# Prompt Engineering Tips

| Tip | DO<br>Example of good use of the tip | DON'T<br>Example of bad use of the tip |
|---|---|---|
| Clear, detailed, non-ambiguous, and concise. | I have 3000 Reddit posts. I want to use topic modeling to find out what the posts are talking about. Before doing the topic modeling, how do I need to prepare the text? | How to pre-process text for NLP? |
| Tell the LLM what role to take, and your own expertise. | You are an expert in computational biology, particularly genomics. What are the key R packages for statistics? | What are the key R packages for statistics? |
| Describe the programming language, libraries, and other technologies. | How do I scrape a website in Python using Selenium? | How do I scrape a website in Python? |
| Explain what the code is for. | I want to scrape a website that relies heavily on JavaScript. I want to get the source code after loading all elements. How can I do that with Python? | I want to scrape a website. How can I do that with Python? |
| Specify any constraints or requirements. | Using ggplot2 in R, I want to iterate over 100 columns, creating the same plot with each. Please write code to do that considering that I'm running the code in my university's high-performance computing cluster. | Using ggplot2 in R, I want to iterate over 100 columns, creating the same plot with each. Please write code to do that. |
| Be specific about what you want the code to do. | Write a function in Python that takes two numbers, adds them, and returns the result. The function should raise a type error if the inputs are not numbers. | Write a function in Python that takes two numbers, adds them, and returns the result. |
| Ask the LLM to explain itself and any assumptions. | Write a regular expression to match salaries in job postings. Explain any assumptions that you're making about the data and what the match should be. | Write a regular expression to match salaries in job postings. |

| Tip | DO<br>Example of good use of the tip | DON'T<br>Example of bad use of the tip |
|---|---|---|
| Provide examples of the desired behavior. | Write a regular expression to match salaries in job postings. These are some examples:<br><br>String: Minimum $60,000 ANNUAL (12 months)<br>Match: $60,000<br><br>String: Rate of Pay $78,035 - $106,517<br>Match: $78,035 and $106,517<br><br>String: starting salary will be from £46,047 up to £61,823<br>Match: £46,047 and £61,823 | Write a regular expression to match salaries in job postings. |
| Ask the LLM to work step-by-step. | What is this code doing?<br>iris %>% group_by(Species) %>% summarize_all(mean) %>% ungroup %>% gather(measure, value, -Species) %>% arrange(value)<br>Work step-by-step. | What is this code doing?<br>iris %>% group_by(Species) %>% summarize_all(mean) %>% ungroup %>% gather(measure, value, -Species) %>% arrange(value) |
| Specify how you want the LLM to give you the information. | What are different packages for supervised learning in R, as well as their pros and cons? Provide the information in a table with only one row per package. | What are different packages for supervised learning in R, as well as their pros and cons? |

**You probably won't get the prompt right the first time. Iterate—reformulate your prompts as needed!**