

Double-click (or enter) to edit

Start coding or [generate with AI](#).

## ▼ 1. Install and Import Essential NLP Libraries: NLTK and spaCy

```
import sys
!{sys.executable} -m pip install nltk spacy
!{sys.executable} -m spacy download en_core_web_sm
```

```
Requirement already satisfied: nltk in /usr/local/lib/python3.12/dist-packages (3.9.1)
Requirement already satisfied: spacy in /usr/local/lib/python3.12/dist-packages (3.8.11)
Requirement already satisfied: click in /usr/local/lib/python3.12/dist-packages (from nltk) (8.3.1)
Requirement already satisfied: joblib in /usr/local/lib/python3.12/dist-packages (from nltk) (1.5.3)
Requirement already satisfied: regex>=2021.8.3 in /usr/local/lib/python3.12/dist-packages (from nltk) (2025.11.3)
Requirement already satisfied: tqdm in /usr/local/lib/python3.12/dist-packages (from nltk) (4.67.1)
Requirement already satisfied: spacy-legacy<3.1.0,>=3.0.11 in /usr/local/lib/python3.12/dist-packages (from spacy) (3.0.12)
Requirement already satisfied: spacy-loggers<2.0.0,>=1.0.0 in /usr/local/lib/python3.12/dist-packages (from spacy) (1.0.5)
Requirement already satisfied: murmurhash<1.1.0,>=0.28.0 in /usr/local/lib/python3.12/dist-packages (from spacy) (1.0.15)
Requirement already satisfied: cymem<2.1.0,>=2.0.2 in /usr/local/lib/python3.12/dist-packages (from spacy) (2.0.13)
Requirement already satisfied: preshed<3.1.0,>=3.0.2 in /usr/local/lib/python3.12/dist-packages (from spacy) (3.0.12)
Requirement already satisfied: thinc<8.4.0,>=8.3.4 in /usr/local/lib/python3.12/dist-packages (from spacy) (8.3.10)
Requirement already satisfied: wasabi<1.2.0,>=0.9.1 in /usr/local/lib/python3.12/dist-packages (from spacy) (1.1.3)
Requirement already satisfied: srsly<3.0.0,>=2.4.3 in /usr/local/lib/python3.12/dist-packages (from spacy) (2.5.2)
Requirement already satisfied: catalogue<2.1.0,>=2.0.6 in /usr/local/lib/python3.12/dist-packages (from spacy) (2.0.10)
Requirement already satisfied: weasel<0.5.0,>=0.4.2 in /usr/local/lib/python3.12/dist-packages (from spacy) (0.4.3)
Requirement already satisfied: typer-slim<1.0.0,>=0.3.0 in /usr/local/lib/python3.12/dist-packages (from spacy) (0.20.0)
Requirement already satisfied: numpy>=1.19.0 in /usr/local/lib/python3.12/dist-packages (from spacy) (2.0.2)
Requirement already satisfied: requests<3.0.0,>=2.13.0 in /usr/local/lib/python3.12/dist-packages (from spacy) (2.32.4)
Requirement already satisfied: pydantic!=1.8,!=1.8.1,<3.0.0,>=1.7.4 in /usr/local/lib/python3.12/dist-packages (from spacy) (2.12.3)
Requirement already satisfied: jinja2 in /usr/local/lib/python3.12/dist-packages (from spacy) (3.1.6)
Requirement already satisfied: setuptools in /usr/local/lib/python3.12/dist-packages (from spacy) (75.2.0)
Requirement already satisfied: packaging>=20.0 in /usr/local/lib/python3.12/dist-packages (from spacy) (25.0)
Requirement already satisfied: annotated-types>=0.6.0 in /usr/local/lib/python3.12/dist-packages (from pydantic!=1.8,!=1.8.1,<3.0.0,>=1)
Requirement already satisfied: pydantic-core==2.41.4 in /usr/local/lib/python3.12/dist-packages (from pydantic!=1.8,!=1.8.1,<3.0.0,>=1)
Requirement already satisfied: typing-extensions>=4.14.1 in /usr/local/lib/python3.12/dist-packages (from pydantic!=1.8,!=1.8.1,<3.0.0)
Requirement already satisfied: typing-inspection>=0.4.2 in /usr/local/lib/python3.12/dist-packages (from pydantic!=1.8,!=1.8.1,<3.0.0,
Requirement already satisfied: charset_normalizer<4,>=2 in /usr/local/lib/python3.12/dist-packages (from requests<3.0.0,>=2.13.0->spac
Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.12/dist-packages (from requests<3.0.0,>=2.13.0->spacy) (3.11)
Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.12/dist-packages (from requests<3.0.0,>=2.13.0->spacy) (2.
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.12/dist-packages (from requests<3.0.0,>=2.13.0->spacy) (20
Requirement already satisfied: blis<1.4.0,>=1.3.0 in /usr/local/lib/python3.12/dist-packages (from thinc<8.4.0,>=8.3.4->spacy) (1.3.3)
Requirement already satisfied: confection<1.0.0,>=0.0.1 in /usr/local/lib/python3.12/dist-packages (from thinc<8.4.0,>=8.3.4->spacy) (
Requirement already satisfied: cloudpathlib<1.0.0,>=0.7.0 in /usr/local/lib/python3.12/dist-packages (from weasel<0.5.0,>=0.4.2->spacy)
```

```
Requirement already satisfied: smart-open<8.0.0,>=5.2.1 in /usr/local/lib/python3.12/dist-packages (from weasel<0.5.0,>=0.4.2->spacy)
Requirement already satisfied: MarkupSafe>=2.0 in /usr/local/lib/python3.12/dist-packages (from jinja2->spacy) (3.0.3)
Requirement already satisfied: wrapt in /usr/local/lib/python3.12/dist-packages (from smart-open<8.0.0,>=5.2.1->weasel<0.5.0,>=0.4.2->
Collecting en-core-web-sm==3.8.0
  Downloading https://github.com/explosion/spacy-models/releases/download/en_core_web_sm-3.8.0/en_core_web_sm-3.8.0-py3-none-any.whl (12.8/12.8 MB 116.3 MB/s eta 0:00:00)
    ✓ Download and installation successful
  You can now load the package via spacy.load('en_core_web_sm')
  △ Restart to reload dependencies
  If you are in a Jupyter or Colab notebook, you may need to restart Python in
  order to load all the package's dependencies. You can do this by selecting the
  'Restart kernel' or 'Restart runtime' option.
```

```
import nltk
import spacy

# Download necessary NLTK data (e.g., for tokenization)
nltk.download('punkt')

# Load the English spaCy model
nlp = spacy.load('en_core_web_sm')

print("NLTK and spaCy are installed and imported successfully!")

[nltk_data] Downloading package punkt to /root/nltk_data...
[nltk_data]  Unzipping tokenizers/punkt.zip.
NLTK and spaCy are installed and imported successfully!
```

## ▼ 2. Load a Short Paragraph

```
paragraph = """Natural Language Processing (NLP) is a field of artificial intelligence that focuses on the interaction between computers and human language. It involves the development of computer programs that can analyze, understand, and generate human language with minimal human intervention. NLP has numerous applications, such as speech recognition, machine translation, sentiment analysis, and text summarization. One of the most common NLP tasks is text classification, where a computer program identifies the category of a given text based on its content. Another important task is named entity recognition, which involves identifying specific entities mentioned in a text, such as people, places, and organizations. NLP is a multidisciplinary field that requires knowledge from various domains, including linguistics, computer science, and mathematics. It is a rapidly growing field with many exciting applications in various industries, such as healthcare, finance, and e-commerce. As the amount of digital data continues to grow, NLP will become even more important for processing and analyzing this data effectively. In conclusion, Natural Language Processing is a fascinating field that has the potential to revolutionize the way we interact with computers and each other through language. Its applications are vast and varied, making it a crucial area of research and development in the field of artificial intelligence. By understanding the principles of NLP, we can develop more intelligent and effective computer programs that can better serve our needs and improve our lives in countless ways. Whether you are a student, researcher, or developer, there is a lot to learn and explore in the world of Natural Language Processing. So why not dive in and start your journey today? You never know what amazing things you might discover along the way!"""

print("Paragraph loaded successfully:")
print(paragraph)

Paragraph loaded successfully:
Natural Language Processing (NLP) is a field of artificial intelligence that focuses on the interaction between computers and human language. It involves the development of computer programs that can analyze, understand, and generate human language with minimal human intervention. NLP has numerous applications, such as speech recognition, machine translation, sentiment analysis, and text summarization. One of the most common NLP tasks is text classification, where a computer program identifies the category of a given text based on its content. Another important task is named entity recognition, which involves identifying specific entities mentioned in a text, such as people, places, and organizations. NLP is a multidisciplinary field that requires knowledge from various domains, including linguistics, computer science, and mathematics. It is a rapidly growing field with many exciting applications in various industries, such as healthcare, finance, and e-commerce. As the amount of digital data continues to grow, NLP will become even more important for processing and analyzing this data effectively. In conclusion, Natural Language Processing is a fascinating field that has the potential to revolutionize the way we interact with computers and each other through language. Its applications are vast and varied, making it a crucial area of research and development in the field of artificial intelligence. By understanding the principles of NLP, we can develop more intelligent and effective computer programs that can better serve our needs and improve our lives in countless ways. Whether you are a student, researcher, or developer, there is a lot to learn and explore in the world of Natural Language Processing. So why not dive in and start your journey today? You never know what amazing things you might discover along the way!
```

## ▼ 3. Perform Simple Preprocessing Tasks

```
# Count the number of words
words = paragraph.split()
```

```
word_count = len(words)
print(f"Number of words in the paragraph: {word_count}")
```

```
Number of words in the paragraph: 58
```

```
# Convert the entire text to lowercase
lowercase_paragraph = paragraph.lower()
print("\nParagraph converted to lowercase:")
print(lowercase_paragraph)
```

```
Paragraph converted to lowercase:
```

```
natural language processing (nlp) is a field of artificial intelligence that focuses on the interaction between computers and human la
```