

Lessons learned* from incorporating generative AI in a new data analytics course in Geography

Dr Becky Dell & Dr Matt Osman

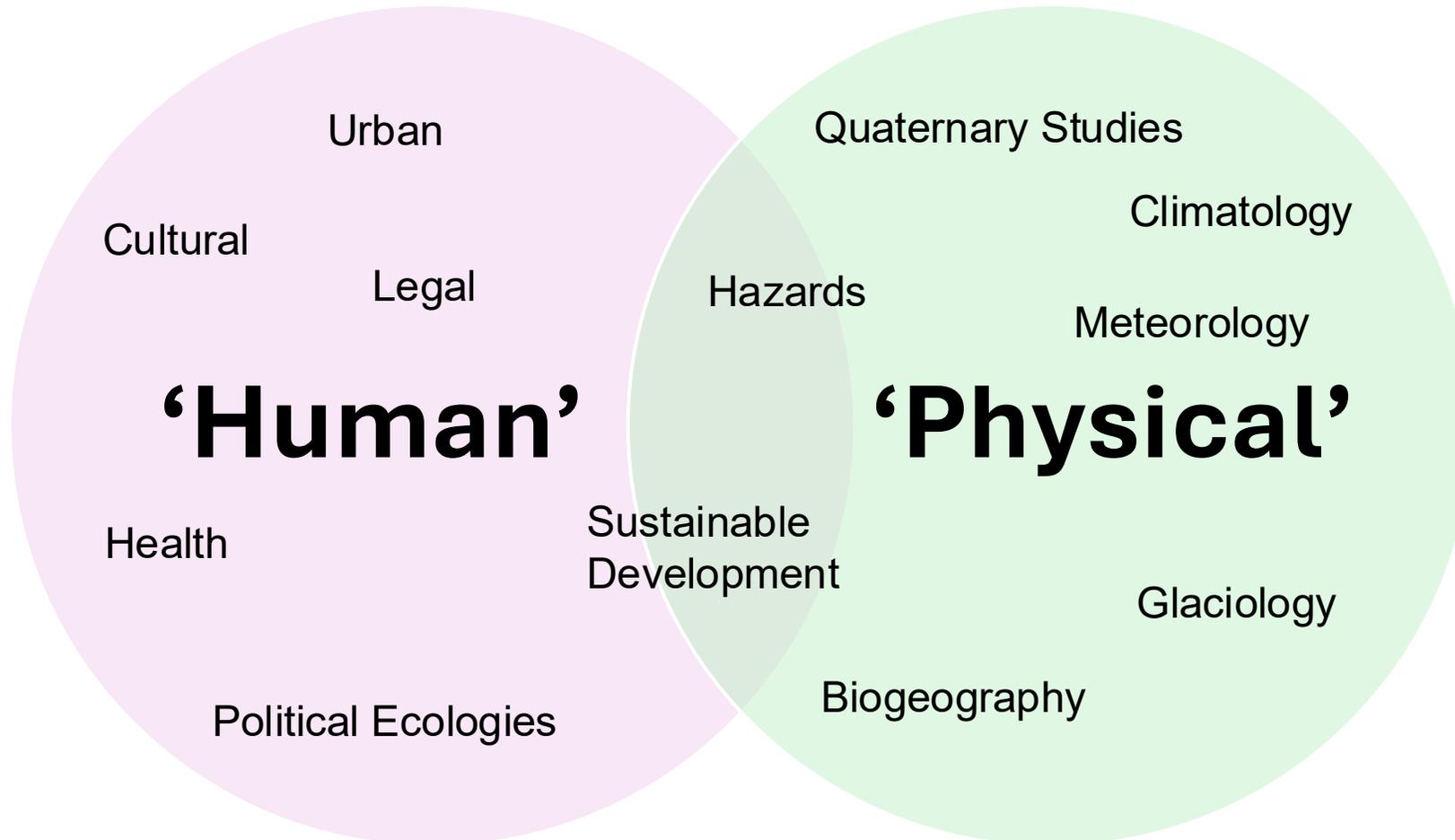
rld46@cam.ac.uk

mo549@cam.ac.uk

Sept 2025

***and uncertainties remaining!**

Geography, the course, & skills



Geography, the course, & skills

Data analytics in climate and environmental science

[Lecture 0]: *Environmental science in the big data era*

 [Lecture 1]: *Intro to scientific coding — PSet1*

  [Lecture 2]: *Working with & visualising environmental data — PSet2*

 [Lecture 3]: *Modelled climate data & spatiotemporal analysis — PSet3*

 [Lecture 4]: *Uncertainty quantification & hypothesis testing — PSet 4*

  [Lecture 5]: *Regression and prediction — PSet5*

 [Lecture 6]: *High dimensionality and data sparsity — PSet6*

 [Lecture 7]: *Spectral analysis of environmental data — PSet7*

 [Lecture 8]: *Assessments of hazard and risk — PSet8*

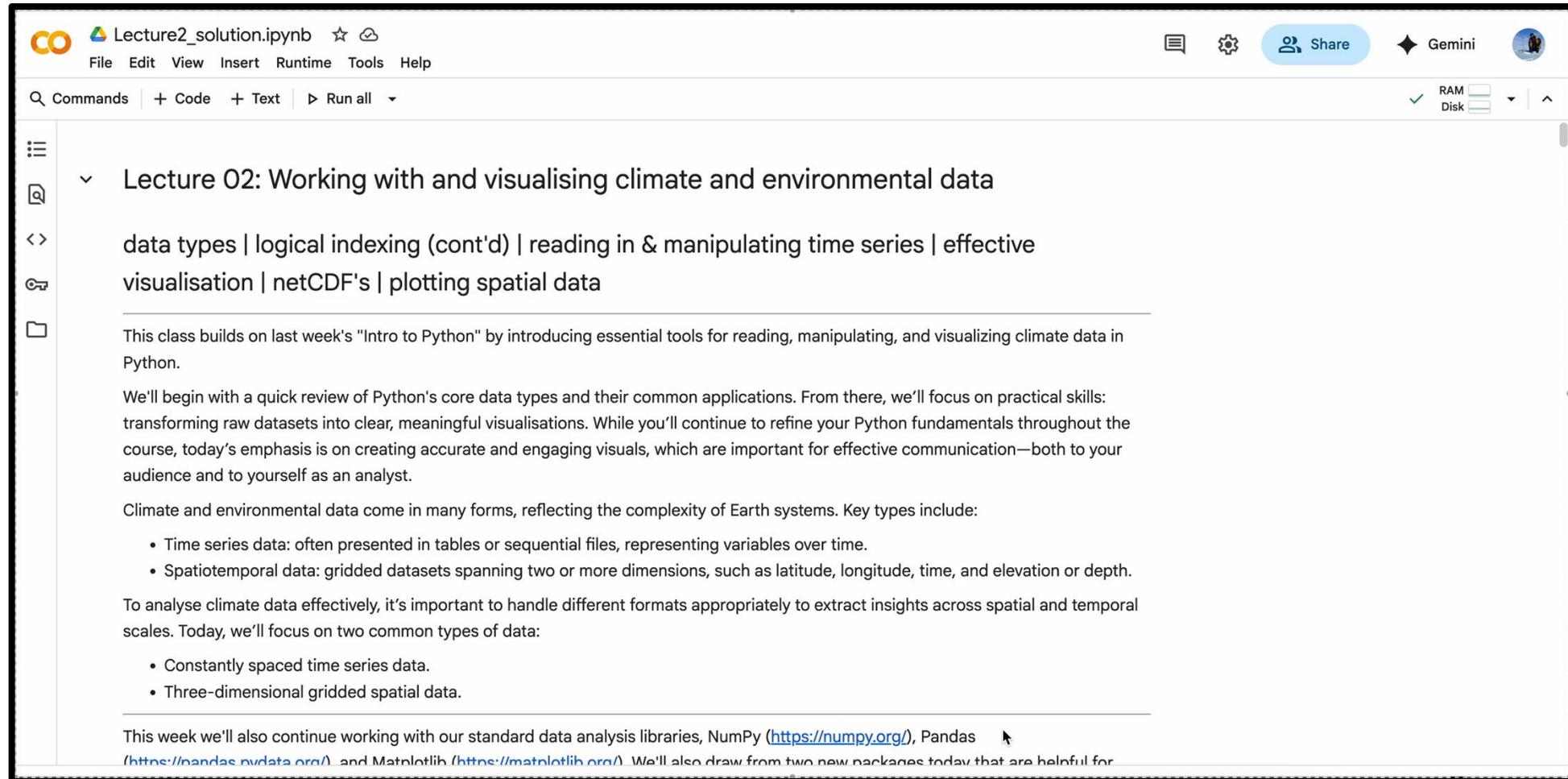
 [Lecture 9]: *Intuitive intro to Bayesian modelling — PSet9*

 [Lecture 10]: *Supervised and unsupervised learning — PSet10*

[Lecture 11]: *Communicating science effectively, and FAIR principles*



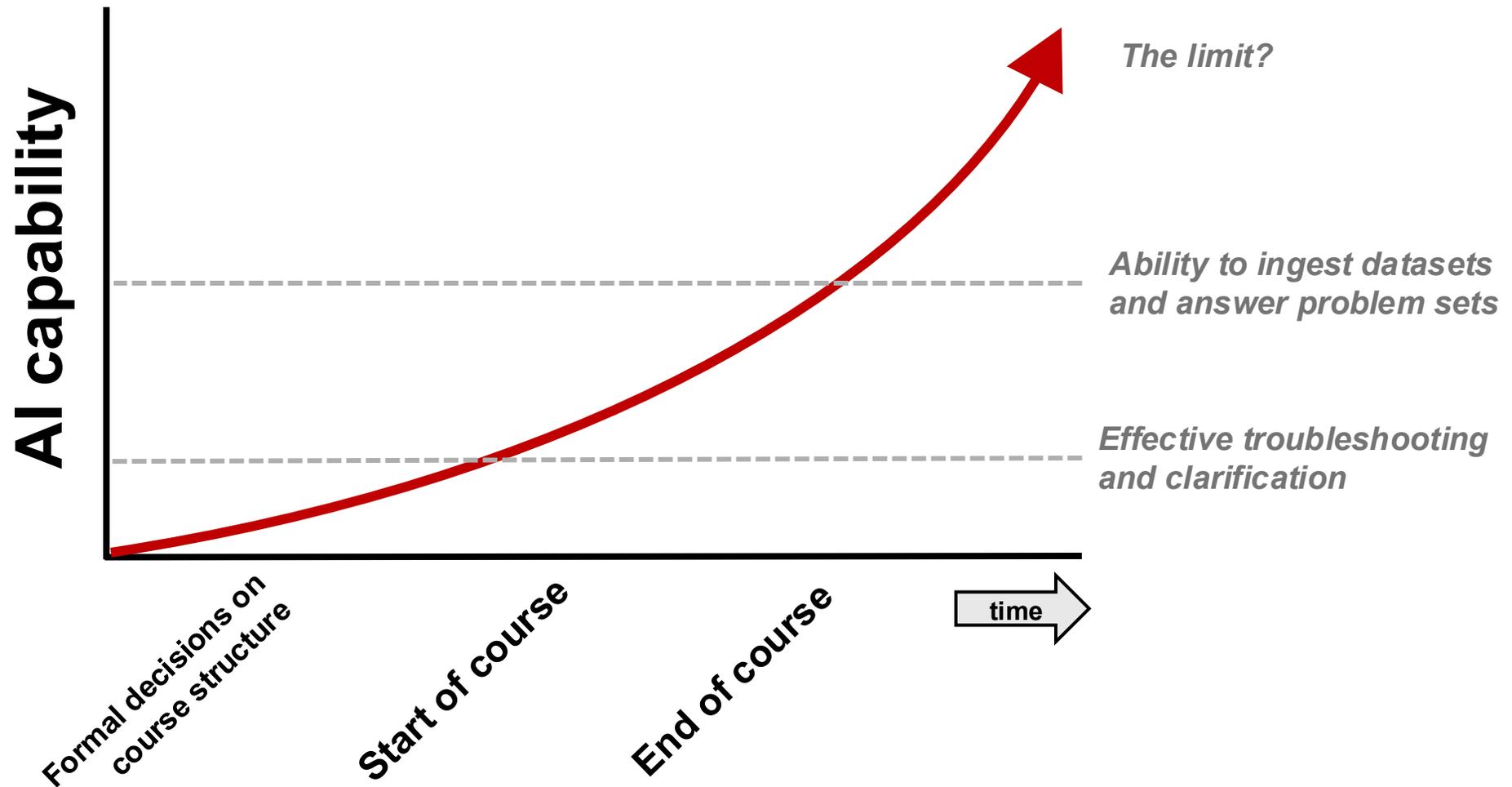
Teaching delivery & AI's role



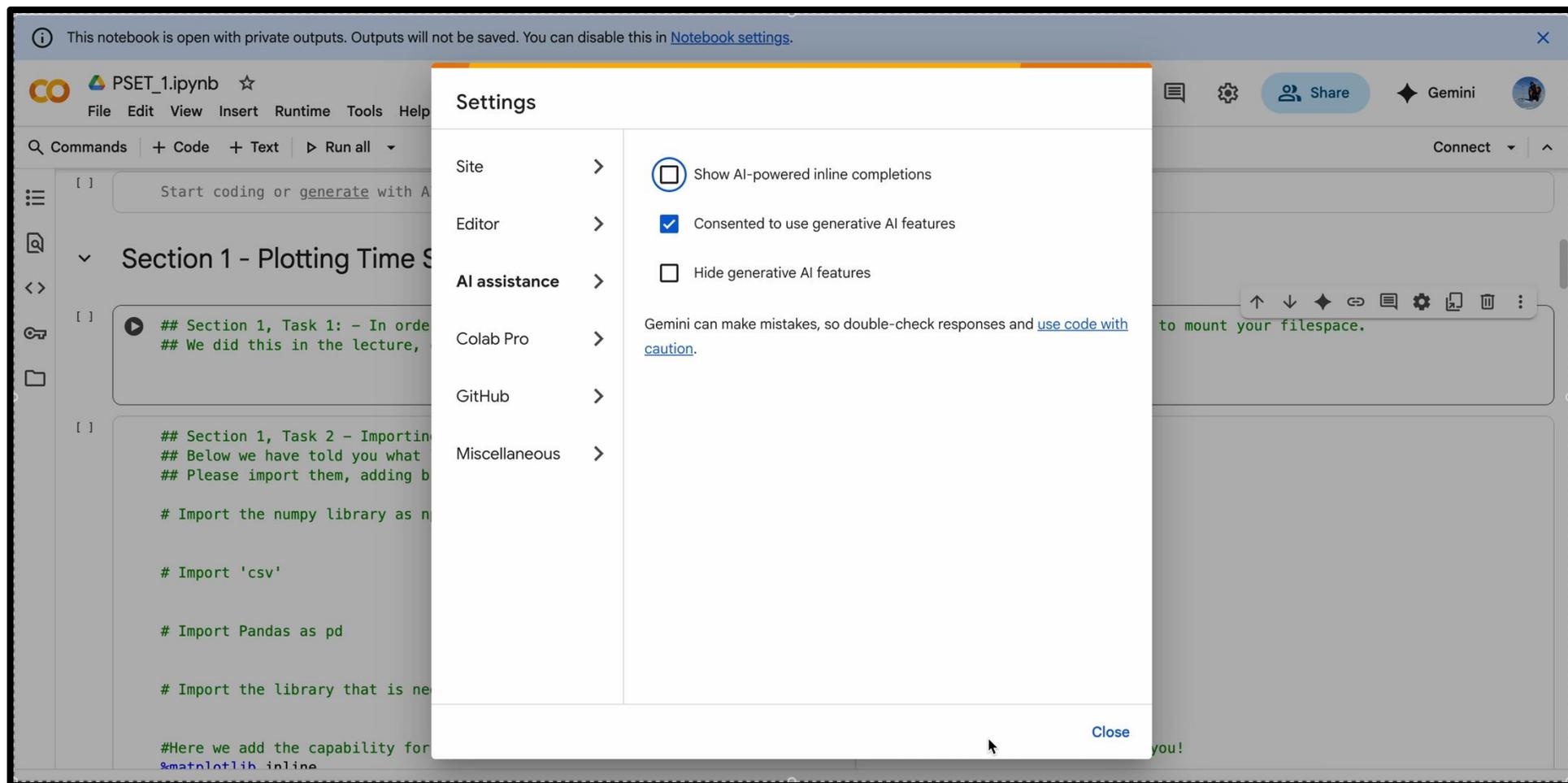
The screenshot shows a Jupyter Notebook interface with the following elements:

- Header:** "Lecture2_solution.ipynb" with a star icon and a share icon. A "Share" button is visible on the right.
- Menu:** "File Edit View Insert Runtime Tools Help".
- Toolbar:** "Commands", "+ Code", "+ Text", "Run all", "RAM", "Disk", and a Gemini icon.
- Left Sidebar:** A navigation menu with icons for home, search, navigation, and a folder.
- Main Content:**
 - Section Header:** "Lecture 02: Working with and visualising climate and environmental data"
 - Sub-headers:** "data types | logical indexing (cont'd) | reading in & manipulating time series | effective visualisation | netCDF's | plotting spatial data"
 - Text:** "This class builds on last week's 'Intro to Python' by introducing essential tools for reading, manipulating, and visualizing climate data in Python."
 - Text:** "We'll begin with a quick review of Python's core data types and their common applications. From there, we'll focus on practical skills: transforming raw datasets into clear, meaningful visualisations. While you'll continue to refine your Python fundamentals throughout the course, today's emphasis is on creating accurate and engaging visuals, which are important for effective communication—both to your audience and to yourself as an analyst."
 - Text:** "Climate and environmental data come in many forms, reflecting the complexity of Earth systems. Key types include:"
 - List-Group:**
 - Time series data: often presented in tables or sequential files, representing variables over time.
 - Spatiotemporal data: gridded datasets spanning two or more dimensions, such as latitude, longitude, time, and elevation or depth.
 - Text:** "To analyse climate data effectively, it's important to handle different formats appropriately to extract insights across spatial and temporal scales. Today, we'll focus on two common types of data:"
 - List-Group:**
 - Constantly spaced time series data.
 - Three-dimensional gridded spatial data.
 - Text:** "This week we'll also continue working with our standard data analysis libraries, NumPy (<https://numpy.org/>), Pandas (<https://pandas.pydata.org/>), and Matplotlib (<https://matplotlib.org/>). We'll also draw from two new packages today that are helpful for

Teaching delivery & AI's role



Assessment, AI, and shortcomings



The screenshot shows the Google Colab interface with a 'Settings' dialog box open. The dialog box is titled 'Settings' and has a 'Close' button at the bottom right. The settings are organized into sections:

- Site**: Show AI-powered inline completions
- Editor**: Consented to use generative AI features
- AI assistance**: Hide generative AI features

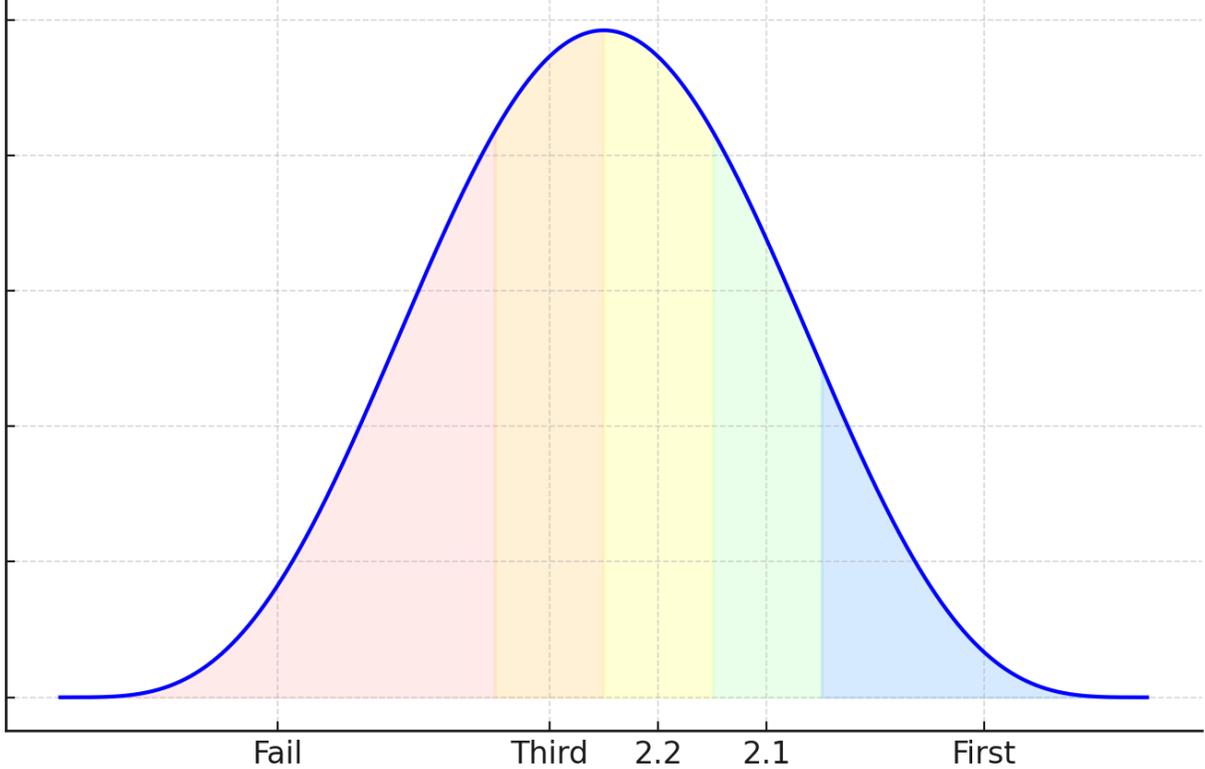
Below these sections, there is a warning: "Gemini can make mistakes, so double-check responses and [use code with caution.](#)"

The background interface shows a notebook titled 'PSET_1.ipynb' with a code cell containing the following text:

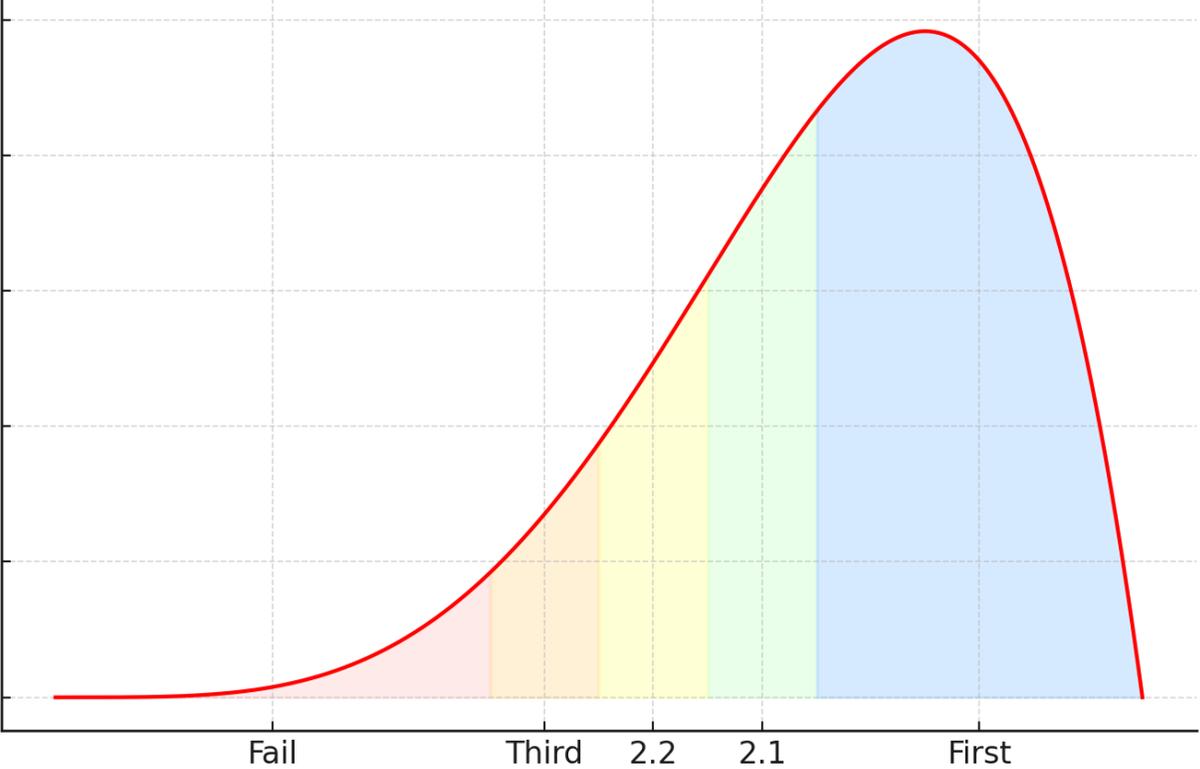
```
## Section 1, Task 1: - In orde  
## We did this in the lecture,  
  
## Section 1, Task 2 - Importin  
## Below we have told you what  
## Please import them, adding b  
  
# Import the numpy library as n  
  
# Import 'csv'  
  
# Import Pandas as pd  
  
# Import the library that is ne  
  
#Here we add the capability for  
&matnplotlib inline
```

Assessment, AI, and shortcomings

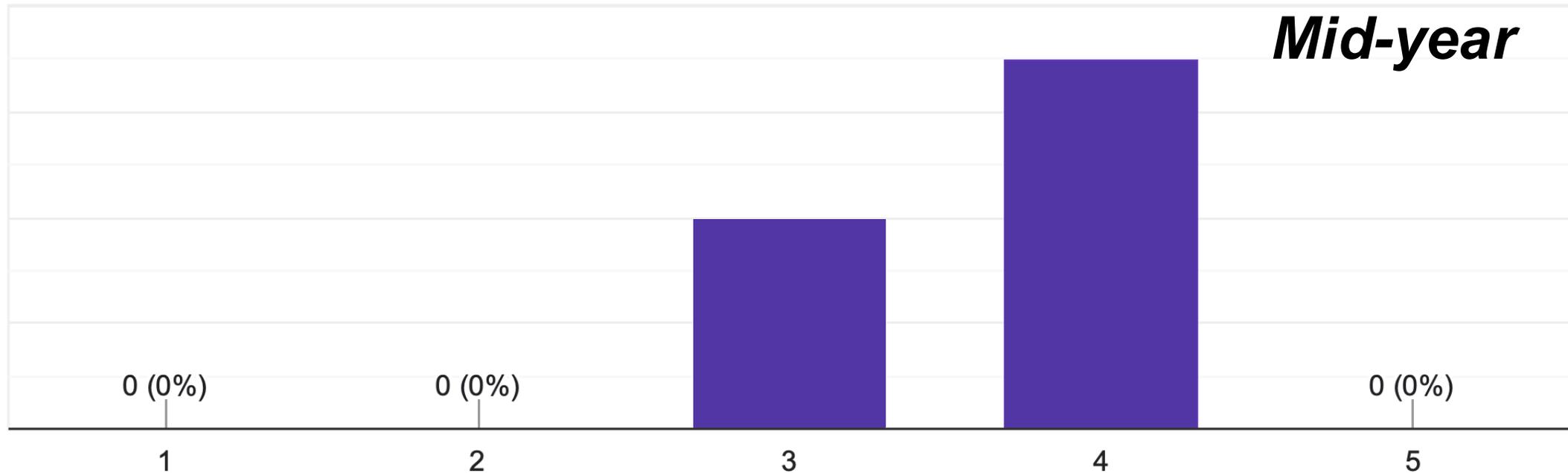
“Typical” mark distribution



Our paper had higher marks!



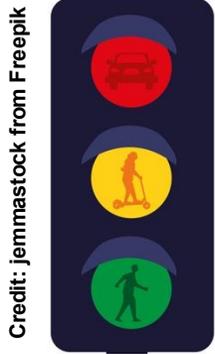
There were clear successes, too



On a scale of 1-5, how familiar are you with coding and environmental data analysis?

Looking forward

Geography's approach this year: course-level autonomy, but little wider consistency



AI use is not allowed.

AI can be used with caution and full disclosure in line with the Department policy.

AI use is encouraged.

Our view:

- **Every student is using AI — consistency and enforcement matter**
- **Granular rules help, but Cambridge needs a unified AI stance and agility**