

regexes

February 9, 2023

```
[9]: import pandas as pd
import numpy as np
import nltk
import re
#do not show warnings
import warnings
warnings.filterwarnings("ignore")
```

0.0.1 prepare data US and clean

```
[10]: df = pd.read_table('https://github.com/blueprints-for-text-analytics-python/
↳blueprints-text/raw/193c79c7d94973f2398e67da8d20bf7a535f7f40/data/
↳un-general-debates/un-general-debates-blueprint.csv.gz', compression='gzip',
↳sep=',', quotechar='"', error_bad_lines=False)

[11]: df_US = df[df['country_name'] == 'United States']

[12]: def remove_special_characters(text):
    return re.sub('[^A-Za-z0-9\,\.\-\_]+', ' ', text)
df_US['text'] = df_US['text'].apply(remove_special_characters).str.strip()
df_US['text'] = df_US['text'].str.lower()
```

1 Barack Obama

Barack Obama is an American politician and attorney who served as the 44th President of the United States from 2009 to 2017. He was the first African American to be elected to the presidency. Obama was born in Honolulu, Hawaii, on August 4, 1961. Prior to his presidency, he served as a U.S. Senator from Illinois from 2005 to 2008. During his presidency, Obama implemented several major policies, including the American Recovery and Reinvestment Act, the Dodd-Frank Wall Street Reform and Consumer Protection Act, and the Affordable Care Act (also known as Obamacare). He also oversaw the end of the U.S. military involvement in Iraq and the killing of Osama bin Laden, the leader of Al-Qaeda. After leaving office, Obama has continued to be an active and influential figure, speaking out on issues such as **climate change**, **racial justice**, and **democracy**.

```
[13]: Obama = df_US[df_US['speaker'] == "Barack Obama"]
Obama
```

```
[13]:      session  year country  country_name      speaker  position \
6336      64   2009     USA   United States Barack Obama President
6526      65   2010     USA   United States Barack Obama President
6719      66   2011     USA   United States Barack Obama President
6914      67   2012     USA   United States Barack Obama President
7107      68   2013     USA   United States Barack Obama President
7301      69   2014     USA   United States Barack Obama President
7495      70   2015     USA   United States Barack Obama President
```

```

                                text
6336  it is my honour to address the assembly for th...
6526  it is a great honour to address this assembly ...
6719  it is a great honour for me to be here today. ...
6914  i would like to begin today by telling the ass...
7107  each year we come together to reaffirm the fou...
7301  we come together at a crossroads between war a...
7495  seventy years after the founding of the united...
```

1.0.1 Climate change

1. Why you want to discover such a pattern?

ANS Based on my search about Barack Obama's background, you anticipated that he would speak about climate change during debates at the United Nations. I use this pattern because I want to see a part of sentence related with climate.

2. why is it interesting to you?

ANS Obama is a well-respected global leader and his opinions on climate change are likely to be of interest at the UN. His advocacy can help bring attention to the importance of addressing this pressing issue.

3. Does the regex work as you expected?

ANS Yes, it does.

```
[14]: #find every year that have the word "climate" in it and print it
for i in range(Obama.shape[0]):
    matches = re.finditer(r'([^\s\.\,][a-zA-Z\s-]+)climate(\.
↵|\,|\s[a-zA-Z\s0-9\,]+\.)', Obama['text'].iloc[i])
    if matches:
        for match in matches:
            print(Obama['year'].iloc[i],match.group())
```

2009 to confront climate change we have invested 80 billion in clean energy.
2009 launched an energy partnership across the americas and moved from being a
bystander to being a leader in international climate negotiations.
2009 and i thank the secretary-general for hosting the summit on climate change
yesterday.
2009 the danger posed by climate change cannot be denied.

2009 and any effort that fails both to help the poorest nations to adapt to the problems that climate change has already wrought and to help them travel a path of clean development simply will not work.

2010 we are also confronting the spectre of climate change.

2011 we must not put off action that climate change demands.

2013 combating climate change, starting businesses, expanding freedom, and leaving behind the old ideological battles of the past.

2014 but the science tells us that we can succeed in combating climate change only if we are joined in that effort by every other nation, by every major power.

2015 we can roll back the pollution that we put in our skies and help economies lift people out of poverty without condemning our children to the ravages of an ever-warming climate.

2015 no country can escape the ravages of climate change.

Notice: According to the results, throughout his presidency, Barack Obama has been an advocate for addressing the issue of climate change. He has spoken out about the importance of taking action on this issue in almost every year, with the exception of 2012. This shows a consistent and persistent commitment to raising awareness about the issue and the need for solutions.

1.0.2 Democracy

1. Why you want to discover such a pattern?

ANS I would like to know the plan of the US for promoting democracy, the direction in which America is pushing for its advancement, and the policies of President Barack Obama that are a strong point in his commitment to democracy.

2. why is it interesting to you?

ANS America is often seen as a symbol of democracy around the world, and its political system and government practices are studied and imitated by many countries.

3. Does the regex work as you expected?

ANS Yes, it does.

```
[7]: #find every year that have the word "democracy" in it and print it
for i in range(Obama.shape[0]):
    matches = re.finditer(r'([^\s\.,][a-zA-Z\s\-\-]+)democracy(\.
↪|\,|\s[a-zA-Z\s0-9]+\.)', Obama['text'].iloc[i])
    if matches:
        for match in matches:
            print(Obama['year'].iloc[i],match.group())
    else:
        print("No matches found")
```

2009 and i admit that america has too often been selective in its promotion of democracy.

2010 make no mistake the ultimate success of democracy in the world will not

come because the united states dictates it it will come because individual citizens demand a say in how they are governed.

2010 just as every democracy reflects the uniqueness of a nation.

2010 which peacefully threw off colonialism and established a thriving democracy of over a billion people.

2010 an ancient culture that found peace and extraordinary development through democracy.

2010 it is time for every member state to open its elections to international monitors and increase the united nations democracy fund.

2010 and so atrocities like sexual violence are prevented and justice is enforced because neither dignity nor democracy can thrive without basic security.

2011 now the people of tunisia are preparing for elections that will move them one step closer to the democracy that they deserve.

2011 some religions and some ethnicities do not desire democracy.

2011 those nations that transition to democracy,

2012 he supported the birth of a new democracy as libyans held elections and built new institutions and began to move forward after decades of dictatorship.

2012 because our support for democracy ultimately put us on the side of the people.

2012 and even as there will be huge challenges that come with a transition to democracy,

2012 and yet the turmoil of recent weeks reminds us that the path to democracy does not end with the casting of a ballot.

2012 true democracy demands that citizens cannot be thrown in jail because of what they believe and that businesses can be opened without paying a bribe.

2012 true democracy real freedom is hard work.

2012 which is moving towards democracy.

2012 to educate our children and create the opportunities that they deserve protect human rights and extend democracy s promise.

2013 and we will continue to promote democracy,

2013 iraq has shown us that democracy cannot simply be imposed by force.

2013 has made decisions inconsistent with inclusive democracy,

2013 for the hard work of forging freedom and democracy is the task of a generation.

2014 america and our allies will support the people of ukraine as they develop their democracy and economy.

2014 where what began as a violent transition has evolved into a genuine democracy.

2014 we address our differences in the open space of democracy,

2015 it is those international principles that have helped constrain bigger countries from imposing their will on smaller ones and have advanced the emergence of democracy and development and individual liberty on every continent.

2015 catastrophes like what we are seeing in syria do not take place in countries where there is genuine democracy and respect for the universal values this institution is supposed to defend.

2015 i recognize that democracy is going to take different forms in different

parts of the world.

2015 i understand that democracy is frustrating.

2015 inclusive democracy,

2015 that is part of democracy.

Notice: According to the results, TThroughout his presidency, Barack Obama emphasized the importance of democracy and supported its promotion globally. He believed that the success of democracy in the world will come from the demand of individual citizens and not from the dictation of the United States. Obama acknowledged that democracy takes different forms in different parts of the world, but emphasized the importance of open and fair elections, protection of human rights, and the rule of law.

1.0.3 China or Russia

1. Why you want to discover such a pattern?

ANS When it comes to the United States' relationships with countries in the world, the only two I am aware of are China and Russia. And I want to see to context in sentence.

2. why is it interesting to you?

ANS China and Russia are certainly two significant countries with which the United States has had a long and complex history.

3. Does the regex work as you expected?

ANS Yes, it does.

```
[8]: #find every year that have the word "china" in it and print it
print('-- China --')
for i in range(Obama.shape[0]):
    matches = re.finditer(r'([^\s\.] [a-zA-Z\s\-\,]+)china(\.
↵|\,|\s[a-zA-Z\s0-9\,]+\.)', Obama['text'].iloc[i])
    if matches:
        for match in matches:
            print(Obama['year'].iloc[i],match.group())
    else:
        print("No matches found")
#find every year that have the word "china" in it and print it
print('-- Russia --')
for i in range(Obama.shape[0]):
    matches = re.finditer(r'([^\s\.] [a-zA-Z\s\-\,]+)russia(\.
↵|\,|\s[a-zA-Z\s0-9\,]+\.)', Obama['text'].iloc[i])
    if matches:
        for match in matches:
            print(Obama['year'].iloc[i],match.group())
    else:
        print("No matches found")
```

-- China --

2013 given president rouhani s stated commitment to reach an agreement, i am

directing john kerry to pursue this effort with the iranian government, in close cooperation with the european union, the united kingdom, france, germany, russia and china.

2015 even as our economy is growing and our troops have largely returned from iraq and afghanistan, we see in our debates about america s role in the world a notion of strength that is defined by opposition to old enemies, perceived adversaries a rising china or resurgent russia, a revolutionary iran or an islam that is incompatible with peace.

2015 for two years, the united states and our partners, including russia and china,

2015 similarly, in the south china sea, the united states makes no claim on territory there.

2015 so we will defend those principles, while encouraging china and other claimants to resolve their differences peacefully.

-- Russia --

2009 in moscow, the united states and russia announced that we would pursue substantial reductions in our strategic warheads and launchers.

2009 we will pursue a new agreement with russia to substantially reduce our strategic warheads and launches.

2010 we have joined with russia to sign the most comprehensive arms control treaty in decades.

2011 the new start treaty between the united states and russia will cut our deployed arsenals to the lowest level in half a century, and our nations are pursuing talks on how to achieve even deeper reductions.

2012 nations have come together to lock down nuclear materials, and america and russia are reducing our arsenals.

2013 petersburg, my preference has always been to work for a diplomatic resolution to this issue, and in the past several weeks, the united states, russia and our allies have reached an agreement to place syria s chemical weapons under international control and then destroy them.

2013 given president rouhani s stated commitment to reach an agreement, i am directing john kerry to pursue this effort with the iranian government, in close cooperation with the european union, the united kingdom, france, germany, russia and china.

2014 against the will of the government in kyiv, crimea was annexed by russia.

2014 when ukraine started to reassert control over its territory, russia gave up the pretence of merely supporting the separatists and moved troops across the border.

2014 we will impose a cost on russia for its aggression, and we will counter falsehoods with the truth.

2014 if russia takes that path a path that for stretches of the post-cold war period resulted in prosperity for the russian people then we will lift our sanctions and welcome russia s role in addressing common challenges.

2014 that is the kind of cooperation we are prepared to pursue again, if russia changes course.

2015 even as our economy is growing and our troops have largely returned from iraq and afghanistan, we see in our debates about america s role in the world a notion of strength that is defined by opposition to old enemies, perceived

adversaries a rising china or resurgent russia,
 2015 for two years, the united states and our partners, including russia and china, stuck together in complex negotiations.
 2015 consider russia s annexation of crimea and further aggression in eastern ukraine.
 2015 america has few economic interests in ukraine, and we recognize the deep and complex history between russia and ukraine.
 2015 that is the basis of the sanctions that the united states and our partners impose on russia.
 2015 now, within russia, state-controlled media may describe those events as an example of a resurgent russia a view shared, by the way, by a number of united states politicians and commentators who have always been deeply sceptical about russia and seem to be convinced that a new cold war is in fact upon us.
 2015 the ukrainian people are more interested than ever in aligning with europe instead of russia.
 2015 imagine if, instead, russia had engaged in true diplomacy and worked with ukraine and the international community to ensure that its interests were protected.
 2015 that would be better for ukraine, but also better for russia and better for the world, which is why we continue to press for this crisis to be resolved in a way that allows a sovereign and democratic ukraine to determine its future and control its territory.
 2015 this is not because we want to isolate russia we do not but because we want a strong russia that is invested in working with us to strengthen the international system as a whole.
 2015 the united states is prepared to work with any nation, including russia and iran, to resolve the conflict.

Notice: According to the results, Barack Obama regarding his stance on China and Russia. From the quotes, it seems that President Obama recognized China as a rising power and Russia as a potential adversary. He sought to work with both nations on various diplomatic efforts, but also imposed sanctions on Russia due to its annexation of Crimea. He believed in the importance of a strong international system, and was willing to work with any nation, including Russia and Iran, to resolve conflicts. However, he also recognized the deep and complex history between Russia and Ukraine and was in favor of Ukraine determining its own future and controlling its own territory.