

Data Transformation

Preparing data for analysis

การแปลงข้อมูลเพื่อเตรียมวิเคราะห์



Clean data is not always ready for analysis.

ข้อมูลที่สะอาดแล้ว
อาจยังไม่พร้อมสำหรับการวิเคราะห์

Today we will learn how to:

-  select only useful data
-  remove unnecessary rows or columns
-  create charts
-  combine information
-  calculate averages
-  group data into categories



From Clean Data to Useful Data

Cleaning fixes problems. Transformation answers questions.

การทำความสะอาดแก้ปัญหา การแปลงข้อมูลช่วยตอบคำถาม



Data Cleaning

Fix wrong or missing data.

แก้ข้อมูลผิดหรือ
ข้อมูลที่หายไป



Data Transformation

Change the data
format for the task.

เปลี่ยนรูปแบบข้อมูล
ให้เหมาะกับงาน



Example:

A large table

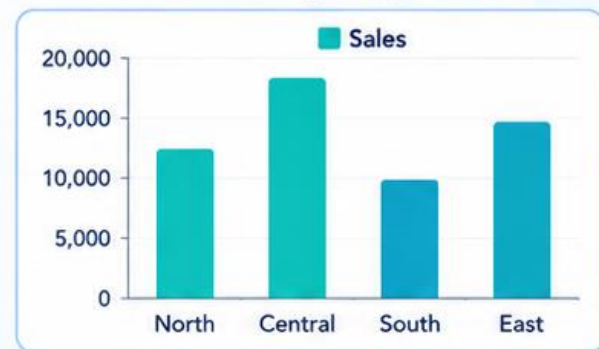


only the year 2015

Year: 2015			
Region	Sales	Quantity	Profit
North	12,500	320	2,150
Central	18,200	450	3,640
South	9,800	280	1,670
East	14,700	380	2,980



chart by region



What Is Data Transformation?

Changing data into a useful format

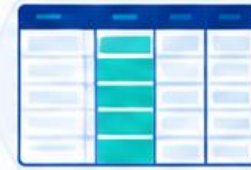
การเปลี่ยนข้อมูลให้อยู่ในรูปแบบที่ใช้ได้



Data transformation means changing data so it is ready for a question, calculation, chart, or analysis.

การแปลงข้อมูล คือ
การปรับข้อมูลให้พร้อมสำหรับคำถาม
การคำนวณ แผนภูมิ หรือการวิเคราะห์

Examples:



1. Keep only needed columns



2. Keep only needed years



3. Combine tables



4. Calculate an average



5. Group values into categories

Spreadsheet-Ready Data

Rows are records. Columns are attributes.

แถวคือข้อมูลแต่ละรายการ คอลัมน์คือคุณลักษณะ

Attributes (Columns)				
	Province	Region	Year	Average Income (THB)
1	Chiang Mai	North	2015	210,000
2	Bangkok	Central	2015	320,000
3	Phuket	South	2015	280,000
4	Khon Kaen	Northeast	2015	185,000
5	Chiang Rai	North	2015	165,000

Record (Row) = 1 item

Attribute (Column)

A good spreadsheet table should have:

- One **record** in each row
- One **attribute** in each column
- Clear **column names**
- The **same format** in each column
- No mixed units

Example: One record (row) looks like this:

Province	Region	Year	Average Income (THB)
Bangkok	Central	2015	320,000



This is now ready for **analysis** or a **chart**!

Types of Data Transformation

Different questions need different changes.

คำถามต่างกัน การแปลงข้อมูลก็ต่างกัน

1

Reduce Information

เลือกเฉพาะข้อมูลที่ต้องการ



2

Increase Information

เพิ่มข้อมูลที่ต้องการ



3

Integrate or Group Information

รวมข้อมูลหรือจัดกลุ่มข้อมูล



Choose the action that matches the question.

เลือกการแปลงข้อมูลที่เหมาะสมกับคำถาม

1. Reducing Information

Keep only what you need.

เลือกเฉพาะข้อมูลที่ต้องการ

Large dataset (too much information)

Province	Region	Year	Population	Average Income	Number of Schools	Area (km ²)
Bangkok	Central	2013	5,686,740	282,500	437	1,568
Chiang Mai	North	2013	1,662,926	169,200	325	20,107
Khon Kaen	Nontheast	2013	1,785,658	156,300	302	10,886
Phuket	South	2013	393,209	212,100	98	543
Chon Buri	East	2013	1,241,761	208,700	245	4,363
...	...	...	...	...	...	...
Bangkok	Central	2015	5,817,427	336,120	462	1,568
Chiang Mai	North	2015	1,679,588	217,450	341	20,107
Khon Kaen	Northeast	2015	1,823,980	189,870	319	10,886
Phuket	South	2015	402,762	248,910	102	543
Chon Buri	East	2015	1,285,792	274,330	255	4,363

Filter



Selected data (only what we need)

Province	Average Income	Year
Bangkok	336,120	2015
Chiang Mai	217,450	2015
Khon Kaen	189,870	2015
Phuket	248,910	2015
Chon Buri	274,330	2015



Now it is ready for analysis or chart.

Examples:



Keep one year



Keep one region



Remove unnecessary provinces



Remove unnecessary columns



Key idea:

Keep only what you need for the question.

เก็บเฉพาะสิ่งที่ต้องใช้ในการตอบคำถาม

Task 1: Compare Income by Region in 2015

Use Excel to reduce data and make a chart.

ใช้ Excel เลือกข้อมูลและสร้างแผนภูมิ



Goal

Compare the average household income of all regions in 2015.

เปรียบเทียบรายได้เฉลี่ยต่อครัวเรือน
ของทุกภูมิภาคในปี 2015



Steps:

- 1 Open the average household income dataset.
- 2 Keep only the region rows.
- 3 Keep only the year 2015.
- 4 Remove unnecessary years and provinces.
- 5 Make a chart.

Before: Large dataset (many provinces and years)

Province	Region	Year	Average Income (THB)
Bangkok	Central	2013	282,500
Bangkok	Central	2014	310,230
Bangkok	Central	2015	336,120
...	...	...	...
Chiang Mai	North	2013	169,200
Chiang Mai	North	2014	193,300
Chiang Mai	North	2015	217,450
...	...	...	...
Songkhla	South	2013	145,800
Songkhla	South	2014	164,100
Songkhla	South	2015	183,220
...	...	...	...

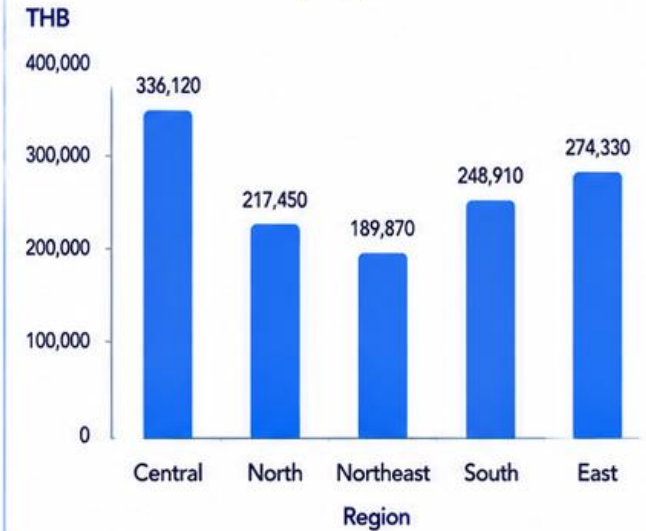


After: Keep only regions in 2015

Region	Year	Average Income (THB)
Central	2015	336,120
North	2015	217,450
Northeast	2015	189,870
South	2015	248,910
East	2015	274,330

Sample Chart (Example)

Average Household Income by Region (2015)



Output:

A chart comparing regions in 2015.

แผนภูมิเปรียบเทียบรายได้ของแต่ละภูมิภาคในปี 2015



Task 2: Compare Income in the North in 2013

Focus on one region and one year.

โฟกัสหนึ่งภูมิภาคและหนึ่งปี

Goal



Compare average household income of all provinces in the North in 2013.

เปรียบเทียบรายได้เฉลี่ยต่อครัวเรือน
ของทุกจังหวัดในภาคเหนือในปี 2013



Steps:

- 1 Keep only provinces in the North.
- 2 Keep only the year 2013.
- 3 Remove unnecessary years and regions.
- 4 Make a chart.

Filter the data: Region = **North** AND Year = **2013**



Region

North



Year

2013

Before: Large dataset

Province	Region	Year	Average Income (THB)
Chiang Mai	North	2012	157,600
Chiang Mai	North	2013	169,200
Chiang Mai	North	2014	193,300
...	...	...	...
Lampang	North	2012	132,400
Lampang	North	2013	142,800
Lampang	North	2014	160,600
...	...	...	...
Phuket	South	2013	212,100
Bangkok	Central	2013	282,500
...	...	...	...

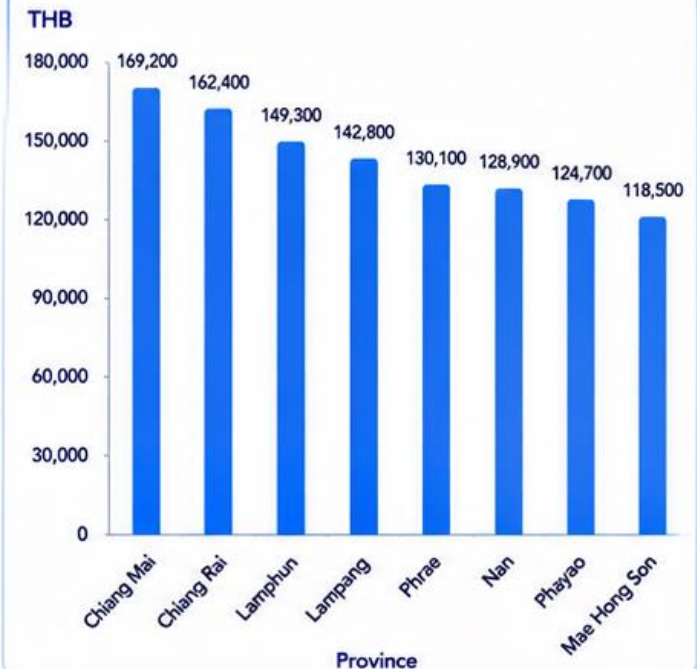


After: North region, Year 2013 only

Province	Average Income (THB)
Chiang Mai	169,200
Lamphun	149,300
Lampang	142,800
Phrae	130,100
Nan	128,900
Phayao	124,700
Chiang Rai	162,400
Mae Hong Son	118,500

Sample Chart (Example)

Average Household Income in the North (2013)



Output:

A chart comparing northern provinces in 2013.
แผนภูมิเปรียบเทียบรายได้ของจังหวัดในภาคเหนือในปี 2013



Task 3: Change in Kingdom Income

Compare values over time.

เปรียบเทียบข้อมูลตามเวลา



Goal

Study the change in average household income in the Kingdom from 1998 to 2015.

ศึกษาการเปลี่ยนแปลงของรายได้เฉลี่ยต่อครัวเรือนทั้งประเทศตั้งแต่ปี 1998 ถึง 2015



Steps:

- 1 Keep only the Kingdom row.
- 2 Keep the needed years from 1998 to 2015.
- 3 Remove unnecessary rows.
- 4 Make a line chart.

Transform: Keep Kingdom row and years 1998–2015

Before: Full dataset (many rows and years)

Geography	1998	1999	...	2013	2014	2015
Bangkok	75,120	77,540	...	282,500	310,230	336,120
Central	58,230	60,810	...	210,200	230,300	248,700
...	...	...	...	...	...	...
North	35,640	36,810	...	145,800	164,100	169,200
South	38,210	39,620	...	156,200	176,800	182,600
Kingdom	48,640	50,820	...	189,870	217,450	248,910



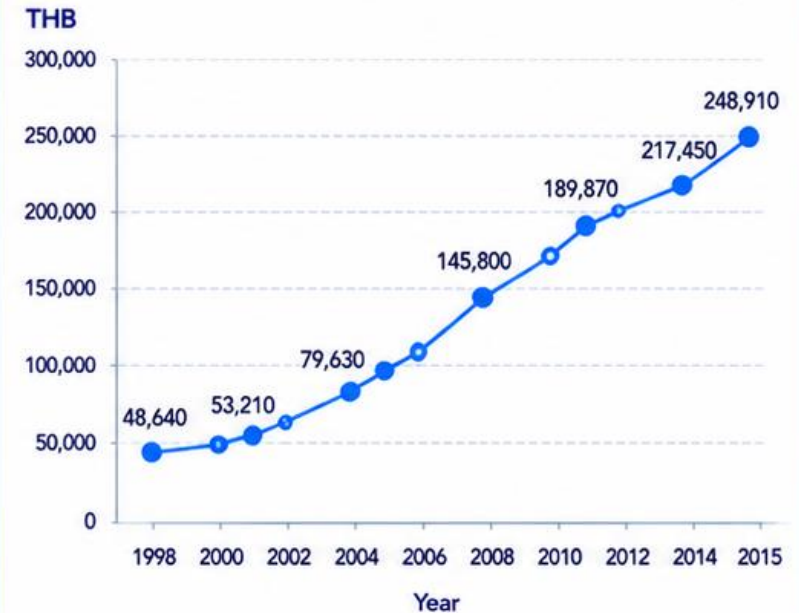
After: Only Kingdom row, years 1998–2015

Year	1998	1999	2000	...	2013	2014	2015
Kingdom	48,640	50,820	53,210	...	189,870	217,450	248,910

Unit: THB

Sample Chart (Example)

Average Household Income in the Kingdom (1998–2015)



Output:

A line chart showing income change over time.

แผนภูมิเส้นแสดงการเปลี่ยนแปลงรายได้ตามเวลา



2. Increasing Information

Add the data needed to answer the question.

เพิ่มข้อมูลที่จำเป็นเพื่อตอบคำถาม

Current data (not enough)

Province	Region	Year	Average Income (THB)
Chiang Mai	North	2015	217,450
Lampang	North	2015	142,800
Phrae	North	2015	130,100
Nan	North	2015	128,900
...	...	...	...



Sometimes one table does not have enough information.

บางครั้งตารางเดียวมีข้อมูลไม่พอ



Increasing information means:



Finding another useful table



Copying needed attributes



Combining matching data



Calculating a new result



How we match data:

Use the same key such as Province and Year.



After: More useful information

Province	Region	Year	Average Income (THB)	Population (people)
Chiang Mai	North	2015	217,450	1,780,900
Lampang	North	2015	142,800	730,700
Phrae	North	2015	130,100	430,900
Nan	North	2015	128,900	472,100
...	...	...	...	...



Now we have more information to answer the question.

ตอนนี้เรามีข้อมูลเพียงพอที่จะตอบคำถามแล้ว



Warning:

Only add correct and relevant information.

เพิ่มเฉพาะข้อมูลที่ถูกต้องและเกี่ยวข้อง



Task 4: Average Weight of Male Students

Combine data and calculate an average.

รวมข้อมูลและคำนวณค่าเฉลี่ย

Goal



Use the collected data of M3/17, M3/18, and M3/19 to find the average weight of male students.

ใช้ข้อมูลของ M3/17, M3/18 และ M3/19 เพื่อหาน้ำหนักเฉลี่ยของนักเรียนชาย



Needed attributes:

- gender
- weight



Steps:

- 1 Create a new Excel table.
- 2 Copy only male students.
- 3 Keep the Weight column.
- 4 Calculate the average weight.

Source data: Three classes (M3/17, M3/18, M3/19)

M3/17				M3/18				M3/19			
ID	Gender	Weight (kg)	Height (cm)	ID	Gender	Weight (kg)	Height (cm)	ID	Gender	Weight (kg)	Height (cm)
17-01	Male	52.0	162	18-01	Male	58.0	168	19-01	Male	57.0	169
17-02	Female	45.0	158	18-02	Male	53.0	163	19-02	Female	46.0	156
17-03	Male	60.0	170	18-03	Female	44.0	157	19-03	Male	63.0	172
17-04	Male	55.5	165	18-04	Male	61.0	171	19-04	Male	54.0	164
17-05	Female	48.0	160	18-05	Female	47.0	159	19-05	Female	45.5	158
...	...	...	...	...	...	...	...	...	...	...	...



Keep only male students

New table (combined and filtered)

Gender	Weight (kg)
Male	52.0
Male	60.0
Male	55.5
Male	58.0
Male	53.0
Male	61.0
Male	57.0
Male	63.0
Male	54.0
...	...
Average weight	56.8 kg



Use **AVERAGE** to calculate the result.
ใช้ฟังก์ชัน AVERAGE เพื่อคำนวณผลลัพธ์

Example formula

=AVERAGE(B2:B100)



We only need Gender to filter and Weight to calculate.
เราใช้ Gender เพื่อคัดกรอง และ Weight เพื่อคำนวณค่าเฉลี่ย

Task 5: Average Height of Female Students

Use the correct attribute for the question.

ใช้คุณลักษณะให้ตรงกับคำถาม

Goal



Use the collected data of M3/17, M3/18, and M3/19 to find the average height of female students.

ใช้ข้อมูลของ M3/17, M3/18 และ M3/19 เพื่อหาส่วนสูงเฉลี่ยของนักเรียนหญิง



Needed attributes:

- gender
- height



Steps:

- 1 Create a new Excel table.
- 2 Copy only female students.
- 3 Keep the Height column.
- 4 Calculate the average height.

Source data: Three classes (M3/17, M3/18, M3/19)

M3/17				M3/18				M3/19			
ID	Gender	Weight (kg)	Height (cm)	ID	Gender	Weight (kg)	Height (cm)	ID	Gender	Weight (kg)	Height (cm)
17-01	Male	52.0	162	18-01	Male	58.0	168	19-01	Male	57.0	169
17-02	Female	45.0	158	18-02	Male	53.0	163	19-02	Female	46.0	156
17-03	Male	60.0	170	18-03	Female	44.0	157	19-03	Male	63.0	172
17-04	Male	55.5	165	18-04	Male	61.0	171	19-04	Male	54.0	164
17-05	Female	48.0	160	18-05	Female	47.0	159	19-05	Female	45.5	158
...	...	...	...	...	...	...	...	...	...	...	...



Keep only female students

New table (combined and filtered)

Gender	Height (cm)
Female	158
Female	160
Female	157
Female	159
Female	156
Female	158
Female	156
Female	157
...	...
Average height	157.6 cm



Use **AVERAGE** to calculate average height.

ใช้ฟังก์ชัน AVERAGE เพื่อคำนวณส่วนสูงเฉลี่ย

Example formula

=AVERAGE(B2:B100)



We only need Gender to filter and Height to calculate.

เราใช้ Gender เพื่อคัดกรอง และใช้ Height เพื่อคำนวณค่าเฉลี่ย

3. Data Integration

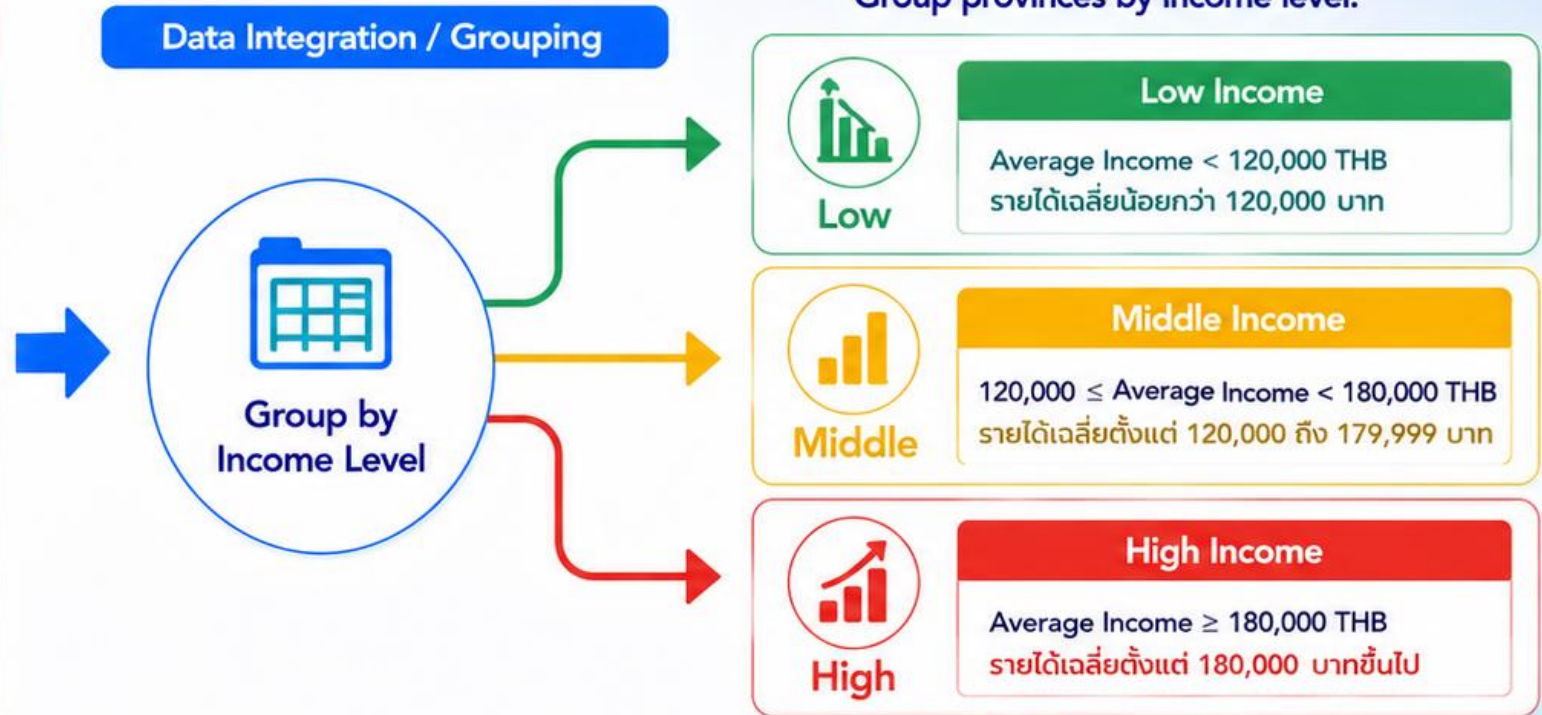
Group data using clear conditions.

จัดกลุ่มข้อมูลโดยใช้เงื่อนไขที่ชัดเจน

Example: Province income data (2015)

Province	Region	Average Income (THB)
Chiang Mai	North	169,200
Lampang	North	142,800
Phrae	North	130,100
Nakhon Sawan	Central	115,300
Chonburi	East	171,900
Bangkok	Central	282,500
Phuket	South	212,100
Yala	South	95,800
...	...	...

Data Integration / Grouping



Data integration can mean grouping data based on an attribute.

การรวมข้อมูลอาจหมายถึงการจัดกลุ่มข้อมูลตามคุณลักษณะ



easier to compare
เปรียบเทียบได้ง่าย



easier to summarize
สรุปได้ง่าย



easier to explain
patterns
อธิบายรูปแบบได้ง่าย

Why group data?

Task 6: Divide Provinces by Income Category

Use conditions to create groups.

ใช้เงื่อนไขเพื่อจัดกลุ่ม

Income Categories (2015)

ช่วงรายได้และกลุ่ม



Low

15,001 - 20,000
baht

รายได้เฉลี่ยต่อครัวเรือน (บาท)



Middle

20,001 - 25,000
baht

รายได้เฉลี่ยต่อครัวเรือน (บาท)



High

25,001 - 30,000
baht

รายได้เฉลี่ยต่อครัวเรือน (บาท)

Source: Average Household Income Table (2015)

ข้อมูลต้นฉบับ

Province	Region	2015 Income (THB)
Bangkok	Central	28,250
Nonthaburi	Central	27,180
Pathum Thani	Central	22,340
Chiang Mai	North	20,150
Lampang	North	18,720
Udon Thani	Northeast	17,480
Khon Kaen	Northeast	19,250
Phuket	South	24,520
Surat Thani	South	21,430
...	...	...



Use conditions
to create
groups
ใช้เงื่อนไข
เพื่อจัดกลุ่ม

Output: Provinces with Income Category (2015)

ผลลัพธ์: เพิ่มกลุ่มรายได้

Province	Region	2015 Income (THB)	Category
Bangkok	Central	28,250	High
Nonthaburi	Central	27,180	High
Pathum Thani	Central	22,340	Middle
Chiang Mai	North	20,150	Middle
Lampang	North	18,720	Low
Udon Thani	Northeast	17,480	Low
Khon Kaen	Northeast	19,250	Low
Phuket	South	24,520	Middle
Surat Thani	South	21,430	Middle
...	...	...	...



Steps:

- 1 Keep province and 2015 income.
- 2 Create a new Category column.
- 3 Use the income ranges to label each province.
- 4 Count or compare the groups.

เลือกเฉพาะชื่อจังหวัดและรายได้ปี 2015

สร้างคอลัมน์ Category

ใช้ช่วงรายได้เพื่อกำหนดกลุ่มให้แต่ละจังหวัด

นับจำนวนหรือเปรียบเทียบแต่ละกลุ่ม



What can we do after grouping?

- Count the number of provinces in each group.
นับจำนวนจังหวัดในแต่ละกลุ่ม
- Compare the groups.
เปรียบเทียบแต่ละกลุ่ม
- Describe patterns.
อธิบายรูปแบบหรือแนวโน้ม



Useful Excel Tools

Tools that help transform data.

เครื่องมือที่ช่วยแปลงข้อมูล

1 Filter



Show only needed rows.

แสดงเฉพาะแถวที่ต้องการ

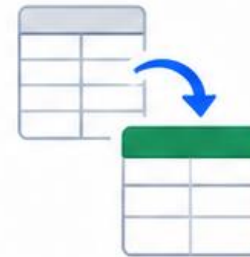
2 Delete/Hide



Remove unnecessary columns.

ลบหรือซ่อนคอลัมน์ที่ไม่จำเป็น

3 Copy/Paste



Create a working table.

คัดลอกและวางเพื่อสร้างตารางสำหรับใช้งาน

4 Sort



Order values (A→Z or Z→A).

จัดเรียงข้อมูลจากน้อยไปมากหรือมากไปน้อย

5 AVERAGE

fx

`=AVERAGE()`

Calculate mean value.

คำนวณค่าเฉลี่ยของข้อมูล

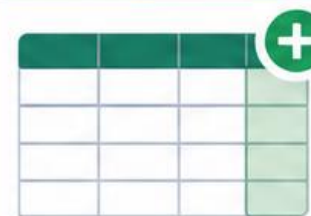
6 Chart



Visualize results.

แสดงผลลัพธ์ในรูปแบบแผนภูมิ

7 New Column



Create categories or new information.

สร้างคอลัมน์ใหม่สำหรับจัดหมวดหมู่หรือข้อมูลเพิ่มเติม



Choose the right tool for the question to transform data correctly and efficiently.

เลือกเครื่องมือให้เหมาะสมกับคำถาม เพื่อแปลงข้อมูลได้อย่างถูกต้องและมีประสิทธิภาพ

Transformation Decisions

What should we keep, remove, add, or group?

เราควรเก็บ ลบ เพิ่ม หรือจัดกลุ่มอะไร?



1 What question am I answering?
ฉันกำลังตอบคำถามอะไร?



Understand the goal.
เข้าใจเป้าหมายของคำถาม



2 Which rows do I need?
ต้องใช้แถวไหน?



Keep only the rows that help
answer the question.
เก็บเฉพาะแถวที่ช่วยตอบคำถาม



3 Which columns do I need?
ต้องใช้คอลัมน์ไหน?



Keep only the columns that
provide useful information.
เก็บเฉพาะคอลัมน์ที่ให้ข้อมูลที่จำเป็น



4 Do I need a new calculation?
ต้องคำนวณค่าใหม่หรือไม่?



Add a formula to calculate
a new result.
เพิ่มสูตรเพื่อคำนวณผลลัพธ์ใหม่



5 Do I need categories?
ต้องจัดกลุ่มหรือไม่?



Group data using clear
conditions.
จัดกลุ่มข้อมูลโดยใช้เงื่อนไขที่ชัดเจน



Let the question guide every transformation.
ให้คำถามเป็นตัวนำทางในการแปลงข้อมูลทุกครั้ง



Mini Check: What Transformation Is Needed?

Choose reduce, increase, or group.

เลือก ลดข้อมูล เพิ่มข้อมูล หรือจัดกลุ่ม



Choose the best transformation action:
เลือกประเภทการแปลงข้อมูลที่เหมาะสมที่สุด

1



Keep only
the year 2015.



4



Remove
unnecessary
provinces.



2



Combine three
class tables.



5



Calculate
average height.



3



Add a new
Category column.



6



Divide income
into Low,
Middle, High.



Options (Choose one)
ตัวเลือก (เลือกหนึ่งข้อ)



Reduce



Increase



Group



Pair Work: 3 minutes
ทำงานเป็นคู่: 3 นาที



Classwork Task Checklist

Use Excel to complete the transformation tasks.

ใช้ Excel ทำงานแปลงข้อมูล



Complete the tasks in order:
ทำงานตามลำดับต่อไปนี้

Task 1



Compare regions in 2015

☐

Task 2



Compare northern provinces in 2013

☐

Task 3



Kingdom income change 1998–2015

☐

Task 4



Average weight of male students

☐

Task 5



Average height of female students

☐

Task 6



Income categories for provinces

☐

Submit:

ส่งงานประกอบด้วย



Transformed table
ตารางที่แปลงแล้ว



Chart or result
กราฟหรือผลลัพธ์



Short answer explaining
what you transformed
คำอธิบายสั้น ๆ ว่าคุณแปลง
ข้อมูลอะไรบ้างและอย่างไร



Important: Save your Excel file regularly.

สำคัญ: บันทึกไฟล์ Excel ของคุณอย่างสม่ำเสมอ



File name example:

M3_Transformation_Tasks_YourName.xlsx



What Did We Learn Today?

Transformation makes data useful for analysis.

การแปลงข้อมูลทำให้ข้อมูลพร้อมวิเคราะห์



Today we learned:

1



Data transformation changes data for a specific task.

การแปลงข้อมูลจะเปลี่ยนข้อมูลเพื่อให้งานที่ต้องการสำเร็จ

2



Spreadsheet data should have clear rows and columns.

ข้อมูลในสเปรดชีตควรมีแถวและคอลัมน์ที่ชัดเจน

3



We can reduce information by keeping only what we need.

เราสามารถลดข้อมูลโดยเก็บเฉพาะสิ่งที่เราต้องการ

4



We can increase information by combining useful data.

เราสามารถเพิ่มข้อมูลโดยการรวมข้อมูลที่มีประโยชน์

5



We can integrate data by grouping values into categories.

เราสามารถรวมข้อมูลโดยการจัดกลุ่มค่าเป็นหมวดหมู่

6



Excel tools help us transform and analyze data.

เครื่องมือของ Excel ช่วยให้เราแปลงและวิเคราะห์ข้อมูล



Question
คำถาม



Transform
แปลงข้อมูล



Chart /
Calculation
กราฟ / การคำนวณ



Insight
ความเข้าใจ

Exit Ticket

Before you leave...

ก่อนจบคาบเรียน...



Answer in your notebook:

ตอบคำถามในสมุดของคุณ

- 1 What is data transformation?
- 2 Give one example of reducing information.
- 3 Give one example of increasing information.
- 4 Why do we need clear rows and columns in a spreadsheet?
- 5 Which Excel tool did you use today, and why?



Transform data to match the question.

แปลงข้อมูลให้ตรงกับคำถาม

