Visualising Data

A chart beats a table because the eye sees patterns at once. Bar charts, line charts and histograms, which chart answers which question, and one matplotlib example you can read step by step.

Introduction

Here is a table of average marks in five subjects: Maths 72, English 81, Science 68, Urdu 85, Islamiat 79. Read it again. Which subject is lowest? You probably had to scan the numbers twice. Now imagine the same five numbers as bars side by side. The shortest bar jumps out in one second, with no reading at all. That is the whole power of visualising data: **the eye finds patterns faster than the brain reads numbers.**

A chart is not decoration. It is an argument made visible. A good chart answers one question clearly. A bad chart — too many colours, no labels, a cut-off axis — hides the answer or, worse, lies about it. In this lesson you will learn the three charts a beginner needs, which question each one answers, and how to make and read one with matplotlib, the standard Python chart library.

Before you make a chart, learn to read one like a sceptic. A chart can mislead even when its numbers are true, and the tricks are always in the same places: an axis that does not start at zero, a title that claims more than the data shows, or a missing label that leaves you guessing what the bars mean. The habit to build is this: read the title, read both axis labels, and check where the axis starts — before you believe anything the chart seems to say. It takes five seconds, and it protects you from most chart lies, including ones in newspapers and advertisements.

(Roman Urdu: Table parhne mein waqt lagta hai, chart dekhte hi samajh aa jata hai — sab se chhoti bar, sab se bara farq, sab kuch ek nazar mein. Chart sajawat nahi, jawab ko nazar aane wali shakal dena hai.)

Key points

- **A chart beats a table when the question is about comparison or change.** Tables win when the exact number matters.
- **Bar chart: compare categories.** Subjects, cities, teams — one bar per group, taller bar means bigger value.
- **Line chart: show change over time.** Marks across terms, sales across months — the line's slope tells the story.
- **Histogram: show how values spread out.** How many students scored 60–70, 70–80, 80–90? Bars here count values in ranges, not categories.
- **One chart, one question.** A chart that tries to answer three questions answers none of them well.
- **Label everything:** a title, names on both axes, and numbers where they matter. An unlabelled chart is a riddle, not an answer.

- **Start a bar axis at zero.** Cutting the axis off makes small differences look huge, and that misleads the reader.
- **Matplotlib's basic pattern is always the same:** make the chart, add labels, show it — `plt.plot()` or `plt.bar()`, then `plt.xlabel()`, `plt.ylabel()`, `plt.title()`, then `plt.show()`.
- **Scatter plot: compare two numbers at once.** Each dot is one student, with study hours on one axis and marks on the other — if dots climb together, more study goes with higher marks. Use it when both axes are numbers.
- **Use few colours, and let colour mean something.** One colour for normal bars and a different colour for the one bar you want eyes on is powerful; a different colour per bar just for decoration is noise.

Examples

Example 1 — A bar chart of average marks by subject

This example needs `pip install matplotlib` first:

```
import matplotlib.pyplot as plt

subjects = ["Maths", "English", "Science", "Urdu", "Islamiat"]
averages = [72, 81, 68, 85, 79]

plt.bar(subjects, averages)
plt.xlabel("Subject")
plt.ylabel("Average marks")
plt.title("Average marks by subject")
plt.show()
```

What this means: five bars, one per subject, each as tall as its average. Reading it takes one second: Urdu's bar is tallest, Science's is shortest. The labels matter — without "Average marks" on the side, a reader cannot tell whether taller means better or just bigger. Try changing one average and re-running: the chart updates, and the story changes with it.

Example 2 — A line chart of marks over time

This example needs `pip install matplotlib` first:

```
import matplotlib.pyplot as plt

terms = ["Term 1", "Term 2", "Term 3", "Term 4"]
maths_average = [64, 68, 71, 78]

plt.plot(terms, maths_average, marker="o")
plt.xlabel("Term")
plt.ylabel("Average maths marks")
plt.title("Maths average is rising term by term")
plt.show()
```

What this means: the line climbs from left to right, so the story is improvement over time — something a bar chart would show less clearly, because bars stress separate groups, not a journey. The dots (`marker="o"`) mark the real data points, so the reader can see where the numbers actually sit and where the line is just joining them.

Example 3 — Which chart for which question?

```
Question: Which subject scored highest?
Answer:  Bar chart — compare the bars, pick the tallest.

Question: Are marks rising or falling over the year?
Answer:  Line chart — follow the slope over time.

Question: Do most students cluster near 70, or spread out?
Answer:  Histogram — bars count how many students fall
         in each marks range (60-70, 70-80, 80-90).
         A tall middle bar means most students are average;
         a flat spread means marks are all over the place.

Question: What exactly was Ali's mark in Maths?
Answer:  A table — charts are bad at exact numbers.
```

What this means: pick the chart from the question, not from habit. Ask first: "Am I comparing groups, tracking change, or looking at a spread?" The answer to that question is the chart type.

Example 4 — A histogram of how marks spread out

This example needs `pip install matplotlib` first:

```
import matplotlib.pyplot as plt

marks = [55, 62, 68, 71, 74, 78, 82, 85, 88, 92]

plt.hist(marks, bins=[50, 60, 70, 80, 90, 100])
plt.xlabel("Marks range")
plt.ylabel("Number of students")
plt.title("How 10 students' marks spread out")
plt.show()
```

What this means: each bar counts students in a 10-mark range — 1 student in 50–60, 2 in 60–70, 3 in 70–80, 3 in 80–90, 1 in 90–100. The shape is a gentle hill in the middle: most students sit in the middle ranges, few at the ends. That hill shape is exactly what lesson DS.6 calls a distribution, and a histogram is how you see it.

Example 5 — Comparing two classes, side by side

This example needs `pip install matplotlib` first:

```
import matplotlib.pyplot as plt

subjects = ["Maths", "English"]
class_a = [72, 81]
class_b = [68, 85]

x = range(len(subjects))
plt.bar([i - 0.2 for i in x], class_a, width=0.4, label="Class A")
plt.bar([i + 0.2 for i in x], class_b, width=0.4, label="Class B")
plt.xticks(list(x), subjects)
plt.ylabel("Average marks")
plt.title("Class A vs Class B, by subject")
plt.legend()
plt.show()
```

What this means: two bars per subject, one per class, standing side by side so the eye compares within each subject. Class A leads in Maths (72 vs 68); Class B leads in English (85 vs 81). The legend is essential here — without it, nobody knows which bar belongs to which class. One chart, one question ("which class is stronger in each subject?"), answered at a glance.

Example 6 — Real-world case: the chart on a shopkeeper's wall

A shopkeeper in a small market puts one chart on the wall behind the counter: weekly sales as bars, one bar per week, for the last two months. He does not need exact numbers — he needs to see, while serving a customer, whether this week is up or down. When a bar drops two weeks in a row, he orders less stock; when bars climb before a festival, he orders more. That is a chart doing its real job: not a report for a boss, but a quick decision aid for the person standing in front of it. The same idea drives dashboards in big companies — sales, visitors, deliveries — all as charts that busy people can read in seconds.

(Roman Urdu: Dukaan ke deewar par laga chart bhi wohi kaam karta hai jo bari company ka dashboard — ek nazar mein faisla.)

Quick recap

- Charts let the eye find patterns faster than reading numbers.
- Bar chart: compare categories. Line chart: change over time. Histogram: spread of values.
- One chart answers one question — keep it simple.
- Always label the title and both axes.
- Start bar axes at zero, or small differences will look huge.

Common mistakes

1. **Making a chart with no labels** — the reader sees shapes but cannot say what they mean. Fix: add a title and both axis labels before you show anyone the chart.
2. **Using a line chart for categories with no order** — joining "Maths, English, Urdu" with a line suggests a journey that does not exist. Fix: use bars for categories; save lines for time or ordered values.
3. **Cutting off the bar axis to exaggerate a gap** — bars from 70 to 80 look twice as tall as bars from 70 to 75 if the axis starts at 65. Fix: start bar charts at zero, or say clearly in the title that the axis is cut.
4. **Making a pie chart with ten slices** — tiny slices are unreadable, and nobody can compare angles accurately anyway. Fix: if there are more than five categories, use a bar chart, or group the smallest categories into one "Other" slice.
5. **Adding 3D effects, shadows or patterned fills** — they look fancy and make values harder to read, because a tilted 3D bar's height no longer matches its value. Fix: flat bars, plain colours, nothing that distorts the height or length.

Practice

1. Take the five subject averages from Example 1 and draw them as bars on paper. Which bar is tallest, and which is shortest?
2. Run Example 1 in Python, then change Science from 68 to 90 and re-run. How does the story change?
3. Write down one question from your own life for each chart type: one comparison, one change over time, one spread.
4. Find any chart online or in a textbook and check it: does it have a title and both axis labels? Does its bar axis start at zero?

1. Run the histogram from Example 4, then add three more marks of your choice and re-run it. Which bar grows, and does the overall shape stay a hill?
2. Draw two classes' averages as side-by-side bars on paper (use Example 5's numbers), then write a two-sentence story the chart tells. Swap one number and rewrite the story — how did it change?

3. Find a chart in a newspaper, a textbook or a social media post. Check it with the sceptic's habit from the Introduction: title, both axis labels, axis start. Write down anything it fails.

Practice quiz

1. Why is a chart often better than a table?

- Charts are prettier
- A pattern jumps out at a glance, while a table hides it in numbers
- Tables are always wrong
- Charts need no data

Answer: A pattern jumps out at a glance, while a table hides it in numbers

2. Which chart is best for comparing categories?

- A line chart
- A bar chart
- A histogram
- A pie chart with 50 slices

Answer: A bar chart

3. Which chart is best for change over time?

- A bar chart
- A line chart
- A scatter plot of unrelated values
- A table

Answer: A line chart

Watch the video lesson: https://youtu.be/SiDytq0Z60w

Read online: https://scopetreasure.com/notes/ds-data-visualisation · More free tools and notes at https://scopetreasure.com