

Plotting a Scatter Chart

Name		Date	
Class / Sec		Roll no.	

1 · AIM

To write a Python program that plots a **scatter chart** using the **Matplotlib** library, to examine whether a **relationship** exists between two variables.

2 · SOFTWARE AND LIBRARIES USED

TOOL	WHAT IT IS FOR
Python 3	The programming language the program is written in
Jupyter Notebook	The environment used to write and run the code (IDLE or Colab also acceptable)
Matplotlib	The plotting library. The scatter() function plots points without joining them

3 · THEORY

A **scatter chart** plots one point for each pair of values, **without joining them**. It is used to see whether two variables are **related** — whether one tends to rise as the other rises.

CORRELATION	WHAT THE POINTS DO	EXAMPLE
Positive	Rise together — points trend upward	Hours studied and marks scored
Negative	One rises as the other falls	Hours on the phone and marks scored
No correlation	Points scattered with no pattern	Shoe size and marks scored

Why the points are not joined: In this data **2 hours appears twice** — with 42 marks and 48 marks. Two different results for the same input. A line cannot pass through both, which is exactly why a scatter chart is the right choice here.

4 · PROGRAM

```
import matplotlib.pyplot as plt

# Hours studied per week and marks scored, for 10 students
hours = [1, 2, 2, 3, 4, 4, 5, 6, 6, 7]
marks = [35, 42, 48, 51, 60, 66, 69, 78, 82, 88]

# Plot the scatter chart — note scatter(), not plot()
plt.scatter(hours, marks, color='green', marker='o')

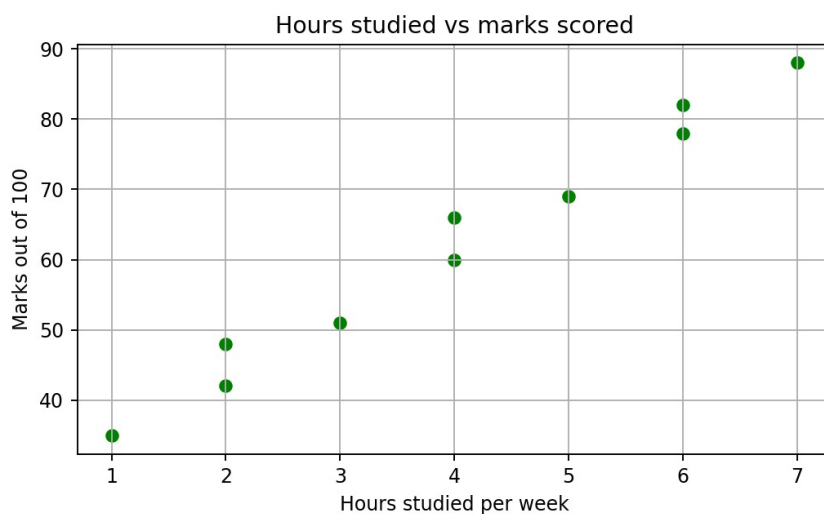
# Label the chart — an unlabelled chart loses marks
plt.title("Hours studied vs marks scored")
plt.xlabel("Hours studied per week")
plt.ylabel("Marks out of 100")
plt.grid(True)

plt.show()
```

scatter() not plot(): Using **plt.plot()** here would join the points into a line and imply a sequence that does not exist. The function name is the difference between a correct answer and a wrong one.

5 • EXPECTED OUTPUT

Running the program opens a chart window that looks like this.



[PASTE SCREENSHOT HERE]

Paste a screenshot of YOUR scatter chart, with the code visible alongside it if possible

6 • READING THE CHART

Before writing your observation, read what the chart actually shows.

QUESTION	WHAT THE CHART SHOWS
Is there a pattern?	Yes — a clear positive correlation . Points rise from lower left to upper right
Range of the data?	From 1 hour / 35 marks up to 7 hours / 88 marks
Do points share an x-value?	Yes — 2, 4 and 6 hours each appear twice , with different marks each time
Why scatter, not a line?	Because the same input gives different outputs. There is no single sequence to join

7 • OBSERVATION AND RESULT

Two or three sentences in your own words. A good observation names the **type of correlation** and is careful about what it claims.

Example observation: The scatter chart shows a strong positive correlation between hours studied and marks scored — students who studied longer generally scored higher. The points do not lie on a perfect line, because two students studying the same hours still scored differently. A scatter chart is correct here because we are examining a relationship, not a sequence.

One caution worth a mark: Correlation is **not** causation. This chart shows the two rise together — it does not prove that studying longer **caused** the higher marks. Saying so in your observation shows real understanding.

8 · VIVA QUESTIONS FOR THIS PROGRAM

LIKELY QUESTION	SHAPE OF A STRONG ANSWER
What is the difference between <code>plot()</code> and <code>scatter()</code> ?	plot() joins points with a line; scatter() leaves them as separate points
What kind of correlation does your chart show?	Positive — as hours studied increase, marks generally increase too
Does this prove studying causes higher marks?	No. Correlation is not causation. The chart shows they rise together, nothing more
Why do two points share the same x-value?	Two students studied the same hours but scored differently — which is normal, and why a line would be wrong
What would no correlation look like?	Points scattered randomly with no upward or downward pattern

9 · HOW THIS ENTRY IS MARKED

FULL MARKS

Chart **fully labelled** — title, x-label, y-label · uses **scatter()** correctly · a screenshot of **your own** chart · names the **type of correlation** · notes that correlation is not causation

LOSES MARKS

Uses **plot()** and joins the points · unlabelled axes · no screenshot, or someone else's · claims the chart **proves** studying causes higher marks · cannot say why scatter rather than line

Before you submit: Collect **your own** data — ask ten classmates their study hours and last test marks. Real data is more defensible in the viva, and your chart will look different from everyone else's.