

```
!pip install nltk spacy
!python -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) (0.21.1)
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) (1.1.3)
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,>=1.7.4) (0.7.0)
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,>=1.7.4) (2.41.4)
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,>=1.7.4) (4.14.1)
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,>=1.7.4) (0.4.2)
Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/python3.12/dist-packages (from requests<3.0.0,>=2.13.0) (3.4.0)
Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.12/dist-packages (from requests<3.0.0,>=2.13.0) (3.10.1)
Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.12/dist-packages (from requests<3.0.0,>=2.13.0) (2.3.0)
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.12/dist-packages (from requests<3.0.0,>=2.13.0) (2025.11.11)
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) (1.3.0)
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) (0.0.1)
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) (0.7.0)
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) (7.0.5)
Requirement already satisfied: MarkupSafe>=2.0 in /usr/local/lib/python3.12/dist-packages (from Jinja2) (3.0.3)
Requirement already satisfied: wrapt in /usr/local/lib/python3.12/dist-packages (from smart-open<8.0.0,>=5.2.1) (1.17.0)
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 109.9 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 the NLTK library for natural language processing tasks.
import nltk
# Import the spaCy library, another powerful NLP tool.
import spacy
# Import pandas for data manipulation and analysis, especially with DataFrames.
import pandas as pd
# Import matplotlib.pyplot for creating static, interactive, and animated visualizations.
import matplotlib.pyplot as plt
# Import Counter from the collections module to easily count hashable objects.
from collections import Counter
```

```
# Download the 'punkt' tokenizer models from NLTK, used for splitting text into sentences and words.
nltk.download('punkt')
# Download the 'averaged_perceptron_tagger' from NLTK, used for part-of-speech tagging.
nltk.download('averaged_perceptron_tagger')
```

```
[nltk_data] Downloading package punkt to /root/nltk_data...
[nltk_data] Package punkt is already up-to-date!
[nltk_data] Downloading package averaged_perceptron_tagger to
[nltk_data] /root/nltk_data...
[nltk_data] Package averaged_perceptron_tagger is already up-to-
[nltk_data] date!
True
```

```
essay_text = """
Artificial Intelligence has transformed modern academic research.
It enables efficient data analysis and supports informed decision making.
Researchers argue that AI improves productivity while raising ethical concerns.
"""

# Define a multi-line string containing a short essay, which will be used for NLP tasks.
# This text serves as our sample data for tokenization and part-of-speech tagging.
```

```
# Print the defined essay text to display its content.
print(essay_text)
```

Artificial Intelligence has transformed modern academic research.
It enables efficient data analysis and supports informed decision making.
Researchers argue that AI improves productivity while raising ethical concerns.

```
# Download the 'punkt_tab' tokenizer models from NLTK, if not already downloaded.
nltk.download('punkt_tab')
# Use NLTK's word_tokenize to split the essay text into a list of individual words and punctuation.
tokens_nltk = nltk.word_tokenize(essay_text)
# Display the list of tokens generated by NLTK.
tokens_nltk
```

```
[nltk_data] Downloading package punkt_tab to /root/nltk_data...
[nltk_data] Package punkt_tab is already up-to-date!
['Artificial',
 'Intelligence',
 'has',
 'transformed',
 'modern',
 'academic',
 'research',
 '.',
 'It',
 'enables',
 'efficient',
 'data',
 'analysis',
 'and',
 'supports',
 'informed',
 'decision',
 'making',
 '.',
 'Researchers',
 'argue',
 'that',
 'AI',
 'improves',
 'productivity',
 'while',
 'raising',
 'ethical',
 'concerns',
 '.']
```

```
# Download the 'averaged_perceptron_tagger_eng' from NLTK for English part-of-speech tagging.
nltk.download('averaged_perceptron_tagger_eng')
# Perform part-of-speech tagging on the NLTK tokens, resulting in a list of (word, tag) tuples.
pos_nltk = nltk.pos_tag(tokens_nltk)
# Display the NLTK part-of-speech tagged list.
pos_nltk
```

```
[nltk_data] Downloading package averaged_perceptron_tagger_eng to
[nltk_data] /root/nltk_data...
[nltk_data] Package averaged_perceptron_tagger_eng is already up-to-
[nltk_data] date!
[('Artificial', 'JJ'),
 ('Intelligence', 'NNP'),
 ('has', 'VBZ'),
 ('transformed', 'VBN'),
 ('modern', 'JJ'),
 ('academic', 'JJ'),
 ('research', 'NN'),
 ('.', '.'),
 ('It', 'PRP'),
 ('enables', 'VBZ'),
 ('efficient', 'JJ'),
 ('data', 'NNS'),
 ('analysis', 'NN'),
 ('and', 'CC'),
 ('supports', 'NNS'),
 ('informed', 'VBN'),
 ('decision', 'NN'),
 ('making', 'NN'),
 ('.', '.'),
 ('Researchers', 'NNP'),
 ('argue', 'VBP'),
 ('that', 'IN'),
 ('AI', 'NNP'),
 ('improves', 'VBZ'),
 ('productivity', 'NN'),
 ('while', 'IN'),
```

```
(('raising', 'VBG'),
 ('ethical', 'JJ'),
 ('concerns', 'NNS'),
 ('.', '.'])
```

```
# Load the small English language model ('en_core_web_sm') from spaCy.
nlp = spacy.load("en_core_web_sm")
# Process the essay text with the loaded spaCy model to create a Doc object.
doc = nlp(essay_text)

# Extract (token text, part-of-speech tag) pairs from the spaCy Doc object.
pos_spacy = [(token.text, token.pos_) for token in doc]
# Display the spaCy part-of-speech tagged list.
pos_spacy
```

```
[('\n', 'SPACE'),
 ('Artificial', 'PROPN'),
 ('Intelligence', 'PROPN'),
 ('has', 'AUX'),
 ('transformed', 'VERB'),
 ('modern', 'ADJ'),
 ('academic', 'ADJ'),
 ('research', 'NOUN'),
 ('.', 'PUNCT'),
 ('\n', 'SPACE'),
 ('It', 'PRON'),
 ('enables', 'VERB'),
 ('efficient', 'ADJ'),
 ('data', 'NOUN'),
 ('analysis', 'NOUN'),
 ('and', 'CCONJ'),
 ('supports', 'VERB'),
 ('informed', 'VERB'),
 ('decision', 'NOUN'),
 ('making', 'NOUN'),
 ('.', 'PUNCT'),
 ('\n', 'SPACE'),
 ('Researchers', 'NOUN'),
 ('argue', 'VERB'),
 ('that', 'SCONJ'),
 ('AI', 'PROPN'),
 ('improves', 'VERB'),
 ('productivity', 'NOUN'),
 ('while', 'SCONJ'),
 ('raising', 'VERB'),
 ('ethical', 'ADJ'),
 ('concerns', 'NOUN'),
 ('.', 'PUNCT'),
 ('\n', 'SPACE')]
```

```
# Filter the NLTK POS-tagged list to extract words that are classified as nouns (tags starting with 'NN').
nouns_nltk = [word for word, tag in pos_nltk if tag.startswith('NN')]
# Filter the NLTK POS-tagged list to extract words that are classified as verbs (tags starting with 'VB').
verbs_nltk = [word for word, tag in pos_nltk if tag.startswith('VB')]
```

```
# Extract tokens from the spaCy Doc object that are classified as nouns (POS tag 'NOUN').
nouns_spacy = [token.text for token in doc if token.pos_ == "NOUN"]
# Extract tokens from the spaCy Doc object that are classified as verbs (POS tag 'VERB').
verbs_spacy = [token.text for token in doc if token.pos_ == "VERB"]
```

```
# Use Counter to calculate the frequency of each noun found by spaCy.
noun_freq = Counter(nouns_spacy)
# Use Counter to calculate the frequency of each verb found by spaCy.
verb_freq = Counter(verbs_spacy)
```

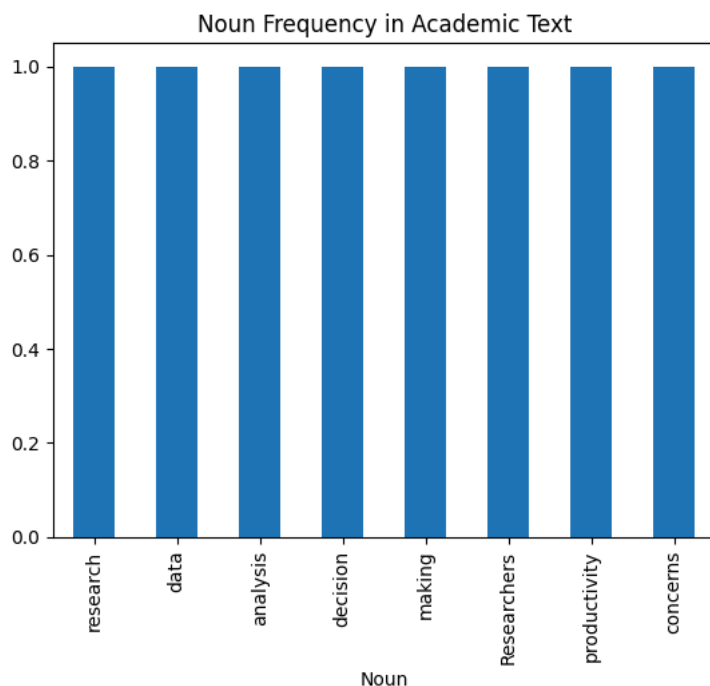
```
# Create a pandas DataFrame from the noun frequencies, with columns 'Noun' and 'Frequency'.
noun_df = pd.DataFrame(noun_freq.items(), columns=["Noun", "Frequency"])
# Create a pandas DataFrame from the verb frequencies, with columns 'Verb' and 'Frequency'.
verb_df = pd.DataFrame(verb_freq.items(), columns=["Verb", "Frequency"])

# Display both the noun frequency DataFrame and the verb frequency DataFrame.
noun_df, verb_df
```

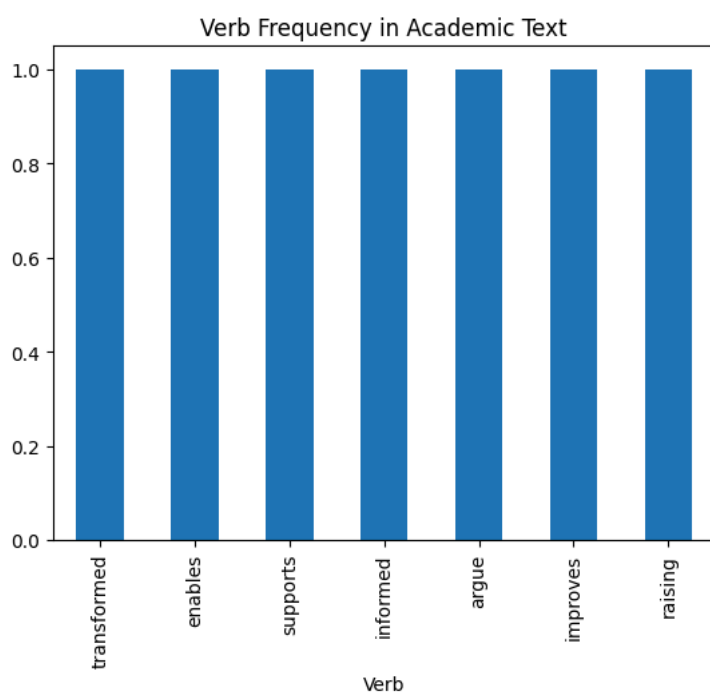
```
(
  Noun  Frequency
0  research      1
1   data      1
2  analysis      1
3  decision      1
4   making      1
5 Researchers      1
6 productivity      1
7   concerns      1,
  Verb  Frequency
```

```
0 transformed      1
1 enables          1
2 supports         1
3 informed         1
4 argue           1
5 improves         1
6 raising         1)
```

```
# Generate a bar plot to visualize the frequency of nouns.
# The x-axis represents the nouns, the y-axis represents their frequency.
# The plot is titled 'Noun Frequency in Academic Text' and the legend is hidden.
noun_df.plot(x="Noun", y="Frequency", kind="bar", title="Noun Frequency in Academic Text", legend=False)
# Display the generated plot.
plt.show()
```



```
# Generate a bar plot to visualize the frequency of verbs.
# The x-axis represents the verbs, the y-axis represents their frequency.
# The plot is titled 'Verb Frequency in Academic Text' and the legend is hidden.
verb_df.plot(x="Verb", y="Frequency", kind="bar", title="Verb Frequency in Academic Text", legend=False)
# Display the generated plot.
plt.show()
```



```
# Install necessary libraries: NLTK for natural language processing and spaCy for advanced NLP.
!pip install nltk spacy
```

```
# Download the English language model for spaCy, which includes resources for tokenization, POS tagging, etc.
!python -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) (0.21.1)
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) (2.10.6)
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)
Requirement already satisfied: pydantic-core==2.41.4 in /usr/local/lib/python3.12/dist-packages (from pydantic!=1.8,!1.8.1)
Requirement already satisfied: typing-extensions>=4.14.1 in /usr/local/lib/python3.12/dist-packages (from pydantic!=1.8,!1.8.1)
Requirement already satisfied: typing-inspection>=0.4.2 in /usr/local/lib/python3.12/dist-packages (from pydantic!=1.8,!1.8.1)
Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/python3.12/dist-packages (from requests<3.0.0,>=2.13.0->spacy)
Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.12/dist-packages (from requests<3.0.0,>=2.13.0->spacy)
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)
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.12/dist-packages (from requests<3.0.0,>=2.13.0->spacy)
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)
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)
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)
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)
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 105.7 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.