

```
In [15]: import pandas as pd
import numpy as np
df1 = pd.read_csv("posttext.csv")
df1.head()
```

```
Out[15]:
```

| | post_date | Post_Text | Post_Embedded_URL | Post_URL |
|---|-----------|---|----------------------------------|---|
| 0 | 22-Jul-20 | Dear Shankra Family, \n first of all, we want ... | http://bit.ly/ShankraTickets2021 | https://m.facebook.com/story.php?story_fbid=30... |
| 1 | 30-Oct-20 | Walking towards the future with the strongest ... | b.link/ShankraTickets2021 | https://m.facebook.com/story.php?story_fbid=33... |
| 2 | 27-Oct-20 | The Mesolcina Mountain Chain, home to Shankra ... | b.link/ShankraTickets2021 | https://m.facebook.com/story.php?story_fbid=33... |
| 3 | 24-Oct-20 | We believe in pure music, freely expressed by ... | b.link/ShankraTickets2021 | https://m.facebook.com/story.php?story_fbid=33... |
| 4 | 21-Oct-20 | The beating heart of our Valley, Shankra Stage... | b.link/ShankraTickets2021 | https://m.facebook.com/story.php?story_fbid=33... |

```
In [16]: df2 = df1.drop(['Post_URL', 'Post_Embedded_URL', 'post_date'], axis = 1)
df2.head()
```

```
Out[16]:
```

| | Post_Text |
|---|---|
| 0 | Dear Shankra Family, \n first of all, we want ... |
| 1 | Walking towards the future with the strongest ... |
| 2 | The Mesolcina Mountain Chain, home to Shankra ... |
| 3 | We believe in pure music, freely expressed by ... |
| 4 | The beating heart of our Valley, Shankra Stage... |

```
In [19]: import nltk
nltk.download('punkt')
nltk.download('wordnet')
from nltk.tokenize import word_tokenize
from nltk.corpus import stopwords
```

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

```
In [20]: df2['Post_Text'].apply(word_tokenize)
```

```
Out[20]: 0    [Dear, Shankra, Family, ,, first, of, all, ,, ...
1    [Walking, towards, the, future, with, the, str...
2    [The, Mesolcina, Mountain, Chain, ,, home, to,...
3    [We, believe, in, pure, music, ,, freely, expr...
4    [The, beating, heart, of, our, Valley,,, Shan...
5    [5, days, of, good, music, and, a, great, comm...
```

```

6   [What, ', s, that, special, song, that, makes,...
7   [Where, do, you, like, to, gather, as, a, comm...
8   [Music, unites, us, ,, connecting, our, indivi...
9   [An, open, mind, towards, music, creates, a, f...
10  [An, open, mind, towards, music, creates, a, f...
11  [It, starts, with, a, simple, action, ,, a, de...
12  [By, challenging, ourselves, and, keeping, our...
13  [Through, memories, ,, every, moment, of, our,...
14  [Pick, colorful, flowers, and, dance, with, yo...
15  [You, can, write, poetry, ,, but, you, can, al...
16  [Shankra, is, a, shared, dream, :, we, build, ...
17  [From, the, smallest, cell, of, our, skin, to,...
18  [Tingling, body, ,, increased, heart, rate, ,,...
19  [Music, speaks, directly, to, our, body, ,, gi...
20  [We, are, born, as, individuals, ,, sharing, t...
21  [Life, expresses, its, energy, through, change...
22  [Dear, Shankra, Family, ,, We, want, to, intro...
23  [Dear, Shankra, Family, ,, We, started, this, ...
24  [Dear, Shankra, Family, ,, as, we, are, facing...
25  [Dear, Shankra, Family, ,, we, are, delighted,...
26  [Dear, Shankra, Family, ,, we, are, delighted,...
27  [Dear, Shankra, Family, ,, we, are, delighted,...
Name: Post_Text, dtype: object

```

```
In [21]: nltk.download('stopwords');
```

```

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

```

```
In [22]: from nltk.corpus import stopwords
stop_words = stopwords.words('english')
print(stop_words)
```

```

['i', 'me', 'my', 'myself', 'we', 'our', 'ours', 'ourselves', 'you', "you're", "you've",
"you'll", "you'd", 'your', 'yours', 'yourself', 'yourselves', 'he', 'him', 'his', 'himse
lf', 'she', "she's", 'her', 'hers', 'herself', 'it', "it's", 'its', 'itself', 'they', 't
hem', 'their', 'theirs', 'themselves', 'what', 'which', 'who', 'whom', 'this', 'that',
"that'll", 'these', 'those', 'am', 'is', 'are', 'was', 'were', 'be', 'been', 'being', 'h
ave', 'has', 'had', 'having', 'do', 'does', 'did', 'doing', 'a', 'an', 'the', 'and', 'bu
t', 'if', 'or', 'because', 'as', 'until', 'while', 'of', 'at', 'by', 'for', 'with', 'abo
ut', 'against', 'between', 'into', 'through', 'during', 'before', 'after', 'above', 'bel
ow', 'to', 'from', 'up', 'down', 'in', 'out', 'on', 'off', 'over', 'under', 'again', 'fu
rther', 'then', 'once', 'here', 'there', 'when', 'where', 'why', 'how', 'all', 'any', 'b
oth', 'each', 'few', 'more', 'most', 'other', 'some', 'such', 'no', 'nor', 'not', 'onl
y', 'own', 'same', 'so', 'than', 'too', 'very', 's', 't', 'can', 'will', 'just', 'don',
"don't", 'should', "should've", 'now', 'd', 'll', 'm', 'o', 're', 've', 'y', 'ain', 'are
n', "aren't", 'couldn', "couldn't", 'didn', "didn't", 'doesn', "doesn't", 'hadn', "had
n't", 'hasn', "hasn't", 'haven', "haven't", 'isn', "isn't", 'ma', 'mightn', "mightn't",
'mustn', "mustn't", 'needn', "needn't", 'shan', "shan't", 'shouldn', "shouldn't", 'was
n', "wasn't", 'weren', "weren't", 'won', "won't", 'wouldn', "wouldn't"]

```

```
In [29]: import io
from nltk.corpus import stopwords
from nltk.tokenize import word_tokenize

# word_tokenize accepts
# a string as an input, not a file.
stop_words = set(stopwords.words('english'))
file1 = open("posttext.csv")

# Use this to read file content as a stream:
line = file1.read()
words = line.split()
```

```

for r in words:
    if not r in stop_words:
        appendFile = open('posttext.csv','a')
        appendFile.write(" "+r)
        appendFile.close()
df2.head(33)

```

Out[29]:

Post_Text

| | Post_Text |
|----|---|
| 0 | Dear Shankra Family, \n first of all, we want ... |
| 1 | Walking towards the future with the strongest ... |
| 2 | The Mesolcina Mountain Chain, home to Shankra ... |
| 3 | We believe in pure music, freely expressed by ... |
| 4 | The beating heart of our Valley, Shankra Stage... |
| 5 | 5 days of good music and a great community. Ar... |
| 6 | What's that special song that makes you dance?... |
| 7 | Where do you like to gather as a community?\nP... |
| 8 | Music unites us, connecting our individual fee... |
| 9 | An open mind towards music creates a feeling o... |
| 10 | An open mind towards music creates a feeling o... |
| 11 | It starts with a simple action, a desire, and ... |
| 12 | By challenging ourselves and keeping our curio... |
| 13 | Through memories, every moment of our life can... |
| 14 | Pick colorful flowers and dance with your chil... |
| 15 | You can write poetry, but you can also act poe... |
| 16 | Shankra is a shared dream: we build each editi... |
| 17 | From the smallest cell of our skin to the fasc... |
| 18 | Tingling body, increased heart rate, nodding h... |
| 19 | Music speaks directly to our body, giving us t... |
| 20 | We are born as individuals, sharing the myster... |
| 21 | Life expresses its energy through change and m... |
| 22 | Dear Shankra Family, \n We want to introduce y... |
| 23 | Dear Shankra Family, \n We started this festiv... |
| 24 | Dear Shankra Family, \n as we are facing an ou... |
| 25 | Dear Shankra Family, \n we are delighted to an... |
| 26 | Dear Shankra Family, \n we are delighted to an... |
| 27 | Dear Shankra Family, \n we are delighted to an... |

In [30]:

```
pd.Series(' '.join(df2['Post_Text']).lower().split()).value_counts()[:45]
```

```
Out[30]: - 150
the 144
to 85
and 81
we 77
of 71
our 53
a 49
in 39
is 38
with 36
switzerland 30
that 30
are 29
your 29
shankra 27
you 26
will 24
festival 22
this 22
for 22
us 20
by 17
all 17
join 16
be 15
from 15
on 14
as 14
israel 13
an 13
*retro 13
it 13
set* 12
what 12
can 12
have 12
germany 11
message 11
their 11
festivals 10
life 10
do 10
france 10
not 9
dtype: int64
```

```
In [34]: df2[df2['Post_Text'].str.match('Mesolcina')]
df2.head()
```

```
Out[34]:
```

| | Post_Text |
|---|---|
| 0 | Dear Shankra Family, \n first of all, we want ... |
| 1 | Walking towards the future with the strongest ... |
| 2 | The Mesolcina Mountain Chain, home to Shankra ... |
| 3 | We believe in pure music, freely expressed by ... |
| 4 | The beating heart of our Valley, Shankra Stage... |

```
In [ ]:
```

