```
In [18]: pip install newspaper3k
```

Requirement already satisfied: newspaper3k in c:\users\connie\anaconda3\lib\s
ite-packages (0.2.8)
Requirement already satisfied: requests>=2.10.0 in c:\users\connie\anaconda3
\lib\site-packages (from newspaper3k) (2.22.0)
Requirement already satisfied: feedparser>=5.2.1 in c:\users\connie\anaconda3
\lib\site-packages (from newspaper3k) (5.2.1)
Requirement already satisfied: lxml>=3.6.0 in c:\users\connie\anaconda3\lib\s
ite-packages (from newspaper3k) (4.5.0)
Requirement already satisfied: Pillow>=3.3.0 in c:\users\connie\anaconda3\lib
\site-packages (from newspaper3k) (7.0.0)
Requirement already satisfied: PyYAML>=3.11 in c:\users\connie\anaconda3\lib
\site-packages (from newspaper3k) (5.3)
Requirement already satisfied: tldextract>=2.0.1 in c:\users\connie\anaconda3
\lib\site-packages (from newspaper3k) (2.2.2)
Requirement already satisfied: feedfinder2>=0.0.4 in c:\users\connie\anaconda
3\lib\site-packages (from newspaper3k) (0.0.4)
Requirement already satisfied: tinysegmenter==0.3 in c:\users\connie\anaconda
3\lib\site-packages (from newspaper3k) (0.3)
Requirement already satisfied: jieba3k>=0.35.1 in c:\users\connie\anaconda3\l
ib\site-packages (from newspaper3k) (0.35.1)
Requirement already satisfied: cssselect>=0.9.2 in c:\users\connie\anaconda3
\lib\site-packages (from newspaper3k) (1.1.0)
Requirement already satisfied: beautifulsoup4>=4.4.1 in c:\users\connie\anaco
nda3\lib\site-packages (from newspaper3k) (4.8.2)
Requirement already satisfied: nltk>=3.2.1 in c:\users\connie\anaconda3\lib\s
ite-packages (from newspaper3k) (3.4.5)
Requirement already satisfied: python-dateutil>=2.5.3 in c:\users\connie\anac
onda3\lib\site-packages (from newspaper3k) (2.8.1)
Requirement already satisfied: chardet<3.1.0,>=3.0.2 in c:\users\connie\anaco
nda3\lib\site-packages (from requests>=2.10.0->newspaper3k) (3.0.4)
Requirement already satisfied: certifi>=2017.4.17 in c:\users\connie\anaconda
3\lib\site-packages (from requests>=2.10.0->newspaper3k) (2019.11.28)
Requirement already satisfied: idna<2.9,>=2.5 in c:\users\connie\anaconda3\li
b\site-packages (from requests>=2.10.0->newspaper3k) (2.8)
Requirement already satisfied: urllib3!=1.25.0,!=1.25.1,<1.26,>=1.21.1 in
c:\users\connie\anaconda3\lib\site-packages (from requests>=2.10.0->newspaper
3k) (1.25.8)
Requirement already satisfied: requests-file>=1.4 in c:\users\connie\anaconda
3\lib\site-packages (from tldextract>=2.0.1->newspaper3k) (1.5.1)
Requirement already satisfied: setuptools in c:\users\connie\anaconda3\lib\si
te-packages (from tldextract>=2.0.1->newspaper3k) (45.2.0.post20200210)
Requirement already satisfied: six in c:\users\connie\anaconda3\lib\site-pack
ages (from feedfinder2>=0.0.4->newspaper3k) (1.14.0)
Requirement already satisfied: soupsieve>=1.2 in c:\users\connie\anaconda3\li
b\site-packages (from beautifulsoup4>=4.4.1->newspaper3k) (1.9.5)
Note: you may need to restart the kernel to use updated packages.

```
In [19]: import nltk
         from newspaper import Article
```

```
In [20]:  url = 'https://www.inc.com/emily-canal/national-coffee-day-startup-funding.htm
          l'
          article = Article(url)
          article.download()
          article.parse()
          nltk.download('punkt')
          article.nlp()
```

```
[nltk_data] Downloading package punkt to
[nltk_data]     C:\Users\Connie\AppData\Roaming\nltk_data...
[nltk_data]   Package punkt is already up-to-date!
```

```
In [21]:  article.publish_date
```

Out[21]: datetime.datetime(2018, 9, 28, 5, 0)

```
In [38]:  article.top_image
```

Out[38]: 'https://www.incimages.com/uploaded_files/image/1024x576/shutterstock_7168884
          58_371272.jpg'

```
In [40]:  #get article text
          print(article.text)
```

The nation officially marks its love affair with caffeine on September 29 eve
ry year. This year, investors' infatuation with coffee startups will reach a
new peak.

```
In [42]:  #article Keywords
          article.keywords
```

Out[42]: ['record',
           'investors',
           'heres',
           'infatuation',
           'officially',
           'pouring',
           'venture',
           'nation',
           'peak',
           'love',
           'marks',
           'reach',
           'coffee',
           'startups',
           'capitalists',
           'billion']

```
In [46]:  print(article.summary)
```

The nation officially marks its love affair with caffeine on September 29 eve
ry year.
This year, investors' infatuation with coffee startups will reach a new peak.

```
In [56]:  #Law 360
          url = 'https://www.law360.com/articles/1281804'
          article = Article(url)
          article.download()
          article.parse()
          nltk.download('punkt')
          article.nlp()
```

```
[nltk_data] Downloading package punkt to
[nltk_data]     C:\Users\Connie\AppData\Roaming\nltk_data...
[nltk_data]   Package punkt is already up-to-date!
```

```
In [59]:  print(article.summary)
```

Corrigan said he has heard from law firm managers who have offices in high-ri
ses that are "pulling their hair out"over how such restrictions will complica
te elevator use in buildings with many floors and many people.
The law firm is headquartered in Ohio and, according to managing partner Scot
t Kadish, began reopening its three Ohio offices June 1.
Chicago, New York and Washington, D.C., will take longer.The firm's Cincinnat
i office typically had around 100 people arriving for work each day before th
e coronavirus hit.
In Ohio, especially in Cincinnati and Columbus, the vast majority of employee
s travel to work by car, Kadish said.
But as more people head back to work, that will be "a little more challengin
g," he said.Traversing elevators is also a concern at the firm's offices in l
arger buildings.

```
In [64]:  article.keywords
```

```
Out[64]:  ['confronting',
           'offices',
           'employees',
           'firm',
           'office',
           'work',
           'obstacles',
           'kadish',
           'chicago',
           'reopening',
           'firms',
           'law']
```