```
!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) (2.1
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.
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
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.
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.
Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.12/dist-packages (from requests<3.0.0,>=2.13.0->spac
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.12/dist-packages (from requests<3.0.0,>=2.13.0->spac
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->sp
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->s
Requirement already satisfied: MarkupSafe>=2.0 in /usr/local/lib/python3.12/dist-packages (from jinja2->spacy) (3.0.3)
Requirement already satisfied: wrap in /usr/local/lib/python3.12/dist-packages (from smart-open<8.0.0,>=5.2.1->weasel<0.5.0,>=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
                                        ━━━━━━━━━━━━━━━━━━━ 12.8/12.8 MB 106.7 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.
```

```
import nltk
import spacy
import pandas as pd
from collections import Counter
import matplotlib.pyplot as plt

nltk.download('punkt')
nltk.download('averaged_perceptron_tagger')
```

```
[nltk_data] Downloading package punkt to /root/nltk_data...
[nltk_data]   Unzipping tokenizers/punkt.zip.
[nltk_data] Downloading package averaged_perceptron_tagger to
[nltk_data]     /root/nltk_data...
[nltk_data]   Unzipping taggers/averaged_perceptron_tagger.zip.
True
```

```
essay_text = """
Academic writing plays a crucial role in higher education.
It allows students to present arguments, analyze concepts,
and communicate research findings in a structured manner.
Effective academic texts rely on clarity, precision, and
appropriate use of terminology.
"""
```

```
import nltk
nltk.download('punkt_tab')
from nltk.tokenize import word_tokenize

tokens_nltk = word_tokenize(essay_text)
tokens_nltk
```

```
[nltk_data] Downloading package punkt_tab to /root/nltk_data...
[nltk_data]   Unzipping tokenizers/punkt_tab.zip.
['Academic',
 'writing',
 'plays',
 'a',
 'crucial',
 'role',
 'in',
 'higher',
 'education',
 '.',
 'It',
 'allows',
 'students',
 'to',
 'present',
 'arguments',
 ',',
 'analyze',
 'concepts',
 ',',
 'and',
 'communicate',
 'research',
 'findings',
 'in',
 'a',
 'structured',
 'manner',
 '.',
 'Effective',
 'academic',
 'texts',
 'rely',
 'on',
 'clarity',
 ',',
 'precision',
 ',',
 'and',
 'appropriate',
 'use',
 'of',
 'terminology',
 '.']
```

```
nlp = spacy.load("en_core_web_sm")
doc = nlp(essay_text)

tokens_spacy = [token.text for token in doc]
tokens_spacy
```

```
['\n',
 'Academic',
 'writing',
 'plays',
 'a',
 'crucial',
 'role',
 'in',
 'higher',
 'education',
 '.',
 '\n',
 'It',
 'allows',
 'students',
 'to',
 'present',
 'arguments',
 ',',
 'analyze',
 'concepts',
```

```
 ',',
 '\n',
 'and',
 'communicate',
 'research',
 'findings',
 'in',
 'a',
 'structured',
 'manner',
 '.',
 '\n',
 'Effective',
 'academic',
 'texts',
 'rely',
 'on',
 'clarity',
 ',',
 'precision',
 ',',
 'and',
 '\n',
 'appropriate',
 'use',
 'of',
 'terminology',
 '.',
 '\n']
```

```
nltk.download('averaged_perceptron_tagger_eng')
pos_nltk = nltk.pos_tag(tokens_nltk)
pos_nltk
```

```
[nltk_data] Downloading package averaged_perceptron_tagger_eng to
[nltk_data]     /root/nltk_data...
[nltk_data]   Unzipping taggers/averaged_perceptron_tagger_eng.zip.
[('Academic', 'NNP'),
 ('writing', 'NN'),
 ('plays', 'VBZ'),
 ('a', 'DT'),
 ('crucial', 'JJ'),
 ('role', 'NN'),
 ('in', 'IN'),
 ('higher', 'JJR'),
 ('education', 'NN'),
 ('.', '.'),
 ('It', 'PRP'),
 ('allows', 'VBZ'),
 ('students', 'NNS'),
 ('to', 'TO'),
 ('present', 'JJ'),
 ('arguments', 'NNS'),
 (',', ','),
 ('analyze', 'JJ'),
 ('concepts', 'NNS'),
 (',', ','),
 ('and', 'CC'),
 ('communicate', 'VB'),
 ('research', 'NN'),
 ('findings', 'NNS'),
 ('in', 'IN'),
 ('a', 'DT'),
 ('structured', 'JJ'),
 ('manner', 'NN'),
 ('.', '.'),
 ('Effective', 'JJ'),
 ('academic', 'JJ'),
 ('texts', 'NN'),
 ('rely', 'RB'),
 ('on', 'IN'),
 ('clarity', 'NN'),
 (',', ','),
 ('precision', 'NN'),
 (',', ','),
 ('and', 'CC'),
 ('appropriate', 'JJ'),
 ('use', 'NN'),
 ('of', 'IN'),
 ('terminology', 'NN'),
 ('.', '.')]
```

```python
pos_spacy = [(token.text, token.pos_) for token in doc]
pos_spacy
```

```
[('\n', 'SPACE'),
 ('Academic', 'ADJ'),
 ('writing', 'NOUN'),
 ('plays', 'VERB'),
 ('a', 'DET'),
 ('crucial', 'ADJ'),
 ('role', 'NOUN'),
 ('in', 'ADP'),
 ('higher', 'ADJ'),
 ('education', 'NOUN'),
 ('.', 'PUNCT'),
 ('\n', 'SPACE'),
 ('It', 'PRON'),
 ('allows', 'VERB'),
 ('students', 'NOUN'),
 ('to', 'PART'),
 ('present', 'ADJ'),
 ('arguments', 'NOUN'),
 (',', 'PUNCT'),
 ('analyze', 'VERB'),
 ('concepts', 'NOUN'),
 (',', 'PUNCT'),
 ('\n', 'SPACE'),
 ('and', 'CCONJ'),
 ('communicate', 'VERB'),
 ('research', 'NOUN'),
 ('findings', 'NOUN'),
 ('in', 'ADP'),
 ('a', 'DET'),
 ('structured', 'ADJ'),
 ('manner', 'NOUN'),
 ('.', 'PUNCT'),
 ('\n', 'SPACE'),
 ('Effective', 'ADJ'),
 ('academic', 'ADJ'),
 ('texts', 'NOUN'),
 ('rely', 'VERB'),
 ('on', 'ADP'),
 ('clarity', 'NOUN'),
 (',', 'PUNCT'),
 ('precision', 'NOUN'),
 (',', 'PUNCT'),
 ('and', 'CCONJ'),
 ('\n', 'SPACE'),
 ('appropriate', 'ADJ'),
 ('use', 'NOUN'),
 ('of', 'ADP'),
 ('terminology', 'NOUN'),
 ('.', 'PUNCT'),
 ('\n', 'SPACE')]
```

```python
nouns_nltk = [word for word, tag in pos_nltk if tag.startswith('NN')]
verbs_nltk = [word for word, tag in pos_nltk if tag.startswith('VB')]
```

```python
nouns_spacy = [token.text for token in doc if token.pos_ == "NOUN"]
verbs_spacy = [token.text for token in doc if token.pos_ == "VERB"]
```
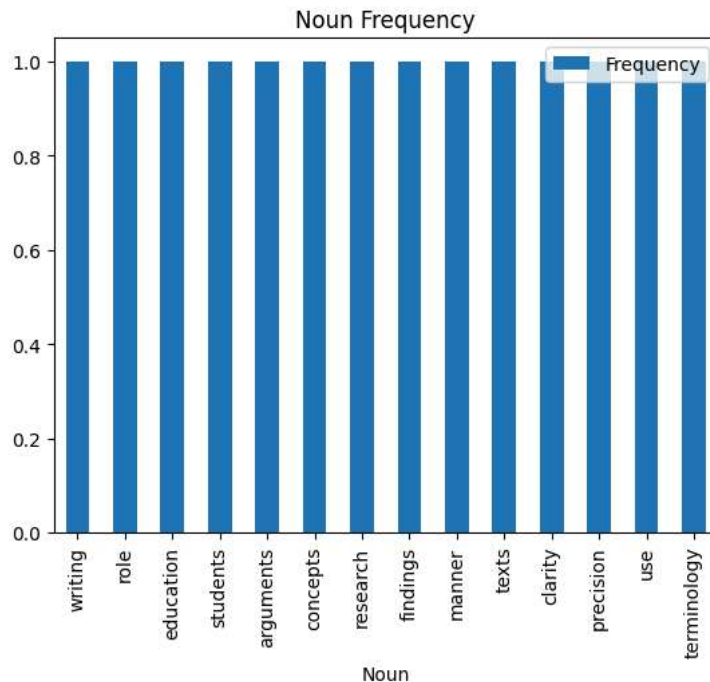
```python
noun_freq = Counter(nouns_spacy)
verb_freq = Counter(verbs_spacy)

noun_freq, verb_freq
```

```
(Counter({'writing': 1,
          'role': 1,
          'education': 1,
          'students': 1,
          'arguments': 1,
          'concepts': 1,
          'research': 1,
          'findings': 1,
          'manner': 1,
          'texts': 1,
          'clarity': 1,
```

```
            'precision': 1,
            'use': 1,
            'terminology': 1}),
 Counter({'plays': 1, 'allows': 1, 'analyze': 1, 'communicate': 1, 'rely': 1}))
```

```python
pd.DataFrame(noun_freq.most_common(), columns=["Noun", "Frequency"]).plot(
    kind="bar", x="Noun", y="Frequency", title="Noun Frequency"
)
plt.show()
```



```python
df = pd.DataFrame(verb_freq.items(), columns=["Verb", "Frequency"])


df.plot(kind="bar", x="Verb", y="Frequency", title="Verb Frequency in Academic Essay")
plt.xlabel("Verbs")
plt.ylabel("Frequency")
plt.show()
```