

NAME: Sai Lokesh Tippabhotla

The image shows a Google Colab notebook interface with two visible code cells. The first cell contains Python code to load a CSV file and display its contents. The second cell shows the installation of NLTK and the download of the 'punkt' tokenizer data.

**Code Cell 1:**

```
[3] import pandas as pd
import matplotlib.pyplot as plt

data=pd.read_csv(r"/content/Corona_NLP_test.csv",encoding= 'unicode_escape')

[4] #extract required column
OriginalTweet= data["OriginalTweet"]

[5] #write to new csv
data_new=OriginalTweet.to_csv(r"C:\Users\sailo\Downloads\data_new.csv",index=False)

[6] data_new=pd.read_csv(r"C:\Users\sailo\Downloads\data_new.csv",encoding= 'unicode_escape')

data_new
```

The output of the first cell shows a preview of the 'OriginalTweet' column with 3793 rows. The visible text includes:

- 0 TRENDING: New Yorkers encounter empty supermar...
- 1 When I couldn't find hand sanitizer at Fred Me...
- 2 Find out how you can protect yourself and love...
- 3 #Panic buying hits #NewYork City as anxious sh...
- 4 #toiletpaper #dunnypaper #coronavirus #coronav...
- ...
- 3793 Meanwhile In A Supermarket in Israel – People...

**Code Cell 2:**

```
[7] 3795 Aast Prof of Economics @conces was on @NBCPhl...
3796 Gov need to do somethings instead of blar je r...
3797 I and @ForestandPaper members are committed to...
3798 rows x 1 columns

[8] data_list=data_new['OriginalTweet'].values.tolist()

A)Convert the text corpus into tokens.

[9] pip install nltk

Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-wheels/public/simple/
Requirement already satisfied: nltk in /usr/local/lib/python3.10/dist-packages (3.8.1)
Requirement already satisfied: click in /usr/local/lib/python3.10/dist-packages (from nltk) (8.1.3)
Requirement already satisfied: tqdm in /usr/local/lib/python3.10/dist-packages (from nltk) (4.65.0)
Requirement already satisfied: joblib in /usr/local/lib/python3.10/dist-packages (from nltk) (1.2.0)
Requirement already satisfied: regex<=2021.8.3 in /usr/local/lib/python3.10/dist-packages (from nltk) (2022.10.31)

[10] import nltk
nltk.download('punkt')
from nltk.tokenize import word_tokenize

[nltk_data] Downloading package punkt to /root/nltk_data...
[nltk_data] Unzipping tokenizers/punkt.zip.

[11] tokens = []
```

The notebook interface includes a file explorer on the left showing 'sample\_data' and 'Corona\_NLP\_test.csv'. The bottom status bar indicates 84.25 GB available disk space and completion times of 10s.





Assignment 3 tweets.ipynb - Colaboratory

colab.research.google.com/drive/1a2bvzNoEPqAwGy4cvhh9PbcV6xslreQ#scrollTo=ce355f17

File Edit View Insert Runtime Tools Help All changes saved

Files

- sample\_data
- C:\Users\sailo\Downloads\data\_n...
- Corona\_NLP\_test.csv

```
[22] "foodrationing": 1,
      "wave": 6,
      "fuelled": 2,
      "Akghxox": 1,

[23] d = dict((k, v) for k, v in word_freq.items() if v >= 100)
      d
      len(d)

82

D)Create word clouds.

[24] pip install wordcloud

Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-wheels/public/simple/
Requirement already satisfied: wordcloud in /usr/local/lib/python3.10/dist-packages (1.8.2.2)
Requirement already satisfied: numpy>=1.6.1 in /usr/local/lib/python3.10/dist-packages (from wordcloud) (1.22.4)
Requirement already satisfied: matplotlib in /usr/local/lib/python3.10/dist-packages (from wordcloud) (3.7.1)
Requirement already satisfied: pillow in /usr/local/lib/python3.10/dist-packages (from wordcloud) (8.4.0)
Requirement already satisfied: packaging>=20.0 in /usr/local/lib/python3.10/dist-packages (from matplotlib->wordcloud) (23.1)
Requirement already satisfied: pyparsing>=2.3.1 in /usr/local/lib/python3.10/dist-packages (from matplotlib->wordcloud) (3.0.9)
Requirement already satisfied: python-dateutil>=2.7 in /usr/local/lib/python3.10/dist-packages (from matplotlib->wordcloud) (2.8.2)
Requirement already satisfied: kiwisolver>=1.0.1 in /usr/local/lib/python3.10/dist-packages (from matplotlib->wordcloud) (1.4.4)
Requirement already satisfied: fonttools>=4.22.0 in /usr/local/lib/python3.10/dist-packages (from matplotlib->wordcloud) (4.39.3)
Requirement already satisfied: cycler>=0.10 in /usr/local/lib/python3.10/dist-packages (from matplotlib->wordcloud) (0.11.0)
Requirement already satisfied: contourpy>=1.0.1 in /usr/local/lib/python3.10/dist-packages (from matplotlib->wordcloud) (1.0.7)
Requirement already satisfied: six>=1.5 in /usr/local/lib/python3.10/dist-packages (from python-dateutil->2.7->matplotlib->wordcloud) (1.16.0)

[25] all_words = ' '.join(delete_empty)
```

10s completed at 11:06 PM

71°F Clear

Assignment 3 tweets.ipynb - Colaboratory

colab.research.google.com/drive/1a2bvzNoEPqAwGy4cvhh9PbcV6xslreQ#scrollTo=ce355f17

File Edit View Insert Runtime Tools Help All changes saved

Files

- sample\_data
- C:\Users\sailo\Downloads\data\_n...
- Corona\_NLP\_test.csv

```
D)Create word clouds.

[24] pip install wordcloud

Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-wheels/public/simple/
Requirement already satisfied: wordcloud in /usr/local/lib/python3.10/dist-packages (1.8.2.2)
Requirement already satisfied: numpy>=1.6.1 in /usr/local/lib/python3.10/dist-packages (from wordcloud) (1.22.4)
Requirement already satisfied: matplotlib in /usr/local/lib/python3.10/dist-packages (from wordcloud) (3.7.1)
Requirement already satisfied: pillow in /usr/local/lib/python3.10/dist-packages (from wordcloud) (8.4.0)
Requirement already satisfied: packaging>=20.0 in /usr/local/lib/python3.10/dist-packages (from matplotlib->wordcloud) (23.1)
Requirement already satisfied: pyparsing>=2.3.1 in /usr/local/lib/python3.10/dist-packages (from matplotlib->wordcloud) (3.0.9)
Requirement already satisfied: python-dateutil>=2.7 in /usr/local/lib/python3.10/dist-packages (from matplotlib->wordcloud) (2.8.2)
Requirement already satisfied: kiwisolver>=1.0.1 in /usr/local/lib/python3.10/dist-packages (from matplotlib->wordcloud) (1.4.4)
Requirement already satisfied: fonttools>=4.22.0 in /usr/local/lib/python3.10/dist-packages (from matplotlib->wordcloud) (4.39.3)
Requirement already satisfied: cycler>=0.10 in /usr/local/lib/python3.10/dist-packages (from matplotlib->wordcloud) (0.11.0)
Requirement already satisfied: contourpy>=1.0.1 in /usr/local/lib/python3.10/dist-packages (from matplotlib->wordcloud) (1.0.7)
Requirement already satisfied: six>=1.5 in /usr/local/lib/python3.10/dist-packages (from python-dateutil->2.7->matplotlib->wordcloud) (1.16.0)

all_words = ' '.join(delete_empty)

from wordcloud import WordCloud

wordcloud = WordCloud(max_words=400,width=800, height=500, colormap='dark2').generate(all_words)
plt.figure(figsize=(12, 8))
plt.imshow(wordcloud, interpolation='bilinear')
plt.axis('off')
plt.show()
```

10s completed at 11:06 PM

71°F Clear



