```
Step 1: Install Necessary Libraries
 In [1]: !pip install -r requirements.txt
       Requirement already satisfied: selenium in c:\users\asus\anaconda3\new folder\envs\20th century\lib\site-packages (from -r requirements.txt (line 1)) (4.25.0)
      Requirement already satisfied: requests in c:\users\asus\anaconda3\new folder\envs\20th_century\lib\site-packages (from -r requirements.txt (line 2)) (2.32.3)
      Requirement already satisfied: bs4 in c:\users\asus\anaconda3\new folder\envs\20th_century\lib\site-packages (from -r requirements.txt (line 3)) (0.0.2)
      Requirement already satisfied: pandas in c:\users\asus\anaconda3\new folder\envs\20th_century\lib\site-packages (from -r requirements.txt (line 4)) (2.2.3)
      Requirement already satisfied: seaborn in c:\users\asus\anaconda3\new folder\envs\20th_century\lib\site-packages (from -r requirements.txt (line 5)) (0.13.2)
      Requirement already satisfied: matplotlib in c:\users\asus\anaconda3\new folder\envs\20th_century\lib\site-packages (from -r requirements.txt (line 6)) (3.9.2)
      Requirement already satisfied: nltk in c:\users\asus\anaconda3\new folder\envs\20th_century\lib\site-packages (from -r requirements.txt (line 7)) (3.9.1)
      Requirement already satisfied: textblob in c:\users\asus\anaconda3\new folder\envs\20th_century\lib\site-packages (from -r requirements.txt (line 8)) (0.18.0.post0)
      Collecting spacy (from -r requirements.txt (line 9))
        Downloading spacy-3.8.2-cp312-cp312-win amd64.whl.metadata (27 kB)
       Requirement already satisfied: urllib3<3,>=1.26 in c:\users\asus\anaconda3\new folder\envs\20th_century\lib\site-packages (from urllib3[socks]<3,>=1.26->selenium->-r requirements.txt (line 1)) (2.2.3)
       Requirement already satisfied: trio~=0.17 in c:\users\asus\anaconda3\new folder\envs\20th_century\lib\site-packages (from selenium->-r requirements.txt (line 1)) (0.27.0)
       Requirement already satisfied: trio-websocket~=0.9 in c:\users\asus\anaconda3\new folder\envs\20th century\lib\site-packages (from selenium->-r requirements.txt (line 1)) (0.11.1)
       Requirement already satisfied: certifi>=2021.10.8 in c:\users\asus\anaconda3\new folder\envs\20th_century\lib\site-packages (from selenium->-r requirements.txt (line 1)) (2024.8.30)
      Requirement already satisfied: typing_extensions~=4.9 in c:\users\asus\anaconda3\new folder\envs\20th_century\lib\site-packages (from selenium->-r requirements.txt (line 1)) (4.12.2)
      Requirement already satisfied: websocket-client~=1.8 in c:\users\asus\anaconda3\new folder\envs\20th_century\lib\site-packages (from selenium->-r requirements.txt (line 1)) (1.8.0)
       Requirement already satisfied: charset-normalizer<4,>=2 in c:\users\asus\anaconda3\new folder\envs\20th_century\lib\site-packages (from requests->-r requirements.txt (line 2)) (3.4.0)
       Requirement already satisfied: idna<4,>=2.5 in c:\users\asus\anaconda3\new folder\envs\20th_century\lib\site-packages (from requests->-r requirements.txt (line 2)) (3.10)
       Requirement already satisfied: beautifulsoup4 in c:\users\asus\anaconda3\new folder\envs\20th_century\lib\site-packages (from bs4->-r requirements.txt (line 3)) (4.12.3)
      Requirement already satisfied: numpy>=1.26.0 in c:\users\asus\anaconda3\new folder\envs\20th_century\lib\site-packages (from pandas->-r requirements.txt (line 4)) (2.1.2)
      Requirement already satisfied: python-dateutil>=2.8.2 in c:\users\asus\anaconda3\new folder\envs\20th_century\lib\site-packages (from pandas->-r requirements.txt (line 4)) (2.9.0.post0)
      Requirement already satisfied: pytz>=2020.1 in c:\users\asus\anaconda3\new folder\envs\20th_century\lib\site-packages (from pandas->-r requirements.txt (line 4)) (2024.2)
      Requirement already satisfied: tzdata>=2022.7 in c:\users\asus\anaconda3\new folder\envs\20th_century\lib\site-packages (from pandas->-r requirements.txt (line 4)) (2024.2)
      Requirement already satisfied: contourpy>=1.0.1 in c:\users\asus\anaconda3\new folder\envs\20th_century\lib\site-packages (from matplotlib->-r requirements.txt (line 6)) (1.3.0)
      Requirement already satisfied: cycler>=0.10 in c:\users\asus\anaconda3\new folder\envs\20th_century\lib\site-packages (from matplotlib->-r requirements.txt (line 6)) (0.12.1)
      Requirement already satisfied: fonttools>=4.22.0 in c:\users\asus\anaconda3\new folder\envs\20th_century\lib\site-packages (from matplotlib->-r requirements.txt (line 6)) (4.54.1)
      Requirement already satisfied: kiwisolver>=1.3.1 in c:\users\asus\anaconda3\new folder\envs\20th_century\lib\site-packages (from matplotlib->-r requirements.txt (line 6)) (1.4.7)
      Requirement already satisfied: packaging>=20.0 in c:\users\asus\anaconda3\new folder\envs\20th_century\lib\site-packages (from matplotlib->-r requirements.txt (line 6)) (24.1)
      Requirement already satisfied: pillow>=8 in c:\users\asus\anaconda3\new folder\envs\20th_century\lib\site-packages (from matplotlib->-r requirements.txt (line 6)) (11.0.0)
      Requirement already satisfied: pyparsing>=2.3.1 in c:\users\asus\anaconda3\new folder\envs\20th_century\lib\site-packages (from matplotlib->-r requirements.txt (line 6)) (3.2.0)
      Requirement already satisfied: click in c:\users\asus\anaconda3\new folder\envs\20th_century\lib\site-packages (from nltk->-r requirements.txt (line 7)) (8.1.7)
      Requirement already satisfied: joblib in c:\users\asus\anaconda3\new folder\envs\20th_century\lib\site-packages (from nltk->-r requirements.txt (line 7)) (1.4.2)
       Requirement already satisfied: regex>=2021.8.3 in c:\users\asus\anaconda3\new folder\envs\20th_century\lib\site-packages (from nltk->-r requirements.txt (line 7)) (2024.9.11)
       Requirement already satisfied: tqdm in c:\users\asus\anaconda3\new folder\envs\20th_century\lib\site-packages (from nltk->-r requirements.txt (line 7)) (4.66.6)
      Collecting spacy-legacy<3.1.0,>=3.0.11 (from spacy->-r requirements.txt (line 9))
        Downloading spacy_legacy-3.0.12-py2.py3-none-any.whl.metadata (2.8 kB)
      Collecting spacy-loggers<2.0.0,>=1.0.0 (from spacy->-r requirements.txt (line 9))
        Downloading spacy_loggers-1.0.5-py3-none-any.whl.metadata (23 kB)
      Collecting murmurhash<1.1.0,>=0.28.0 (from spacy->-r requirements.txt (line 9))
        Downloading murmurhash-1.0.10-cp312-cp312-win_amd64.whl.metadata (2.0 kB)
      Collecting cymem<2.1.0,>=2.0.2 (from spacy->-r requirements.txt (line 9))
        Downloading cymem-2.0.8-cp312-cp312-win_amd64.whl.metadata (8.6 kB)
      Collecting preshed<3.1.0,>=3.0.2 (from spacy->-r requirements.txt (line 9))
        Downloading preshed-3.0.9-cp312-cp312-win_amd64.whl.metadata (2.2 kB)
      Collecting thinc<8.4.0,>=8.3.0 (from spacy->-r requirements.txt (line 9))
        Downloading thinc-8.3.2-cp312-cp312-win_amd64.whl.metadata (15 kB)
      Collecting wasabi<1.2.0,>=0.9.1 (from spacy->-r requirements.txt (line 9))
        Downloading wasabi-1.1.3-py3-none-any.whl.metadata (28 kB)
       Collecting srsly<3.0.0,>=2.4.3 (from spacy->-r requirements.txt (line 9))
        Downloading srsly-2.4.8-cp312-cp312-win_amd64.whl.metadata (20 kB)
      Collecting catalogue<2.1.0,>=2.0.6 (from spacy->-r requirements.txt (line 9))
        Downloading catalogue-2.0.10-py3-none-any.whl.metadata (14 kB)
      Collecting weasel<0.5.0,>=0.1.0 (from spacy->-r requirements.txt (line 9))
        Downloading weasel-0.4.1-py3-none-any.whl.metadata (4.6 kB)
      Collecting typer<1.0.0,>=0.3.0 (from spacy->-r requirements.txt (line 9))
        Downloading typer-0.12.5-py3-none-any.whl.metadata (15 kB)
      Collecting pydantic!=1.8,!=1.8.1,<3.0.0,>=1.7.4 (from spacy->-r requirements.txt (line 9))
        Downloading pydantic-2.9.2-py3-none-any.whl.metadata (149 kB)
       Collecting jinja2 (from spacy->-r requirements.txt (line 9))
        Downloading jinja2-3.1.4-py3-none-any.whl.metadata (2.6 kB)
       Requirement already satisfied: setuptools in c:\users\asus\anaconda3\new folder\envs\20th_century\lib\site-packages (from spacy->-r requirements.txt (line 9)) (75.1.0)
      Collecting langcodes<4.0.0,>=3.2.0 (from spacy->-r requirements.txt (line 9))
        Downloading langcodes-3.4.1-py3-none-any.whl.metadata (29 kB)
      Collecting language-data>=1.2 (from langcodes<4.0.0,>=3.2.0->spacy->-r requirements.txt (line 9))
        Downloading language_data-1.2.0-py3-none-any.whl.metadata (4.3 kB)
      Collecting annotated-types>=0.6.0 (from pydantic!=1.8,!=1.8.1,<3.0.0,>=1.7.4->spacy->-r requirements.txt (line 9))
        Downloading annotated_types-0.7.0-py3-none-any.whl.metadata (15 kB)
      Collecting pydantic-core=2.23.4 (from pydantic!=1.8,!=1.8.1,<3.0.0,>=1.7.4->spacy->-r requirements.txt (line 9))
        Downloading pydantic_core-2.23.4-cp312-none-win_amd64.whl.metadata (6.7 kB)
      Requirement already satisfied: six>=1.5 in c:\users\asus\anaconda3\new folder\envs\20th_century\lib\site-packages (from python-dateutil>=2.8.2->pandas->-r requirements.txt (line 4)) (1.16.0)
      Collecting blis<1.1.0,>=1.0.0 (from thinc<8.4.0,>=8.3.0->spacy->-r requirements.txt (line 9))
        Downloading blis-1.0.1-cp312-cp312-win_amd64.whl.metadata (7.8 kB)
      Collecting confection<1.0.0,>=0.0.1 (from thinc<8.4.0,>=8.3.0->spacy->-r requirements.txt (line 9))
        Downloading confection-0.1.5-py3-none-any.whl.metadata (19 kB)
      Collecting numpy>=1.26.0 (from pandas->-r requirements.txt (line 4))
        Downloading numpy-2.0.2-cp312-cp312-win_amd64.whl.metadata (59 kB)
      Requirement already satisfied: colorama in c:\users\asus\anaconda3\new folder\envs\20th_century\lib\site-packages (from tqdm->nltk->-r requirements.txt (line 7)) (0.4.6)
      Requirement already satisfied: attrs>=23.2.0 in c:\users\asus\anaconda3\new folder\envs\20th_century\lib\site-packages (from trio~=0.17->selenium->-r requirements.txt (line 1)) (24.2.0)
      Requirement already satisfied: sortedcontainers in c:\users\asus\anaconda3\new folder\envs\20th_century\lib\site-packages (from trio~=0.17->selenium->-r requirements.txt (line 1)) (2.4.0)
      Requirement already satisfied: outcome in c:\users\asus\anaconda3\new folder\envs\20th_century\lib\site-packages (from trio~=0.17->selenium->-r requirements.txt (line 1)) (1.3.0.post0)
       Requirement already satisfied: sniffio>=1.3.0 in c:\users\asus\anaconda3\new folder\envs\20th_century\lib\site-packages (from trio~=0.17->selenium->-r requirements.txt (line 1)) (1.3.1)
       Requirement already satisfied: cffi>=1.14 in c:\users\asus\anaconda3\new folder\envs\20th_century\lib\site-packages (from trio~=0.17->selenium->-r requirements.txt (line 1)) (1.17.1)
      Requirement already satisfied: wsproto>=0.14 in c:\users\asus\anaconda3\new folder\envs\20th_century\lib\site-packages (from trio-websocket~=0.9->selenium->-r requirements.txt (line 1)) (1.2.0)
      Collecting shellingham>=1.3.0 (from typer<1.0.0,>=0.3.0->spacy->-r requirements.txt (line 9))
        Downloading shellingham-1.5.4-py2.py3-none-any.whl.metadata (3.5 kB)
      Collecting rich>=10.11.0 (from typer<1.0.0,>=0.3.0->spacy->-r requirements.txt (line 9))
        Downloading rich-13.9.3-py3-none-any.whl.metadata (18 kB)
       Requirement already satisfied: pysocks!=1.5.7,<2.0,>=1.5.6 in c:\users\asus\anaconda3\new folder\envs\20th_century\lib\site-packages (from urllib3[socks]<3,>=1.26->selenium->-r requirements.txt (line 1)) (1.7.1)
      Collecting cloudpathlib<1.0.0,>=0.7.0 (from weasel<0.5.0,>=0.1.0->spacy->-r requirements.txt (line 9))
        Downloading cloudpathlib-0.20.0-py3-none-any.whl.metadata (14 kB)
      Collecting smart-open < 8.0.0, >=5.2.1 (from weasel < 0.5.0, >=0.1.0 -> spacy -> -r requirements.txt (line 9))
        Downloading smart_open-7.0.5-py3-none-any.whl.metadata (24 kB)
       Requirement already satisfied: soupsieve>1.2 in c:\users\asus\anaconda3\new folder\envs\20th_century\lib\site-packages (from beautifulsoup4->bs4->-r requirements.txt (line 3)) (2.6)
      Collecting MarkupSafe>=2.0 (from jinja2->spacy->-r requirements.txt (line 9))
        Downloading MarkupSafe-3.0.2-cp312-cp312-win_amd64.whl.metadata (4.1 kB)
      Requirement already satisfied: pycparser in c:\users\asus\anaconda3\new folder\envs\20th_century\lib\site-packages (from cffi>=1.14->trio~=0.17->selenium->-r requirements.txt (line 1)) (2.22)
      Collecting marisa-trie>=0.7.7 (from language-data>=1.2->langcodes<4.0.0,>=3.2.0->spacy->-r requirements.txt (line 9))
        Downloading marisa_trie-1.2.1-cp312-cp312-win_amd64.whl.metadata (9.3 kB)
      Collecting markdown-it-py>=2.2.0 (from rich>=10.11.0->typer<1.0.0,>=0.3.0->spacy->-r requirements.txt (line 9))
        Downloading markdown_it_py-3.0.0-py3-none-any.whl.metadata (6.9 kB)
      Requirement already satisfied: pygments<3.0.0,>=2.13.0 in c:\users\asus\anaconda3\new folder\envs\20th_century\lib\site-packages (from rich>=10.11.0->typer<1.0.0,>=0.3.0->spacy->-r requirements.txt (line 9)) (2.15.1)
      Collecting wrapt (from smart-open<8.0.0,>=5.2.1->weasel<0.5.0,>=0.1.0->spacy->-r requirements.txt (line 9))
        Downloading wrapt-1.16.0-cp312-cp312-win_amd64.whl.metadata (6.8 kB)
      Requirement already satisfied: h11<1,>=0.9.0 in c:\users\asus\anaconda3\new folder\envs\20th_century\lib\site-packages (from wsproto>=0.14->trio-websocket~=0.9->selenium->-r requirements.txt (line 1)) (0.14.0)
       \texttt{Collecting mdurl} \sim \texttt{0.1} \ (\texttt{from markdown-it-py} > \texttt{2.2.0-} \times \texttt{rich} > \texttt{10.11.0-} \times \texttt{typer} < \texttt{1.0.0,} > \texttt{0.3.0-} \times \texttt{spacy} - \texttt{r requirements.txt} \ (\texttt{line 9}) ) 
        Downloading mdurl-0.1.2-py3-none-any.whl.metadata (1.6 kB)
      Downloading spacy-3.8.2-cp312-cp312-win_amd64.whl (11.8 MB)
         ----- 0.0/11.8 MB ? eta -:--:-
          ----- 0.3/11.8 MB ? eta -:--:-
         - ----- 0.5/11.8 MB 1.7 MB/s eta 0:00:07
         -- ----- 0.8/11.8 MB 1.3 MB/s eta 0:00:09
         --- 1.0/11.8 MB 1.6 MB/s eta 0:00:07
         ---- 1.3/11.8 MB 1.5 MB/s eta 0:00:07
         ----- 1.6/11.8 MB 1.5 MB/s eta 0:00:07
         ----- 2.1/11.8 MB 1.5 MB/s eta 0:00:07
         ----- 2.1/11.8 MB 1.5 MB/s eta 0:00:07
         ------ 2.4/11.8 MB 1.3 MB/s eta 0:00:08
         ----- 2.6/11.8 MB 1.3 MB/s eta 0:00:08
         ----- 2.9/11.8 MB 1.3 MB/s eta 0:00:07
         ----- 3.4/11.8 MB 1.4 MB/s eta 0:00:07
         ----- 3.7/11.8 MB 1.4 MB/s eta 0:00:06
         ----- 3.9/11.8 MB 1.4 MB/s eta 0:00:06
         ----- 4.2/11.8 MB 1.4 MB/s eta 0:00:06
         ----- 4.7/11.8 MB 1.4 MB/s eta 0:00:05
         ----- 5.0/11.8 MB 1.4 MB/s eta 0:00:05
         ----- 5.5/11.8 MB 1.5 MB/s eta 0:00:05
         ----- 5.8/11.8 MB 1.5 MB/s eta 0:00:05
         ----- 6.3/11.8 MB 1.5 MB/s eta 0:00:04
         ----- 6.6/11.8 MB 1.5 MB/s eta 0:00:04
         ----- 6.8/11.8 MB 1.5 MB/s eta 0:00:04
         ----- 7.3/11.8 MB 1.5 MB/s eta 0:00:03
         ----- 8.1/11.8 MB 1.6 MB/s eta 0:00:03
         ----- 8.4/11.8 MB 1.6 MB/s eta 0:00:03
         ------ 8.9/11.8 MB 1.6 MB/s eta 0:00:02
         ----- 9.2/11.8 MB 1.6 MB/s eta 0:00:02
         ----- 9.4/11.8 MB 1.6 MB/s eta 0:00:02
         ----- 10.0/11.8 MB 1.6 MB/s eta 0:00:02
         ----- 10.2/11.8 MB 1.6 MB/s eta 0:00:01
         ----- 10.5/11.8 MB 1.6 MB/s eta 0:00:01
         ----- -- 11.0/11.8 MB 1.6 MB/s eta 0:00:01
         ------ - 11.3/11.8 MB 1.6 MB/s eta 0:00:01
         ----- 11.5/11.8 MB 1.6 MB/s eta 0:00:01
         ----- 11.8/11.8 MB 1.6 MB/s eta 0:00:00
      Downloading catalogue-2.0.10-py3-none-any.whl (17 kB)
      Downloading cymem-2.0.8-cp312-cp312-win_amd64.whl (39 kB)
      Downloading langcodes-3.4.1-py3-none-any.whl (182 kB)
      Downloading murmurhash-1.0.10-cp312-cp312-win_amd64.whl (25 kB)
      Downloading preshed-3.0.9-cp312-cp312-win amd64.whl (122 kB)
      Downloading pydantic-2.9.2-py3-none-any.whl (434 kB)
      Downloading pydantic_core-2.23.4-cp312-none-win_amd64.whl (1.9 MB)
         ----- 0.0/1.9 MB ? eta -:--:-
         ---- 0.3/1.9 MB ? eta -:--:-
         ----- 0.5/1.9 MB 1.9 MB/s eta 0:00:01
         ----- 1.0/1.9 MB 1.8 MB/s eta 0:00:01
         ----- 1.3/1.9 MB 1.8 MB/s eta 0:00:01
         ----- 1.8/1.9 MB 1.9 MB/s eta 0:00:01
         ----- 1.9/1.9 MB 1.8 MB/s eta 0:00:00
      Downloading spacy_legacy-3.0.12-py2.py3-none-any.whl (29 kB)
      Downloading spacy_loggers-1.0.5-py3-none-any.whl (22 kB)
      Downloading srsly-2.4.8-cp312-cp312-win_amd64.whl (478 kB)
      Downloading thinc-8.3.2-cp312-cp312-win amd64.whl (1.5 MB)
         ----- 0.0/1.5 MB ? eta -:--:-
         ----- 0.3/1.5 MB ? eta -:--:-
         ----- 0.8/1.5 MB 2.0 MB/s eta 0:00:01
         ----- 1.0/1.5 MB 2.0 MB/s eta 0:00:01
         ----- 1.5/1.5 MB 2.0 MB/s eta 0:00:00
      Downloading numpy-2.0.2-cp312-cp312-win_amd64.whl (15.6 MB)
         ----- 0.0/15.6 MB ? eta -:--:-
         ----- 0.3/15.6 MB ? eta -:--:-
         - ----- 0.5/15.6 MB 1.9 MB/s eta 0:00:09
         - ----- 0.5/15.6 MB 1.9 MB/s eta 0:00:09
         -- ----- 1.0/15.6 MB 1.2 MB/s eta 0:00:13
         --- 1.3/15.6 MB 1.2 MB/s eta 0:00:12
         ---- 1.8/15.6 MB 1.4 MB/s eta 0:00:10
         ----- 2.1/15.6 MB 1.5 MB/s eta 0:00:10
         ----- 2.6/15.6 MB 1.6 MB/s eta 0:00:09
         ----- 3.1/15.6 MB 1.7 MB/s eta 0:00:08
         ----- 3.1/15.6 MB 1.7 MB/s eta 0:00:08
         ----- 3.4/15.6 MB 1.5 MB/s eta 0:00:08
         ----- 3.9/15.6 MB 1.6 MB/s eta 0:00:08
         ----- 3.9/15.6 MB 1.6 MB/s eta 0:00:08
         ----- 4.5/15.6 MB 1.5 MB/s eta 0:00:08
         ----- 4.7/15.6 MB 1.6 MB/s eta 0:00:07
         ----- 5.2/15.6 MB 1.6 MB/s eta 0:00:07
         ----- 5.8/15.6 MB 1.6 MB/s eta 0:00:07
         ----- 6.0/15.6 MB 1.6 MB/s eta 0:00:06
         ----- 6.6/15.6 MB 1.7 MB/s eta 0:00:06
         ----- 7.1/15.6 MB 1.7 MB/s eta 0:00:05
         ----- 7.3/15.6 MB 1.7 MB/s eta 0:00:05
         ----- 7.3/15.6 MB 1.7 MB/s eta 0:00:05
         ----- 7.9/15.6 MB 1.6 MB/s eta 0:00:05
         ----- 7.9/15.6 MB 1.6 MB/s eta 0:00:05
         ----- 8.4/15.6 MB 1.6 MB/s eta 0:00:05
         ----- 8.9/15.6 MB 1.7 MB/s eta 0:00:05
         ----- 9.7/15.6 MB 1.7 MB/s eta 0:00:04
         ------ 10.0/15.6 MB 1.7 MB/s eta 0:00:04
         ------ 10.0/15.6 MB 1.7 MB/s eta 0:00:04
         ------ 10.0/15.6 MB 1.7 MB/s eta 0:00:04
         ------ 10.2/15.6 MB 1.6 MB/s eta 0:00:04
         ----- 10.5/15.6 MB 1.6 MB/s eta 0:00:04
           ----- 11.0/15.6 MB 1.6 MB/s eta 0:00:03
         ----- 11.5/15.6 MB 1.6 MB/s eta 0:00:03
         ----- 12.3/15.6 MB 1.7 MB/s eta 0:00:02
         ----- 12.8/15.6 MB 1.7 MB/s eta 0:00:02
         ----- 13.6/15.6 MB 1.8 MB/s eta 0:00:02
         ----- 14.2/15.6 MB 1.8 MB/s eta 0:00:01
         ----- 14.4/15.6 MB 1.8 MB/s eta 0:00:01
         ----- -- 14.7/15.6 MB 1.8 MB/s eta 0:00:01
         ----- - 15.2/15.6 MB 1.8 MB/s eta 0:00:01
         ----- 15.6/15.6 MB 1.8 MB/s eta 0:00:00
      Downloading typer-0.12.5-py3-none-any.whl (47 kB)
      Downloading wasabi-1.1.3-py3-none-any.whl (27 kB)
      Downloading weasel-0.4.1-py3-none-any.whl (50 kB)
      Downloading jinja2-3.1.4-py3-none-any.whl (133 kB)
      Downloading annotated_types-0.7.0-py3-none-any.whl (13 kB)
      Downloading blis-1.0.1-cp312-cp312-win_amd64.whl (6.4 MB)
         ----- 0.0/6.4 MB ? eta -:--:-
         - ----- 0.3/6.4 MB ? eta -:--:-
         ---- 0.8/6.4 MB 3.7 MB/s eta 0:00:02
         ----- 1.3/6.4 MB 2.7 MB/s eta 0:00:02
         ----- 1.8/6.4 MB 2.5 MB/s eta 0:00:02
         ------ 1.8/6.4 MB 2.5 MB/s eta 0:00:02
         ----- 2.6/6.4 MB 2.3 MB/s eta 0:00:02
         ----- 2.9/6.4 MB 2.3 MB/s eta 0:00:02
         ----- 3.1/6.4 MB 2.2 MB/s eta 0:00:02
         ----- 3.7/6.4 MB 2.1 MB/s eta 0:00:02
         ----- 3.9/6.4 MB 2.1 MB/s eta 0:00:02
         ----- 4.2/6.4 MB 1.9 MB/s eta 0:00:02
         ----- 4.5/6.4 MB 1.9 MB/s eta 0:00:01
         ----- 4.7/6.4 MB 1.9 MB/s eta 0:00:01
         ----- 5.0/6.4 MB 1.9 MB/s eta 0:00:01
         ----- 5.5/6.4 MB 1.8 MB/s eta 0:00:01
         ----- 5.8/6.4 MB 1.8 MB/s eta 0:00:01
         ----- -- 6.0/6.4 MB 1.8 MB/s eta 0:00:01
         ----- 6.4/6.4 MB 1.7 MB/s eta 0:00:00
      Downloading cloudpathlib-0.20.0-py3-none-any.whl (52 kB)
      Downloading confection-0.1.5-py3-none-any.whl (35 kB)
      Downloading language_data-1.2.0-py3-none-any.whl (5.4 MB)
         ----- 0.0/5.4 MB ? eta -:--:-
         - ----- 0.3/5.4 MB ? eta -:--:-
         --- 0.5/5.4 MB 1.5 MB/s eta 0:00:04
         ---- 0.8/5.4 MB 1.5 MB/s eta 0:00:04
         ----- 1.3/5.4 MB 1.6 MB/s eta 0:00:03
         ----- 1.6/5.4 MB 1.6 MB/s eta 0:00:03
         ----- 1.8/5.4 MB 1.6 MB/s eta 0:00:03
         ----- 2.1/5.4 MB 1.6 MB/s eta 0:00:03
         ----- 2.4/5.4 MB 1.6 MB/s eta 0:00:02
         ----- 2.9/5.4 MB 1.6 MB/s eta 0:00:02
         ----- 3.1/5.4 MB 1.6 MB/s eta 0:00:02
         ----- 3.4/5.4 MB 1.6 MB/s eta 0:00:02
         ----- 3.9/5.4 MB 1.6 MB/s eta 0:00:01
         ----- 4.2/5.4 MB 1.6 MB/s eta 0:00:01
         ----- 4.5/5.4 MB 1.6 MB/s eta 0:00:01
         ----- 5.0/5.4 MB 1.6 MB/s eta 0:00:01
         ----- - 5.2/5.4 MB 1.6 MB/s eta 0:00:01
         ----- 5.4/5.4 MB 1.6 MB/s eta 0:00:00
      Downloading MarkupSafe-3.0.2-cp312-cp312-win_amd64.whl (15 kB)
      Downloading rich-13.9.3-py3-none-any.whl (242 kB)
      Downloading shellingham-1.5.4-py2.py3-none-any.whl (9.8 kB)
      Downloading smart_open-7.0.5-py3-none-any.whl (61 kB)
      Downloading marisa_trie-1.2.1-cp312-cp312-win_amd64.whl (150 kB)
      Downloading markdown_it_py-3.0.0-py3-none-any.whl (87 kB)
      Downloading wrapt-1.16.0-cp312-cp312-win_amd64.whl (37 kB)
      Downloading mdurl-0.1.2-py3-none-any.whl (10.0 kB)
      Installing collected packages: cymem, wrapt, wasabi, spacy-loggers, spacy-legacy, shellingham, pydantic-core, numpy, murmurhash, mdurl, MarkupSafe, marisa-trie, cloudpathlib, catalogue, annotated-types, srsly, smart-open, pydantic, pre
      shed, markdown-it-py, language-data, jinja2, blis, rich, langcodes, confection, typer, thinc, weasel, spacy
       Attempting uninstall: numpy
         Found existing installation: numpy 2.1.2
         Uninstalling numpy-2.1.2:
           Successfully uninstalled numpy-2.1.2
      Successfully installed MarkupSafe-3.0.2 annotated-types-0.7.0 blis-1.0.1 catalogue-2.0.10 cloudpathlib-0.20.0 confection-0.1.5 cymem-2.0.8 jinja2-3.1.4 langcodes-3.4.1 language-data-1.2.0 marisa-trie-1.2.1 markdown-it-py-3.0.0 mdurl-0.
      1.2 murmurhash-1.0.10 numpy-2.0.2 preshed-3.0.9 pydantic-core-2.23.4 rich-13.9.3 shellingham-1.5.4 smart-open-7.0.5 spacy-legacy-3.0.12 spacy-legacy-3.0.12 spacy-legacy-3.0.5 srsly-2.4.8 thinc-8.3.2 typer-0.12.5 wasabi-1.1.3 we
      asel-0.4.1 wrapt-1.16.0
       Step 2: Import Libraries
 In [2]: import pandas as pd
       import numpy as np
       import matplotlib.pyplot as plt
       import seaborn as sns
       import spacy
       import datetime as dt
       import os
       import glob
       import re
       Step 3: Load the Text File
 In [4]: # Load the text file
       with open('key_events_20th_century.txt', 'r') as file:
          text = file.read()
       Step 4: Evaluate and Clean the Text
 In [5]: print(text[:1000]) # Display the first 1000 characters for a quick check
      Jump to content
      Main menu
      Search
      Donate
      Appearance
      Create account
      Log in
      Personal tools
      Key Events of the 20th Century
      Add languages
      Article
      Talk
      From Wikipedia, the free encyclopedia
      Look for Key Events of the 20th Century on one of Wikipedia's sister projects:
      Wiktionary (dictionary)
      Wikibooks (textbooks)
      Wikiquote (quotations)
      Wikisource (library)
      Wikiversity (learning resources)
      Commons (media)
      Wikivoyage (travel guide)
      Wikinews (news source)
      Wikidata (linked database)
      Wikispecies (species directory)
      Wikipedia does not have an article with this exact name. Please search for Key Events of the 20th Century in Wikipedia to check for alternative titles or spellings.
      You need to log in or create an account and be autoconfirmed to create new articles. Alternatively, you can use the article wizard to submit a draft for review, or request a new article.
      Search for "Key Events of the 20th Century" in existing articles.
      Look for pages within Wikipedia that link to this title.
      Other rea
 In [6]: # Check Country Names:
       country_list = ["United States", "Canada", "United Kingdom", "France", "Germany"]
 In [7]: #Text Cleaning:
       # Example: Replace unusual characters
       text = text.replace("-", "-").replace("'", "'").replace("'", "'")
 In [8]: # Save the cleaned text (if modified):
       with open('twentieth-century-cleaned.txt', 'w') as file:
          file.write(text)
       Step 5: Create a Named Entity Recognition (NER) Object
In [10]: !python -m spacy download en_core_web_sm
      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 MB)
          ----- 0.0/12.8 MB ? eta -:--:-
            ----- 0.3/12.8 MB ? eta -:--:-
          - ----- 0.5/12.8 MB 1.9 MB/s eta 0:00:07
          --- 1.0/12.8 MB 1.9 MB/s eta 0:00:07
          ---- 1.3/12.8 MB 1.8 MB/s eta 0:00:07
           ----- 1.8/12.8 MB 1.8 MB/s eta 0:00:07
          ----- 2.1/12.8 MB 1.8 MB/s eta 0:00:07
          ----- 2.4/12.8 MB 1.8 MB/s eta 0:00:06
           ----- 2.9/12.8 MB 1.7 MB/s eta 0:00:06
           ----- 3.1/12.8 MB 1.8 MB/s eta 0:00:06
           ----- 3.4/12.8 MB 1.7 MB/s eta 0:00:06
           ----- 3.7/12.8 MB 1.7 MB/s eta 0:00:06
           ----- 3.9/12.8 MB 1.7 MB/s eta 0:00:06
           ----- 4.5/12.8 MB 1.7 MB/s eta 0:00:05
          ----- 4.7/12.8 MB 1.7 MB/s eta 0:00:05
          ----- 5.0/12.8 MB 1.6 MB/s eta 0:00:05
           ----- 5.2/12.8 MB 1.6 MB/s eta 0:00:05
          ----- 5.5/12.8 MB 1.6 MB/s eta 0:00:05
          ----- 6.0/12.8 MB 1.6 MB/s eta 0:00:05
          ----- 6.3/12.8 MB 1.6 MB/s eta 0:00:05
          ----- 6.6/12.8 MB 1.6 MB/s eta 0:00:04
          ----- 6.8/12.8 MB 1.5 MB/s eta 0:00:04
          ----- 7.1/12.8 MB 1.5 MB/s eta 0:00:04
          ----- 7.1/12.8 MB 1.5 MB/s eta 0:00:04
           ----- 7.3/12.8 MB 1.5 MB/s eta 0:00:04
           ----- 7.6/12.8 MB 1.5 MB/s eta 0:00:04
           ----- 7.9/12.8 MB 1.5 MB/s eta 0:00:04
           ------ 8.1/12.8 MB 1.4 MB/s eta 0:00:04
           ----- 8.1/12.8 MB 1.4 MB/s eta 0:00:04
           ----- 8.4/12.8 MB 1.4 MB/s eta 0:00:04
           ------ 8.7/12.8 MB 1.4 MB/s eta 0:00:03
           ------ 8.9/12.8 MB 1.4 MB/s eta 0:00:03
           ----- 9.2/12.8 MB 1.4 MB/s eta 0:00:03
           ----- 9.2/12.8 MB 1.4 MB/s eta 0:00:03
          ----- 9.4/12.8 MB 1.3 MB/s eta 0:00:03
          ----- 9.7/12.8 MB 1.3 MB/s eta 0:00:03
           ----- 10.0/12.8 MB 1.3 MB/s eta 0:00:03
          ----- 10.2/12.8 MB 1.3 MB/s eta 0:00:02
          ----- 10.5/12.8 MB 1.3 MB/s eta 0:00:02
          ----- 10.7/12.8 MB 1.3 MB/s eta 0:00:02
           ----- 10.7/12.8 MB 1.3 MB/s eta 0:00:02
           ----- 11.0/12.8 MB 1.3 MB/s eta 0:00:02
          ----- 11.3/12.8 MB 1.3 MB/s eta 0:00:02
          ----- 11.3/12.8 MB 1.3 MB/s eta 0:00:02
           ----- 11.3/12.8 MB 1.3 MB/s eta 0:00:02
          ----- 11.5/12.8 MB 1.2 MB/s eta 0:00:02
          ----- 11.5/12.8 MB 1.2 MB/s eta 0:00:02
          ----- 11.8/12.8 MB 1.2 MB/s eta 0:00:01
          ----- 11.8/12.8 MB 1.2 MB/s eta 0:00:01
           ----- -- 12.1/12.8 MB 1.2 MB/s eta 0:00:01
           ----- -- 12.1/12.8 MB 1.2 MB/s eta 0:00:01
           ------ - 12.3/12.8 MB 1.2 MB/s eta 0:00:01
           ----- - 12.3/12.8 MB 1.2 MB/s eta 0:00:01
          ----- 12.6/12.8 MB 1.1 MB/s eta 0:00:01
           ------ 12.8/12.8 MB 1.1 MB/s eta 0:00:00
      Installing collected packages: en-core-web-sm
      Successfully installed en-core-web-sm-3.8.0
      [+] Download and installation successful
      You can now load the package via spacy.load('en_core_web_sm')
In [11]: # Load SpaCy's English NER model
       nlp = spacy.load("en_core_web_sm")
       # Create the NER object
       doc = nlp(text)
       Step 6: Split Sentence Entities from the NER Object
In [12]: # List to store sentences and entities
       sentences_entities = []
       # Iterate through sentences in the NER document
       for sentence in doc.sents:
          # Extract entities in each sentence
          entities = [(ent.text, ent.label_) for ent in sentence.ents]
          sentences_entities.append((sentence.text, entities))
       # Display the first few sentences with their entities
       sentences_entities[:5]
Out[12]: [("Jump to content\nMain menu\nSearch\nDonate\nAppearance\nCreate account\nLog in\nPersonal tools\nKey Events of the 20th Century\nAdd languages\nArticle\nTalk\nTools\nFrom Wikipedia, the free encyclopedia\nLook for Key Events of the
       20th Century on one of Wikipedia's sister projects:\n",
         [('the 20th Century\nAdd', 'DATE'),
          ('Wikipedia', 'ORG'),
          ('Look for Key Events of', 'WORK_OF_ART'),
          ('the 20th Century', 'DATE'),
          ('one', 'CARDINAL'),
          ('Wikipedia', 'ORG')]),
         ('Wiktionary (dictionary)\nWikibooks (textbooks)\nWikiquote (quotations)\nWikivoyage (travel guide)\nWikinews (news source)\nWikidata (linked database)\nWikisp
        ecies (species directory) \nWikipedia does not have an article with this exact name.',
         [('Wikivoyage', 'PERSON'), ('Wikipedia', 'ORG')]),
         ('Please search for Key Events of the 20th Century in Wikipedia to check for alternative titles or spellings.\n',
         [('Key Events', 'ORG'), ('the 20th Century', 'DATE'), ('Wikipedia', 'GPE')]),
         ('You need to log in or create an account and be autoconfirmed to create new articles.',
         ('Alternatively, you can use the article wizard to submit a draft for review, or request a new article.\n',
       Step 7: Filter Entities for Country Names
In [13]: # Filter sentences to keep only entities that are in the country list
       filtered_entities = [
           (sentence, [(entity, label) for entity, label in entities if entity in country_list and label == 'GPE'])
          for sentence, entities in sentences_entities
       # Remove sentences with no relevant entities
       filtered_entities = [(sentence, entities) for sentence, entities in filtered_entities if entities]
       # Display filtered results
       filtered_entities[:5]
Out[13]: []
       Step 8: Create the Relationships DataFrame
In [14]: # Create DataFrame for relationships
       relationships = []
```

Process the filtered sentences and entities
for sentence, entities in filtered_entities:

relationships_df = pd.DataFrame(relationships)

Convert to DataFrame

In [15]: # Save DataFrame to CSV

Out[14]: —

Show the first few rows
relationships_df.head()

for i in range(len(entities) - 1):

relationships.append({

for j in range(i + 1, len(entities)):

'Sentence': sentence,

'Country1': entities[i][0],
'Country2': entities[j][0]

Step 9: Save and Export the DataFrame

relationships_df.to_csv('country_relationships.csv', index=False)

if len(entities) > 1: # Ensure there are multiple countries in the sentence