```
from bs4 import BeautifulSoup
import pandas as pd
import requests
import json
import time
import re
```

⌄ Business Insider

```
url = "https://www.businessinsider.in/international/news/23"
source = requests.get(url).text
soup = BeautifulSoup(source, 'html.parser')
```

```
print(soup.prettify())
```

```
<!DOCTYPE html>
<html xmlns:ext="http://exslt.org/common" xmlns:str="http://exslt.org/strings">
 <head>
  <meta content="text/html; charset=utf-8" http-equiv="Content-Type"/>
  <title>
   Page 23: International News - read latest and trending International news and headlines
  </title>
  <meta content="Wednesday, 31 January, 2024  08:51:20PM" http-equiv="Last-Modified"/>
  <meta content="Page 23: International News - read breaking and trending International news, stories, headlines. Business insider Ind
  <meta content="International news" name="keywords"/>
  <meta content="NOODP" name="robots"/>
  <link href="https://www.businessinsider.in/international/news/22" rel="prev"/>
  <link href="https://www.businessinsider.in/international/news/24" rel="next"/>
  <link href="https://www.businessinsider.in/international/news/23" rel="canonical">
   <link href="//www.google-analytics.com" rel="preconnect"/>
   <link href="//connect.facebook.net" rel="preconnect"/>
   <link href="//cm.g.doubleclick.net" rel="preconnect"/>
   <link href="//googleads.g.doubleclick.net" rel="preconnect"/>
   <link href="//googleads.g.doubleclick.net" rel="preconnect"/>
   <link href="https://www.businessinsider.in" rel="preconnect"/>
   <link href="https//sb.scorecardresearch.com" rel="preconnect"/>
   <link href="//ade.clmbtech.com" rel="preconnect"/>
   <link href="//static.clmbtech.com" rel="preconnect"/>
   <link href="https://www.googletagmanager.com" rel="preconnect"/>
   <link href="https://www.googletagservices.com" rel="preconnect"/>
   <link href="https://tpc.googlesyndication.com" rel="preconnect"/>
   <link href="https://pagead2.googlesyndication.com" rel="preconnect"/>
   <link href="https://ade.clmbtech.com" rel="preconnect"/>
   <link href="https://www.googletagservices.com" rel="dns-prefetch"/>
   <link href="//www.google-analytics.com" rel="dns-prefetch"/>
   <link href="//connect.facebook.net" rel="dns-prefetch"/>
   <link href="//cm.g.doubleclick.net" rel="dns-prefetch"/>
   <link href="//googleads.g.doubleclick.net" rel="dns-prefetch"/>
   <link href="https://www.businessinsider.in" rel="dns-prefetch"/>
   <link href="https://sb.scorecardresearch.com" rel="dns-prefetch"/>
   <link href="https://ade.clmbtech.com" rel="dns-prefetch"/>
   <link href="https://static.clmbtech.com" rel="dns-prefetch"/>
   <link href="https://www.googletagmanager.com" rel="dns-prefetch"/>
   <link href="https://tpc.googlesyndication.com" rel="dns-prefetch"/>
   <link href="https://pagead2.googlesyndication.com" rel="dns-prefetch"/>
   <link href="https://ade.clmbtech.com" rel="dns-prefetch"/>
   <link href="/pn_manifest.cms" rel="manifest"/>
   <meta content="NOODP" name="robots"/>
   <meta content="text/html; charset=utf-8" http-equiv="Content-Type"/>
   <meta content="NpZv5Qfl0TuFiebbwMualP_h2a3QtmQGFjE9hhw9um8" name="google-site-verification"/>
   <meta content="True" name="HandheldFriendly"/>
   <meta content="320" name="MobileOptimized"/>
   <meta content="width=device-width, initial-scale=1.0, minimum-scale=1.0, maximum-scale=1.0, user-scalable=no" name="viewport"/>
   <script type="application/ld+json">
    {
   "@context" : "http://schema.org",
   "@type" : "Organization",
   "name" : "Business Insider India",
   "url" : "https://www.businessinsider.in",
   "logo":{
        "@type":"ImageObject",
        "width":"600",
```

```python
import json
from bs4 import BeautifulSoup
import requests

url = "https://www.businessinsider.in/international/news/23"
source = requests.get(url).text
soup = BeautifulSoup(source, 'html.parser')

# Find the specific script containing the URLs based on context or other attributes
target_script = None
for script in soup.find_all('script', {'type': 'application/ld+json'}):
    if '"@type" : "ItemList"' in script.text:
        target_script = script
        break

if target_script:
    # Extract the JSON-like data
    script_content = target_script.contents[0].strip()

    # Load JSON data
    data = json.loads(script_content)

    # Extract the URLs from the JSON data
    urls = [
        item['url']
        for item in data.get('itemListElement', [])
    ]

    # Print the URLs with their indices
    for index, url in enumerate(urls, start=1):
        print(f"{index}. {url}")
else:
    print("Target script not found.")
```

1. https://www.businessinsider.in/international/news/ukrainian-partisans-say-they-killed-24-russian-soldiers-in-crimea-after-poisoning-t
2. https://www.businessinsider.in/international/news/ukraine-is-using-its-advanced-leopard-tanks-like-long-range-artillery-instead-of-pe
3. https://www.businessinsider.in/international/news/at-least-30-american-troops-have-died-in-a-spate-of-us-military-helicopter-and-tilt
4. https://www.businessinsider.in/international/news/ukraine-launched-a-barrage-of-attack-drones-at-crimea-overnight-russia-says/article
5. https://www.businessinsider.in/international/news/few-things-worry-nato-like-russias-yasen-class-submarines-and-moscow-may-be-plannin
6. https://www.businessinsider.in/international/news/israel-may-try-to-drive-hamas-out-of-its-tunnels-in-gaza-by-pumping-them-full-of-se
7. https://www.businessinsider.in/international/news/ukraine-is-on-track-to-make-72-howitzers-next-year-but-will-struggle-to-get-the-amm
8. https://www.businessinsider.in/international/news/ukrainian-soldiers-feel-they-have-been-abandoned-on-a-crucial-part-of-the-frontline
9. https://www.businessinsider.in/international/news/indiana-lottery-approved-50000-win-for-a-man-whose-ticket-was-torn-into-pieces-by-a
10. https://www.businessinsider.in/international/news/sheryl-sandberg-hillary-clinton-sen-kirsten-gillibrand-and-others-issue-a-global-c
11. https://www.businessinsider.in/international/news/former-fox-employee-alleges-in-a-new-lawsuit-that-tucker-carlsons-right-hand-man-j
12. https://www.businessinsider.in/international/news/biden-nominates-indian-american-vc-for-key-post/articleshow/105744147.cms
13. https://www.businessinsider.in/international/news/a-us-navy-warship-had-to-gun-down-more-threats-as-american-forces-get-pulled-deepe
14. https://www.businessinsider.in/international/news/a-missing-botticelli-masterpiece-worth-109-million-was-hanging-in-an-italian-famil
15. https://www.businessinsider.in/international/news/ukrainian-soldier-admits-some-of-the-marines-trying-to-defend-a-key-river-against-
16. https://www.businessinsider.in/international/news/oil-titan-sultan-al-jaber-backtracks-on-comments-that-there-is-no-science-around-f
17. https://www.businessinsider.in/international/news/russia-is-ramping-up-its-use-of-long-range-glide-bombs-because-ukraine-keeps-shoot
18. https://www.businessinsider.in/international/news/a-new-era-for-the-market-means-a-new-way-to-think-about-investing/articleshow/1057
19. https://www.businessinsider.in/international/news/russia-likely-doesnt-know-how-many-troops-its-lost-because-its-military-has-a-cult
20. https://www.businessinsider.in/international/news/as-world-leaders-gather-for-cop28-one-question-looms-who-will-pay-for-the-green-tr
21. https://www.businessinsider.in/international/news/a-florida-lottery-ticket-worth-44-million-is-unclaimed-and-set-to-expire-in-just-a
22. https://www.businessinsider.in/international/news/the-air-force-is-trying-to-save-2-severely-damaged-f-35s-by-merging-them-into-a-si
23. https://www.businessinsider.in/international/news/islamic-resistance-vows-retaliation-after-us-airstrike-killed-5-iraqi-militants/ar
24. https://www.businessinsider.in/international/news/rizzlers-rejoice-the-oxford-english-dictionary-has-announced-its-word-of-the-year/
25. https://www.businessinsider.in/international/news/the-oxycontin-peddling-sackler-family-has-tried-to-get-immunity-from-civil-lawsuit

```python
import requests
from bs4 import BeautifulSoup
import re
import json

bi_dict1 = {}
i=0
# Assuming filtered_href_values is a list of href values

for idx in range(1, 23): #as it contains till december i.e one month
    url1 = f"https://www.businessinsider.in/international/news/{idx}"
    source = requests.get(url1).text
    soup = BeautifulSoup(source, 'html.parser')

    # Initialize a list to store text values for the current iteration
    a = []
```

```python
# Find the specific script containing the URLs based on context or other attributes
target_script = None
for script in soup.find_all('script', {'type': 'application/ld+json'}):
    if '"@type" : "ItemList"' in script.text:
        target_script = script
        break

if target_script:
    # Extract the JSON-like data
    script_content = target_script.contents[0].strip()

    # Load JSON data
    data = json.loads(script_content)

    # Extract the URLs from the JSON data
    urls = [
        item['url']
        for item in data.get('itemListElement', [])
    ]

    # Print the URLs with their indices
    for index, url in enumerate(urls, start=1):
        print(f"{index}. {url}")
        bi_dict1[i]=url
        i += 1
else:
    print("Target script not found.")
```

Target script not found.
1. https://www.businessinsider.in/international/news/a-woman-in-china-has-decided-to-leave-her-2-8m-fortune-to-her-dogs-and-cats-and-r
2. https://www.businessinsider.in/international/news/us-special-operations-forces-are-searching-for-a-new-extreme-long-range-rifle-for
3. https://www.businessinsider.in/international/news/i-dont-mind-giving-my-kids-screen-time-but-teaching-them-how-to-find-what-they-wa
4. https://www.businessinsider.in/international/news/kremlin-denies-report-that-putin-reached-out-to-the-us-about-talks-on-ending-the-
5. https://www.businessinsider.in/international/news/russia-is-changing-tactics-to-try-to-capture-the-city-of-avdiivka-and-plans-to-fi
6. https://www.businessinsider.in/international/news/ukraine-appears-to-be-attacking-russias-oil-and-gas-industry-with-small-cheap-dro
7. https://www.businessinsider.in/international/news/the-father-of-an-idf-soldier-killed-on-october-7-claimed-his-sons-decapitated-hea
8. https://www.businessinsider.in/international/news/ukrainian-forces-are-using-flocks-of-fpv-drones-led-by-queen-drone-to-attack-russ
9. https://www.businessinsider.in/international/news/im-an-older-dad-so-ill-be-a-senior-citizen-when-my-2-toddlers-graduate-from-high-
10. https://www.businessinsider.in/international/news/i-missed-my-child-free-life-after-having-my-baby-even-though-i-love-her-it-took-
11. https://www.businessinsider.in/international/news/i-was-mom-shamed-for-leaving-my-kids-to-go-back-to-school-my-partner-was-the-one
12. https://www.businessinsider.in/international/news/ukrainian-soldier-torn-apart-by-tank-fire-fighting-for-mariupol-faced-further-he
13. https://www.businessinsider.in/international/news/chinas-air-force-is-far-from-ready-for-a-crucial-mission-it-needs-to-seize-taiwa
14. https://www.businessinsider.in/international/news/ukraine-takes-out-one-of-russias-most-feared-drone-aces-say-reports/articleshow/
15. https://www.businessinsider.in/international/news/video-appears-to-show-ukraines-new-ironclad-drone-vehicle-machine-gunning-a-russ
16. https://www.businessinsider.in/international/news/vince-mcmahon-resigns-from-wwe-parent-company-board-following-sex-abuse-lawsuit/
17. https://www.businessinsider.in/international/news/russia-has-the-advantage-and-ukraine-needs-to-dig-in-if-its-going-to-fend-off-th
18. https://www.businessinsider.in/international/news/russia-says-its-scientists-created-a-new-nuclear-blast-simulator-with-flash-and-
19. https://www.businessinsider.in/international/news/photos-show-the-iranian-missile-parts-us-navy-seals-seized-in-a-deadly-nighttime
20. https://www.businessinsider.in/international/news/ukrainian-drone-operator-says-there-are-so-many-drones-in-the-war-that-nobody-kn
21. https://www.businessinsider.in/international/news/police-arrested-a-thai-woman-whose-pet-lion-was-seen-on-video-cruising-around-in
22. https://www.businessinsider.in/international/news/a-us-air-force-fighter-wing-is-asking-the-public-to-please-stop-pointing-lasers-
23. https://www.businessinsider.in/international/news/taiwan-says-china-is-still-sending-balloons-to-the-island-which-it-calls-an-act-
24. https://www.businessinsider.in/international/news/stop-frying-toothpicks-and-eating-them-south-korea-tells-mukbang-influencers/art
25. https://www.businessinsider.in/international/news/read-the-explosive-resignation-letter-sent-by-the-chairman-of-sports-illustrated
1. https://www.businessinsider.in/international/news/a-man-robbed-a-bank-days-after-being-found-not-guilty-of-robbing-a-different-bank
2. https://www.businessinsider.in/international/news/nato-vastly-underestimated-russias-ability-to-resupply-its-army-with-troops-and-a
3. https://www.businessinsider.in/international/news/increased-us-military-presence-in-the-middle-east-has-cost-1-6-billion-a-bill-it-
4. https://www.businessinsider.in/international/news/an-amish-family-had-their-horse-and-buggy-stolen-from-a-walmart-parking-lot-while
5. https://www.businessinsider.in/international/news/im-an-american-in-south-korea-living-here-allowed-me-to-have-a-bigger-family-/art
6. https://www.businessinsider.in/international/news/ukraine-floats-the-theory-it-may-have-been-tricked-into-shooting-down-the-plane-
7. https://www.businessinsider.in/international/news/nawaz-sharif-supporters-bring-lion-tiger-to-a-rally-in-lahore/articleshow/1071336
8. https://www.businessinsider.in/international/news/netflixs-password-crackdown-worked-it-could-have-more-tricks-to-lure-subscribers-
9. https://www.businessinsider.in/international/news/a-us-military-base-was-pummeled-with-rogue-waves-from-a-storm-surge-on-the-marsha
10. https://www.businessinsider.in/international/news/war-in-ukraine-shows-a-need-to-rethink-air-superiority-as-it-may-only-be-possibl
11. https://www.businessinsider.in/international/news/attacks-by-irans-militia-allies-are-placing-the-biden-administration-in-a-bind/a
12. https://www.businessinsider.in/international/news/the-houthis-are-still-trying-to-attack-ships-but-most-of-the-time-they-cant-get-
13. https://www.businessinsider.in/international/news/ukraine-looks-to-be-losing-the-battle-for-ammunition-and-its-forces-are-being-ma
14. https://www.businessinsider.in/international/news/zelenskyys-aides-kept-an-emergency-escape-train-on-standby-for-him-at-the-start-
15. https://www.businessinsider.in/international/news/woman-accused-of-murdering-her-doppelgnger-wanted-to-stage-her-death-to-get-away
16. https://www.businessinsider.in/international/news/russia-made-a-flashy-upgrade-to-its-cruise-missiles-since-ukraine-shot-down-so-m
17. https://www.businessinsider.in/international/news/russia-said-a-military-transport-plane-crashed-with-65-ukrainian-pows-on-board-r
18. https://www.businessinsider.in/international/news/china-doesnt-have-the-military-power-to-successfully-invade-taiwan-the-majority-
19. https://www.businessinsider.in/international/news/powerful-weapons-the-us-sent-to-ukraine-are-breaking-down-but-the-pentagon-says-
20. https://www.businessinsider.in/international/news/i-was-a-teen-mom-now-im-44-my-2-kids-are-adults-and-i-have-my-life-back-/article
21. https://www.businessinsider.in/international/news/why-the-west-cant-stop-the-houthi-rebels-missile-attacks-on-ships/articleshow/10
22. https://www.businessinsider.in/international/news/netflix-is-thrilled-with-its-crackdown-on-shared-passwords-dont-expect-it-to-end
23. https://www.businessinsider.in/international/news/a-ukrainian-soldier-says-he-escaped-from-behind-enemy-lines-hidden-inside-a-sofa
24. https://www.businessinsider.in/international/news/the-b-21-raider-took-its-first-flight-just-2-months-ago-but-the-pentagon-is-alre
25. https://www.businessinsider.in/international/news/after-my-divorce-i-only-get-my-kids-half-the-time-i-make-the-most-when-they-are-
1. https://www.businessinsider.in/international/news/ukraine-is-running-out-of-ammo-for-its-us-supplied-howitzers-as-it-tries-to-fend-
2. https://www.businessinsider.in/international/news/anna-wintour-kept-her-sunglasses-on-while-firing-pitchfork-staff-writer-says/arti
3. https://www.businessinsider.in/international/news/hamas-inflicted-a-huge-loss-on-israel-killing-21-soldiers-with-a-single-devastati

4. https://www.businessinsider.in/international/news/russias-meat-wave-assaults-leave-frozen-corpses-uncollected-as-russian-soldiers-
5. https://www.businessinsider.in/international/news/the-cia-is-posting-videos-to-persuade-russians-to-betray-their-leaders-and-says-
6. https://www.businessinsider.in/international/news/ukraine-appears-to-have-developed-an-exploding-drone-with-a-jet-engine-that-coul

```
bi_dict1[0]
```

    'https://www.businessinsider.in/international/news/a-woman-in-china-has-decided-to-leave-her-2-8m-fortune-to-her-dogs-and-cats-and-noth
    ing-to-her-children-says-report/articleshow/107209472.cms'

```
# Filter URLs containing the word "Hamas" (case-insensitive)
filtered_urls = {key: value for key, value in bi_dict1.items() if 'hamas' in value.lower()}

# Create a new dictionary with consecutive integer keys
filtered_urls_dict = {i + 1: url for i, url in enumerate(filtered_urls.values())}

# Print the filtered URLs with new keys
for key, value in filtered_urls_dict.items():
    print(f"{key}. {value}")
```

1. https://www.businessinsider.in/international/news/hamas-inflicted-a-huge-loss-on-israel-killing-21-soldiers-with-a-single-devastating
2. https://www.businessinsider.in/international/news/israeli-military-leaders-say-israel-has-a-choice-defeat-hamas-or-save-the-hostages/
3. https://www.businessinsider.in/international/news/10-year-old-boy-hospitalized-after-being-attacked-by-a-shark-during-a-shark-tank-ex
4. https://www.businessinsider.in/international/news/photos-show-subdued-christmas-celebrations-in-the-west-bank-town-of-bethlehem-amid-
5. https://www.businessinsider.in/international/news/hamasapos-oct-7-terror-attacks-left-destruction-in-their-wake-see-photos-of-what-is
6. https://www.businessinsider.in/international/news/the-3000-year-history-of-tunnel-warfare-shows-how-hard-itaposs-going-to-be-for-isra
7. https://www.businessinsider.in/international/news/powerful-photos-show-an-israeli-light-installation-projecting-224-beams-one-for-eac
8. https://www.businessinsider.in/international/news/hamas-using-top-grade-chinese-weapons-in-gaza-including-assault-rifles-and-grenade-
9. https://www.businessinsider.in/international/news/hamas-killed-in-an-apparent-israeli-hit-in-beirut-recently-said-he-was-living-on-bo
10. https://www.businessinsider.in/international/news/new-years-eve-terrorist-threat-in-us-and-europe-as-israel-hamas-war-fuels-rage/art
11. https://www.businessinsider.in/international/news/hamas-used-horrific-sexual-violence-raping-and-mutilating-israeli-women-and-girls-
12. https://www.businessinsider.in/international/news/a-gen-z-israeli-explained-why-hes-refusing-to-fight-hamas-now-israel-has-sent-him-
13. https://www.businessinsider.in/international/news/israel-drove-suitcases-stuffed-with-cash-into-gaza-to-keep-hamas-in-power-nyt/arti
14. https://www.businessinsider.in/international/news/former-idf-chiefs-nephew-killed-in-battle-in-southern-gaza-just-days-after-his-son
15. https://www.businessinsider.in/international/news/dramatic-video-of-a-passenger-plane-landing-at-tel-aviv-airport-amid-a-barrage-as-
16. https://www.businessinsider.in/international/news/idf-special-forces-raid-to-rescue-hamas-hostages-in-gaza-fails-say-israel-media-re
17. https://www.businessinsider.in/international/news/israeli-minister-said-his-son-killed-by-hamas-in-gaza-believed-he-was-fighting-a-j
18. https://www.businessinsider.in/international/news/freed-israeli-captives-held-by-hamas-were-terrified-idf-airstrikes-would-kill-them
19. https://www.businessinsider.in/international/news/israels-war-with-hamas-shows-why-even-50-years-of-upgrades-cant-make-tanks-invulne

```
bi_dict = {}

# Assuming filtered_urls_dict is a dictionary with indices as keys and URLs as values

for idx, url in filtered_urls_dict.items():
    # Fetch HTML content
    source = requests.get(url).text
    soup = BeautifulSoup(source, 'html.parser')

    # Initialize a list to store text values for the current iteration
    a = []

    # Find the script tag with the id '__NEXT_DATA__'
    script_tags = soup.find_all('script', type='application/ld+json')
    for script_tag in script_tags:
        try:
            json_data = json.loads(script_tag.string)
            text_data = json_data.get('articleBody', '')

            # Append the "text" values to the list
            if isinstance(text_data, str):
                a.append(text_data)

        except json.JSONDecodeError as e:
            print(f"Error decoding JSON: {e}")

    # Now, 'a' contains the stored text values for the current iteration
    print(f"Stored Text Values in 'a' for iteration {idx}:")
    for text_value in a:
        bi_dict.setdefault(idx, []).append(text_value)
```

    Stored Text Values in 'a' for iteration 1:
    Stored Text Values in 'a' for iteration 2:
    Stored Text Values in 'a' for iteration 3:
    Stored Text Values in 'a' for iteration 4:
    Stored Text Values in 'a' for iteration 5:

```
Stored Text Values in 'a' for iteration 6:
Stored Text Values in 'a' for iteration 7:
Stored Text Values in 'a' for iteration 8:
Stored Text Values in 'a' for iteration 9:
Stored Text Values in 'a' for iteration 10:
Stored Text Values in 'a' for iteration 11:
Stored Text Values in 'a' for iteration 12:
Stored Text Values in 'a' for iteration 13:
Stored Text Values in 'a' for iteration 14:
Stored Text Values in 'a' for iteration 15:
Stored Text Values in 'a' for iteration 16:
Stored Text Values in 'a' for iteration 17:
Stored Text Values in 'a' for iteration 18:
Stored Text Values in 'a' for iteration 19:
```

bi_dict[2]

```
['',
 '',
 "Top Israeli military officials say defeating Hamas could come at the cost of the lives of hostages.Some officials told NYT they don't
think Israel's two goals are both feasible.High-ranking Israeli military officials are increasingly concerned that if Israel has any
hope of defeating Hamas, it may have to do so at the cost of the lives of remaining Israeli hostages in Gaza, The New York Times
reported.The Times reported that some Israeli commanders are worried the country's twin goals of destroying Hamas and freeing the
hostages are not both feasible.Instead, The Times reported, some Israeli military officials believe that Israel may have to use
diplomatic channels — as opposed to military strategy — to recover the hostages from Gaza.Israel vowed to annihilate Hamas after the
militant group in October carried out the worst terrorist attack on Israeli soil in the nation's history, killing nearly 1,200 people
and taking more than 200 civilians hostage.The Israeli armed forces have carried out a brutal and sustained ground invasion across the
Gaza Strip in the months since, but The Times reported that they've made less progress than they hoped for.Hamas has proven more
difficult to erase than originally anticipated, in part because the group operates a vast network of underground tunnels and
infrastructure and is backed, both financially and militarily, by Iran.The Israel Defense Forces have taken out many Hamas militants
and commanders since launching their counterassault last year. But as of December, three of its top leaders remain at large, including
Mohammed Deif, the head of Hamas' military wing who masterminded the October 7 attacks.In addition to Israeli military sources
interviewed by The Times, European officials have also criticized Israel's military strategy and urged it to reconsider engaging in
peace talks.Their concerns are echoed by the families of hostages still believed to be alive, who say the safe return of their loved
ones should take precedence over destroying Hamas.The Hamas-run Gaza Health Ministry said over the weekend that 25,105 Gazans have died
since the war started, and 62,681 have been wounded."]
```

⌄ TIME

```python
from bs4 import BeautifulSoup
import pandas as pd
import requests
import json
```

```python
url = "https://time.com/tag/israel-hamas-war/?page=5"
source = requests.get(url).text
soup = BeautifulSoup(source, 'html.parser')
```

```python
print(soup.prettify())
```

```
<!DOCTYPE html>
<!--
://///////////////////////: ..-////-..         .//`              :/-          ..-/////////////////-
-`````````./////.`````````:  `////.          -///            -///          /////-`````````````.-
-         ./////.         -   `////.         ////:           ./////`        /////-            -
          ./////.             `////.        `/////-          `/////-        /////-
          ./////.             `////.       :///////.         `///////       /////-
          ./////.             `////.      //../////.        :/:-////.       /////-            `.
          ./////.             `////.    ./// :////.       :/: `////:       ////:..........:-
          ./////.             `////.     //:  :////`     -/:` -////`       ////:..........:-
          ./////.             `////.   `//`  ://///  -//`  `////-        /////-             `.
          ./////.             `////.    ://   `/////` -//`   :////        /////-
          ./////.             `////.    `//.   `////:-//`   .////.        /////-
          ./////.             `////.    -//`    `//////`    ////:        /////-             -
          ./////.             `////.    //:      `////.     :////`        /////-                 :
        ..-////-..          ..:////-.. ..-//:..        .//.         ..:////:-.`..-/////////////////:
-->
<html lang="en">
 <head>
  <title>
   Israel-Hamas War | Page 5 | Time
  </title>
  <meta charset="utf-8"/>
  <link href="/favicon.ico" rel="shortcut icon" type="image/vnd.microsoft.icon"/>
  <link href="/img/favicons/favicon-32.png" rel="icon" sizes="32x32"/>
  <meta content="IE=edge" http-equiv="X-UA-Compatible"/>
  <meta content="noindex, nofollow" name="robots"/>
  <link href="/img/favicons/favicon-57.png" rel="apple-touch-icon"/>
```

```
<link href="/img/favicons/favicon-57.png" rel="apple-touch-icon" sizes="57x57"/>
<link href="/img/favicons/favicon-72.png" rel="apple-touch-icon" sizes="72x72"/>
<link href="/img/favicons/favicon-76.png" rel="apple-touch-icon" sizes="76x76"/>
<link href="/img/favicons/favicon-114.png" rel="apple-touch-icon" sizes="114x114"/>
<link href="/img/favicons/favicon-120.png" rel="apple-touch-icon" sizes="120x120"/>
<link href="/img/favicons/favicon-144.png" rel="apple-touch-icon" sizes="144x144"/>
<link href="/img/favicons/favicon-152.png" rel="apple-touch-icon" sizes="152x152"/>
<link href="/img/favicons/favicon-180.png" rel="apple-touch-icon" sizes="180x180"/>
<link href="/img/favicons/favicon-192.png" rel="icon" sizes="192x192"/>
<link href="/manifest.json" rel="manifest"/>
<meta content="telephone=no" name="format-detection"/>
<meta content="width=device-width, initial-scale=1" name="viewport"/>
<meta content="Page 5 | Read the latest stories about Israel-Hamas War on Time" name="description"/>
<link href="https://time.com/tag/israel-hamas-war/?page=5" rel="canonical"/>
<meta content="Israel-Hamas War | Page 5 | Time" property="og:title"/>
<meta content="website" property="og:type"/>
<meta content="Time" property="og:site_name"/>
<meta content="https://time.com/tag/israel-hamas-war/?page=5" property="og:url"/>
<meta content="Page 5 | Read the latest stories about Israel-Hamas War on Time" property="og:description"/>
<meta content="https://time.com/tag/israel-hamas-war/?page=5" name="pinterest:url"/>
<meta content="summary_large_image" name="twitter:card"/>
<meta content="https://time.com/tag/israel-hamas-war/?page=5" name="twitter:url"/>
<meta content="Israel-Hamas War" name="twitter:title"/>
<meta content="Page 5 | Read the latest stories about Israel-Hamas War on Time" name="twitter:description"/>
<meta content="Page 5 | Read the latest stories about Israel-Hamas War on Time" name="pinterest:description"/>
<meta content="@TIME" name="twitter:site"/>
<meta content="http://time.com" name="pinterest:media"/>
<meta content="http://time.com" name="twitter:image"/>
<meta content="http://time.com" property="og:image"/>
```

```python
# Assuming 'soup' contains your BeautifulSoup object

links = soup.find_all('a', href=lambda href: href and re.match(r'^/\d', href))

# Initialize an empty list to store the complete URLs
complete_urls = []

# Prepend "https://time.com" and add the complete URLs to the list
for link in links:
    complete_url = 'https://time.com' + link['href']
    complete_urls.append(complete_url)

# Print or further process the list of complete URLs
print(complete_urls)
```

['https://time.com/6432223/biden-meeting-american-hamas-hostage-families/', 'https://time.com/6384946/biden-netanyahu-zionist-zelensky/'

```python
url = "https://time.com/6343715/israel-hamas-war-journalists-gaza/"
source = requests.get(url).text
soup = BeautifulSoup(source, 'html.parser')
```

```python
print(soup.prettify())
```

```
<!DOCTYPE html>
<html lang="en">
 <head>
  <meta charset="utf-8"/>
  <meta content="width=device-width, initial-scale=1" name="viewport"/>
  <link as="font" crossorigin="" href="/_next/static/media/010f47b240fbdab5-s.p.otf" rel="preload" type="font/otf"/>
  <link as="font" crossorigin="" href="/_next/static/media/0e4fe491bf84089c-s.p.woff2" rel="preload" type="font/woff2"/>
  <link as="font" crossorigin="" href="/_next/static/media/1d6159b7489fdf23-s.p.woff2" rel="preload" type="font/woff2"/>
  <link as="font" crossorigin="" href="/_next/static/media/26de8d18dde0c48f-s.p.woff2" rel="preload" type="font/woff2"/>
  <link as="font" crossorigin="" href="/_next/static/media/2b0140c2962edc97-s.p.woff2" rel="preload" type="font/woff2"/>
  <link as="font" crossorigin="" href="/_next/static/media/52db2a4aafb407ce-s.p.woff2" rel="preload" type="font/woff2"/>
  <link as="font" crossorigin="" href="/_next/static/media/5bb441026facce1b-s.p.woff2" rel="preload" type="font/woff2"/>
  <link as="font" crossorigin="" href="/_next/static/media/7119bfceaba8cfa6-s.p.woff2" rel="preload" type="font/woff2"/>
  <link as="font" crossorigin="" href="/_next/static/media/7d8c9b0ca4a64a5a-s.p.woff2" rel="preload" type="font/woff2"/>
  <link as="font" crossorigin="" href="/_next/static/media/84fc76f029a83dc5-s.p.woff2" rel="preload" type="font/woff2"/>
  <link as="font" crossorigin="" href="/_next/static/media/886f446b96dc7734-s.p.woff2" rel="preload" type="font/woff2"/>
  <link as="font" crossorigin="" href="/_next/static/media/92fbc930bebf0d1a-s.p.otf" rel="preload" type="font/otf"/>
  <link as="font" crossorigin="" href="/_next/static/media/934c4b7cb736f2a3-s.p.woff2" rel="preload" type="font/woff2"/>
  <link as="font" crossorigin="" href="/_next/static/media/94c2e74712c096dd-s.p.woff2" rel="preload" type="font/woff2"/>
  <link as="font" crossorigin="" href="/_next/static/media/9d37a504cc748d39-s.p.otf" rel="preload" type="font/otf"/>
  <link as="font" crossorigin="" href="/_next/static/media/a012f8576f047c44-s.p.woff2" rel="preload" type="font/woff2"/>
  <link as="font" crossorigin="" href="/_next/static/media/a3eaebd93ddd2233-s.p.woff2" rel="preload" type="font/woff2"/>
  <link as="font" crossorigin="" href="/_next/static/media/bf29a4e700232a12-s.p.otf" rel="preload" type="font/otf"/>
  <link as="font" crossorigin="" href="/_next/static/media/ddcf1e4d937d5abc-s.p.woff2" rel="preload" type="font/woff2"/>
```

```
        <link as="font" crossorigin="" href="/_next/static/media/e693e841d50dcf2f-s.p.woff2" rel="preload" type="font/woff2"/>
        <link as="font" crossorigin="" href="/_next/static/media/f8a669a999c483ca-s.p.woff2" rel="preload" type="font/woff2"/>
        <link as="font" crossorigin="" href="/_next/static/media/fe8ceef9917cb35a-s.p.woff2" rel="preload" type="font/woff2"/>
        <link as="image" fetchpriority="high" imagesizes="100vw" imagesrcset="https://api.time.com/wp-content/uploads/2023/12/GettyImages-1$
        <link data-precedence="next" href="/_next/static/css/a7a8d2bccda70abd.css" rel="stylesheet"/>
        <link data-precedence="next" href="/_next/static/css/ff0d02eb0cbd5a6e.css" rel="stylesheet"/>
        <link as="script" fetchpriority="low" href="/_next/static/chunks/webpack-b824b5bcade98f8e.js" rel="preload"/>
        <script async="" src="/_next/static/chunks/fd9d1056-b78f42e200247fb4.js">
        </script>
        <script async="" src="/_next/static/chunks/472-4023c6bb05930885.js">
        </script>
        <script async="" src="/_next/static/chunks/main-app-ce281d8555633725.js">
        </script>
        <script async="" src="/_next/static/chunks/33-df7fa13cfaa8bbd4.js">
        </script>
        <script async="" src="/_next/static/chunks/679-99d1092783bf9ca9.js">
        </script>
        <script async="" src="/_next/static/chunks/app/(content)/article/%5BarticleId%5D/%5B...slug%5D/layout-43b587317459618c.js">
        </script>
        <script async="" src="/_next/static/chunks/app/(content)/article/%5BarticleId%5D/%5B...slug%5D/error-ca3dd8d5534e2df7.js">
        </script>
        <link as="script" href="https://www.googletagmanager.com/gtag/js?id=UA-97981691-1" rel="preload"/>
        <link as="script" href="https://cdn.speedcurve.com/js/lux.js?id=4265275546" rel="preload"/>
        <link as="script" href="https://cdn.jwplayer.com/libraries/1BWlFwPe.js" rel="preload"/>
        <link as="script" href="//assets.time.com/draper/draper-next/0.6.0.js" rel="preload"/>
        <link as="script" href="//micro.rubiconproject.com/prebid/dynamic/20996.js" rel="preload"/>
        <link as="script" href="https://c.amazon-adsystem.com/aax2/apstag.js" rel="preload"/>
        <link as="script" href="https://securepubads.g.doubleclick.net/tag/js/gpt.js" rel="preload"/>
        <link as="script" href="https://pub.doubleverify.com/dvtag/21226187/DV759520/pub.js" rel="preload"/>
        <link as="script" href="https://fundingchoicesmessages.google.com/i/pub-6285516971701018?ers=1" rel="preload"/>
        <link as="script" href="https://socialcanvas-cdn.kargo.com/js/rh.min.js" rel="preload"/>
        <link as="script" href="//s.ntv.io/serve/load.js" rel="preload"/>
```

```python
# Find all 'p' tags with the specified class
p_elements = soup.find_all('p', class_='self-baseline px-0 font-pt-serif text-17px leading-7 tracking-0.5px')

# Extract and print the text content from all 'p' elements
for p_element in p_elements:
    text_content = p_element.get_text(strip=True)
    print(text_content)
```

For weeks, Motaz Azaiza's Instagram feed has exclusively borne witness to the horrors unfolding in Gaza. Since Oct. 7, the 24-year-old pho
But as Israel's punishing military campaign to root out Hamas from the Strip enters its third month, and as its ground invasion pushes f
The fate of journalists like Azaiza matters—not only for Gaza's Palestinians, many of whom have come to rely on their local press to rep
None of these journalists are neutral observers, nor do they claim to be. Each and every one of them is simultaneously covering and livi
"To be honest, I never imagined I would ever report on all of this violence," Hind Khoudary, a 28-year-old freelance reporter for Turkey
If there's one thing that keeps her going, Khoudary says, it's "the fact that people are listening and seeing and interacting and this i
The question is how long they can continue like this. "I no longer have any hope of survival," Bisan Owda, a 25-year-old Gazan filmmaker
Some of the most prominent Palestinian journalists to have emerged from the war have been forced to step back from their work. Plestia A
Her fears are not unfounded. At least 63 journalists have been killed covering the war, according to the Committee to Protect Journalists
Sherif Mansour, CPJ's Middle East and North Africa program coordinator, tells TIME that although this is hardly the first time Palestini
"What we're seeing in this war is that this deadly pattern is becoming more deadly," Mansour says, noting that some Palestinian journali
Protections are unlikely to be forthcoming. Israel and Egypt have prevented most international journalists from entering Gaza, and the I
As for Gaza's Palestinian journalists, they have largely been left to fend for themselves. For them, "infrastructure and protection and
"They are on the frontlines and in many ways they are the most needed," says Mansour. "But they are also the most vulnerable."

```
time_dict1 = {}
i = 0

# Assuming filtered_href_values is a list of href values
for idx in range(2, 6):  # as it contains till December i.e one month
    url1 = f"https://time.com/tag/israel-hamas-war/?page={idx}"
    source = requests.get(url1).text
    soup = BeautifulSoup(source, 'html.parser')

    # Find all 'a' tags with href starting with "/" and followed by a number
    links = soup.find_all('a', href=lambda href: href and re.match(r'^/\d', href))

    # Prepend "https://time.com" and print the complete URLs
    # Initialize an empty list to store the complete URLs
    complete_urls = []

    # Prepend "https://time.com" and add the complete URLs to the list
    for link in links:
        complete_url = 'https://time.com' + link['href']
        complete_urls.append(complete_url)

    # Extract the URLs from the list
    urls = [
        item for item in complete_urls
    ]

    # Print the URLs with their indices
    for index, url in enumerate(urls, start=1):
        print(f"{index}. {url}")
        time_dict1[i] = url
        i += 1
else:
    print("Target script not found.")
```

1. https://time.com/6564983/saudi-arabia-israel-palestinian-state/
2. https://time.com/6564827/israeli-hostages-call-to-action-davos-encounter/
3. https://time.com/6564776/women-children-victims-israel-hamas-united-nations/
4. https://time.com/6564535/gaza-ukraine-davos-world-economic-forum/
5. https://time.com/6552304/israel-united-nations-us/
6. https://time.com/6559293/morning-consult-israel-global-opinion/
7. https://time.com/6556113/curiosity-propaganda-hala-alyan-essay/
8. https://time.com/6556358/israel-hamas-war-hostages-qatar-france-send-medicine/
9. https://time.com/6556311/houthi-specially-designated-global-terrorists-us/
10. https://time.com/6556020/malta-flagged-carrier-hit-by-missile-as-houthi-red-sea-shipping-attacks-continue/
11. https://time.com/6555620/us-strikes-houthis-yemen/
12. https://time.com/6555634/houthi-yemen-missile-american-cargo/
13. https://time.com/6555577/hamas-israeli-hostages-video/
14. https://time.com/6555453/sagiv-jehezkel-antalyaspor-turkey-soccer-israel-hamas/
15. https://time.com/6555242/pro-palestinian-marches-london-paris-europe-demonstrations/
16. https://time.com/6554560/lesson-on-human-suffering-kibbutz-kfar-aza/
17. https://time.com/6554861/yemen-houthi-rebels-history-cause-israel-hamas-war/
18. https://time.com/6554472/gaza-starved-hunger-airdrop-food-aid/
19. https://time.com/6554381/poet-palestine-gaza-war/
20. https://time.com/6554003/israel-hamas-war-lessons/
1. https://time.com/6553708/gaza-end-of-western-hypocrisy-essay/
2. https://time.com/6553912/israel-south-africa-icj-genocide/
3. https://time.com/6549715/history-israel-hamas-war-end/
4. https://time.com/6553822/blinken-palestinian-governance-reform-postwar-gaza/
5. https://time.com/6553787/video-israeli-army-palestinians-west-bank-shot/
6. https://time.com/6553444/us-democrats-minority-outreach/
7. https://time.com/6553202/resurgence-global-anti-americanism-essay/
8. https://time.com/6553226/al-jazeera-journalist-son-killed-airstrike-al-dahdouh/
9. https://time.com/6553142/netanyahu-israel-hamas-war-plans-gaza/
10. https://time.com/6552740/gaza-israel-starvation-hunger/
11. https://time.com/6552732/blinken-mideast-tour-turkey-gaza-war-fears/
12. https://time.com/6552530/u-s-intervention-peace-middle-east/
13. https://time.com/6552144/israel-assassination-hamas-strategy/
14. https://time.com/6551721/biden-must-push-netanyahu-gaza/
15. https://time.com/6551540/saleh-arouri-hamas-official-killed-beirut/
16. https://time.com/6551363/israel-supreme-court-netanyahu/
17. https://time.com/6551345/israel-pulls-troops-from-gaza/
18. https://time.com/6551106/israeli-strikes-gaza-north-kill-palestinians/
19. https://time.com/6551044/israeli-teen-refuses-to-enlist-in-idf/
20. https://time.com/6550851/bethlehem-christmas-sermon-nativity-rubble/
1. https://time.com/6550772/un-resolution-gaza-aid-vote-ceasefire-criticism/
2. https://time.com/6550698/un-security-council-gaza-aid-resolution-ceasefire/
3. https://time.com/6549544/israel-and-hamas-the-media-war/
4. https://time.com/6549002/what-we-left-in-gaza-essay/
5. https://time.com/6549448/gaza-shelter-unlivable/
6. https://time.com/6549535/israel-hamas-war-ceasefire-hostage-swap-egypt/

7. https://time.com/6549112/u-s-coalition-force-red-sea/
8. https://time.com/6548968/houthi-rebels-shipping-attacks-red-sea-disrupt-global-economy-costs/
9. https://time.com/6548632/israel-hamas-war-civilians-hostages-killed/
10. https://time.com/6548183/yemen-houthi-rebels-red-sea-attacks-impact/
11. https://time.com/6548504/sustainable-ceasefire-gaza-european-global-ministers/
12. https://time.com/6548455/us-britain-navy-attack-drones-red-sea/
13. https://time.com/6358280/letters-from-gaza-iron-man-essay/
14. https://time.com/6548277/idf-kills-israeli-hostages-gaza/
15. https://time.com/6548160/wael-dahdouh-injured-israel-hamas-war/
16. https://time.com/6548068/palestinian-children-israeli-prison-arrested/
17. https://time.com/6548051/6548051/
18. https://time.com/6516866/terrorism-campus-free-speech-essay/

```python
len(time_dict1)
```

```
80
```

```python
time_dict1[1]
```

```
'https://time.com/6564827/israeli-hostages-call-to-action-davos-encounter/'
```

```python
time_dict = {}

# Assuming filtered_urls_dict is a dictionary with indices as keys and URLs as values

for idx, url in time_dict1.items():
    # Fetch HTML content
    source = requests.get(url).text
    soup = BeautifulSoup(source, 'html.parser')

    # Initialize a list to store text values for the current iteration
    a = []

    # Find the script tag with the id '__NEXT_DATA__'
    p_elements = soup.find_all('p', class_='self-baseline px-0 font-pt-serif text-17px leading-7 tracking-0.5px')
    for p_element in p_elements:
        try:
            text_content = p_element.get_text(strip=True)

            # Append the "text" values to the list
            if isinstance(text_content, str):
                a.append(text_content)

        except json.JSONDecodeError as e:
            print(f"Error decoding JSON: {e}")

    # Now, 'a' contains the stored text values for the current iteration
    print(f"Stored Text Values in 'a' for iteration {idx}:")
    for text_value in a:
        time_dict.setdefault(idx, []).append(text_value)
```

```
Stored Text Values in 'a' for iteration 0:
Stored Text Values in 'a' for iteration 1:
Stored Text Values in 'a' for iteration 2:
Stored Text Values in 'a' for iteration 3:
Stored Text Values in 'a' for iteration 4:
Stored Text Values in 'a' for iteration 5:
Stored Text Values in 'a' for iteration 6:
Stored Text Values in 'a' for iteration 7:
Stored Text Values in 'a' for iteration 8:
Stored Text Values in 'a' for iteration 9:
Stored Text Values in 'a' for iteration 10:
Stored Text Values in 'a' for iteration 11:
Stored Text Values in 'a' for iteration 12:
Stored Text Values in 'a' for iteration 13:
Stored Text Values in 'a' for iteration 14:
Stored Text Values in 'a' for iteration 15:
Stored Text Values in 'a' for iteration 16:
Stored Text Values in 'a' for iteration 17:
Stored Text Values in 'a' for iteration 18:
Stored Text Values in 'a' for iteration 19:
Stored Text Values in 'a' for iteration 20:
Stored Text Values in 'a' for iteration 21:
Stored Text Values in 'a' for iteration 22:
Stored Text Values in 'a' for iteration 23:
Stored Text Values in 'a' for iteration 24:
Stored Text Values in 'a' for iteration 25:
Stored Text Values in 'a' for iteration 26:
```

```
Stored Text Values in 'a' for iteration 27:
Stored Text Values in 'a' for iteration 28:
Stored Text Values in 'a' for iteration 29:
Stored Text Values in 'a' for iteration 30:
Stored Text Values in 'a' for iteration 31:
Stored Text Values in 'a' for iteration 32:
Stored Text Values in 'a' for iteration 33:
Stored Text Values in 'a' for iteration 34:
Stored Text Values in 'a' for iteration 35:
Stored Text Values in 'a' for iteration 36:
Stored Text Values in 'a' for iteration 37:
Stored Text Values in 'a' for iteration 38:
Stored Text Values in 'a' for iteration 39:
Stored Text Values in 'a' for iteration 40:
Stored Text Values in 'a' for iteration 41:
Stored Text Values in 'a' for iteration 42:
Stored Text Values in 'a' for iteration 43:
Stored Text Values in 'a' for iteration 44:
Stored Text Values in 'a' for iteration 45:
Stored Text Values in 'a' for iteration 46:
Stored Text Values in 'a' for iteration 47:
Stored Text Values in 'a' for iteration 48:
Stored Text Values in 'a' for iteration 49:
Stored Text Values in 'a' for iteration 50:
Stored Text Values in 'a' for iteration 51:
Stored Text Values in 'a' for iteration 52:
Stored Text Values in 'a' for iteration 53:
Stored Text Values in 'a' for iteration 54:
Stored Text Values in 'a' for iteration 55:
Stored Text Values in 'a' for iteration 56:
```

```
time_dict[16] #random check
```

['The U.S. and U.K. launched strikes attacking multiple Houthi rebel targets in Yemen on Thursday, Jan. 11, marking a significant escalation in the broader conflict of the Middle East. In October, shortly after the start of theIsrael-Hamas war, the Houthis began launching missile attacks at cargo ships passing through the Red Sea, headed towards Israeli ports.',
 'Here's what you need to know about Yemen's Houthi rebels and their place in the current conflict.',
 'The Houthi rebels are one of two main factions controlling territory amidst Yemen's ongoing civil war. Currently, they control the Western coast of the country, including its capital city, Sana'a.',
 'The Houthis initially began as a cultural revivalist movement in the 1990s for the Zaydi sect of Islam which, as of 2022, was practiced by approximately35% of the Yemeni population. According to Stacey Philbrick Yadav, the chair of international relations at Hobart and William Smith Colleges and author ofYemen in the Shadow of Transition, many Zaydi muslims were becoming increasingly frustrated with the growing prominence of Saudi-Arabian Salafi Islam in the country, which they felt repressed their Zaydi cultural and religious heritage.',
 'Additionally, the Houthis resented what they perceived to be widespread corruption and mismanagement in the Yemeni government throughout the 2000s. This led the Houthis to launch several insurgencies against the Yemeni government between 2004 and 2010.',
 'In 2011, during the Arab Spring, the grievances of the Yemeni population reached a boiling point, and mass protests erupted calling for the resignation of President Ali Abdullah Saleh, who had been ruling the country for over thirty years. After Saleh resigned, Saudi Arabia backed Abdrabbuh Mansur Hadi as the new leader of the country. The Houthis resented this choice, and launched a military campaign against the new government, which led to a civil war that is still ongoing.',
 'Their conflict with the Saudi-backed government has led the Houthis to become more closely allied with Iran, from which they receive some support. Nevertheless, some experts say it's not entirely accurate to call the Houthi rebels a direct proxy.',
 '"They do have a relationship with and support from Iran, but are not a straightforward proxy of Iranian interests. They have their own locally defined interests and so I think that their actions in the past two months have reflected that,"Philbrick Yadav told TIMEin December.',
 'Read More:How Congress Is Reacting to Biden's Military Attack on the Houthis in Yemen',
 'In the Arab world, the Palestinian cause is extremely popular and often a signifier of progressive values. By attacking Israeli ships, the Houthis may expand their support base across Yemen and the Arab world. Additionally, it is believed the group is hoping to sabotage theSaudi normalization with Israel, which had been in the works.',
 'Over80% of all internationally traded goods are transported via cargo ships, since air travel is a much less cost effective way to transport large items or huge amounts of goods. Ships must travel through the Red Sea in order to access the Suez Canal, which is the only waterway that allows for direct passage between Europe and Asia.',
 'Otherwise, the ships traveling between Europe and Asia must journey around the African continent, which can add over 30 days in travel time. Due to the attacks by Houthi rebels, insurance prices on ships have risen dramatically and many shipping companies have chosen to take the longer routes, as a safety precaution. This is expected toincrease the pricesof many consumer goods, from clothing to coffee.',
 '"These strikes are in direct response to unprecedented Houthi attacks against international maritime vessels in the Red Sea—including the use of anti-ship ballistic missiles for the first time in history," President Bidensaid in a statement. "I will not hesitate to direct further measures to protect our people and the free flow of international commerce as necessary."']

Start coding or generate with AI.

⌄ Telegraph

```
#one random page
url = "https://www.telegraphindia.com/world/page-53"
source = requests.get(url).text
soup = BeautifulSoup(source, 'html.parser')
```

```
print(soup.prettify())
```

```html
<!DOCTYPE html>
<html lang="en">
 <head>
  <script>
   (function (w, d, s, l, i) {
        w[l] = w[l] || [];
        w[l].push({
            'gtm.start':
                new Date().getTime(), event: 'gtm.js'
        });
        var f = d.getElementsByTagName(s)[0],
            j = d.createElement(s), dl = l != 'dataLayer' ? '&l=' + l : '';
        j.async = true;
        j.src =
            'https://www.googletagmanager.com/gtm.js?id=' + i + dl;
        f.parentNode.insertBefore(j, f);
    })(window, document, 'script', 'dataLayer', 'GTM-T9HRQZR');


    var THIS_PAGE_NAME = 'sectionpage';
  </script>
  <meta charset="utf-8"/>
  <meta content="IE=edge" http-equiv="X-UA-Compatible"/>
  <meta content="width=device-width, initial-scale=1.0" name="viewport"/>
  <title>
   World News Today - Latest International News, World News Headlines - Telegraph India
  </title>
  <meta content="World News- Get latest international top news headlines from around the world. Find the breaking news, top world stor
" name="description">
  <meta content="" name="keywords">
   <meta content="https://www.telegraphindia.com/world" property="og:url">
    <meta content="World News Today - Latest International News, World News Headlines - Telegraph India" property="og:title"/>
    <meta content="website" property="og:type"/>
    <meta content="World News- Get latest international top news headlines from around the world. Find the breaking news, top world s
" property="og:description">
    <meta content="https://www.telegraphindia.com/images/tt-og-v-1.png" property="og:image"/>
    <meta content="534248160342700" property="fb:app_id"/>
    <meta content="https://www.telegraphindia.com/images/tt-og-v-1.png" name="twitter:image">
   <meta content="summary_large_image" name="twitter:card"/>
   <meta content="https://www.telegraphindia.com/world" name="twitter:url"/>
   <meta content="@ttindia" name="twitter:site"/>
   <meta content="@ttindia" name="twitter:creator"/>
   <meta content="World News Today - Latest International News, World News Headlines - Telegraph India" name="twitter:title"/>
   <meta content="World News- Get latest international top news headlines from around the world. Find the breaking news, top world
" name="twitter:description"/>
   <meta content="index, follow" name="robots"/>
   <meta content="fKpCKm4JyeZ6rB3vWUu8mnf8G2s7YBrdizSWDEvPxGA" name="google-site-verification">
    <meta content="origin" name="referrer"/>
    <meta content="16C1B4A8FD9A29E3D5AFC3EF251C77A8" name="msvalidate.01">
     <link href="/images/favicon.png" rel="icon"/>
     <link href="/manifest.json" rel="manifest"/>
     <link href="https://www.telegraphindia.com/world" rel="canonical">
      <link href="https://fonts.googleapis.com" rel="preconnect"/>
      <link href="https://fonts.googleapis.com" rel="dns-prefetch"/>
      <link href="//fonts.googleapis.com" rel="preconnect"/>
      <link href="//fonts.googleapis.com" rel="dns-prefetch"/>
      <link href="//www.googletagmanager.com" rel="preconnect"/>
```

```python
# Assuming 'soup' contains your BeautifulSoup object

# Find all 'a' tags with href starting with '/world' and containing 'cid'
links = soup.find_all('a', href=lambda href: href and re.match(r'^/world', href) and 'cid' in href)

# Initialize an empty list to store the complete URLs
complete_urls = []

# Prepend "https://www.telegraphindia.com" and add the complete URLs to the list
for link in links:
    complete_url = 'https://www.telegraphindia.com' + link['href']
    complete_urls.append(complete_url)

# Print or further process the list of complete URLs
print(complete_urls)
```

```
['https://www.telegraphindia.com/world/israeli-forces-and-hamas-fighters-engage-in-house-to-house-gunbattles-in-khan-younis/cid/1985099'
```

```
print("Number of URLs:", len(complete_urls)) #single page contains 16-18 urls
```

```
print( Number or URLS: , ien(complete_urls)) #single page contains 16-18 urls
```

```
        Number of URLs: 19
```

```
#check one random url
url = "https://www.telegraphindia.com/world/dubai-india-proposes-to-host-un-climate-conference-in-2028-launches-green-credit-initiative/cid/19
source = requests.get(url).text
soup = BeautifulSoup(source, 'html.parser')
```

```
print(soup.prettify())
```

```html
    <!DOCTYPE html>
    <html lang="en">
     <head>
      <script>
       (function (w, d, s, l, i) {
            w[l] = w[l] || [];
            w[l].push({
                'gtm.start':
                    new Date().getTime(), event: 'gtm.js'
            });
            var f = d.getElementsByTagName(s)[0],
                j = d.createElement(s), dl = l != 'dataLayer' ? '&l=' + l : '';
            j.async = true;
            j.src =
                'https://www.googletagmanager.com/gtm.js?id=' + i + dl;
            f.parentNode.insertBefore(j, f);
       })(window, document, 'script', 'dataLayer', 'GTM-T9HRQZR');


       var THIS_PAGE_NAME = 'regular';
      </script>
      <meta charset="utf-8"/>
      <meta content="IE=edge" http-equiv="X-UA-Compatible"/>
      <meta content="width=device-width, initial-scale=1.0" name="viewport"/>
      <title>
       Dubai | Dubai: India proposes to host UN climate conference in 2028, launches Green Credit Initiative - Telegraph India
      </title>
      <meta content="At COP28, PM Modi proposes Green Credit Initiative which focuses on creating carbon sinks through people's participat
       <meta content="Dubai, PM Narendra Modi, COP28, Climate Summit" name="keywords">
        <meta content="Dubai, PM Narendra Modi, COP28, Climate Summit" name="news_keywords"/>
        <meta content="https://www.telegraphindia.com/world/dubai-india-proposes-to-host-un-climate-conference-in-2028-launches-green-cred
         <meta content="India at COP28 summit in Dubai; PM Modi launches 'Green Credit Initiative' at UN climate talks, proposes to host 2
          <meta content="article" property="og:type"/>
          <meta content="At COP28, PM Modi proposes Green Credit Initiative which focuses on creating carbon sinks through people's partic
          <meta content="https://assets.telegraphindia.com/telegraph/2023/Dec/1701428454_modi-53.jpg" property="og:image"/>
          <meta content="534248160342700" property="fb:app_id"/>
          <meta content="https://assets.telegraphindia.com/telegraph/2023/Dec/1701428454_modi-53.jpg" name="twitter:image">
           <meta content="summary_large_image" name="twitter:card"/>
           <meta content="https://www.telegraphindia.com/world/dubai-india-proposes-to-host-un-climate-conference-in-2028-launches-green-
           <meta content="@ttindia" name="twitter:site"/>
           <meta content="@ttindia" name="twitter:creator"/>
           <meta content="India at COP28 summit in Dubai; PM Modi launches 'Green Credit Initiative' at UN climate talks, proposes to host
           <meta content="At COP28, PM Modi proposes Green Credit Initiative which focuses on creating carbon sinks through people's parti
           <meta content="index, follow" name="robots"/>
           <meta content="max-image-preview:large" name="robots"/>
           <meta content="fKpCKm4JyeZ6rB3vWUu8mnf8G2s7YBrdizSWDEvPxGA" name="google-site-verification">
            <meta content="origin" name="referrer"/>
            <meta content="16C1B4A8FD9A29E3D5AFC3EF251C77A8" name="msvalidate.01">
             <link href="/images/favicon.png" rel="icon"/>
             <link href="/manifest.json" rel="manifest"/>
             <link href="https://www.telegraphindia.com/world/dubai-india-proposes-to-host-un-climate-conference-in-2028-launches-green-cr
              <link href="https://www.telegraphindia.com/amp/world/dubai-india-proposes-to-host-un-climate-conference-in-2028-launches-gre
              <link href="https://fonts.googleapis.com" rel="preconnect"/>
              <link href="https://fonts.googleapis.com" rel="dns-prefetch"/>
              <link href="//fonts.googleapis.com" rel="preconnect"/>
              <link href="//fonts.googleapis.com" rel="dns-prefetch"/>
              <link href="//www.googletagmanager.com" rel="preconnect"/>
```

```
# Find all <p> tags and extract the text content
paragraphs = soup.find_all('p')
```

```
# Print the text content
for paragraph in paragraphs:
    print(paragraph.get_text())
```

```
    Prime Minister Narendra Modi speaks during a plenary session at the COP28 U.N. Climate Summit.
                                                   PTI picture.

    Asserting that the world does not have much time to correct the mistakes of the last century, Prime Minister Narendra Modi on Friday ann
```

Addressing the high-level segment for heads of state and governments during the UN climate conference (COP28) here, Modi called for a pr
"It focuses on creating carbon sinks through people's participation and I invite all of you to join this initiative," he said, stressing
This initiative is similar to the Green Credit Programme, notified domestically in October. It is an innovative market-based mechanism d
Asserting that India has presented a great example to the world of striking balance between development and environment conservation, th
Modi was the only leader to join COP28 President Sultan Al Jaber on the stage along with the UN Climate Change Executive Secretary Simon
"Over the past century, a small section of humanity has indiscriminately exploited nature. However, the entire humanity is paying the pr
"Thinking only about our own interests will only lead the world into darkness," the prime minister added.
Modi's statement came in the context that the poor and developing nations bear the brunt of extreme climate events such as floods, droug
The Prime Minister called for maintaining a balance between mitigation and adaptation and said that energy transition across the world m
Modi has been championing the Lifestyle for Environment (LiFE movement), which he had announced at the Glasgow COP in 2021), urging coun
Citing a study by the International Energy Agency, Modi said, "This approach (LiFE) can reduce carbon emissions by 2 billion tonnes." He
"We shall cooperate with each other and shall support each other. We need to give all developing countries our fair share in the global
If India's proposal to host COP33 is accepted, it would be the next big global conference in the country after the G20 Summit earlier th
India hosted COP8 in New Delhi in 2002 where countries adopted the Delhi Ministerial Declaration which called for efforts by developed c
Climate science defines carbon budget as the amount of greenhouse gases that can be emitted for a given level of global warming (1.5 deg
Developed countries have already consumed more than 80 per cent of the global carbon budget, leaving developing and poor countries with
"India is one of the very few economies in the world that is on track to achieve its NDC targets," he said.
India achieved its emissions intensity-related targets 11 years ahead of the committed time frame and non-fossil fuel targets nine years
"And India did not just stop there, we remain ambitious," he said.
The country aims to reduce emissions intensity of gross domestic product by 45 per cent by 2030 from 2005 levels and achieve 50 per cent
It has also committed to become a net zero economy by 2070.
As part of its G20 Presidency this year, India drew consensus from the world's major economies for a Green Development Pact seeking to b
The Pact shifted the conversations from the billions to the trillions needed for the energy transition. It noted that developing countri
According to an official release, prohibitory orders under Section 144 of the CrPC, which was earlier made effective from 9 am to 10pm o

∨ Iteratting over pages to get links

```python
import requests
from bs4 import BeautifulSoup
import re

telegraph_dict= {}
i = 0

# Assuming filtered_href_values is a list of href values
for idx in range(1, 53):  # as it contains till December i.e one month
    url1 = f"https://www.telegraphindia.com/world/page-{idx}"
    source = requests.get(url1).text
    soup = BeautifulSoup(source, 'html.parser')

    # Find all 'a' tags with href starting with "/" and followed by a number
    links = soup.find_all('a', href=lambda href: href and re.match(r'^/world', href) and 'cid' in href)

    # Prepend "https://time.com" and print the complete URLs
    # Initialize an empty list to store the complete URLs
    complete_urls = []

    # Prepend "https://time.com" and add the complete URLs to the list
    for link in links:
        complete_url = 'https://www.telegraphindia.com' + link['href']
        complete_urls.append(complete_url)

    # Extract the URLs from the list
    urls = [
        item for item in complete_urls
    ]

    # Print the URLs with their indices
    for index, url in enumerate(urls, start=1):
        print(f"{index}. {url}")
        telegraph_dict[i] = url
        i += 1
else:
    print("Target script not found.")
```

1. https://www.telegraphindia.com/world/pakistan-media-regulatory-body-ordered-to-provide-unrestricted-media-coverage-to-imran-khan-ot
2. https://www.telegraphindia.com/world/united-states-navy-destroyer-shoots-down-houthi-claimed-missile-in-red-sea/cid/1997429
3. https://www.telegraphindia.com/world/10000-indian-construction-workers-to-reach-israel-soon-in-batches-starting-next-week/cid/19974
4. https://www.telegraphindia.com/world/main-opposition-in-maldives-accuses-president-mohamed-muizzu-of-protecting-criminals-after-att
5. https://www.telegraphindia.com/world/coroners-office-suspect-no-foul-play-in-indian-origin-purdue-university-students-death-report/
6. https://www.telegraphindia.com/world/israeli-soldiers-dressed-in-medical-gear-kill-a-hamas-commander-two-others-at-a-west-bank-hosp
7. https://www.telegraphindia.com/world/president-joe-biden-says-hes-decided-on-response-to-killing-of-3-united-states-troops-plans-to
8. https://www.telegraphindia.com/world/former-pakistan-prime-minister-imran-khan-and-his-wife-bushra-bibi-sentenced-to-14-years-in-ja
9. https://www.telegraphindia.com/world/united-states-announces-overhaul-of-its-annual-lottery-for-h-1b-visas-to-reduce-potential-for-
10. https://www.telegraphindia.com/world/united-states-launches-five-week-pilot-programme-to-renew-h-1b-visas-domestically/cid/199732!

11. https://www.telegraphindia.com/world/indian-origin-couple-jailed-for-33-years-in-united-kingdom-for-exporting-over-half-tonne-coca
12. https://www.telegraphindia.com/world/us-elections-2024-nikki-haley-coins-new-slogan-make-america-normal-again/cid/1997324
13. https://www.telegraphindia.com/world/h-1b-visa-application-submission-process-for-fiscal-year-2025-to-begin-march-6/cid/1997322
14. https://www.telegraphindia.com/world/israeli-prime-minister-benjamin-netanyahu-rules-out-gaza-pullout-and-release-of-thousands-of-
15. https://www.telegraphindia.com/world/undercover-attack-kills-3-militants-israeli-troops-disguised-as-women-storm-hospital/cid/1997
16. https://www.telegraphindia.com/world/lets-save-agriculture-french-farmers-block-highways-to-paris-amid-protests-against-government
17. https://www.telegraphindia.com/world/pakistan-former-pm-imran-khan-sentenced-10-years-in-prison-a-week-before-polls/cid/1997253
18. https://www.telegraphindia.com/world/israeli-pm-benjamin-netanyahu-rules-out-gaza-pullout-or-release-of-thousands-of-militants/cid
19. https://www.telegraphindia.com/world/maldives-government-approaches-supreme-court-over-the-revision-of-rules-that-enable-oppositio
20. https://www.telegraphindia.com/world/president-emmanuel-macron-launches-special-french-language-programme-for-indian-students/cid/
1. https://www.telegraphindia.com/world/china-will-influence-our-neighbourhood-india-shouldnt-be-scared-of-competition-external-affair
2. https://www.telegraphindia.com/world/china-threat-is-an-opportunity-for-india-us-relationship-to-make-genuine-free-trade-agreement-
3. https://www.telegraphindia.com/world/suspended-prison-sentences-for-10-indian-fishermen-for-poaching-in-lanka-waters/cid/1997162
4. https://www.telegraphindia.com/world/pakistani-court-sentences-former-prime-minister-imran-khan-to-10-years-for-exposing-official-s
5. https://www.telegraphindia.com/world/universe-chants-jai-shree-ram-aerial-banner-thrills-indian-americans-in-houston/cid/1997132
6. https://www.telegraphindia.com/world/indian-origin-couple-convicted-in-united-kingdom-for-exporting-over-half-tonne-cocaine-to-aust
7. https://www.telegraphindia.com/world/elon-musks-neuralink-implants-brain-chip-in-first-human/cid/1997115
8. https://www.telegraphindia.com/world/israel-expresses-gratefulness-for-indias-friendship-and-support-discusses-regional-security-is
9. https://www.telegraphindia.com/world/indian-consulate-in-atlanta-strongly-condemns-fatal-attack-on-student-from-india/cid/1997102
10. https://www.telegraphindia.com/world/indian-american-congressman-ro-khanna-campaigns-for-us-president-joe-biden-in-south-carolina-
11. https://www.telegraphindia.com/world/united-states-doesnt-seek-another-war-white-house-on-drone-strike-by-iran-backed-group/cid/19
12. https://www.telegraphindia.com/world/united-states-national-security-advisor-jake-sullivan-and-his-chinese-counterpart-wang-yi-ho
13. https://www.telegraphindia.com/world/indian-student-in-united-states-who-went-missing-at-purdue-university-in-indiana-confirmed-de
14. https://www.telegraphindia.com/world/over-50-per-cent-of-pakistan-poll-stations-declared-sensitive-ahead-of-february-8-general-ele
15. https://www.telegraphindia.com/world/iran-denies-accusations-of-its-involvement-in-jordan-strike-that-killed-three-americans/cid/1
16. https://www.telegraphindia.com/world/israeli-intelligence-dossier-showcases-involvement-of-un-staffs-in-militant-groups/cid/199704
17. https://www.telegraphindia.com/world/enemy-drone-responsible-for-three-americans-death-and-injuries-mistaken-as-us-drone/cid/19976
1. https://www.telegraphindia.com/world/border-issue-with-india-to-be-resolved-through-political-consensus-says-nepal-foreign-minister
2. https://www.telegraphindia.com/world/china-invites-bangladesh-prime-minister-sheikh-hasina-for-official-visit-to-beijing/cid/199699
3. https://www.telegraphindia.com/world/maldives-parliament-approves-only-one-cabinet-member-among-four-amid-oppositions-red-whip/cid/
4. https://www.telegraphindia.com/world/opposition-party-in-maldives-readies-to-move-impeachment-motion-against-president-mohamed-muiz
5. https://www.telegraphindia.com/world/indian-warship-deployed-along-somalias-east-coast-responds-to-iranian-flagged-vessels-distress
6. https://www.telegraphindia.com/world/pakistan-iran-agree-to-expand-security-cooperation-after-missile-strikes/cid/1996943
7. https://www.telegraphindia.com/world/militants-located-in-border-areas-of-pakistan-and-iran-led-and-supported-by-third-countries-ir
8. https://www.telegraphindia.com/world/indian-pleads-guilty-in-us-to-running-dark-web-drug-enterprise-forfeits-150-mn-in-cryptocurren
9. https://www.telegraphindia.com/world/iran-says-it-has-no-link-to-drone-strike-in-jordan-that-killed-united-states-soldiers/cid/1996
10. https://www.telegraphindia.com/world/indian-student-in-united-states-hammered-to-death-by-homeless-drug-addict/cid/1996854
11. https://www.telegraphindia.com/world/joe-biden-vows-to-retaliate-after-strike-against-us-forces-in-jordan-kills-3-service-members/
12. https://www.telegraphindia.com/world/climate-activists-throw-soup-at-glass-protecting-mona-lisa-in-paris-as-french-farmers-aim-to-
13. https://www.telegraphindia.com/world/iran-executes-four-people-it-says-are-linked-to-israeli-intelligence-report/cid/1996849
14. https://www.telegraphindia.com/world/joe-biden-says-united-states-shall-respond-after-drone-strike-by-iran-backed-group-kills-3-us
15. https://www.telegraphindia.com/world/clashes-erupt-in-maldives-parliament-during-crucial-vote-on-president-mohamed-muizzus-cabinet
16. https://www.telegraphindia.com/world/un-chief-urges-major-donors-not-to-suspend-aid-to-its-agency-for-palestinians/cid/1996754
17. https://www.telegraphindia.com/world/after-two-years-of-bloody-fighting-ukraine-wrestles-with-conscription/cid/1996753
18. https://www.telegraphindia.com/world/significant-number-of-weapons-used-by-hamas-come-from-unlikely-source-israel/cid/1996752
19. https://www.telegraphindia.com/world/us-helmed-gaza-deal-in-the-works-hostage-pact-could-halt-war-for-two-months/cid/1996751
20. https://www.telegraphindia.com/world/three-us-service-members-killed-34-wounded-in-jordan-drone-attack-linked-to-iran/cid/1996745

⌄ Filter URLs containing both "Hamas" and "Israel"

```
# Filter URLs containing both "Hamas" and "Israel" (case-insensitive)
filtered_urls = {key: value for key, value in telegraph_dict.items() if 'hamas' in value.lower() or 'israel' in value.lower()}

# Create a new dictionary with consecutive integer keys
filtered_urls_dict = {i + 1: url for i, url in enumerate(filtered_urls.values())}

# Print the filtered URLs with new keys
for key, value in filtered_urls_dict.items():
    print(f"{key}. {value}")
```

1. https://www.telegraphindia.com/world/10000-indian-construction-workers-to-reach-israel-soon-in-batches-starting-next-week/cid/19974
2. https://www.telegraphindia.com/world/israeli-soldiers-dressed-in-medical-gear-kill-a-hamas-commander-two-others-at-a-west-bank-hosp
3. https://www.telegraphindia.com/world/israeli-prime-minister-benjamin-netanyahu-rules-out-gaza-pullout-and-release-of-thousands-of-
4. https://www.telegraphindia.com/world/undercover-attack-kills-3-militants-israeli-troops-disguised-as-women-storm-hospital/cid/1997
5. https://www.telegraphindia.com/world/israeli-pm-benjamin-netanyahu-rules-out-gaza-pullout-or-release-of-thousands-of-militants/cid/
6. https://www.telegraphindia.com/world/israel-expresses-gratefulness-for-indias-friendship-and-support-discusses-regional-security-is
7. https://www.telegraphindia.com/world/israeli-intelligence-dossier-showcases-involvement-of-un-staffs-in-militant-groups/cid/1997043
8. https://www.telegraphindia.com/world/iran-executes-four-people-it-says-are-linked-to-israeli-intelligence-report/cid/1996849
9. https://www.telegraphindia.com/world/significant-number-of-weapons-used-by-hamas-come-from-unlikely-source-israel/cid/1996752
10. https://www.telegraphindia.com/world/un-fires-12-gaza-strip-employees-and-begins-investigation-into-them-after-accusations-by-isra
11. https://www.telegraphindia.com/world/israeli-prime-minister-benjamin-netanyahu-declines-charge-of-genocide-as-outrageous/cid/19964
12. https://www.telegraphindia.com/world/international-court-of-justice-omits-ceasefire-directive-for-israels-actions-in-gaza/cid/1996
13. https://www.telegraphindia.com/world/families-of-israeli-hostages-block-emergency-aids-seeking-return-of-their-loved-ones/cid/1996
14. https://www.telegraphindia.com/world/israeli-military-issues-evacuation-order-of-homeless-palestinians-from-south-gaza/cid/1996384
15. https://www.telegraphindia.com/world/world-court-orders-israel-to-prevent-acts-of-genocide-in-gaza-fails-to-order-ceasefire/cid/19
16. https://www.telegraphindia.com/world/israel-braces-for-world-court-ruling-focuses-attack-on-south-gaza/cid/1996315
17. https://www.telegraphindia.com/world/20-killed-in-israeli-strike-on-crowd-seeking-aid-says-gazas-health-ministry/cid/1996250
18. https://www.telegraphindia.com/world/israel-fends-off-charge-of-genocide-jerusalem-declassifies-cabinet-decisions/cid/1996249

19. https://www.telegraphindia.com/world/mass-casualties-israeli-tanks-strike-united-nations-compound-sheltering-palestinians/cid/1996
20. https://www.telegraphindia.com/world/china-rejects-supplying-weapons-to-hamas-says-always-adopted-responsible-attitude-to-defence-
21. https://www.telegraphindia.com/world/plan-for-one-month-gaza-truce-makes-progress-israel-hits-khan-younis/cid/1995907
22. https://www.telegraphindia.com/world/israel-hamas-war-risk-of-pockets-of-famine-in-gaza-says-world-food-programme/cid/1995806
23. https://www.telegraphindia.com/world/israeli-army-in-deadliest-day-of-war-24-soldiers-killed-in-gaza-including-21-in-single-blast/c
24. https://www.telegraphindia.com/world/israeli-forces-storm-hospital-as-khan-younis-hit-by-bloodiest-fighting-of-2024/cid/1995551
25. https://www.telegraphindia.com/world/saudi-arabias-top-diplomat-says-it-wont-recognise-israel-without-a-path-to-a-palestinian-stat
26. https://www.telegraphindia.com/world/israel-hamas-war-us-official-heads-to-middle-east-for-talks-on-hostages/cid/1995381
27. https://www.telegraphindia.com/world/palestinian-death-toll-in-gaza-surpasses-25000-with-no-end-in-sight-to-israel-hamas-war/cid/1
28. https://www.telegraphindia.com/world/soldiers-uncover-gaza-tunnel-hamas-once-used-to-hold-hostages-israeli-military/cid/1995264
29. https://www.telegraphindia.com/world/israeli-women-fight-on-gaza-front-line-for-first-time-since-establishment-of-israel-in-1948/c
30. https://www.telegraphindia.com/world/israel-pounds-targets-across-gaza-strip-while-planes-drop-leaflets-to-locate-captives/cid/199
31. https://www.telegraphindia.com/world/israels-progress-is-limited-in-eradicating-hamas-and-also-liberating-israeli-hostages-still-i
32. https://www.telegraphindia.com/world/israeli-strike-on-syria-destroys-building-used-by-iranian-paramilitary-revolutionary-guard-ki
33. https://www.telegraphindia.com/world/israeli-strike-on-syrias-damascus-kills-four-iranian-revolutionary-guards/cid/1995162
34. https://www.telegraphindia.com/world/israeli-strike-on-damascus-kills-iranian-guards-official-report/cid/1995092
35. https://www.telegraphindia.com/world/israel-war-cabinet-member-bares-rift-urges-for-ceasefire-with-hamas-and-free-hostages/cid/199
36. https://www.telegraphindia.com/world/north-gaza-residents-walk-freely-as-israel-dials-back-its-intense-military-campaign/cid/19949
37. https://www.telegraphindia.com/world/patients-and-shelter-seekers-flee-southern-gaza-hospital-siege-by-israeli-forces/cid/1994723
38. https://www.telegraphindia.com/world/israel-stepped-up-its-assault-on-khan-younis-in-southern-gaza-damaging-field-hospital-in-city
39. https://www.telegraphindia.com/world/israel-and-hamas-reach-deal-to-allow-medications-to-be-delivered-to-israeli-hostages-and-pale
40. https://www.telegraphindia.com/world/at-davos-blinken-calls-pathway-to-palestinian-state-a-necessity-for-israeli-security/cid/1994
41. https://www.telegraphindia.com/world/end-israel-hamas-war-to-stop-houthi-raids-qatars-prime-minister-sheikh-mohammed-bin-abdulrahm
42. https://www.telegraphindia.com/world/new-hamas-video-shows-bodies-of-two-israeli-hostages-killed-in-captivity/cid/1994276
43. https://www.telegraphindia.com/world/irans-revolutionary-guards-attack-israels-espionage-centers-in-iraq-state-media-says/cid/1994
44. https://www.telegraphindia.com/world/israeli-soccer-player-sagiv-jehezkel-leaves-turkey-on-monday-after-detention/cid/1994008
45. https://www.telegraphindia.com/world/israeli-forces-bombard-targets-in-south-north-and-centre-of-gaza/cid/1994006
46. https://www.telegraphindia.com/world/white-house-says-its-the-right-time-for-israel-to-scale-back-gaza-war-as-fighting-hits-100-da
47. https://www.telegraphindia.com/world/students-sue-harvard-university-over-hatred-of-jews-post-hamas-attack-on-israel/cid/1993430
48. https://www.telegraphindia.com/world/israel-taps-top-legal-minds-including-a-holocaust-survivor-to-battle-genocide-claim-at-world-
49. https://www.telegraphindia.com/world/israel-will-defend-itself-at-united-nations-top-court-against-allegations-of-genocide-against
50. https://www.telegraphindia.com/world/un-court-starts-hearings-on-israels-unstoppable-actions-on-gaza-for-hamas-removal/cid/1993193
51. https://www.telegraphindia.com/world/un-top-court-hears-genocide-allegation-as-israel-focuses-fighting-in-central-gaza/cid/1993137
52. https://www.telegraphindia.com/world/top-united-nations-court-opens-hearings-in-gaza-genocide-case-against-israel/cid/1993072
53. https://www.telegraphindia.com/world/at-world-court-israel-to-confront-south-africas-accusations-it-committed-genocide-in-the-gaza
54. https://www.telegraphindia.com/world/israel-south-africa-trade-barbs-on-eve-of-genocide-case/cid/1992920
55. https://www.telegraphindia.com/world/israeli-vehicles-drove-over-palestinian-killed-in-army-raid-camera-footage-shows/cid/1992915
56. https://www.telegraphindia.com/world/united-states-secretary-of-state-antony-blinken-calls-on-israel-to-work-with-moderate-palesti

```python
tg_dict = {}

# Assuming filtered_urls_dict is a dictionary with indices as keys and URLs as values

for idx, url in filtered_urls_dict.items():
    # Fetch HTML content
    source = requests.get(url).text
    soup = BeautifulSoup(source, 'html.parser')

    # Initialize a list to store text values for the current iteration
    a = []

    # Find the script tag with the id '__NEXT_DATA__'
    paragraphs = soup.find_all('p')

    # Print the text content
    for paragraph in paragraphs:
        print(paragraph.get_text())
        try:
            text_content = paragraph.get_text()

            # Append the "text" values to the list
            if isinstance(text_content, str):
                a.append(text_content)

        except json.JSONDecodeError as e:
            print(f"Error decoding JSON: {e}")

    # Now, 'a' contains the stored text values for the current iteration
    print(f"Stored Text Values in 'a' for iteration {idx}:")
    for text_value in a:
        tg_dict.setdefault(idx, []).append(text_value)
```

Representational image.

Shutterstock

With Israel's construction sector suffering a severe manpower crisis post the October 7 conflict with Hamas, some 10,000 workers from
These 10,000 workers will reach in batches of 700 to 1,000 a week, a source in the Israel's Builders Association (IBA) told PTI here
With Israel's latest conflict with Hamas in Gaza a little short of four months, and a ban on the entry of Palestinian workers and the
Following the conflict, Israel has banned the entry of Palestinian workers. The departure of several thousand other foreign workers,

Israeli business daily The Calcalist in a report in Hebrew last week said that the quota of foreign manpower for the construction ind
The IBA source confirmed to PTI that the details in the Calcalist report were correct.
When asked about the arrival of the first batch of workers, the source said, "We hope they'll come next week."   The workers
The IBA is looking to hire workers also from other countries such as Mexico, Kenya and Malawi. Screening of workers is said to have b
"To date, out of about 8,000 workers examined, about 5,500 have been found suitable for work in Israel – the vast majority Indian," T
"The Indian construction workers are at a high professional level. Many of them were previously employed in the Gulf countries, so th
The selection process is led by IBA's CEO Igal Slovik and Izchak Gurvitz, who heads the association's division dealing with workers'
Israel's Ministry of Housing and Construction is currently said to be working on a proposal to increase by another 10,000 the hiring
This will increase the total quota of foreign workers employed in the industry from 50,000 workers today to 60,000.
Israeli Prime Minister Benjamin Netanyahu during a telephonic conversation with his Indian counterpart Narendra Modi in December had
The discussions then revolved around bringing in almost 160,000 people in various sectors.
About 18,000 Indians are working in Israel, mostly as caregivers. Most of them decided to stay back in Israel and did not leave the c
Israel and India also inked an agreement in May last year during the then Israeli Foreign Minister Eli Cohen's visit to New Delhi tha
A statement released by the Israeli Foreign Ministry then said that 34,000 workers will be engaged in the construction field and anot
Except for the headline, this story has not been edited by The Telegraph Online staff and has been published from a syndicated feed.
According to an official release, prohibitory orders under Section 144 of the CrPC, which was earlier made effective from 9 am to 10pm
Stored Text Values in 'a' for iteration 1:
Ambulances evacuating premature babies who had been in intensive care at Al-Shifa Hospital head to the border crossing with Egypt in F
                                                Samar Abu Elouf/The New York Times

One of the Israeli soldiers wore medical scrubs, another a white coat and surgical mask. Their team swept into the West Bank hospital
Fifteen minutes later they were gone.
The raid on Tuesday took place as the Israeli military battled Hamas on multiple fronts: with the dramatic operation in the West Bank,
Israel's latest efforts in the nearly four-month war came amid a renewed push by multiple peace brokers, including the United States,
But even as those talks continued behind closed doors, Prime Minister Benjamin Netanyahu asserted anew that Israel would continue figh
Israeli forces have tried to strike Hamas leaders and their allies both inside Gaza and outside the territory. This month, Hamas blame
Israeli forces have escalated efforts against Palestinian militant activity in the West Bank, arresting more than 2,980 Palestinians s
Surveillance video released by the Palestinian Authority Health Ministry showed several gunmen in civilian garb — including one dresse
They went to a room where the Hamas commander, Mohammad Jalamneh, 27, was visiting a friend and shot him and two other men dead, said
"They assassinated these three people, including a patient," Sbeihat said in a phone interview.
In a statement, Hamas mourned Jalamneh, describing him as a leader in the Al-Qassam Brigades, the faction's armed wing. A local militi
Nazzal said that Basil Ghazawi had been receiving treatment in the hospital's rehabilitation ward since late October, when he was para
The raid at the hospital raises questions under international law, experts said. Hospitals require special protection and respect unde
If Ghazawi was paralyzed and incapable of defending himself, he should not have been subject to attack under international law, Liebli
The Israeli military said all three men killed in the raid had been involved in militant activities, including attacks against Israeli
Israel's raids in the West Bank escalated as violence spiraled out of the war in Gaza, which began after the Hamas-led Oct. 7 attacks
Israeli ground troops moved into northern Gaza in late October, and have since turned their focus to the territory's south, where fier
A major part of the challenge, Israeli officials say, is the network of tunnels that Hamas has built over years beneath Gaza.
The military's statement on Tuesday that it was flooding some of the tunnels was its first public acknowledgment of a contentious proj
The military said it had begun pumping seawater to flush militants out of the labyrinthine underground network, which Hamas has used t
Senior Israeli military officials, who spoke on the condition of anonymity to discuss intelligence matters, estimated this month that
On the whole, the flooding project has had limited success, the officials said. Large volumes of water have been pumped in, but many c
The New York Times News Service
Advocate Yadav said the district administration will make arrangements for performing puja and it will be facilitated by the Kashi Vis
Stored Text Values in 'a' for iteration 2:
Benjamin Netanyahu

                                                File image

Israeli Prime Minister Benjamin Netanyahu said on Tuesday that there will be no military withdrawal from Gaza or the release of thousa
During an event in the occupied West Bank, he once again vowed that the war will not end without "absolute victory" over Hamas. "We wi

tg_dict[11]

['Israeli Prime Minister Benjamin Netanyahu\n   File image\n',
 'Israeli Prime Minister Benjamin Netanyahu welcomed a decision by the World Court on Friday not to order a ceasefire in the war in Gaza but rejected a charge of genocide as "outrageous", and said it would continue to defend itself.',
 'The World Court ordered Israel to take action to prevent acts of genocide as it wages war against Hamas militants in the Gaza Strip, but it stopped short of calling for an immediate ceasefire.',
 "On the court's demand that Israel try to contain death and damage in its military offensive in the tiny coastal enclave, Israel says it seeks to protect citizens in Gaza and also accuses Hamas of using Palestinians as human shields — an accusation that Hamas denies.",
 "It says it must have the right to defend itself after Hamas' cross-border attack on October 7 that killed 1,200 people.',
 'Responding to a case brought by South Africa at the International Court of Justice in The Hague, Netanyahu said Israel's commitment to international law was "unwavering".',
 'But he added in a statement: "Like every country, Israel has a basic right to defend itself ... The World Court in the Hague justly rejected the outrageous demand to deprive us of this right."',
 'He appeared to be referring to the fact that the ICJ did not call for an immediate ceasefire in Hamas-run Gaza.',
 '"But the mere claim that Israel is committing genocide against Palestinians is not only false, it's outrageous, and the willingness of the court to even discuss this is a disgrace that will not be erased for generations," Netanyahu said.',
 'Defence minister Yoav Gallant said Israel did not need to be lectured on morality. In the days after the October 7 attack, Gallant had said Israel would impose a total blockade on Gaza as part of a battle against "human animals".',
 '"The International Court of Justice in The Hague went above and beyond, when it granted South Africa's antisemitic request to discuss the claim of genocide in Gaza, and now refuses to reject the petition outright," he said in a statement.',
 'Reuters',
 '',
 'Advocate Yadav said the district administration will make arrangements for performing puja and it will be facilitated by the Kashi Vishwanath Trust']

Statesman

```
url = "https://www.thestatesman.com/world/page/67"
source = requests.get(url).text
soup = BeautifulSoup(source, 'html.parser')
```

```
print(soup.prettify())
```

```
<!DOCTYPE html>
<html class="no-js" lang="en">
 <head>
  <meta charset="utf-8"/>
  <script>
   if(navigator.userAgent.match(/MSIE|Internet Explorer/i)||navigator.userAgent.match(/Trident\/7\..*?rv:11/i)){var href=document.loca
  </script>
  <script>
   class RocketLazyLoadScripts{constructor(){this.v="1.2.4",this.triggerEvents=["keydown","mousedown","mousemove","touchmove","touchst
  </script>
  <meta content="ie=edge" http-equiv="x-ua-compatible"/>
  <meta content="width=device-width, initial-scale=1" name="viewport"/>
  <link href="https://www.thestatesman.com/wp-content/themes/statesmantheme/images/favicon.ico" rel="shortcut icon" type="image/x-icon
  <link data-minify="1" href="https://www.thestatesman.com/wp-content/cache/min/1/wp-content/themes/statesmantheme/assets/css/style.cs
  <link data-minify="1" href="https://www.thestatesman.com/wp-content/cache/min/1/wp-content/themes/statesmantheme/assets/css/widgets
  <link data-minify="1" href="https://www.thestatesman.com/wp-content/cache/min/1/wp-content/themes/statesmantheme/assets/css/responsi
  <script type="rocketlazyloadscript">
   (function(html){html.className = html.className.replace(/\bno-js\b/,'js')})(document.documentElement);
  </script>
  <meta content="index, follow, max-image-preview:large, max-snippet:-1, max-video-preview:-1" name="robots">
  <!-- This site is optimized with the Yoast SEO plugin v21.1 - https://yoast.com/wordpress/plugins/seo/ -->
  <title>
   World News: Get all Latest Breaking Global News Stories, Photos &amp; Videos
  </title>
  <meta content="World News: get the latest exclusive breaking news stories, photos and videos from around the world at statesman.com
   <link href="https://www.thestatesman.com/world/page/67" rel="canonical">
    <link href="https://www.thestatesman.com/world/page/66" rel="prev">
     <link href="https://www.thestatesman.com/world/page/68" rel="next">
      <meta content="en_US" property="og:locale">
      <meta content="article" property="og:type"/>
      <meta content="World News: Get all Latest Breaking Global News Stories, Photos &amp; Videos" property="og:title"/>
      <meta content="World News: get the latest exclusive breaking news stories, photos and videos from around the world at statesma
      <meta content="https://www.thestatesman.com/world" property="og:url"/>
      <meta content="The Statesman" property="og:site_name"/>
      <meta content="https://www.thestatesman.com/wp-content/uploads/2019/06/statesman-logo-default-pic.jpg" property="og:image"/>
      <meta content="1200" property="og:image:width"/>
      <meta content="800" property="og:image:height"/>
      <meta content="image/jpeg" property="og:image:type"/>
      <meta content="summary_large_image" name="twitter:card"/>
      <meta content="@TheStatesmanLtd" name="twitter:site"/>
      <script class="yoast-schema-graph" type="application/ld+json">
       {"@context":"https://schema.org","@graph":[{"@type":"CollectionPage","@id":"https://www.thestatesman.com/world","url":"https
      </script>
      <!-- / Yoast SEO plugin. -->
      <link href="https://www.thestatesman.com/web-stories/feed/" rel="alternate" title="The Statesman » Stories Feed" type="applica
      <style type="text/css">
       img.wp-smiley,
img.emoji {
        display: inline !important;
        border: none !important;
        box-shadow: none !important;
        height: 1em !important;
        width: 1em !important;
        margin: 0 0.07em !important;
        vertical-align: -0.1em !important;
        background: none !important;
        padding: 0 !important;
```

```
# Assuming 'soup' is your BeautifulSoup object
world_links = [a['href'] for a in soup.find_all('a', href=lambda href: href and 'world' in href.lower() and 'https' in href and any(char.isd
```

```
# Print or use the extracted links
for link in world_links:
    print(link)
```

https://www.thestatesman.com/world/saudi-arabia-to-host-largest-ever-un-conference-on-land-drought-1503265221.html
https://www.thestatesman.com/world/saudi-arabia-to-host-largest-ever-un-conference-on-land-drought-1503265221.html
https://www.thestatesman.com/world/everything-ok-says-maldives-envoy-on-ties-with-india-1503265207.html
https://www.thestatesman.com/world/everything-ok-says-maldives-envoy-on-ties-with-india-1503265207.html
https://www.thestatesman.com/world/indian-navy-in-joint-op-thwarts-hijacking-of-sri-lankan-fishing-vessel-1503265027.html
https://www.thestatesman.com/world/indian-navy-in-joint-op-thwarts-hijacking-of-sri-lankan-fishing-vessel-1503265027.html
https://www.thestatesman.com/world/prenatal-exposure-to-pm-2-5-may-up-severe-respiratory-distress-in-newborns-1503265032.html
https://www.thestatesman.com/world/prenatal-exposure-to-pm-2-5-may-up-severe-respiratory-distress-in-newborns-1503265032.html
https://www.thestatesman.com/world/australia-sees-40-rise-in-sexual-assault-cases-since-2013-1503265035.html

https://www.thestatesman.com/world/australia-sees-40-rise-in-sexual-assault-cases-since-2013-1503265035.html
https://www.thestatesman.com/world/infant-killed-by-familys-hybrid-wolf-dog-pet-in-us-state-1503246866.html
https://www.thestatesman.com/world/infant-killed-by-familys-hybrid-wolf-dog-pet-in-us-state-1503246866.html
https://www.thestatesman.com/world/biden-nominates-indian-american-vc-for-key-post-1503246863.html
https://www.thestatesman.com/world/biden-nominates-indian-american-vc-for-key-post-1503246863.html
https://www.thestatesman.com/world/we-will-operate-in-maximum-force-israel-warns-as-truce-with-hamas-reaches-dead-end-1503246814.html
https://www.thestatesman.com/world/we-will-operate-in-maximum-force-israel-warns-as-truce-with-hamas-reaches-dead-end-1503246814.html
https://www.thestatesman.com/world/bigg-boss-17-housemates-go-on-an-unusual-coffee-date-on-nomination-task-1503246666.html
https://www.thestatesman.com/world/bigg-boss-17-housemates-go-on-an-unusual-coffee-date-on-nomination-task-1503246666.html
https://www.thestatesman.com/world/5-bodies-found-underwater-in-search-for-missing-us-osprey-crew-in-japan-1503246647.html
https://www.thestatesman.com/world/5-bodies-found-underwater-in-search-for-missing-us-osprey-crew-in-japan-1503246647.html

⌄ checking one url

```
url = "https://www.thestatesman.com/world/genocide-charge-against-israel-false-outrageous-prime-minister-benjamin-netanyahu-1503263746.html"
source = requests.get(url).text
soup = BeautifulSoup(source, 'html.parser')


print(soup.prettify())
```

```
<!DOCTYPE html>
<html class="no-js" lang="en">
 <head>
  <meta charset="utf-8"/>
  <script>
   if(navigator.userAgent.match(/MSIE|Internet Explorer/i)||navigator.userAgent.match(/Trident\/7\..*?rv:11/i)){var href=document.loca
  </script>
  <script>
   class RocketLazyLoadScripts{constructor(){this.v="1.2.4",this.triggerEvents=["keydown","mousedown","mousemove","touchmove","touchst
  </script>
  <meta content="ie=edge" http-equiv="x-ua-compatible"/>
  <meta content="width=device-width, initial-scale=1" name="viewport"/>
  <link href="https://www.thestatesman.com/wp-content/themes/statesmantheme/images/favicon.ico" rel="shortcut icon" type="image/x-icon
  <link data-minify="1" href="https://www.thestatesman.com/wp-content/cache/min/1/wp-content/themes/statesmantheme/assets/css/style.cs
  <link data-minify="1" href="https://www.thestatesman.com/wp-content/cache/min/1/wp-content/themes/statesmantheme/assets/css/widgets_
  <link data-minify="1" href="https://www.thestatesman.com/wp-content/cache/min/1/wp-content/themes/statesmantheme/assets/css/responsi
  <script type="rocketlazyloadscript">
   (function(html){html.className = html.className.replace(/\bno-js\b/,'js')})(document.documentElement);
  </script>
  <meta content="index, follow, max-image-preview:large, max-snippet:-1, max-video-preview:-1" name="robots">
  <!-- This site is optimized with the Yoast SEO plugin v21.1 - https://yoast.com/wordpress/plugins/seo/ -->
  <title>
   Genocide charge against Israel "false", "outrageous": Prime Minister Benjamin Netanyahu - The Statesman
  </title>
  <meta content="Israeli PM said this in response to the International Court of Justice's (ICJ) decision on South Africa's genocide d
   <link href="https://www.thestatesman.com/world/genocide-charge-against-israel-false-outrageous-prime-minister-benjamin-netanyahu-1
    <meta content="en_US" property="og:locale">
    <meta content="article" property="og:type"/>
    <meta content='Genocide charge against Israel "false", "outrageous": Prime Minister Benjamin Netanyahu - The Statesman' property
    <meta content="Israeli PM said this in response to the International Court of Justice's (ICJ) decision on South Africa's genocid
    <meta content="https://www.thestatesman.com/world/genocide-charge-against-israel-false-outrageous-prime-minister-benjamin-netany
    <meta content="The Statesman" property="og:site_name"/>
    <meta content="https://www.facebook.com/thestatesman1875/" property="article:publisher"/>
    <meta content="2024-01-27T01:58:02+00:00" property="article:published_time"/>
    <meta content="2024-01-27T03:21:42+00:00" property="article:modified_time"/>
    <meta content="https://www.thestatesman.com/wp-content/uploads/2020/08/Netanyahu.jpg" property="og:image"/>
    <meta content="1200" property="og:image:width"/>
    <meta content="800" property="og:image:height"/>
    <meta content="image/jpeg" property="og:image:type"/>
    <meta content="ANI" name="author"/>
    <meta content="summary_large_image" name="twitter:card"/>
    <meta content="@TheStatesmanLtd" name="twitter:creator"/>
    <meta content="@TheStatesmanLtd" name="twitter:site"/>
    <meta content="Written by" name="twitter:label1"/>
    <meta content="ANI" name="twitter:data1"/>
    <meta content="Est. reading time" name="twitter:label2"/>
    <meta content="2 minutes" name="twitter:data2"/>
    <script class="yoast-schema-graph" type="application/ld+json">
     {"@context":"https://schema.org","@graph":[{"@type":"WebPage","@id":"https://www.thestatesman.com/world/genocide-charge-against
    </script>
    <!-- / Yoast SEO plugin. -->
    <link href="https://www.thestatesman.com/web-stories/feed/" rel="alternate" title="The Statesman » Stories Feed" type="applicati
    <style type="text/css">
     img.wp-smiley,
img.emoji {
        display: inline !important;
        border: none !important;
```

```
# Find all <p> tags and extract the text content, excluding specific sections
for paragraph in paragraphs:
    text_content = paragraph.get_text()

    # Exclude sections containing specific phrases
    if 'PEOPLE\'S PARLIAMENT' not in text_content and 'Address' not in text_content and 'All rights reserved' not in text_content and 'Advert
        print(text_content)
```

Representational image.

File picture

The Biden administration is imposing visa bans on dozens of Israeli settlers who have committed acts of violence in the West Bank, takin
In a statement on Tuesday, secretary of state Antony Blinken announced a new policy "targeting individuals believed to have been involve
US officials are most concerned about a recent surge in attacks by Israeli settlers against Palestinians, but the new policy will also d
The restrictions may also be extended to their immediate family members, Blinken said. The move comes as the US tries to revive efforts
New York Times News Service
According to an official release, prohibitory orders under Section 144 of the CrPC, which was earlier made effective from 9 am to 10pm o

```
import requests
from bs4 import BeautifulSoup

b = {}  # Dictionary to store URLs
i = 0

# Loop through the pages
for idx in range(1, 67):  # assuming it contains data till December (one month)
    url1 = f"https://www.thestatesman.com/world/page/{idx}"
    source = requests.get(url1).text
    soup = BeautifulSoup(source, 'html.parser')

    # Find all 'a' tags with href containing 'world', 'https', at least one digit, and not containing 'page'
    links = [a['href'] for a in soup.find_all('a', href=lambda href: href and 'world' in href.lower() and 'https' in href and any(char.isdigi

    # Prepend the base URL if needed, and print the complete URLs
    # Initialize an empty list to store the complete URLs
    complete_urls = []

    # Prepend the base URL and add the complete URLs to the list
    for link in links:
        complete_url = link
        complete_urls.append(complete_url)

    # Extract the URLs from the list
    urls = complete_urls

    # Print the URLs with their indices
    for index, url in enumerate(urls, start=1):
        print(f"{index}. {url}")
        b[i] = url
        i += 1
else:
    print("Target script not found.")
```

1. https://www.thestatesman.com/world/saudi-arabia-to-host-largest-ever-un-conference-on-land-drought-1503265221.html
2. https://www.thestatesman.com/world/saudi-arabia-to-host-largest-ever-un-conference-on-land-drought-1503265221.html
3. https://www.thestatesman.com/world/everything-ok-says-maldives-envoy-on-ties-with-india-1503265207.html
4. https://www.thestatesman.com/world/everything-ok-says-maldives-envoy-on-ties-with-india-1503265207.html
5. https://www.thestatesman.com/world/indian-navy-in-joint-op-thwarts-hijacking-of-sri-lankan-fishing-vessel-1503265027.html
6. https://www.thestatesman.com/world/indian-navy-in-joint-op-thwarts-hijacking-of-sri-lankan-fishing-vessel-1503265027.html
7. https://www.thestatesman.com/world/prenatal-exposure-to-pm-2-5-may-up-severe-respiratory-distress-in-newborns-1503265032.html
8. https://www.thestatesman.com/world/prenatal-exposure-to-pm-2-5-may-up-severe-respiratory-distress-in-newborns-1503265032.html
9. https://www.thestatesman.com/world/australia-sees-40-rise-in-sexual-assault-cases-since-2013-1503265035.html
10. https://www.thestatesman.com/world/australia-sees-40-rise-in-sexual-assault-cases-since-2013-1503265035.html
11. https://www.thestatesman.com/world/now-indian-pilgrims-can-take-mountain-flight-across-kailash-mansarovar-1503265038.html
12. https://www.thestatesman.com/world/now-indian-pilgrims-can-take-mountain-flight-across-kailash-mansarovar-1503265038.html
13. https://www.thestatesman.com/world/us-court-orders-reversal-of-elon-musks-usd-56-bn-tesla-compensation-package-1503265079.html
14. https://www.thestatesman.com/world/us-court-orders-reversal-of-elon-musks-usd-56-bn-tesla-compensation-package-1503265079.html
15. https://www.thestatesman.com/world/russia-beating-western-sanction-but-its-long-term-prospects-for-the-economy-looks-grim-as-war-
16. https://www.thestatesman.com/world/russia-beating-western-sanction-but-its-long-term-prospects-for-the-economy-looks-grim-as-war-
17. https://www.thestatesman.com/world/uns-palestine-relief-agency-faces-collapse-after-complaints-staff-joined-terror-attack-15032646
18. https://www.thestatesman.com/world/uns-palestine-relief-agency-faces-collapse-after-complaints-staff-joined-terror-attack-15032646
19. https://www.thestatesman.com/world/punjab-man-guilty-of-fatal-car-crash-in-canada-deported-to-india-1503264678.html
20. https://www.thestatesman.com/world/punjab-man-guilty-of-fatal-car-crash-in-canada-deported-to-india-1503264678.html
1. https://www.thestatesman.com/world/saudi-arabia-to-host-largest-ever-un-conference-on-land-drought-1503265221.html
2. https://www.thestatesman.com/world/saudi-arabia-to-host-largest-ever-un-conference-on-land-drought-1503265221.html
3. https://www.thestatesman.com/world/everything-ok-says-maldives-envoy-on-ties-with-india-1503265207.html
4. https://www.thestatesman.com/world/everything-ok-says-maldives-envoy-on-ties-with-india-1503265207.html
5. https://www.thestatesman.com/world/indian-navy-in-joint-op-thwarts-hijacking-of-sri-lankan-fishing-vessel-1503265027.html

6. https://www.thestatesman.com/world/indian-navy-in-joint-op-thwarts-hijacking-of-sri-lankan-fishing-vessel-1503265027.html
7. https://www.thestatesman.com/world/prenatal-exposure-to-pm-2-5-may-up-severe-respiratory-distress-in-newborns-1503265032.html
8. https://www.thestatesman.com/world/prenatal-exposure-to-pm-2-5-may-up-severe-respiratory-distress-in-newborns-1503265032.html
9. https://www.thestatesman.com/world/australia-sees-40-rise-in-sexual-assault-cases-since-2013-1503265035.html
10. https://www.thestatesman.com/world/australia-sees-40-rise-in-sexual-assault-cases-since-2013-1503265035.html
11. https://www.thestatesman.com/world/missing-indian-student-found-dead-on-purdue-university-campus-in-us-1503264675.html
12. https://www.thestatesman.com/world/missing-indian-student-found-dead-on-purdue-university-campus-in-us-1503264675.html
13. https://www.thestatesman.com/world/elon-musk-announces-neuralinks-successful-brain-implant-procedure-in-first-human-recipient-1503
14. https://www.thestatesman.com/world/elon-musk-announces-neuralinks-successful-brain-implant-procedure-in-first-human-recipient-1503
15. https://www.thestatesman.com/world/security-beefed-up-as-maldivian-opposition-party-readies-impeachment-motion-against-president-r
16. https://www.thestatesman.com/world/security-beefed-up-as-maldivian-opposition-party-readies-impeachment-motion-against-president-r
17. https://www.thestatesman.com/world/uk-orders-review-into-nhs-trust-which-treated-british-indian-teens-killer-1503264331.html
18. https://www.thestatesman.com/world/uk-orders-review-into-nhs-trust-which-treated-british-indian-teens-killer-1503264331.html
19. https://www.thestatesman.com/world/suspension-of-funding-to-unrwa-deplorable-palestine-authority-1503264323.html
20. https://www.thestatesman.com/world/suspension-of-funding-to-unrwa-deplorable-palestine-authority-1503264323.html
1. https://www.thestatesman.com/world/saudi-arabia-to-host-largest-ever-un-conference-on-land-drought-1503265221.html
2. https://www.thestatesman.com/world/saudi-arabia-to-host-largest-ever-un-conference-on-land-drought-1503265221.html
3. https://www.thestatesman.com/world/everything-ok-says-maldives-envoy-on-ties-with-india-1503265207.html
4. https://www.thestatesman.com/world/everything-ok-says-maldives-envoy-on-ties-with-india-1503265207.html
5. https://www.thestatesman.com/world/indian-navy-in-joint-op-thwarts-hijacking-of-sri-lankan-fishing-vessel-1503265027.html
6. https://www.thestatesman.com/world/indian-navy-in-joint-op-thwarts-hijacking-of-sri-lankan-fishing-vessel-1503265027.html
7. https://www.thestatesman.com/world/prenatal-exposure-to-pm-2-5-may-up-severe-respiratory-distress-in-newborns-1503265032.html
8. https://www.thestatesman.com/world/prenatal-exposure-to-pm-2-5-may-up-severe-respiratory-distress-in-newborns-1503265032.html
9. https://www.thestatesman.com/world/australia-sees-40-rise-in-sexual-assault-cases-since-2013-1503265035.html
10. https://www.thestatesman.com/world/australia-sees-40-rise-in-sexual-assault-cases-since-2013-1503265035.html
11. https://www.thestatesman.com/world/funding-suspension-forces-unrwa-to-halt-all-gaza-activities-in-weeks-official-1503264327.html
12. https://www.thestatesman.com/world/funding-suspension-forces-unrwa-to-halt-all-gaza-activities-in-weeks-official-1503264327.html
13. https://www.thestatesman.com/world/peace-talks-constructive-says-israel-pmo-1503264319.html
14. https://www.thestatesman.com/world/peace-talks-constructive-says-israel-pmo-1503264319.html
15. https://www.thestatesman.com/world/trump-slams-us-mexico-border-deal-as-catastrophe-waiting-to-happen-1503264254.html
16. https://www.thestatesman.com/world/trump-slams-us-mexico-border-deal-as-catastrophe-waiting-to-happen-1503264254.html
17. https://www.thestatesman.com/world/lot-of-juggling-replacements-needed-for-india-to-get-ideal-test-batting-order-1503264037.html

```python
# Filter URLs containing both "Hamas" and "Israel" (case-insensitive)
filtered_urls = {key: value for key, value in b.items() if 'hamas' in value.lower() or 'israel' in value.lower()}

# Create a new dictionary with consecutive integer keys
filtered_urls_dict = {i + 1: url for i, url in enumerate(filtered_urls.values())}

# Print the filtered URLs with new keys
for key, value in filtered_urls_dict.items():
    print(f"{key}. {value}")
```

1. https://www.thestatesman.com/world/peace-talks-constructive-says-israel-pmo-1503264319.html
2. https://www.thestatesman.com/world/peace-talks-constructive-says-israel-pmo-1503264319.html
3. https://www.thestatesman.com/world/genocide-charge-against-israel-false-outrageous-prime-minister-benjamin-netanyahu-1503263746.htm
4. https://www.thestatesman.com/world/genocide-charge-against-israel-false-outrageous-prime-minister-benjamin-netanyahu-1503263746.htm
5. https://www.thestatesman.com/world/us-pauses-additional-funding-for-unrwa-amid-allegations-of-agencys-involvement-in-hamas-attack-c
6. https://www.thestatesman.com/world/us-pauses-additional-funding-for-unrwa-amid-allegations-of-agencys-involvement-in-hamas-attack-c
7. https://www.thestatesman.com/world/israel-hamas-war-cia-mossad-chiefs-to-meet-qatar-pm-for-temporary-ceasefire-1503263508.html
8. https://www.thestatesman.com/world/israel-hamas-war-cia-mossad-chiefs-to-meet-qatar-pm-for-temporary-ceasefire-1503263508.html
9. https://www.thestatesman.com/world/icj-to-deliver-interim-ruling-on-genocide-case-against-israel-on-friday-1503263148.html
10. https://www.thestatesman.com/world/icj-to-deliver-interim-ruling-on-genocide-case-against-israel-on-friday-1503263148.html
11. https://www.thestatesman.com/world/death-toll-rises-to-13-in-israeli-missile-strike-on-damascus-1503262007.html
12. https://www.thestatesman.com/world/death-toll-rises-to-13-in-israeli-missile-strike-on-damascus-1503262007.html
13. https://www.thestatesman.com/world/israel-to-maintain-security-control-over-gaza-after-conflict-pm-netanyahu-1503261048.html
14. https://www.thestatesman.com/world/israel-to-maintain-security-control-over-gaza-after-conflict-pm-netanyahu-1503261048.html
15. https://www.thestatesman.com/world/hamass-battalion-framework-in-northern-gaza-dismantled-israeli-defence-min-1503260045.html
16. https://www.thestatesman.com/world/hamass-battalion-framework-in-northern-gaza-dismantled-israeli-defence-min-1503260045.html
17. https://www.thestatesman.com/world/idf-claims-to-arrest-sisters-of-slain-hamas-leader-saleh-al-arouri-1503259655.html
18. https://www.thestatesman.com/world/idf-claims-to-arrest-sisters-of-slain-hamas-leader-saleh-al-arouri-1503259655.html
19. https://www.thestatesman.com/world/international-court-of-justice-hears-south-africa-genocide-case-against-israel-1503258634.html
20. https://www.thestatesman.com/world/international-court-of-justice-hears-south-africa-genocide-case-against-israel-1503258634.html
21. https://www.thestatesman.com/world/israel-reports-new-cases-of-monkeypox-measles-1503258078.html
22. https://www.thestatesman.com/world/israel-reports-new-cases-of-monkeypox-measles-1503258078.html
23. https://www.thestatesman.com/world/senior-hezbollah-commander-killed-in-israeli-airstrike-1503258072.html
24. https://www.thestatesman.com/world/senior-hezbollah-commander-killed-in-israeli-airstrike-1503258072.html
25. https://www.thestatesman.com/world/loss-of-lives-clearly-unacceptable-dialogue-only-way-forward-india-on-israel-hamas-war-15032586
26. https://www.thestatesman.com/world/loss-of-lives-clearly-unacceptable-dialogue-only-way-forward-india-on-israel-hamas-war-15032586
27. https://www.thestatesman.com/world/israeli-delegation-in-cairo-for-hostage-release-negotiations-1503257725.html
28. https://www.thestatesman.com/world/israeli-delegation-in-cairo-for-hostage-release-negotiations-1503257725.html
29. https://www.thestatesman.com/world/israel-to-present-harrowing-video-accounts-of-oct-7-hamas-attack-before-icj-1503257370.html
30. https://www.thestatesman.com/world/israel-to-present-harrowing-video-accounts-of-oct-7-hamas-attack-before-icj-1503257370.html
31. https://www.thestatesman.com/world/palestinians-will-run-civilian-affairs-in-post-conflict-gaza-israel-1503256492.html
32. https://www.thestatesman.com/world/palestinians-will-run-civilian-affairs-in-post-conflict-gaza-israel-1503256492.html
33. https://www.thestatesman.com/world/time-for-diplomatic-solution-with-hezbollah-is-short-israeli-defence-minister-1503256489.html
34. https://www.thestatesman.com/world/time-for-diplomatic-solution-with-hezbollah-is-short-israeli-defence-minister-1503256489.html
35. https://www.thestatesman.com/world/israels-top-court-delays-law-thwarting-removal-of-netanyahu-from-office-1503256054.html
36. https://www.thestatesman.com/world/israels-top-court-delays-law-thwarting-removal-of-netanyahu-from-office-1503256054.html
37. https://www.thestatesman.com/world/hezbollah-chief-vows-to-wage-war-against-israel-if-it-expands-attacks-on-lebanon-1503256051.htm
38. https://www.thestatesman.com/world/hezbollah-chief-vows-to-wage-war-against-israel-if-it-expands-attacks-on-lebanon-1503256051.htm

39. https://www.thestatesman.com/world/biden-at-risk-of-losing-young-black-hispanic-voters-over-not-halting-israel-palestine-war-1503z
40. https://www.thestatesman.com/world/biden-at-risk-of-losing-young-black-hispanic-voters-over-not-halting-israel-palestine-war-1503z
41. https://www.thestatesman.com/world/israels-supreme-court-strikes-down-netanyahus-judicial-overhaul-1503255367.html
42. https://www.thestatesman.com/world/israels-supreme-court-strikes-down-netanyahus-judicial-overhaul-1503255367.html
43. https://www.thestatesman.com/world/biden-admin-sidesteps-congress-again-approves-military-equipment-sale-to-israel-1503254501.html
44. https://www.thestatesman.com/world/biden-admin-sidesteps-congress-again-approves-military-equipment-sale-to-israel-1503254501.html
45. https://www.thestatesman.com/world/israel-to-expand-operations-in-southern-gaza-report-1503254292.html
46. https://www.thestatesman.com/world/israel-to-expand-operations-in-southern-gaza-report-1503254292.html
47. https://www.thestatesman.com/world/american-israeli-missing-since-hamas-attack-declared-dead-1503254149.html
48. https://www.thestatesman.com/world/american-israeli-missing-since-hamas-attack-declared-dead-1503254149.html
49. https://www.thestatesman.com/world/israel-issues-travel-advisory-for-its-nationals-in-india-following-explosion-outside-mission-15
50. https://www.thestatesman.com/world/israel-issues-travel-advisory-for-its-nationals-in-india-following-explosion-outside-mission-15
51. https://www.thestatesman.com/world/israel-expands-ground-offensive-in-central-gaza-un-warns-of-catastrophic-humanitarian-crisis-19
52. https://www.thestatesman.com/world/israel-expands-ground-offensive-in-central-gaza-un-warns-of-catastrophic-humanitarian-crisis-19
53. https://www.thestatesman.com/world/hamas-denies-rejecting-egyptian-ceasefire-deal-1503253247.html
54. https://www.thestatesman.com/world/hamas-denies-rejecting-egyptian-ceasefire-deal-1503253247.html
55. https://www.thestatesman.com/world/defence-minister-gallant-warns-hezbollah-israel-can-do-in-beirut-what-it-did-in-gaza-150325324
56. https://www.thestatesman.com/world/defence-minister-gallant-warns-hezbollah-israel-can-do-in-beirut-what-it-did-in-gaza-150325324
57. https://www.thestatesman.com/world/us-expects-israel-hamas-conflict-to-move-to-lower-intensity-phase-says-blinken-1503251810.html

```python
sm_dict = {}

# Assuming filtered_urls_dict is a dictionary with indices as keys and URLs as values

for idx, url in filtered_urls_dict.items():
    # Fetch HTML content
    source = requests.get(url).text
    soup = BeautifulSoup(source, 'html.parser')

    # Initialize a list to store text values for the current iteration
    a = []

    # Find all <p> tags and extract the text content
    paragraphs = soup.find_all('p')

    # Print the text content
    for paragraph in paragraphs:
        try:
            text_content = paragraph.get_text()
            # Exclude paragraphs containing specific phrases
            if 'PEOPLE\'S PARLIAMENT' not in text_content and 'Address' not in text_content and 'All rights reserved' not in text_content an
                # Append the "text" values to the list
                if isinstance(text_content, str):
                    a.append(text_content)

        except json.JSONDecodeError as e:
            print(f"Error decoding JSON: {e}")

    # Now, 'a' contains the stored text values for the current iteration
    print(f"Stored Text Values in 'a' for iteration {idx}:")
    for text_value in a:
        sm_dict.setdefault(idx, []).append(text_value)
```

```
Stored Text Values in 'a' for iteration 1:
Stored Text Values in 'a' for iteration 2:
Stored Text Values in 'a' for iteration 3:
Stored Text Values in 'a' for iteration 4:
Stored Text Values in 'a' for iteration 5:
Stored Text Values in 'a' for iteration 6:
Stored Text Values in 'a' for iteration 7:
Stored Text Values in 'a' for iteration 8:
Stored Text Values in 'a' for iteration 9:
Stored Text Values in 'a' for iteration 10:
Stored Text Values in 'a' for iteration 11:
Stored Text Values in 'a' for iteration 12:
Stored Text Values in 'a' for iteration 13:
Stored Text Values in 'a' for iteration 14:
Stored Text Values in 'a' for iteration 15:
Stored Text Values in 'a' for iteration 16:
Stored Text Values in 'a' for iteration 17:
Stored Text Values in 'a' for iteration 18:
Stored Text Values in 'a' for iteration 19:
Stored Text Values in 'a' for iteration 20:
Stored Text Values in 'a' for iteration 21:
Stored Text Values in 'a' for iteration 22:
Stored Text Values in 'a' for iteration 23:
Stored Text Values in 'a' for iteration 24:
Stored Text Values in 'a' for iteration 25:
Stored Text Values in 'a' for iteration 26:
```

```
Stored Text Values in 'a' for iteration 27:
Stored Text Values in 'a' for iteration 28:
Stored Text Values in 'a' for iteration 29:
Stored Text Values in 'a' for iteration 30:
Stored Text Values in 'a' for iteration 31:
Stored Text Values in 'a' for iteration 32:
Stored Text Values in 'a' for iteration 33:
Stored Text Values in 'a' for iteration 34:
Stored Text Values in 'a' for iteration 35:
Stored Text Values in 'a' for iteration 36:
Stored Text Values in 'a' for iteration 37:
Stored Text Values in 'a' for iteration 38:
Stored Text Values in 'a' for iteration 39:
Stored Text Values in 'a' for iteration 40:
Stored Text Values in 'a' for iteration 41:
Stored Text Values in 'a' for iteration 42:
Stored Text Values in 'a' for iteration 43:
Stored Text Values in 'a' for iteration 44:
Stored Text Values in 'a' for iteration 45:
Stored Text Values in 'a' for iteration 46:
Stored Text Values in 'a' for iteration 47:
Stored Text Values in 'a' for iteration 48:
Stored Text Values in 'a' for iteration 49:
Stored Text Values in 'a' for iteration 50:
Stored Text Values in 'a' for iteration 51:
Stored Text Values in 'a' for iteration 52:
Stored Text Values in 'a' for iteration 53:
Stored Text Values in 'a' for iteration 54:
Stored Text Values in 'a' for iteration 55:
Stored Text Values in 'a' for iteration 56:
Stored Text Values in 'a' for iteration 57:
Stored Text Values in 'a' for iteration 58:
```

```
sm_dict[12]
```

```
['',
 'The death toll from Israeli missile attack last week on Iran's targets in the Syrian capital of Damascus has risen to 13, according
to a war monitor.',
 '\n                              IANS | New Delhi | January 22, 2024 2:28 pm ',
 ' Representation image (Photo: REUTERS)    ',
 'The death toll from Israeli missile attack last week on Iran's targets in the Syrian capital of Damascus has risen to 13, according
to a war monitor.',
 'The Syrian Observatory for Human Rights said on Sunday that among the 13 casualties were five Iranians, three of whom held commanding
roles in Iran's Islamic Revolutionary Guard Corps (IRGC), four Syrians contracted with Iranian militias, two Lebanese individuals, one
Iraqi national, and one Syrian civilian worker.',
 'The Israeli missiles on Saturday struck the affluent Mazzeh Western Villas neighbourhood, specifically targetting a residential
building where a meeting of IRGC commanders was reportedly taking place, Xinhua news agency reported.',
 '',
 'On Saturday, the Syrian Foreign Ministry strongly condemned the Israeli attack, and called on the international community to take
immediate action to halt such "atrocities".',
 '',
 ''Human rights' is probably the most abused phrase in the international political lexicon. The US, which is mindful of human rights
violations in places as far afield as India, China and Venezuela, keeps mum when small children die on its border with Mexico or when
the Israeli Army kills more than 20,000 civilians in Gaza, with active US support. ',
 "The US military carried out an airstrike on a facility in eastern Syria used by Iran's Islamic Revolutionary Guard Corps (IRGC) and
groups backed by Tehran, the Defense Department said in a statement.",
 'US President Joe Biden has said that he raised the issues of human rights and freedom of press during his… ',
 '\n\n']
```

∨ BBC

```
url = "https://www.bbc.com/news/topics/c2vdnvdg6xxt?page=2" #a random page,
source = requests.get(url).text
soup = BeautifulSoup(source, 'html.parser')
```

```
print(soup.prettify())
```

```
<!DOCTYPE html>
<html>
 <head>
  <meta charset="utf-8"/>
  <meta content="width=device-width" name="viewport"/>
  <title>
   Israel Gaza war: Latest news from BBC News
  </title>
  <meta content="Israel Gaza war: Latest news from BBC News" property="og:title"/>
  <meta content="Israel Gaza war: Latest news from BBC News" name="twitter:title"/>
  <meta content="Israel Gaza war which started after armed attack by Hamas on 7 October 2023" name="description"/>
  <meta content="Israel Gaza war which started after armed attack by Hamas on 7 October 2023" property="og:description"/>
```

```
<meta content="Israel Gaza war which started after armed attack by Hamas on 7 October 2023" name="twitter:description"/>
<meta content="#da532c" name="msapplication-TileColor"/>
<meta content="#ffffff" name="theme-color"/>
<meta content="NOODP, NOYDIR" name="robots"/>
<link href="/bbcx/apple-touch-icon.png" rel="apple-touch-icon" sizes="180x180"/>
<link href="/bbcx/favicon-32x32.png" rel="icon" sizes="32x32" type="image/png"/>
<link href="/bbcx/favicon-16x16.png" rel="icon" sizes="16x16" type="image/png"/>
<link href="/bbcx/favicon.ico" rel="alternate icon"/>
<link href="/bbcx/site.webmanifest" rel="manifest"/>
<link color="#000000" href="/bbcx/safari-pinned-tab.svg" rel="mask-icon"/>
<link href="https://www.bbc.com/news/topics/c2vdnvdg6xxt" rel="canonical"/>
<link data-testid="en-hreflang-tag" href="https://www.bbc.com/news/topics/c2vdnvdg6xxt" hreflang="en" rel="alternate"/>
<link data-testid="en-gb-hreflang-tag" href="https://www.bbc.co.uk/news/topics/c2vdnvdg6xxt" hreflang="en-gb" rel="alternate"/>
<meta content="1.1.1+1" name="version"/>
<script type="application/ld+json">
 {"@context":"http://schema.org","@type":"WebPage","description":"Israel Gaza war which started after armed attack by Hamas on 7 Oct
</script>
<meta content="22" name="next-head-count"/>
<script async="" data-nscript="beforeInteractive" id="inline-dotcom">
 window.dotcom = window.dotcom || { cmd: [] }; window.dotcomConfig = { pageAds: true, playerAds: true, xPlatform: true };
</script>
<script async="" data-nscript="beforeInteractive" id="inline-googletag">
 window.googletag = window.googletag || { cmd: [] };
</script>
<script async="" data-nscript="beforeInteractive" id="inline-analytics">
 window.analytics = window.analytics || { echoClient: null, buffer:[] };
</script>
<script async="" data-nscript="beforeInteractive" id="inline-dotcom-config">
 window.dotcom = window.dotcom || { cmd: [] }; window.dotcomConfig = {"pageAds":true,"playerAds":true,"xPlatform":true,"opsFlags":{"
</script>
<noscript data-n-css="">
</noscript>
<script defer="" nomodule="" src="/bbcx/_next/static/chunks/polyfills-c67a75d1b6f99dc8.js">
</script>
<script async="" data-nscript="beforeInteractive" defer="" id="requirejs-script" src="https://static.bbci.co.uk/frameworks/requirej
</script>
<script async="" data-nscript="beforeInteractive" defer="" id="ads-service-script" src="https://gn-web-assets.api.bbc.com/ngas/late
</script>
<script async="" data-nscript="beforeInteractive" defer="" id="bump-script" src="https://emp.bbci.co.uk/emp/bump-4/bump-4.js" type="
</script>
<script defer="" src="/bbcx/_next/static/chunks/333.2bb044bf09346690.js">
</script>
<script defer="" src="/bbcx/_next/static/chunks/786.2c4b170e93e6bfe1.js">
</script>
```

```python
import requests
from bs4 import BeautifulSoup

# Specify the URL of the page you want to scrape
url = "https://www.bbc.com/news/topics/c2vdnvdg6xxt?page=2"

# Send a GET request to the URL
response = requests.get(url)

# Parse the HTML content of the page
soup = BeautifulSoup(response.text, 'html.parser')

# Find all anchor tags with class 'sc-4befc967-1' (adjust the class accordingly)
anchor_tags = soup.find_all('a', class_='sc-4befc967-1')

# Extract href values starting with "/news/world-middle-east"
filtered_href_values = [tag['href'] for tag in anchor_tags if tag.get('href', '').startswith('/news/world-middle-east')]

# Now you can use the filtered href values as needed
print(filtered_href_values)
```

```
['/news/world-middle-east-68141039', '/news/world-middle-east-68006607', '/news/world-middle-east-68137050', '/news/world-middle-east-67
```

```python
len(filtered_href_values)
```

```
26
```

```python
import requests
from bs4 import BeautifulSoup

bbc1 = {}
i = 0

# Iterate over pages 1 to 25
for page_number in range(1, 26):
    # Construct the URL for the current page
    url = f"https://www.bbc.com/news/topics/c2vdnvdg6xxt?page={page_number}"
    source = requests.get(url).text
    soup = BeautifulSoup(source, 'html.parser')

    # Find all anchor tags with class 'sc-4befc967-1'
    anchor_tags = soup.find_all('a', class_='sc-4befc967-1')

    # Extract href values that start with '/news/world-middle-east'
    filtered_href_values = [tag['href'] for tag in anchor_tags if tag.get('href', '').startswith('/news/world-middle-east')]

    # Print the URLs with their indices
    for index, url in enumerate(filtered_href_values, start=1):
        print(f"{index}. {url}")
        bbc1[i] = url
        i += 1

    if not filtered_href_values:
        print(f"No URLs found on page {page_number}")

# Print the collected URLs
print("Collected URLs:")
for index, url in bbc1.items():
    print(f"{index}. {url}")

#this code gets changed when run from different windows
```

```
 1. /news/world-middle-east-68141039
 2. /news/world-middle-east-68006607
 3. /news/world-middle-east-68137050
 4. /news/world-middle-east-67983018
 5. /news/world-middle-east-68006607
 6. /news/world-middle-east-68150519
 7. /news/world-middle-east-68138853
 8. /news/world-middle-east-68137050
 9. /news/world-middle-east-67039975
10. /news/world-middle-east-20415675
11. /news/world-middle-east-67614911
12. /news/world-middle-east-68097993
13. /news/world-middle-east-68113223
14. /news/world-middle-east-68017444
15. /news/world-middle-east-67972962
16. /news/world-middle-east-67949952
17. /news/world-middle-east-68059132
18. /news/world-middle-east-68034128
19. /news/world-middle-east-67958260
20. /news/world-middle-east-67952301
21. /news/world-middle-east-67039975
22. /news/world-middle-east-68141039
23. /news/world-middle-east-68150519
24. /news/world-middle-east-68138853
25. /news/world-middle-east-68137050
26. /news/world-middle-east-68006607
 1. /news/world-middle-east-68141039
 2. /news/world-middle-east-68006607
 3. /news/world-middle-east-68137050
 4. /news/world-middle-east-67983018
 5. /news/world-middle-east-68006607
 6. /news/world-middle-east-68150519
 7. /news/world-middle-east-68138853
 8. /news/world-middle-east-68137050
 9. /news/world-middle-east-67039975
10. /news/world-middle-east-20415675
11. /news/world-middle-east-67614911
12. /news/world-middle-east-68097993
13. /news/world-middle-east-68113223
14. /news/world-middle-east-68017444
15. /news/world-middle-east-67972962
16. /news/world-middle-east-67949952
17. /news/world-middle-east-68059132
18. /news/world-middle-east-68034128
19. /news/world-middle-east-67958260
20. /news/world-middle-east-67952301
```

```
21. /news/world-middle-east-67039975
22. /news/world-middle-east-68141039
23. /news/world-middle-east-68150519
24. /news/world-middle-east-68138853
25. /news/world-middle-east-68137050
26. /news/world-middle-east-68006607
1. /news/world-middle-east-68141039
2. /news/world-middle-east-68006607
3. /news/world-middle-east-68137050
4. /news/world-middle-east-67983018
5. /news/world-middle-east-68006607
6. /news/world-middle-east-68150519
```

```
len(bbc1)
```

```
650
```

```
bbc_dict1=bbc1
```

```python
import requests
from bs4 import BeautifulSoup

# Assuming you have a dictionary bbc_dict1 with href values
bbc_dict = {}

for idx, href_value in enumerate(bbc_dict1.values()):
    url = f"https://www.bbc.com{href_value}"
    source = requests.get(url).text
    soup = BeautifulSoup(source, 'html.parser')

    # Extract and concatenate all text content from each <p> tag
    paragraphs = soup.find_all('p')
    text_content = ' '.join(paragraph.get_text(strip=True) for paragraph in paragraphs)

    # Check if the text content contains "Hamas" or "conflict"
    if "Hamas" in text_content or "conflict" in text_content:
        # Store the text content in the dictionary using append
        for text_value in text_content.split('. '):
            bbc_dict.setdefault(idx, []).append(text_value)

# Print the dictionary
for key, value in bbc_dict.items():
    print(f"Key: {key}")
    print(f"  Text Content:")
    for text_value in value:
        print(f"    {text_value}")
    print()
```

```
    Streaming output truncated to the last 5000 lines.
      "We feel that the ground is shaking under us, something like an earthquake, and this is repeated almost every evening, the bombing
      "When we were in Gaza [City], most of the attacks were from the air." Manifestations of security breakdown are spreading throughou
      A journalist who requested anonymity told the BBC: "I was returning from my job in Rafah late at night
      I had three masked people stop me - they were carrying knives, and one of them was carrying a gun
      "They searched the car for anything valuable
      One of them noticed that I was a journalist and allowed me to leave." On the streets of Rafah masked police officers with guns in
      "I bought a bag of wheat marked Unrwa [the UN agency for Palestinian refugees] for $100 (£79; €92) - 10 times the original price
      Why are the police not arresting the war traders who sell wheat most likely stolen from the UN aid agencies?" complained Mohammed
      "I can't find milk for my baby, my son has autism and has not received treatment for months
      I cannot calm him down, and his condition has deteriorated after we made great strides in treating him before the war." Hamas alwa
      The Israeli army has destroyed the government system, killed a large number of security and police officers, and destroyed securit
      Neveen Imadedin, a mother displaced from her home in Gaza City and now living in the southern town of Rafah, said: "We want the wa
      We want our state, not to return to the no-peace, no-war situation
      "Look what happened to us, thieves stole everything from my house, my clothes, home furniture and even solar panels." "My house is
      "We are displaced here in Rafah and before that [we were] in Khan Younis
      We have no money and no home, the unstoppable wars between Hamas and Israel have destroyed our lives." A businessman called Mohamm
      "We have lived under unjust rule for more than 16 years
      Hamas imposed heavy taxes on us and now they have dragged us into a devastating war in which we lost our property, money and homes
      "I used to live in a good house and work in trade
      Now what will we do? The crossings are closed, the company has been destroyed, and the house is not fit for habitation due to the
      "The worst thing that could happen to us is that we return under Hamas rule when the war ends." No injuries are reported, as the I
      Kataib Hezbollah is believed to have been behind the attack in Jordan that killed three US troops
      Why did Hamas attack Israel, and what has happened in Gaza since the 7 October raid? Some 19,000 children have been orphaned by th
      Top officials issue a joint appeal after countries halt funding for the territory's main aid agency
      Copyright 2024 BBC
      All rights reserved.TheBBCisnot responsible for the content of external sites.Read about our approach to external linking


    Key: 330
      Text Content:
```

They are known as Israel's eyes on the Gaza border
For years, units of young female conscripts had one job here
It was to sit in surveillance bases for hours, looking for signs of anything suspicious
In the months leading up to the 7 October attacks by Hamas, they did begin to see things: practice raids, mock hostage-taking, and
Noa, not her real name, says they would pass information about what they were seeing to intelligence and higher-ranking officers,
"We were just the eyes," she says
It was clear to some of these women that Hamas was planning something big - that there was, in Noa's words, a "balloon that was go
The BBC has now spoken to these young women about the escalation in suspicious activity they observed, the reports they filed, and
We have also seen WhatsApp messages the women sent in the months before 7 October, talking about incidents at the border
To some of them it became a dark joke: who would be on duty when the inevitable attack came? These women were not the only ones ra
The BBC has spoken also to the grieving families who have now lost their daughters, and to experts who see the IDF's response to t
The IDF said it was "currently focused on eliminating the threat from the terrorist organisation Hamas" and declined to answer the
"The problem is that they [the military] didn't connect the dots," a former commander at one of the border units tells the BBC
If they had, she says, they would have realised that Hamas was preparing something unprecedented
Shai Ashram, 19, was one of the women on duty on 7 October
In a call with her family, where they could hear gunshots ringing in the background, she said there were "terrorists in the base a
She was one of more than a dozen surveillance soldiers killed
Others were taken hostage
As Hamas attacked, the women at Nahal Oz, a base about a kilometre from the Gaza border, began to say goodbye to one another on th
Noa, who was not on duty and was reading the messages from home, remembers thinking "this is it"
The attack they had long feared was now actually happening
Because of the locations of their bases, the women of this military unit - known as tatzpitaniyot in Hebrew - were among the first
The women sit inside rooms close to the border, staring for hours every day at live surveillance footage captured by cameras along
There are several of these units next to the Gaza fence, and others at different positions along Israel's borders
They are all made up of young women, aged in their late teens to early 20s
They do not carry guns

bbc_dict[1]

["More than half of Gaza's buildings have been damaged or destroyed since Israel launched its retaliation for the Hamas attacks of 7 October, new analysis seen by the BBC reveals",
 "Detailed before-and-after imagery also shows how the bombardment of southern and central Gaza has intensified since the start of December, with the city of Khan Younis bearing much of the brunt of Israel's military action",
 'Israel has repeatedly told Gazans to move south for their own safety',
 'Across Gaza, residential areas have been left ruined, previously busy shopping streets reduced to rubble, universities destroyed and farmlands churned up, with tent cities springing up on the southern border to house many thousands of people left homeless',
 "About 1.7 million people - more than 80% of Gaza's population - are displaced, with nearly half crammed in the far southern end of the strip,according to the United Nations",
 'Further analysis, by BBC Verify, reveals the scale of destruction of farmland, identifying multiple areas of extensive damage',
 'The Israel Defense Forces (IDF) has said it is targeting both Hamas fighters and "terror infrastructure", when challenged over the scale of damage',
 'Now, satellite data analysis obtained by the BBC shows the true extent of the destruction',
 'The analysis suggests between 144,000 and 175,000 buildings across the whole Gaza Strip have been damaged or destroyed',
 "That's between 50% and 61% of Gaza's buildings",
 'The analysis, carried out by Corey Scher of City University of New York and Jamon Van Den Hoek of Oregon State University, compares images to reveal sudden changes in the height or structure of buildings which indicate damage',
 'The southern city of Khan Younis has been particularly badly hit in recent weeks, with more than 38,000 (or more than 46%) of buildings now destroyed or damaged, according to the analysis',
 'Over the past fortnight, more than 1,500 buildings have been destroyed or damaged there',
 "Al-Farra Tower - a 16-storey residential block in the centre of the city, the tallest building in the area - was flattened on 9 January as can be seen in before-and-after images of the city's skyline",
 'Much of the neighbourhood in which it sits has been levelled by Israeli attacks since late December',
 '"Israeli forces targeted residential complexes, especially in the downtown Khan Younis area," said Rawan Qaddah, a 20-year-old resident, who has been displaced and has lost contact with her family',
 'She named schools among the many buildings which had been damaged',
 'Some were now being used to house displaced people temporarily',
 'You can clearly see the level of damage from street level',
 'Once bustling high streets have been left derelict or destroyed',
 "These images show the front of the Shawarma Sanabel restaurant before Israel's invasion, and how the same junction looked in a composite image from January after intense bombardment of the area",
 'The IDF has repeatedly justified its actions by noting that Hamas deliberately embeds itself in civilian areas and explained destruction of buildings in the light of targeting fighters',
 'But questions have been asked about destruction of buildings seemingly firmly in the control of the IDF',
 'One example was the Israa University, in northern Gaza - initially badly damaged shortly before being blown up completely in what looked like a massive controlled explosion',
 'The videowas widely shared on social mediaand the IDF says the approval process for the blast is now being investigated',
 "Many of Gaza's historic sites have suffered extensive damage,including the al-Omari Mosqueoriginally built in the 7th Century",
 "Mr Scher, one of the academics who worked on the Gaza damage assessment, said it stands out compared with other war zones he's analysed",
 '"We\'ve done work over Ukraine, we\'ve also looked at Aleppo and other cities, but the extent and the pace of damage is remarkable',
 'I\'ve never seen this much damage appear so quickly." Further analysis, carried out by BBC Verify, shows large areas of previously cultivated land across Gaza have been extensively damaged',
 'As you can see from the satellite image below, several parts of Gaza show the effects of Israeli incursions and bombardment',
 'Although Gaza was heavily dependent on imports before the start of the war, a lot of its food came from farming and food production inside the strip',
 "Aid agencies say half of Gaza's population is now facing starvation",
 'BBC Arabic spoke to one farmer, Saeed, who fled south from Beit Lahia, in the north of Gaza, in mid-November',
 'The 33-year-old grew guava, figs, lemons, oranges, mint, and basil and earned about $6,000 (£5,535) from these crops every year - the only source of income for him, his father and his sister',
 'He had tended to the farm, inherited from his grandparents, for 15 years',
 'But days after fleeing, he says he was told by a relative that the farm had been destroyed by the IDF, along with five surrounding homes which belonged to his relatives',

'In the north and centre of Gaza, where most agriculture took place before the war, large areas of land appear ruined',
'In many places the damage corresponds with the construction of temporary Israeli defences, earth banks to protect armoured
vehicles, and the clearing of surrounding land',
'Some farmers have lost their crops even though their land was not directly hit, the BBC understands',

I have all data of newspaper in dictionaries:

1. bi_dict [Business Insider]

2. time_dict [TIME]

3. tg_dict [Telegraph]

4. sm_dict [StatesMan]

5. bbc_dict [BBC]

```
import nltk
from nltk.corpus import stopwords
from nltk.tokenize import word_tokenize
from nltk.stem import WordNetLemmatizer
from gensim import corpora, models
import string

nltk.download('punkt')
nltk.download('stopwords')
nltk.download('wordnet')
```

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

```
# Printing the structure of articles_dict for each source
for source, articles_dict in zip([ 'Business Insider', 'TIME', 'Telegraph', 'StatesMan', "BBC"],
                                 [ bi_dict, time_dict, tg_dict, sm_dict, bbc_dict]):
    print(f"{source}: {type(next(iter(articles_dict.values())))}")
```

```
Business Insider: <class 'list'>
TIME: <class 'list'>
Telegraph: <class 'list'>
StatesMan: <class 'list'>
BBC: <class 'list'>
```

```python
from nltk.tokenize import word_tokenize
from nltk.corpus import stopwords
from nltk.stem import WordNetLemmatizer
from gensim import corpora, models


# Additional custom stop words, including 'say' and 'u'
custom_stop_words = ['said', 'along', 'say', 'kumar', 'israel', 'u', 'hamas', '2023', '2024', 'sunday', 'house', '7', 'u', 'also', 'would']


# Function to preprocess text
def preprocess_text(text_list):
    stop_words = set(stopwords.words('english') + custom_stop_words)
    lemmatizer = WordNetLemmatizer()
    tokens = [word_tokenize(sentence.lower()) for sentence in text_list]
    tokens = [
        lemmatizer.lemmatize(token)
        for sentence_tokens in tokens
        for token in sentence_tokens
        if token.isalnum() and token.lower() not in stop_words
    ]
    return tokens


# Function to split keys and perform topic modeling
def split_keys_and_topic_modeling(news_dict, dict_name):
    mid_point = len(news_dict) // 2
    first_set_articles = list(news_dict.values())[:mid_point]
    second_set_articles = list(news_dict.values())[mid_point:]

    # Preprocess text data for the first set of articles
    processed_texts_first_set = [preprocess_text(article) for article in first_set_articles]

    # Create dictionary and corpus for the first set
    dictionary_first_set = corpora.Dictionary(processed_texts_first_set)
    corpus_first_set = [dictionary_first_set.doc2bow(text) for text in processed_texts_first_set]

    # Build the LDA model for the first set
    lda_model_first_set = models.LdaModel(corpus_first_set, num_topics=5, id2word=dictionary_first_set, passes=15)

    # Print the topics for the first set along with dictionary name
    print(f"Topics for the First Set of Articles - {dict_name}:")
    for idx, topic in lda_model_first_set.print_topics(-1):
        print(f'Topic #{idx + 1}: {topic}')

    # Preprocess text data for the second set of articles
    processed_texts_second_set = [preprocess_text(article) for article in second_set_articles]

    # Create dictionary and corpus for the second set
    dictionary_second_set = corpora.Dictionary(processed_texts_second_set)
    corpus_second_set = [dictionary_second_set.doc2bow(text) for text in processed_texts_second_set]

    # Build the LDA model for the second set
    lda_model_second_set = models.LdaModel(corpus_second_set, num_topics=5, id2word=dictionary_second_set, passes=15)

    # Print the topics for the second set along with dictionary name
    print(f"\nTopics for the Second Set of Articles - {dict_name}:")
    for idx, topic in lda_model_second_set.print_topics(-1):
        print(f'Topic #{idx + 1}: {topic}')

# Assuming you have dictionaries bbc_dict, bi_dict, time_dict, tg_dict, sm_dict
# Apply the function to each news dictionary
for dict_name, news_dict in [("bbc_dict", bbc_dict), ("bi_dict", bi_dict), ("time_dict", time_dict), ("tg_dict", tg_dict), ("sm_dict", sm_di
    split_keys_and_topic_modeling(news_dict, dict_name)
```

```
Topics for the First Set of Articles - bbc_dict:
Topic #1: 0.016*"iran" + 0.014*"say" + 0.012*"gaza" + 0.010*"mosque" + 0.009*"war" + 0.008*"attack" + 0.008*"church" + 0.007*"many" +
Topic #2: 0.024*"gaza" + 0.011*"child" + 0.010*"say" + 0.009*"u" + 0.009*"attack" + 0.008*"un" + 0.008*"war" + 0.008*"killed" + 0.008*
Topic #3: 0.032*"gaza" + 0.018*"palestinian" + 0.017*"killed" + 0.016*"israeli" + 0.015*"attack" + 0.012*"west" + 0.012*"bank" + 0.011
Topic #4: 0.021*"say" + 0.010*"attack" + 0.010*"gaza" + 0.009*"hostage" + 0.009*"woman" + 0.009*"one" + 0.008*"border" + 0.008*"octobe
Topic #5: 0.026*"gaza" + 0.010*"area" + 0.009*"people" + 0.009*"israeli" + 0.008*"damage" + 0.008*"idf" + 0.008*"building" + 0.006*"ci

Topics for the Second Set of Articles - bbc_dict:
Topic #1: 0.022*"israeli" + 0.019*"hospital" + 0.017*"gaza" + 0.016*"attack" + 0.016*"killed" + 0.015*"palestinian" + 0.010*"member" +
Topic #2: 0.032*"gaza" + 0.014*"attack" + 0.013*"people" + 0.013*"un" + 0.013*"war" + 0.012*"unrwa" + 0.011*"israeli" + 0.011*"october
Topic #3: 0.014*"gaza" + 0.013*"say" + 0.009*"attack" + 0.008*"u" + 0.008*"israeli" + 0.007*"iran" + 0.006*"one" + 0.006*"day" + 0.005
Topic #4: 0.034*"gaza" + 0.010*"area" + 0.010*"people" + 0.009*"building" + 0.009*"damage" + 0.008*"war" + 0.008*"say" + 0.007*"bbc" +
Topic #5: 0.018*"gaza" + 0.016*"say" + 0.015*"child" + 0.009*"many" + 0.008*"one" + 0.008*"killed" + 0.007*"family" + 0.007*"mosque" +
Topics for the First Set of Articles - bi_dict:
```

```
Topic #1: 0.021*"israeli" + 0.016*"military" + 0.011*"gaza" + 0.009*"attack" + 0.009*"since" + 0.009*"hostage" + 0.009*"time" + 0.009*
Topic #2: 0.018*"attack" + 0.012*"strike" + 0.012*"arouri" + 0.009*"killed" + 0.009*"official" + 0.009*"drone" + 0.009*"hit" + 0.009*
Topic #3: 0.046*"tunnel" + 0.008*"war" + 0.008*"cong" + 0.008*"viet" + 0.008*"army" + 0.007*"roman" + 0.007*"christmas" + 0.007*"city"
Topic #4: 0.012*"safe" + 0.012*"resident" + 0.010*"body" + 0.010*"heard" + 0.008*"attack" + 0.008*"time" + 0.008*"community" + 0.008*
Topic #5: 0.020*"idf" + 0.018*"shark" + 0.016*"soldier" + 0.010*"tank" + 0.010*"monday" + 0.010*"per" + 0.008*"time" + 0.008*"israeli"

Topics for the Second Set of Articles - bi_dict:
Topic #1: 0.013*"hostage" + 0.012*"gaza" + 0.010*"idf" + 0.010*"eisenkot" + 0.008*"rocket" + 0.007*"killed" + 0.007*"force" + 0.007*"s
Topic #2: 0.017*"new" + 0.011*"time" + 0.011*"killed" + 0.010*"attack" + 0.010*"year" + 0.009*"israeli" + 0.009*"netanyahu" + 0.008*"g
Topic #3: 0.030*"tank" + 0.020*"israeli" + 0.009*"woman" + 0.009*"october" + 0.007*"gaza" + 0.007*"war" + 0.007*"merkava" + 0.006*"idf
Topic #4: 0.019*"mitnick" + 0.015*"military" + 0.011*"israeli" + 0.011*"war" + 0.011*"people" + 0.009*"palestinian" + 0.009*"prison" +
Topic #5: 0.001*"israeli" + 0.001*"gaza" + 0.001*"attack" + 0.001*"killed" + 0.001*"time" + 0.001*"new" + 0.001*"war" + 0.001*"eisenko
Topics for the First Set of Articles - time_dict:
Topic #1: 0.013*"gaza" + 0.010*"israeli" + 0.010*"palestinian" + 0.008*"killed" + 0.007*"military" + 0.007*"war" + 0.007*"hostage" + 0
Topic #2: 0.008*"gaza" + 0.008*"houthis" + 0.007*"war" + 0.006*"israeli" + 0.006*"people" + 0.005*"biden" + 0.005*"attack" + 0.005*"pa
Topic #3: 0.010*"gaza" + 0.007*"world" + 0.007*"war" + 0.005*"country" + 0.005*"time" + 0.004*"israeli" + 0.004*"one" + 0.004*"aid" +
Topic #4: 0.013*"israeli" + 0.008*"netanyahu" + 0.007*"war" + 0.007*"gaza" + 0.007*"military" + 0.006*"journalist" + 0.005*"court" + 0
Topic #5: 0.012*"palestinian" + 0.010*"gaza" + 0.008*"israeli" + 0.005*"war" + 0.004*"force" + 0.004*"say" + 0.004*"poetry" + 0.003*"p

Topics for the Second Set of Articles - time_dict:
Topic #1: 0.016*"child" + 0.010*"israeli" + 0.009*"palestinian" + 0.006*"military" + 0.004*"tiktok" + 0.004*"content" + 0.004*"many" +
Topic #2: 0.014*"gaza" + 0.011*"resolution" + 0.010*"humanitarian" + 0.009*"ceasefire" + 0.007*"civilian" + 0.007*"council" + 0.007*"v
Topic #3: 0.008*"nuclear" + 0.007*"palestinian" + 0.006*"iran" + 0.006*"state" + 0.005*"woman" + 0.005*"israeli" + 0.005*"gaza" + 0.00
Topic #4: 0.010*"israeli" + 0.009*"gaza" + 0.008*"journalist" + 0.007*"military" + 0.007*"palestinian" + 0.006*"killed" + 0.006*"hosta
Topic #5: 0.010*"gaza" + 0.008*"israeli" + 0.007*"war" + 0.007*"people" + 0.007*"say" + 0.006*"palestinian" + 0.005*"day" + 0.004*"one
Topics for the First Set of Articles - tg_dict:
Topic #1: 0.013*"court" + 0.008*"case" + 0.007*"palestinian" + 0.007*"gaza" + 0.007*"order" + 0.006*"genocide" + 0.006*"official" + 0
Topic #2: 0.013*"gaza" + 0.011*"israeli" + 0.008*"palestinian" + 0.007*"war" + 0.007*"official" + 0.006*"security" + 0.006*"people" +
Topic #3: 0.014*"court" + 0.013*"gaza" + 0.012*"israeli" + 0.010*"south" + 0.008*"palestinian" + 0.008*"international" + 0.008*"africa
Topic #4: 0.008*"gaza" + 0.008*"unrwa" + 0.007*"palestinian" + 0.007*"syria" + 0.007*"israeli" + 0.006*"medication" + 0.005*"un" + 0.0
Topic #5: 0.024*"gaza" + 0.021*"israeli" + 0.012*"palestinian" + 0.010*"killed" + 0.010*"military" + 0.009*"official" + 0.009*"war" +

Topics for the Second Set of Articles - tg_dict:
Topic #1: 0.020*"israeli" + 0.007*"gaza" + 0.006*"new" + 0.006*"embassy" + 0.006*"security" + 0.006*"military" + 0.005*"delhi" + 0.004
Topic #2: 0.024*"gaza" + 0.016*"israeli" + 0.012*"palestinian" + 0.009*"war" + 0.008*"people" + 0.008*"civilian" + 0.008*"killed" + 0
Topic #3: 0.012*"israeli" + 0.008*"gaza" + 0.008*"october" + 0.007*"official" + 0.006*"war" + 0.006*"military" + 0.006*"time" + 0.005
Topic #4: 0.015*"israeli" + 0.014*"gaza" + 0.014*"military" + 0.011*"hostage" + 0.008*"official" + 0.008*"killed" + 0.007*"wednesday"
Topic #5: 0.008*"tunnel" + 0.006*"gaza" + 0.006*"voter" + 0.006*"biden" + 0.006*"israeli" + 0.005*"among" + 0.005*"military" + 0.004*
Topics for the First Set of Articles - sm_dict:
Topic #1: 0.030*"gaza" + 0.022*"israeli" + 0.018*"minister" + 0.016*"defence" + 0.016*"military" + 0.014*"control" +
Topic #2: 0.019*"state" + 0.017*"u" + 0.017*"unrwa" + 0.017*"allegation" + 0.014*"united" + 0.013*"attack" + 0.012*"october" + 0.012*
Topic #3: 0.015*"israeli" + 0.013*"killed" + 0.013*"idf" + 0.013*"hezbollah" + 0.013*"attack" + 0.011*"war" + 0.011*"hostage" + 0.009
Topic #4: 0.019*"hostage" + 0.014*"gaza" + 0.014*"case" + 0.013*"south" + 0.013*"africa" + 0.012*"israeli" + 0.010*"ceasefire" + 0.01
```

Part 2

```
pip install wordcloud matplotlib nltk gensim
```

```
Requirement already satisfied: wordcloud in /usr/local/lib/python3.10/dist-packages (1.9.3)
Requirement already satisfied: matplotlib in /usr/local/lib/python3.10/dist-packages (3.7.1)
Requirement already satisfied: nltk in /usr/local/lib/python3.10/dist-packages (3.8.1)
Requirement already satisfied: gensim in /usr/local/lib/python3.10/dist-packages (4.3.2)
Requirement already satisfied: numpy>=1.6.1 in /usr/local/lib/python3.10/dist-packages (from wordcloud) (1.23.5)
Requirement already satisfied: pillow in /usr/local/lib/python3.10/dist-packages (from wordcloud) (9.4.0)
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: pyparsing>=2.3.1 in /usr/local/lib/python3.10/dist-packages (from matplotlib) (3.1.1)
Requirement already satisfied: python-dateutil>=2.7 in /usr/local/lib/python3.10/dist-packages (from matplotlib) (2.8.2)
Requirement already satisfied: click in /usr/local/lib/python3.10/dist-packages (from nltk) (8.1.7)
Requirement already satisfied: joblib in /usr/local/lib/python3.10/dist-packages (from nltk) (1.3.2)
Requirement already satisfied: regex>=2021.8.3 in /usr/local/lib/python3.10/dist-packages (from nltk) (2023.6.3)
Requirement already satisfied: tqdm in /usr/local/lib/python3.10/dist-packages (from nltk) (4.66.1)
Requirement already satisfied: scipy>=1.7.0 in /usr/local/lib/python3.10/dist-packages (from gensim) (1.11.4)
Requirement already satisfied: smart-open>=1.8.1 in /usr/local/lib/python3.10/dist-packages (from gensim) (6.4.0)
Requirement already satisfied: six>=1.5 in /usr/local/lib/python3.10/dist-packages (from python-dateutil>=2.7->matplotlib) (1.16.0)
```

```python
from nltk.tokenize import word_tokenize
from nltk.corpus import stopwords
from wordcloud import WordCloud

# Define custom stop words
custom_stop_words = ['said', 'along', 'say', 'kumar', 'israel', 'u', 'hamas', '2023', '2024', 'sunday', 'house', '7', 'u', 'also', 'would']

# Tokenize and remove stopwords for each dictionary
tokenized_dicts = {}

for source, articles_dict in zip(['Business Insider', 'TIME', 'Telegraph', 'StatesMan', 'BBC'],
                                 [bi_dict, time_dict, tg_dict, sm_dict, bbc_dict]):
    tokenized_articles = []
    for articles_list in articles_dict.values():
        for article in articles_list:
            tokens = word_tokenize(article.lower())  # Lowercasing for consistency
            filtered tokens   [word for word in tokens if word isalpha() and word not in custom stop words and word not in stopwords words('
```

```python
import matplotlib.pyplot as plt

plt.figure(figsize=(10, 5))
plt.imshow(wordcloud, interpolation='bilinear')
plt.axis('off')
plt.show()
```



```python
from gensim.models import Word2Vec

# Train Word2Vec model on each dictionary
word2vec_models = {}

for source, tokenized_articles in tokenized_dicts.items():
    model = Word2Vec([tokenized_articles], vector_size=100, window=5, min_count=1, workers=4)
    word2vec_models[source] = model

# Find 5 most similar words with 'economy' in each dictionary
for source, model in word2vec_models.items():
    similar_words = model.wv.most_similar('war', topn=5)
    print(f"{source} - Similar words with 'war': {similar_words}")
```

```
    Business Insider - Similar words with 'war': [('night', 0.3189409673213959), ('background', 0.29267844557762146), ('display', 0.28168058
    TIME - Similar words with 'war': [('food', 0.4354916512966156), ('theory', 0.3917715847492218), ('populations', 0.356428861618042), ('pe
    Telegraph - Similar words with 'war': [('israeli', 0.6901981234550476), ('people', 0.6637016534805298), ('military', 0.6286309361457825)
    StatesMan - Similar words with 'war': [('gaza', 0.9218995571136475), ('law', 0.9040326476097107), ('government', 0.8981812000274658), ('
    BBC - Similar words with 'war': [('gaza', 0.6964634656906128), ('un', 0.6858299970626831), ('israeli', 0.6649380326271057), ('says', 0.6
```

```python
# Combine all tokenized articles for a global Word2Vec model
global_model = Word2Vec([all_tokens], vector_size=100, window=5, min_count=1, workers=4)

# Find 5 most similar words with 'economy' across all dictionaries
x = "currency"
```