

Jory Fhel R. Benedicto  
BSIT - 3

Question: **What makes spinach the healthiest vegetables in the world?**

## DATA CLEANING

### Vegetables Dataframe (CORPUS)

Out[21]:

	transcript	full_names
<b>beets</b>	People often describe beetroot as a superfood and have used it for centuries to treat fever, constipation, and skin complaints. Researchers are no...	Beetroot
<b>broccoli</b>	Broccoli has a reputation as a superfood. It is low in calories but contains a wealth of nutrients and antioxidants that support many aspects of h...	Brassica oleracea var. italica
<b>cabbage</b>	Cabbage, which is often lumped into the same category as lettuce because of their similar appearance, is actually a part of the cruciferous vegeta...	Brassica oleracea var. capitata
<b>carrots</b>	We include products we think are useful for our readers. If you buy through links on this page, we may earn a small commission. Here's our process...	Daucus carota subsp. sativus
<b>cauliflower</b>	Cauliflower is a cruciferous vegetable that is naturally high in fiber and B-vitamins. It provides antioxidants and phytonutrients that can protec...	Brassica oleracea var. botrytis
<b>garlic</b>	Garlic ( <i>Allium sativum</i> ), is used widely as a flavoring in cooking, but it has also been used as a medicine throughout ancient and modern history; ...	<i>Allium sativum</i>
<b>kale</b>	We include products we think are useful for our readers. If you buy through links on this page, we may earn a small commission. Here's our process...	Kale
<b>onions</b>	We include products we think are useful for our readers. If you buy through links on this page, we may earn a small commission. Here's our process...	<i>Allium cepa</i>
<b>peas</b>	A recent review and meta-analysis focus on the role of legumes in heart health. Taking data from multiple studies and earlier analyses, the author...	<i>Pisum sativum</i>
<b>seaweed</b>	Seaweed grows in or near salty waters. There are several types, and they generally contain many healthful minerals that are easy for the body to b...	Seaweed
<b>spinach</b>	Spinach is a superfood. It is loaded with tons of nutrients in a low-calorie package. Dark, leafy greens like spinach are important for skin, hair...	<i>Spinacia oleracea</i>
<b>sprouts</b>	Brussels sprouts are a staple of Christmas and Thanksgiving dinner. However, including more of these low calorie, nutritionally dense vegetables i...	Sprouts
<b>sweet potatoes</b>	We include products we think are useful for our readers. If you buy through links on this page, we may earn a small commission. Here's our process...	<i>Ipomoea batatas</i>
<b>tomato</b>	We include products we think are useful for our readers. If you buy through links on this page, we may earn a small commission. Here's our process...	<i>Solanum lycopersicum</i>

## DOCUMENT-TERM MATRIX

```
In [24]: # We are going to create a document-term matrix using CountVectorizer, and exclude common English stop words
from sklearn.feature_extraction.text import CountVectorizer

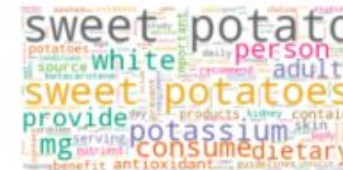
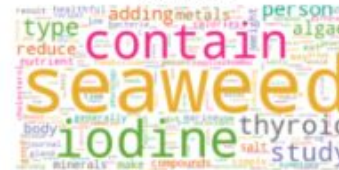
cv = CountVectorizer(stop_words='english')
data_cv = cv.fit_transform(data_clean.transcript)
data_dtm = pd.DataFrame(data_cv.toarray(), columns=cv.get_feature_names())
data_dtm.index = data_clean.index
data_dtm
```

Out[24]:

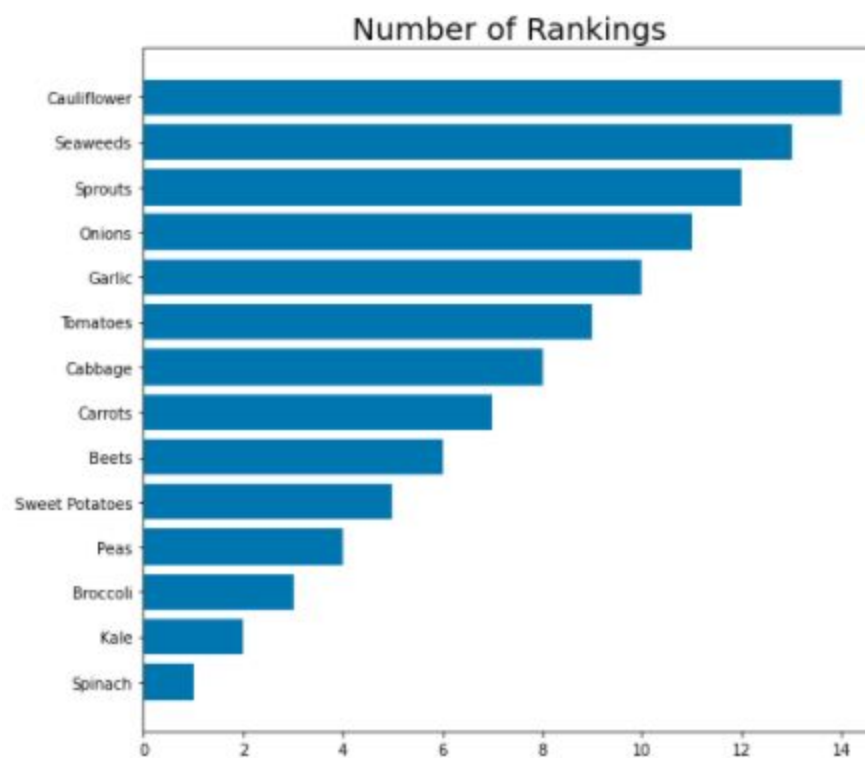
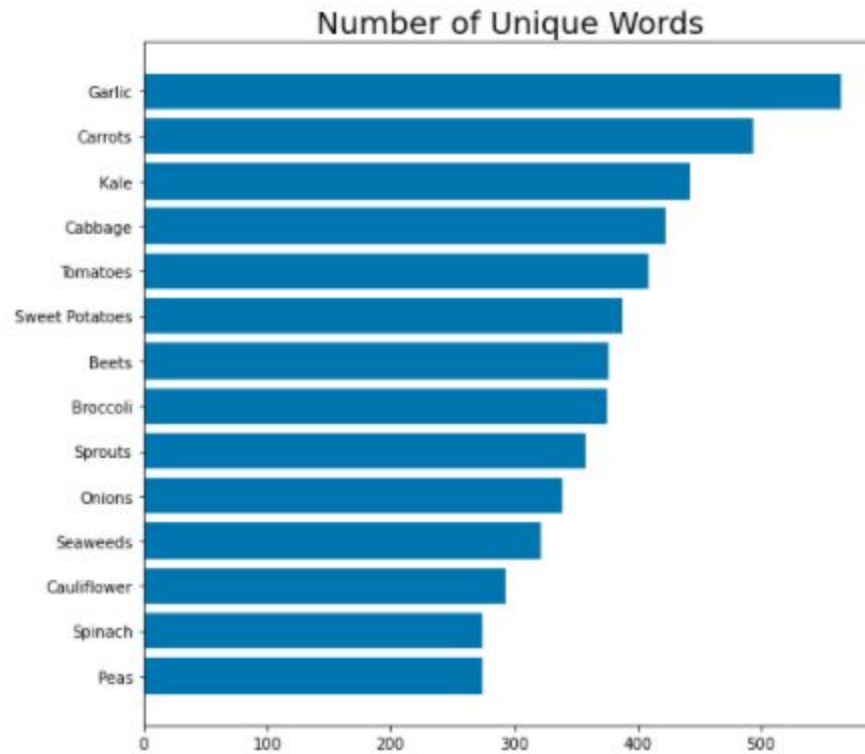
	ability	able	abnormally	absorb	absorbed	absorbing	absorbs	absorption	academic	accompaniment	...	yams	year	yearround	year
beets	2	0	0	0	0	0	0	0	0	1	...	0	0	0	0
broccoli	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0
cabbage	2	2	0	1	0	0	0	0	0	0	...	0	0	0	0
carrots	0	0	0	2	0	0	1	0	0	0	...	0	2	0	0
cauliflower	0	0	0	0	0	0	0	2	0	0	...	0	0	0	0
garlic	0	0	1	0	0	0	0	0	1	0	...	0	1	0	0
kale	0	0	0	1	0	2	0	0	0	0	...	0	1	0	0
onions	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0
peas	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0
seaweed	0	1	0	1	0	0	1	0	0	0	...	0	0	0	0
spinach	0	0	0	0	1	0	0	2	0	0	...	0	0	0	0
sprouts	1	1	0	0	0	0	0	0	0	0	...	0	1	1	1
sweet potatoes	0	0	0	0	0	0	0	1	0	0	...	1	1	0	0
tomato	0	0	0	0	0	0	0	0	0	0	...	0	1	0	0

14 rows × 2296 columns

## Word Cloud



## Plotting the Unique Words and the Vegetable rankings



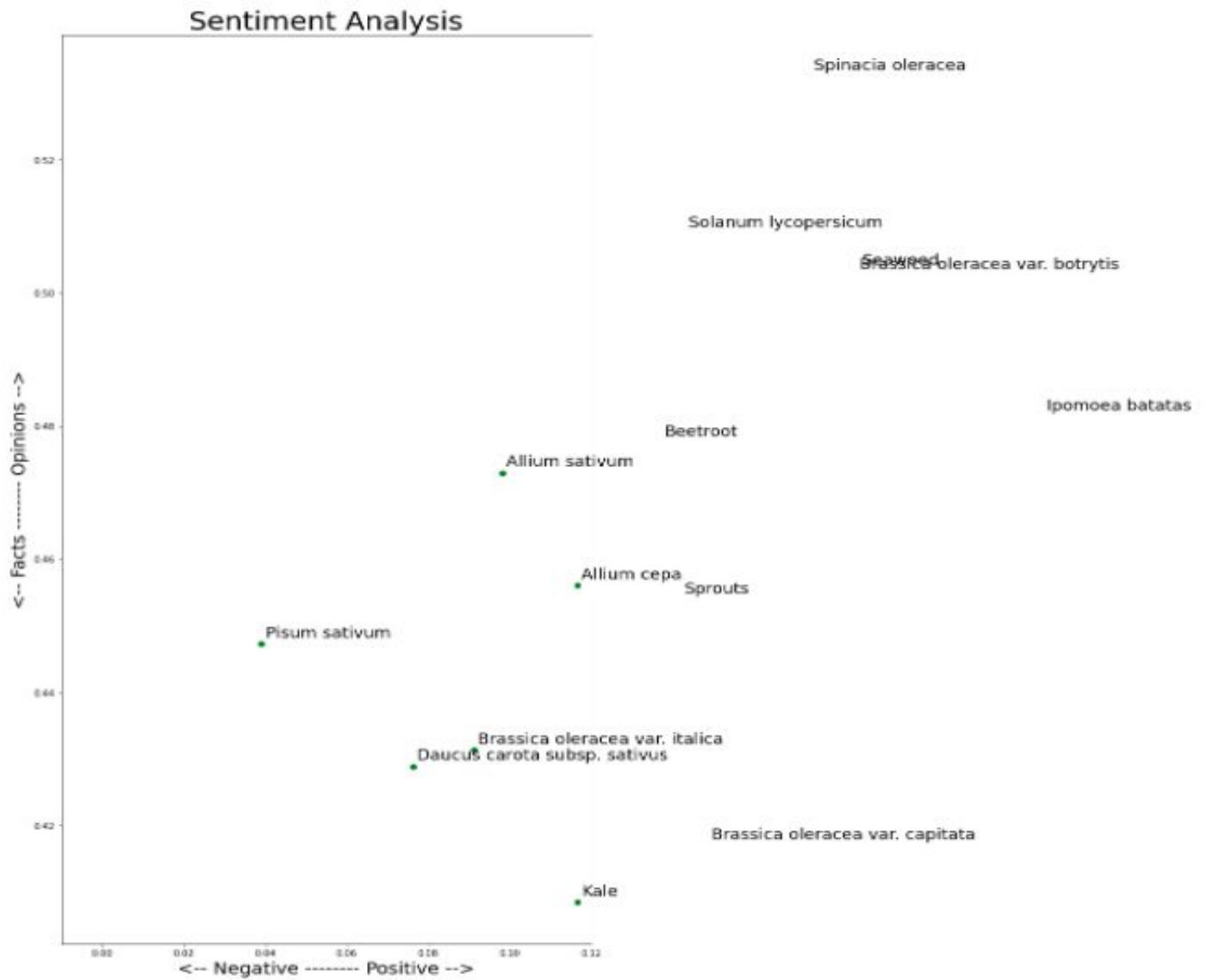
## SENTIMENT ANALYSIS

### Plotting The Polarity and Subjectivity

Out[3]:

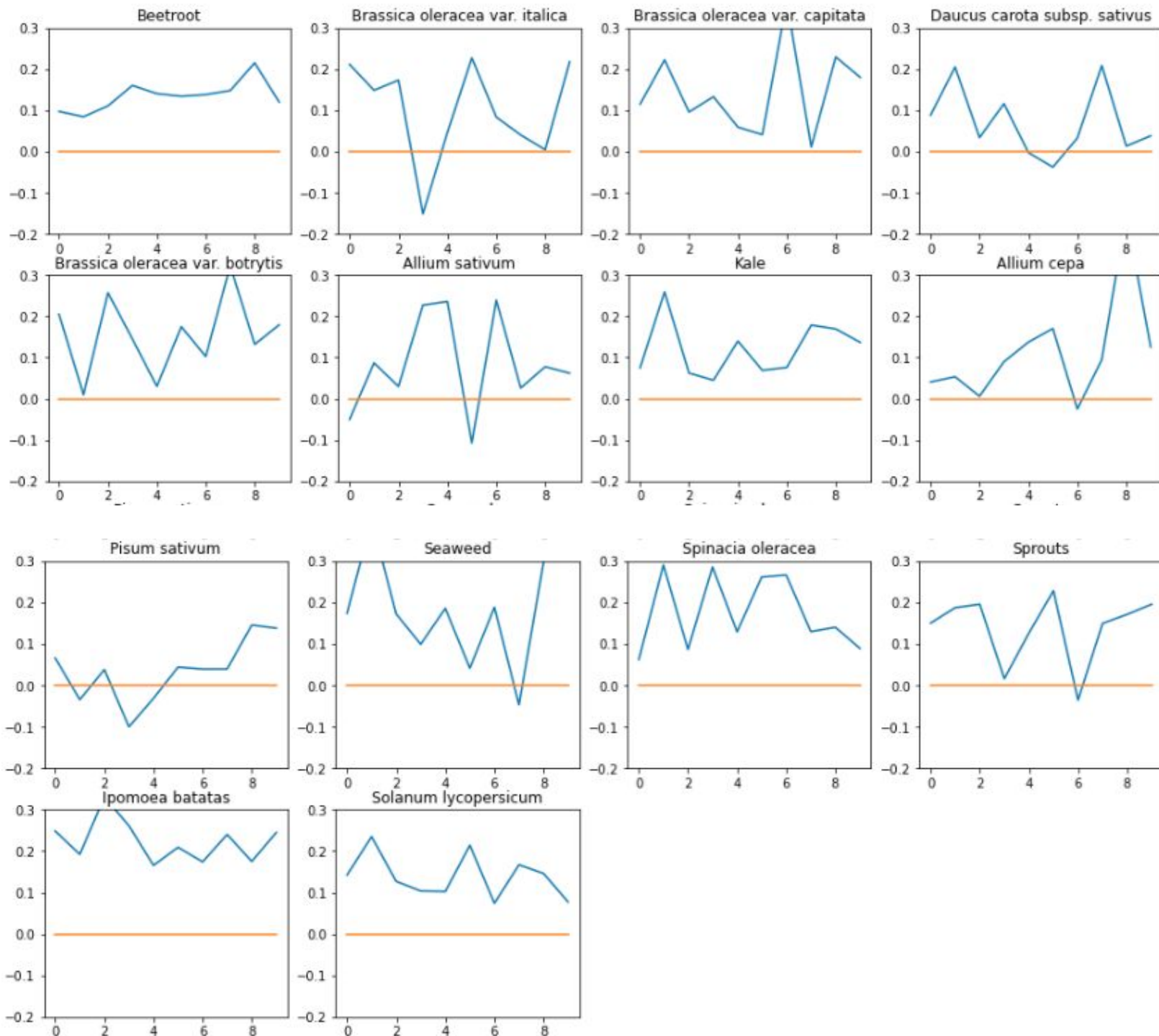
	transcript	full_names	polarity	subjectivity
<b>beets</b>	People often describe beetroot as a superfood ...	Beetroot	0.136863	0.477575
<b>broccoli</b>	Broccoli has a reputation as a superfood. It i...	Brassica oleracea var. italica	0.091176	0.431289
<b>cabbage</b>	Cabbage, which is often lumped into the same c...	Brassica oleracea var. capitata	0.148545	0.416935
<b>carrots</b>	We include products we think are useful for ou...	Daucus carota subsp. sativus	0.076316	0.428896
<b>cauliflower</b>	Cauliflower is a cruciferous vegetable that is...	Brassica oleracea var. botrytis	0.184764	0.502676
<b>garlic</b>	Garlic ( <i>Allium sativum</i> ), is used widely as a f...	Allium sativum	0.098121	0.472986
<b>kale</b>	We include products we think are useful for ou...	Kale	0.116494	0.408436
<b>onions</b>	We include products we think are useful for ou...	Allium cepa	0.116492	0.456074
<b>peas</b>	A recent review and meta-analysis focus on the...	Pisum sativum	0.039132	0.447280
<b>seaweed</b>	Seaweed grows in or near salty waters. There a...	Seaweed	0.185491	0.503219
<b>spinach</b>	Spinach is a superfood. It is loaded with tons...	Spinacia oleracea	0.173506	0.532588
<b>sprouts</b>	Brussels sprouts are a staple of Christmas and...	Sprouts	0.141923	0.453961
<b>sweet potatoes</b>	We include products we think are useful for ou...	Ipomoea batatas	0.230795	0.481400
<b>tomato</b>	We include products we think are useful for ou...	Solanum lycopersicum	0.142864	0.508868

## Plotting The Results





## Plot Of All Vegetables



All of these results, creating a natural language processing using sentiment analysis shows that the vegetable Spinach is the most healthiest vegetable in the world.

**Reference/Web Page links:**

[https://www.medicalnewstoday.com/articles/270609'](https://www.medicalnewstoday.com/articles/270609),  
[https://www.medicalnewstoday.com/articles/270435'](https://www.medicalnewstoday.com/articles/270435),  
[https://www.medicalnewstoday.com/articles/266765'](https://www.medicalnewstoday.com/articles/266765),  
[https://www.medicalnewstoday.com/articles/327090'](https://www.medicalnewstoday.com/articles/327090),  
[https://www.medicalnewstoday.com/articles/281438'](https://www.medicalnewstoday.com/articles/281438),  
[https://www.medicalnewstoday.com/articles/311343'](https://www.medicalnewstoday.com/articles/311343),  
[https://www.medicalnewstoday.com/articles/270191'](https://www.medicalnewstoday.com/articles/270191),  
[https://www.medicalnewstoday.com/articles/284823#1'](https://www.medicalnewstoday.com/articles/284823#1),  
[https://www.medicalnewstoday.com/articles/273031'](https://www.medicalnewstoday.com/articles/273031),  
[https://www.medicalnewstoday.com/articles/265853'](https://www.medicalnewstoday.com/articles/265853),  
[https://www.medicalnewstoday.com/articles/276714'](https://www.medicalnewstoday.com/articles/276714),  
[https://www.medicalnewstoday.com/articles/284765'](https://www.medicalnewstoday.com/articles/284765),  
[https://www.medicalnewstoday.com/articles/323916'](https://www.medicalnewstoday.com/articles/323916),  
[https://www.medicalnewstoday.com/articles/282844'](https://www.medicalnewstoday.com/articles/282844)