In [5]:
```python
import bs4
fh=open('TrumpLies.html',encoding='utf-8')
soup = bs4.BeautifulSoup(fh,'html.parser')
soup
```

Out[5]:
```
<!DOCTYPE html>

<!--[if (gt IE 9)|!(IE)]> <!-->
<html class="no-js page-interactive section-opinion page-theme-standard tone-
opinion page-interactive-default limit-small layout-xlarge app-interactive" i
temid="https://www.nytimes.com/interactive/2017/06/23/opinion/trumps-lies.htm
l" itemscope="" itemtype="https://schema.org/NewsArticle" lang="en" xmlns:og
="http://opengraphprotocol.org/schema/">
<!--<![endif]-->
<!--[if IE 9]> <html lang="en" class="no-js ie9 lt-ie10 page-interactive sect
ion-opinion page-theme-standard tone-opinion page-interactive-default limit-s
mall layout-xlarge app-interactive" xmlns:og="http://opengraphprotocol.org/sc
hema/"> <![endif]-->
<!--[if IE 8]> <html lang="en" class="no-js ie8 lt-ie10 lt-ie9 page-interacti
ve section-opinion page-theme-standard tone-opinion page-interactive-default
limit-small layout-xlarge app-interactive" xmlns:og="http://opengraphprotoco
l.org/schema/"> <![endif]-->
<!--[if (lt IE 8)]> <html lang="en" class="no-js lt-ie10 lt-ie9 lt-ie8 page-i
nteractive section-opinion page-theme-standard tone-opinion page-interactive-
```

In [11]:
```python
data=soup.title.text
data.strip()
```

Out[11]: `'Opinion | President Trump's Lies, the Definitive List - The New York Times'`

```
In [13]:    1  soup.find('p',class_='g-body').get_text().strip().split()
```

```
Out[13]:  ['Many',
           'Americans',
           'have',
           'become',
           'accustomed',
           'to',
           'President',
           'Trump's',
           'lies.',
           'But',
           'as',
           'regular',
           'as',
           'they',
           'have',
           'become,',
           'the',
           'country',
           'should',
           'not',
           'allow',
           'itself',
           'to',
           'become',
           'numb',
           'to',
           'them.',
           'So',
           'we',
           'have',
           'catalogued',
           'nearly',
           'every',
           'outright',
           'lie',
           'he',
           'has',
           'told',
           'publicly',
           'since',
           'taking',
           'the',
           'oath',
           'of',
           'office.',
           'Updated',
           ':',
           'The',
           'president',
           'is',
           'still',
           'lying,',
           'so',
           "we've",
           'added',
           'to',
           'this',
```

```
 'list,',
 'taking',
 'it',
 'through',
 'Nov.',
 '11,',
 'and',
 'provided',
 'links',
 'to',
 'the',
 'facts',
 'in',
 'each',
 'case.']
```

In [19]:
```python
1  ps = soup.findAll('p',class_='sccret-word')
2  ps
```

Out[19]:
```
[<p class="sccret-word">inscription</p>,
 <p class="sccret-word">Ampex</p>,
 <p class="sccret-word">bumblebee</p>,
 <p class="sccret-word">nettle</p>,
 <p class="sccret-word">helix</p>,
 <p class="sccret-word">administrable</p>,
 <p class="sccret-word">edict</p>,
 <p class="sccret-word">Greenbelt</p>,
 <p class="sccret-word">ibis</p>,
 <p class="sccret-word">sinkhole</p>,
 <p class="sccret-word">sonar</p>,
 <p class="sccret-word">seminary</p>,
 <p class="sccret-word">econometric</p>,
 <p class="sccret-word">stampede</p>,
 <p class="sccret-word">illegitimacy</p>,
 <p class="sccret-word">Andorra</p>,
 <p class="sccret-word">sworn</p>,
 <p class="sccret-word">flatland</p>,
 <p class="sccret-word">marksman</p>,
```

In [30]:
```python
ps = soup.find_all('span',class_='short-desc')
for i in ps:
    print(i.find('a')['href'])
```

https://www.buzzfeed.com/andrewkaczynski/in-2002-donald-trump-said-he-support
ed-invading-iraq-on-the (https://www.buzzfeed.com/andrewkaczynski/in-2002-don
ald-trump-said-he-supported-invading-iraq-on-the)
http://nation.time.com/2013/11/06/10-things-you-didnt-know-about-time/ (htt
p://nation.time.com/2013/11/06/10-things-you-didnt-know-about-time/)
https://www.nytimes.com/2017/01/23/us/politics/donald-trump-congress-democrat
s.html (https://www.nytimes.com/2017/01/23/us/politics/donald-trump-congress-
democrats.html)
https://www.nytimes.com/2017/01/21/us/politics/trump-white-house-briefing-ina
uguration-crowd-size.html (https://www.nytimes.com/2017/01/21/us/politics/tru
mp-white-house-briefing-inauguration-crowd-size.html)
https://www.nytimes.com/2017/01/24/us/politics/unauthorized-immigrant-voting-
trump-lie.html (https://www.nytimes.com/2017/01/24/us/politics/unauthorized-i
mmigrant-voting-trump-lie.html)
https://www.nytimes.com/2017/03/13/us/politics/fact-check-trump-obamacare-hea
lth-care.html (https://www.nytimes.com/2017/03/13/us/politics/fact-check-trum
p-obamacare-health-care.html)
https://www.dnainfo.com/chicago/2017-chicago-murders/timeline?mon=1 (https://
www.dnainfo.com/chicago/2017-chicago-murders/timeline?mon=1)

- ## 1. Extract all the quotes with authors name

- ## 2. Authors whose push quotes appeared more than once on the page

- ## 3. extract quotes that has word "Love"

- ## 4. list out all author name with their link (if available)

- ## 5. printout Quotes longer than 100 charaters

- ## 6. extracrt all quotes with no. of likes and sort then it in desceding order

- ## 7. Quotes author name and text make Dataframe

In [2]:
```python
import bs4
file = open('Popular_Quotes.html', encoding='utf-8')
soup = bs4.BeautifulSoup(file, 'html.parser')
soup
```

Out[2]:
```
<!DOCTYPE html>

<html class="desktop withSiteHeaderTopFullImage">
<head>
<title>Popular Quotes</title>
<meta content="Popular quotes from Goodreads members. Oscar Wilde: 'Be yourse
lf; everyone else is already taken.', Marilyn Monroe: 'I'm selfish, impatient
and a little ..." name="description"/>
<meta content="telephone=no" name="format-detection"/>
<link href="https://www.goodreads.com/quotes" rel="canonical"/>
<!-- * Copied from https://info.analytics.a2z.com/#/docs/data_collection/csa/
onboard (https://info.analytics.a2z.com/#/docs/data_collection/csa/onboard)
*/ -->
<script>
  //<![CDATA[
    !function(){function n(n,t){var r=i(n);return t&&(r=r("instance",t)),r}va
r r=[],c=0,i=function(t){return function(){var n=c++;return r.push([t,[].slic
e.call(arguments,0),n,{time:Date.now()}]),i(n)}};n._s=r,this.csa=n}();
```

In [39]:
```python
1  d = soup.find_all('div',class_='quoteDetails')
2  for i in d:
3      print(i.find('div',class_='quoteText').text.strip())
```

"Be yourself; everyone else is already taken."

    —

    Oscar Wilde
"I'm selfish, impatient and a little insecure. I make mistakes, I am out of c
ontrol and at times hard to handle. But if you can't handle me at my worst, t
hen you sure as hell don't deserve me at my best."

    —

    Marilyn Monroe
"So many books, so little time."

    —

    Frank Zappa
"Two things are infinite: the universe and human stupidity; and I'm not sure
about the universe."

In [46]:
```python
1  t = soup.findAll('span',class_='authorOrTitle')
2  l = []
3  n = []
4  for i in t:
5      a = i.text.strip()
6      l.append(a)
7      if(a in l):
8          n.append(a)
9  n
```

Out[46]: ['Oscar Wilde',
 'Marilyn Monroe',
 'Frank Zappa',
 'Albert Einstein',
 'Marcus Tullius Cicero',
 'Bernard M. Baruch',
 'William W. Purkey',
 'Dr. Seuss',
 'Mae West',
 'Mahatma Gandhi',
 'Robert Frost',
 'J.K. Rowling,',
 'Albert Camus',
 'Mark Twain',
 'C.S. Lewis,',
 'Maya Angelou',
 'Elbert Hubbard',
 'Oscar Wilde',
 'Oscar Wilde',
 'Martin Luther King Jr.,',
 'Mahatma Gandhi',
 'Stephen Chbosky,',
 'Friedrich Nietzsche,',
 'Oscar Wilde,',
 'Ralph Waldo Emerson',
 'Steve Jobs',
 'Narcotics Anonymous',
 'Marilyn Monroe',
 'Andre Gide,',
 'H. Jackson Brown Jr.,']

In [51]:
```python
d = soup.find_all('div',class_='quoteText')
for i in d:
    if("Love" or 'love' in i):
        print(i.text.strip())
```

"Be yourself; everyone else is already taken."

―

       Oscar Wilde
"I'm selfish, impatient and a little insecure. I make mistakes, I am out of c
ontrol and at times hard to handle. But if you can't handle me at my worst, t
hen you sure as hell don't deserve me at my best."

―

       Marilyn Monroe
"So many books, so little time."

―

       Frank Zappa
"Two things are infinite: the universe and human stupidity; and I'm not sure
about the universe."

In [66]:

```python
from bs4 import BeautifulSoup

with open("Popular_Quotes.html", "r", encoding="utf-8") as file:
    soup = BeautifulSoup(file, "html.parser")

quote_blocks = soup.find_all("div", class_="quote")

for block in quote_blocks:
    title_tag = block.find("a", class_="authorOrTitle")
    author_tag = block.find("span", class_="authorOrTitle")

    if title_tag and author_tag:
        title = title_tag.get_text().strip()
        link = title_tag['href'].strip()
        author = author_tag.get_text().strip()

        print(f"Title: {title}")
        print(f"Link: {link}")
        print(f"Author: {author}")
        print("-" * 40)
```

```
Title: Harry Potter and the Goblet of Fire
Link: /work/quotes/3046572
Author: J.K. Rowling,
----------------------------------------
Title: The Four Loves
Link: /work/quotes/14816053
Author: C.S. Lewis,
----------------------------------------
Title: A Testament of Hope: The Essential Writings and Speeches
Link: /work/quotes/52037
Author: Martin Luther King Jr.,
----------------------------------------
Title: The Perks of Being a Wallflower
Link: /work/quotes/2236198
Author: Stephen Chbosky,
----------------------------------------
Title: Twilight of the Idols
Link: /work/quotes/19234046
Author: Friedrich Nietzsche,
----------------------------------------
Title: The Happy Prince and Other Stories
Link: /work/quotes/23684401
Author: Oscar Wilde,
----------------------------------------
Title: Autumn Leaves
Link: /work/quotes/6810684
Author: Andre Gide,
----------------------------------------
Title: P.S. I Love You
Link: /work/quotes/41658
Author: H. Jackson Brown Jr.,
----------------------------------------
```

```
In [ ]:    1    if(i.['href']):
           2
```

In [71]:
```python
quotes = soup.find_all('div', class_='quoteText')

for quote in quotes:
    text = quote.get_text(strip=True)
    if len(text) > 100:
        print(text)
        print("-" * 50)
```

"I'm selfish, impatient and a little insecure. I make mistakes, I am out of con
trol and at times hard to handle. But if you can't handle me at my worst, then
you sure as hell don't deserve me at my best."—Marilyn Monroe
----
"Two things are infinite: the universe and human stupidity; and I'm not sure ab
out the universe."—Albert Einstein

"Be who you are and say what you feel, because those who mind don't matter, and
those who matter don't mind."—Bernard M. Baruch

"You've gotta dance like there's nobody watching,Love like you'll never be hur
t,Sing like there's nobody listening,And live like it's heaven on earth."—Willi
am W. Purkey

"You know you're in love when you can't fall asleep because reality is finally
better than your dreams."—Dr. Seuss

"If you want to know what a man's like, take a good look at how he treats his i
nferiors, not his equals."—J.K. Rowling,Harry Potter and the Goblet of Fire
----
"Don't walk in front of me… I may not followDon't walk behind me… I may not lea
dWalk beside me… just be my friend"—Albert Camus

"Friendship ... is born at the moment when one man says to another "What! You t
oo? I thought that no one but myself . . ."—C.S. Lewis,The Four Loves
----
"I've learned that people will forget what you said, people will forget what yo
u did, but people will never forget how you made them feel."—Maya Angelou
----
"Darkness cannot drive out darkness: only light can do that. Hate cannot drive
out hate: only love can do that."—Martin Luther King Jr.,A Testament of Hope: T
he Essential Writings and Speeches

"I am so clever that sometimes I don't understand a single word of what I am sa
ying."—Oscar Wilde,The Happy Prince and Other Stories

"To be yourself in a world that is constantly trying to make you something else
is the greatest accomplishment."—Ralph Waldo Emerson

"Here's to the crazy ones. The misfits. The rebels. The troublemakers. The roun
d pegs in the square holes. The ones who see things differently. They're not fo
nd of rules. And they have no respect for the status quo. You can quote them, d
isagree with them, glorify or vilify them. About the only thing you can't do is
ignore them. Because they change things. They push the human race forward. And
while some may see them as the crazy ones, we see genius. Because the people wh
o are crazy enough to think they can change the world, are the ones who do."—St
eve Jobs

"Insanity is doing the same thing, over and over again, but expecting different
results."—Narcotics Anonymous

"I believe that everything happens for a reason. People change so that you can
learn to let go, things go wrong so that you appreciate them when they're righ
t, you believe lies so you eventually learn to trust no one but yourself, and s
ometimes good things fall apart so better things can fall together."—Marilyn Mo
nroe

"It is better to be hated for what you are than to be loved for what you are no
t."—Andre Gide,Autumn Leaves

"Twenty years from now you will be more disappointed by the things that you did
n't do than by the ones you did do. So throw off the bowlines. Sail away from t
he safe harbor. Catch the trade winds in your sails. Explore. Dream. Discove
r."—H. Jackson Brown Jr.,P.S. I Love You

In [72]:
```python
quotes = soup.find_all('div', class_='quote')
q_list = []

for q in quotes:
    text = q.find('div', class_='quoteText').get_text(strip=True)
    like_tag = q.find('a', class_='smallText')
    if like_tag:
        likes = int(like_tag.text.strip().split()[0].replace(',', ''))
        q_list.append((text, likes))

# Sort by likes descending
q_list.sort(key=lambda x: x[1], reverse=True)

for t, l in q_list:
    print(f"{l} likes - {t[:100]}...")
```

172090 likes - "Be yourself; everyone else is already taken."—Oscar Wilde...
162993 likes - "I'm selfish, impatient and a little insecure. I make mistakes, I am out of control and at times har...
147069 likes - "So many books, so little time."—Frank Zappa...
146082 likes - "Two things are infinite: the universe and human stupidity; and I'm not sure about the universe."—Al...
133363 likes - "A room without books is like a body without a soul."—Marcus Tul lius Cicero...
128599 likes - "Be who you are and say what you feel, because those who mind do n't matter, and those who matter don...
127116 likes - "You've gotta dance like there's nobody watching,Love like you'l l never be hurt,Sing like there's no...
124987 likes - "You know you're in love when you can't fall asleep because real ity is finally better than your drea...
117240 likes - "You only live once, but if you do it right, once is enough."—Ma e West...
113213 likes - "Be the change that you wish to see in the world."—Mahatma Gandh i...
108002 likes - "In three words I can sum up everything I've learned about life: it goes on."—Robert Frost...
100983 likes - "If you want to know what a man's like, take a good look at how he treats his inferiors, not his equ...
92274 likes - "Don't walk in front of me… I may not followDon't walk behind me… I may not leadWalk beside me… just...
90869 likes - "If you tell the truth, you don't have to remember anything."—Mar k Twain...
87553 likes - "Friendship ... is born at the moment when one man says to anothe r "What! You too? I thought that no...
87398 likes - "I've learned that people will forget what you said, people will forget what you did, but people wil...
83514 likes - "A friend is someone who knows all about you and still loves yo u."—Elbert Hubbard...
83168 likes - "To live is the rarest thing in the world. Most people exist, tha t is all."—Oscar Wilde...
80923 likes - "Always forgive your enemies; nothing annoys them so much."—Oscar Wilde...
77451 likes - "Darkness cannot drive out darkness: only light can do that. Hate cannot drive out hate: only love c...
77061 likes - "Live as if you were to die tomorrow. Learn as if you were to liv e forever."—Mahatma Gandhi...
74875 likes - "We accept the love we think we deserve."—Stephen Chbosky,The Per ks of Being a Wallflower...
73251 likes - "Without music, life would be a mistake."—Friedrich Nietzsche,Twi light of the Idols...
71126 likes - "I am so clever that sometimes I don't understand a single word o f what I am saying."—Oscar Wilde,Th...
69009 likes - "To be yourself in a world that is constantly trying to make you something else is the greatest acco...
67728 likes - "Here's to the crazy ones. The misfits. The rebels. The troublema kers. The round pegs in the square ...
64173 likes - "Insanity is doing the same thing, over and over again, but expec ting different results."—Narcotics ...
61795 likes - "I believe that everything happens for a reason. People change so that you can learn to let go, thin...
61028 likes - "It is better to be hated for what you are than to be loved for w hat you are not."—Andre Gide,Autumn...

59792 likes - "Twenty years from now you will be more disappointed by the thing
s that you didn't do than by the on...

In [3]:
```python
import pandas as pd

quotes = soup.find_all('div', class_='quote')
data = []

for q in quotes:
    text = q.find('div', class_='quoteText').get_text(strip=True)
    author = q.find('span', class_='authorOrTitle').get_text(strip=True)
    data.append([author, text])

df = pd.DataFrame(data, columns=['Author', 'Quote'])
print(df.head())
```

```
                    Author                                              Quote
0              Oscar Wilde  "Be yourself; everyone else is already taken."...
1           Marilyn Monroe  "I'm selfish, impatient and a little insecure....
2              Frank Zappa       "So many books, so little time."—Frank Zappa
3          Albert Einstein  "Two things are infinite: the universe and hum...
4   Marcus Tullius Cicero  "A room without books is like a body without a...
```

In [ ]:
```python

```