PREMIUM EDITION

B.PHARMA · IST YEAR · 2026-27

Data Visualization with Matplotlib

UNIT-V · BASICS OF PYTHON IN PHARMACEUTICAL SCIENCES

Conc-time curves, dissolution and ADR plots – every figure built live in matplotlib, read like an examiner.

Line Plots | Histograms | Scatter + Box | Dissolution | Smart Revision | Exam Points

Inside Unit V

5.1 Figure anatomy + first plot

5.2 Line: conc-time oral vs IV

5.3 Histogram + scatter

5.4 Box plots: ADR rates

5.5 Dissolution profiles

Levels: Remember to Analyze

Pills and Python · PCI-aligned · Remember to Analyze

§ UNIT OVERVIEW

S Unit-5 at a Glance

Roadmap

5.1 Anatomy of a figure Figure, Axes, labels, legend, ticks – then your first `plot()`

5.2 Line plots conc-time curves oral vs IV; read Cmax / Tmax off the curves

5.3 Histogram + scatter dissolution-% spread; dose-vs-response trend slope

5.4 Box plots ADR rates across 3 drugs via median + IQR

5.5 Dissolution profiles multi-batch release curves; which batch looks similar?

Why plots win marks

Examiners forgive a missing import; they never forgive an ***unlabeled axis***. Every plot in this unit follows one habit: ***label legend title interpret***.

All figures in these notes are real matplotlib output from seeded synthetic teaching data – read values off them exactly as you would in the lab record.

Professor's Note

A plot without interpretation scores half. After every figure write one line: what rose, where it peaked, and what that means pharmaceutically.

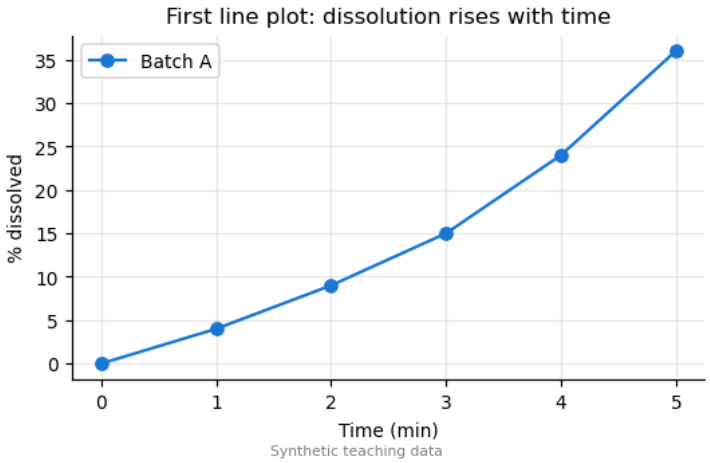
PHARMACY • DATA • AI • HEALTHCARE

§ S.1 ANATOMY OF A FIGURE

1 Figure, Axes, Labels

Definition

Figure is the whole window (`plt.subplots()` makes it). *Axes* is the box where data lives. Labels name the axes with units, legend names the curves, title states the conclusion.



Parts checklist

1 *Title* – one-line result
2 *xlabel / ylabel* – with units
3 *Curves + markers* – one style each
4 *Legend* – never cover data
5 *Grid* – light, reading aid only

first_plot.py

```
# first line plot - dissolution rises with time
import matplotlib.pyplot as plt
x = [0, 1, 2, 3, 4, 5] # time, min
y = [0, 4, 9, 15, 24, 36] # % dissolved
plt.plot(x, y, marker='o', label='Batch A')
plt.xlabel('Time (min)'); plt.ylabel('% dissolved')
plt.title('Dissolution rises with time'); plt.legend()
plt.savefig('first.png', dpi=110)
```

Output

Figure saved: one rising curve with markers, axes labelled, legend top-left. *Reading:* release climbs from 0 to 36% over 5 min – monotonic rise, no plateau yet.

§ 5.2 LINE PLOTS - CODE

2 Conc-Time: Oral vs IV

Pharma story

IV bolus starts at its *peak* and falls (elimination only). Oral starts at zero, *rises to Cmax at Tmax* (absorption beats elimination), then falls. Two curves on one Axes make the comparison instant.

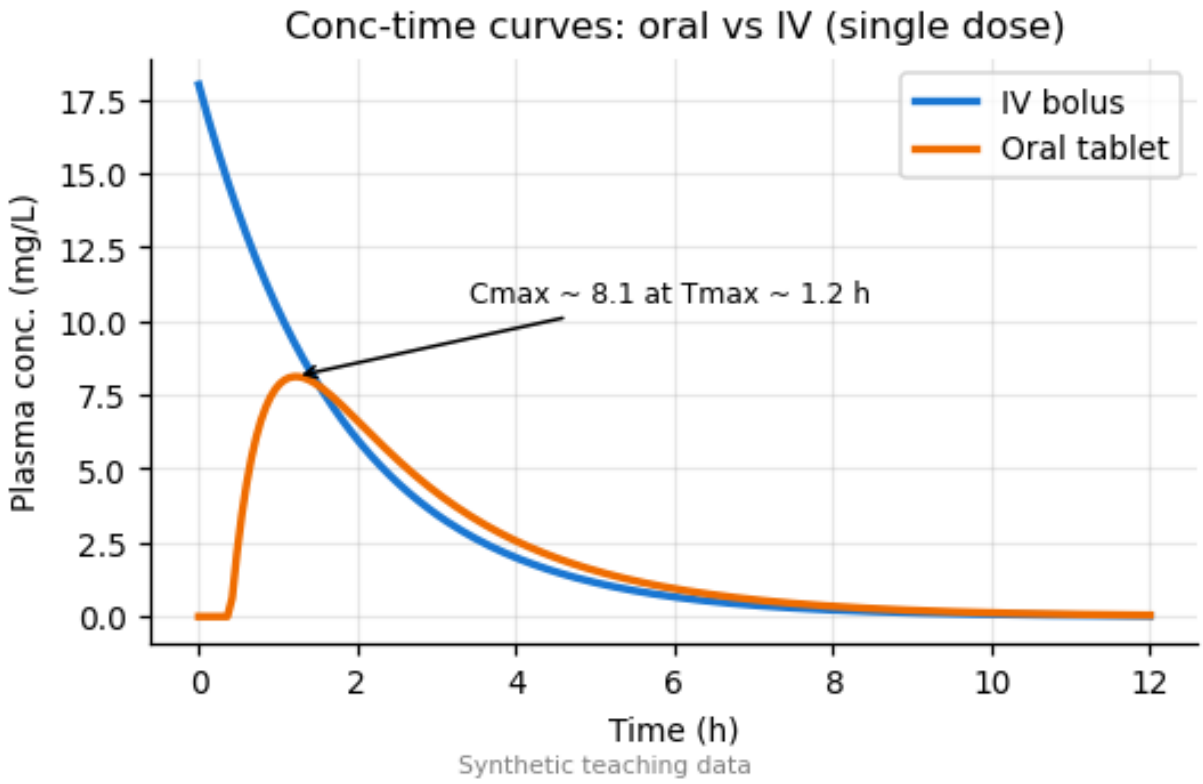
conc_time.py

```
# conc-time curves: oral vs IV (synthetic teaching data)
import numpy as np; import matplotlib.pyplot as plt
t = np.linspace(0, 12, 200) # hours
iv = 18 * np.exp(-0.55 * t)
to = t - 0.4; oral = np.where(to>0, 15.5*(np.exp(-0.5*to)-np.exp(-2.4*to)), 0)
fig, ax = plt.subplots(figsize=(5, 3.2), dpi=110)
ax.plot(t, iv, label='IV bolus'); ax.plot(t, oral, label='Oral tablet')
ax.set_xlabel('Time (h)'); ax.set_ylabel('Plasma conc. (mg/L)')
ax.legend(); fig.savefig('conc.png', bbox_inches='tight')
```

Why two curves, one Axes?

Same x (time) and same y (mg/L) means the eye compares *shapes* directly: IV peak at t = 0 vs oral delayed peak. Separate figures would hide that contrast.

§ 5.2 LINE PLOTS – FIGURE



Reading the curves (exam lines)

IV: C0 ~ 18 mg/L at t = 0, exponential fall – no absorption phase.

Oral: Cmax ~ 8.1 mg/L at Tmax ~ 1.2 h – then declines parallel to IV (same elimination).

One-line interpretation: oral lags IV, peaks lower and later, then eliminates similarly – exactly what first-pass + absorption delay predicts.

Output

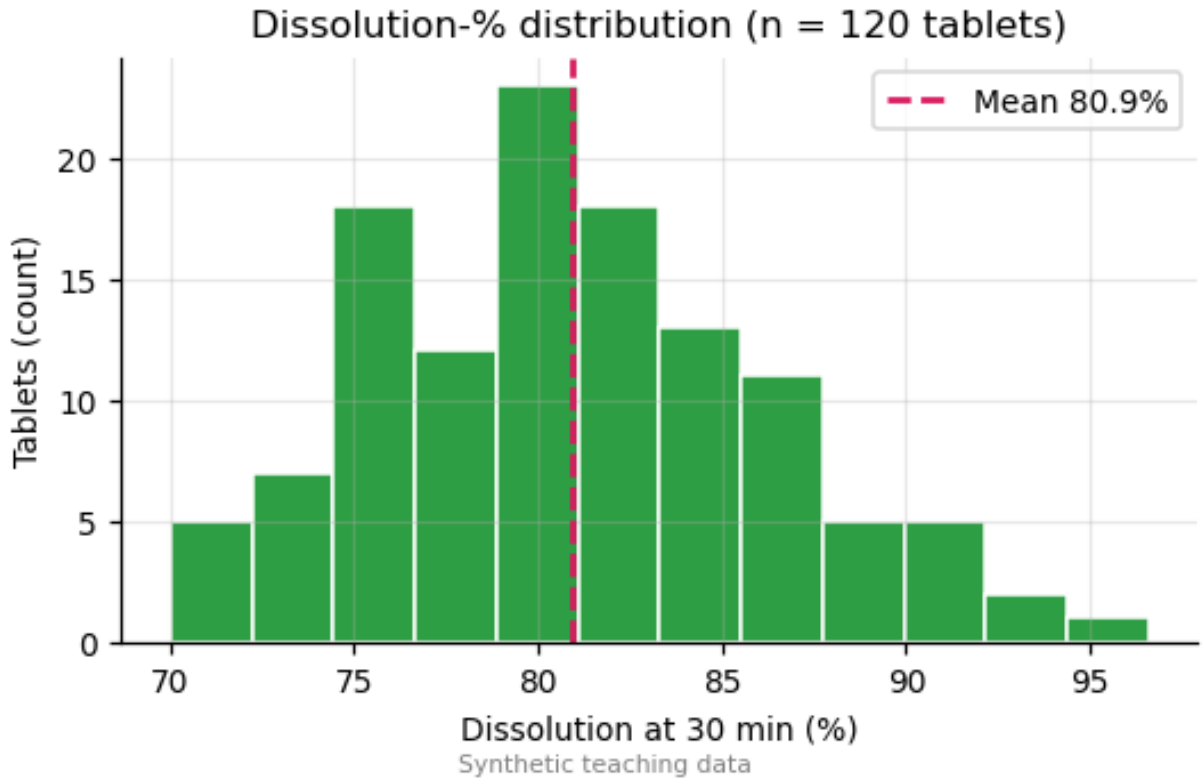
Synthetic teaching data. IV curve always starts highest; oral curve crosses it only on the way down – state both facts for full marks.

§ 5.3A HISTOGRAMS

3 Dissolution-% Distribution

histogram.py

```
# distribution of dissolution at 30 min, n=120 tablets
rng = np.random.default_rng(5)
d = np.clip(rng.normal(82, 6, 120), 60, 100)
ax.hist(d, bins=12, edgecolor='white')
ax.axvline(d.mean(), ls='--', label='Mean') # dashed pink line
ax.set_xlabel('Dissolution at 30 min (%)')
```



What the histogram says

Mean ~ 80.9%, bulk of tablets 75–90% – a *tight, near-symmetric mound*. Bars far left of 70% are suspect tablets worth retesting.

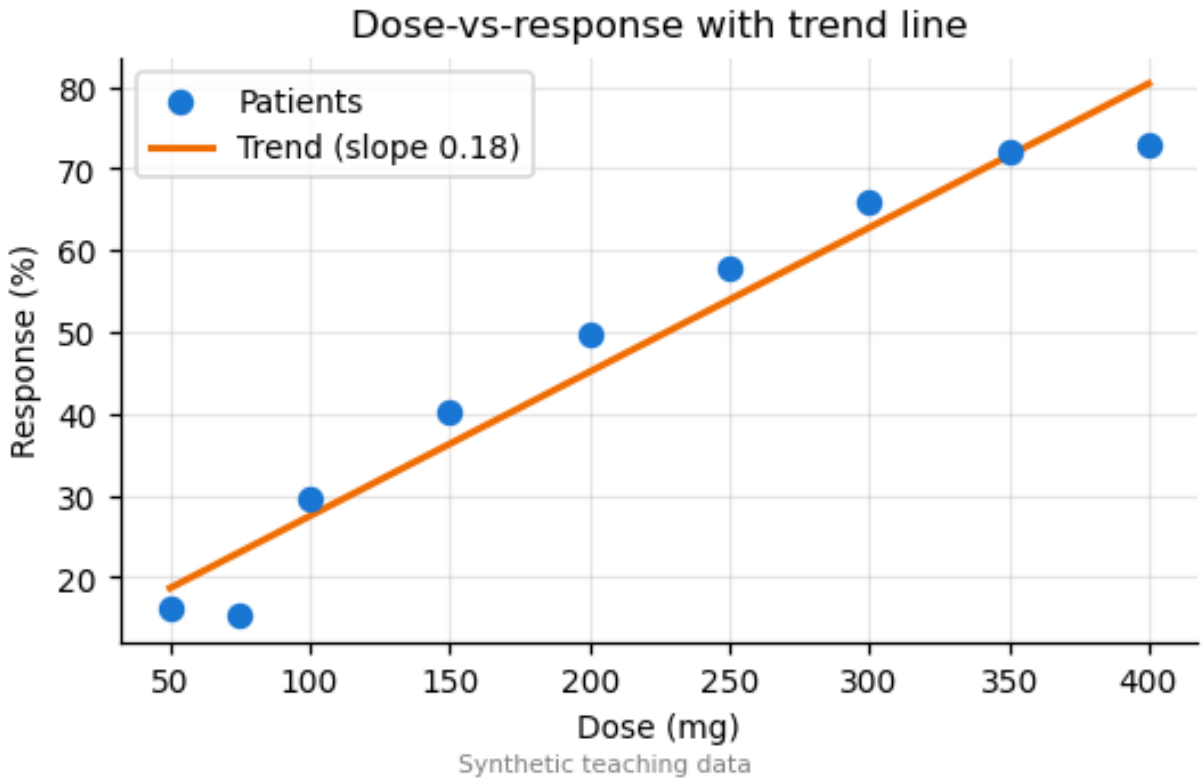
Exam line: histograms show spread, line plots show trend – if asked why not a line plot here, say the x-axis is value-bins, not time.

§ 5.3B SCATTER PLOTS

4 Dose-vs-Response + Trend

scatter.py

```
# dose-response scatter with fitted trend line
dose = np.array([50, 75, 100, 150, 200, 250, 300, 350, 400])
m, b = np.polyfit(dose, resp, 1) # straight-line fit
ax.scatter(dose, resp, s=42, label='Patients')
ax.plot(xx, m*xx+b, label='Trend') # orange line
ax.set_xlabel('Dose (mg)'); ax.set_ylabel('Response (%)')
```



Trend reading

Positive slope ~ 0.18 response-% per mg - response rises with dose but *flattens past ~300 mg* (diminishing returns).

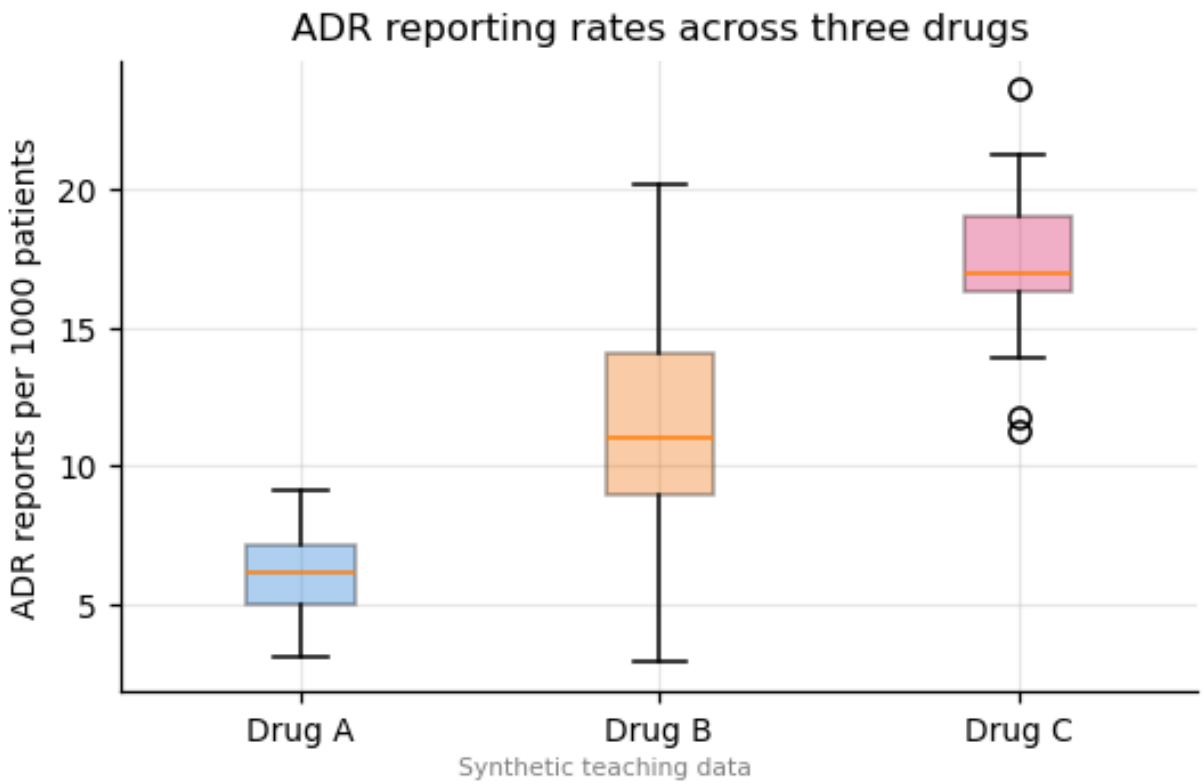
Exam line: a scatter proves association, not causation - always add that caveat before claiming efficacy.

§ 5.4 BOX PLOTS

5 ADR Rates Across 3 Drugs

boxplot.py

```
# ADR reports per 1000 patients, 3 drugs (synthetic teaching data)
ax.boxplot([A, B, C], labels=['Drug A', 'Drug B', 'Drug C'])
ax.set_ylabel('ADR reports per 1000 patients')
# box = IQR (Q1-Q3), line = median, whiskers = range, dots = outliers
```



Median / IQR reading

Drug A lowest median (~6) + narrow box – safest, most consistent. **Drug B** middle median (~11) but widest box – variable reporting. **Drug C** highest median (~17) – needs safety review.

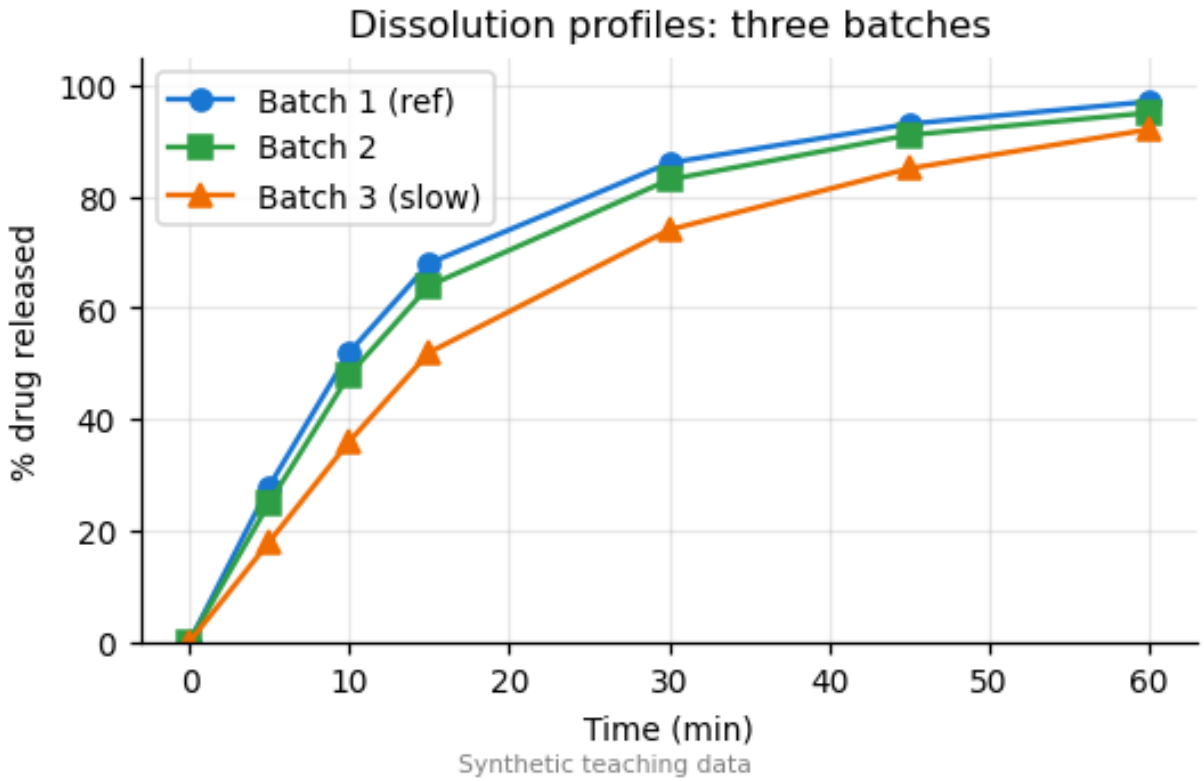
Exam definitions: median = middle line, IQR = box height, outliers = dots. Wide box = inconsistent; high median = frequent ADRs.

§ 5.5 DISSOLUTION PROFILES

6 Multi-Batch Release Curves

dissolution.py

```
# % released vs time for three batches (synthetic teaching data)
tm = np.array([0, 5, 10, 15, 30, 45, 60])
ax.plot(tm, b1, marker='o', label='Batch 1 (ref)')
ax.plot(tm, b2, marker='s', label='Batch 2')
ax.plot(tm, b3, marker='..', label='Batch 3 (slow)')
ax.set_ylim(0, 105); ax.legend()
```



Similarity reading (no jargon needed)

Batch 2 *hugs* Batch 1 at every time point – similar release. Batch 3 runs *10–15 points lower* throughout – slower release, likely different granulation or coating.

Exam line: compare whole curves, not one time point – curves that stay close at all points are similar; one that sags everywhere is not.

§ 5.6 LABELS, TITLES, LEGENDS

7 The 4-Line Finish

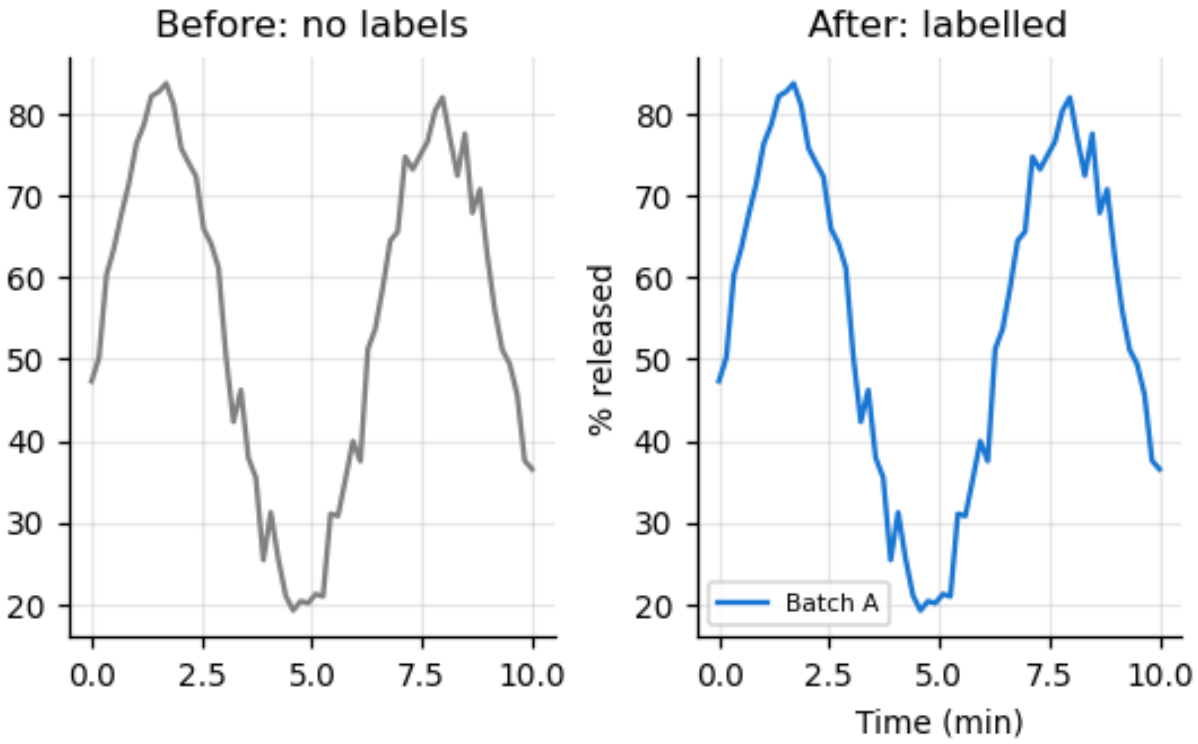
Master checklist

xlabel / ylabel – quantity + unit, e.g. `Time (min)`. No unit = half mark gone.

title – states the finding, not the method: 'Oral peaks later than IV' beats 'Conc vs time'.

legend – names every curve; move it off the data with `loc='best'`.

order – label everything *before* `savefig()`; anything after saving never appears.



Before – After

Left: bare gray curve, no labels – examiner guesses the axes. Right: same data with labels + legend – the figure explains itself. Four extra lines of code, full marks.

§ 5.7 INTERPRETATION MASTERCLASS

8 What Examiners Ask

Plot type	Examiner asks	Model answer stem
Line	Where is the peak? What falls faster?	Peaks at … then declines; A falls faster than B
Histogram	Where is the bulk? Any outliers?	Bulk at …; bars below … are suspect
Scatter	Trend? Strength?	Positive trend, slope ~…; association, not proof
Box	Which median highest? Which IQR widest?	C highest median; B widest IQR = most variable
Dissolution	Which batch matches ref?	Batch 2 hugs ref; Batch 3 sags throughout

Three-sentence template

Sentence 1: ***describe*** (rises/falls/peaks at …). Sentence 2: ***compare*** (higher/wider/closer than …). Sentence 3: ***conclude*** (so Batch 2 is similar / Drug C needs review). Practice it under every figure.

Output Numbers beat adjectives: quote Cmax, Tmax, mean, median, slope – then interpret them in one pharma sentence.

§ 5.8 ERRORS YOU WILL MEET

9 Fix It Fast

oops.py

```
# ERROR 1: x and y lengths differ
ax.plot([0,1,2], [5, 9]) # ValueError: x and y must have same first dimension
# FIX: check len(x) == len(y) before plotting
# ERROR 2: labels after savefig() never appear
fig.savefig('a.png'); ax.set_xlabel('Time') # too late!
# FIX: xlabel, legend, title FIRST, savefig() LAST
```

Mismatched lengths

ValueError: x and y must have same first dimension. Cause: lists of different lengths. Fix: `len(x) == len(y)` check, or slice to the shorter one.

Missing labels + show() order

Blank-looking figure? You called `savefig()` before `xlabel()` / `legend()`. Rule: plot - label - legend - title - save, in that order.

Legend covering data? Add `loc='best'` or `bbox_to_anchor` - never let the box hide the peak you are asked about.

Notebook shows nothing? In scripts use `savefig()`; `plt.show()` is for interactive windows and blocks scripts after it.

§ 5.9 PLOT CHOOSER

10 Which Plot When?

Question	Plot	Function
Trend over time (conc, release)	Line	`plt.plot / ax.plot`
Spread of values (dissolution %)	Histogram	`ax.hist`
Relation of two variables	Scatter + trend	`ax.scatter + polyfit`
Compare groups (ADR by drug)	Box plot	`ax.boxplot`
Several batches vs time	Multi-line	`ax.plot x N + legend`

Decision rule

One variable over *time* line. One variable's *spread* histogram. Two variables' *relation* scatter. Group *comparison* box plot.

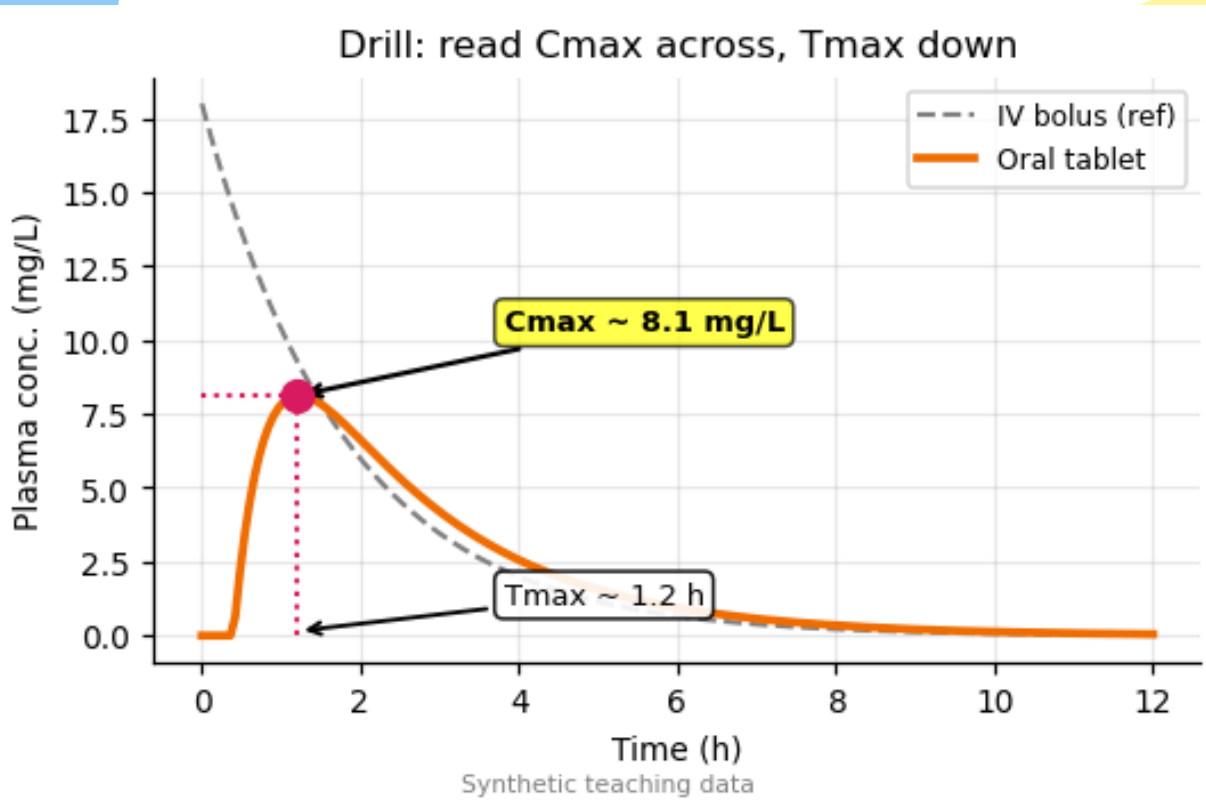
If the examiner's question has the word 'distribution', reach for a histogram; 'trend' or 'profile', a line; 'across groups', a box plot.

Memory hook

Lines move, bars pile, dots relate, boxes compare. Say it once before choosing and you will rarely pick wrong.

§ DEEP-DIVE CMAX READING DRILL

II Annotate the Peak



3-step drill (do it on every conc-time Q)

1 **Top the peak** – highest point of the oral curve: Cmax ~ 8.1 mg/L (dot + arrow).

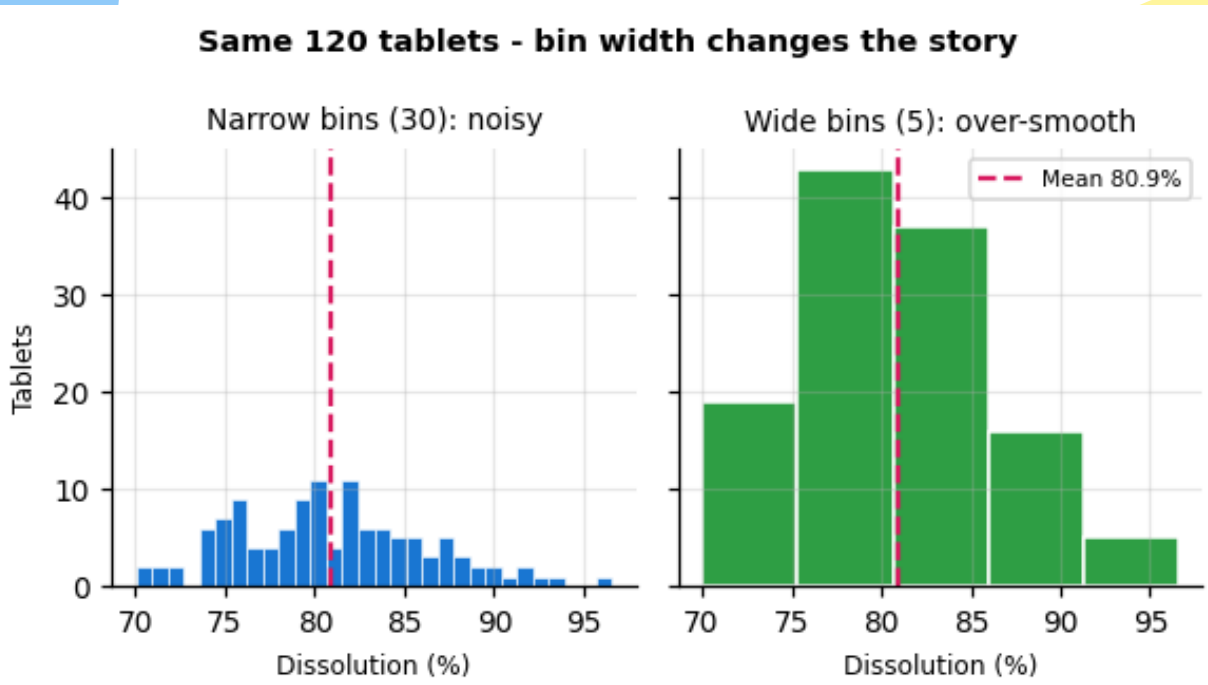
2 **Drop to x-axis** – dotted line down: Tmax ~ 1.2 h. Read ***across for Cmax, down for Tmax***.

3 **Trace across** – dotted line to y-axis confirms the value; quote both numbers, then add: oral lags IV, peaks lower and later .

Output Code behind the arrows: ***ax.annotate(..., arrowprops)*** marks the peak, dotted ***ax.plot*** guides the eye to each axis.

§ DEEP-DIVE BINS + RESCUE

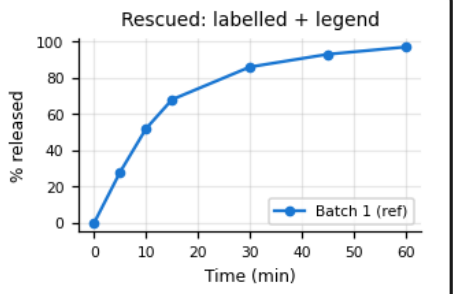
12 Bins Lie, Labels Rescue



Bin-width verdict

30 bins: spiky, gaps distract - noise poses as structure. *5 bins:* one blob - skew and suspects vanish. *12 bins (p.6):* just right for n = 120.

Exam line: too many bins = noise, too few = hides shape - defend your bin count in one sentence.



rescue.py

```
# rescue: 4 lines save the figure
ax.plot(tm, rel, marker='o', label='Batch 1')
ax.set_xlabel('Time (min)'); ax.set_ylabel('% released')
ax.legend(loc='lower right') # off the rise
fig.savefig('rescued.png') # LAST
```

Output Same curve, rescued: units on both axes + legend off the data. Copy these 4 lines onto every bare plot.

§ EXAM MUST-MEMORIZE

13 Exam Points

#	Point
1	Line plot: IV peaks at t = 0; oral peaks later at Cmax / Tmax – state both.
2	Histogram shows spread (bins), never use it for time trends.
3	Scatter slope = response per unit dose; association is not causation.
4	Box: line = median, box = IQR, dots = outliers; wide box = variable.
5	Dissolution: compare whole curves; hugging curves are similar.
6	Every figure needs xlabel + unit, ylabel + unit, legend, title.
7	Order: plot – label – legend – title – savefig() LAST.
8	ValueError on plot = mismatched x/y lengths; check len() first.

Mark-eaters

Unlabeled axes, legend hiding the peak, and one-time-point dissolution verdicts – each costs a full mark. The before-after demo on p.11 is the antidote.

§ VIVA ONE-MINUTE DRILL

14 Say It Aloud

Examiner fires – you answer

Q: Why does oral lag IV? A: Absorption + first-pass delay; Cmax lower, Tmax later, then same elimination slope.

Q: Histogram vs bar chart? A: Histogram bins one continuous variable (dissolution %); bar chart compares categories (drugs).

Q: Box plot without whiskers? A: Whiskers show range beyond IQR; without them you cannot spot outliers or true spread.

Q: Batch 3 sags – accept it? A: No – consistently lower release at all points signals a process deviation; investigate, do not average it away.

Q: Scatter slopes down – meaning? A: Negative association: higher dose, lower response – check toxicity or data error before concluding anything.

Delivery tip

Point at the figure while answering: ***trace the curve with your finger*** as you quote Cmax, mean, or median. Examiners credit answers anchored to visible values.

§ REVISION GRID + COS

15 Revision Grid

Concept	Function	One-line use
Line plot	`ax.plot`	Conc-time, release profiles
Histogram	`ax.hist`	Dissolution-% spread
Scatter	`ax.scatter`	Dose-response relation
Box plot	`ax.boxplot`	ADR medians + IQRs
Labels	`xlabel/legend`	Units, names, title - always
Save	`savefig`	Last call; dpi=110, tight

COS mapping

CO5: visualize and interpret pharmaceutical data with Python. Evidence in this unit: *conc-time oral vs IV*, *dissolution histogram + profiles*, *dose-response scatter*, *ADR box plots* - each with code, figure, and interpretation.

Output Revision mantra: *label, legend, title, interpret* - four lines that convert any correct plot into full marks.

§ SELF-CHECK SUMMARY

16 Check Yourself

5 questions (answers in brackets)

1. IV vs oral: which peaks at t = 0? [IV; oral peaks at Tmax]
2. Histogram answers spread or trend? [spread; trends need line plots]
3. Box-plot box spans what? [IQR, Q1–Q3; line inside = median]
4. Two dissolution curves hug – verdict? [similar release profiles]
5. savefig() goes before or after legend()? [after – last call of all]

Summary

- Line for trends, histogram for spread, scatter for relations, box for groups.
- Quote the number (Cmax, mean, slope, median) then say what it means.
- Label–legend–title–save: finish every figure the same way.

Takeaway

Matplotlib turns pharma tables into arguments: a labelled curve ***shows*** what a paragraph only claims. Master the four plot types here and CO5 is secured.

§ PINPOINTS MUST REMEMBER

17 Pinpoints – Must Remember

10 hooks

1 Question says trend/profile – line; distribution/spread – histogram; relation – scatter; across groups – box.

2 Every figure: xlabel + unit, ylabel + unit, legend, finding-title. No unit = half mark gone.

3 Cmax = peak height, Tmax = peak time; read across then down (drill p.12).

4 Histogram words: mound, symmetric/skewed, bulk at ..., suspect bars below ...

5 Scatter: slope sign = direction; flattening = diminishing returns; association *is not* causation.

6 Box: line = median, box = IQR, whiskers = range, dots = outliers; wide box = variable.

7 Dissolution: compare whole curves; hugging = similar; sagging everywhere = deviation.

8 Order: plot – label – legend – title – *savefig() LAST*.

9 Bins: too many = noise, too few = hides shape; n = 120 likes ~12 bins.

10 Interpret in 3 sentences: describe – compare – conclude, with numbers quoted.

3 exam predictions

P1. Conc-time: label Cmax/Tmax on oral vs IV + one pharma sentence (absorption delay).

P2. Box plot: highest median? widest IQR? verdict on which drug needs review.

P3. Dissolution: which batch matches ref and why – whole-curve argument, not one point.

Mnemonic – LABEL-PLOT

Label axes, **A**nnotate peaks, **B**ins justified, **E**xplain numbers, **L**egend off data – **PLOT** like it matters.