Data Science

Using Data to Understand PM2.5 Air Pollution

Can data help us understand real problems?

From Last Lesson

Technology changed how we live, communicate, and collect information.

Quick Review

1. Technology helps us collect information faster.
2. Digital systems create large amounts of data.
3. Data can help people understand real-world problems.

What problem could we study using data?

What is Data Science?

A simple definition

Data Science means using data to:

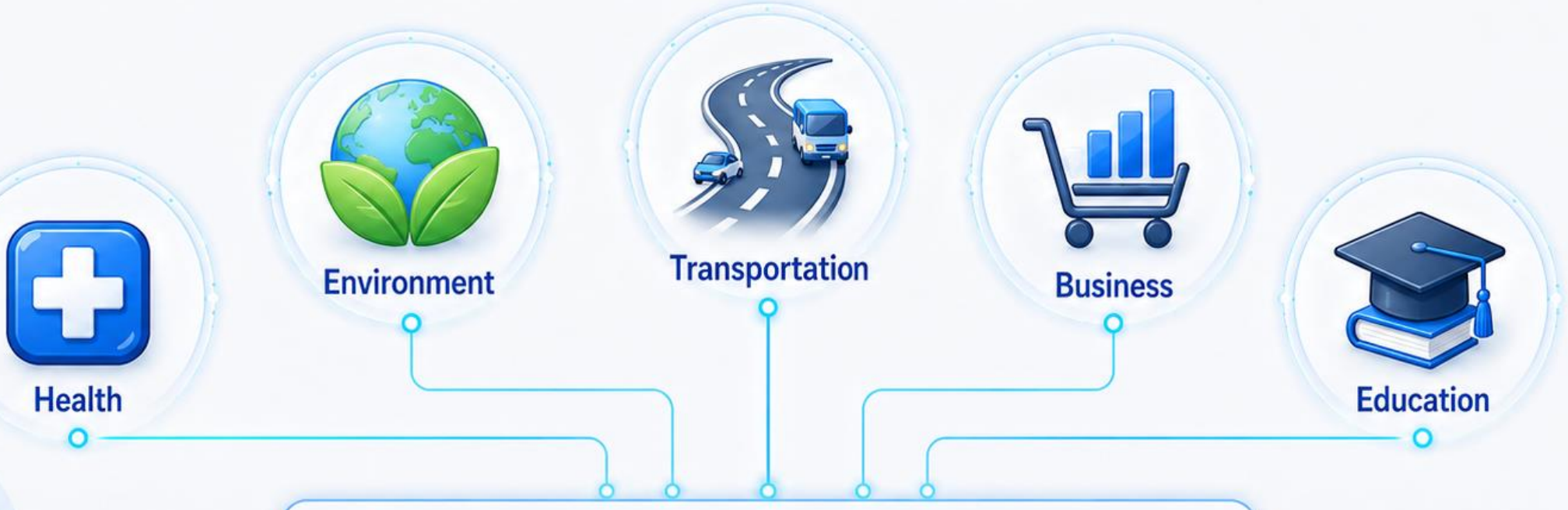
- find patterns
- understand problems
- make better decisions

We do not guess. We use evidence.

Why Do We Need Data Science?

Some problems are too big to solve by guessing.

Data Science helps people understand problems in:



Big problems need evidence, not guesses.

Case Study: PM2.5 Air Pollution

Can data help us understand the air we breathe?

Problem

PM2.5 air pollution can affect daily life.

Question

When is the air most polluted?

Goal

Use data to find patterns.

What Data Do We Need?

Before analyzing, we must know what was collected.

PM2.5 Investigation Data	
Data Collected	**Example**
PM2.5 level	35 µg/m³
Time	8:00 AM
Date	25–27 June
Location	Road / intersection

Good analysis starts with understanding the data.

How Is the Data Collected?

IoT sensors can measure air quality automatically.

IoT = Internet of Things

Sensor

Internet

Database

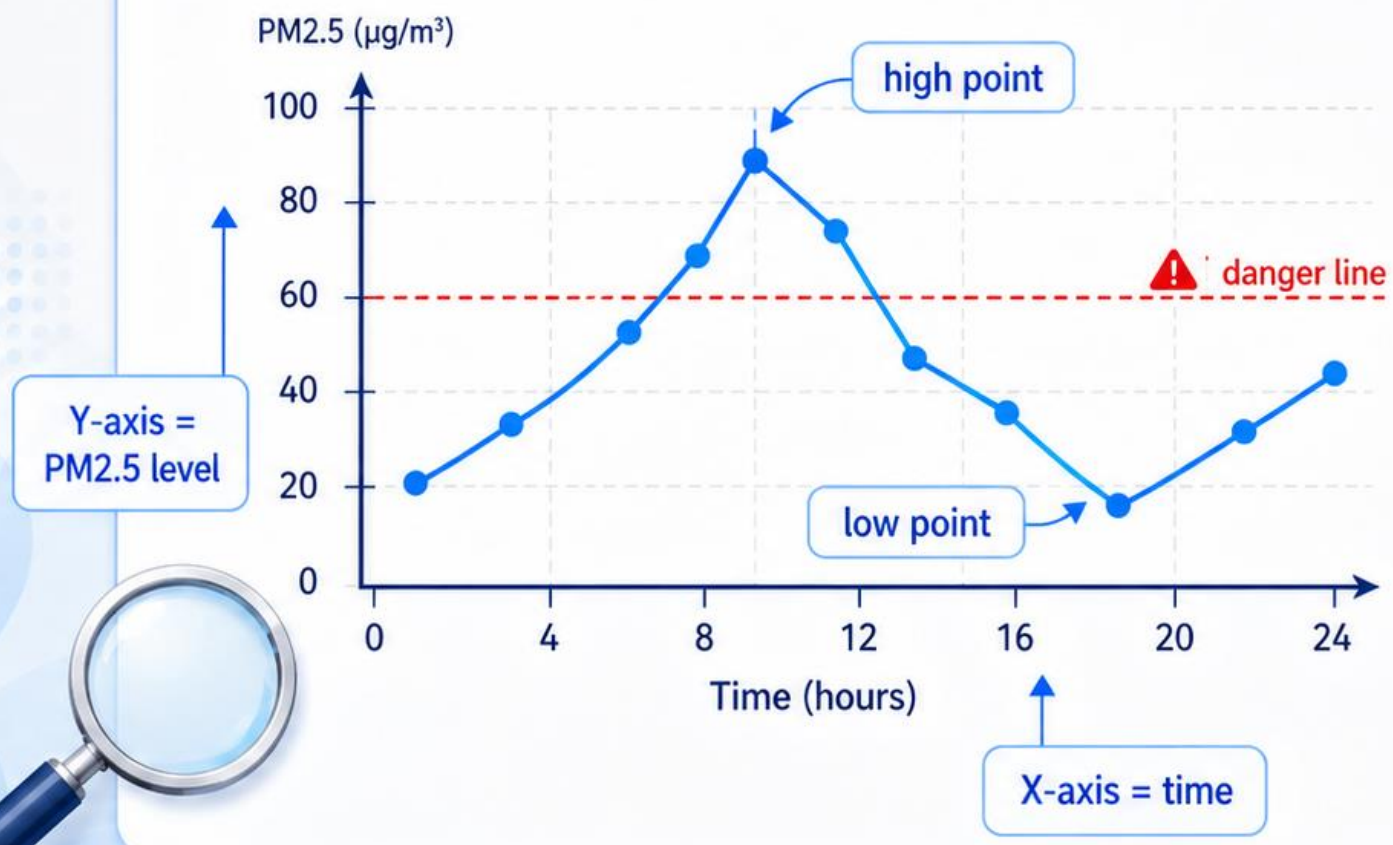
Dashboard

- automatically
- repeatedly
- real time
- many locations

Before Reading the Graph

Every graph has important parts.

Example: PM2.5 Over Time



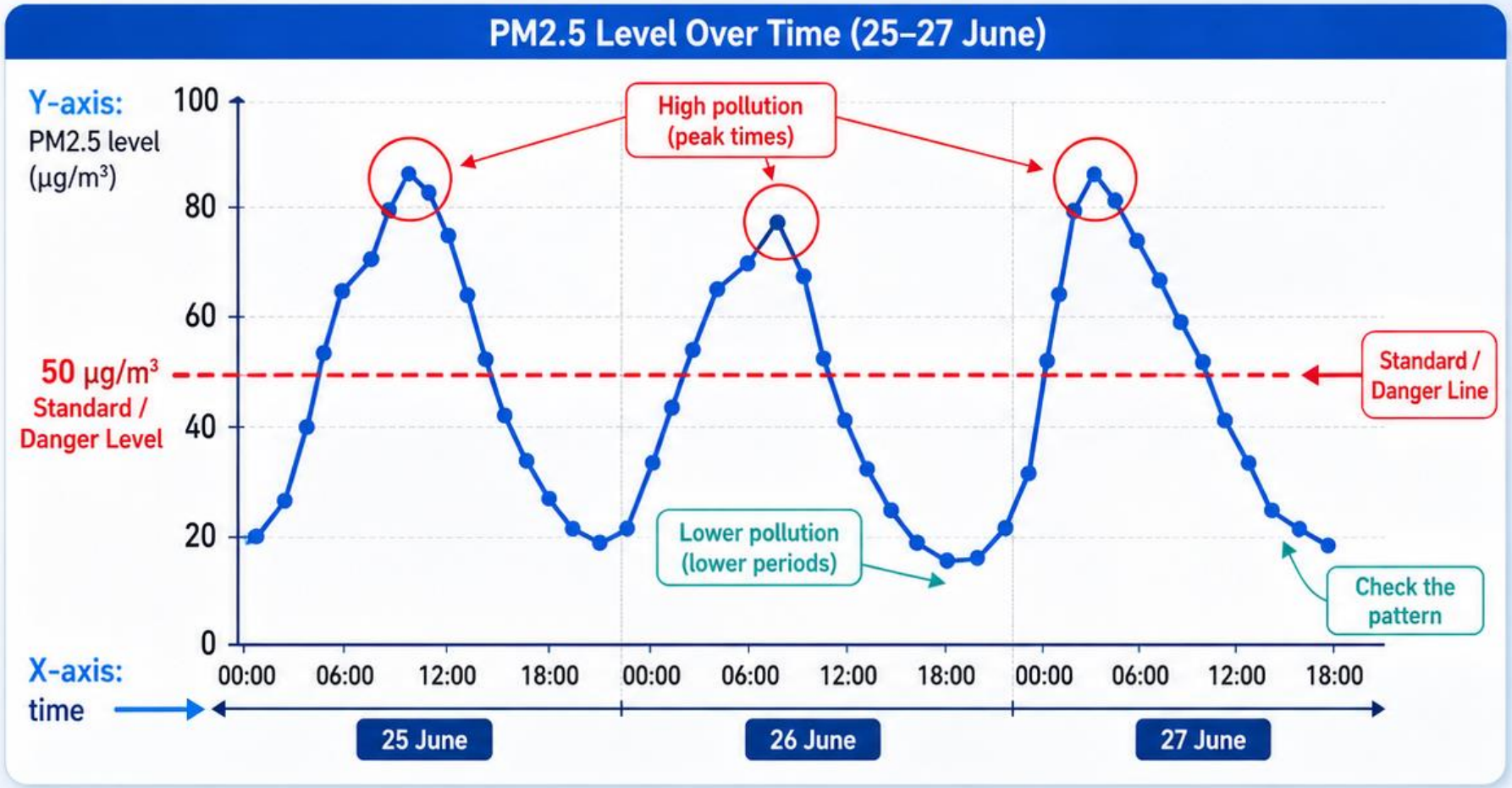
Look for:

- X-axis = time
- Y-axis = PM2.5 level
- high points
- low points
- repeated patterns
- danger line

What should we check first?

PM2.5 Data Graph

Dust level between 25–27 June



X-axis: time

Y-axis: PM2.5 level (µg/m³)

Red line: standard / danger level

Peaks: high pollution times

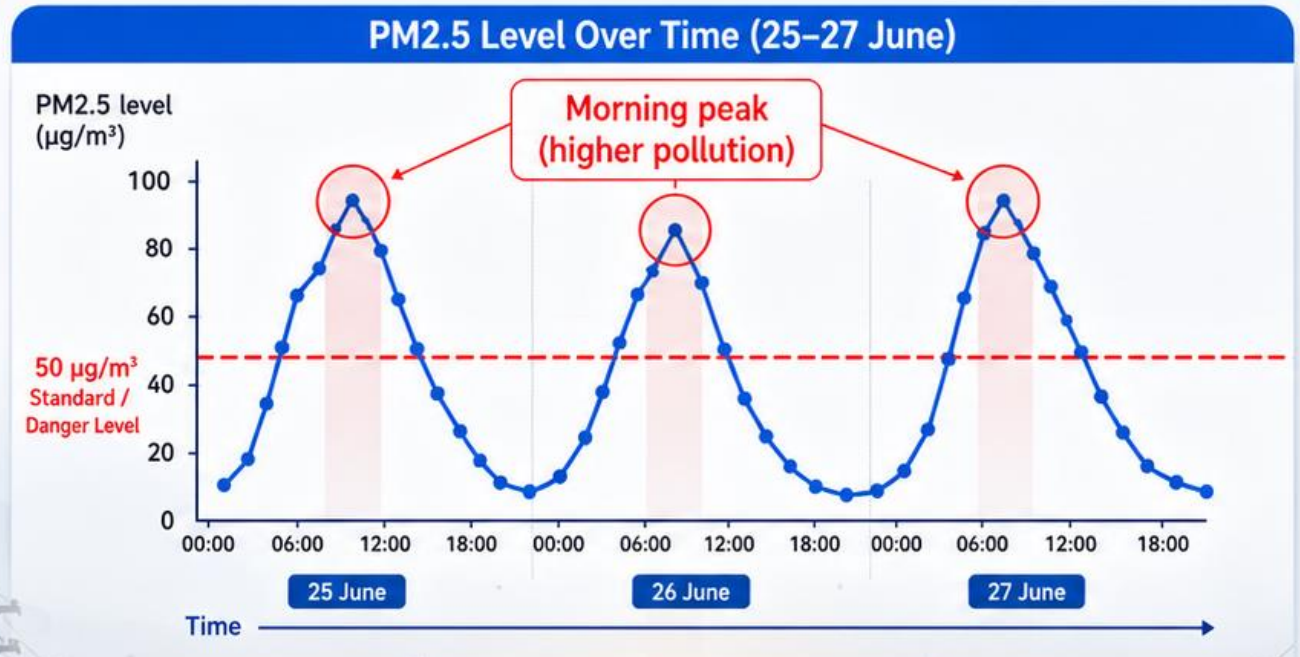
Graph Detective Questions

1. When is PM2.5 highest?
2. When is it lowest?
3. What pattern repeats?

Main question:
What pattern can you see?

What Pattern Can We Find?

Data becomes useful when we notice patterns.



PM2.5 increases at certain times.

Morning traffic may affect pollution.

Some times of day are safer than others.

Discussion question:

Why might PM2.5 increase in the morning?

From Pattern to Decision

Data Science helps people choose better actions.

1. Sensor Data
2. Find Pattern
3. Understand Risk
4. Make Decision
5. Take Action

Examples of Actions

- wear a mask
- avoid outdoor activity
- change travel time
- warn students
- plan traffic solutions

PM2.5 Data Detective

Pair Activity

Work with a partner. Answer:

1. What problem are we studying?
2. What data was collected?
3. What pattern can you see?
4. What decision could people make?
5. What extra data would make the answer better?

Output:
Write short answers in your notebook.

Share Your Findings

What did your data team discover?

Group Sharing Prompts

Our pattern was...

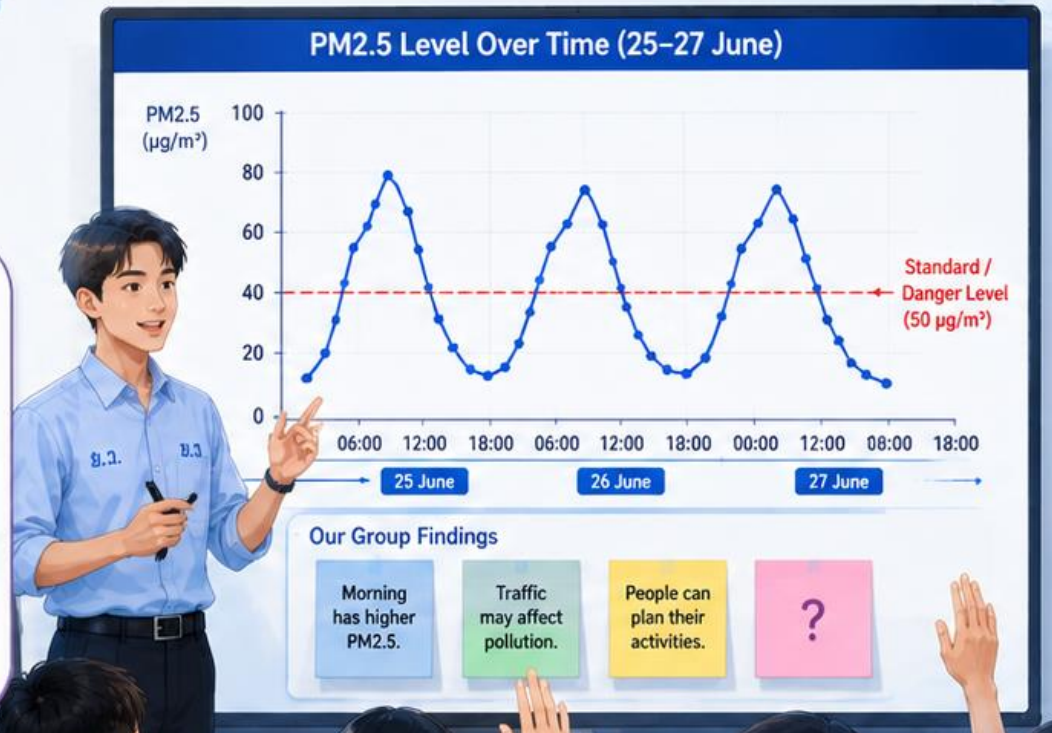
We think this happens because...

One decision people could make is...

One extra data type we need is...

Teacher Question

Is this conclusion based on evidence?



How Could We Improve the Solution?

Better data can lead to better decisions.

Extra data that could help:

- traffic amount
- weather
- wind direction
- temperature
- location
- time of day
- number of sensors

What data would make our conclusion stronger?

What Did We Learn Today?

Data Science helps us understand real problems.

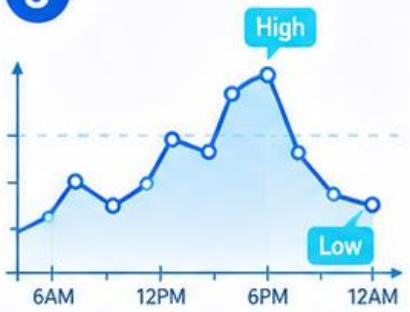
1

Data Science uses data to find patterns.

2

IoT sensors collect PM2.5 data.

3



Graphs show high and low pollution times.

4

Patterns support better decisions.

5

Better data creates stronger conclusions.

Exit Ticket

Before you leave

1. What is Data Science?
2. What pattern did we find in the PM2.5 data?
3. How can this data help people make decisions?

Use evidence from the graph.