

```
# Extracting news headline and date from cnbc.com
# I am using selenium instead of BeautifulSoup because the website is loading content dynamically using JavaScript
```

```
!pip install selenium
```

```
Collecting selenium
  Downloading selenium-4.17.2-py3-none-any.whl (9.9 MB)
    9.9/9.9 MB 22.6 MB/s eta 0:00:00
Requirement already satisfied: urllib3[socks]<3,=>1.26 in /usr/local/lib/python3.10/dist-packages (from selenium) (2.0.7)
Collecting trio~=0.17 (from selenium)
  Downloading trio-0.24.0-py3-none-any.whl (460 kB)
    460.2/460.2 kB 34.9 MB/s eta 0:00:00
Collecting trio-websocket~=0.9 (from selenium)
  Downloading trio_websocket-0.11.1-py3-none-any.whl (17 kB)
Requirement already satisfied: certifi>=2021.10.8 in /usr/local/lib/python3.10/dist-packages (from selenium) (2024.2.2)
Requirement already satisfied: typing_extensions>=4.9.0 in /usr/local/lib/python3.10/dist-packages (from selenium) (4.9.0)
Requirement already satisfied: attrs>=20.1.0 in /usr/local/lib/python3.10/dist-packages (from trio~=0.17->selenium) (23.2.0)
Requirement already satisfied: sortedcontainers in /usr/local/lib/python3.10/dist-packages (from trio~=0.17->selenium) (2.4.0)
Requirement already satisfied: idna in /usr/local/lib/python3.10/dist-packages (from trio~=0.17->selenium) (3.6)
Collecting outcome (from trio~=0.17->selenium)
  Downloading outcome-1.3.0.post0-py2.py3-none-any.whl (10 kB)
Requirement already satisfied: sniffio>=1.3.0 in /usr/local/lib/python3.10/dist-packages (from trio~=0.17->selenium) (1.3.0)
Requirement already satisfied: exceptiongroup in /usr/local/lib/python3.10/dist-packages (from trio~=0.17->selenium) (1.2.0)
Collecting wsproto>=0.14 (from trio-websocket~=0.9->selenium)
  Downloading wsproto-1.2.0-py3-none-any.whl (24 kB)
Requirement already satisfied: pysocks!=1.5.7,<2.0,>=1.5.6 in /usr/local/lib/python3.10/dist-packages (from urllib3[socks]<3,=>1.26->selenium) (1.7.1)
Collecting h11<1,>=0.9.0 (from wsproto>=0.14->trio-websocket~=0.9->selenium)
  Downloading h11-0.14.0-py3-none-any.whl (58 kB)
    58.3/58.3 kB 5.9 MB/s eta 0:00:00
Installing collected packages: outcome, h11, wsproto, trio, trio-websocket, selenium
Successfully installed h11-0.14.0 outcome-1.3.0.post0 selenium-4.17.2 trio-0.24.0 trio-websocket-0.11.1 wsproto-1.2.0
```

```
!apt-get update
!apt-get install -y chromium-chromedriver
!cp /usr/lib/chromium-browser/chromedriver /usr/bin
```

```
Get:1 http://security.ubuntu.com/ubuntu jammy-security InRelease [110 kB]
Get:2 https://developer.download.nvidia.com/compute/cuda/repos/ubuntu2204/x86\_64 InRelease [1,581 B]
Hit:3 http://archive.ubuntu.com/ubuntu jammy InRelease
Get:4 https://cloud.r-project.org/bin/linux/ubuntu jammy-cran40/ InRelease [3,626 B]
Get:5 http://archive.ubuntu.com/ubuntu jammy-updates InRelease [119 kB]
Hit:6 http://archive.ubuntu.com/ubuntu jammy-backports InRelease
Hit:7 https://ppa.launchpadcontent.net/c2d4u.team/c2d4u4.0+/ubuntu jammy InRelease
Get:8 https://developer.download.nvidia.com/compute/cuda/repos/ubuntu2204/x86\_64 Packages [671 kB]
Hit:9 https://ppa.launchpadcontent.net/deadsnakes/ppa/ubuntu jammy InRelease
Hit:10 https://ppa.launchpadcontent.net/graphics-drivers/ppa/ubuntu jammy InRelease
Hit:11 https://ppa.launchpadcontent.net/ubuntugis/ppa/ubuntu jammy InRelease
Get:12 http://security.ubuntu.com/ubuntu jammy-security/restricted amd64 Packages [1,746 kB]
Get:13 http://archive.ubuntu.com/ubuntu jammy-updates/main amd64 Packages [1,723 kB]
Get:14 http://security.ubuntu.com/ubuntu jammy-security/main amd64 Packages [1,436 kB]
Get:15 http://archive.ubuntu.com/ubuntu jammy-updates/universe amd64 Packages [1,334 kB]
Get:16 http://archive.ubuntu.com/ubuntu jammy-updates/restricted amd64 Packages [1,800 kB]
Fetched 8,944 kB in 5s (1,931 kB/s)
Reading package lists... Done
Building dependency tree... Done
Reading state information... Done
The following additional packages will be installed:
  apparmor chromium-browser libfuse3-3 liblzo2-2 libudev1 snapd squashfs-tools systemd-hwe-hwdb
  udev
Suggested packages:
  apparmor-profiles-extra apparmor-utils fuse3 zenity | kdialog
The following NEW packages will be installed:
  apparmor chromium-browser chromium-chromedriver libfuse3-3 liblzo2-2 snapd squashfs-tools
  systemd-hwe-hwdb udev
The following packages will be upgraded:
  libudev1
1 upgraded, 9 newly installed, 0 to remove and 32 not upgraded.
Need to get 26.4 MB of archives.
After this operation, 116 MB of additional disk space will be used.
Get:1 http://archive.ubuntu.com/ubuntu jammy-updates/main amd64 apparmor amd64 3.0.4-2ubuntu2.3 [595 kB]
Get:2 http://archive.ubuntu.com/ubuntu jammy/main amd64 liblzo2-2 amd64 2.10-2build3 [53.7 kB]
Get:3 http://archive.ubuntu.com/ubuntu jammy/main amd64 squashfs-tools amd64 1:4.5-3build1 [159 kB]
Get:4 http://archive.ubuntu.com/ubuntu jammy-updates/main amd64 libudev1 amd64 249.11-0ubuntu3.12 [78.2 kB]
Get:5 http://archive.ubuntu.com/ubuntu jammy-updates/main amd64 udev amd64 249.11-0ubuntu3.12 [1,557 kB]
Get:6 http://archive.ubuntu.com/ubuntu jammy/main amd64 libfuse3-3 amd64 3.10.5-1build1 [81.2 kB]
Get:7 http://archive.ubuntu.com/ubuntu jammy-updates/main amd64 snapd amd64 2.58+22.04.1 [23.8 MB]
Get:8 http://archive.ubuntu.com/ubuntu jammy-updates/universe amd64 chromium-browser amd64 1:85.0.4183.83-0ubuntu2.22.04.1 [49.2 kB]
Get:9 http://archive.ubuntu.com/ubuntu jammy-updates/universe amd64 chromium-chromedriver amd64 1:85.0.4183.83-0ubuntu2.22.04.1 [2,36
```

```

Get:10 http://archive.ubuntu.com/ubuntu jammy-updates/main amd64 systemd-hwe-hwdb all 249.11.5 [3,228 B]
Fetched 26.4 MB in 1s (20.1 MB/s)
Preconfiguring packages ...
Selecting previously unselected package apparmor.
(Reading database ... 121747 files and directories currently installed.)
Preparing to unpack .../apparmor_3.0.4-2ubuntu2.3_amd64.deb ...
Unpacking apparmor (3.0.4-2ubuntu2.3) ...
Selecting previously unselected package liblzo2-2:amd64.
Preparing to unpack .../liblzo2-2_2.10-2build3_amd64.deb ...
Unpacking liblzo2-2:amd64 (2.10-2build3) ...
Selecting previously unselected package squashfs-tools.
Preparing to unpack .../squashfs-tools_1%3a4.5-3build1_amd64.deb ...
Unpacking squashfs-tools (1:4.5-3build1) ...

```

```

from selenium.webdriver.common.by import By
from selenium.webdriver.chrome.options import Options
from selenium.webdriver.support.ui import WebDriverWait
from selenium.webdriver.support import expected_conditions as EC
from selenium.common.exceptions import TimeoutException
from selenium import webdriver

# Specify the path to the ChromeDriver binary
chrome_options = webdriver.ChromeOptions()
chrome_options.add_argument('--headless')
chrome_options.add_argument('--no-sandbox')
chrome_options.add_argument('--disable-dev-shm-usage')

# Set the path to the ChromeDriver binary
chrome_driver_path = "/usr/lib/chromium-browser/chromedriver"

# Initialize the Chrome driver with options
driver = webdriver.Chrome(options=chrome_options)

def scrape_cnbc_news():
    global news_data
    url = "https://www.cnbc.com/search/?query=green%20hydrogen&qsearchterm=green%20hydrogen"
    driver.get(url)
    try:
        # Wait for the dynamic content to load (adjust the wait time as needed)
        WebDriverWait(driver, 10).until(EC.presence_of_element_located((By.CLASS_NAME, 'SearchResult-searchResultTitle')))

        # Extract headlines and dates
        news_items = driver.find_elements(By.CLASS_NAME, 'SearchResult-searchResultContent')
        news_data = []
        for item in news_items:
            headline = item.find_element(By.CSS_SELECTOR, 'div.SearchResult-searchResultTitle a.resultlink span.Card-title').text.strip()
            date = item.find_element(By.CSS_SELECTOR, 'span.SearchResult-byline span.SearchResult-publishedDate').text.strip()
            news_data.append({'headline': headline, 'date': date})

        return news_data
    except TimeoutException:
        print("Timed out waiting for page to load")
        return []
    finally:
        # Close the WebDriver
        driver.quit()

scrape_cnbc_news()

```

```

[{'headline': 'Saudi Arabia's ACWA Power to develop largest green hydrogen project in Indonesia worth over $1 ...',
  'date': '1/29/2024 8:10:42 AM'},
{'headline': 'TES CEO: 2024 must be a year of delivery for green goals',
  'date': '1/15/2024 12:37:50 PM'},
{'headline': 'Europe does not have an incentive program for green hydrogen, says Clariant CEO',
  'date': '1/17/2024 3:07:08 PM'},
{'headline': 'The hydrogen stock bubble burst, but Bernstein says it's time for a second look at ...',
  'date': '8/11/2023 7:18:35 PM'},
{'headline': 'Here are Wednesday's biggest analyst calls: Apple, Tesla, Alphabet, Microsoft, AMD, Meta, Spotify and more',
  'date': '1/31/2024 1:20:15 PM'},
{'headline': 'We need everything: CEO on why natural gas infrastructure needs to be ready for clean ...',
  'date': '11/9/2023 10:00:18 AM'},
{'headline': 'There's a lot of talk about hydrogen's potential. But transportation costs represent a big challenge',
  'date': '7/4/2023 8:20:24 AM'},
{'headline': 'Biden's Inflation Reduction Act makes green hydrogen profitable at scale, Goldman Sachs says',
  'date': '11/30/2022 1:11:51 PM'},
{'headline': 'In Australia, firms plan 'super hub' to produce green hydrogen using wind and solar',
  'date': '11/16/2022 12:02:17 AM'},

```

```
{'headline': 'This green hydrogen stock can rally more than 20%, says Citi',  
'date': '7/6/2023 9:45:47 AM']}
```

```
# Extracting news headline, date and source from news.google.com
```

```
import requests  
from bs4 import BeautifulSoup  
  
def fetch_google_news_rss():  
    global news_data  
    url = "https://news.google.com/rss/search?q=green%20hydrogen&hl=en-IN&gl=IN&ceid=IN:en"  
    response = requests.get(url)  
    soup = BeautifulSoup(response.text, 'xml')  
    news_data = []  
    for item in soup.find_all('item'):  
        headline = item.title.text.strip()  
        date = item.pubDate.text.strip()  
        source = item.source.text.strip() if item.source else "N/A" # Check if source is present  
        news_data.append({'headline': headline, 'date': date, 'source': source})  
    return news_data
```

```
fetch_google_news_rss()
```

```
{'source': 'Mercom India'},  
{'headline': 'Govt shares guidelines for Green Hydrogen Pilot Projects in steel sector - Moneycontrol',  
'date': 'Sun, 04 Feb 2024 10:13:01 GMT',  
'source': 'Moneycontrol'},  
{'headline': 'NGEL Partners with Maharashtra Government for Green Hydrogen Project Development - SolarQuarter',  
'date': 'Mon, 05 Feb 2024 08:09:57 GMT',  
'source': 'SolarQuarter'},  
{'headline': 'Govt allocates Rs 115 crore for green hydrogen pilot in shipping sector: MNRE - ETEnergyWorld',  
'date': 'Fri, 02 Feb 2024 17:35:36 GMT',  
'source': 'ETEnergyWorld'},  
{'headline': 'India's Plans to Achieve Cost-Effective Green Hydrogen Production - India Briefing',  
'date': 'Fri, 26 Jan 2024 08:00:00 GMT',  
'source': 'India Briefing'},  
{'headline': 'Study unveils scalable and efficient photoelectrode modules for green hydrogen production - Tech Xplore',  
'date': 'Wed, 07 Feb 2024 14:49:03 GMT',  
'source': 'Tech Xplore'},  
{'headline': 'India plans green hydrogen pilot projects in shipping and steel - H2 View',  
'date': 'Mon, 05 Feb 2024 11:53:26 GMT',  
'source': 'H2 View'},  
{'headline': 'SunHydrogen, COTEC step up production of innovative green hydrogen panels - ETN News',  
'date': 'Tue, 06 Feb 2024 09:34:58 GMT',  
'source': 'ETN News'},  
{'headline': 'MNRE issues guidelines for pilot projects implementing Green Hydrogen in steel sector - CNBCTV18',  
'date': 'Fri, 02 Feb 2024 17:51:25 GMT',  
'source': 'CNBCTV18'},  
{'headline': 'Mangaluru: Green hydrogen, coal bleaching are best choice for clean energy - Minister K J George - Daijiworld.com',  
'date': 'Mon, 05 Feb 2024 14:14:14 GMT',  
'source': 'Daijiworld.com'},  
{'headline': 'Government Initiates Plans to Promote Green Hydrogen for Round-the-Clock RE - Asia Pacific - Energetica India',  
'date': 'Wed, 07 Feb 2024 06:56:11 GMT',  
'source': 'Energetica India'},  
{'headline': 'Green Hydrogen Revolution in Indian Steel: National Mission Unleashes Promising Pilot Initiatives - SolarQuarter',  
'date': 'Sat, 03 Feb 2024 05:29:53 GMT',  
'source': 'SolarQuarter'},  
{'headline': 'New industry-academia collab wants to reduce cost of green hydrogen production - Offshore Energy',  
'date': 'Wed, 07 Feb 2024 10:46:44 GMT',  
'source': 'Offshore Energy'},  
{'headline': 'Tecnimont (MAIRE) awarded a front-end-engineering design for an integrated green hydrogen and ammonia plant in Portugal - World Fertilizer',  
'date': 'Wed, 07 Feb 2024 09:26:13 GMT',  
'source': 'World Fertilizer'},  
{'headline': 'Support for Ports and Portability to Green Hydrogen - ETAuto',  
'date': 'Sat, 03 Feb 2024 03:38:07 GMT',  
'source': 'ETAuto'},  
{'headline': 'Budget 2024: Green hydrogen's allocation rises by whopping 102 per cent - ETEnergyWorld',  
'date': 'Fri, 02 Feb 2024 01:30:00 GMT',  
'source': 'ETEnergyWorld'},  
{'headline': 'EU awards HyFive €28.7m to develop three green hydrogen plants in Spain and Germany - H2 View',  
'date': 'Thu, 08 Feb 2024 12:10:10 GMT',  
'source': 'H2 View'},  
{'headline': 'Oracle Power Advances Green Hydrogen Production in Pakistan With Feasibility Study - Hydrogen Central',  
'date': 'Wed, 07 Feb 2024 18:41:23 GMT',  
'source': 'Hydrogen Central'},  
{'headline': 'India rolls out program to support green hydrogen ship retrofits - Offshore Energy',  
'date': 'Mon, 05 Feb 2024 13:59:55 GMT',  
'source': 'Offshore Energy'},  
{'headline': 'Green Hydrogen Mission gets whopping 102% hike in allocation in Interim Budget - The Economic Times',  
'date': 'Thu, 01 Feb 2024 15:38:12 GMT',  
'source': 'The Economic Times'}
```

```
!pip install pandas matplotlib wordcloud
```

```
Requirement already satisfied: pandas in /usr/local/lib/python3.10/dist-packages (1.5.3)
Requirement already satisfied: matplotlib in /usr/local/lib/python3.10/dist-packages (3.7.1)
Requirement already satisfied: wordcloud in /usr/local/lib/python3.10/dist-packages (1.9.3)
Requirement already satisfied: python-dateutil>=2.8.1 in /usr/local/lib/python3.10/dist-packages (from pandas) (2.8.2)
Requirement already satisfied: pytz>=2020.1 in /usr/local/lib/python3.10/dist-packages (from pandas) (2023.4)
Requirement already satisfied: numpy>=1.21.0 in /usr/local/lib/python3.10/dist-packages (from pandas) (1.23.5)
Requirement already satisfied: contourpy>=1.0.1 in /usr/local/lib/python3.10/dist-packages (from matplotlib) (1.2.0)
Requirement already satisfied: cycler>=0.10 in /usr/local/lib/python3.10/dist-packages (from matplotlib) (0.12.1)
Requirement already satisfied: fonttools>=4.22.0 in /usr/local/lib/python3.10/dist-packages (from matplotlib) (4.47.2)
Requirement already satisfied: kiwisolver>=1.0.1 in /usr/local/lib/python3.10/dist-packages (from matplotlib) (1.4.5)
Requirement already satisfied: packaging>=20.0 in /usr/local/lib/python3.10/dist-packages (from matplotlib) (23.2)
Requirement already satisfied: pillow>=6.2.0 in /usr/local/lib/python3.10/dist-packages (from matplotlib) (9.4.0)
Requirement already satisfied: pyparsing>=2.3.1 in /usr/local/lib/python3.10/dist-packages (from matplotlib) (3.1.1)
Requirement already satisfied: six>=1.5 in /usr/local/lib/python3.10/dist-packages (from python-dateutil>=2.8.1->pandas) (1.16.0)
```

```
import pandas as pd
from wordcloud import WordCloud
import matplotlib.pyplot as plt
import datetime
from transformers import pipeline

# create dataframe
df = pd.DataFrame(news_data)

# load the sentiment analysis model
sentiment_analysis = pipeline("sentiment-analysis")

# Add a sentiment score column
df['sentiment_score'] = df['headline'].apply(lambda x: sentiment_analysis(x)[0]['score'])
print(df)
```

No model was supplied, defaulted to distilbert-base-uncased-finetuned-sst-2-english and Using a pipeline without specifying a model name and revision in production is not recommended
/usr/local/lib/python3.10/dist-packages/huggingface_hub/utils/_token.py:88: UserWarning: The secret `HF_TOKEN` does not exist in your Colab secrets.
To authenticate with the Hugging Face Hub, create a token in your settings tab (<https://huggingface.co/settings/tokens>)
You will be able to reuse this secret in all of your notebooks.
Please note that authentication is recommended but still optional to access public model

```
warnings.warn(
config.json: 100%                               629/629 [00:00<00:00, 16.5kB/s]
model.safetensors: 100%                         268M/268M [00:05<00:00, 33.5MB/s]
tokenizer_config.json: 100%                     48.0/48.0 [00:00<00:00, 1.58kB/s]
vocab.txt: 100%                                232k/232k [00:00<00:00, 2.51MB/s]
```

```

      headline \
0  MNRE Launches Pilot Projects to Implement Gree...
1  India Awards Tenders for 4.12 Lakh Tonnes of G...
2  Centre working to promote use of green hydroge...
3  "Once Volume Increases, Green Hydrogen Will Co...
4  5 ways to make green hydrogen available around...
..
95  Green Hydrogen - Shipping and Steel Sectors - ...
96  Government of India Unveils Guidelines for Gre...
97  SAF, Green Hydrogen, CCUS likely to be in spot...
98  Kraft Heinz Partners With Carlton Power to Del...
99  Govt to provide 50% financial aid to pilot pro...

      date                source  sentiment_score
0  Mon, 05 Feb 2024 10:00:10 GMT  Mercom India      0.993594
1  Thu, 08 Feb 2024 11:24:20 GMT  SolarQuarter     0.807997
2  Wed, 07 Feb 2024 08:32:14 GMT   ThePrint      0.983792
3  Tue, 06 Feb 2024 12:52:16 GMT    NDTV          0.996130
4  Mon, 05 Feb 2024 17:30:45 GMT  World Economic Forum  0.997399
..
95  Mon, 05 Feb 2024 08:47:52 GMT   Insights IAS    0.994936
96  Sat, 03 Feb 2024 07:47:00 GMT  Energetica India  0.995704
97  Sat, 03 Feb 2024 16:54:39 GMT    PSU Watch    0.964729
98  Mon, 05 Feb 2024 08:00:00 GMT   Business Wire   0.993533
99  Fri, 02 Feb 2024 16:10:57 GMT    PSU Watch    0.717441
```

```
[100 rows x 4 columns]
```

```
# using spacy for performing named entity recognition in order to extract organization names
```

```
import spacy
```

```

# Load the spaCy NER model
ner = spacy.load("en_core_web_sm")

def extract_organization_names(text):
    doc = ner(text)
    organizations = [ent.text for ent in doc.ents if ent.label_ == "ORG"]
    return organizations

# Add an 'organization_names' column
df['organization_names'] = df['headline'].apply(extract_organization_names)

# Display the DataFrame
print(df)

```

```

                                headline \
0  MNRE Launches Pilot Projects to Implement Gree...
1  India Awards Tenders for 4.12 Lakh Tonnes of G...
2  Centre working to promote use of green hydroge...
3  "Once Volume Increases, Green Hydrogen Will Co...
4  5 ways to make green hydrogen available around...
..
95  Green Hydrogen - Shipping and Steel Sectors - ...
96  Government of India Unveils Guidelines for Gre...
97  SAF, Green Hydrogen, CCUS likely to be in spot...
98  Kraft Heinz Partners With Carlton Power to Del...
99  Govt to provide 50% financial aid to pilot pro...

                                date          source  sentiment_score \
0  Mon, 05 Feb 2024 10:00:10 GMT      Mercom India      0.993594
1  Thu, 08 Feb 2024 11:24:20 GMT      SolarQuarter      0.807997
2  Wed, 07 Feb 2024 08:32:14 GMT          ThePrint      0.983792
3  Tue, 06 Feb 2024 12:52:16 GMT          NDTV          0.996130
4  Mon, 05 Feb 2024 17:30:45 GMT  World Economic Forum      0.997399
..
95  Mon, 05 Feb 2024 08:47:52 GMT      Insights IAS      0.994936
96  Sat, 03 Feb 2024 07:47:00 GMT      Energetica India      0.995704
97  Sat, 03 Feb 2024 16:54:39 GMT          PSU Watch      0.964729
98  Mon, 05 Feb 2024 08:00:00 GMT      Business Wire      0.993533
99  Fri, 02 Feb 2024 16:10:57 GMT          PSU Watch      0.717441

                                organization_names
0                                     []
1          [India Awards Tenders, Green Hydrogen]
2                                     []
3                                     []
4          [World Economic Forum - World Economic Forum]
..
95                                [Green Hydrogen]
96                                     []
97          [SAF, Green Hydrogen, CCUS, EIL]
98  [Kraft Heinz Partners, Carlton Power to Delive...
99                                [Green Hydrogen]

[100 rows x 5 columns]

```

```

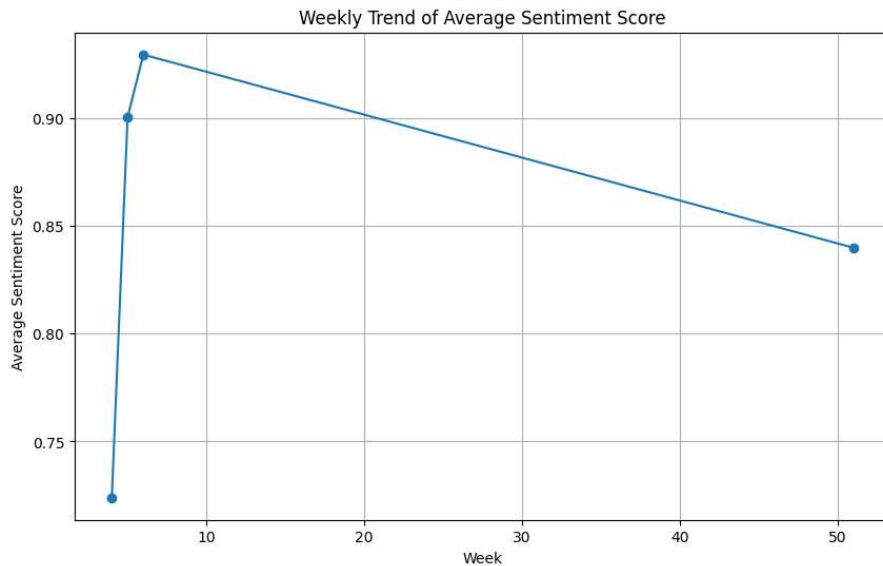
# Convert 'date' column to datetime format
df['date'] = pd.to_datetime(df['date'])

# Extract week number from the date
df['week'] = df['date'].dt.isocalendar().week

# Group by week and calculate average sentiment score
weekly_sentiment = df.groupby('week')['sentiment_score'].mean()

# Plotting the trend
plt.figure(figsize=(10, 6))
plt.plot(weekly_sentiment, marker='o')
plt.title('Weekly Trend of Average Sentiment Score')
plt.xlabel('Week')
plt.ylabel('Average Sentiment Score')
plt.grid(True)
plt.show()

```



```
# Concatenate organization names from all headlines
all_organizations = ' '.join(df['organization_names'].explode().dropna())

# Generate and display the word cloud
wordcloud = WordCloud(width=800, height=400, background_color='white').generate(all_organizations)
plt.figure(figsize=(10, 6))
plt.imshow(wordcloud, interpolation='bilinear')
plt.axis('off')
plt.title('Word Cloud of Organization Names in Headlines')
plt.show()
```



```
df = pd.DataFrame(news_data)
df.to_csv('news_data.csv', index=False)
```

```
from google.colab import files
from google.colab import drive

drive.mount('/content/drive')
!cp news_data.csv "/content/drive/My Drive/"
```

Mounted at /content/drive

```
from google.colab import auth
auth.authenticate_user()

import gspread

# Authenticate google sheet
from google.auth import default
creds, _ = default()

gs = gspread.authorize(creds)

# create google sheet in google drive
sh = gs.create("news_data")

# get google sheet
worksheet = gs.open("news_data").sheet1

df = pd.DataFrame(news_data)

# Update the worksheet with DataFrame values
worksheet.update([df.columns.values.tolist()] + df.values.tolist())

# Share the Google Sheet with access rights to "Anyone with the Link"
sh.share("", perm_type="anyone", role="reader")

# Get the link to the shared Google Sheet
sheet_link = f"https://docs.google.com/spreadsheets/d/{sh.id}"
print(f"Google Sheet Link: {sheet_link}")
```

Google Sheet Link: https://docs.google.com/spreadsheets/d/1CGmnHXSSpBHCm3DSmyZhp0fX1jC60xwX4VEuFW9s2_A