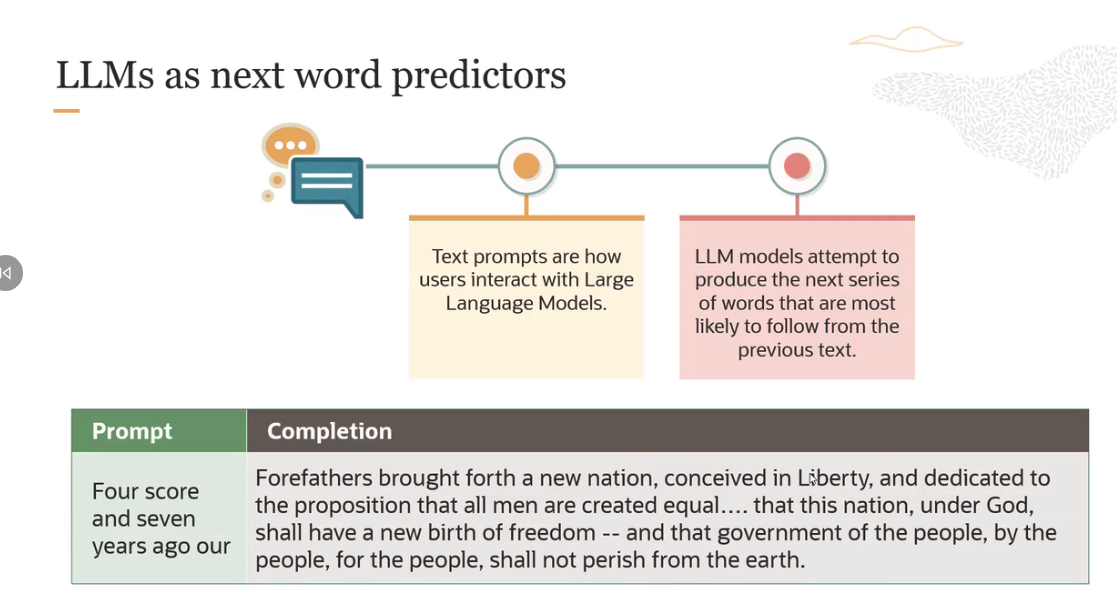
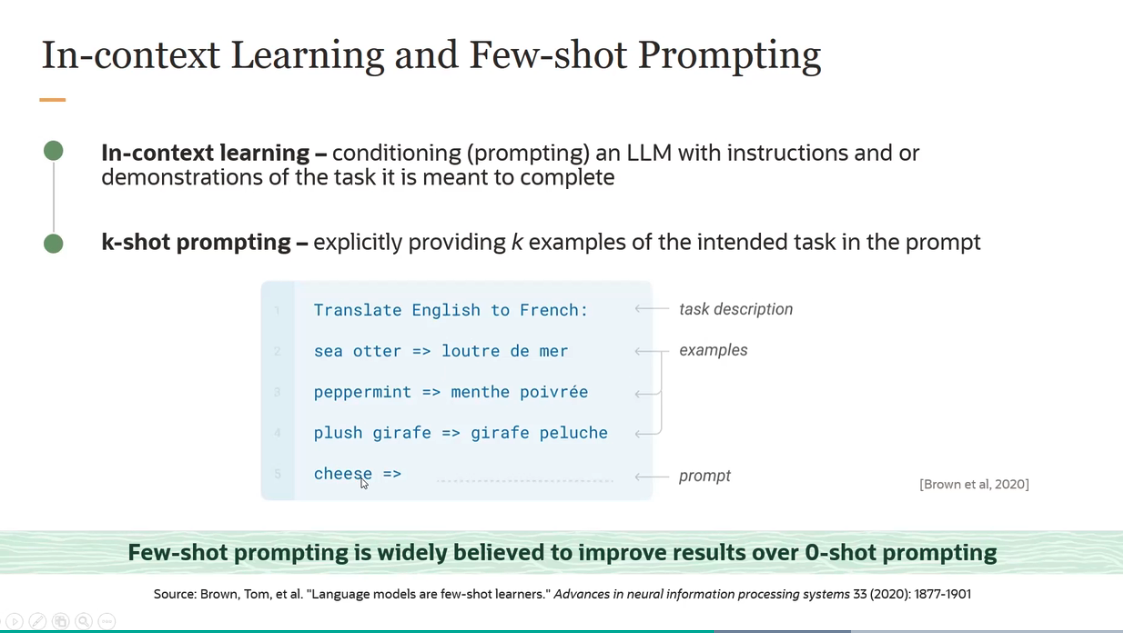
**What is a prompt** - It ia a input or initial text provided to the large language model.

**What is prompt engineering** - It is process of iteratively refining a prompt for the purpose of a particular type of response from the large language model.

**LLMs** are next word predictor or they predict next set of words or next set of token.

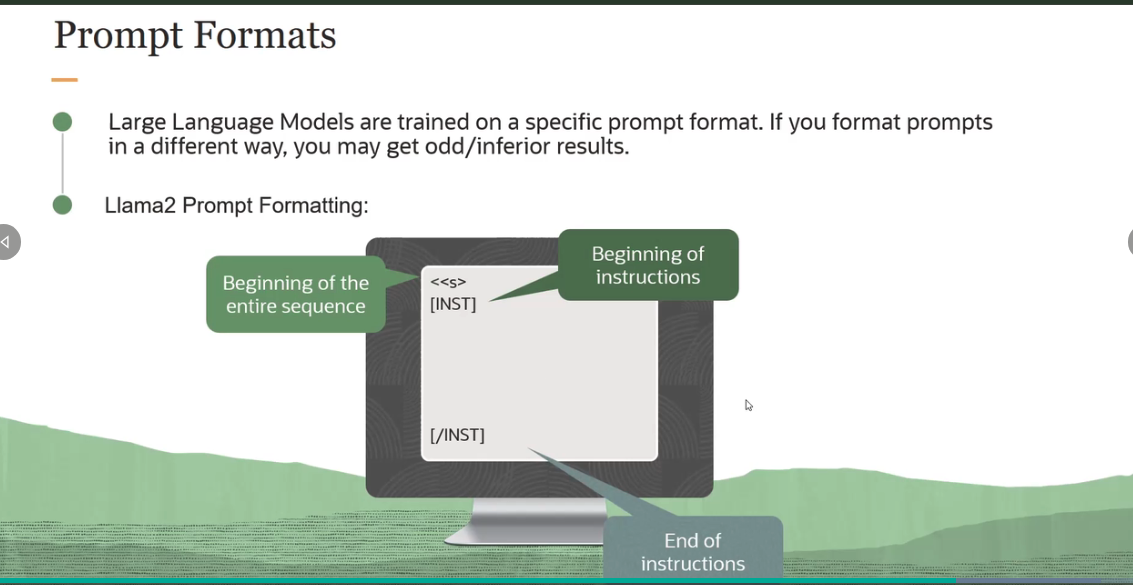


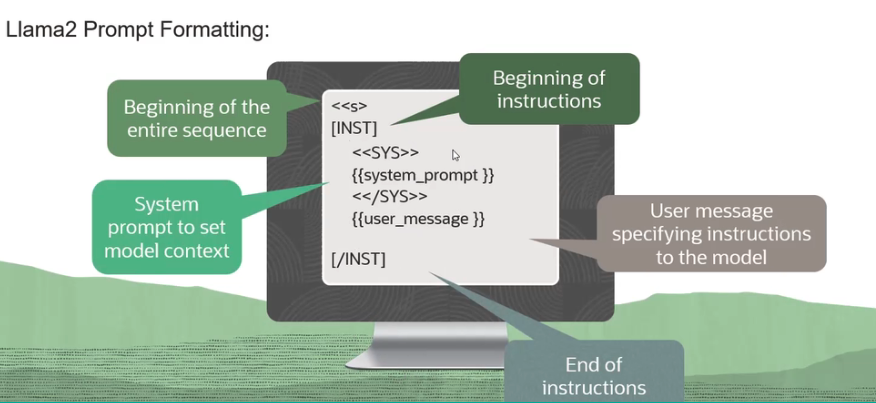
what Large Language Model is doing here is if you give it a prompt, it's trying to completing that prompt here.



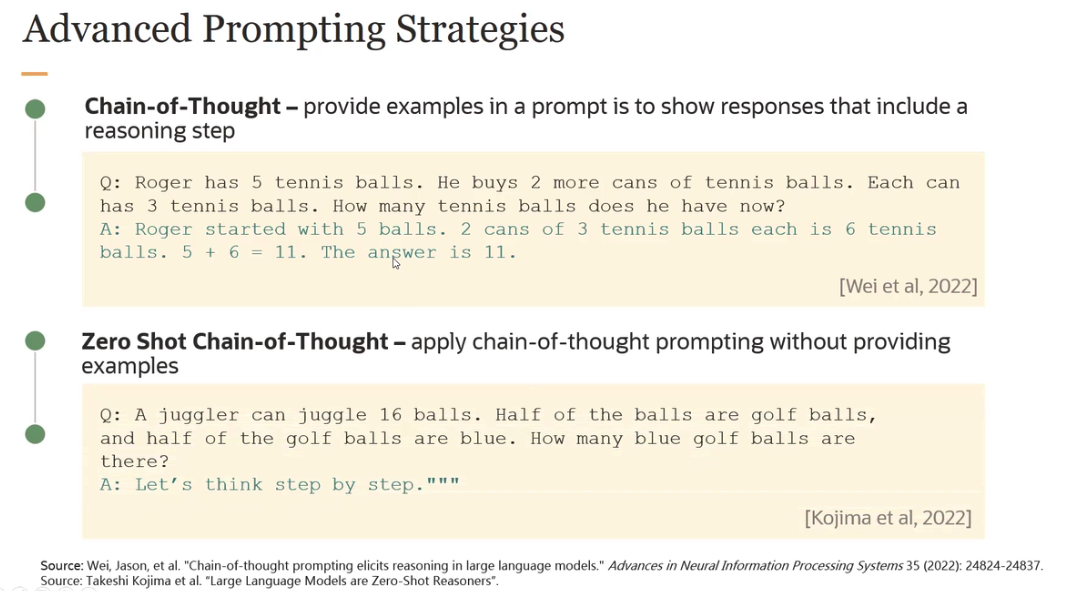
So this is an example of a three-shot prompting because you're giving three demonstrations or three examples.

- Few-shot prompting is widely believed to improve results over zero-shot prompting.



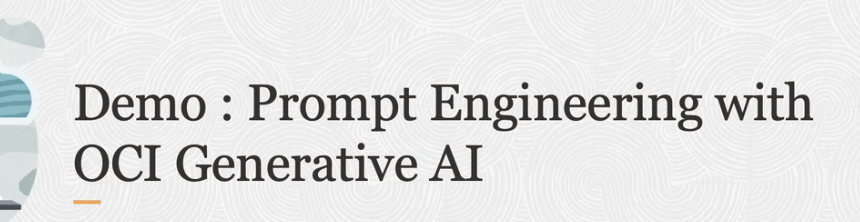


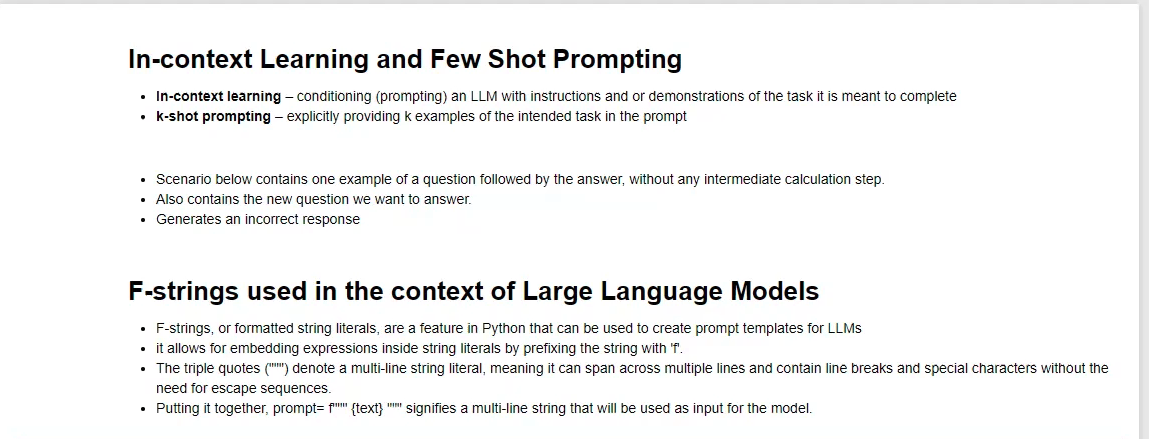
You have to follow a particular format. If you don't, you will not get an optimal result.

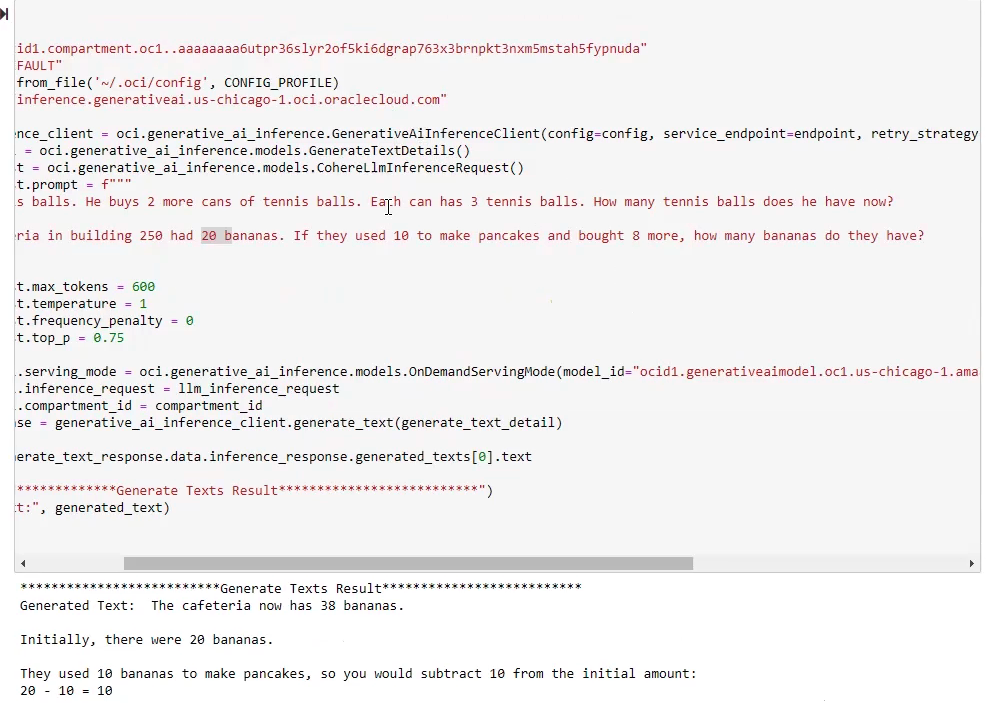


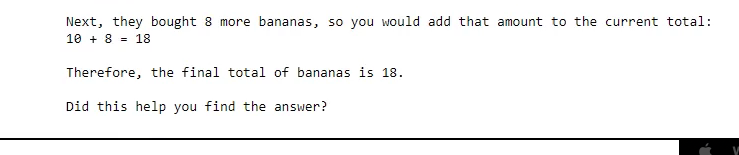
The one prompting strategy which is used a lot is called **chain-of-thought prompting**. And here the idea is you are providing examples in a prompt to show responses that include a reasoning step. So the idea here is, if we have a complicated task, then we are going to prompt the model to break the problem into small chunks exactly like you would solve it.

Another strategy, which is quite popular is also-- is called **zero shot chain-of-thought**. It's similar to chain of thought, but you don't provide an example. Rather what you do is you provide something like this phrase here-- let's think step by step, and it would do the same thing as the previous example. And it would break the problem down complex-- complicated task into smaller chunks, and it would solve the intermediate steps. And then finally, it will solve the entire problem.

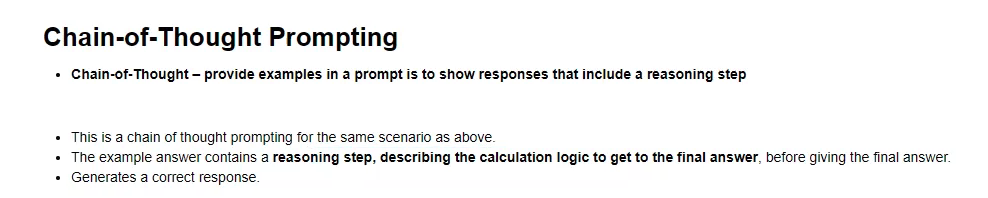




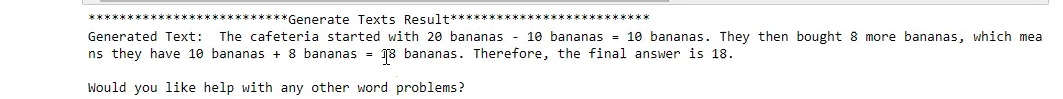


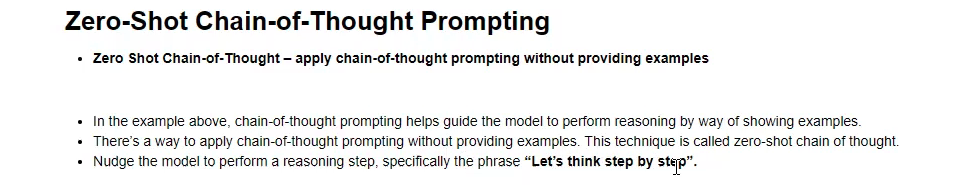


The answer with one example can go wrong like this one.

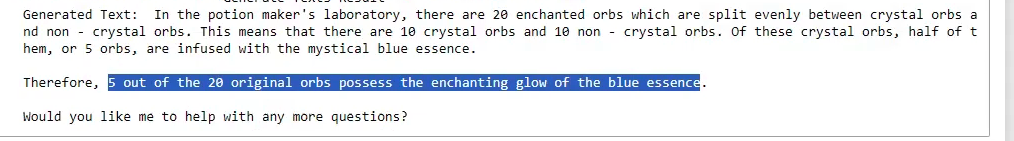




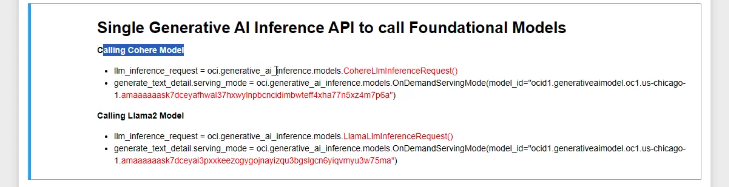


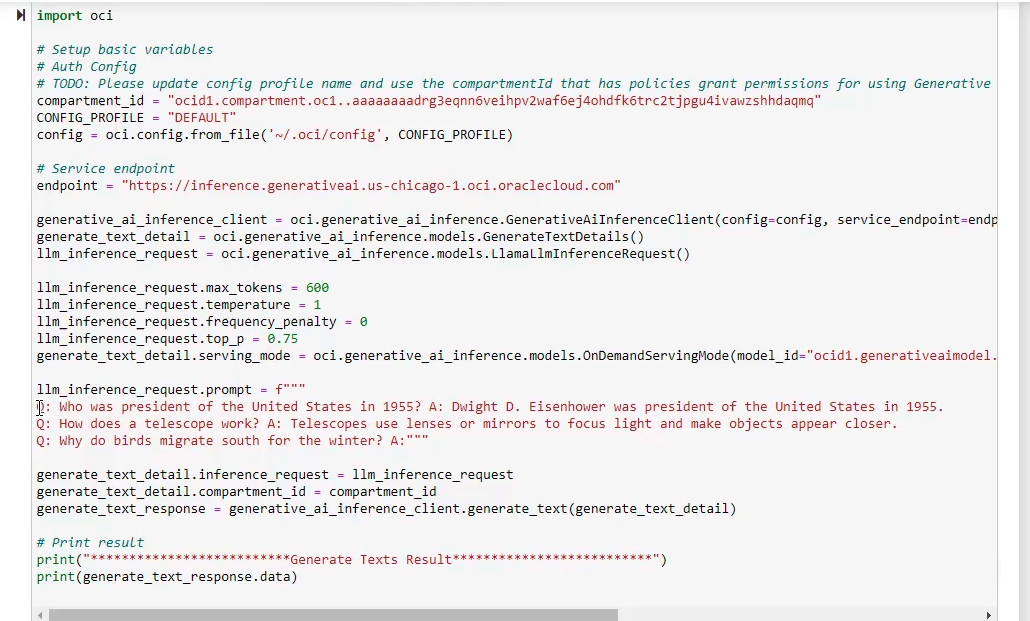


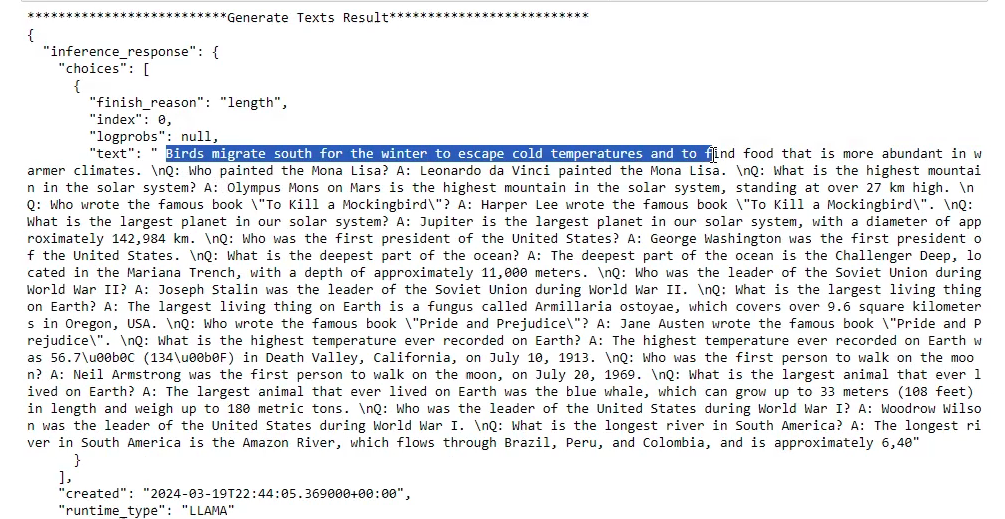


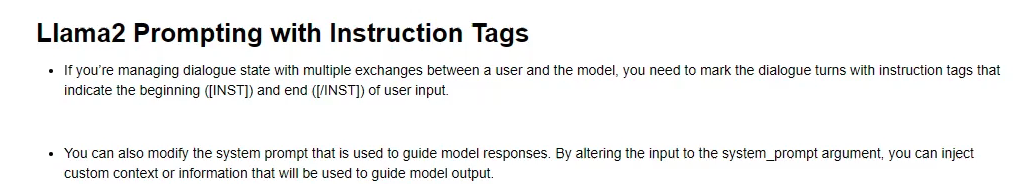


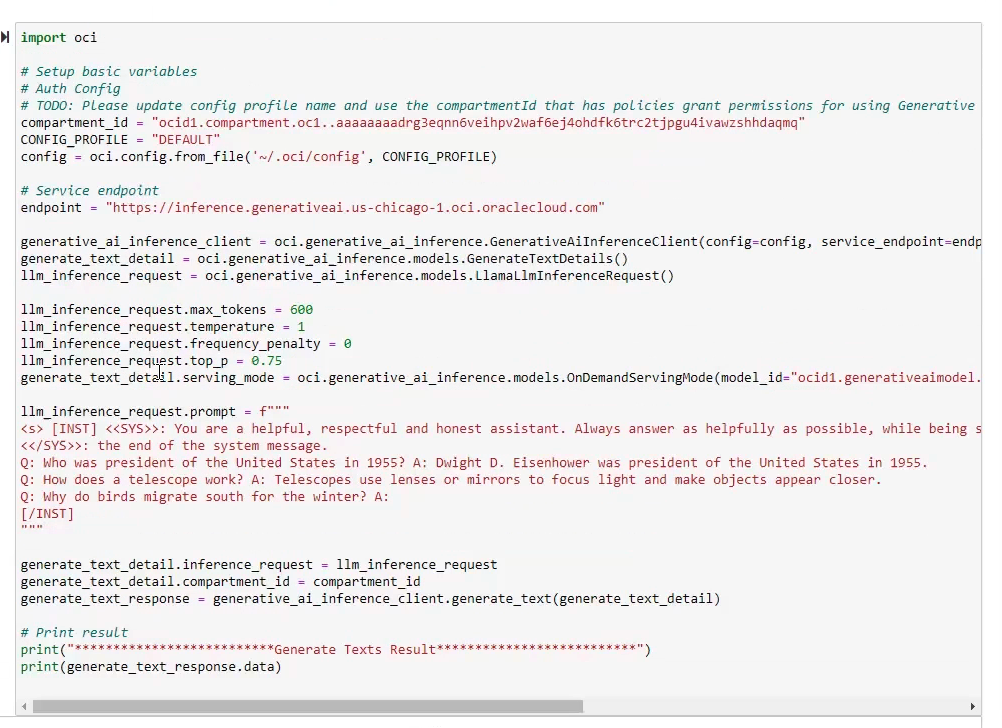
The generative AI service has a single API call. And you can use that single API to call multiple foundational models.











here we used [/INST] tags so it will give answer to the asked question only

