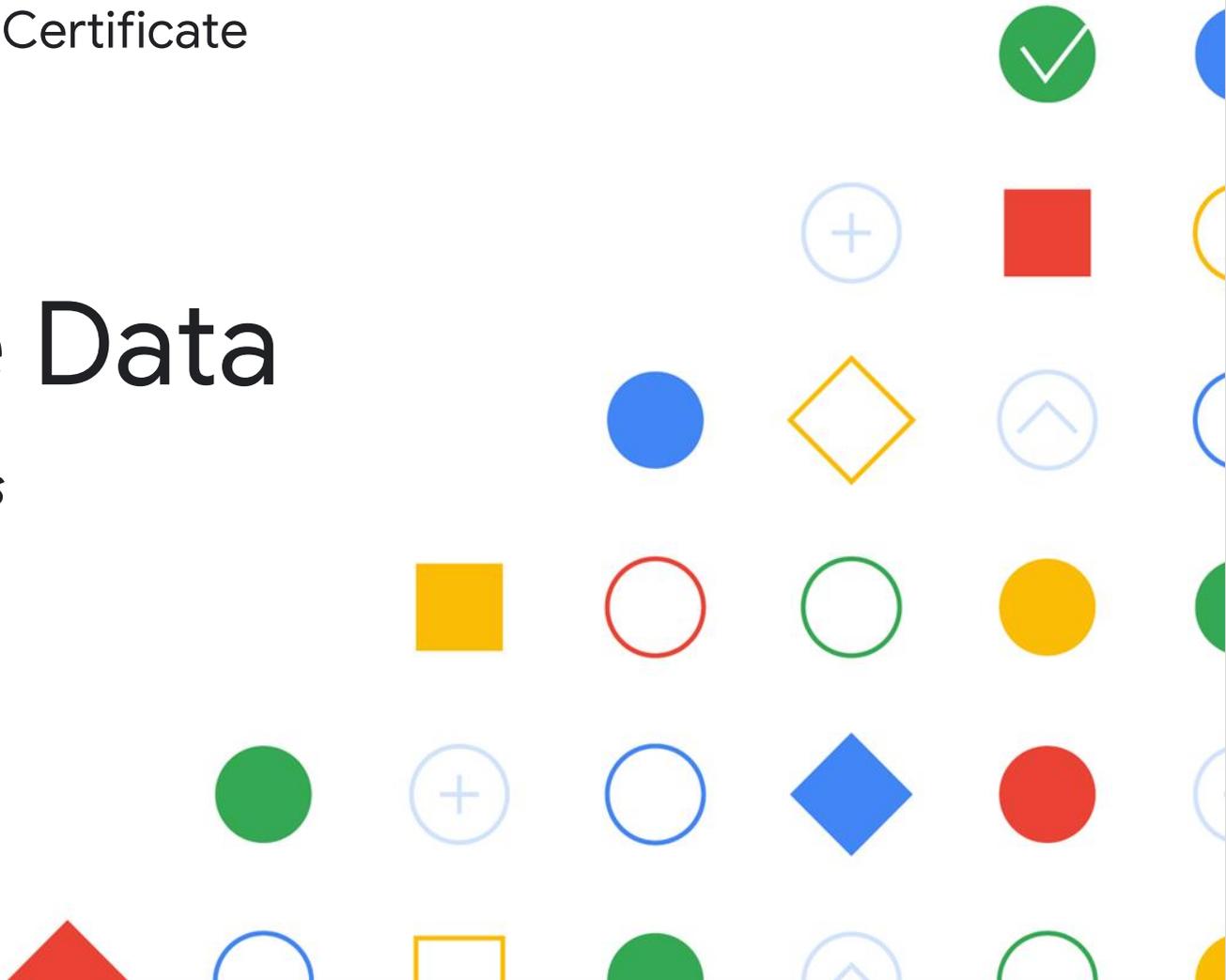


5. Analyze Data

to Answer Questions



Overview:

01

Organizing Data to Begin Analysis

02

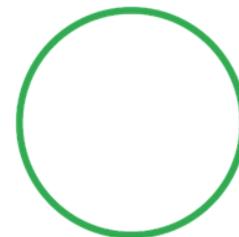
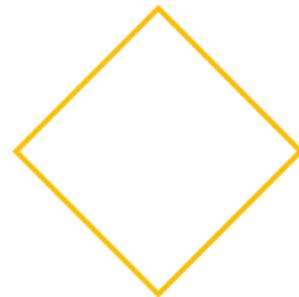
Formatting and Adjusting Data

03

Aggregating Data for Analysis

04

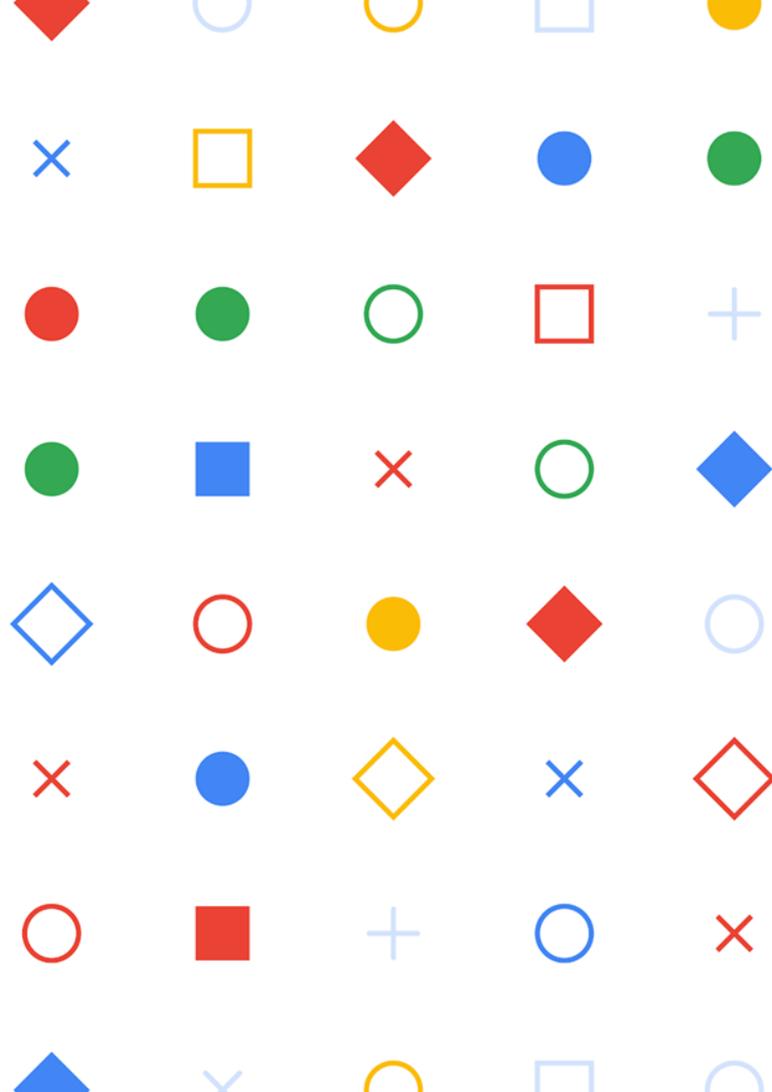
Performing Data Calculations



Organizing Data to Begin Analysis

— *Part 1:*

1. Four Phases of Data Analysis
2. Sorting and Filtering: Definitions
3. Sorting and Filtering: Spreadsheets
4. Sorting and Filtering: SQL



4 Phases of Data Analysis

4 ระยะของการวิเคราะห์ข้อมูลหาแนวโน้ม (Trends) และความสัมพันธ์ (Relationships) เพื่อตอบโจทย์ที่ต้องการ

1

Organize data

จัดระเบียบและรวบรวมข้อมูลที่จะใช้วิเคราะห์ให้เรียบบร้อยว่าตัวแปร (Variables) และประเภทข้อมูล (Data types) เหมาะสม

2

Format and adjust data

ปรับประเภทหรือค่าของข้อมูลให้เหมาะสม รวมถึงเรียงลำดับข้อมูล (Sorting) และกรองข้อมูลที่ไม่จำเป็นออก (Filtering)

3

Get input from others

รับความคิดเห็นจากคนอื่น เพื่อช่วยเปิดมุมมองหรือป้องกันปัญหาที่อาจเกิดขึ้นในการวิเคราะห์ข้อมูล

4

Transform data

หาความสัมพันธ์ระหว่างจุดข้อมูล (Relationships or correlations between data points) และทำการคำนวณ (Calculations)

Sorting and Filtering: Definitions

การเรียงลำดับ (Sorting) มี 2 แบบ: จากน้อยไปมาก (Ascending: A-Z) และจากมากไปน้อย (Descending: Z-A)

การกรองค่า (Filtering): กรองเอาแถวที่มีค่าที่สอดคล้องเงื่อนไข (Criteria) คงไว้ในตาราง

	A	B	C	D	E
1	Name	Age	Height	Gender	Birthday
2	Great	31	178	M	1997-05-02
3	Cherprang	26	160	F	1996-05-02
4	Jennie	26	163	F	1996-01-16
5	Bambam	25	172	M	1991-01-07
6	Sergey	48	173	M	1973-06-21
7	Larry	49	181	M	1973-03-26
8	Elon	50	187	M	1971-06-28
9	Jisoo	50	162	F	1995-01-03

Sort by Age in ascending order

	A	B	C	D	E
1	Name	Age	Height	Gender	Birthday
2	Bambam	25	178	M	1997-05-02
3	Cherprang	26	160	F	1996-05-02
4	Jennie	26	163	F	1996-01-16
5	Great	31	172	M	1991-01-07
6	Sergey	48	173	M	1973-06-21
7	Larry	49	181	M	1973-03-26
8	Elon	50	187	M	1971-06-28
9	Jisoo	50	162	F	1995-01-03

Filter *by* Age > 30
(Filter *out* rows with Age <= 30)

	A	B	C	D	E
1	Name	Age	Height	Gender	Birthday
5	Great	31	172	M	1991-01-07
6	Sergey	48	173	M	1973-06-21
7	Larry	49	181	M	1973-03-26
8	Elon	50	187	M	1971-06-28

```
RUN SAVE SHARE SCHEDULE
1 SELECT make, engine_type, price
2 FROM linejenniebot01-hfyb.cars.car_info
3 WHERE make = 'audi'
4 ORDER BY price DESC
```

Query results

	JOB INFORMATION	RESULTS	JSON	EXECUTION
Row	make	engine_type	price	
1	audi	ohc	23875	
2	audi	ohc	18920	
3	audi	ohc	17710	
4	audi	ohc	17450	
5	audi	ohc	15250	
6	audi	ohc	13950	
7	audi	ohc	0	

Sorting and Filtering: Spreadsheets

การเรียงลำดับ (Sorting) ใน Google Sheets มี 2 แบบ:

1. Sort sheet: ข้อมูลในตารางทั้งหน้า (Sheet) ถูกเรียงแถวตามลำดับค่าในคอลัมน์ที่ต้องการ
2. Sort range: ข้อมูลถูกเรียงลำดับเฉพาะในช่วงข้อมูล (Range) ที่เลือก ตามลำดับค่าในคอลัมน์ที่ต้องการ

Ex. เรียงลำดับสมาชิกจากอายุน้อยไปมาก

New Members of Our Company (2022) ☆ 🗖 🌐
File Edit View Insert Format Data Tools Extensions Help Last edit was seconds ago

Sort sheet
Sort range

Sort sheet by column B (A to Z)
Sort sheet by column B (Z to A)

1	Name	Age	He
2	Bambam	25	1
3	Cherprang	26	1
4	Elon	50	1
5	Great	31	1
6	Jennie	26	1
7	Jisoo	26	1
8	Larry	49	1
9	Sergey	48	1
10			

Result

1	Name	Age	Height	Gender	Birthday	Nationality
2	Bambam	25	178	M	1997-05-02	Thai
3	Cherprang	26	160	F	1996-05-02	Thai
4	Jennie	26	163	F	1996-01-16	Korean
5	Great	31	172	M	1991-01-07	Thai
6	Sergey	48	173	M	1973-06-21	American
7	Larry	49	181	M	1973-03-26	American
8	Elon	50	187	M	1971-06-28	
9	Jisoo	26	162	F	1995-01-03	

Ex. เรียงเฉพาะคอลัมน์อายุจากน้อยไปมากโดยเรียงคอลัมน์อื่นไว้เหมือนเดิม

New Members of Our Company (2022) ☆ 🗖 🌐
File Edit View Insert Format Data Tools Extensions Help Last edit was seconds ago

Sort sheet
Sort range

Sort range by column B (A to Z)
Sort range by column B (Z to A)

Advanced range sorting options

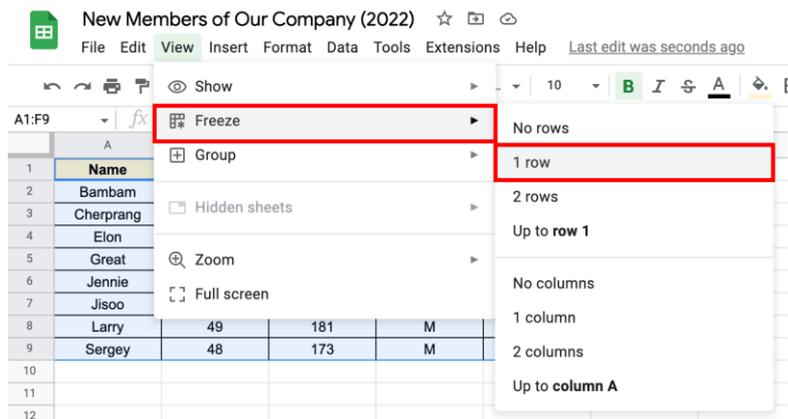
1	Name	Age	He
2	Bambam	25	1
3	Cherprang	26	1
4	Elon	50	1
5	Great	31	1
6	Jennie	26	1
7	Jisoo	26	1
8	Larry	49	1
9	Sergey	48	1
10			

Result

1	Name	Age	Height	Gender	Birthday	Nationality
2	Bambam	50	178	M	1997-05-02	Thai
3	Cherprang	49	160	F	1996-05-02	Thai
4	Elon	48	187	M	1971-06-28	
5	Great	31	172	M	1991-01-07	Thai
6	Jennie	26	163	F	1996-01-16	Korean
7	Jisoo	26	162	F	1995-01-03	
8	Larry	25	181	M	1973-03-26	American
9	Sergey		173	M	1973-06-21	American

Sorting and Filtering: Spreadsheets

Note: ก่อนจะ Sort sheet หรือ Sort range เราต้องตั้งค่าให้โปรแกรมตรึงแถวบนสุดซึ่งเป็นหัวตาราง (Freeze header row) ก่อนที่จะทำการ Sort ไม่เช่นนั้น ชื่อคอลัมน์จะถูกนำมาเรียงลำดับด้วย!



New Members of Our Company (2022) ☆ 📁 ☁
File Edit View Insert Format Data Tools Extensions Help Last edit was seconds ago

Freeze

- No rows
- 1 row**
- 2 rows
- Up to row 1
- No columns
- 1 column
- 2 columns
- Up to column A

	A	B	C	D	E	F
1	Name					
2	Bambam					
3	Cherprang					
4	Elon					
5	Great					
6	Jennie					
7	Jisoo					
8	Larry	49	181	M		
9	Sergey	48	173	M		
10						
11						
12						



	A	B	C	D	E	F
1	Name	Age	Height	Gender	Birthday	Nationality
2	Bambam	25	178	M	1997-05-02	Thai
3	Cherprang	26	160	F	1996-05-02	Thai
4	Elon	50	187	M	1971-06-28	
5	Great	31	172	M	1991-01-07	Thai
6	Jennie	26	163	F	1996-01-16	Korean
7	Jisoo		162	F	1995-01-03	
8	Larry	49	181	M	1973-03-26	American
9	Sergey	48	173	M	1973-06-21	American

Sorting and Filtering: Spreadsheets

อีก 1 วิธีในการ Sort คือการใช้ฟังก์ชัน SORT

Syntax:

=SORT (ขอบเขตข้อมูลตารางที่จะเอา, หมายเลขคอลัมน์ที่ใช้เรียงลำดับ, เรียงจากน้อยไปมากใช้หรือไม่?)

TRUE = Ascending
FALSE = Descending

Ex. เรียงลำดับสมาชิกจากอายุมากไปน้อย เอาเฉพาะชื่อกับอายุ

	A	B	C	D	E	F
1	Name	Age	Height	Gender	Birthday	Nationality
2	Bambam	25	178	M	1997-05-02	Thai
3	Cherprang	26	160	F	1996-05-02	Thai
4	Elon	50	187	M	1971-06-28	
5	Great	31	172	M	1991-01-07	Thai
6	Jennie	26	163	F	1996-01-16	Korean
7	Jisoo		162	F	1995-01-03	
8	Larry	49	181	M	1973-03-26	American
9	Sergey	48	173	M	1973-06-21	American
10						
11	=SORT(A2:B9, 2, FALSE)					
12						
13	SORT(range, sort_column, is_ascending,					
14	[sort_column2, ...], [is_ascending2, ...])					
15						
16						



	A	B	C
1	Name	Age	Height
2	Bambam	25	178
3	Cherprang	26	160
4	Elon	50	187
5	Great	31	172
6	Jennie	26	163
7	Jisoo		162
8	Larry	49	181
9	Sergey	48	173
10			
11	Elon	50	
12	Larry	49	
13	Sergey	48	
14	Great	31	
15	Cherprang	26	
16	Jennie	26	
17	Bambam	25	
18	Jisoo		

Sorting and Filtering: Spreadsheets

เราสามารถเรียงลำดับข้อมูลโดยการพิจารณามากกว่า 1 คอลัมน์ได้

Ex. เรียงลำดับสมาชิกตามลำดับประเทศตามพจนานุกรม และเรียงลำดับส่วนสูงจากมากไปน้อยของคนในแต่ละประเทศ

Sort sheet

Sort range

Sort range by column F (A to Z)

Sort range by column F (Z to A)

Advanced range sorting options

Create a filter

Filter views

Add a slicer

Protect sheets and ranges

Named ranges

Randomize range

Sort range from A1 to F9

Data has header row

Sort by: Nationality A → Z Z → A

then 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	Elon	50	187	M	1971-06-28	American
3	Larry	49	181	M	1973-03-26	American
4	Sergey	48	173	M	1973-06-21	American
5	Jennie	26	163	F	1996-01-16	Korean
6	Jisoo	27	162	F	1995-01-03	Korean
7	Bambam	25	178	M	1997-05-02	Thai
8	Great	31	172	M	1991-01-07	Thai
9	Cherprang	26	160	F	1996-05-02	Thai

หมายเหตุ: สังเกตว่าเราใช้ไลต์ตารางทั้งหมดก่อนที่จะใช้ Sort range แปลว่า Range ในที่นี้ของเราคือตารางทั้งหมด และข้อมูลแต่ละแถวจะไม่เกิดการสลับกัน

Sorting and Filtering: SQL

กรองและเรียงลำดับข้อมูล

Syntax: กรองเรียงจากน้อยไปมาก (Ascending)

SELECT	รายชื่อคอลัมน์ที่ต้องการ
FROM	ชื่อตาราง
WHERE	เงื่อนไขกรองข้อมูล
ORDER BY	ชื่อคอลัมน์

Syntax: เรียงจากมากไปน้อย (Descending)

SELECT	รายชื่อคอลัมน์ที่ต้องการ
FROM	ชื่อตาราง
WHERE	เงื่อนไขกรองข้อมูล
ORDER BY	ชื่อคอลัมน์ DESC

Ex. เรียงลำดับราคารถยนต์ยี่ห้อ audi จากมากไปน้อย (Descending)

[▶ RUN](#) [📄 SAVE](#) [+ 👤 SHARE](#) [🕒 SCHEDULE](#)

```
1 SELECT make, engine_type, price
2 FROM linejenniebot01-hfyb.cars.car_info
3 WHERE make = 'audi'
4 ORDER BY price DESC
```

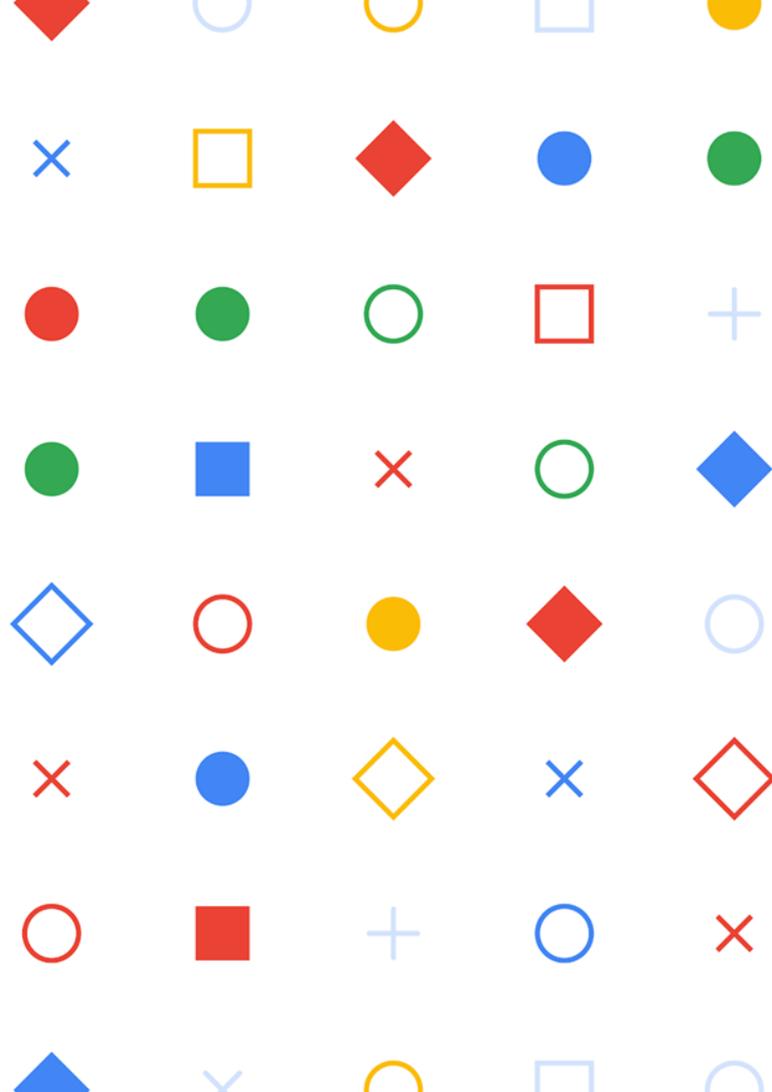
Query results

JOB INFORMATION		RESULTS	JSON	EXECUTION
Row	make	engine_type	price	
1	audi	ohc	23875	
2	audi	ohc	18920	
3	audi	ohc	17710	
4	audi	ohc	17450	
5	audi	ohc	15250	
6	audi	ohc	13950	
7	audi	ohc	0	

Formatting and Adjusting Data

— *Part2:*

1. Converting Units: Spreadsheets
2. Formatting Data: Spreadsheets
3. Formatting Data: SQL
4. Data Validation: Spreadsheets
5. Conditional Formatting, Concatenating Data and Text Functions (Already covered in Course 4)



Converting Units: Spreadsheets

CONVERT

เปลี่ยนหน่วยข้อมูล

Syntax:

=CONVERT (ค่าตั้งต้น, หน่วยตั้งต้น, หน่วยใหม่)

ดูหน่วยที่แปลงได้ทั้งหมดได้ที่:

<https://support.google.com/docs/answer/6055540?hl=en>

Ex. แปลงหน่วยในชีวิตประจำวัน

D2	A	B	C	D	E
	ค่าตั้งต้น	หน่วยตั้งต้น	หน่วยใหม่	ค่าใหม่	
1	77	F	C	25	=CONVERT(77,"F","C")
2	80	yr	day	29220	=CONVERT(80,"yr","day")
3	58	kg	lbm	127.8681121	=CONVERT(58,"kg","lbm")

Formatting Data: Spreadsheets

Ex. แปลงประเภทข้อมูลจาก Number เป็น Currency (Thai Baht)

1

order_id	product_name	product_price
A0001	Playstation 5	23,000
A0002	PS5 Joystick	2,000
A0003	Lactasoy 125ml	5
A0004	Lactasoy 125ml	5
A0005	Lactasoy 125ml	5
A0001	Playstation 5	23,000
A0006	Lactasoy 150ml	5
A0007	Lactasoy 125ml	5
A0008	Lactasoy 125ml	5
A0009	Lactasoy 125ml	5

2

Thai Baht ฿1,000.12



	A	B	C	D
1	order_id	product_name	product_price	amount
2	A0001	Playstation 5	฿23,000.00	1
3	A0002	PS5 Joystick	฿2,000.00	2
4	A0003	Lactasoy 125ml	฿5.00	10,000
5	A0004	Lactasoy 125ml	฿5.00	20,000
6	A0005	Lactasoy 125ml	฿5.00	30,000
7	A0001	Playstation 5	฿23,000.00	1
8	A0006	Lactasoy 150ml	฿5.00	300
9	A0007	Lactasoy 125ml	฿5.00	40,000
10	A0008	Lactasoy 125ml	฿5.00	50,000
11	A0009	Lactasoy 125ml	฿5.00	100,000

หมายเหตุ: ถ้าไม่เจอประเภทข้อมูลในแบบที่ต้องการ
สามารถเลือกและปรับแต่งเองได้ที่เมนูเหล่านี้

Formatting Data: SQL

เปลี่ยนประเภทข้อมูลด้วย CAST ... AS ...

Syntax:

CAST(ชื่อคอลัมน์ AS ประเภทข้อมูลใหม่)

Note: ตัวอย่างประเภทข้อมูลที่ใช้บ่อยในเบื้องต้น

- STRING
- INT64, FLOAT64
- DATE, DATETIME, TIME, TIMESTAMP
- BOOL

ดูประเภทข้อมูลทั้งหมดได้ที่:

<https://cloud.google.com/bigquery/docs/reference/standard-sql/data-types>

Ex. แก้ไขข้อมูลวันสั่งซื้อ จากข้อความ (STRING) เป็น วัน (DATE) และ ข้อมูลจำนวนสั่งซื้อ จากจำนวนเต็ม (INT64) เป็นจำนวนทศนิยม

 RUN  SAVE  SHARE  SCHEDULE

```
1 SELECT
2   order_id,
3   customer_id,
4   CAST(order_date AS DATE),
5   product_id,
6   CAST(amount AS FLOAT64)
7 FROM projectonlineshop.transactions.txn_202206
```

Data Validation: Spreadsheets

เราสามารถ "บังคับ" ค่าที่เป็นไปได้ของข้อมูลที่กรอกเพื่อช่วยยืนยันความถูกต้องของข้อมูลได้

Ex. Drop-down list

The image shows a Google Sheets interface with a spreadsheet titled "New Members of Our Company (2022)". The spreadsheet has columns for Name, Role, and Go Outing. A red box labeled '1' highlights the 'Role' column header. A red box labeled '2' highlights the 'Data validation' option in the 'Data' menu. A red box labeled '3' highlights the 'Data validation' dialog box, which is open and shows the following settings:

- Cell range: Sheet30!B2:B9
- Criteria: List of items (Captain, Staff, Participant)
- Show dropdown list in cell
- On invalid data: Show warning Reject input
- Appearance: Show validation help text

The dialog box has 'Cancel', 'Remove validation', and 'Save' buttons. An arrow points from the dialog box to a table showing the result of the validation:

	A	B	C
1	Name	Role	Go Outing
2	Bambam	Staff	
3	Cherprang		
4	Elon	Captain	
5	Great	Staff	
6	Jennie	Participant	
7	Jisoo		
8	Larry		
9	Sergey		

Data Validation: Spreadsheets

เราสามารถ "บังคับ" ค่าที่เป็นไปได้ของข้อมูลที่กรอกเพื่อช่วยยืนยันความถูกต้องของข้อมูลได้

Ex. Checkbox

Data validation

Cell range:

Criteria:

Use custom cell values

Checked:

Unchecked:

On invalid data: Show warning Reject input

Appearance: Show validation help text:

[Reset](#)

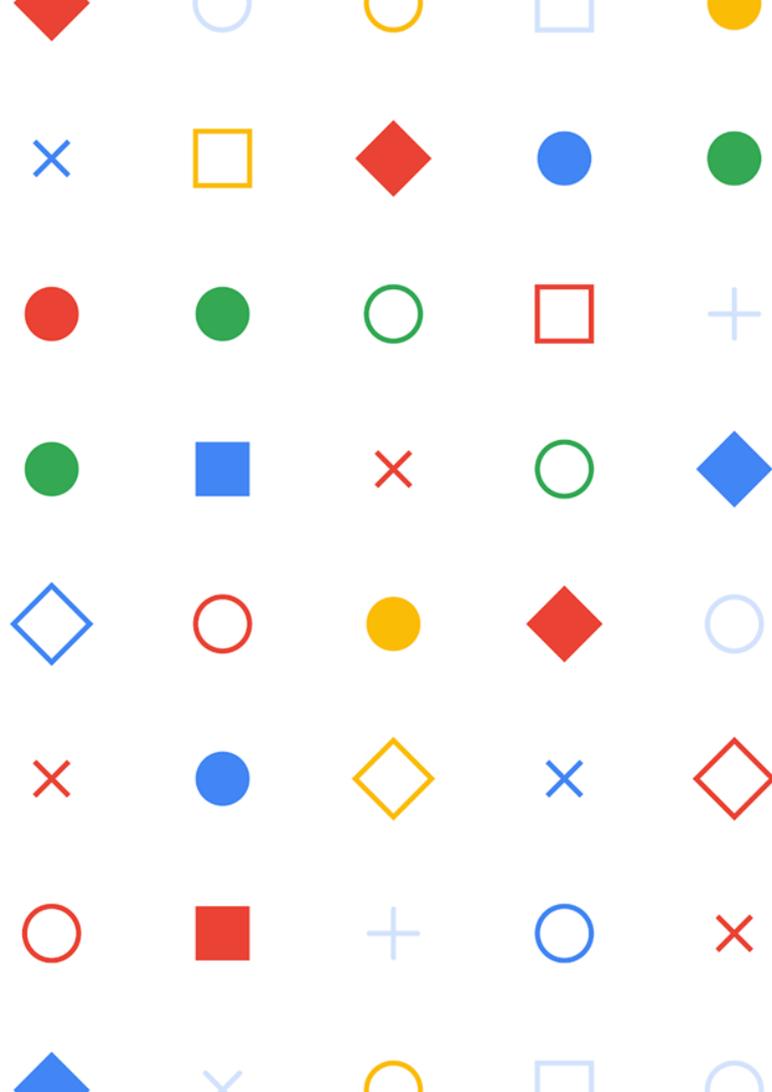


	A	B	C
1	Name	Role	Go Outing
2	Bambam	Staff	<input checked="" type="checkbox"/>
3	Cherprang	Captain	<input checked="" type="checkbox"/>
4	Elon	Staff	<input type="checkbox"/>
5	Great	Participant	<input checked="" type="checkbox"/>
6	Jennie	Staff	<input checked="" type="checkbox"/>
7	Jisoo	Staff	<input checked="" type="checkbox"/>
8	Larry	Participant	<input checked="" type="checkbox"/>
9	Sergey	Participant	<input type="checkbox"/>

Aggregating Data for Analysis

— *Part3:*

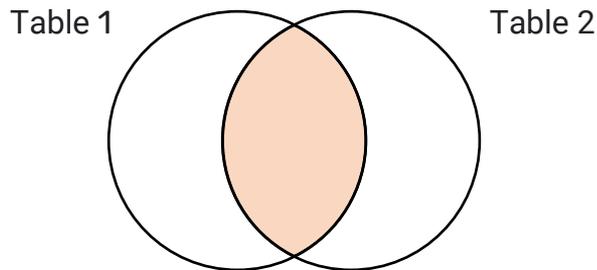
1. Four Types of Joins in SQL
2. Joining Tables in Spreadsheets
3. Group By in SQL
3. Subqueries



4 Types of Joins in SQL

เราสามารถนำตารางมาเชื่อมกันได้หลัก ๆ 4 แบบ

1. Inner Join : เอาเฉพาะ key ที่มีทั้ง 2 ตาราง



Ex. ต้องการตารางที่มีชื่อคน ชื่อวง และปีที่เดบิวต์ อยู่ในตารางเดียวกัน

Row	name	band	Row	band	debut
1	Bambam	IGOT7	1	IGOT7	2014
2	Cherprang	BNK48	2	BNK48	2017
3	Elon	SHIBA	3	SKD69	2016
4	Great	SKD69	4	BLPNK	2016
5	Jennie	BLPNK	5	BTS	2013
6	Jisoo	BLPNK	6	4EVE	2020
7	Larry	GOOGL			
8	Sergey	GOOGL			



Row	name	band	debut
1	Bambam	IGOT7	2014
2	Cherprang	BNK48	2017
3	Great	SKD69	2016
4	Jennie	BLPNK	2016
5	Jisoo	BLPNK	2016

```
1 SELECT m.name, m.band, b.debut
2 FROM
3     awesomeproject.company.members AS m
4     INNER JOIN awesomeproject.company.bands AS b
5     ON m.band = b.band
```

Notes: เราเรียกการใช้ "นามแฝง" แทนชื่อตารางยาว ๆ ว่าการ "Aliasing" และจริง ๆ แล้วไม่ต้องมีคำว่า AS ก็ได้ (แต่ควรใส่เพื่อความชัดเจน)

4 Types of Joins in SQL

เราสามารถนำตารางมาเชื่อมกันได้
หลัก ๆ 4 แบบ

2. Left Join : เอาเฉพาะ key ที่มีในตารางซ้าย

Table 1

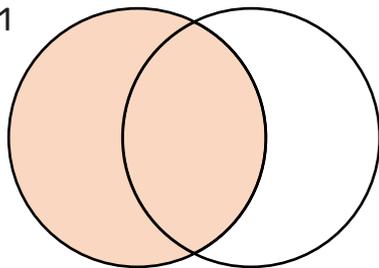


Table 2

Row	name	band	Row	band	debut
1	Bambam	IGOT7	1	IGOT7	2014
2	Cherprang	BNK48	2	BNK48	2017
3	Elon	SHIBA	3	SKD69	2016
4	Great	SKD69	4	BLPNK	2016
5	Jennie	BLPNK	5	BTS	2013
6	Jisoo	BLPNK	6	4EVE	2020
7	Larry	GOOGL			
8	Sergey	GOOGL			



Row	name	band	debut
1	Bambam	IGOT7	2014
2	Cherprang	BNK48	2017
3	Elon	SHIBA	<i>null</i>
4	Great	SKD69	2016
5	Jennie	BLPNK	2016
6	Jisoo	BLPNK	2016
7	Larry	GOOGL	<i>null</i>
8	Sergey	GOOGL	<i>null</i>

```
1 SELECT m.name, m.band, b.debut
2 FROM
3 awesomeproject.company.members AS m
4 LEFT JOIN awesomeproject.company.bands AS b
5 ON m.band = b.band
```

4 Types of Joins in SQL

เราสามารถนำตารางมาเชื่อมกันได้
หลัก ๆ 4 แบบ

3. Right Join : เอาเฉพาะ key ที่มีในตารางขวา

Table 1

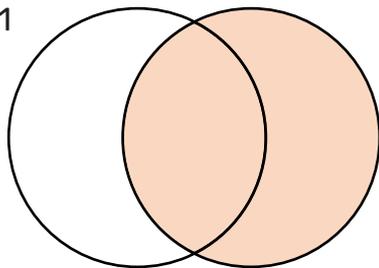


Table 2

Row	name	band	Row	band	debut
1	Bambam	IGOT7	1	IGOT7	2014
2	Cherprang	BNK48	2	BNK48	2017
3	Elon	SHIBA	3	SKD69	2016
4	Great	SKD69	4	BLPNK	2016
5	Jennie	BLPNK	5	BTS	2013
6	Jisoo	BLPNK	6	4EVE	2020
7	Larry	GOOGL			
8	Sergey	GOOGL			



Row	name	band	debut
1	Bambam	IGOT7	2014
2	Cherprang	BNK48	2017
3	Great	SKD69	2016
4	Jennie	BLPNK	2016
5	Jisoo	BLPNK	2016
6	<i>null</i>	BTS	2013
7	<i>null</i>	4EVE	2020

```
1 SELECT m.name, b.band, b.debut
2 FROM
3     awesomeproject.company.members AS m
4     RIGHT JOIN awesomeproject.company.bands AS b
5     ON m.band = b.band
```

4 Types of Joins in SQL

เราสามารถนำตารางมาเชื่อมกันได้
หลัก ๆ 4 แบบ

4. (Full) Outer Join : เอา key ทั้งหมด

Table 1

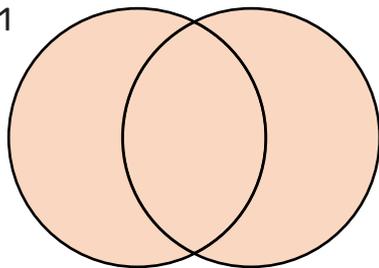


Table 2

Row	name	band	Row	band	debut
1	Bambam	IGOT7	1	IGOT7	2014
2	Cherprang	BNK48	2	BNK48	2017
3	Elon	SHIBA	3	SKD69	2016
4	Great	SKD69	4	BLPNK	2016
5	Jennie	BLPNK	5	BTS	2013
6	Jisoo	BLPNK	6	4EVE	2020
7	Larry	GOOGL			
8	Sergey	GOOGL			



Row	name	band	debut
1	Cherprang	BNK48	2017
2	<i>null</i>	BTS	2013
3	Jennie	BLPNK	2016
4	Jisoo	BLPNK	2016
5	Bambam	IGOT7	2014
6	Elon	SHIBA	<i>null</i>
7	Great	SKD69	2016
8	<i>null</i>	4EVE	2020
9	Larry	GOOGL	<i>null</i>
10	Sergey	GOOGL	<i>null</i>

```
1 SELECT m.name, COALESCE(m.band, b.band) AS band, b.debut
2 FROM
3 awesomeproject.company.members AS m
4 FULL OUTER JOIN awesomeproject.company.bands AS b
5 ON m.band = b.band
```

Joining Tables in Spreadsheets

VLOOKUP

หาข้อมูลในแนวดิ่ง (Vertical look-up) ถ้าเจอค่าที่ต้องการให้เอาค่าในคอลัมน์ทางขวามาเติมใส่ตารางอันใหม่

Syntax:

=VLOOKUP(ค่าค้นหา, data, หมายเลขคอลัมน์ที่ต้องการเอาค่ามาใส่ตารางใหม่, FALSE)

Notes:

- FALSE อันสุดท้ายจำเป็นต้องใส่เพื่อบอกว่าเราไม่ต้องการให้โปรแกรมค้นหาค่าใกล้เคียง แต่ต้องการให้ค้นหาค่าที่เหมือนเป๊ะตามค่าค้นหา
- ไม่สามารถเอาค่าทางซ้ายของค่าค้นหามาได้

<https://support.google.com/docs/table/25273?hl=en>

Ex. รวมชื่อจริงกับชื่อวงเข้าด้วยกัน

ใช้ Absolute Reference เพื่อใช้ data เดิมในการค้นหา

=VLOOKUP(E3, \$A\$3:\$B\$6, 2, FALSE)

	A	B	C	D	E	F	G
1							
2	Band	Debut		Name	Band	2014 x but	
3	IGOT7	2014		Bambam	IGOT7	=VLOOKUP(E3, \$A\$3:\$B\$6, 2, FALSE)	
4	BNK48	2017		Cherprang	BNK48		
5	SKD69	2016		Elon	SHIBA		
6	BLPNK	2016		Great	SKD69		
7				Jennie	BLPNK		
8			คอลัมน์ 1	คอลัมน์ 2	Jisoo	BLPNK	
9				Larry	GOOGL		
10				Sergey	GOOGL		

Result

Name	Band	Debut
Bambam	IGOT7	2014
Cherprang	BNK48	2017
Elon	SHIBA	#N/A
Great	SKD69	2016
Jennie	BLPNK	2016
Jisoo	BLPNK	2016
Larry	GOOGL	#N/A
Sergey	GOOGL	#N/A

← หาค่าไม่เจอ



Group By

การสรุปค่าทางสถิติแบบแบ่งตามกลุ่ม

Syntax:

SELECT ชื่อคอลัมน์ (Optional), การคำนวณ
ค่าสถิติแบ่งตามกลุ่ม

FROM ชื่อตาราง

GROUP BY ชื่อคอลัมน์ที่ใช้จัดกลุ่ม

HAVING เงื่อนไขค่าสถิติของกลุ่ม (Optional)

Ex. สรุปจำนวนครั้งที่ซื้อและจำนวนรหัสผลิตภัณฑ์ที่ซื้อ เอาเฉพาะคนที่ซื้อมากกว่า 1 ครั้ง

Row	order_id	customer_name	order_date	product_id	quantity
1	A0001	Great	2022-05-31	P0069	1
2	A0002	Great	2022-06-01	P0070	2
3	A0003	Elon	2022-06-01	P4234	10000
4	A0004	Elon	2022-06-02	P4234	20000
5	A0005	Great	2022-06-03	P4234	30000
6	A0006	Larry	2022-05-31	P0069	1
7	A0007	Great	2022-06-04	P4234	300
8	A0008	Elon	2022-06-04	P4234	40000
9	A0009	Elon	2022-06-05	P4234	50000
10	A0010	Great	2022-06-06	P4234	100000

← Larry มี num_orders
ไม่เกิน 1 ครั้ง

```
1 SELECT
2   customer_name,
3   COUNT(order_id) AS num_orders,
4   COUNT(DISTINCT product_id) AS num_products
5 FROM
6   awesomeproject.company.orders
7 GROUP BY customer_name
8 HAVING num_orders > 1 ←
```

Result

Row	customer_name	num_orders	num_products
1	Great	5	3
2	Elon	4	1

Subqueries

เราสามารถ query ภายใน query อีกที่หนึ่ง
ได้ โดยครอบ query ด้านใน (Inner query /
Inner select) ด้วยเครื่องหมายวงเล็บ

orders

Row	order_id	customer_name	order_date	product_id	quantity
1	A0001	Great	2022-05-31	P0069	1
2	A0002	Great	2022-06-01	P0070	2
3	A0003	Elon	2022-06-01	P4234	10000
4	A0004	Elon	2022-06-02	P4234	20000
5	A0005	Great	2022-06-03	P4234	30000
6	A0006	Larry	2022-05-31	P0069	1
7	A0007	Great	2022-06-04	P4234	300
8	A0008	Elon	2022-06-04	P4234	40000
9	A0009	Elon	2022-06-05	P4234	50000
10	A0010	Great	2022-06-06	P4234	100000

Ex. คำนวณยอดขายรวมจำนวนชิ้น (quantity) ที่ถูกสั่งซื้อของแต่ละ product_id

products

Row	product_id	product_name	product_price
1	P0069	Playstation 5	23000.0
2	P0070	PS5 Joystick	2000.0
3	P4234	Lactasoy 125ml	5.0
4	P5555	Gucci Belt	55555.0

```
1 SELECT
2   p.*,
3   q.total_quantity
4 FROM
5   awesomeproject.company.products AS p
6  LEFT JOIN
7  (
8    SELECT product_id, SUM(quantity) AS total_quantity
9    FROM awesomeproject.company.orders
10   GROUP BY product_id
11  ) AS q
12 ON p.product_id = q.product_id
```

คำนวณ Inner query
ให้จบก่อน แล้วค่อย
เอาผลไปใช้ต่อใน
Outer query ด้านนอก

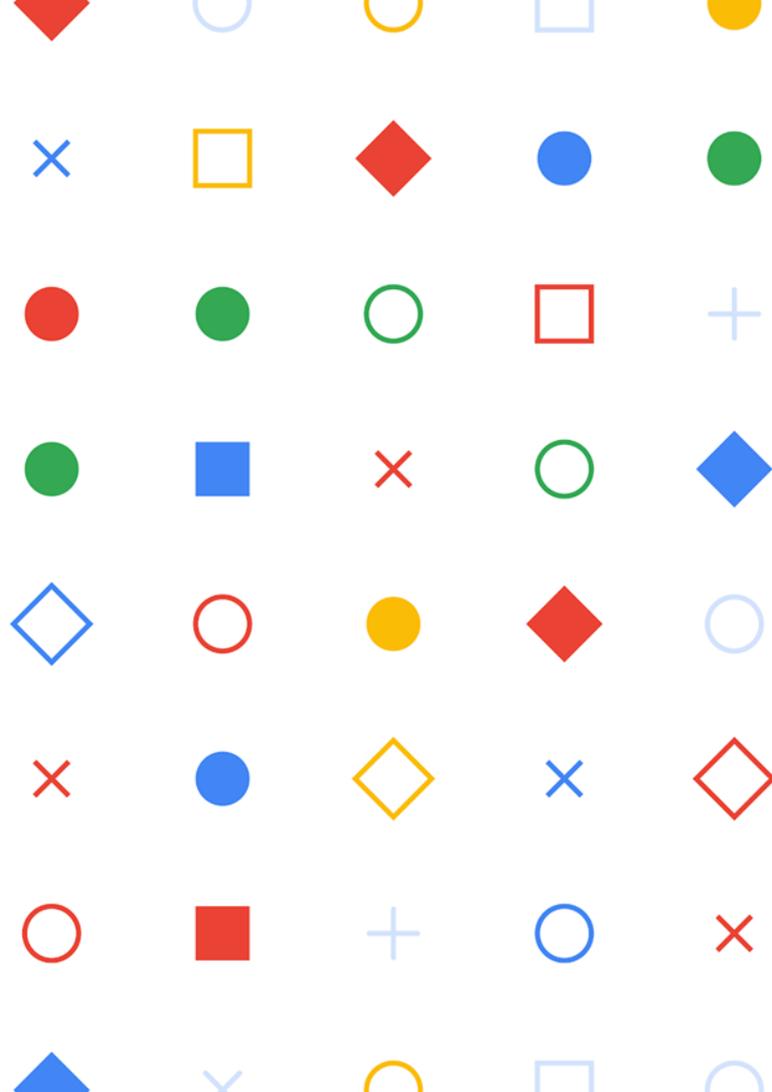
Result

Row	product_id	product_name	product_price	total_quantity
1	P0069	Playstation 5	23000.0	2
2	P0070	PS5 Joystick	2000.0	2
3	P4234	Lactasoy 125ml	5.0	250300
4	P5555	Gucci Belt	55555.0	null

Performing Data Calculations

— *Part4:*

1. Data Calculations with Conditions
2. Pivot Tables
3. Data Calculations in SQL
4. Six Types of Data Validation
5. Temporary Tables in SQL



Data Calculations with Conditions

COUNTIF

นับจำนวนช่องถ้าค่าในช่องตรงกับเงื่อนไขที่ต้องการ

Syntax:

`=COUNTIF(range, "criterion")`

Notes:

- เงื่อนไข (Criterion) ต้องใส่เครื่องหมาย quotes คร่อม
- มีฟังก์ชันที่คล้ายๆ กัน เช่น
 - COUNTIFS: ใส่ได้หลายเงื่อนไข
 - SUMIF: เปลี่ยนจากนับช่องเป็นหาผลรวม

<https://support.google.com/docs/table/25273?hl=en>

Ex. นับจำนวนคนที่มีอายุเกิน 30 ปี, จำนวนคนที่มีอายุน้อยกว่า 30 ปีและเป็นผู้หญิง, ความสูงมากที่สุดของคนไทยในตาราง

C11	A	B	C	D	E	F
	Name	Age	Height	Gender	Birthday	Nationality
1	Bambam	25	178	M	1997-05-02	Thai
2	Cherprang	26	160	F	1996-05-02	Thai
3	Elon	50	187	M	1971-06-28	
4	Great	31	172	M	1991-01-07	Thai
5	Jennie	26	163	F	1996-01-16	Korean
6	Jisoo		162	F	1995-01-03	
7	Larry	49	181	M	1973-03-26	American
8	Sergey	48	173	M	1973-06-21	American
9						
10						
11	No. of people with Age > 30		4			=COUNTIF(B2:B9, ">30")
12	No. of people with Age < 30 AND Gender = "F"		2			=COUNTIFS(B2:B9,"<30",D2:D9,"F")
13	Maximum Height of people with Nationality = "Thai"		178			=MAXIFS(C2:C9,F2:F9,"Thai")

Pivot Tables

ตารางสรุปข้อมูลสถิติแบบแบ่งตามกลุ่ม

Ex. สรุปพฤติกรรมการสั่งซื้อของลูกค้าแต่ละท่าน

	A	B	C	D	E	F
1	order_id	customer_name	order_date	product_id	product_price	quantity
2	A0001	Great	2022-05-31	P0069	฿23,000	1
3	A0002	Great	2022-06-01	P0070	฿2,000	2
4	A0003	Elon	2022-06-01	P4234	฿5	10,000
5	A0004	Elon	2022-06-02	P4234	฿5	20,000
6	A0005	Great	2022-06-03	P4234	฿5	30,000
7	A0006	Larry	2022-05-31	P0069	฿23,000	1
8	A0007	Great	2022-06-04	P4234	฿5	300
9	A0008	Elon	2022-06-04	P4234	฿5	40,000
10	A0009	Elon	2022-06-05	P4234	฿5	50,000
11	A0010	Great	2022-06-06	P4234	฿5	100,000

Result

customer_name	COUNTA of order_id	COUNTUNIQUE of product_id	SUM of quantity	Total Spent
Elon	4	1	120,000	฿600,000
Great	5	3	130,303	฿678,500
Larry	1	1	1	฿23,000
Grand Total	10	3	250,304	฿1,301,500

Pivot table editor

Rows

Add

customer_name

X

Order

Ascending

Sort by

customer_na...



Show totals

Values as: Columns

Add

order_id

X

Summarize by

COUNTA

Show as

Default

product_id

X

Summarize by

COUNTUNIQ...

Show as

Default

quantity

X

Summarize by

SUM

Show as

Default

Total Spent

X

Formula

=SUMPRODUCT(product_price, quantity)

Summarize by

Custom

Show as

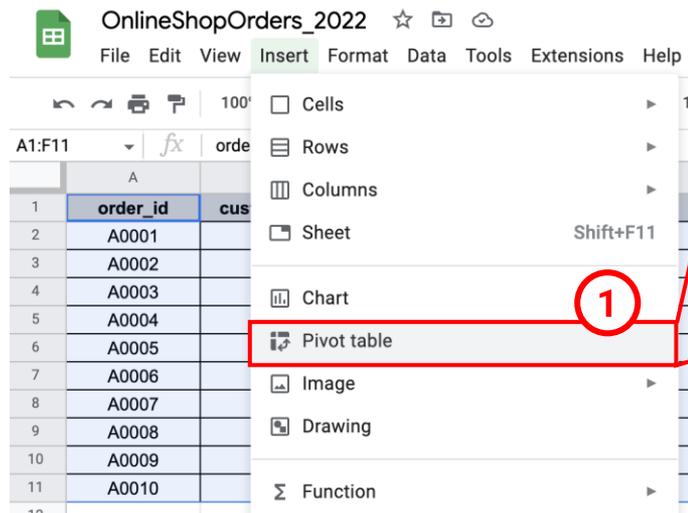
Default

← ใช้ Calculated Field

← คุณแล้วรวมผล

Pivot Tables

ขั้นตอนการสร้าง Pivot Table เบื้องต้นใน Google Sheets



OnlineShopOrders_2022

File Edit View Insert Format Data Tools Extensions Help

100%

A1:F11

	A	
1	order_id	cus
2	A0001	
3	A0002	
4	A0003	
5	A0004	
6	A0005	
7	A0006	
8	A0007	
9	A0008	
10	A0009	
11	A0010	
12		

Cells

Rows

Columns

Sheet

Chart

Pivot table

Image

Drawing

Function

Shift+F11

1

Create pivot table

Data range

Sheet1!A1:F11

Insert to

New sheet

Existing sheet

Cancel

Create

Pivot table editor

Sheet1!A1:F11

Suggested

Average of product_price for each customer_name

Count of order_id for each product_id

Sum of quantity for each customer_name

Rows

Add

order_id

customer_name

order_date

product_id

product_price

quantity

Values

Add

Calculated Field

Filters

order_id

customer_name

order_date

product_id

product_price

quantity

Data Calculations in SQL

เราสามารถนำค่าในคอลัมน์มา

- บวก + ลบ - คูณ * หาร /
- หาเศษจากการหาร % (เช่น 5 % 2 ได้ 1)
- คำนวณค่าสถิติอื่นๆ

Syntax:

```
SELECT
```

```
(col1 + col2)/2 AS xxx,  
(col2 - col1)*100/col1 AS yyy,  
SUM(col1*col2) AS zzz
```

```
FROM ชื่อตาราง
```

Ex. สรุปพฤติกรรมการสั่งซื้อของลูกค้าแต่ละคน แต่ไม่เอาคนชื่อ Larry

Row	order_id	customer_name	order_date	product_id	product_price	quantity
1	A0001	Great	2022-05-31	P0069	23000.0	1
2	A0002	Great	2022-06-01	P0070	2000.0	2
3	A0003	Elon	2022-06-01	P4234	5.0	10000
4	A0004	Elon	2022-06-02	P4234	5.0	20000
5	A0005	Great	2022-06-03	P4234	5.0	30000
6	A0006	Larry	2022-05-31	P0069	23000.0	1
7	A0007	Great	2022-06-04	P4234	5.0	300
8	A0008	Elon	2022-06-04	P4234	5.0	40000
9	A0009	Elon	2022-06-05	P4234	5.0	50000
10	A0010	Great	2022-06-06	P4234	5.0	100000

```
1 SELECT  
2   customer_name,  
3   COUNT(order_id) AS num_orders,  
4   COUNT(DISTINCT product_id) AS num_products,  
5   SUM(quantity) AS total_quantity,  
6   SUM(product_price*quantity) AS total_spent  
7 FROM awesomeproject.company.orders_detailed  
8 WHERE customer_name != 'Larry'  
9 GROUP BY customer_name
```

Result

Row	customer_name	num_orders	num_products	total_quantity	total_spent
1	Elon	4	1	120000	600000.0
2	Great	5	3	130303	678500.0

6 Types of Data Validation

การตรวจสอบข้อมูล (Data Validation) มีเป้าหมายเพื่อยืนยันความครบถ้วน (Complete) แม่นยำ (Accurate) สม่ำเสมอ (Consistent) และปลอดภัย (Secure) ซึ่งมีการตรวจสอบหลัก ๆ 6 อย่าง (พร้อมตัวอย่าง) คือ

1. Data type

ประเภทข้อมูล

ตรวจสอบว่าประเภทข้อมูลของแต่ละคอลัมน์ตรงตามที่จะควรจะเป็น

รายได้ควรเป็น FLOAT ไม่ใช่ STRING

2. Data range

ช่วงของข้อมูล

ตรวจสอบว่าข้อมูลมีค่าอยู่ในช่วงที่ควรจะเป็น

อายุมนุษย์ต้องมากกว่า 0 และไม่่าจะเกิน 120 ปี

3. Data constraints

ข้อจำกัดของข้อมูล

ตรวจสอบว่าข้อมูลมีค่าตรงตามเงื่อนไขหรือข้อจำกัด

จำนวนหลักบัตรประชาชน = 13 หลักพอดี

4. Data consistency

ความสม่ำเสมอของข้อมูล

ตรวจสอบว่าข้อมูล make sense ในบริบท (context) ที่มันอยู่

วันที่สั่งซื้อสินค้าครั้งแรกควรมากกว่าวันที่สมัครสมาชิก

5. Data structure

โครงสร้างข้อมูล

ตรวจสอบว่าข้อมูลเป็นไปตามโครงสร้างที่ได้ออกแบบไว้

คอลัมน์ครบ ชื่อคอลัมน์กับข้อมูลตรงกัน ไม่เหลือมกัน

6. Code validation

ความถูกต้องของโค้ด

ตรวจสอบว่าโค้ดของเราตรวจสอบข้อมูลได้ครบถ้วนทุกแบบ

โค้ดควรมี comment ที่บ่งชี้ว่าส่วนไหนตรวจสอบอะไร

Temporary Tables in SQL

เราสามารถสร้างตารางชั่วคราวขึ้นมาได้โดยไม่ต้อง CREATE TABLE ด้วยคำสั่ง WITH

Syntax:

```
WITH ชื่อตารางชั่วคราว AS (  
    SELECT ...  
    FROM ...  
)  
SELECT ...  
FROM ชื่อตารางชั่วคราว
```

Note: ตารางชั่วคราวจะถูกลบออกจากหน่วยความจำหลังจากจบ SQL session แต่ละครึ่ง

Ex. สรุปจำนวนครั้งที่ซื้อและปริมาณที่ซื้อของสินค้าแต่ละรหัสใน 3 วันแรกของเดือน มิ.ย.

temp table

Row	order_id	customer_name	order_date	product_id	product_price	quantity
1	A0002	Great	2022-06-01	P0070	2000.0	2
2	A0003	Elon	2022-06-01	P4234	5.0	10000
3	A0004	Elon	2022-06-02	P4234	5.0	20000
4	A0005	Great	2022-06-03	P4234	5.0	30000

```
1 WITH temp_table AS (  
2     SELECT *  
3     FROM awesomeproject.company.orders_detailed  
4     WHERE order_date BETWEEN "2022-06-01" AND "2022-06-03"  
5 )  
6  
7 SELECT  
8     product_id,  
9     COUNT(order_id) AS num_orders,  
10    SUM(quantity) AS total_quantity,  
11 FROM temp_table  
12 GROUP BY product_id  
13 ORDER BY total_quantity DESC
```

Result

Row	product_id	num_orders	total_quantity
1	P4234	3	60000
2	P0070	1	2

