

# NLP

Find your favorite news source and grab the article text.

1. Show the most common words in the article.
2. Show the most common words under a part of speech. (i.e. NOUN: {'Bob':12, 'Alice':4,})
3. Find a subject/object relationship through the dependency parser in any sentence.
4. Show the most common Entities and their types.
5. Find Entities and their dependency (hint: entity.root.head)
6. Find the most similar words in the article

Note: Yes, the notebook from the video is not provided, I leave it to you to make your own :) it's your final assignment for the semester. Enjoy!

```
In [1]: import spacy
        nlp = spacy.load("en_core_web_sm")
```

```
In [2]: txt = open('./economist.txt', newline='\n', encoding='utf-8').read()
```

```
In [3]: doc = nlp(txt)
        doc
```

Out[3]: On what terms could the war in Ukraine stop?

Russia's lightning attack on Ukraine's capital, Kyiv, was a failure. Its creeping artillery war to seize the eastern region of Donbas has ground to a bloody halt. It has lost a chunk of stolen territory south of the city of Kharkiv, and this week announced a retreat from Kherson, the only provincial capital it had captured since its invasion in February. With each setback, Vladimir Putin, Russia's president, has sought new ways to torment Ukraine. The latest is a relentless bombardment that seeks to wreck Ukraine's infrastructure. Residents of the capital have been told they may have to evacuate if the power grid collapses, halting water and sewage services.

Power cuts have not sapped Ukraine's will to fight. But they are a reminder that, eight months after his unprovoked invasion, Mr Putin keeps looking for ways to raise the stakes. Some worry he might blow up a dam on the Dnieper river, as Stalin did in 1941, to slow his adversaries' advance.

The ever-evolving Russian assault also raises an awkward question: how long will America and Europe keep providing Ukraine with the billions of dollars' worth of military and economic aid it needs every month to fend Russia off? "For as long as it takes," say Western leaders. But many of their citizens reject the idea of bankrolling an indefinite conflict with Russia. Tens of thousands of people took to the streets of Rome on November 5th, calling for an end to the fighting. "We don't want war. No weapons, no sanctions. Where is diplomacy?" read one placard.

In America, too, questions have been raised. Hard-left Democrats recently issued a call, swiftly retracted, for negotiations. Gains by America-first Republicans in the midterm elections on November 8th, although smaller than expected, are a reminder that American politics might change dramatically after the next presidential election, in two years' time, and with it policy on Ukraine.

Jake Sullivan, Mr Biden's national-security adviser, made an unannounced trip to Kyiv on November 4th to promise "unwavering" support. But he also urged Ukraine to think about future peace terms. It has since emerged that he has been in touch with his Russian counterparts, to warn them not to use nuclear weapons. On November 9th Mr Biden said Russia and Ukraine would "lick their wounds" after the battle for Kherson, and might then be ready for compromise. He insisted he would not tell Ukraine what to do.

In private, Western and Ukrainian officials are starting to ponder what a stable outcome might look like. Will Ukraine become a new Finland, forced to cede land to its invaders and to remain neutral for decades? Or another West Germany, with its national territory partitioned by war and its democratic half absorbed into NATO? A much-discussed template is Israel, a country under constant threat that has been able to defend itself without formal alliances but with extensive military help from America.

The precise terms of any negotiated settlement depend on what happens on the battlefield. There is likely to be a lot more fighting before either side is ready to end the war. Russia and Ukraine have each lost, by one estimate, roughly 100,000 soldiers, killed and wounded, but both still hope to manoeuvre to a more favourable position.

The retreat from Kherson is a humiliation for Mr Putin. But it will give Russian forces a more easily defended line along the Dnieper river. Mr Putin shows no sign of throwing in the towel. He has mobilised hundreds of thousands more recruits. Some have been rushed into battle with little training or equipment to hold the line; the rest may be used for a renewed push next year.

Ukraine, for its part, hopes to maintain its momentum. Its army is getting reinforcements this winter, in the form of thousands of recruits trained by Britain and other Western countries. Western arms continue to arrive. On November 4th the Pentagon announced another arms package, worth \$400m, including 45 refurbished t-72b tanks and 1,100 drones. The first new nasams anti-aircraft batteries were deployed this week.

The West's stocks of weapons are not unlimited. European armies have eaten deep into theirs; even mighty America worries about eroding its own ability to fight future wars. It is Russia, however, that seems to face the most immediate shortages. It has used up most of its precision bombs and missiles, and is struggling to replace them because of sanctions. It is obtaining fresh weapons from the likes of Iran and perhaps North Korea. (China has so far heeded American warnings to stay out of the war.)

Cold calculation

Mr Putin is hoping his campaign to destroy Ukraine's electricity grid will freeze the country into submission, or at least turn it into a weak, failing state. But the evidence of past conflicts is that aerial bombing of civilians, in the absence of an effective gr

ound campaign, rarely secures victory. Nearly 90% of Ukrainians want the country to keep fighting.

In Russia, according to the Levada Centre, a pollster, only 36% want to press on with the war, whereas 57% favour peace talks. At the same time, support for Mr Putin remains at 79%. Russians, it seems, would like the war to end but, starved of impartial news, do not blame Mr Putin for it. Still, the more he tries to drag them into fighting, the more he risks losing popular support.

Ukraine's more avid Western supporters think that, with time, Ukraine will become stronger, and Russia weaker. But Mr Putin is hoping that "General Winter" will somehow revive his fortunes, if not by weakening Ukraine's will to fight then by gnawing at the West's readiness to support it, as the heating bills balloon in Europe.

Mr Putin claims that he is ready to negotiate (from the starting point that the West should recognise his theft of Ukrainian territory) but that Ukraine's Western "masters" have prevented it from talking. The two sides held lengthy talks after Russia seized the Crimean peninsula and part of Donbas in 2014. They talked again in the spring, as Russia besieged Kyiv. But Ukraine set its face against further negotiations after Russia's retreat from Kyiv in April revealed widespread atrocities against civilians. Volodymyr Zelensky, Ukraine's president, suggested this week that talks could be revived, but only if Russia was willing to give back Ukrainian land, pay compensation and accept responsibility for war crimes.

The West is vague about its own aims. Mr Biden has at times mused about wanting to see Mr Putin ousted from power; at others he has talked about finding "off-ramps" for the Russian leader. He defined his goals most clearly in a guest article in the New York Times in May: "a democratic, independent, sovereign and prosperous Ukraine with the means to deter and defend itself against further aggression". Notably, that left out the question of Ukraine's borders. Western leaders say this is for Ukraine to decide; their aim is to strengthen its negotiating hand.

More recently, though, Ukraine's backers have sounded more specific. In a statement on October 11th, leaders of the G7 group of industrialised countries offered their "full support to Ukraine's independence, territorial integrity and sovereignty in its internationally recognised borders". They demanded that Russia "completely and unconditionally withdraw" from all seized lands. Among other things, they pledged to find ways of using seized Russian assets to help fund Ukraine's reconstruction.

In need of a break

"The G7 statement is basically a demand for total Russian surrender, which is not a plausible diplomatic outcome. Diplomacy by definition involves give and take. The expectation should not be another Treaty of Versailles," says Samuel Charap of the RAND Corporation, an American think-tank, referring to the punitive terms imposed on Germany at the end of the first world war. The West, Ukraine and Russia, he argues, should start talking, if only to set the groundwork for more substantive negotiations in future: "Fighting and talking at the same time should be the norm."

Many disagree. "Keep up the pressure. Don't be in a hurry to draw lines on a map. It would be bureaucratic suicide. Somebody will put it up on Twitter alongside the Molotov-Ribbentrop pact," retorts Dan Fried of the Atlantic Council, another American think-tank, alluding to Nazi Germany's and the Soviet Union's carve-up of Poland in 1939.

Few Western leaders question Ukraine's ambition to recapture the territory lost since Russia invaded in February. Many would support efforts to reclaim the parts of Donbas seized in 2014. But opinion is more divided when it comes to reconquering Crimea. Many worry that the prospect of losing the peninsula might prompt a dangerous escalation from Mr Putin.

To some in the Biden administration, the war is a matter of principle: territory should never be seized by force, so all Russian gains must be reversed. Others, doubting Ukraine's ability to reconquer much more, think the time for diplomacy is soon. Either way, America is in no rush to spell out diplomatic positions that might cause rifts in the pro-Ukraine camp.

Another pressing concern is the nature of future Western security guarantees for Ukraine. They will need to be robust given that Russia will probably remain a threat to Ukraine for as long as Mr Putin is in power, if not longer. Several central and eastern European countries favour Ukraine's rapid admission to NATO, on the grounds that the alliance's commitment to mutual defence would firmly deter Russia. For all its nuclear menaces, it has so far refrained from overtly striking NATO territory.

The Biden administration, though, is wary of having to extend its nuclear umbrella to a

country in a state of latent or actual conflict with Russia. Throughout, Mr Biden has been careful to minimise the risk of a direct nato-Russia conflict for fear that it would lead to "World War III". Several nato members in western Europe are similarly sceptical. So attention has turned to interim, or alternative, arrangements. In September Anders Fogh Rasmussen, a former secretary-general of nato, and Andriy Yermak, Mr Zelensky's chief of staff, proposed a "Kyiv Security Compact" which would offer security assistance short of a mutual-defence pact. Some in Ukraine considered it a betrayal. Modelled on Western support for Israel, about which Mr Zelensky has spoken, the compact would strengthen Ukraine's armed forces—in effect turning the current ad hoc support into a systematic, long-term commitment.

Ukraine's partners would promise "multi-decade" investments in the country's defence industry, massive weapons transfers, training, joint exercises and intelligence support. The compact would require neither Russia's assent nor Ukraine's neutrality. It would not preclude nato membership. In some circumstances, there could be military intervention to help Ukraine. If it were attacked, the signatories would "use all elements of their national and collective power and take appropriate measures—which may include diplomatic, economic and military means." A wider group of countries, including Asian allies, would reinforce such military assistance with sanctions on Russia, including provisions to "snap back" any of the current penalties that may be lifted as part of a deal.

Even this may be too ambitious for Team Biden. Some ask, for instance, what commitments Ukraine would take on, by way of reforms to strengthen democracy, say, or to fight corruption. The parallel with Israel is imperfect. Among other things, Israel is a nuclear power and occupies Arab land. For Mykola Bielieskov of the National Institute for Strategic Studies, a think-tank in Kyiv, the Israeli model "is not only about mobilising our partners externally; it's also about explaining to people what it means to live next door to a crazy neighbour, to existential threats."

The sky must be the limit

Whatever the diplomatic template, Russia's blitz has proved that the West will need to help Ukraine create a properly integrated and layered air-defence system, mixing fighter aircraft, surface-to-air batteries and shoulder-launched weapons. At the moment weapons are arriving piecemeal and often cannot exchange data. There are also worries about Ukraine running out of certain types of air-defence munitions. If that were to happen Russia could then deploy much more air power in support of ground troops.

Uncertain harvest

Ukraine's mix-and-match weapons—a "Mr Potato Head" arsenal, as some call it—cause problems elsewhere. For instance, it has no fewer than 14 different types of artillery pieces, with an average brigade operating four different sorts. "This is a logistical nightmare for them, especially when we talk about ammunition," says Nick Reynolds of Rusi, a British think-tank. Some of the weapons are wearing out badly, and Europe's defence industry, sapped by decades of low spending, is poorly placed to produce spare parts. "Red lights are flashing in terms of that support being available," adds Mr Reynolds.

How long the war goes on depends mainly on Mr Putin. He is in a bind, both in Ukraine and at home. Moderate technocrats are worried about the strains on the economy; "national patriots" such as Yevgeny Prigozhin, who commands the Wagner mercenary group, have called for purges of supposedly treacherous generals.

A pause to play at diplomacy may suit Mr Putin for a time—particularly if it allows him to consolidate some territorial gains. That may explain his recent toning down of nuclear rhetoric and his sudden casting of Ukrainians as victims of Western aggression. "The West is throwing Ukrainians into a furnace"; Russia, in contrast, "has always treated Ukrainian people with respect," Mr Putin declared on November 4th. (His propagandists and officials, though, still talk of "de-Satanising" Ukraine.) The shift fits another of Mr Putin's guises, as the champion of a global movement to cast off Western dominance.

In all this Mr Putin is seeking to woo waverers, especially in the global South. He also wants to reassure friends such as China and India, which have made plain their disapproval of his nuclear recklessness. Above all Mr Putin is interested in reaching one rich-world listener: Donald Trump, whose allies in Congress question American aid to Ukraine and who may soon announce another run for president.

For all his setbacks, Mr Putin is not yet out of options to persecute Ukraine and try to divide the West. Militarily, he could commit more of his air force and mobilise more troops. In the covert "grey zone" he could sabotage undersea gas pipelines and internet connections to the West, conduct bigger cyber-attacks, interfere with communications satellites and step up disinformation campaigns. He could also sink ships carrying grain from

Ukraine. Ultimately, he could use tactical nuclear weapons. All this, though, would come at heavy cost: it would make Russia even more of a pariah, weaken him at home and might provoke harsh retaliation.

The stakes are higher for Mr Putin than for the West. But they are highest for Ukrainians, many of whom mistrust the very idea of talks with Russia and see military victory as their only option—even if it takes years to achieve. The more land it can regain, Ukraine reckons, the greater the chance of getting rid of Mr Putin. Yet that same prospect alarms many in the West: a rout of the Russian army might be what pushes Mr Putin to go nuclear. That is one reason why Team Biden long ago stopped talking about helping Ukraine “win”.

As it has often done with Israel, America may at some point try to limit Ukraine’s ambitions. It need not do so overtly; it can simply withhold the arms Ukraine needs, as it already does to an extent. It declines to provide Western aircraft, Patriot air-defence missiles and longer-range attack strike missiles for fear that they might goad Russia to use nuclear weapons.

All this explains why some Ukrainians have been circulating a bittersweet message issued to Finnish troops in 1940 at the end of the “Winter War” with the Soviet Union by their commander, Carl Gustaf Mannerheim. The vastly outnumbered Finns had inflicted heavy losses on Soviet forces but nonetheless had to cede territory because help from their friends had dried up, Mannerheim wrote, signing off with the words, “We are proudly conscious of the historic duty which we shall continue to fulfil; the defence of that Western civilisation which has been our heritage for centuries, but we know also that we have paid to the very last penny any debt we may have owed the West.”

Ukraine’s fate depends not just on the valour of its soldiers or resilience of its people, but also on external factors it cannot control: the inscrutable calculations of Russia’s despotic ruler and the fortitude of its friends. The benefits to the West of the war are already clear. Russia has been enormously weakened, making Europe’s flank much easier to defend. For Ukraine, which has suffered horrific losses, the outcome looks much less certain.

## 1. Show the most common words in the article.

```
In [4]: from collections import Counter
```

```
In [26]: token_list = []
for token in doc:
    if token.is_stop == False:
        if token.is_punct == False:
            if token.is_space == False:
                lem = token.lemma_
                token_list.append(lem)
```

```
In [6]: counted = Counter(token_list)
counted.most_common(15)
```

```
Out[6]: [('Ukraine', 52),
          ('Russia', 30),
          ('Mr', 29),
          ('Putin', 22),
          ('western', 15),
          ('war', 14),
          ('West', 14),
          ('talk', 12),
          ('weapon', 11),
          ('support', 11),
          ('nuclear', 9),
          ('country', 9),
          ('russian', 8),
          ('long', 8),
          ('Biden', 8)]
```

## 2. Show the most common words under a part of speech. (i.e. NOUN: {'Bob':12, 'Alice':4,})

```
In [7]: test = doc[50]
        [test.text, test.pos_, test.lemma_]
```

```
Out[7]: ['stolen', 'VERB', 'steal']
```

```
In [8]: import pandas as pd
        df = pd.DataFrame(columns=['token', 'pos', 'lemma'])
        df['token'] = [token for token in doc]
        df['pos'] = [token.pos_ for token in doc]
        df['lemma'] = [token.lemma_ for token in doc]
```

```
In [9]: df.head()
```

```
Out[9]:
```

	token	pos	lemma
0	On	ADP	on
1	what	PRON	what
2	terms	NOUN	term
3	could	AUX	could
4	the	DET	the

```
In [10]: noun_sub = df[df['pos']=='NOUN']
         noun_sub.head()
```

```
Out[10]:
```

	token	pos	lemma
2	terms	NOUN	term
5	war	NOUN	war
8	stop	NOUN	stop
13	lightning	NOUN	lightning
14	attack	NOUN	attack

```
In [11]: noun_count = Counter(noun_sub['lemma'])
         print('NOUN:', noun_count.most_common(10))
```

```
NOUN: [('war', 14), ('weapon', 11), ('s', 9), ('support', 9), ('country', 9), ('defence', 8), ('territory', 7), ('power', 7), ('time', 7), ('term', 6)]
```

## 3. Find a subject/object relationship through the dependency parser in any sentence.

```
In [12]: def pr_tree(word, level):
         if word.is_punct:
             return
         for child in word.lefts:
             pr_tree(child, level+1)
         print('\t'* level + word.text + ' - ' + word.dep_)
```

```
    for child in word.rights:
        pr_tree(child, level+1)
```

```
In [13]: sent = doc[48:66]
```

```
In [14]: pr_tree(sent.root, 0)
```

```
      a - det
chunk - dobj
      of - prep
              stolen - amod
              territory - pobj
```

Based on the above parsed sentence, "region" is the dependent object of the verb "to seize," which modifies the main subject, "war." The main verb, "ground," has no dependent object.

## 4. Show the most common Entities and their types.

```
In [15]: for entity in doc.ents:
          print(entity, entity.label_)
```

Ukraine GPE  
Russia GPE  
Ukraine GPE  
Kyiv PERSON  
Donbas GPE  
Kharkiv GPE  
this week DATE  
Kherson PERSON  
February DATE  
Vladimir Putin PERSON  
Russia GPE  
Ukraine GPE  
Ukraine GPE  
Ukraine GPE  
eight months DATE  
Putin PERSON  
Stalin PERSON  
1941 DATE  
Russian NORP  
America GPE  
Europe LOC  
Ukraine GPE  
the billions of dollars MONEY  
every month DATE  
Russia GPE  
Western NORP  
Russia GPE  
Tens of thousands CARDINAL  
Rome GPE  
November 5th DATE  
one CARDINAL  
America GPE  
Democrats NORP  
America GPE  
Republicans NORP  
November 8th DATE  
American NORP  
two years DATE  
Ukraine GPE  
Jake Sullivan PERSON  
Biden PERSON  
November 4th DATE  
Ukraine GPE  
Russian NORP  
November 9th DATE  
Biden PERSON  
Russia GPE  
Ukraine GPE  
Kherson PERSON  
Ukraine GPE  
Western NORP  
Ukrainian GPE  
Ukraine GPE  
Finland GPE  
decades DATE  
West Germany GPE  
half CARDINAL  
Israel GPE  
America GPE  
Russia GPE  
Ukraine GPE  
one CARDINAL



roughly 100,000 CARDINAL  
Kherson PERSON  
Putin PERSON  
Russian NORP  
Dnieper ORG  
Putin PERSON  
hundreds of thousands CARDINAL  
next year DATE  
Ukraine GPE  
this winter DATE  
thousands CARDINAL  
Britain GPE  
November 4th DATE  
Pentagon ORG  
400 MONEY  
45 CARDINAL  
1,100 CARDINAL  
first ORDINAL  
this week DATE  
European NORP  
America GPE  
Russia GPE  
Iran GPE  
North Korea GPE  
China GPE  
American NORP  
Putin PERSON  
Ukraine GPE  
Nearly 90% PERCENT  
Ukrainians NORP  
Russia GPE  
the Levada Centre ORG  
only 36% PERCENT  
57% PERCENT  
Putin PERSON  
79% PERCENT  
Russians NORP  
Mr Putin PERSON  
Ukraine GPE  
Ukraine GPE  
Russia GPE  
Putin PERSON  
Ukraine GPE  
Europe LOC  
Putin PERSON  
West LOC  
Ukrainian NORP  
Ukraine GPE  
two CARDINAL  
Russia GPE  
Crimean NORP  
Donbas GPE  
2014 DATE  
the spring DATE  
Russia GPE  
Kyiv PERSON  
Ukraine GPE  
Russia GPE  
Kyiv PERSON  
April DATE  
Volodymyr Zelensky PERSON  
Ukraine GPE

this week DATE  
Russia GPE  
Ukrainian GPE  
West LOC  
Biden PERSON  
Mr Putin PERSON  
Russian NORP  
the New York Times ORG  
May DATE  
Ukraine GPE  
Ukraine GPE  
Ukraine GPE  
Ukraine GPE  
October 11th DATE  
Ukraine GPE  
Russia GPE  
Russian NORP  
Ukraine GPE  
Russian NORP  
Samuel Charap PERSON  
American NORP  
Germany GPE  
the end of the first world war DATE  
West LOC  
Ukraine GPE  
Russia GPE  
Twitter PERSON  
Dan Fried PERSON  
the Atlantic Council ORG  
American NORP  
Nazi Germany's GPE  
the Soviet Union's GPE  
Poland GPE  
1939 DATE  
Ukraine GPE  
Russia GPE  
February DATE  
Donbas ORG  
2014 DATE  
Crimea ORG  
Putin PERSON  
Biden PERSON  
Russian NORP  
Ukraine GPE  
America GPE  
pro-Ukraine NORP  
Ukraine GPE  
Russia GPE  
Ukraine GPE  
Putin PERSON  
European NORP  
Ukraine GPE  
nato ORG  
Russia GPE  
nato ORG  
Biden PERSON  
Russia GPE  
Mr Biden PERSON  
Russia GPE  
World War III EVENT  
nato ORG  
Europe LOC

September DATE  
nato ORG  
Andriy Yermak PERSON  
Zelensky PERSON  
a "Kyiv Security Compact WORK\_OF\_ART  
Ukraine GPE  
Israel GPE  
Zelensky PERSON  
Ukraine GPE  
Ukraine GPE  
multi-decade DATE  
Russia GPE  
Ukraine GPE  
nato ORG  
Ukraine GPE  
Asian NORP  
Russia GPE  
Team Biden ORG  
Ukraine GPE  
Israel GPE  
Israel GPE  
Arab NORP  
Mykola Bielineskov PERSON  
the National Institute for Strategic Studies ORG  
Kyiv GPE  
Israeli NORP  
Russia GPE  
West LOC  
Ukraine GPE  
Ukraine GPE  
Russia GPE  
Ukraine GPE  
Mr Potato Head" WORK\_OF\_ART  
14 CARDINAL  
four CARDINAL  
Nick Reynolds PERSON  
British NORP  
Europe LOC  
decades DATE  
Mr Reynolds PERSON  
Mr Putin PERSON  
Ukraine GPE  
Yevgeny Prigozhin PERSON  
Wagner ORG  
Mr Putin PERSON  
Ukrainians NORP  
West LOC  
Ukrainians NORP  
Russia GPE  
Ukrainian NORP  
Putin PERSON  
November 4th DATE  
Ukraine GPE  
Putin PERSON  
Western NORP  
Putin PERSON  
South LOC  
China GPE  
India GPE  
Putin PERSON  
one CARDINAL  
Donald Trump PERSON

Congress ORG  
American NORP  
Ukraine GPE  
Putin PERSON  
Ukraine GPE  
Ukraine GPE  
Russia GPE  
Putin PERSON  
Ukrainians NORP  
Russia GPE  
years DATE  
Ukraine GPE  
Putin PERSON  
Russian NORP  
Mr Putin PERSON  
one CARDINAL  
Team Biden PERSON  
long ago DATE  
Ukraine GPE  
Israel GPE  
America GPE  
Ukraine GPE  
Ukraine GPE  
Patriot ORG  
Russia GPE  
Ukrainians NORP  
Finnish NORP  
1940 DATE  
Winter War EVENT  
the Soviet Union GPE  
Carl Gustaf Mannerheim PERSON  
Finns NORP  
Soviet NORP  
Mannerheim PERSON  
Western NORP  
centuries DATE  
last penny MONEY  
West LOC  
Ukraine GPE  
Russia GPE  
Russia GPE  
Europe LOC  
Ukraine GPE

```
In [16]: ent_list = []  
        label_list = []  
        for entity in doc.ents:  
            ent_list.append(entity.lemma_)  
            label_list.append(entity.label_)  
        ent_list[0:5]
```

```
Out[16]: ['Ukraine', 'Russia', 'Ukraine', 'Kyiv', 'Donbas']
```

```
In [17]: label_list[0:5]
```

```
Out[17]: ['GPE', 'GPE', 'GPE', 'PERSON', 'GPE']
```

```
In [18]: ent_df = pd.DataFrame(columns=['entity', 'label'])  
        ent_df['entity'], ent_df['label'] = ent_list, label_list  
        ent_df.head()
```

```
Out [18]:
```

	entity	label
0	Ukraine	GPE
1	Russia	GPE
2	Ukraine	GPE
3	Kyiv	PERSON
4	Donbas	GPE

```
In [19]: ent_df.groupby(['entity', 'label']).size().sort_values(ascending=False)[0:10]
```

```
Out[19]:
```

entity	label	
Ukraine	GPE	52
Russia	GPE	30
Putin	PERSON	16
russian	NORP	8
America	GPE	7
West	LOC	6
american	NORP	5
Ukrainians	NORP	5
Europe	LOC	5
Mr Putin	PERSON	5

dtype: int64

## 5. Find Entites and their dependency (hint: entity.root.head)

```
In [20]: for entity in doc.ents[0:10]:
          print (entity,entity.root.head)
```

```
Ukraine stop
Russia attack
Ukraine capital
Kyiv attack
Donbas of
Kharkiv of
this week announced
Kherson from
February in
Vladimir Putin sought
```

## 6. Find the most similar words in the article

```
In [33]: import numpy as np
token_list = []
for token in doc:
    if token.is_stop == False:
        if token.is_punct == False:
            if token.is_space == False:
                token_list.append(token)
```

```
In [34]: token_list[0:5]
```

```
Out[34]: [terms, war, Ukraine, stop, Russia]
```

```
In [56]: from spacy.tokens import Token, Doc
import warnings
```

```
warnings.filterwarnings('ignore')
```

```
In [71]: sim = pd.DataFrame(columns=['token1', 'token2', 'sim_score'])
count = 0
for token in token_list:
    count += 1
    scores = pd.DataFrame(columns=['token1', 'token2', 'sim_score'])
    for token2 in token_list:
        score = token.similarity(token2)
        scores = scores.append([{'token1': token.text,
                                'token2': token2.text,
                                'sim_score': score}])

    scores_sub = scores[scores['token1'] != scores['token2']]
    score_max = scores_sub.loc[scores_sub['sim_score'].idxmax()]
    sim = sim.append(score_max)
    print(count, ' of ', len(token_list))
```

1 of 1495  
2 of 1495  
3 of 1495  
4 of 1495  
5 of 1495  
6 of 1495  
7 of 1495  
8 of 1495  
9 of 1495  
10 of 1495  
11 of 1495  
12 of 1495  
13 of 1495  
14 of 1495  
15 of 1495  
16 of 1495  
17 of 1495  
18 of 1495  
19 of 1495  
20 of 1495  
21 of 1495  
22 of 1495  
23 of 1495  
24 of 1495  
25 of 1495  
26 of 1495  
27 of 1495  
28 of 1495  
29 of 1495  
30 of 1495  
31 of 1495  
32 of 1495  
33 of 1495  
34 of 1495  
35 of 1495  
36 of 1495  
37 of 1495  
38 of 1495  
39 of 1495  
40 of 1495  
41 of 1495  
42 of 1495  
43 of 1495  
44 of 1495  
45 of 1495  
46 of 1495  
47 of 1495  
48 of 1495  
49 of 1495  
50 of 1495  
51 of 1495  
52 of 1495  
53 of 1495  
54 of 1495  
55 of 1495  
56 of 1495  
57 of 1495  
58 of 1495  
59 of 1495  
60 of 1495  
61 of 1495  
62 of 1495

63 of 1495  
64 of 1495  
65 of 1495  
66 of 1495  
67 of 1495  
68 of 1495  
69 of 1495  
70 of 1495  
71 of 1495  
72 of 1495  
73 of 1495  
74 of 1495  
75 of 1495  
76 of 1495  
77 of 1495  
78 of 1495  
79 of 1495  
80 of 1495  
81 of 1495  
82 of 1495  
83 of 1495  
84 of 1495  
85 of 1495  
86 of 1495  
87 of 1495  
88 of 1495  
89 of 1495  
90 of 1495  
91 of 1495  
92 of 1495  
93 of 1495  
94 of 1495  
95 of 1495  
96 of 1495  
97 of 1495  
98 of 1495  
99 of 1495  
100 of 1495  
101 of 1495  
102 of 1495  
103 of 1495  
104 of 1495  
105 of 1495  
106 of 1495  
107 of 1495  
108 of 1495  
109 of 1495  
110 of 1495  
111 of 1495  
112 of 1495  
113 of 1495  
114 of 1495  
115 of 1495  
116 of 1495  
117 of 1495  
118 of 1495  
119 of 1495  
120 of 1495  
121 of 1495  
122 of 1495  
123 of 1495  
124 of 1495



125 of 1495  
126 of 1495  
127 of 1495  
128 of 1495  
129 of 1495  
130 of 1495  
131 of 1495  
132 of 1495  
133 of 1495  
134 of 1495  
135 of 1495  
136 of 1495  
137 of 1495  
138 of 1495  
139 of 1495  
140 of 1495  
141 of 1495  
142 of 1495  
143 of 1495  
144 of 1495  
145 of 1495  
146 of 1495  
147 of 1495  
148 of 1495  
149 of 1495  
150 of 1495  
151 of 1495  
152 of 1495  
153 of 1495  
154 of 1495  
155 of 1495  
156 of 1495  
157 of 1495  
158 of 1495  
159 of 1495  
160 of 1495  
161 of 1495  
162 of 1495  
163 of 1495  
164 of 1495  
165 of 1495  
166 of 1495  
167 of 1495  
168 of 1495  
169 of 1495  
170 of 1495  
171 of 1495  
172 of 1495  
173 of 1495  
174 of 1495  
175 of 1495  
176 of 1495  
177 of 1495  
178 of 1495  
179 of 1495  
180 of 1495  
181 of 1495  
182 of 1495  
183 of 1495  
184 of 1495  
185 of 1495  
186 of 1495

187 of 1495  
188 of 1495  
189 of 1495  
190 of 1495  
191 of 1495  
192 of 1495  
193 of 1495  
194 of 1495  
195 of 1495  
196 of 1495  
197 of 1495  
198 of 1495  
199 of 1495  
200 of 1495  
201 of 1495  
202 of 1495  
203 of 1495  
204 of 1495  
205 of 1495  
206 of 1495  
207 of 1495  
208 of 1495  
209 of 1495  
210 of 1495  
211 of 1495  
212 of 1495  
213 of 1495  
214 of 1495  
215 of 1495  
216 of 1495  
217 of 1495  
218 of 1495  
219 of 1495  
220 of 1495  
221 of 1495  
222 of 1495  
223 of 1495  
224 of 1495  
225 of 1495  
226 of 1495  
227 of 1495  
228 of 1495  
229 of 1495  
230 of 1495  
231 of 1495  
232 of 1495  
233 of 1495  
234 of 1495  
235 of 1495  
236 of 1495  
237 of 1495  
238 of 1495  
239 of 1495  
240 of 1495  
241 of 1495  
242 of 1495  
243 of 1495  
244 of 1495  
245 of 1495  
246 of 1495  
247 of 1495  
248 of 1495

249 of 1495  
250 of 1495  
251 of 1495  
252 of 1495  
253 of 1495  
254 of 1495  
255 of 1495  
256 of 1495  
257 of 1495  
258 of 1495  
259 of 1495  
260 of 1495  
261 of 1495  
262 of 1495  
263 of 1495  
264 of 1495  
265 of 1495  
266 of 1495  
267 of 1495  
268 of 1495  
269 of 1495  
270 of 1495  
271 of 1495  
272 of 1495  
273 of 1495  
274 of 1495  
275 of 1495  
276 of 1495  
277 of 1495  
278 of 1495  
279 of 1495  
280 of 1495  
281 of 1495  
282 of 1495  
283 of 1495  
284 of 1495  
285 of 1495  
286 of 1495  
287 of 1495  
288 of 1495  
289 of 1495  
290 of 1495  
291 of 1495  
292 of 1495  
293 of 1495  
294 of 1495  
295 of 1495  
296 of 1495  
297 of 1495  
298 of 1495  
299 of 1495  
300 of 1495  
301 of 1495  
302 of 1495  
303 of 1495  
304 of 1495  
305 of 1495  
306 of 1495  
307 of 1495  
308 of 1495  
309 of 1495  
310 of 1495

311 of 1495  
312 of 1495  
313 of 1495  
314 of 1495  
315 of 1495  
316 of 1495  
317 of 1495  
318 of 1495  
319 of 1495  
320 of 1495  
321 of 1495  
322 of 1495  
323 of 1495  
324 of 1495  
325 of 1495  
326 of 1495  
327 of 1495  
328 of 1495  
329 of 1495  
330 of 1495  
331 of 1495  
332 of 1495  
333 of 1495  
334 of 1495  
335 of 1495  
336 of 1495  
337 of 1495  
338 of 1495  
339 of 1495  
340 of 1495  
341 of 1495  
342 of 1495  
343 of 1495  
344 of 1495  
345 of 1495  
346 of 1495  
347 of 1495  
348 of 1495  
349 of 1495  
350 of 1495  
351 of 1495  
352 of 1495  
353 of 1495  
354 of 1495  
355 of 1495  
356 of 1495  
357 of 1495  
358 of 1495  
359 of 1495  
360 of 1495  
361 of 1495  
362 of 1495  
363 of 1495  
364 of 1495  
365 of 1495  
366 of 1495  
367 of 1495  
368 of 1495  
369 of 1495  
370 of 1495  
371 of 1495  
372 of 1495

373 of 1495  
374 of 1495  
375 of 1495  
376 of 1495  
377 of 1495  
378 of 1495  
379 of 1495  
380 of 1495  
381 of 1495  
382 of 1495  
383 of 1495  
384 of 1495  
385 of 1495  
386 of 1495  
387 of 1495  
388 of 1495  
389 of 1495  
390 of 1495  
391 of 1495  
392 of 1495  
393 of 1495  
394 of 1495  
395 of 1495  
396 of 1495  
397 of 1495  
398 of 1495  
399 of 1495  
400 of 1495  
401 of 1495  
402 of 1495  
403 of 1495  
404 of 1495  
405 of 1495  
406 of 1495  
407 of 1495  
408 of 1495  
409 of 1495  
410 of 1495  
411 of 1495  
412 of 1495  
413 of 1495  
414 of 1495  
415 of 1495  
416 of 1495  
417 of 1495  
418 of 1495  
419 of 1495  
420 of 1495  
421 of 1495  
422 of 1495  
423 of 1495  
424 of 1495  
425 of 1495  
426 of 1495  
427 of 1495  
428 of 1495  
429 of 1495  
430 of 1495  
431 of 1495  
432 of 1495  
433 of 1495  
434 of 1495

435 of 1495  
436 of 1495  
437 of 1495  
438 of 1495  
439 of 1495  
440 of 1495  
441 of 1495  
442 of 1495  
443 of 1495  
444 of 1495  
445 of 1495  
446 of 1495  
447 of 1495  
448 of 1495  
449 of 1495  
450 of 1495  
451 of 1495  
452 of 1495  
453 of 1495  
454 of 1495  
455 of 1495  
456 of 1495  
457 of 1495  
458 of 1495  
459 of 1495  
460 of 1495  
461 of 1495  
462 of 1495  
463 of 1495  
464 of 1495  
465 of 1495  
466 of 1495  
467 of 1495  
468 of 1495  
469 of 1495  
470 of 1495  
471 of 1495  
472 of 1495  
473 of 1495  
474 of 1495  
475 of 1495  
476 of 1495  
477 of 1495  
478 of 1495  
479 of 1495  
480 of 1495  
481 of 1495  
482 of 1495  
483 of 1495  
484 of 1495  
485 of 1495  
486 of 1495  
487 of 1495  
488 of 1495  
489 of 1495  
490 of 1495  
491 of 1495  
492 of 1495  
493 of 1495  
494 of 1495  
495 of 1495  
496 of 1495

497 of 1495  
498 of 1495  
499 of 1495  
500 of 1495  
501 of 1495  
502 of 1495  
503 of 1495  
504 of 1495  
505 of 1495  
506 of 1495  
507 of 1495  
508 of 1495  
509 of 1495  
510 of 1495  
511 of 1495  
512 of 1495  
513 of 1495  
514 of 1495  
515 of 1495  
516 of 1495  
517 of 1495  
518 of 1495  
519 of 1495  
520 of 1495  
521 of 1495  
522 of 1495  
523 of 1495  
524 of 1495  
525 of 1495  
526 of 1495  
527 of 1495  
528 of 1495  
529 of 1495  
530 of 1495  
531 of 1495  
532 of 1495  
533 of 1495  
534 of 1495  
535 of 1495  
536 of 1495  
537 of 1495  
538 of 1495  
539 of 1495  
540 of 1495  
541 of 1495  
542 of 1495  
543 of 1495  
544 of 1495  
545 of 1495  
546 of 1495  
547 of 1495  
548 of 1495  
549 of 1495  
550 of 1495  
551 of 1495  
552 of 1495  
553 of 1495  
554 of 1495  
555 of 1495  
556 of 1495  
557 of 1495  
558 of 1495

559 of 1495  
560 of 1495  
561 of 1495  
562 of 1495  
563 of 1495  
564 of 1495  
565 of 1495  
566 of 1495  
567 of 1495  
568 of 1495  
569 of 1495  
570 of 1495  
571 of 1495  
572 of 1495  
573 of 1495  
574 of 1495  
575 of 1495  
576 of 1495  
577 of 1495  
578 of 1495  
579 of 1495  
580 of 1495  
581 of 1495  
582 of 1495  
583 of 1495  
584 of 1495  
585 of 1495  
586 of 1495  
587 of 1495  
588 of 1495  
589 of 1495  
590 of 1495  
591 of 1495  
592 of 1495  
593 of 1495  
594 of 1495  
595 of 1495  
596 of 1495  
597 of 1495  
598 of 1495  
599 of 1495  
600 of 1495  
601 of 1495  
602 of 1495  
603 of 1495  
604 of 1495  
605 of 1495  
606 of 1495  
607 of 1495  
608 of 1495  
609 of 1495  
610 of 1495  
611 of 1495  
612 of 1495  
613 of 1495  
614 of 1495  
615 of 1495  
616 of 1495  
617 of 1495  
618 of 1495  
619 of 1495  
620 of 1495



621 of 1495  
622 of 1495  
623 of 1495  
624 of 1495  
625 of 1495  
626 of 1495  
627 of 1495  
628 of 1495  
629 of 1495  
630 of 1495  
631 of 1495  
632 of 1495  
633 of 1495  
634 of 1495  
635 of 1495  
636 of 1495  
637 of 1495  
638 of 1495  
639 of 1495  
640 of 1495  
641 of 1495  
642 of 1495  
643 of 1495  
644 of 1495  
645 of 1495  
646 of 1495  
647 of 1495  
648 of 1495  
649 of 1495  
650 of 1495  
651 of 1495  
652 of 1495  
653 of 1495  
654 of 1495  
655 of 1495  
656 of 1495  
657 of 1495  
658 of 1495  
659 of 1495  
660 of 1495  
661 of 1495  
662 of 1495  
663 of 1495  
664 of 1495  
665 of 1495  
666 of 1495  
667 of 1495  
668 of 1495  
669 of 1495  
670 of 1495  
671 of 1495  
672 of 1495  
673 of 1495  
674 of 1495  
675 of 1495  
676 of 1495  
677 of 1495  
678 of 1495  
679 of 1495  
680 of 1495  
681 of 1495  
682 of 1495

683 of 1495  
684 of 1495  
685 of 1495  
686 of 1495  
687 of 1495  
688 of 1495  
689 of 1495  
690 of 1495  
691 of 1495  
692 of 1495  
693 of 1495  
694 of 1495  
695 of 1495  
696 of 1495  
697 of 1495  
698 of 1495  
699 of 1495  
700 of 1495  
701 of 1495  
702 of 1495  
703 of 1495  
704 of 1495  
705 of 1495  
706 of 1495  
707 of 1495  
708 of 1495  
709 of 1495  
710 of 1495  
711 of 1495  
712 of 1495  
713 of 1495  
714 of 1495  
715 of 1495  
716 of 1495  
717 of 1495  
718 of 1495  
719 of 1495  
720 of 1495  
721 of 1495  
722 of 1495  
723 of 1495  
724 of 1495  
725 of 1495  
726 of 1495  
727 of 1495  
728 of 1495  
729 of 1495  
730 of 1495  
731 of 1495  
732 of 1495  
733 of 1495  
734 of 1495  
735 of 1495  
736 of 1495  
737 of 1495  
738 of 1495  
739 of 1495  
740 of 1495  
741 of 1495  
742 of 1495  
743 of 1495  
744 of 1495

745 of 1495  
746 of 1495  
747 of 1495  
748 of 1495  
749 of 1495  
750 of 1495  
751 of 1495  
752 of 1495  
753 of 1495  
754 of 1495  
755 of 1495  
756 of 1495  
757 of 1495  
758 of 1495  
759 of 1495  
760 of 1495  
761 of 1495  
762 of 1495  
763 of 1495  
764 of 1495  
765 of 1495  
766 of 1495  
767 of 1495  
768 of 1495  
769 of 1495  
770 of 1495  
771 of 1495  
772 of 1495  
773 of 1495  
774 of 1495  
775 of 1495  
776 of 1495  
777 of 1495  
778 of 1495  
779 of 1495  
780 of 1495  
781 of 1495  
782 of 1495  
783 of 1495  
784 of 1495  
785 of 1495  
786 of 1495  
787 of 1495  
788 of 1495  
789 of 1495  
790 of 1495  
791 of 1495  
792 of 1495  
793 of 1495  
794 of 1495  
795 of 1495  
796 of 1495  
797 of 1495  
798 of 1495  
799 of 1495  
800 of 1495  
801 of 1495  
802 of 1495  
803 of 1495  
804 of 1495  
805 of 1495  
806 of 1495

807 of 1495  
808 of 1495  
809 of 1495  
810 of 1495  
811 of 1495  
812 of 1495  
813 of 1495  
814 of 1495  
815 of 1495  
816 of 1495  
817 of 1495  
818 of 1495  
819 of 1495  
820 of 1495  
821 of 1495  
822 of 1495  
823 of 1495  
824 of 1495  
825 of 1495  
826 of 1495  
827 of 1495  
828 of 1495  
829 of 1495  
830 of 1495  
831 of 1495  
832 of 1495  
833 of 1495  
834 of 1495  
835 of 1495  
836 of 1495  
837 of 1495  
838 of 1495  
839 of 1495  
840 of 1495  
841 of 1495  
842 of 1495  
843 of 1495  
844 of 1495  
845 of 1495  
846 of 1495  
847 of 1495  
848 of 1495  
849 of 1495  
850 of 1495  
851 of 1495  
852 of 1495  
853 of 1495  
854 of 1495  
855 of 1495  
856 of 1495  
857 of 1495  
858 of 1495  
859 of 1495  
860 of 1495  
861 of 1495  
862 of 1495  
863 of 1495  
864 of 1495  
865 of 1495  
866 of 1495  
867 of 1495  
868 of 1495

869 of 1495  
870 of 1495  
871 of 1495  
872 of 1495  
873 of 1495  
874 of 1495  
875 of 1495  
876 of 1495  
877 of 1495  
878 of 1495  
879 of 1495  
880 of 1495  
881 of 1495  
882 of 1495  
883 of 1495  
884 of 1495  
885 of 1495  
886 of 1495  
887 of 1495  
888 of 1495  
889 of 1495  
890 of 1495  
891 of 1495  
892 of 1495  
893 of 1495  
894 of 1495  
895 of 1495  
896 of 1495  
897 of 1495  
898 of 1495  
899 of 1495  
900 of 1495  
901 of 1495  
902 of 1495  
903 of 1495  
904 of 1495  
905 of 1495  
906 of 1495  
907 of 1495  
908 of 1495  
909 of 1495  
910 of 1495  
911 of 1495  
912 of 1495  
913 of 1495  
914 of 1495  
915 of 1495  
916 of 1495  
917 of 1495  
918 of 1495  
919 of 1495  
920 of 1495  
921 of 1495  
922 of 1495  
923 of 1495  
924 of 1495  
925 of 1495  
926 of 1495  
927 of 1495  
928 of 1495  
929 of 1495  
930 of 1495

931 of 1495  
932 of 1495  
933 of 1495  
934 of 1495  
935 of 1495  
936 of 1495  
937 of 1495  
938 of 1495  
939 of 1495  
940 of 1495  
941 of 1495  
942 of 1495  
943 of 1495  
944 of 1495  
945 of 1495  
946 of 1495  
947 of 1495  
948 of 1495  
949 of 1495  
950 of 1495  
951 of 1495  
952 of 1495  
953 of 1495  
954 of 1495  
955 of 1495  
956 of 1495  
957 of 1495  
958 of 1495  
959 of 1495  
960 of 1495  
961 of 1495  
962 of 1495  
963 of 1495  
964 of 1495  
965 of 1495  
966 of 1495  
967 of 1495  
968 of 1495  
969 of 1495  
970 of 1495  
971 of 1495  
972 of 1495  
973 of 1495  
974 of 1495  
975 of 1495  
976 of 1495  
977 of 1495  
978 of 1495  
979 of 1495  
980 of 1495  
981 of 1495  
982 of 1495  
983 of 1495  
984 of 1495  
985 of 1495  
986 of 1495  
987 of 1495  
988 of 1495  
989 of 1495  
990 of 1495  
991 of 1495  
992 of 1495

993 of 1495  
994 of 1495  
995 of 1495  
996 of 1495  
997 of 1495  
998 of 1495  
999 of 1495  
1000 of 1495  
1001 of 1495  
1002 of 1495  
1003 of 1495  
1004 of 1495  
1005 of 1495  
1006 of 1495  
1007 of 1495  
1008 of 1495  
1009 of 1495  
1010 of 1495  
1011 of 1495  
1012 of 1495  
1013 of 1495  
1014 of 1495  
1015 of 1495  
1016 of 1495  
1017 of 1495  
1018 of 1495  
1019 of 1495  
1020 of 1495  
1021 of 1495  
1022 of 1495  
1023 of 1495  
1024 of 1495  
1025 of 1495  
1026 of 1495  
1027 of 1495  
1028 of 1495  
1029 of 1495  
1030 of 1495  
1031 of 1495  
1032 of 1495  
1033 of 1495  
1034 of 1495  
1035 of 1495  
1036 of 1495  
1037 of 1495  
1038 of 1495  
1039 of 1495  
1040 of 1495  
1041 of 1495  
1042 of 1495  
1043 of 1495  
1044 of 1495  
1045 of 1495  
1046 of 1495  
1047 of 1495  
1048 of 1495  
1049 of 1495  
1050 of 1495  
1051 of 1495  
1052 of 1495  
1053 of 1495  
1054 of 1495

1055 of 1495  
1056 of 1495  
1057 of 1495  
1058 of 1495  
1059 of 1495  
1060 of 1495  
1061 of 1495  
1062 of 1495  
1063 of 1495  
1064 of 1495  
1065 of 1495  
1066 of 1495  
1067 of 1495  
1068 of 1495  
1069 of 1495  
1070 of 1495  
1071 of 1495  
1072 of 1495  
1073 of 1495  
1074 of 1495  
1075 of 1495  
1076 of 1495  
1077 of 1495  
1078 of 1495  
1079 of 1495  
1080 of 1495  
1081 of 1495  
1082 of 1495  
1083 of 1495  
1084 of 1495  
1085 of 1495  
1086 of 1495  
1087 of 1495  
1088 of 1495  
1089 of 1495  
1090 of 1495  
1091 of 1495  
1092 of 1495  
1093 of 1495  
1094 of 1495  
1095 of 1495  
1096 of 1495  
1097 of 1495  
1098 of 1495  
1099 of 1495  
1100 of 1495  
1101 of 1495  
1102 of 1495  
1103 of 1495  
1104 of 1495  
1105 of 1495  
1106 of 1495  
1107 of 1495  
1108 of 1495  
1109 of 1495  
1110 of 1495  
1111 of 1495  
1112 of 1495  
1113 of 1495  
1114 of 1495  
1115 of 1495  
1116 of 1495



1117 of 1495  
1118 of 1495  
1119 of 1495  
1120 of 1495  
1121 of 1495  
1122 of 1495  
1123 of 1495  
1124 of 1495  
1125 of 1495  
1126 of 1495  
1127 of 1495  
1128 of 1495  
1129 of 1495  
1130 of 1495  
1131 of 1495  
1132 of 1495  
1133 of 1495  
1134 of 1495  
1135 of 1495  
1136 of 1495  
1137 of 1495  
1138 of 1495  
1139 of 1495  
1140 of 1495  
1141 of 1495  
1142 of 1495  
1143 of 1495  
1144 of 1495  
1145 of 1495  
1146 of 1495  
1147 of 1495  
1148 of 1495  
1149 of 1495  
1150 of 1495  
1151 of 1495  
1152 of 1495  
1153 of 1495  
1154 of 1495  
1155 of 1495  
1156 of 1495  
1157 of 1495  
1158 of 1495  
1159 of 1495  
1160 of 1495  
1161 of 1495  
1162 of 1495  
1163 of 1495  
1164 of 1495  
1165 of 1495  
1166 of 1495  
1167 of 1495  
1168 of 1495  
1169 of 1495  
1170 of 1495  
1171 of 1495  
1172 of 1495  
1173 of 1495  
1174 of 1495  
1175 of 1495  
1176 of 1495  
1177 of 1495  
1178 of 1495

1179 of 1495  
1180 of 1495  
1181 of 1495  
1182 of 1495  
1183 of 1495  
1184 of 1495  
1185 of 1495  
1186 of 1495  
1187 of 1495  
1188 of 1495  
1189 of 1495  
1190 of 1495  
1191 of 1495  
1192 of 1495  
1193 of 1495  
1194 of 1495  
1195 of 1495  
1196 of 1495  
1197 of 1495  
1198 of 1495  
1199 of 1495  
1200 of 1495  
1201 of 1495  
1202 of 1495  
1203 of 1495  
1204 of 1495  
1205 of 1495  
1206 of 1495  
1207 of 1495  
1208 of 1495  
1209 of 1495  
1210 of 1495  
1211 of 1495  
1212 of 1495  
1213 of 1495  
1214 of 1495  
1215 of 1495  
1216 of 1495  
1217 of 1495  
1218 of 1495  
1219 of 1495  
1220 of 1495  
1221 of 1495  
1222 of 1495  
1223 of 1495  
1224 of 1495  
1225 of 1495  
1226 of 1495  
1227 of 1495  
1228 of 1495  
1229 of 1495  
1230 of 1495  
1231 of 1495  
1232 of 1495  
1233 of 1495  
1234 of 1495  
1235 of 1495  
1236 of 1495  
1237 of 1495  
1238 of 1495  
1239 of 1495  
1240 of 1495

1241 of 1495  
1242 of 1495  
1243 of 1495  
1244 of 1495  
1245 of 1495  
1246 of 1495  
1247 of 1495  
1248 of 1495  
1249 of 1495  
1250 of 1495  
1251 of 1495  
1252 of 1495  
1253 of 1495  
1254 of 1495  
1255 of 1495  
1256 of 1495  
1257 of 1495  
1258 of 1495  
1259 of 1495  
1260 of 1495  
1261 of 1495  
1262 of 1495  
1263 of 1495  
1264 of 1495  
1265 of 1495  
1266 of 1495  
1267 of 1495  
1268 of 1495  
1269 of 1495  
1270 of 1495  
1271 of 1495  
1272 of 1495  
1273 of 1495  
1274 of 1495  
1275 of 1495  
1276 of 1495  
1277 of 1495  
1278 of 1495  
1279 of 1495  
1280 of 1495  
1281 of 1495  
1282 of 1495  
1283 of 1495  
1284 of 1495  
1285 of 1495  
1286 of 1495  
1287 of 1495  
1288 of 1495  
1289 of 1495  
1290 of 1495  
1291 of 1495  
1292 of 1495  
1293 of 1495  
1294 of 1495  
1295 of 1495  
1296 of 1495  
1297 of 1495  
1298 of 1495  
1299 of 1495  
1300 of 1495  
1301 of 1495  
1302 of 1495

1303 of 1495  
1304 of 1495  
1305 of 1495  
1306 of 1495  
1307 of 1495  
1308 of 1495  
1309 of 1495  
1310 of 1495  
1311 of 1495  
1312 of 1495  
1313 of 1495  
1314 of 1495  
1315 of 1495  
1316 of 1495  
1317 of 1495  
1318 of 1495  
1319 of 1495  
1320 of 1495  
1321 of 1495  
1322 of 1495  
1323 of 1495  
1324 of 1495  
1325 of 1495  
1326 of 1495  
1327 of 1495  
1328 of 1495  
1329 of 1495  
1330 of 1495  
1331 of 1495  
1332 of 1495  
1333 of 1495  
1334 of 1495  
1335 of 1495  
1336 of 1495  
1337 of 1495  
1338 of 1495  
1339 of 1495  
1340 of 1495  
1341 of 1495  
1342 of 1495  
1343 of 1495  
1344 of 1495  
1345 of 1495  
1346 of 1495  
1347 of 1495  
1348 of 1495  
1349 of 1495  
1350 of 1495  
1351 of 1495  
1352 of 1495  
1353 of 1495  
1354 of 1495  
1355 of 1495  
1356 of 1495  
1357 of 1495  
1358 of 1495  
1359 of 1495  
1360 of 1495  
1361 of 1495  
1362 of 1495  
1363 of 1495  
1364 of 1495

1365 of 1495  
1366 of 1495  
1367 of 1495  
1368 of 1495  
1369 of 1495  
1370 of 1495  
1371 of 1495  
1372 of 1495  
1373 of 1495  
1374 of 1495  
1375 of 1495  
1376 of 1495  
1377 of 1495  
1378 of 1495  
1379 of 1495  
1380 of 1495  
1381 of 1495  
1382 of 1495  
1383 of 1495  
1384 of 1495  
1385 of 1495  
1386 of 1495  
1387 of 1495  
1388 of 1495  
1389 of 1495  
1390 of 1495  
1391 of 1495  
1392 of 1495  
1393 of 1495  
1394 of 1495  
1395 of 1495  
1396 of 1495  
1397 of 1495  
1398 of 1495  
1399 of 1495  
1400 of 1495  
1401 of 1495  
1402 of 1495  
1403 of 1495  
1404 of 1495  
1405 of 1495  
1406 of 1495  
1407 of 1495  
1408 of 1495  
1409 of 1495  
1410 of 1495  
1411 of 1495  
1412 of 1495  
1413 of 1495  
1414 of 1495  
1415 of 1495  
1416 of 1495  
1417 of 1495  
1418 of 1495  
1419 of 1495  
1420 of 1495  
1421 of 1495  
1422 of 1495  
1423 of 1495  
1424 of 1495  
1425 of 1495  
1426 of 1495

1427 of 1495  
1428 of 1495  
1429 of 1495  
1430 of 1495  
1431 of 1495  
1432 of 1495  
1433 of 1495  
1434 of 1495  
1435 of 1495  
1436 of 1495  
1437 of 1495  
1438 of 1495  
1439 of 1495  
1440 of 1495  
1441 of 1495  
1442 of 1495  
1443 of 1495  
1444 of 1495  
1445 of 1495  
1446 of 1495  
1447 of 1495  
1448 of 1495  
1449 of 1495  
1450 of 1495  
1451 of 1495  
1452 of 1495  
1453 of 1495  
1454 of 1495  
1455 of 1495  
1456 of 1495  
1457 of 1495  
1458 of 1495  
1459 of 1495  
1460 of 1495  
1461 of 1495  
1462 of 1495  
1463 of 1495  
1464 of 1495  
1465 of 1495  
1466 of 1495  
1467 of 1495  
1468 of 1495  
1469 of 1495  
1470 of 1495  
1471 of 1495  
1472 of 1495  
1473 of 1495  
1474 of 1495  
1475 of 1495  
1476 of 1495  
1477 of 1495  
1478 of 1495  
1479 of 1495  
1480 of 1495  
1481 of 1495  
1482 of 1495  
1483 of 1495  
1484 of 1495  
1485 of 1495  
1486 of 1495  
1487 of 1495  
1488 of 1495

```
1489 of 1495
1490 of 1495
1491 of 1495
1492 of 1495
1493 of 1495
1494 of 1495
1495 of 1495
```

```
In [76]: sim.sort_values('sim_score',ascending=False).reset_index().drop('index',axis=1).loc[0]
```

```
Out[76]: token1      British
token2      American
sim_score    0.936294
Name: 0, dtype: object
```