

# Using ChatGPT's Advanced Data Analysis Tool

2023-11-02



# Get started

- Go to [chat.openai.com](https://chat.openai.com) and sign up & in
- The Advanced Data Analysis tool is only available for ChatGPT Plus or Enterprise subscribers
- If you do not have it, you can just watch
- Also sign up & in for DataCamp to access the instructions workspace



# What is ChatGPT Advanced Data Analysis?

- An extension to ChatGPT that has access to Python
- Allows you to upload & download files and perform operations on them



# Because ChatGPT is a really inefficient calculator



What is  $2 * 3 / 7 * 42.5 * 0.01 * 1.53 * 2.11 / 2.56 * 0.45$



To calculate this expression, follow the order of operations, which is parentheses, exponents, multiplication and division (from left to right), and addition and subtraction (from left to right). This expression only involves multiplication and division, so we'll just go from left to right:

$$2 * 3 / 7 * 42.5 * 0.01 * 1.53 * 2.11 / 2.56 * 0.45$$

$$= 6 / 7 * 42.5 * 0.01 * 1.53 * 2.11 / 2.56 * 0.45$$

$$= 0.8571428571 * 42.5 * 0.01 * 1.53 * 2.11 / 2.56 * 0.45$$

$$= 36.42857142 * 0.01 * 1.53 * 2.11 / 2.56 * 0.45$$

$$= 0.3642857142 * 1.53 * 2.11 / 2.56 * 0.45$$

$$= 0.5575769428 * 2.11 / 2.56 * 0.45$$

$$= 1.1762064505 / 2.56 * 0.45$$

$$= 0.4594548635 * 0.45$$

$$= 0.2066546886$$

So, the result of the calculation is approximately 0.2067.

I asked ChatGPT to compute

$$2 * 3 / 7 * 42.5 * 0.01 * 1.53 * 2.11 / 2.56 * 0.45$$

The correct answer is 0.2067229

- Python can compute this in microseconds.
- GPT has recently dramatically improved in its math capabilities but...
- It needs to explain its working to get a reasonable approximation to the answer
- So it takes several seconds to run, and the answer is only correct to 4 decimal places.





# ChatGPT and DataCamp Workspace are converging



DALLE-3 drew this cartoon of two people running towards each other from opposite directions


- ChatGPT's Advanced Data Analysis tool is an LLM with data analysis capabilities added
- DataCamp Workspace is a data analysis environment with an LLM added

Either way, the future of data analysis by mixing code and natural language.



# Enabling Advanced Data Analysis

 My plan

 Custom instructions

 Settings & Beta

 Log out



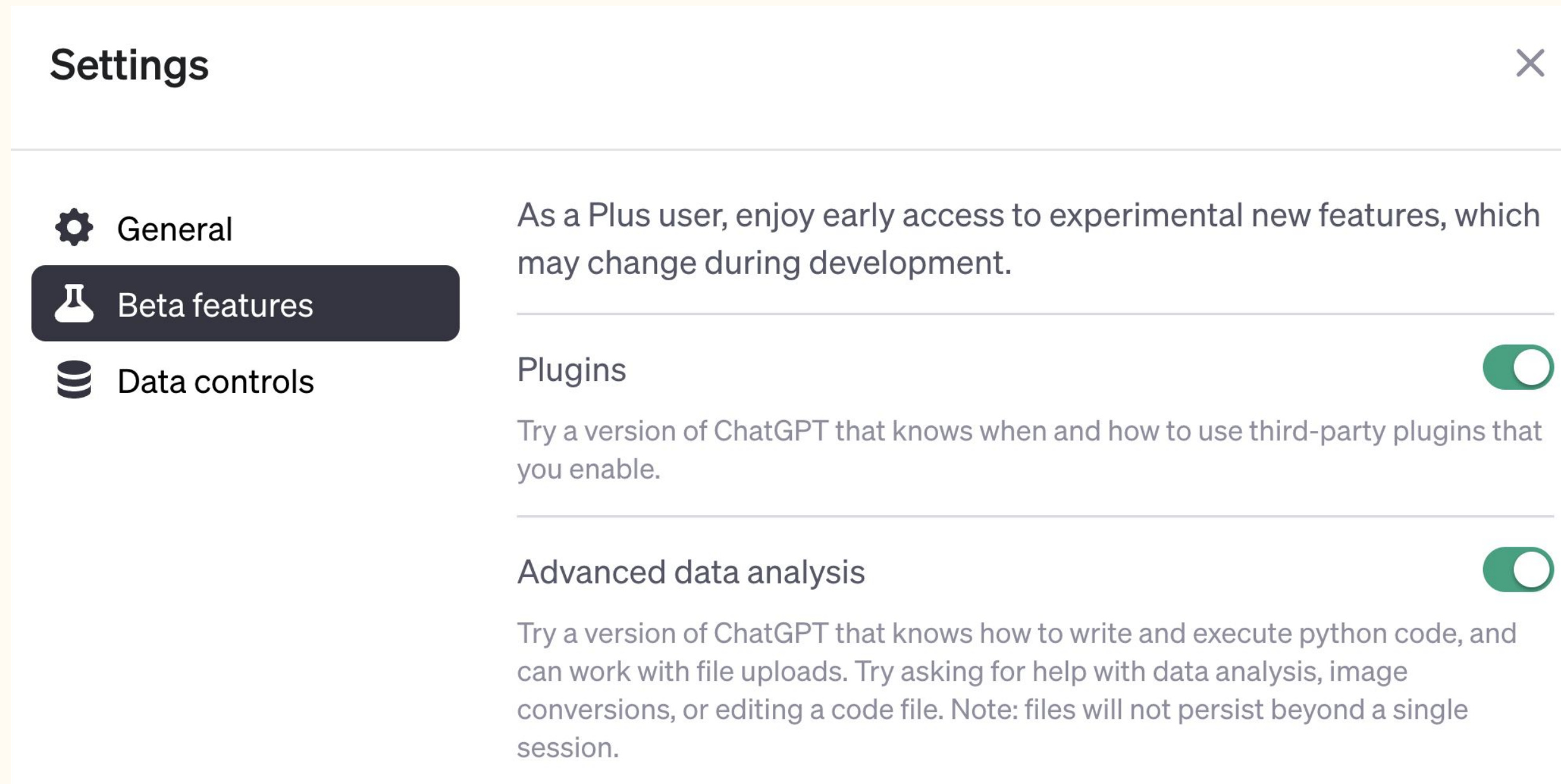
**Richie Cotton**



- In the bottom left, click the ellipsis next to your name
- Click 'Settings & Beta'



# Enabling Advanced Data Analysis

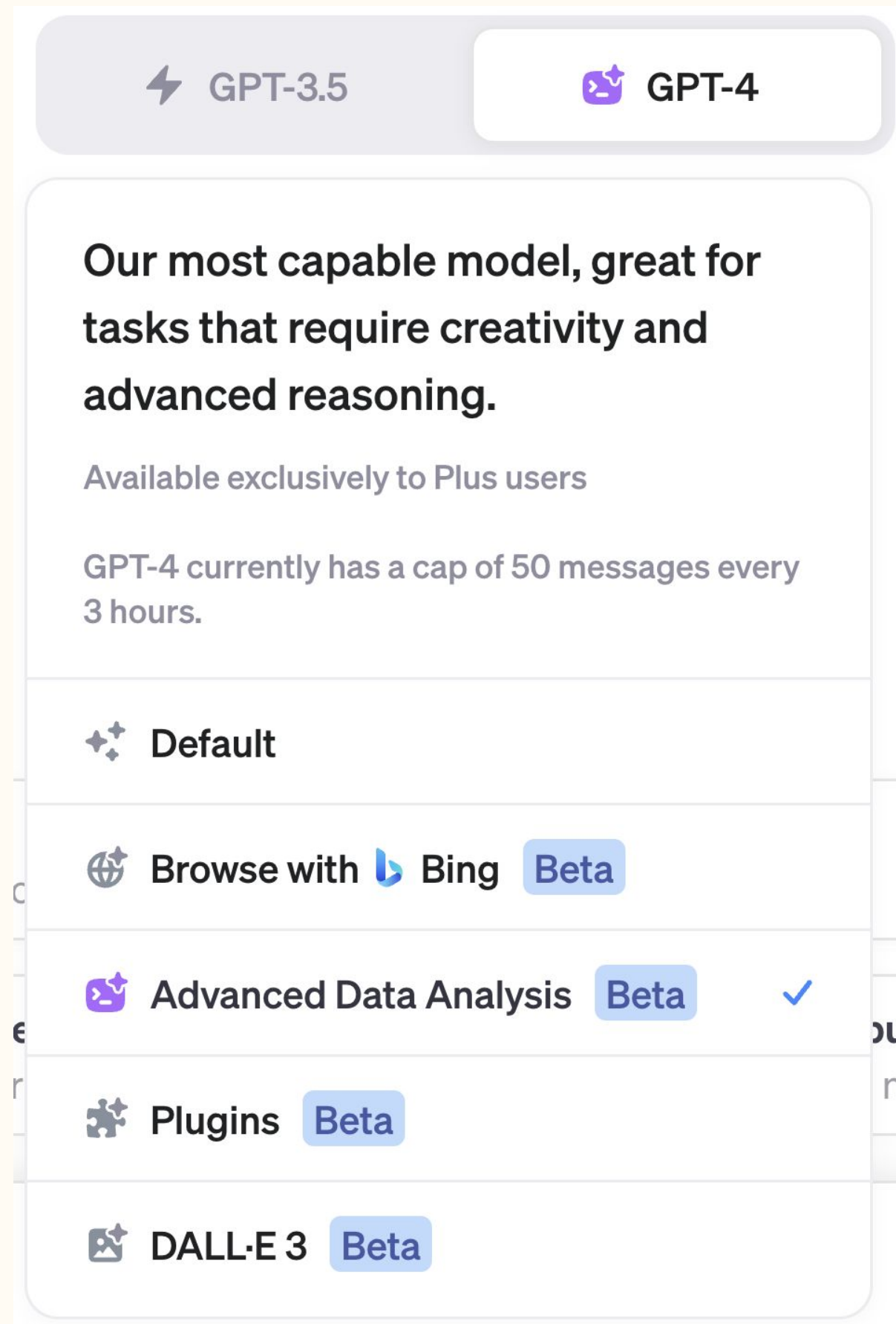


- Select the 'Beta features' pane
- Toggle 'Advanced data analysis' to on





# Enabling Advanced Data Analysis



- Start a new chat
- Select the GPT-4 model
- Hover over GPT-4 and in the dropdown menu select 'Advanced Data Analysis'

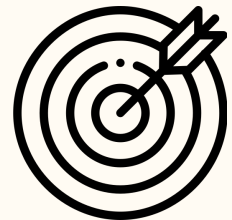




# The first dataset



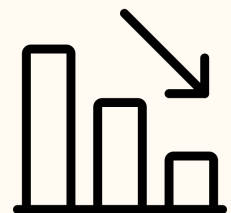
**Data:** The world's highest grossing media franchises



**Primary Goal:** Understand how to use ChatGPT and the Advanced Data Analysis tool for data analysis.



**Secondary Goal:** Understand data upload & download, plus prompts for analysis.



**Why?** This is the simplest dataset I could think of



# The second dataset



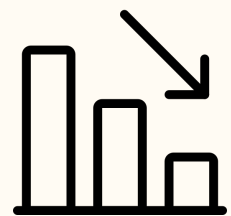
**Data:** This week's NFL American Football results



**Primary Goal:** Learn how to perform web scraping with ChatGPT



**Secondary Goal:** Tips for dealing with ChatGPT failing to write Python code.



**Why?** pro-football-reference.com is pathologically hard to scrape, so it's best to pass this task to AI.

