```
@author: michellekim
"""

# check count of unique words
from urllib.request import urlopen
from bs4 import BeautifulSoup

def array_expander(array):
    big_array = [0,0] * len(array)
    count = 0
    for x in array:
        big_array[count] = array[count]
        count = count + 1
    return big_array


url = "https://education.nationalgeographic.org/resource/air-pollution/"
html = urlopen(url).read() # returns a sequence of bytes
soup = BeautifulSoup(html, "html.parser") # assign html as name, pass to python's html parser

# kill all script and style elements
for script in soup(["script", "style"]):
    script.extract()    # rip it out

# get human-readable text inside a document or tag
text = soup.get_text()

# break into lines and remove leading and trailing space on each
# .splitlines() Splits the string at line breaks such as '\n'
# .strip() removes whitespace
lines = (line.strip() for line in text.splitlines())

# break multi-headlines into a line each
chunks = (phrase.strip() for line in lines for phrase in line.split("  "))

# drop blank lines
text = '\n'.join(chunk for chunk in chunks if chunk)

print(text)

words = [[0,0]]

to_words = text.split() # text split into words


for x in to_words: # going through each word of the text
    for y in words: # going through each element of previously seen words array
        if y[0] == x:
            y[1] = y[1] + 1
            break
    words.append([x,1])

print(words)
```