

```
import nltk
import re
from IPython.display import Markdown, display

def printmd(string):
    display(Markdown(string))
```

Class Task

Read random wikipedia pages, tokenize, lemmatize and create a word cloud.

STEPS:

1. Read Data
2. Make string lower
3. Clean Data (Select only words)
4. Tokenize
5. Remove Stop words
6. Lemmatize
7. Generate Frequency distribution plot (code provided)

```
!pip install wikipedia
import wikipedia
import wordcloud
import matplotlib.pyplot as plt
```

```
# title = wikipedia.random()
title='Kargil'
wikipedia.search(title)
page=wikipedia.page(title)
content = page.content
```

```
print(title)
```

```
Looking in indexes: https://pypi.org/simple, https://us-
python.pkg.dev/colab-wheels/public/simple/
Requirement already satisfied: wikipedia in
/usr/local/lib/python3.7/dist-packages (1.4.0)
Requirement already satisfied: beautifulsoup4 in
/usr/local/lib/python3.7/dist-packages (from wikipedia) (4.6.3)
Requirement already satisfied: requests<3.0.0,>=2.0.0 in
/usr/local/lib/python3.7/dist-packages (from wikipedia) (2.23.0)
Requirement already satisfied: urllib3!=1.25.0,!=1.25.1,<1.26,>=1.21.1
in /usr/local/lib/python3.7/dist-packages (from
requests<3.0.0,>=2.0.0->wikipedia) (1.24.3)
Requirement already satisfied: certifi>=2017.4.17 in
/usr/local/lib/python3.7/dist-packages (from requests<3.0.0,>=2.0.0-
>wikipedia) (2022.6.15)
Requirement already satisfied: idna<3,>=2.5 in
```

```
/usr/local/lib/python3.7/dist-packages (from requests<3.0.0,>=2.0.0-
>wikipedia) (2.10)
Requirement already satisfied: chardet<4,>=3.0.2 in
/usr/local/lib/python3.7/dist-packages (from requests<3.0.0,>=2.0.0-
>wikipedia) (3.0.4)
Kargil
```

content

```
{"type": "string"}
```

STRING LOWER CASE

```
content=content.lower()
content
```

```
{"type": "string"}
```

TOKENIZATION

```
nlTK.download('punkt')
tokenizer=nlTK.tokenize.word_tokenize
tokens= tokenizer(content)
print(tokens)
```

```
['kargil', 'is', 'a', 'city', 'and', 'a', 'joint', 'capital', 'of',
'the', 'union', 'territory', 'of', 'ladakh', 'in', 'india', '.', 'it',
'is', 'also', 'the', 'headquarters', 'of', 'the', 'kargil',
'district', '.', 'it', 'is', 'the', 'second-largest', 'city', 'in',
'ladakh', 'after', 'leh', '.', 'kargil', 'is', 'located', '204',
'kilometres', '(', '127', 'mi', ')', 'to', 'the', 'east', 'of',
'srinagar', 'in', 'jammu', 'and', 'kashmir', ',', 'and', '234',
'kilometres', '(', '145', 'mi', ')', 'to', 'the', 'west', 'of', 'leh',
.', 'it', 'is', 'the', 'centre', 'point', 'of', 'the', 'suru',
'river', '.', '==', 'etymology', '==', 'the', 'name', 'kargil', 'is',
'said', 'to', 'derive', 'from', 'the', 'words', 'khar', ',',
'meaning', 'castle', ',', 'and', 'rkil', 'meaning', '``', 'centre',
''', '``', 'kargil', 'denotes', 'a', 'place', 'between', 'many',
'forts', ',', 'a', 'central', 'place', 'where', 'people', 'could',
'stay', '``', '.', 'it', 'appears', 'to', 'be', 'a', 'fitting',
'description', 'for', 'a', 'place', 'that', 'is', 'equidistant',
'from', 'srinagar', ',', 'leh', 'and', 'skardu', 'historically', ',',
'the', 'region', 'around', 'kargil', 'was', 'called', 'purig', '.',
'a', 'major', 'study', 'of', 'the', 'history', 'of', 'purig', 'is',
'included', 'in', 'the', '1987', 'book', 'qadeem', 'ladakh', 'by',
'kacho', 'sikander', 'khan', ',', 'which', 'includes', 'genealogies',
'of', 'various', 'dynasties', 'that', 'ruled', 'the', 'region', '.',
```

'==', 'history', '==', 'kargil', 'is', 'the', 'main', 'town', 'in', 'the', 'historical', 'region', 'of', 'purig', '.,', 'which', 'consists', 'of', 'the', 'suru', 'river', 'basin', '.,', 'it', 'was', 'not', 'the', 'historical', 'capital', 'of', 'the', 'region', '.,', 'earlier', '.,', 'purig', 'consisted', 'of', 'a', 'number', 'of', 'small', 'but', 'independent', 'kingdoms', '.,', 'which', 'included', 'pashkum', '.,', 'chiktan', '.,', 'fokar', '.,', 'soth', 'and', 'the', 'suru', 'valley', '.,', 'gasho', '.,', 'thatha', 'khan', '.,', 'an', 'exiled', 'prince', 'in', 'the', 'ad', '9th', 'century', '.,', 'is', 'perhaps', 'the', 'first', 'ruler', 'who', 'brought', 'together', 'all', 'the', 'territories', 'under', 'a', 'united', 'administration', '.,', 'another', 'sultan', 'of', 'purig', 'extended', 'his', 'kingdom', 'to', 'include', 'zanskar', '.,', 'sot', '.,', 'barso', '.,', 'sankoo', 'pretty', 'much', 'the', 'territory', 'of', 'the', 'present', 'kargil', 'district', '.,', 'he', 'is', 'referred', 'to', 'as', '.,', 'the', 'purig', 'sultan', '.,', 'his', 'capital', 'was', 'based', 'at', 'karpokhar', 'in', 'the', 'suru', 'valley', '.,', 'the', 'other', 'famous', 'kings', 'of', 'kargil', 'were', 'boti', 'khan', '.,', 'abdal', 'khan', '.,', 'amrood', 'choo', '.,', 'tsering', 'malik', '.,', 'kunchok', 'sherab', 'stan', 'and', 'thi', 'sultan.it', 'is', 'said', 'that', 'it', 'was', 'the', 'period', 'of', 'ali', 'sher', 'khan', 'anchan', '.,', 'the', 'famous', 'ruler', 'of', 'skardu', '.,', 'in', 'the', 'late', '16th', 'and', 'early', '17th', 'centuries', 'which', 'had', 'a', 'great', 'influence', 'on', 'the', 'area', '.,', 'this', 'prince', 'from', 'baltistan', 'conquered', 'most', 'of', 'the', 'principalities', 'of', 'purig', 'and', 'introduced', 'balti', 'culture', 'in', 'the', 'kargil', 'district', '.,', 'subsequently', '.,', 'it', 'was', 'the', 'dogras', 'who', 'united', 'baltistan', '.,', 'purig', '.,', 'zanskar', 'and', 'present-day', 'leh', 'district', 'in', 'the', 'first', 'half', 'of', 'the', '19th', 'century', 'under', 'a', 'single', 'administrative', 'unit', '.,', 'which', 'lasted', 'until', '1947', 'when', 'a', 'new', 'line', 'of', 'control', 'was', 'demarcated', 'between', 'india', 'and', 'pakistan', 'separating', 'skardu', 'and', 'kargil', '.,', 'before', 'the', 'partition', 'of', 'india', 'in', '1947', '.,', 'kargil', 'was', 'part', 'of', 'the', 'wazarat', '(', 'district', ')', 'of', 'ladakh', '.,', 'a', 'sparsely', 'populated', 'region', 'with', 'diverse', 'linguistic', '.,', 'ethnic', 'and', 'religious', 'groups', '.,', 'living', 'in', 'isolated', 'valleys', 'separated', 'by', 'some', 'of', 'the', 'world', '"s", 'highest', 'mountains', '.,', 'the', 'ladakh', 'wazarat', 'had', 'three', 'tehsils', '(', 'sub-districts', ')', '.,', 'named', 'after', 'the', 'cities', 'of', 'their', 'headquarters', '.,', 'leh', '.,', 'skardu', 'and', 'kargil', '.,', 'the', 'district', 'headquarters', 'shifted', 'between', 'the', 'three', 'locations', 'each', 'year.the', 'first', 'kashmir', 'war', '(', '1947-48', ')', 'concluded', 'with', 'a', 'ceasefire', 'line', 'that', 'divided', 'the', 'ladakh', 'wazarat', '.,', 'putting', 'roughly', 'the', 'kargil', 'and', 'leh', 'tehsils', 'on', 'the', 'indian', 'side', '.,', 'and', 'the', 'skardu', 'tehsil', 'on', 'the', 'pakistan', 'side', '.,', 'the', 'two', 'indian',

'tehsils', 'were', 'soon', 'promoted', 'to', 'districts', 'and', 'ladakh', 'was', 'named', 'a', 'division', 'on', 'par', 'with', 'the', 'jammu', 'and', 'kashmir', 'divisions', 'in', 'the', 'indian', 'state', 'of', 'jammu', 'and', 'kashmir', 'pakistan', 'renamed', 'the', 'skardu', 'tehsil', 'baltistan', 'and', 'divided', 'it', 'into', 'further', 'districts', 'at', 'the', 'end', 'of', 'indo-pakistani', 'war', 'of', '1971', 'the', 'two', 'nations', 'signed', 'the', 'simla', 'agreement', 'converting', 'the', 'former', 'ceasefire', 'line', 'with', 'some', 'adjustments', 'into', 'a', 'line', 'of', 'control', 'and', 'promising', 'not', 'to', 'engage', 'in', 'armed', 'conflict', 'with', 'respect', 'to', 'that', 'boundary.in', '1999', 'the', 'area', 'saw', 'infiltration', 'by', 'pakistani', 'forces', 'leading', 'to', 'the', 'kargil', 'war', 'fighting', 'occurred', 'along', 'a', '160', 'km', 'long', 'stretch', 'of', 'ridges', 'overlooking', 'the', 'only', 'road', 'linking', 'srinagar', 'and', 'leh', 'the', 'military', 'outposts', 'on', 'the', 'ridges', 'above', 'the', 'highway', 'were', 'generally', 'around', '5,000', 'metres', '(', '16,000', 'ft', ')', 'high', 'with', 'a', 'few', 'as', 'high', 'as', '5,485', 'metres', '(', '18,000', 'ft', ')', 'after', 'several', 'months', 'of', 'fighting', 'and', 'diplomatic', 'activity', 'the', 'pakistani', 'forces', 'were', 'forced', 'to', 'withdraw', 'to', 'their', 'side', 'of', 'the', 'line', 'of', 'control', 'by', 'their', 'prime', 'minister', 'nawaz', 'sharif', 'after', 'his', 'visit', 'to', 'the', 'usa', '==', 'geography', '==', 'kargil', 'has', 'an', 'average', 'elevation', 'of', '2,676', 'metres', '(', '8,780', 'feet', ')', 'and', 'is', 'situated', 'along', 'the', 'banks', 'of', 'the', 'suru', 'river', '(', 'indus', ')', 'the', 'town', 'of', 'kargil', 'is', 'located', '205', 'km', '(', '127', 'mi', ')', 'from', 'srinagar', 'facing', 'the', 'northern', 'areas', 'across', 'the', 'loc', 'like', 'other', 'areas', 'in', 'the', 'himalayas', 'kargil', 'has', 'a', 'temperate', 'climate', 'summers', 'are', 'hot', 'with', 'cool', 'nights', 'while', 'winters', 'are', 'long', 'and', 'chilly', 'with', 'temperatures', 'often', 'dropping', 'below', '-20', '°c', '(', '-4', '°f', ')', '==', 'demographics', '==', 'during', 'the', '2011', 'census', 'the', 'population', 'of', 'kargil', 'town', 'was', 'recorded', 'at', '16,338.', 'a', 'majority', 'of', 'the', 'population', '(', '11,496', ')', 'is', 'classified', 'as', 'scheduled', 'tribes', 'the', 'literacy', 'rate', 'is', '75', '%', '==', 'religion', '==', 'islam', 'is', 'the', 'largest', 'religion', 'in', 'kargil', 'city', 'followed', 'by', 'over', '77', '%', 'of', 'people', 'hinduism', 'is', 'the', 'second-largest', 'religion', 'with', '19.21', '%', 'adherents', 'buddhism', 'and', 'sikhism', 'form', '0.54', '%', 'and', '2.2', '%', 'of', 'the', 'population', 'respectively', '==', 'media', 'and', 'communications', '==', 'all', 'india', 'radio', "'s", 'channel', 'air', 'kargil', 'am', '684', 'is', 'broadcast', 'from', 'a', 'radio', 'station', 'at', 'kargil', 'town', 'greater', 'ladakh', 'is', 'the', 'largest', 'circulated', 'bi-lingual', 'newspaper', 'in', 'whole', 'ladakh',

'union', 'territory', 'that', 'publishes', 'once', 'in', 'a', 'week',
'.', '==', 'transportation', '==', '==', 'air', '==', 'kargil',
'airport', 'is', 'a', 'non-operational', 'airport', 'located', '8',
'kilometres', 'from', 'the', 'town', '.,', 'the', 'airport', 'is',
'included', 'in', 'udan', 'scheme', 'and', 'is', 'proposed', 'to',
'be', 'operational', 'in', 'the', 'near', 'future', '.,', 'the',
'nearest', 'operational', 'airport', 'is', 'the', 'srinagar',
'international', 'airport', '.,', '==', 'rail', '==', 'there', 'is',
'no', 'rail-connectivity', 'to', 'kargil', 'yet', '.,', 'the',
'srinagar-kargil-leh', 'railway', 'line', 'is', 'proposed', 'which',
'will', 'connect', 'srinagar', 'and', 'leh', 'via', 'kargil', '.,',
'the', 'nearest', 'major', 'railway', 'station', 'to', 'kargil', 'is',
'jammu', 'tawi', 'railway', 'station', 'located', 'at', 'a',
'distance', 'of', '472', 'kilometres', '.,', '==', 'road', '==',
'an', 'indian', 'national', 'highway', '(', 'nh', '1', ')',
'connecting', 'srinagar', 'to', 'leh', 'cuts', 'through', 'kargil',
'.,', '==', 'kargil-skardu', 'road', '==', 'a', 'famous', 'all-
weather', 'road', '.,', 'the', 'kargil-skardu', 'road', 'linked',
'kargil', 'with', 'skardu', '.,', 'a', 'city', 'in', 'gilgit-
baltistan', '.,', 'since', 'the', '1948', 'kashmir', 'war', '.,', 'the',
'road', 'has', 'been', 'closed', '.,', 'whilst', 'the', 'indian',
'government', 'has', 'been', 'interested', 'in', 'opening', 'the',
'road', 'as', 'a', 'humanitarian', 'gesture', '.,', 'this', 'has',
'been', 'refused', 'by', 'the', 'pakistani', 'government', '.,', '==',
'see', 'also', '==', 'kargil', 'war', 'ladakh', 'leh', 'jammu', 'and',
'kashmir', '(', 'state', ')', 'jammu', 'and', 'kashmir', '(', 'union',
'territory', ')', '1988', 'gilgit', 'massacre', '==', 'references',
'==', '==', 'bibliography', '==', 'aggarwal', '.,', 'ravina', '(',
'2004', ')', '.,', 'beyond', 'lines', 'of', 'control', '.,',
'performance', 'and', 'politics', 'on', 'the', 'disputed', 'borders',
'of', 'ladakh', '.,', 'india', '.,', 'duke', 'university', 'press', '.,',
'isbn', '0-8223-3414-3', 'gupta', '.,', 'radhika', '(', '2013', ')',
'.,', '.,', 'allegiance', 'and', 'alienation', '.,', 'border',
'dynamics', 'in', 'kargil', '""', '.,', 'in', 'david', 'n.', 'gellner',
'(', 'ed.', ')', '.,', 'borderland', 'lives', 'in', 'northern',
'south', 'asia', '.,', 'duke', 'university', 'press', '.,', 'pp', '.,',
'47-71', '.,', 'isbn', '978-0-8223-7730-6.', 'huttenback', '.,',
'robert', 'a', '.,', '(', '1961', ')', '.,', '.,', 'gulab', 'singh',
'and', 'the', 'creation', 'of', 'the', 'dogra', 'state', 'of',
'jammu', '.,', 'kashmir', '.,', 'and', 'ladakh', '""', '(', 'pdf', ')',
'.,', 'the', 'journal', 'of', 'asian', 'studies', '.,', '20', '(', '4',
'.)', '.,', '477-488', '.,', 'doi:10.2307/2049956', '.,', 'jstor',
'2049956', 'karim', '.,', 'maj', 'gen', 'afsir', '(', '2013', ')', '.,',
'kashmir', 'the', 'troubled', 'frontiers', '.,', 'lancer',
'publishers', 'llc', '.,', 'pp', '.,', '30-', '.,', 'isbn', '978-1-
935501-76-3', 'panikkar', '.,', 'k.', 'm.', '(', '1930', ')', '.,',
'gulab', 'singh', '.,', 'london', '.,', 'martin', 'hopkinson', 'ltd',
'==', 'further', 'reading', '==', 'hussain', '.,', 'javed', '(', '21',
'october', '2006', ')', '.,', '.,', 'kargil', '.,', 'what', 'might',
'have', 'happened', '""', '.,', 'dawn', '.,', 'archived', 'from', 'the',

```
'original', 'on', '2', 'december', '2008.', 'retrieved', '20', 'may',  
'2009.', 'yukiyasu', 'osada', ';', 'gavin', 'allwright', ';',  
'atsushi', 'kanamaru', '(', '2000', ')', ',', 'mapping', 'the',  
'tibetan', 'world', ',', 'tokyo', ':', 'kotan', 'publishing', '(',  
'published', '2004', ')', ',', 'isbn', '0-9701716-0-9', 'paul',  
'beersmans', '(', '13', 'june', '1998', ')', ',', 'jammu', 'and',  
'kashmir', 'state', '1998', ',', 'belgian', 'association', 'for',  
'solidarity', 'with', 'jammu', 'and', 'kashmir', ',', 'archived',  
'from', 'the', 'original', 'on', '6', 'october', '2007', '==',  
'external', 'links', '==', 'kargil', 'kargil', 'official', 'website']
```

```
[nltk_data] Downloading package punkt to /root/nltk_data...
```

```
[nltk_data] Package punkt is already up-to-date!
```

DATA CLEANING

```
import re  
abc = re.findall(r'[a-z]+', content)  
content = abc
```

REMOVING STOPWORDS

```
nltk.download('stopwords')  
en=set(nltk.corpus.stopwords.words('english'))
```

```
[nltk_data] Downloading package stopwords to /root/nltk_data...
```

```
[nltk_data] Package stopwords is already up-to-date!
```

```
final_content=[]  
for i in content:  
    if i not in en:  
        final_content.append(i)
```

PERFORMING LEMITIZATION

```
nltk.download('wordnet')  
import nltk  
nltk.download('omw-1.4')  
from nltk.stem import WordNetLemmatizer
```

```
lemmatizer = WordNetLemmatizer()
```

```
words = []
```

```
for w in final_content:
    words.append(lemmatizer.lemmatize(w))
```

```
[nltk_data] Downloading package wordnet to /root/nltk_data...
[nltk_data] Package wordnet is already up-to-date!
[nltk_data] Downloading package omw-1.4 to /root/nltk_data...
[nltk_data] Package omw-1.4 is already up-to-date!
```

FREQUENCY DISTRIBUTION

```
from nltk import FreqDist
```

```
freq=FreqDist(words)
print(freq.most_common())
```

```
freq.plot(20)
```

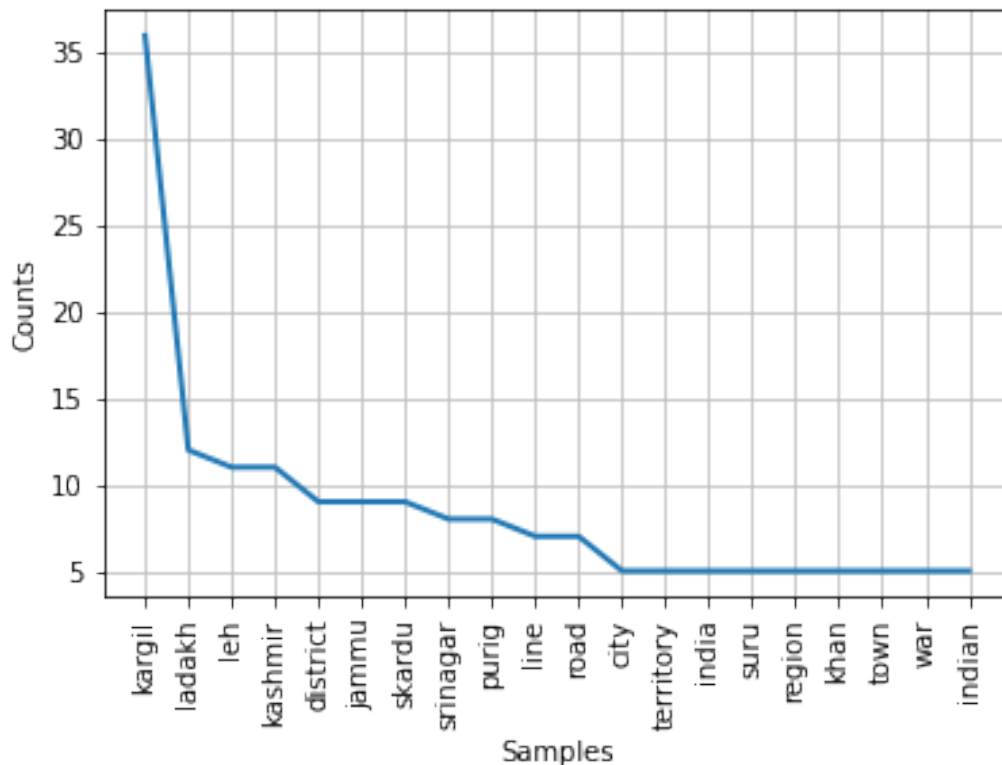
```
[('kargil', 36), ('ladakh', 12), ('leh', 11), ('kashmir', 11),
('district', 9), ('jammu', 9), ('skardu', 9), ('srinagar', 8),
('purig', 8), ('line', 7), ('road', 7), ('city', 5), ('territory', 5),
('india', 5), ('suru', 5), ('region', 5), ('khan', 5), ('town', 5),
('war', 5), ('indian', 5), ('airport', 5), ('largest', 4), ('located',
4), ('kilometre', 4), ('th', 4), ('area', 4), ('baltistan', 4),
('control', 4), ('state', 4), ('pakistani', 4), ('isbn', 4),
('capital', 3), ('union', 3), ('headquarters', 3), ('mi', 3),
('river', 3), ('place', 3), ('included', 3), ('valley', 3),
('century', 3), ('first', 3), ('sultan', 3), ('famous', 3),
('pakistan', 3), ('wazarat', 3), ('tehsils', 3), ('side', 3),
('metre', 3), ('population', 3), ('religion', 3), ('station', 3),
('operational', 3), ('railway', 3), ('also', 2), ('second', 2),
('centre', 2), ('said', 2), ('meaning', 2), ('people', 2), ('around',
2), ('major', 2), ('study', 2), ('history', 2), ('historical', 2),
('kingdom', 2), ('prince', 2), ('ruler', 2), ('united', 2),
('zanskar', 2), ('present', 2), ('world', 2), ('three', 2), ('named',
2), ('ceasefire', 2), ('divided', 2), ('tehsil', 2), ('two', 2),
('division', 2), ('force', 2), ('fighting', 2), ('along', 2), ('km',
2), ('long', 2), ('ridge', 2), ('highway', 2), ('ft', 2), ('high', 2),
('northern', 2), ('radio', 2), ('air', 2), ('proposed', 2),
('nearest', 2), ('rail', 2), ('gilgit', 2), ('government', 2),
('border', 2), ('duke', 2), ('university', 2), ('press', 2), ('pp',
2), ('gulab', 2), ('singh', 2), ('october', 2), ('archived', 2),
('original', 2), ('joint', 1), ('east', 1), ('west', 1), ('point', 1),
('etymology', 1), ('name', 1), ('derive', 1), ('word', 1), ('khar',
1), ('castle', 1), ('rkil', 1), ('denotes', 1), ('many', 1), ('fort',
1), ('central', 1), ('could', 1), ('stay', 1), ('appears', 1),
('fitting', 1), ('description', 1), ('equidistant', 1),
('historically', 1), ('called', 1), ('book', 1), ('qadeem', 1),
('kacho', 1), ('sikander', 1), ('includes', 1), ('genealogy', 1),
('various', 1), ('dynasty', 1), ('ruled', 1), ('main', 1),
('consists', 1), ('basin', 1), ('earlier', 1), ('consisted', 1),
```

('number', 1), ('small', 1), ('independent', 1), ('pashkum', 1),
('chiktan', 1), ('fokar', 1), ('soth', 1), ('gasho', 1), ('thatha',
1), ('exiled', 1), ('ad', 1), ('perhaps', 1), ('brought', 1),
('together', 1), ('administration', 1), ('another', 1), ('extended',
1), ('include', 1), ('sot', 1), ('barso', 1), ('sankoo', 1),
('pretty', 1), ('much', 1), ('referred', 1), ('based', 1),
('karpokhar', 1), ('king', 1), ('boti', 1), ('abdal', 1), ('amrood',
1), ('choo', 1), ('tsering', 1), ('malik', 1), ('kunchok', 1),
('sherab', 1), ('stan', 1), ('thi', 1), ('period', 1), ('ali', 1),
('sher', 1), ('anchan', 1), ('late', 1), ('early', 1), ('great', 1),
('influence', 1), ('conquered', 1), ('principality', 1),
('introduced', 1), ('balti', 1), ('culture', 1), ('subsequently', 1),
('dogras', 1), ('day', 1), ('half', 1), ('single', 1),
('administrative', 1), ('unit', 1), ('lasted', 1), ('new', 1),
('demarcated', 1), ('separating', 1), ('partition', 1), ('part', 1),
('sparsely', 1), ('populated', 1), ('diverse', 1), ('linguistic', 1),
('ethnic', 1), ('religious', 1), ('group', 1), ('living', 1),
('isolated', 1), ('separated', 1), ('highest', 1), ('mountain', 1),
('sub', 1), ('shifted', 1), ('location', 1), ('year', 1),
('concluded', 1), ('putting', 1), ('roughly', 1), ('soon', 1),
('promoted', 1), ('par', 1), ('renamed', 1), ('end', 1), ('indo', 1),
('nation', 1), ('signed', 1), ('simla', 1), ('agreement', 1),
('converting', 1), ('former', 1), ('adjustment', 1), ('promising', 1),
('engage', 1), ('armed', 1), ('conflict', 1), ('respect', 1),
('boundary', 1), ('saw', 1), ('infiltration', 1), ('leading', 1),
('occurred', 1), ('stretch', 1), ('overlooking', 1), ('linking', 1),
('military', 1), ('outpost', 1), ('generally', 1), ('several', 1),
('month', 1), ('diplomatic', 1), ('activity', 1), ('forced', 1),
('withdraw', 1), ('prime', 1), ('minister', 1), ('nawaz', 1),
('sharif', 1), ('visit', 1), ('usa', 1), ('geography', 1), ('average',
1), ('elevation', 1), ('foot', 1), ('situated', 1), ('bank', 1),
('indus', 1), ('facing', 1), ('across', 1), ('loc', 1), ('like', 1),
('himalaya', 1), ('temperate', 1), ('climate', 1), ('summer', 1),
('hot', 1), ('cool', 1), ('night', 1), ('winter', 1), ('chilly', 1),
('temperature', 1), ('often', 1), ('dropping', 1), ('c', 1), ('f', 1),
('demographic', 1), ('census', 1), ('recorded', 1), ('majority', 1),
('classified', 1), ('scheduled', 1), ('tribe', 1), ('literacy', 1),
('rate', 1), ('islam', 1), ('followed', 1), ('hinduism', 1),
('adherent', 1), ('buddhism', 1), ('sikhism', 1), ('form', 1),
('respectively', 1), ('medium', 1), ('communication', 1), ('channel',
1), ('broadcast', 1), ('greater', 1), ('circulated', 1), ('bi', 1),
('lingual', 1), ('newspaper', 1), ('whole', 1), ('publishes', 1),
('week', 1), ('transportation', 1), ('non', 1), ('udan', 1),
('scheme', 1), ('near', 1), ('future', 1), ('international', 1),
('connectivity', 1), ('yet', 1), ('connect', 1), ('via', 1), ('tawi',
1), ('distance', 1), ('national', 1), ('nh', 1), ('connecting', 1),
('cut', 1), ('weather', 1), ('linked', 1), ('since', 1), ('closed',
1), ('whilst', 1), ('interested', 1), ('opening', 1), ('humanitarian',
1), ('gesture', 1), ('refused', 1), ('see', 1), ('massacre', 1),
('reference', 1), ('bibliography', 1), ('aggarwal', 1), ('ravina', 1),


```

('beyond', 1), ('performance', 1), ('politics', 1), ('disputed', 1),
('gupta', 1), ('radhika', 1), ('allegiance', 1), ('alienation', 1),
('dynamic', 1), ('david', 1), ('n', 1), ('gellner', 1), ('ed', 1),
('borderland', 1), ('life', 1), ('south', 1), ('asia', 1),
('huttenback', 1), ('robert', 1), ('creation', 1), ('dogra', 1),
('pdf', 1), ('journal', 1), ('asian', 1), ('doi', 1), ('jstor', 1),
('karim', 1), ('maj', 1), ('gen', 1), ('afsir', 1), ('troubled', 1),
('frontier', 1), ('lancer', 1), ('publisher', 1), ('llc', 1),
('panikkar', 1), ('k', 1), ('london', 1), ('martin', 1), ('hopkinson',
1), ('ltd', 1), ('reading', 1), ('hussain', 1), ('javed', 1),
('might', 1), ('happened', 1), ('dawn', 1), ('december', 1),
('retrieved', 1), ('may', 1), ('yukiyasu', 1), ('osada', 1), ('gavin',
1), ('allwright', 1), ('atsushi', 1), ('kanamaru', 1), ('mapping', 1),
('tibetan', 1), ('tokyo', 1), ('kotan', 1), ('publishing', 1),
('published', 1), ('paul', 1), ('beersmans', 1), ('june', 1),
('belgian', 1), ('association', 1), ('solidarity', 1), ('external',
1), ('link', 1), ('official', 1), ('website', 1)]

```



```

<matplotlib.axes._subplots.AxesSubplot at 0x7f7e52291d90>

```

```

wc=wordcloud.WordCloud(width = 800, height = 800,
                        background_color = 'white',
min_font_size=10).generate_from_frequencies(freq)
plt.figure()

plt.imshow(wc)

```

```
<matplotlib.image.AxesImage at 0x7f7e4993c110>
```

