NLP Test

The following is a task to demonstrate your understanding and experience of some of the major components of Natural Language Processing. There isn’t necessarily a right answer.

Guidelines:

* We seek well-presented ideas and code, not complicated and/or messy solutions.
* The code should be well-named, structured, and commented on.
* Don’t spend more than 3-4 hours creating your solution.
* Usage of ChatGPT and similar tools is allowed but not encouraged unless you can explain in detail your code and solution regardless.
* Submission should be in .ipynb (code output must be included)
* **Explain your decisions** using markdown cells in the notebook.
* The file should be named *nlptask.ipynb*

**Task**: Write two short NLP pipelines in Python. The first one should identify tokens that refer to **gold** and **silver** (with separate identifiers for each) as financial commodities. Print out the number of documents in the sample file containing at least one reference to gold and the same for silver. The second pipeline should compute the sentiment of sentences where gold or silver are mentioned as commodities. Print out one overall (full sample file) sentiment value for gold and for silver separately.