

```
In [1]: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
%matplotlib inline
from sklearn import tree
from sklearn.tree import DecisionTreeClassifier

from sklearn.datasets import load_iris
from sklearn import tree
from sklearn.tree import export_graphviz
import io
from scipy import misc

#reading real newsdataset (Buzzfeed)
buzzr= pd.read_csv("/Users/sonia/Downloads/BuzzFeed_Real_1-Webpage.csv", e
```

```
In [2]: np.random.seed(10)
```

```
In [3]: buzzr
```

Out[3]:

	top_img	text	meta_data_theme	meta_data
0	http://a.abcnews.com/images/Politics/AP_donald...	Less than a day after protests over the police...	NaN	
1	NaN	NaN	NaN	
2	NaN	NaN	NaN	
3	NaN	NaN	NaN	
4	http://rightwingnews.com/wp-content/uploads/20...	Obama To UN: 'Giving Up Liberty, Enhances	NaN	

```
In [4]: missing_values_count = buzzr.isnull().sum()
```

```
In [5]: missing_values_count[0:106] #gives total amount of missing value
```

```
Out[5]: top_img          907
text          907
```

meta_data__theme	996
meta_data__msapplication-TileColor	995
authors__001	914
meta_data__publisher	992
meta_data__googlebot	983
authors__002	946
authors__003	982
meta_data__generator	987
authors__004	987
authors__005	992
authors__006	993
authors__007	994
authors__008	995
authors__009	996
authors__010	996
meta_data__fb_title	988
meta_data__msApplication-ID	996
meta_data__description	920
meta_data__title	995
meta_data__modernizr-version	996
meta_data__foundation-version	996
meta_data__jquery-version	996
meta_data__pubdate	975
meta_data__copyright	988
meta_data__author	960
meta_data__vr_category	996
meta_data__vr_type	996
meta_data__og_site_name	908
...	
meta_data__article_publisher	990
meta_data__theme-color	996
meta_data__fb_pages	915
meta_data__thumbnail	975
meta_data__template-top	975
meta_data__vr_canonical	975
meta_data__build	965
meta_data__article_section	982
meta_data__article_tag	992
meta_data__article_published_time	982
meta_data__article_modified_time	986
meta_data__google-site-verification	983
meta_data__verify-v1	996
meta_data__Last-Modified	989
meta_data__keywords	958
meta_data__ov_id	996
meta_data__lastmod	975
meta_data__article_author	968
meta_data__viewport	907
meta_data__format-detection	996
meta_data__news_keywords	986

```

canonical_link      908
images              0
title              907
url                907
summary            997
movies__-           993
publish_date__$date 942
publish_date        997
source             907
Length: 106, dtype: int64

```

In [7]: buzzr

1		NaN	NaN	NaN
2		NaN	NaN	NaN
3		NaN	NaN	NaN
4	http://rightwingnews.com/wp-content/uploads/20...		Obama To UN: 'Giving Up Liberty, Enhances Secu...	NaN
5		NaN	NaN	NaN
6		NaN	NaN	NaN
7		NaN	NaN	NaN
8		NaN	NaN	NaN
9		NaN	NaN	NaN

In [8]: buzzr=buzzr.drop(columns='images')*#dropping unnecessary columns*

In [9]: buzzr

24		NaN	NaN	NaN
25		NaN	NaN	NaN
26	http://static.politico.com/e9/11/6144cdc24e319...		Getty Images Wealth Of Nations Trump vs. Clint...	NaN
27		NaN	NaN	NaN
28		NaN	NaN	NaN
29	http://a.abcnews.com/images/US/AP_Obama_BM_201...		President Obama today vetoed a bill that would...	NaN
...

In [10]: buzzr=buzzr.dropna(axis=0, how='all', thresh=None, subset=None, inplace=F

In [11]: buzzr

Out[11]:

		top_img	text	meta_data_theme
0	http://a.abcnews.com/images/Politics/AP_donald...		Less than a day after protests over the police...	NaN
1		NaN	NaN	NaN
2		NaN	NaN	NaN
3		NaN	NaN	NaN
4	http://rightwingnews.com/wp-content/uploads/20...		Obama To UN: 'Giving Up Liberty, Enhances Secu...	NaN
			Gettv Images	

```
In [12]: 1, how='all', thresh=None, subset=None, inplace=False) #dropping rows that
```

```
In [13]: buzzr
```

1		NaN	NaN	NaN
2		NaN	NaN	NaN
3		NaN	NaN	NaN
4	http://rightwingnews.com/wp-content/uploads/20...		Obama To UN: 'Giving Up Liberty, Enhances Secu...	NaN
26	http://static.politico.com/e9/11/6144cdc24e319...		Getty Images Wealth Of Nations Trump vs. Clint...	NaN
29	http://a.abcnews.com/images/US/AP_Obama_BM_201...		President Obama today vetoed a bill that would...	NaN

```
In [14]: buzzr.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 116 entries, 0 to 994
Columns: 103 entries, top_img to source
dtypes: float64(8), object(95)
memory usage: 94.2+ KB
```

```
In [17]: , 'meta_data__viewport', 'meta_data__format-detection', 'movies__-' ] ) #droppi
```

In [18]: buzzr

Out[18]:

	top_img	text	meta_data__msapplic Tile
0	http://a.abcnews.com/images/Politics/AP_donald...	Less than a day after protests over the police...	
1	NaN	NaN	
2	NaN	NaN	
3	NaN	NaN	
4	http://rightwingnews.com/wp-content/uploads/20...	Obama To UN: 'Giving Up Liberty, Enhances Secu...	
Getty Images			

In [19]: buzzr=buzzr.drop(columns=['meta_data__msapplication-TileColor'])

In [20]: buzzr

Out[20]:

	top_img	text	authors__001	meta_
0	http://a.abcnews.com/images/Politics/AP_donald...	Less than a day after protests over the police...	More Candace	
1	NaN	NaN	NaN	
2	NaN	NaN	NaN	
3	NaN	NaN	NaN	
4	http://rightwingnews.com/wp-content/uploads/20...	Obama To UN: 'Giving Up Liberty, Enhances Secu...	Cassy Fiano	
26	http://static.politico.com/e9/11/6144cdc24e319...	Getty Images Wealth Of Nations	Jack Shafer	

In [21]: buzzr.loc[:, 'title']

Out[21]: 0 Donald Trump: Drugs a 'Very, Very Big Factor' ...

1 NaN
2 NaN
3 NaN
4 Obama To UN: 'Giving Up Liberty, Enhances Secu...
26 Trump vs. Clinton: A Fundamental Clash over Ho...
29 President Obama Vetoes 9/11 Victims Bill, Sett...
30 NaN
31 NaN
32 NaN
33 CHAOS! NC Protest MOB Ambushes Female Truck Dr...
56 10 Moments That Mattered From Hillary Clinton ...
57 NaN
58 NaN
59 NaN
62 Young Girl's Emotional Council Speech Laments ...
63 NaN
64 NaN
65 NaN
69 Jeb Bush to lecture at Harvard this fall
71 Clinton vs. Trump: The debate before the debate
74 Is Hillary in a No-Win Situation?
77 Hillary Clinton on police shootings: 'too many...
98 Johnson: 'I'm going to sleep well at the end o...
101 Anthony Weiner Under Federal Investigation for...
102 NaN
103 NaN
104 NaN
105 Senate Dems Move To Force Trump To Release His...
121 Dem congressman to IRS: Audit the Trump Founda...
...
799 Everything you need to know about Trump and Cl...
802 U.S. attorney sees no rush to arraign Rahami i...
805 Donald Trump's rise puts Ted Cruz in a bind
810 Trump hits Clinton on Islamic State: 'It is ti...
813 Bridgegate: Port Authority officials planned t...
815 Bono: 'Trump is potentially the worst idea tha...
817 2016 presidential race tightens in Florida
819 HuffPo Just Dropped The Mic On Internet Polls ...
828 Tim Kaine on Libertarian ticket: 'Everybody kn...
830 Watch Joe Biden Take Trump's Own Words And Tea...
839 U.S. mistakenly granted citizenship to hundred...
842 Insiders: The heat is on Hillary
845 Donald Trump Jr. Just Pinned A Tweet So Stupid...
855 More Milestone Moments for Donald Trump! - Eag...
883 Donald Trump retreats to friendly media ground
897 Pence gives no specifics on why Trump Foundati...
899 Details Emerge About NYC Bomb Suspect Ahmad Kh...
900 NaN
901 NaN
902 NaN

```

907             Trump's puzzling pitch to black voters
910     Joe Arpaio: Trump didn't call Obama's birth ce...
913     Barack Obama at Benjamin Netanyahu meeting: 'H...
932     McAuliffe hangs with Willie Nelson -- and his ...
935     The Mystery of Trump's Man in Moscow: Carter Page
938     It's "Trump is HITLER" Month at the Washington...
965             Obama's team isn't laughing at Trump anymore
967     Georgia poll: Donald Trump, Hillary Clinton in...
985     Chelsea Handler Gets The Last Word After RNC C...
994             Is Donald Trump qualified to be president?
Name: title, Length: 116, dtype: object

```

```
In [22]: buzzr.info()
```

```

<class 'pandas.core.frame.DataFrame'>
Int64Index: 116 entries, 0 to 994
Data columns (total 98 columns):
top_img                90 non-null object
text                   90 non-null object
authors__001           83 non-null object
meta_data__publisher   5 non-null object
meta_data__googlebot   14 non-null object
authors__002           51 non-null object
authors__003           15 non-null object
meta_data__generator   10 non-null object
authors__004           10 non-null object
authors__005           5 non-null object
authors__006           4 non-null object
authors__007           3 non-null object
authors__008           2 non-null object
authors__009           1 non-null object
authors__010           1 non-null object
meta_data__fb_title     9 non-null object
meta_data__msApplication-ID 1 non-null object
meta_data__description  77 non-null object
meta_data__title        2 non-null object
meta_data__modernizr-version 1 non-null object
meta_data__foundation-version 1 non-null object
meta_data__jquery-version 1 non-null object
meta_data__pubdate      22 non-null object
meta_data__copyright     9 non-null object
meta_data__author       37 non-null object
meta_data__vr__category  1 non-null object
meta_data__vr__type      1 non-null object
meta_data__og__site_name 89 non-null object
meta_data__og__description 88 non-null object
meta_data__og__pubdate   22 non-null object
meta_data__og__title     89 non-null object
meta_data__og__locale    21 non-null object
meta_data__og__image__width 35 non-null float64

```


meta_data__og__image__identifier	35 non-null object
meta_data__og__image__type	1 non-null object
meta_data__og__image__height	35 non-null float64
meta_data__og__updated_time	11 non-null object
meta_data__og__url	89 non-null object
meta_data__og__image	54 non-null object
meta_data__og__fb_appid	32 non-null float64
meta_data__og__type	89 non-null object
meta_data__section	23 non-null object
meta_data__referrer	8 non-null object
meta_data__twitter__description	39 non-null object
meta_data__twitter__title	40 non-null object
meta_data__twitter__url	16 non-null object
meta_data__twitter__image	42 non-null object
meta_data__twitter__creator	45 non-null object
meta_data__object-hash	1 non-null float64
meta_data__host	32 non-null object
meta_data__twitter__site	50 non-null object
meta_data__twitter__app__	27 non-null object
meta_data__twitter__app__ __ipad	27 non-null object
meta_data__twitter__app__ __iphone	27 non-null object
meta_data__twitter__app__ __googleplay	24 non-null object
meta_data__twitter__card	74 non-null object
meta_data__robots	24 non-null object
meta_data__apple-itunes-app	1 non-null object
meta_data__DC.date.issued	11 non-null object
meta_data__theme-version	1 non-null object
meta_data__publish_date	1 non-null object
meta_data__al__	34 non-null object
meta_data__al__ __url	34 non-null object
meta_data__al__ __app_store_id	18 non-null float64
meta_data__al__ __app_name	26 non-null object
meta_data__msvalidate.01	1 non-null object
meta_data__al__ __package	8 non-null object
meta_data__Date	8 non-null object
meta_data__fb__admins	11 non-null object
meta_data__fb__app_id	56 non-null float64
meta_data__date	1 non-null object
meta_data__msapplication-TileImage	5 non-null object
meta_data__application-name	1 non-null object
meta_data__msApplication-PackageFamilyName	1 non-null object
meta_data__article__publisher	7 non-null object
meta_data__theme-color	1 non-null object
meta_data__fb__pages	82 non-null object
meta_data__thumbnail	22 non-null object
meta_data__template-top	22 non-null object
meta_data__vr__canonical	22 non-null object
meta_data__build	32 non-null object
meta_data__article__section	15 non-null object
meta_data__article__tag	5 non-null object

```

meta_data__article__published_time    15 non-null object
meta_data__article__modified_time     11 non-null object
meta_data__google-site-verification    14 non-null object
meta_data__verify-v1                   1 non-null object
meta_data__Last-Modified                8 non-null object
meta_data__keywords                    39 non-null object
meta_data__ov__id                      1 non-null float64
meta_data__lastmod                     22 non-null object
meta_data__article__author             29 non-null object
meta_data__news_keywords                11 non-null object
canonical_link                         89 non-null object
title                                  90 non-null object
url                                    90 non-null object
publish_date__$date                    55 non-null float64
source                                 90 non-null object
dtypes: float64(8), object(90)
memory usage: 89.7+ KB

```

```
In [24]: buzzr=buzzr.drop(columns=['meta_data__googlebot','meta_data__msApplicatio
```

```
In [25]: buzzr
```

```
Out[25]:
```

	top_img	text	authors__001	meta_
0	http://a.abcnews.com/images/Politics/AP_donald...	Less than a day after protests over the police...	More Candace	
1	NaN	NaN	NaN	
2	NaN	NaN	NaN	
3	NaN	NaN	NaN	
4	http://rightwingnews.com/wp-content/uploads/20...	Obama To UN: 'Giving Up Liberty, Enhances Secu...	Cassy Fiano	
26	http://static.politico.com/e9/11/6144cdc24e319...	Getty Images Wealth Of Nations	Jack Shafer	

```
In [26]: buzzr.info()
```

```

<class 'pandas.core.frame.DataFrame'>
Int64Index: 116 entries, 0 to 994
Data columns (total 92 columns):
top_img                                90 non-null object

```

text	90 non-null object
authors__001	83 non-null object
meta_data__publisher	5 non-null object
authors__002	51 non-null object
authors__003	15 non-null object
meta_data__generator	10 non-null object
authors__004	10 non-null object
authors__005	5 non-null object
authors__006	4 non-null object
authors__007	3 non-null object
authors__008	2 non-null object
authors__009	1 non-null object
authors__010	1 non-null object
meta_data__fb_title	9 non-null object
meta_data__description	77 non-null object
meta_data__title	2 non-null object
meta_data__modernizr-version	1 non-null object
meta_data__foundation-version	1 non-null object
meta_data__jquery-version	1 non-null object
meta_data__pubdate	22 non-null object
meta_data__copyright	9 non-null object
meta_data__author	37 non-null object
meta_data__vr__category	1 non-null object
meta_data__vr__type	1 non-null object
meta_data__og__site_name	89 non-null object
meta_data__og__description	88 non-null object
meta_data__og__pubdate	22 non-null object
meta_data__og__title	89 non-null object
meta_data__og__locale	21 non-null object
meta_data__og__updated_time	11 non-null object
meta_data__og__url	89 non-null object
meta_data__og__image	54 non-null object
meta_data__og__fb_appid	32 non-null float64
meta_data__og__type	89 non-null object
meta_data__section	23 non-null object
meta_data__referrer	8 non-null object
meta_data__twitter__description	39 non-null object
meta_data__twitter__title	40 non-null object
meta_data__twitter__url	16 non-null object
meta_data__twitter__image	42 non-null object
meta_data__twitter__creator	45 non-null object
meta_data__object-hash	1 non-null float64
meta_data__host	32 non-null object
meta_data__twitter__site	50 non-null object
meta_data__twitter__app__	27 non-null object
meta_data__twitter__app__ __ipad	27 non-null object
meta_data__twitter__app__ __iphone	27 non-null object
meta_data__twitter__app__ __googleplay	24 non-null object
meta_data__twitter__card	74 non-null object
meta_data__robots	24 non-null object

```

meta_data__apple-itunes-app      1 non-null object
meta_data__DC.date.issued        11 non-null object
meta_data__theme-version         1 non-null object
meta_data__publish_date          1 non-null object
meta_data__al__|                 34 non-null object
meta_data__al__|__url             34 non-null object
meta_data__al__|__app_store_id    18 non-null float64
meta_data__al__|__app_name       26 non-null object
meta_data__msvalidate.01         1 non-null object
meta_data__al__|__package        8 non-null object
meta_data__Date                  8 non-null object
meta_data__fb_admins             11 non-null object
meta_data__fb_app_id             56 non-null float64
meta_data__date                  1 non-null object
meta_data__msapplication-TileImage 5 non-null object
meta_data__application-name      1 non-null object
meta_data__msApplication-PackageFamilyName 1 non-null object
meta_data__article_publisher     7 non-null object
meta_data__theme-color           1 non-null object
meta_data__fb_pages              82 non-null object
meta_data__thumbnail             22 non-null object
meta_data__template-top          22 non-null object
meta_data__vr_canonical          22 non-null object
meta_data__build                 32 non-null object
meta_data__article__section      15 non-null object
meta_data__article__tag          5 non-null object
meta_data__article__published_time 15 non-null object
meta_data__article__modified_time 11 non-null object
meta_data__google-site-verification 14 non-null object
meta_data__verify-v1            1 non-null object
meta_data__Last-Modified         8 non-null object
meta_data__keywords              39 non-null object
meta_data__ov__id                1 non-null float64
meta_data__lastmod               22 non-null object
meta_data__article__author       29 non-null object
meta_data__news_keywords         11 non-null object
canonical_link                   89 non-null object
title                           90 non-null object
url                              90 non-null object
publish_date__$date              55 non-null float64
source                           90 non-null object
dtypes: float64(6), object(86)
memory usage: 84.3+ KB

```

```
In [29]: buzzr.loc[:, 'text']
```

```

Out[29]: 0      Less than a day after protests over the police...
         1      NaN
         2      NaN
         3      NaN

```

4 Obama To UN: 'Giving Up Liberty, Enhances Secu...
26 Getty Images Wealth Of Nations Trump vs. Clint...
29 President Obama today vetoed a bill that would...
30 NaN
31 NaN
32 NaN
33 CHAOS! NC Protest MOB Ambushes Female Truck Dr...
56 Hillary Clinton and Donald Trump ushered the 2...
57 NaN
58 NaN
59 NaN
62 Peaceful protesters crowded Charlotte's first ...
63 NaN
64 NaN
65 NaN
69 Story highlights Bush will deliver his first l...
71 Hillary Clinton's campaign is making one plea ...
74 I've watched every presidential debate ever br...
77 Story highlights "This, though, is certain: to...
98 "I take great pride in the fact that I am actu...
101 The U.S. Attorney's Office in Manhattan is inv...
102 NaN
103 NaN
104 NaN
105 11.7k SHARES SHARE THIS STORY\n\nThe Senate re...
121 Story highlights Trump reportedly used \$258,00...
...
799 Lester Holt, who anchors "NBC Nightly News," w...
802 Under the Radar Blog Archives Select Date... Jul...
805 Story highlights Ted Cruz refused to endorse T...
810 "Crooked Hillary has been fighting ISIS, or wh...
813 Newark (CNN) Former Port Authority executives ...
815 Bono: 'Trump is potentially the worst idea tha...
817 Story highlights The campaigns of Hillary Clin...
819 The earliest actual poll to come out of last n...
828 Story highlights Democratic vice presidential ...
830 If there's anyone in this day and age who most...
839 The Department of Homeland Security's inspecto...
842 The pressure is on Hillary Clinton.\n\nInsider...
845 \n\nPosted by Frank Wilkenmeyer on 19 Sep 2016...
855 Last week I wrote an article titled "Donald Tr...
883 Donald Trump's reputation for being always ava...
897 Pence gives no specifics on why Trump Foundati...
899 As police today captured the man wanted for qu...
900 NaN
901 NaN
902 NaN
907 CLEVELAND — Donald Trump campaigned Wednesday ...
910 "He didn't say the document was legitimate," J...
913 Story highlights Obama has taken steps toward ...

```

932 Story highlights Virginia Gov. Terry McAuliffe...
935 In March, in a bold "Oh yeah?" moment during a...
938 Like much of the mainstream media, the Washing...
965 2016 Obama's team isn't laughing at Trump anym...
967 Story highlights Trump has 45%, Clinton 42% an...
985 There may be a few women out there who enjoy a...
994 Off Message Is Donald Trump qualified to be pr...
Name: text, Length: 116, dtype: object

```

```
In [34]: buzzr=buzzr.drop(columns=['meta_data__theme-version','meta_data__al__'],'')
```

```
In [49]: buzzr
```

```
Out[49]:
```

	top_img	text	authors__001	author
0	http://a.abcnews.com/images/Politics/AP_donald...	Less than a day after protests over the police...	More Candace	Adam
4	http://rightwingnews.com/wp-content/uploads/20...	Obama To UN: 'Giving Up Liberty, Enhances Secu...	Cassy Fiano	
26	http://static.politico.com/e9/11/6144cdc24e319...	Getty Images Wealth Of Nations Trump vs. Clint...	Jack Shafer	Erick
29	http://a.abcnews.com/images/US/AP_Obama_RM_201...	President Obama today vetoed a bill	John	Mo

```
In [51]: buzzr=buzzr.drop([795,796,797])
```

In [52]: buzzr

Out[52]:

	top_img	text	authors_001	author
0	http://a.abcnews.com/images/Politics/AP_donald...	Less than a day after protests over the police...	More Candace	Adam
4	http://rightwingnews.com/wp-content/uploads/20...	Obama To UN: 'Giving Up Liberty, Enhances Secu...	Cassy Fiano	
26	http://static.politico.com/e9/11/6144cdc24e319...	Getty Images Wealth Of Nations Trump vs. Clint...	Jack Shafer	Erick
29	http://a.abcnews.com/images/US/AP_Obama_RM_201...	President Obama today vetoed a bill	John	Mo

In [53]: buzzr=buzzr.fillna('missing')

In [54]: buzzr

Out[54]:

	top_img	text	authors_001	author
0	http://a.abcnews.com/images/Politics/AP_donald...	Less than a day after protests over the police...	More Candace	Adam
4	http://rightwingnews.com/wp-content/uploads/20...	Obama To UN: 'Giving Up Liberty, Enhances Secu...	Cassy Fiano	
26	http://static.politico.com/e9/11/6144cdc24e319...	Getty Images Wealth Of Nations Trump vs. Clint...	Jack Shafer	Erick
29	http://a.abcnews.com/images/US/AP_Obama_RM_201...	President Obama today vetoed a bill	John	Mo

In [55]: buzzr.info()

<class 'pandas.core.frame.DataFrame'>

Int64Index: 95 entries, 0 to 994

Data columns (total 85 columns):

top_img	95 non-null object
text	95 non-null object
authors__001	95 non-null object
authors__002	95 non-null object
authors__003	95 non-null object
meta_data__generator	95 non-null object
authors__004	95 non-null object
authors__005	95 non-null object
authors__006	95 non-null object
authors__007	95 non-null object
authors__008	95 non-null object
authors__009	95 non-null object
authors__010	95 non-null object
meta_data__fb_title	95 non-null object
meta_data__description	95 non-null object
meta_data__title	95 non-null object
meta_data__modernizr-version	95 non-null object
meta_data__foundation-version	95 non-null object
meta_data__jquery-version	95 non-null object
meta_data__pubdate	95 non-null object
meta_data__copyright	95 non-null object
meta_data__author	95 non-null object
meta_data__vr__category	95 non-null object
meta_data__vr__type	95 non-null object
meta_data__og__site_name	95 non-null object
meta_data__og__description	95 non-null object
meta_data__og__pubdate	95 non-null object
meta_data__og__title	95 non-null object
meta_data__og__locale	95 non-null object
meta_data__og__updated_time	95 non-null object
meta_data__og__url	95 non-null object
meta_data__og__image	95 non-null object
meta_data__og__fb_appid	95 non-null object
meta_data__og__type	95 non-null object
meta_data__section	95 non-null object
meta_data__referrer	95 non-null object
meta_data__twitter__description	95 non-null object
meta_data__twitter__title	95 non-null object
meta_data__twitter__url	95 non-null object
meta_data__twitter__image	95 non-null object
meta_data__twitter__creator	95 non-null object
meta_data__object-hash	95 non-null object
meta_data__host	95 non-null object
meta_data__twitter__site	95 non-null object
meta_data__twitter__app__	95 non-null object
meta_data__twitter__app__ __ipad	95 non-null object
meta_data__twitter__app__ __iphone	95 non-null object
meta_data__twitter__app__ __googleplay	95 non-null object


```

meta_data__twitter__card      95 non-null object
meta_data__robots              95 non-null object
meta_data__apple-itunes-app   95 non-null object
meta_data__DC.date.issued     95 non-null object
meta_data__publish_date       95 non-null object
meta_data__al__|__url         95 non-null object
meta_data__Date               95 non-null object
meta_data__fb__admins         95 non-null object
meta_data__fb__app_id         95 non-null object
meta_data__date               95 non-null object
meta_data__msapplication-TileImage 95 non-null object
meta_data__application-name   95 non-null object
meta_data__msApplication-PackageFamilyName 95 non-null object
meta_data__article__publisher 95 non-null object
meta_data__theme-color        95 non-null object
meta_data__fb__pages          95 non-null object
meta_data__thumbnail          95 non-null object
meta_data__template-top       95 non-null object
meta_data__vr__canonical      95 non-null object
meta_data__build              95 non-null object
meta_data__article__section   95 non-null object
meta_data__article__tag       95 non-null object
meta_data__article__published_time 95 non-null object
meta_data__article__modified_time 95 non-null object
meta_data__google-site-verification 95 non-null object
meta_data__verify-v1          95 non-null object
meta_data__Last-Modified      95 non-null object
meta_data__keywords           95 non-null object
meta_data__ov__id             95 non-null object
meta_data__lastmod            95 non-null object
meta_data__article__author    95 non-null object
meta_data__news_keywords      95 non-null object
canonical_link                 95 non-null object
title                          95 non-null object
url                            95 non-null object
publish_date__$date           95 non-null object
source                         95 non-null object
dtypes: object(85)
memory usage: 63.8+ KB

```

```
In [56]: buzzr.loc[:, 'title']
```

```

Out[56]: 0      Donald Trump: Drugs a 'Very, Very Big Factor' ...
         4      Obama To UN: 'Giving Up Liberty, Enhances Secu...
         26     Trump vs. Clinton: A Fundamental Clash over Ho...
         29     President Obama Vetoes 9/11 Victims Bill, Sett...
         33     CHAOS! NC Protest MOB Ambushes Female Truck Dr...
         56     10 Moments That Mattered From Hillary Clinton ...
         62     Young Girl's Emotional Council Speech Laments ...
         69           Jeb Bush to lecture at Harvard this fall

```

71 Clinton vs. Trump: The debate before the debate
74 Is Hillary in a No-Win Situation?
77 Hillary Clinton on police shootings: 'too many...
98 Johnson: 'I'm going to sleep well at the end o...
101 Anthony Weiner Under Federal Investigation for...
105 Senate Dems Move To Force Trump To Release His...
121 Dem congressman to IRS: Audit the Trump Founda...
124 Armed Intruders Break in House...Armed Resident ...
152 Post-debate, Trump team hits moderator Holt fo...
155 One Critically Wounded As Charlotte Police End...
196 Rouhani: Iran deal able to weather the election
198 "How Can we LEAD When we Can't Even Control Ou...
226 9 times Ted Cruz insulted Donald Trump before ...
229 Critical counties: Wake County, NC, could put ...
233 Obama wears hat, breaking 'Politics 101' rule ...
235 Clinton: 'The next 50 days will determine the ...
238 I Was RFK's Speechwriter. Now I'm Voting for T...
241 NC State Troopers Take A Knee Before Protectin...
267 Sacramento mayor punches pie thrower
269 Megan Rapinoe Punished Over Not Standing For A...
277 Liberals Are Furious When Reality Star Brody J...
299 Donna Brazile slams Trump for not apologizing ...
...
787 Magistrate won't order federal arraignment for...
789 US: ISIS did not use mustard agent in base attack
794 Terrorist Attacks Will Likely Affect 2016 Pres...
799 Everything you need to know about Trump and Cl...
802 U.S. attorney sees no rush to arraign Rahami i...
805 Donald Trump's rise puts Ted Cruz in a bind
810 Trump hits Clinton on Islamic State: 'It is ti...
813 Bridgegate: Port Authority officials planned t...
815 Bono: 'Trump is potentially the worst idea tha...
817 2016 presidential race tightens in Florida
819 HuffPo Just Dropped The Mic On Internet Polls ...
828 Tim Kaine on Libertarian ticket: 'Everybody kn...
830 Watch Joe Biden Take Trump's Own Words And Tea...
839 U.S. mistakenly granted citizenship to hundred...
842 Insiders: The heat is on Hillary
845 Donald Trump Jr. Just Pinned A Tweet So Stupid...
855 More Milestone Moments for Donald Trump! - Eag...
883 Donald Trump retreats to friendly media ground
897 Pence gives no specifics on why Trump Foundati...
899 Details Emerge About NYC Bomb Suspect Ahmad Kh...
907 Trump's puzzling pitch to black voters
910 Joe Arpaio: Trump didn't call Obama's birth ce...
913 Barack Obama at Benjamin Netanyahu meeting: 'H...
932 McAuliffe hangs with Willie Nelson -- and his ...
935 The Mystery of Trump's Man in Moscow: Carter Page
938 It's "Trump is HITLER" Month at the Washington...
965 Obama's team isn't laughing at Trump anymore

```

967    Georgia poll: Donald Trump, Hillary Clinton in...
985    Chelsea Handler Gets The Last Word After RNC C...
994        Is Donald Trump qualified to be president?
Name: title, Length: 95, dtype: object

```

```
In [57]: buzzr.head(5)
```

```
Out[57]:
```

	canonical_link	title	url	publish_date_
	abcnews.go.com/Politics/donald-trump-dr...	Donald Trump: Drugs a 'Very, Very Big Factor' ...	http://abcn.ws/2d4lNn9	n
	/rightwingnews.com/barack-obama/obama-un...	Obama To UN: 'Giving Up Liberty, Enhances Secu...	http://rightwingnews.com/barack-obama/obama-un...	1.4744
	www.politico.com/magazine/story/2016/09...	Trump vs. Clinton: A Fundamental Clash over Ho...	http://politi.co/2de2qs0	1.4749
	http://abcnews.go.com/Politics/president-obama...	President Obama Vetoes 9/11 Victims Bill, Sett...	http://abcn.ws/2dh2NFs	n
	/rightwingnews.com/black-lives-matter/ch...	CHAOS! NC Protest MOB Ambushes Female Truck Dr...	http://rightwingnews.com/black-lives-matter/ch...	1.4744

```
In [58]: buzzr.loc[:, 'text']
```

```

Out[58]: 0    Less than a day after protests over the police...
         4    Obama To UN: 'Giving Up Liberty, Enhances Secu...
         26   Getty Images Wealth Of Nations Trump vs. Clint...
         29   President Obama today vetoed a bill that would...
         33   CHAOS! NC Protest MOB Ambushes Female Truck Dr...

```

56 Hillary Clinton and Donald Trump ushered the 2...
62 Peaceful protesters crowded Charlotte's first ...
69 Story highlights Bush will deliver his first l...
71 Hillary Clinton's campaign is making one plea ...
74 I've watched every presidential debate ever br...
77 Story highlights "This, though, is certain: to...
98 "I take great pride in the fact that I am actu...
101 The U.S. Attorney's Office in Manhattan is inv...
105 11.7k SHARES SHARE THIS STORY\n\nThe Senate re...
121 Story highlights Trump reportedly used \$258,00...
124 Friday morning around 4:00, three men – some o...
152 "He didn't ask her about a lot of things she s...
155 One Critically Wounded As Charlotte Police End...
196 Rouhani: Iran deal able to weather the electio...
198 Donald Trump was in Pittsburgh, Pennsylvania T...
226 "Donald, you're a sniveling coward," Ted Cruz ...
229 Critical Counties is a CNN series exploring 11...
233 Obama wears hat, breaking 'Politics 101' rule ...
235 Hillary Clinton said this year's election will...
238 I was a Democrat all my life. I came to Washin...
241 NC State Troopers Take A Knee Before Protectin...
267 Story highlights A protester pied Sacramento's...
269 U.S. women's soccer star Megan Rapinoe could f...
277 Liberals Are Furious When Reality Star Brody J...
299 Donna Brazile said Trump's hectoring of Americ...

...

787 Under the Radar Blog Archives Select Date... Jul...
789 Story highlights A final lab test of ISIS weap...
794 The recent connected bombings in New York and ...
799 Lester Holt, who anchors "NBC Nightly News," w...
802 Under the Radar Blog Archives Select Date... Jul...
805 Story highlights Ted Cruz refused to endorse T...
810 "Crooked Hillary has been fighting ISIS, or wh...
813 Newark (CNN) Former Port Authority executives ...
815 Bono: 'Trump is potentially the worst idea tha...
817 Story highlights The campaigns of Hillary Clin...
819 The earliest actual poll to come out of last n...
828 Story highlights Democratic vice presidential ...
830 If there's anyone in this day and age who most...
839 The Department of Homeland Security's inspecto...
842 The pressure is on Hillary Clinton.\n\nInsider...
845 \n\nPosted by Frank Wilkenmeyer on 19 Sep 2016...
855 Last week I wrote an article titled "Donald Tr...
883 Donald Trump's reputation for being always ava...
897 Pence gives no specifics on why Trump Foundati...
899 As police today captured the man wanted for qu...
907 CLEVELAND – Donald Trump campaigned Wednesday ...
910 "He didn't say the document was legitimate," J...
913 Story highlights Obama has taken steps toward ...
932 Story highlights Virginia Gov. Terry McAuliffe...

```

935 In March, in a bold "Oh yeah?" moment during a...
938 Like much of the mainstream media, the Washing...
965 2016 Obama's team isn't laughing at Trump anym...
967 Story highlights Trump has 45%, Clinton 42% an...
985 There may be a few women out there who enjoy a...
994 Off Message Is Donald Trump qualified to be pr...
Name: text, Length: 95, dtype: object

```

```
In [59]: buzzr.loc[:, 'url']
```

```

Out[59]: 0          http://abcn.ws/2d4lNn9
          (http://abcn.ws/2d4lNn9)
          4          http://rightwingnews.com/barack-obama/obama-un...
          (http://rightwingnews.com/barack-obama/obama-un...)
          26         http://politi.co/2de2qs0
          (http://politi.co/2de2qs0)
          29         http://abcn.ws/2dh2NFs
          (http://abcn.ws/2dh2NFs)
          33         http://rightwingnews.com/black-lives-matter/ch...
          (http://rightwingnews.com/black-lives-matter/ch...)
          56         http://abcn.ws/2d56mMD
          (http://abcn.ws/2d56mMD)
          62         http://abcn.ws/2cTj7ap
          (http://abcn.ws/2cTj7ap)
          69         http://cnn.it/2d7qa20
          (http://cnn.it/2d7qa20)
          71         http://politi.co/2d0l760
          (http://politi.co/2d0l760)
          74         http://politi.co/2d0FTlN
          (http://politi.co/2d0FTlN)
          77         http://cnn.it/2cRxlv5
          (http://cnn.it/2cRxlv5)
          98         http://politi.co/2cx5AZI
          (http://politi.co/2cx5AZI)
          101        http://abcn.ws/2cVpv23
          (http://abcn.ws/2cVpv23)
          105        http://occupydemocrats.com/2016/09/23/senate-d...
          (http://occupydemocrats.com/2016/09/23/senate-d...)
          121        http://cnn.it/2diatmB
          (http://cnn.it/2diatmB)
          124        http://eaglerising.com/36992/caught-on-video-a...
          (http://eaglerising.com/36992/caught-on-video-a...)
          152        http://politi.co/2dojlvI
          (http://politi.co/2dojlvI)
          155        http://rightwingnews.com/black-lives-matter/on...
          (http://rightwingnews.com/black-lives-matter/on...)
          196        http://politi.co/2da5pR0
          (http://politi.co/2da5pR0)
          198        http://eaglerising.com/36966/trump-how-can-we-...
          (http://eaglerising.com/36966/trump-how-can-we-...)

```

226 <http://politi.co/2cSUHmo>
 (<http://politi.co/2cSUHmo>)
 229 <http://cnn.it/2dnoYpa>
 (<http://cnn.it/2dnoYpa>)
 233 <http://politi.co/2dbESny>
 (<http://politi.co/2dbESny>)
 235 <http://politi.co/2cmo9Qd>
 (<http://politi.co/2cmo9Qd>)
 238 <http://politi.co/2cQNTTT>
 (<http://politi.co/2cQNTTT>)
 241 <http://rightwingnews.com/black-lives-matter/nc...>
 (<http://rightwingnews.com/black-lives-matter/nc...>)
 267 <http://cnn.it/2de6xHA>
 (<http://cnn.it/2de6xHA>)
 269 <http://www.opposingviews.com/i/society/megan-r...>
 (<http://www.opposingviews.com/i/society/megan-r...>)
 277 <http://rightwingnews.com/top-news/liberals-fur...>
 (<http://rightwingnews.com/top-news/liberals-fur...>)
 299 <http://politi.co/2cSX97k>
 (<http://politi.co/2cSX97k>)
 ...
 787 <http://politi.co/2cS4PJJ>
 (<http://politi.co/2cS4PJJ>)
 789 <http://cnn.it/2dwOp8d>
 (<http://cnn.it/2dwOp8d>)
 794 <http://abcn.ws/2cDjJ92>
 (<http://abcn.ws/2cDjJ92>)
 799 <http://politi.co/2d3mmAD>
 (<http://politi.co/2d3mmAD>)
 802 <http://politi.co/2cS8sh4>
 (<http://politi.co/2cS8sh4>)
 805 <http://cnn.it/2dlbk5X>
 (<http://cnn.it/2dlbk5X>)
 810 <http://politi.co/2cEsAnL>
 (<http://politi.co/2cEsAnL>)
 813 <http://cnn.it/2dn8vWp>
 (<http://cnn.it/2dn8vWp>)
 815 <http://politi.co/2cN5r3A>
 (<http://politi.co/2cN5r3A>)
 817 <http://cnn.it/2ddxgjJ>
 (<http://cnn.it/2ddxgjJ>)
 819 <http://addictinginfo.org/2016/09/27/huffpo-jus...>
 (<http://addictinginfo.org/2016/09/27/huffpo-jus...>)
 828 <http://cnn.it/2cX5Nof>
 (<http://cnn.it/2cX5Nof>)
 830 <http://addictinginfo.org/2016/09/27/watch-joe-...>
 (<http://addictinginfo.org/2016/09/27/watch-joe-...>)
 839 <http://politi.co/2cKqILg>
 (<http://politi.co/2cKqILg>)
 842 <http://politi.co/2cP9WaV>

```

(http://politi.co/2cP9WaV)
845 http://winningdemocrats.com/donald-trump-jr-ju...
(http://winningdemocrats.com/donald-trump-jr-ju...)
855 http://eaglerising.com/36829/more-milestone-mo...
(http://eaglerising.com/36829/more-milestone-mo...)
883 http://cnn.it/2dcjQUU
(http://cnn.it/2dcjQUU)
897 http://politi.co/2d1BwU6
(http://politi.co/2d1BwU6)
899 http://abcn.ws/2cMLIxB
(http://abcn.ws/2cMLIxB)
907 http://politi.co/2cRQmvU
(http://politi.co/2cRQmvU)
910 http://politi.co/2cLospd
(http://politi.co/2cLospd)
913 http://cnn.it/2dgYW6Z
(http://cnn.it/2dgYW6Z)
932 http://cnn.it/2dbrw9X
(http://cnn.it/2dbrw9X)
935 http://politi.co/2cYpJ8p
(http://politi.co/2cYpJ8p)
938 http://eaglerising.com/36955/its-trump-is-hitl...
(http://eaglerising.com/36955/its-trump-is-hitl...)
965 http://politi.co/2cW2vAD
(http://politi.co/2cW2vAD)
967 http://cnn.it/2cynaZx
(http://cnn.it/2cynaZx)
985 http://www.addictinginfo.org/2016/09/19/chelse...
(http://www.addictinginfo.org/2016/09/19/chelse...)
994 http://politi.co/2dkim5e
(http://politi.co/2dkim5e)
Name: url, Length: 95, dtype: object

```

```
In [60]: buzzr.loc[:, 'canonical_link']
```

```

Out[60]: 0 http://abcnews.go.com/Politics/donald-trump-dr...
(http://abcnews.go.com/Politics/donald-trump-dr...)
4 http://rightwingnews.com/barack-obama/obama-un...
(http://rightwingnews.com/barack-obama/obama-un...)
26 http://www.politico.com/magazine/story/2016/09...
(http://www.politico.com/magazine/story/2016/09...)
29 http://abcnews.go.com/Politics/president-obama...
(http://abcnews.go.com/Politics/president-obama...)
33 http://rightwingnews.com/black-lives-matter/ch...
(http://rightwingnews.com/black-lives-matter/ch...)
56 http://abcnews.go.com/Politics/presidential-de...
(http://abcnews.go.com/Politics/presidential-de...)
62 http://abcnews.go.com/US/young-girls-emotional...
(http://abcnews.go.com/US/young-girls-emotional...)
69 http://www.cnn.com/2016/09/27/politics/jeb-bus...

```

(<http://www.cnn.com/2016/09/27/politics/jeb-bus...>)
71 <http://www.politico.com/story/2016/09/clinton-...>
(<http://www.politico.com/story/2016/09/clinton-...>)
74 <http://www.politico.com/magazine/story/2016/09...>
(<http://www.politico.com/magazine/story/2016/09...>)
77 <http://www.cnn.com/2016/09/21/politics/hillary...>
(<http://www.cnn.com/2016/09/21/politics/hillary...>)
98 <http://www.politico.com/story/2016/09/gary-joh...>
(<http://www.politico.com/story/2016/09/gary-joh...>)
101 <http://abcnews.go.com/Politics/anthony-weiner-...>
(<http://abcnews.go.com/Politics/anthony-weiner-...>)
105 <http://occupydemocrats.com/2016/09/23/senate-d...>
(<http://occupydemocrats.com/2016/09/23/senate-d...>)
121 <http://www.cnn.com/2016/09/21/politics/levin-i...>
(<http://www.cnn.com/2016/09/21/politics/levin-i...>)
124 <http://eaglerising.com/36992/caught-on-video-a...>
(<http://eaglerising.com/36992/caught-on-video-a...>)
152 <http://www.politico.com/story/2016/09/trump-re...>
(<http://www.politico.com/story/2016/09/trump-re...>)
155 <http://rightwingnews.com/black-lives-matter/on...>
(<http://rightwingnews.com/black-lives-matter/on...>)
196 <http://www.politico.com/story/2016/09/iran-dea...>
(<http://www.politico.com/story/2016/09/iran-dea...>)
198 <http://eaglerising.com/36966/trump-how-can-we-...>
(<http://eaglerising.com/36966/trump-how-can-we-...>)
226 <http://www.politico.com/story/2016/09/ted-cruz...>
(<http://www.politico.com/story/2016/09/ted-cruz...>)
229 <http://www.cnn.com/2016/09/23/politics/critica...>
(<http://www.cnn.com/2016/09/23/politics/critica...>)
233 <http://www.politico.com/story/2016/09/barack-o...>
(<http://www.politico.com/story/2016/09/barack-o...>)
235 <http://www.politico.com/story/2016/09/clinton-...>
(<http://www.politico.com/story/2016/09/clinton-...>)
238 <http://www.politico.com/magazine/story/2016/09...>
(<http://www.politico.com/magazine/story/2016/09...>)
241 <http://rightwingnews.com/black-lives-matter/nc...>
(<http://rightwingnews.com/black-lives-matter/nc...>)
267 <http://www.cnn.com/2016/09/22/politics/kevin-j...>
(<http://www.cnn.com/2016/09/22/politics/kevin-j...>)
269 <http://www.opposingviews.com/i/society/megan-r...>
(<http://www.opposingviews.com/i/society/megan-r...>)
277 <http://rightwingnews.com/top-news/liberals-fur...>
(<http://rightwingnews.com/top-news/liberals-fur...>)
299 <http://www.politico.com/story/2016/09/trump-de...>
(<http://www.politico.com/story/2016/09/trump-de...>)
...
787 <http://www.politico.com/blogs/under-the-radar/...>
(<http://www.politico.com/blogs/under-the-radar/...>)
789 <http://www.cnn.com/2016/09/27/politics/lab-tes...>
(<http://www.cnn.com/2016/09/27/politics/lab-tes...>)

794 <http://abcnews.go.com/Politics/terrorist-attac...>
(<http://abcnews.go.com/Politics/terrorist-attac...>)

799 <http://www.politico.com/story/2016/09/everythi...>
(<http://www.politico.com/story/2016/09/everythi...>)

802 <http://www.politico.com/blogs/under-the-radar/...>
(<http://www.politico.com/blogs/under-the-radar/...>)

805 <http://www.cnn.com/2016/09/22/politics/donald-...>
(<http://www.cnn.com/2016/09/22/politics/donald-...>)

810 <http://www.politico.com/story/2016/09/trump-is...>
(<http://www.politico.com/story/2016/09/trump-is...>)

813 <http://www.cnn.com/2016/09/26/politics/bridgeg...>
(<http://www.cnn.com/2016/09/26/politics/bridgeg...>)

815 <http://www.politico.com/story/2016/09/bono-tru...>
(<http://www.politico.com/story/2016/09/bono-tru...>)

817 <http://www.cnn.com/2016/09/20/politics/electio...>
(<http://www.cnn.com/2016/09/20/politics/electio...>)

819 <http://addictinginfo.com/2016/09/27/huffpo-jus...>
(<http://addictinginfo.com/2016/09/27/huffpo-jus...>)

828 <http://www.cnn.com/2016/09/23/politics/tim-kai...>
(<http://www.cnn.com/2016/09/23/politics/tim-kai...>)

830 <http://addictinginfo.com/2016/09/27/watch-joe-...>
(<http://addictinginfo.com/2016/09/27/watch-joe-...>)

839 <http://www.politico.com/story/2016/09/immigran...>
(<http://www.politico.com/story/2016/09/immigran...>)

842 <http://www.politico.com/story/2016/09/debate-t...>
(<http://www.politico.com/story/2016/09/debate-t...>)

845 <http://oppositionreport.com/donald-trump-jr-ju...>
(<http://oppositionreport.com/donald-trump-jr-ju...>)

855 <http://eaglerising.com/36829/more-milestone-mo...>
(<http://eaglerising.com/36829/more-milestone-mo...>)

883 <http://money.cnn.com/2016/09/19/media/donald-t...>
(<http://money.cnn.com/2016/09/19/media/donald-t...>)

897 <http://www.politico.com/story/2016/09/trump-fo...>
(<http://www.politico.com/story/2016/09/trump-fo...>)

899 <http://abcnews.go.com/US/source-suspect-wanted...>
(<http://abcnews.go.com/US/source-suspect-wanted...>)

907 <http://www.politico.com/story/2016/09/is-trump...>
(<http://www.politico.com/story/2016/09/is-trump...>)

910 <http://www.politico.com/story/2016/09/joe-arpa...>
(<http://www.politico.com/story/2016/09/joe-arpa...>)

913 <http://www.cnn.com/2016/09/21/politics/obama-n...>
(<http://www.cnn.com/2016/09/21/politics/obama-n...>)

932 <http://www.cnn.com/2016/09/19/politics/terry-m...>
(<http://www.cnn.com/2016/09/19/politics/terry-m...>)

935 <http://www.politico.com/magazine/story/2016/09...>
(<http://www.politico.com/magazine/story/2016/09...>)

938 <http://eaglerising.com/36955/its-trump-is-hitl...>
(<http://eaglerising.com/36955/its-trump-is-hitl...>)

965 <http://www.politico.com/story/2016/09/barack-o...>
(<http://www.politico.com/story/2016/09/barack-o...>)

```

967     http://www.cnn.com/2016/09/19/politics/georgia...
      (http://www.cnn.com/2016/09/19/politics/georgia...)
985     http://addictinginfo.com/2016/09/19/chelsea-ha...
      (http://addictinginfo.com/2016/09/19/chelsea-ha...)
994     http://www.politico.com/story/2016/09/is-donal...
      (http://www.politico.com/story/2016/09/is-donal...)
Name: canonical_link, Length: 95, dtype: object

```

```
In [61]: buzzr.loc[:, 'source']
```

```

Out[61]: 0          http://abcn.ws (http://abcn.ws)
4          http://rightwingnews.com (http://rightwingnews.com)
26         http://politi.co (http://politi.co)
29         http://abcn.ws (http://abcn.ws)
33         http://rightwingnews.com (http://rightwingnews.com)
56         http://abcn.ws (http://abcn.ws)
62         http://abcn.ws (http://abcn.ws)
69         http://cnn.it (http://cnn.it)
71         http://politi.co (http://politi.co)
74         http://politi.co (http://politi.co)
77         http://cnn.it (http://cnn.it)
98         http://politi.co (http://politi.co)
101        http://abcn.ws (http://abcn.ws)
105        http://occupydemocrats.com (http://occupydemocrats.com)
121        http://cnn.it (http://cnn.it)
124        http://eaglerising.com (http://eaglerising.com)
152        http://politi.co (http://politi.co)
155        http://rightwingnews.com (http://rightwingnews.com)
196        http://politi.co (http://politi.co)
198        http://eaglerising.com (http://eaglerising.com)
226        http://politi.co (http://politi.co)
229        http://cnn.it (http://cnn.it)
233        http://politi.co (http://politi.co)
235        http://politi.co (http://politi.co)
238        http://politi.co (http://politi.co)
241        http://rightwingnews.com (http://rightwingnews.com)
267        http://cnn.it (http://cnn.it)
269        http://www.opposingviews.com (http://www.opposingviews.com)
277        http://rightwingnews.com (http://rightwingnews.com)
299        http://politi.co (http://politi.co)
...
787        http://politi.co (http://politi.co)
789        http://cnn.it (http://cnn.it)
794        http://abcn.ws (http://abcn.ws)
799        http://politi.co (http://politi.co)
802        http://politi.co (http://politi.co)
805        http://cnn.it (http://cnn.it)
810        http://politi.co (http://politi.co)
813        http://cnn.it (http://cnn.it)
815        http://politi.co (http://politi.co)

```

```

817             http://cnn.it (http://cnn.it)
819         http://addictinginfo.org (http://addictinginfo.org)
828             http://cnn.it (http://cnn.it)
830         http://addictinginfo.org (http://addictinginfo.org)
839             http://politi.co (http://politi.co)
842             http://politi.co (http://politi.co)
845         http://winningdemocrats.com (http://winningdemocrats.com)
855         http://eaglerising.com (http://eaglerising.com)
883             http://cnn.it (http://cnn.it)
897             http://politi.co (http://politi.co)
899             http://abcn.ws (http://abcn.ws)
907             http://politi.co (http://politi.co)
910             http://politi.co (http://politi.co)
913             http://cnn.it (http://cnn.it)
932             http://cnn.it (http://cnn.it)
935             http://politi.co (http://politi.co)
938         http://eaglerising.com (http://eaglerising.com)
965             http://politi.co (http://politi.co)
967             http://cnn.it (http://cnn.it)
985         http://www.addictinginfo.org (http://www.addictinginfo.org)
994             http://politi.co (http://politi.co)
Name: source, Length: 95, dtype: object

```

In [63]: buzzr

Out[63]:

	top_img	text	authors__001	author
0	http://a.abcnews.com/images/Politics/AP_donald...	Less than a day after protests over the police...	More Candace	Adam
4	http://rightwingnews.com/wp-content/uploads/20...	Obama To UN: 'Giving Up Liberty, Enhances Secu...	Cassy Fiano	
26	http://static.politico.com/e9/11/6144cdc24e319...	Getty Images Wealth Of Nations Trump vs. Clint...	Jack Shafer	Erick
29	http://a.abcnews.com/images/US/AP_Obama_RM_201...	President Obama today vetoed a bill	John	Mo

```

In [64]: # Adding label to the data
buzzr['real'] = 0

```

```
In [140]: buzzr=buzzr.rename(columns={'real':'fakeness'})
```

```
In [141]: buzzr
```

canonical_link	title	url	publish_date	\$date
Politics/donald-trump-dr...	Donald Trump: Drugs a 'Very, Very Big Factor' ...	http://abcn.ws/2d4lNn9		missing
m/barack-obama/obama-un...	Obama To UN: 'Giving Up Liberty, Enhances Secu...	http://rightwingnews.com/barack-obama/obama-un...	1.47448e+12	http
magazine/story/2016/09...	Trump vs. Clinton: A Fundamental Clash over Ho...	http://politi.co/2de2qs0	1.47497e+12	
jo.com/Politics/president-obama...	President Obama Vetoes 9/11 Victims Bill, S...	http://abcn.ws/2dh2NFs		missing

```
In [143]: # Saving it for the future use
buzzr.to_csv("/Users/sonia/Downloads/finalyearproject/Realnews_clean_all_
```

```
In [67]: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
%matplotlib inline
from sklearn import tree
from sklearn.tree import DecisionTreeClassifier

from sklearn.datasets import load_iris
from sklearn import tree
from sklearn.tree import export_graphviz
import io
from scipy import misc

#read the real news dataset-Politifact
polir= pd.read_csv("/Users/sonia/Downloads/PolitiFact_Real_1-Webpage.csv")
```

In [68]: polir

28	NaN	NaN	NaN
29	NaN	NaN	NaN
...
1608	NaN	NaN	NaN
1609	NaN	NaN	NaN
1610	NaN	NaN	NaN
1611	NaN	NaN	NaN
1612	NaN	NaN	NaN
1613	NaN	NaN	NaN
1614	NaN	NaN	NaN
1615	NaN	NaN	NaN
1616	NaN	NaN	NaN

In [69]: polir=polir.drop(columns='images')

In [70]: polir

Out[70]:

	top_img	text	meta_data__msapplication- TileColor	auth
0	http://rightwingnews.com/wp-content/uploads/20...	Famous dog killed in spot she waited a year fo...	NaN	
1	NaN	NaN	NaN	
2	NaN	NaN	NaN	
3	NaN	NaN	NaN	
4	NaN	NaN	NaN	
5	NaN	NaN	NaN	
6	NaN	NaN	NaN	

In [71]: polir=polir.dropna(axis=0, how='all', thresh=None, subset=None, inplace=F

In [72]: polir

Out[72]:

	top_img	text	meta_data_msapplicatio TileCo
0	http://rightwingnews.com/wp-content/uploads/20...	Famous dog killed in spot she waited a year fo...	N
22	http://i2.cdn.cnn.com/cnnnext/dam/assets/16091...	Story highlights The House Oversight panel vot...	N
43	http://newsbake.com/wp-content/uploads/2016/05...	We are absolutely heartbroken to hear about th...	N
		Nine years ago, a	

In [73]: polir=polir.dropna(axis=1, how='all', thresh=None, subset=None, inplace=F

In [74]: polir

Out[74]:

	top_img	text	meta_data_msapplicatio TileCo
0	http://rightwingnews.com/wp-content/uploads/20...	Famous dog killed in spot she waited a year fo...	N
22	http://i2.cdn.cnn.com/cnnnext/dam/assets/16091...	Story highlights The House Oversight panel vot...	N
43	http://newsbake.com/wp-content/uploads/2016/05...	We are absolutely heartbroken to hear about th...	N
		Nine years ago, a	

In [75]: `polir.info()`

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 141 entries, 0 to 1637
Columns: 113 entries, top_img to source
dtypes: float64(7), object(106)
memory usage: 125.6+ KB
```

In [76]: `polir=polir.drop(columns=['meta_data__msapplication-TileColor','meta_data`

In [77]: `polir`

Out[77]:

	top_img	text	authors__001	meta_dat
0	http://rightwingnews.com/wp-content/uploads/20...	Famous dog killed in spot she waited a year fo...	NaN	
22	http://i2.cdn.cnn.com/cnnnext/dam/assets/16091...	Story highlights The House Oversight panel vot...	Tom Lobianco	
43	http://newsbake.com/wp-content/uploads/2016/05...	We are absolutely heartbroken to hear about th...	Nancy Wells	
		Nine years ago, a		

In [78]: `polir=polir.fillna('missing')`

In [79]: `polir`

Out[79]:

	top_img	text	authors_001	meta_dat
0	http://rightwingnews.com/wp-content/uploads/20...	Famous dog killed in spot she waited a year fo...	missing	
22	http://i2.cdn.cnn.com/cnnnext/dam/assets/16091...	Story highlights The House Oversight panel vot...	Tom Lobianco	
43	http://newsbake.com/wp-content/uploads/2016/05...	We are absolutely heartbroken to hear about th...	Nancy Wells	
		Nine years ago, a		

In [81]: `polir=polir.drop([103,104,105,1635,1636,1637])`

In [82]: `polir`

Out[82]:

	top_img	text	authors_001	meta_dat
0	http://rightwingnews.com/wp-content/uploads/20...	Famous dog killed in spot she waited a year fo...	missing	
22	http://i2.cdn.cnn.com/cnnnext/dam/assets/16091...	Story highlights The House Oversight panel vot...	Tom Lobianco	
43	http://newsbake.com/wp-content/uploads/2016/05...	We are absolutely heartbroken to hear about th...	Nancy Wells	
		Nine years ago, a		

In [83]: `polir=polir.drop([1066,1067])`

In [84]: polir

Out[84]:

	top_img	text	authors_001	meta_dat
0	http://rightwingnews.com/wp-content/uploads/20...	Famous dog killed in spot she waited a year fo...	missing	
22	http://i2.cdn.cnn.com/cnnnext/dam/assets/16091...	Story highlights The House Oversight panel vot...	Tom Lobianco	
43	http://newsbake.com/wp-content/uploads/2016/05...	We are absolutely heartbroken to hear about th...	Nancy Wells	
		Nine years ago, a		

In [85]: polir.loc[:, 'title']

Out[85]:

```

0      Famous dog killed in spot she waited a year fo...
22     House oversight panel votes Clinton IT chief i...
43     America Just Tragically Lost A Country Music I...
92           Monuments to the Battle for the New South
95     RACE TIED on debate day; Colo. and Pa. have tu...
97     Giuliani slams Lester Holt for fact-checking l...
99           Kaine doesn't feel bad for Pence
102    Anti-Trump Super PAC Launches Pre-Debate Video...
106    Glenn Beck shares buyer's remorse on Cruz endo...
108    NY Gov. Andrew Cuomo disputes Trump's claim co...
111    Clinton economy speech to focus on valuing peo...
132    Trump: 'I was going to say something extremely...
134      The 8 Biggest Unforced Errors in Debate History
137    Trump Calls For This Racist Policy To Be Force...
146    Obama Claims It Will Be A Personal Insult If Y...
167      Robert Kennedy Would Have Hated Donald Trump
170    Is West Virginia Holding America's Weirdest El...
173    This Was the Trump Hillary Wanted America to S...
176    Pence: There's too much talk of police departm...
179    Trump Jr. Once Again Shows His Lack Of Humanit...
189    Clinton demands Charlotte police video release...
192    Muslim Afghani Rapes 2-YEAR-OLD, Gets 2 Years ...
221    Trump needles Clinton: 'Sleep well Hillary - s...
224    TRUTH COMES OUT! Keith Lamont Scott's Did This...
285    Trump Won't Tell America Where His Products Ar...
```

```

296     Black Caucus marches to DOJ to protest shootings
299     Trump Silent As Police Credit A Sikh Immigrant...
315     Clinton on Trump's birther claim: A 'racist lie'
318         When does early voting start in every state?
321     Rep. Peter King: We shouldn't be 'pandering' t...
        ...
1001     Twitter Just Wrecked Donald Trump After Horrib...
1010     Meet This Week's White Texan Terrorist Trump I...
1031         McConnell plays hardball in spending fight
1034     What the US government can and cannot do to US...
1062         Trump received $1.6 million from Secret Service
1065     ABC News – Breaking News, Latest News, Headlin...
1068         Ivanka Trump, Paul Ryan meet
1086     Schumer transfers millions to Dems in bid for ...
1089     Trump Said Not Paying Taxes Makes Him "Smart."...
1105     Transparent creator Jill Soloway compares 'mon...
1228         Airstrikes Bury Syrians in Aleppo – ABC News
1235     America's relationships are on the line – no m...
1238     The Word Liberals are TERRIFIED Of – Eagle Rising
1265     Clint Eastwood Had Enough, Blows Up On Hillary...
1325     Italian prime minister jokes about Bill Clinto...
1328     WWII Veteran Pens Brutal Op-Ed Slamming Draft-...
1344     Louisiana Cop Claims Murdering A 6-Year Old Ch...
1353     Paul Ryan defends Donald Trump, dodges on Mach...
1371     Trump launches 'Crooked Hillary' Snapchat filt...
1383     Trump Said Not Paying Taxes Makes Him "Smart."...
1399     Lawyer Jailed For Wearing Black Lives Matter P...
1407     Charity: Clinton Foundation Distributed "Water...
1435         Trump's High-Energy War on American Politics
1438     Donald Trump is RIGHT on Profiling [VIDEO] – E...
1466     Hillary Clinton prepares for unpredictable Tru...
1473     Donald Trump, Germany's disfavored son – POLITICO
1488     BREAKING: Hollywood Legend Just Died Of Terrib...
1538         Worst. President. Ever.
1541     Don King drops N-word while introducing Donald...
1634     Donald Trump Jr. Compares Syrian Refugees to S...
Name: title, Length: 133, dtype: object

```

```
In [86]: polir.loc[:, 'text']
```

```

Out[86]: 0     Famous dog killed in spot she waited a year fo...
22     Story highlights The House Oversight panel vot...
43     We are absolutely heartbroken to hear about th...
92     Nine years ago, a driver lost control of his p...
95     RACE TIED on debate day; Colo. and Pa. have tu...
97     Giuliani slams Lester Holt for fact-checking l...
99     Tim Kaine said he doesn't feel bad for his cou...
102     An anti-Donald Trump super PAC is going on the...
106     Glenn Beck shares buyer's remorse on Cruz endo...
108     Story highlights "I don't know how you could h...

```

111 Story highlights Clinton is offering voters an...
132 Trump: 'I was going to say something extremely...
134 On the night of September 26, as the house lig...
137 Donald Trump is calling for one of the most co...
146 Obama Claims It Will Be A Personal Insult If Y...
167 AP Photo Backtalk Robert Kennedy Would Have Ha...
170 A billionaire who built up a corporate empire ...
173 If I hadn't seen the last fifteen months of th...
176 Mike Pence said "we ought to set aside this ta...
179 We've all seen the horrific photos of children...
189 Charlotte-Mecklenburg Police Chief Kerr Putney...
192 A 22-year-old Muslim man from Afghanistan prou...
221 Donald Trump mocked Hillary Clinton's planes a...
224 2.1k SHARES Facebook Twitter\n\nFor the past w...
285 Donald Trump constantly talks about how he'll ...
296 Reps. John Lewis (D-Ga.), Brenda Lawrence (D-M...
299 55.6k SHARES SHARE THIS STORY\n\nOnce again, D...
315 "He has a long record of engaging in racist be...
318 In total, 37 states and the District of Columb...
321 Rep. Peter King: We shouldn't be 'pandering' t...
...
1001 \n\nPosted by John Prager on 20 Sep 2016\n\nTh...
1010 21.2k SHARES SHARE THIS STORY\n\nWhile the med...
1031 Sen. Mitch McConnell said a filed spending mea...
1034 Story highlights Ahmad Khan Raham is alleged t...
1062 Donald Trump's campaign isn't alone in patroni...
1065 External links are provided for reference purp...
1068 Story highlights House Speaker Paul Ryan and D...
1086 Sen. Chuck Schumer's money could give Democrat...
1089 72k SHARES SHARE THIS STORY\n\nThe first debat...
1105 Story highlights "Transparent" creator Jill So...
1228 A Syrian family leaves the area following a re...
1235 NEW YORK – All politics is geopolitical in thi...
1238 Terrorism is terrorism, regardless of the sour...
1265 7.6k SHARES Facebook Twitter\n\nHollywood lege...
1325 Italian Prime Minister Matteo Renzi said his w...
1328 7.8k SHARES SHARE THIS STORY\n\nA ninety-six-y...
1344 A Louisiana deputy city marshal is in hot wate...
1353 Story highlights Trump on Tuesday defended cal...
1371 Donald Trump is turning to Snapchat to drum up...
1383 72k SHARES SHARE THIS STORY\n\nThe first debat...
1399 An Ohio judge ordered an attorney jailed for f...
1407 Former President Bill Clinton and his Clinton ...
1435 There is a moment at every Donald Trump rally ...
1438 HBO's Real Time host Bill Maher – someone not ...
1466 New York (CNN) Hillary Clinton is visiting onl...
1473 KALLSTADT, Germany – Few places in Germany are...
1488 Hollywood loses yet another one of their deare...
1538 As my 25th wedding anniversary approached, I t...
1541 Story highlights Trump was sitting in a chair ...

1634 Donald Trump Jr., a son of the Republican pres...
 Name: text, Length: 133, dtype: object

```
In [87]: polir.loc[:, 'url']
```

```
Out[87]: 0      http://rightwingnews.com/top-news/famous-dog-k...
         (http://rightwingnews.com/top-news/famous-dog-k...)
         22      http://cnn.it/2deaH2d
         (http://cnn.it/2deaH2d)
         43      http://newsbake.com/entertainment-news/music-e...
         (http://newsbake.com/entertainment-news/music-e...)
         92      http://politi.co/2dd9U1x
         (http://politi.co/2dd9U1x)
         95      http://politi.co/2dkzgKC
         (http://politi.co/2dkzgKC)
         97      http://politi.co/2dd2Ajf
         (http://politi.co/2dd2Ajf)
         99      http://politi.co/2cRoo3f
         (http://politi.co/2cRoo3f)
         102     http://abcn.ws/2dgPtRo
         (http://abcn.ws/2dgPtRo)
         106     http://politi.co/2dmOPx8
         (http://politi.co/2dmOPx8)
         108     http://cnn.it/2d3tTeL
         (http://cnn.it/2d3tTeL)
         111     http://cnn.it/2deB1Fu
         (http://cnn.it/2deB1Fu)
         132     http://politi.co/2cYOq5I
         (http://politi.co/2cYOq5I)
         134     http://politi.co/2dlg4MU
         (http://politi.co/2dlg4MU)
         137     http://www.addictinginfo.org/2016/09/21/trump-...
         (http://www.addictinginfo.org/2016/09/21/trump-...)
         146     http://rightwingnews.com/top-news/obama-claims...
         (http://rightwingnews.com/top-news/obama-claims...)
         167     http://politi.co/2cGb5l5
         (http://politi.co/2cGb5l5)
         170     http://politi.co/2d2EDuM
         (http://politi.co/2d2EDuM)
         173     http://politi.co/2dcfCgO
         (http://politi.co/2dcfCgO)
         176     http://politi.co/2cUFU8G
         (http://politi.co/2cUFU8G)
         179     http://www.addictinginfo.org/2016/09/20/trump-...
         (http://www.addictinginfo.org/2016/09/20/trump-...)
         189     http://politi.co/2d6lgRT
         (http://politi.co/2d6lgRT)
         192     http://eaglerising.com/36944/muslim-afghani-ra...
         (http://eaglerising.com/36944/muslim-afghani-ra...)
         221     http://politi.co/2cGp6EH
```

```

(http://politi.co/2cGp6EH)
224      http://freedomdaily.com/truth-comes-keith-lamo...
(http://freedomdaily.com/truth-comes-keith-lamo...)
285      http://addictinginfo.org/2016/09/26/trump-wont...
(http://addictinginfo.org/2016/09/26/trump-wont...)
296      http://politi.co/2cTzPJz
(http://politi.co/2cTzPJz)
299      http://occupydemocrats.com/2016/09/20/trump-si...
(http://occupydemocrats.com/2016/09/20/trump-si...)
315      http://politi.co/2d25FGy
(http://politi.co/2d25FGy)
318      http://politi.co/2d3R8Jw
(http://politi.co/2d3R8Jw)
321      http://politi.co/2cNmBKW
(http://politi.co/2cNmBKW)

...

1001     http://winningdemocrats.com/twitter-just-wreck...
(http://winningdemocrats.com/twitter-just-wreck...)
1010     http://occupydemocrats.com/2016/09/20/meet-wee...
(http://occupydemocrats.com/2016/09/20/meet-wee...)
1031     http://politi.co/2cGf4x0
(http://politi.co/2cGf4x0)
1034     http://cnn.it/2d6Dcug
(http://cnn.it/2d6Dcug)
1062     http://politi.co/2cQ4Ssc
(http://politi.co/2cQ4Ssc)
1065     http://abcn.ws/2cLOLtg
(http://abcn.ws/2cLOLtg)
1068     http://cnn.it/2d6HcPU
(http://cnn.it/2d6HcPU)
1086     http://politi.co/2cYJWvx
(http://politi.co/2cYJWvx)
1089     http://occupydemocrats.com/2016/09/28/trump-sa...
(http://occupydemocrats.com/2016/09/28/trump-sa...)
1105     http://cnn.it/2d5RPSS
(http://cnn.it/2d5RPSS)
1228     http://abcn.ws/2d6AS7G
(http://abcn.ws/2d6AS7G)
1235     http://politi.co/2dlxPwp
(http://politi.co/2dlxPwp)
1238     http://eaglerising.com/36863/terrorism-the-wor...
(http://eaglerising.com/36863/terrorism-the-wor...)
1265     http://freedomdaily.com/clint-eastwood-had-eno...
(http://freedomdaily.com/clint-eastwood-had-eno...)
1325     http://politi.co/2d208t0
(http://politi.co/2d208t0)
1328     http://occupydemocrats.com/2016/09/26/wwii-vet...
(http://occupydemocrats.com/2016/09/26/wwii-vet...)
1344     http://author.addictinginfo.org/2016/09/23/lou...
(http://author.addictinginfo.org/2016/09/23/lou...)

```

```

1353 http://cnn.it/2dx08Ud
(http://cnn.it/2dx08Ud)
1371 http://cnn.it/2du7wQg
(http://cnn.it/2du7wQg)
1383 http://occupydemocrats.com/2016/09/28/trump-sa...
(http://occupydemocrats.com/2016/09/28/trump-sa...)
1399 http://www.opposingviews.com/i/society/judge-t...
(http://www.opposingviews.com/i/society/judge-t...)
1407 http://eaglerising.com/36899/charity-clinton-f...
(http://eaglerising.com/36899/charity-clinton-f...)
1435 http://politi.co/2cZBGip
(http://politi.co/2cZBGip)
1438 http://eaglerising.com/37033/bill-maher-donald...
(http://eaglerising.com/37033/bill-maher-donald...)
1466 http://cnn.it/2cMNF03
(http://cnn.it/2cMNF03)
1473 http://politi.co/2csN1WG
(http://politi.co/2csN1WG)
1488 http://newsbake.com/entertainment-news/tv/brea...
(http://newsbake.com/entertainment-news/tv/brea...)
1538 http://politi.co/2dlqa5t
(http://politi.co/2dlqa5t)
1541 http://cnn.it/2dh5vq9
(http://cnn.it/2dh5vq9)
1634 http://abcn.ws/2cFVuaq
(http://abcn.ws/2cFVuaq)
Name: url, Length: 133, dtype: object

```

```
In [88]: polir.loc[:, 'source']
```

```

Out[88]: 0 http://rightwingnews.com (http://rightwingnews.com)
22 http://cnn.it (http://cnn.it)
43 http://newsbake.com (http://newsbake.com)
92 http://politi.co (http://politi.co)
95 http://politi.co (http://politi.co)
97 http://politi.co (http://politi.co)
99 http://politi.co (http://politi.co)
102 http://abcn.ws (http://abcn.ws)
106 http://politi.co (http://politi.co)
108 http://cnn.it (http://cnn.it)
111 http://cnn.it (http://cnn.it)
132 http://politi.co (http://politi.co)
134 http://politi.co (http://politi.co)
137 http://www.addictinginfo.org (http://www.addictinginfo.org)
146 http://rightwingnews.com (http://rightwingnews.com)
167 http://politi.co (http://politi.co)
170 http://politi.co (http://politi.co)
173 http://politi.co (http://politi.co)
176 http://politi.co (http://politi.co)
179 http://www.addictinginfo.org (http://www.addictinginfo.org)

```

```

189          http://politi.co (http://politi.co)
192          http://eaglerising.com (http://eaglerising.com)
221          http://politi.co (http://politi.co)
224          http://freedomdaily.com (http://freedomdaily.com)
285          http://addictinginfo.org (http://addictinginfo.org)
296          http://politi.co (http://politi.co)
299          http://occupydemocrats.com (http://occupydemocrats.com)
315          http://politi.co (http://politi.co)
318          http://politi.co (http://politi.co)
321          http://politi.co (http://politi.co)
...
1001         http://winningdemocrats.com (http://winningdemocrats.com)
1010         http://occupydemocrats.com (http://occupydemocrats.com)
1031         http://politi.co (http://politi.co)
1034         http://cnn.it (http://cnn.it)
1062         http://politi.co (http://politi.co)
1065         http://abcn.ws (http://abcn.ws)
1068         http://cnn.it (http://cnn.it)
1086         http://politi.co (http://politi.co)
1089         http://occupydemocrats.com (http://occupydemocrats.com)
1105         http://cnn.it (http://cnn.it)
1228         http://abcn.ws (http://abcn.ws)
1235         http://politi.co (http://politi.co)
1238         http://eaglerising.com (http://eaglerising.com)
1265         http://freedomdaily.com (http://freedomdaily.com)
1325         http://politi.co (http://politi.co)
1328         http://occupydemocrats.com (http://occupydemocrats.com)
1344         http://author.addictinginfo.org
(http://author.addictinginfo.org)
1353         http://cnn.it (http://cnn.it)
1371         http://cnn.it (http://cnn.it)
1383         http://occupydemocrats.com (http://occupydemocrats.com)
1399         http://www.opposingviews.com (http://www.opposingviews.com)
1407         http://eaglerising.com (http://eaglerising.com)
1435         http://politi.co (http://politi.co)
1438         http://eaglerising.com (http://eaglerising.com)
1466         http://cnn.it (http://cnn.it)
1473         http://politi.co (http://politi.co)
1488         http://newsbake.com (http://newsbake.com)
1538         http://politi.co (http://politi.co)
1541         http://cnn.it (http://cnn.it)
1634         http://abcn.ws (http://abcn.ws)
Name: source, Length: 133, dtype: object

```

```
In [89]: polir.loc[:, 'canonical_link']
```

```

Out[89]: 0      http://rightwingnews.com/top-news/famous-dog-k...
(http://rightwingnews.com/top-news/famous-dog-k...)
22      http://www.cnn.com/2016/09/22/politics/bryan-p...
(http://www.cnn.com/2016/09/22/politics/bryan-p...)

```

43 <http://newsbake.com/entertainment-news/music-e...>
(<http://newsbake.com/entertainment-news/music-e...>)

92 <http://www.politico.com/magazine/story/2016/09...>
(<http://www.politico.com/magazine/story/2016/09...>)

95 <http://www.politico.com/tipsheets/playbook/201...>
(<http://www.politico.com/tipsheets/playbook/201...>)

97 <http://www.politico.com/story/2016/09/giuliani...>
(<http://www.politico.com/story/2016/09/giuliani...>)

99 <http://www.politico.com/story/2016/09/tim-kain...>
(<http://www.politico.com/story/2016/09/tim-kain...>)

102 <http://abcnews.go.com/Politics/anti-trump-supe...>
(<http://abcnews.go.com/Politics/anti-trump-supe...>)

106 <http://www.politico.com/story/2016/09/glenn-be...>
(<http://www.politico.com/story/2016/09/glenn-be...>)

108 <http://www.cnn.com/2016/09/19/politics/andrew-...>
(<http://www.cnn.com/2016/09/19/politics/andrew-...>)

111 <http://www.cnn.com/2016/09/20/politics/hillary...>
(<http://www.cnn.com/2016/09/20/politics/hillary...>)

132 <http://www.politico.com/story/2016/09/what-was...>
(<http://www.politico.com/story/2016/09/what-was...>)

134 <http://www.politico.com/magazine/story/2016/09...>
(<http://www.politico.com/magazine/story/2016/09...>)

137 <http://addictinginfo.com/2016/09/21/trump-call...>
(<http://addictinginfo.com/2016/09/21/trump-call...>)

146 <http://rightwingnews.com/top-news/obama-claims...>
(<http://rightwingnews.com/top-news/obama-claims...>)

167 <http://www.politico.com/magazine/story/2016/09...>
(<http://www.politico.com/magazine/story/2016/09...>)

170 <http://www.politico.com/magazine/story/2016/09...>
(<http://www.politico.com/magazine/story/2016/09...>)

173 <http://www.politico.com/magazine/story/2016/09...>
(<http://www.politico.com/magazine/story/2016/09...>)

176 <http://www.politico.com/story/2016/09/mike-pen...>
(<http://www.politico.com/story/2016/09/mike-pen...>)

179 <http://addictinginfo.com/2016/09/20/trump-jr-o...>
(<http://addictinginfo.com/2016/09/20/trump-jr-o...>)

189 <http://www.politico.com/story/2016/09/keith-sc...>
(<http://www.politico.com/story/2016/09/keith-sc...>)

192 <http://eaglerising.com/36944/muslim-afghani-ra...>
(<http://eaglerising.com/36944/muslim-afghani-ra...>)

221 <http://www.politico.com/story/2016/09/trump-cl...>
(<http://www.politico.com/story/2016/09/trump-cl...>)

224 <http://freedomdaily.com/truth-comes-keith-lamo...>
(<http://freedomdaily.com/truth-comes-keith-lamo...>)

285 <http://addictinginfo.com/2016/09/26/trump-wont...>
(<http://addictinginfo.com/2016/09/26/trump-wont...>)

296 <http://www.politico.com/story/2016/09/congress...>
(<http://www.politico.com/story/2016/09/congress...>)

299 <http://occupydemocrats.com/2016/09/20/trump-si...>
(<http://occupydemocrats.com/2016/09/20/trump-si...>)

315 <http://www.politico.com/story/2016/09/clinton-...>
(<http://www.politico.com/story/2016/09/clinton-...>)

318 <http://www.politico.com/story/2016/09/early-vo...>
(<http://www.politico.com/story/2016/09/early-vo...>)

321 <http://www.politico.com/story/2016/09/peter-ki...>
(<http://www.politico.com/story/2016/09/peter-ki...>)

...

1001 <http://oppositionreport.com/twitter-just-wreck...>
(<http://oppositionreport.com/twitter-just-wreck...>)

1010 <http://occupydemocrats.com/2016/09/20/meet-wee...>
(<http://occupydemocrats.com/2016/09/20/meet-wee...>)

1031 <http://www.politico.com/story/2016/09/mitch-mc...>
(<http://www.politico.com/story/2016/09/mitch-mc...>)

1034 <http://www.cnn.com/2016/09/20/politics/legal-r...>
(<http://www.cnn.com/2016/09/20/politics/legal-r...>)

1062 <http://www.politico.com/story/2016/09/donald-t...>
(<http://www.politico.com/story/2016/09/donald-t...>)

1065 <http://abcnews.go.com>
(<http://abcnews.go.com>)

1068 <http://www.cnn.com/2016/09/19/politics/ivanka-...>
(<http://www.cnn.com/2016/09/19/politics/ivanka-...>)

1086 <http://www.politico.com/story/2016/09/chuck-sc...>
(<http://www.politico.com/story/2016/09/chuck-sc...>)

1089 <http://occupydemocrats.com/2016/09/28/trump-sa...>
(<http://occupydemocrats.com/2016/09/28/trump-sa...>)

1105 <http://www.cnn.com/2016/09/19/politics/jill-so...>
(<http://www.cnn.com/2016/09/19/politics/jill-so...>)

1228 <http://abcnews.go.com/International/photos/air...>
(<http://abcnews.go.com/International/photos/air...>)

1235 <http://www.politico.com/story/2016/09/america...>
(<http://www.politico.com/story/2016/09/america...>)

1238 <http://eaglerising.com/36863/terrorism-the-wor...>
(<http://eaglerising.com/36863/terrorism-the-wor...>)

1265 <http://freedomdaily.com/clint-eastwood-had-eno...>
(<http://freedomdaily.com/clint-eastwood-had-eno...>)

1325 <http://www.politico.com/story/2016/09/matteo-r...>
(<http://www.politico.com/story/2016/09/matteo-r...>)

1328 <http://occupydemocrats.com/2016/09/26/wwii-vet...>
(<http://occupydemocrats.com/2016/09/26/wwii-vet...>)

1344 <http://addictinginfo.com/2016/09/23/louisiana-...>
(<http://addictinginfo.com/2016/09/23/louisiana-...>)

1353 <http://www.cnn.com/2016/09/27/politics/paul-ry...>
(<http://www.cnn.com/2016/09/27/politics/paul-ry...>)

1371 <http://money.cnn.com/2016/09/26/technology/tru...>
(<http://money.cnn.com/2016/09/26/technology/tru...>)

1383 <http://occupydemocrats.com/2016/09/28/trump-sa...>
(<http://occupydemocrats.com/2016/09/28/trump-sa...>)

1399 <http://www.opposingviews.com/i/society/judge-t...>
(<http://www.opposingviews.com/i/society/judge-t...>)

1407 <http://eaglerising.com/36899/charity-clinton-f...>

```
(http://eaglerising.com/36899/charity-clinton-f...)
1435 http://www.politico.com/magazine/story/2016/09...
(http://www.politico.com/magazine/story/2016/09...)
1438 http://eaglerising.com/37033/bill-maher-donald...
(http://eaglerising.com/37033/bill-maher-donald...)
1466 http://www.cnn.com/2016/09/20/politics/hillary...
(http://www.cnn.com/2016/09/20/politics/hillary...)
1473 http://www.politico.eu/article/donald-trump-an...
(http://www.politico.eu/article/donald-trump-an...)
1488 http://newsbake.com/entertainment-news/tv/brea...
(http://newsbake.com/entertainment-news/tv/brea...)
1538 http://www.politico.com/magazine/story/2016/09...
(http://www.politico.com/magazine/story/2016/09...)
1541 http://www.cnn.com/2016/09/21/politics/don-kin...
(http://www.cnn.com/2016/09/21/politics/don-kin...)
1634 http://abcnews.go.com/Politics/donald-trump-jr...
(http://abcnews.go.com/Politics/donald-trump-jr...)
Name: canonical_link, Length: 133, dtype: object
```

```
In [95]: polir.head(5)
```

```
Out[95]:
```

	top_img	text	authors__001	meta_data__publishe
0	http://rightwingnews.com/wp-content/uploads/20...	Famous dog killed in spot she waited a year fo...	missing	missin
22	http://i2.cdn.cnn.com/cnnnext/dam/assets/16091...	Story highlights The House Oversight panel vot...	Tom Lobianco	missin
43	http://newsbake.com/wp-content/uploads/2016/05...	We are absolutely heartbroken to hear about th...	Nancy Wells	missin
92	http://static.politico.com/25/ed/85332de14c45b...	Nine years ago, a driver lost control of his p...	Jack Shafer	missin
95	http://static.politico.com/da/f5/44342c424c68b...	RACE TIED on debate day; Colo. and Pa. have tu...	Daniel Lippman	missin

5 rows × 108 columns



```
In [134]: polir = polir.rename(columns={'real': 'fakeness'})
```

In [135]: polir

Out[135]:

	top_img	text	authors__001	meta_dat
0	http://rightwingnews.com/wp-content/uploads/20...	Famous dog killed in spot she waited a year fo...	missing	
22	http://i2.cdn.cnn.com/cnnnext/dam/assets/16091...	Story highlights The House Oversight panel vot...	Tom Lobianco	
43	http://newsbake.com/wp-content/uploads/2016/05...	We are absolutely heartbroken to hear about th...	Nancy Wells	
		Nine years ago, a		

In [136]: polir

Out[136]:

	top_img	text	authors__001	meta_dat
0	http://rightwingnews.com/wp-content/uploads/20...	Famous dog killed in spot she waited a year fo...	missing	
22	http://i2.cdn.cnn.com/cnnnext/dam/assets/16091...	Story highlights The House Oversight panel vot...	Tom Lobianco	
43	http://newsbake.com/wp-content/uploads/2016/05...	We are absolutely heartbroken to hear about th...	Nancy Wells	
		Nine years ago, a		

In [137]: `polir.info()`

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 133 entries, 0 to 1634
Columns: 108 entries, top_img to fakeness
dtypes: int64(1), object(107)
memory usage: 113.3+ KB
```

In [138]: *#saving it as a csv file*

```
polir.to_csv("/Users/sonia/Downloads/finalyearproject/Realnews_clean_all_
```

In [139]: `polir`

Out[139]:

	top_img	text	authors__001	meta_dat
0	http://rightwingnews.com/wp-content/uploads/20...	Famous dog killed in spot she waited a year fo...	missing	
22	http://i2.cdn.cnn.com/cnnnext/dam/assets/16091...	Story highlights The House Oversight panel vot...	Tom Lobianco	
43	http://newsbake.com/wp-content/uploads/2016/05...	We are absolutely heartbroken to hear about th...	Nancy Wells	
		Nine years ago, a		

In [107]:

In []: