

```
In [51]: import nltk, re, string, collections
from nltk import ngrams
import numpy as np
import pandas as pd
from pandas import DataFrame
import seaborn as sns
import matplotlib.pyplot as plt
nltk.download('punkt')
```

[nltk_data] Downloading package punkt to /Users/medha/nltk_data...
[nltk_data] Package punkt is already up-to-date!

Out[51]: True

```
In [52]: path = "/Users/medha/NLP/assignment1/The Time Machine.txt"
with open(path) as file:
    text = file.read()
```

```
In [53]: tokens = text.split()

bigrams = ngrams(tokens, 2)
```

```
In [54]: frequency = nltk.FreqDist(bigrams)
common = frequency.most_common(1000)

print(common)
```

[(('of', 'the'), 337), (('in', 'the'), 176), (('I', 'had'), 127), (('I', 'was'), 108), (('and', 'the'), 102), (('to', 'the'), 99), (('of', 'a'), 76), (('the', 'Time'), 69), (('in', 'a'), 60), (('I', 'could'), 60), (('I', 'saw'), 59), (('as', 'I'), 58), (('it', 'was'), 56), (('and', 'I'), 55), (('Project', 'Gutenberg-tm'), 53), (('for', 'the'), 52), (('that', 'I'), 52), (('at', 'the'), 49), (('from', 'the'), 47), (('of', 'my'), 47), (('on', 'the'), 47), (('into', 'the'), 46), (('with', 'the'), 45), (('upon', 'the'), 44), (('that', 'the'), 44), (('was', 'a'), 43), (('I', 'felt'), 43), (('said', 'the'), 38), (('by', 'the'), 38), (('was', 'the'), 38), (('in', 'my'), 38), (('with', 'a'), 37), (('seemed', 'to'), 37), (('to', 'be'), 36), (('Time', 'Traveller'), 35), (('for', 'a'), 34), (('out', 'of'), 33), (('had', 'been'), 33), (('the', 'same'), 33), (('It', 'was'), 31), (('of', 'this'), 29), (('a', 'little'), 28), (('I', 'found'), 28), (('the', 'Project'), 27), (('the', 'little'), 27), (('I', 'did'), 27), (('The', 'Time'), 26), (('through', 'the'), 26), (('I', 'have'), 25), (('all', 'the'), 25), (('Then', 'I'), 25), (('and', 'then'), 24), (('there', 'was'), 23), (('to', 'me'), 23), (('to', 'a'), 23), (('Project', 'Gutenberg'), 22), (('under', 'the'), 22), (('must', 'have'), 22), (('have', 'been'), 22), (('began', 'to'), 22), (('like', 'a'), 22), (('and', 'a'), 21), (('down', 'the'), 21), (('towards', 'the'), 21), (('thought', 'of'), 21), (('over', 'the'), 20), (('as', 'it'), 20), (('with', 'my'), 20), (('they', 'were'), 20), (('I', 'thought'), 20), (('when', 'I'), 20), (('the', 'Morlocks'), 20), (('of', 'our'), 19), (('did', 'not'), 19), (('had', 'a'), 19), (('as', 'the'), 19), (('me', 'and'), 19), (('me.', 'I'), 19), (('of', 'that'), 18), (('up', 'the'), 18), (('among', 'the'), 18), (('Gutenberg-tm', 'electronic'), 18), (('Time', 'Machine'), 17), (('in', 'this'), 17), (('which', 'I'), 17), (('about', 'the'), 17), (('one', 'of'), 17), (('of', 'his'), 17), (('the', 'Medical'), 17), (('my', 'own'), 17), (('and', 'in'), 17), (('at', 'my'), 17), (('is', 'a'), 16), (('it', 'seemed'), 16), (('saw', 'the'), 16), (('I', 'went'), 16), (('to', 'my'), 16), (('the', 'sun'), 16), (('of', 'these'), 16), (('I', 'began'), 16), (('terms', 'of'), 15), (('and', 'there'), 15), (('do', 'not'), 15), (('kind', 'of'), 15), (('that', 'it'), 15), (('was', 'in'), 15), (('as', 'a'), 15), (('and', 'it'), 15), (('could', 'not'), 15), (('came', 'to'), 15), (('I', 'stood'), 15), (('the', 'world'), 15), (('But', 'I'), 15), (('along', 'the'), 14), (('tried', 'to'), 14), (('to', 'see'), 14), (('to', 'me'), 14), (('I', 'must'), 14), (('was', 'not'), 14), (('of', 'an'), 14), (('into', 'my'), 14), (('across', 'the'), 14), (('me', 'to'), 14), (('and', 'that'), 13), (('of', 'its'), 13), (('in', 'his'), 13), (('the', 'machine'), 13), (('I', 'looked'), 13), (('a', 'moment'), 13), (('I', 'heard'), 13), (('the', 'thing'), 13), (('I', 'came'), 13), (('our', 'own'), 13), (('saw', 'a'), 13), (('of', 'them'), 13), (('little', 'people'), 13), (('that', 'was'), 13), (('of', 'their'), 13), (('I', 'tried'), 13), (('Gutenberg', 'Literary'), 13), (('Literary', 'Archive'), 13), (('the', 'terms'), 12), (('In', 'the'), 12), (('the', 'White'), 12), (('was', 'so'), 12), (('up', 'and'), 12), (('it', 'in'), 12), (('There', 'were'), 12), (('I', 'should'), 12), (('but', 'the'), 12), (('a', 'certain'), 12), (('the', 'first'), 12), (('and', 'went'), 12), (('As', 'I'), 12), (('could', 'see'), 12), (('in', 'their'), 12), (('of', 'some'), 12), (('the', 'United'), 12), (('Time', 'Machine'), 11), (('on', 'a'), 11), (('it', 'is'), 11), (('only', 'a'), 11), (('and', 'so'), 11), (('a', 'kind'), 11), (('such', 'a'), 11), (('the', 'whole'), 11), (('the', 'great'), 11), (('I', 'am'), 11), (('Time', 'Traveller'), 11), (('would', 'be'), 11), (('there', 'is'), 11), (('of', 'it'), 11), (('had', 'seen'), 11), (('sense', 'of'), 11), (('my', 'mind'), 11), (('that', 'had'), 11), (('but', 'I'), 11), (('that', 'my'), 11), (('like', 'the'), 11), (('had', 'come'), 11), (('into', 'a'), 11), (('to', 'make'), 11), (('against', 'the'), 11), (('Palace', 'of'), 10), (('of', 'Green'), 10), (('back', 'to'), 10), (('see', 'the'), 10), (('looked', 'at'), 10), (('in', 'which'), 10), (('I', 'took'), 10), (('And', 'the'), 10), (('and', 'as'), 10), (('came', 'into'), 10), (('moment', 'I'), 10), (('my', 'mind'), 10), (('was', 'still'), 10), (('at', 'last'), 10), (('and', 'my'), 10), (('was', 'to'), 10), (('the', 'full'), 10), (('I', 'made'), 10), (('from', 'my'), 10), (('had', 'the'), 10), (('by', 'a'), 10), (('here', 'and'), 10), (('the', 'old'), 10), (('the', 'darkness'), 10), (('electronic', 'works'), 10), (('If', 'you'), 10), (('and', 'with'), 9), (('a', 'very'), 9), (('Medical', 'Man'), 9), (('at', 'a'), 9), (('away', 'from'), 9), (('There', 'was'), 9), (('me', 'that'), 9), (('he', 'said'), 9), (('Then', 'the'), 9), (('which', 'the'), 9), (('the', 'door'), 9), (('For', 'a'), 9), (('some', 'of'), 9), (('had', 'to'), 9), (('to', 'get'), 9), (('I', 'put'), 9), (('of', 'any'), 9), (('was', 'very'), 9), (('upon', 'me'), 9), (('came', 'a'), 9), (('what', 'I'), 9), (('a', 'vast'), 9), (('determined', 'to'), 9), (('at', 'first'), 9), (('the', 'hill'), 9), (('the', 'bronze'), 9), (('part', 'of'), 9), (('my', 'hand'), 9), (('the', 'Palace'), 9), (('the', 'work'), 9), (('Archive', 'Foundation'), 9), (('Time', 'Traveller's'), 8), (('it', 'to'), 8), (('Time', 'Traveller'), 8), (('you', 'are'), 8), (('is', 'the'), 8), (('means', 'of'), 8), (('may', 'be'), 8), (('his', 'hand'), 8), (('and', 'sat'), 8), (('I', 'don't'), 8), (('was', 'no'), 8), (('should', 'have'), 8), (('I', 'remember'), 8), (('hand', 'and'), 8), (('me', 'in'), 8), (('the', 'Editor'), 8), (('who', 'had'), 8), (('as', 'my'), 8), (('and', 'saw'), 8), (('They', 'were'), 8), (('you', 'must'), 8), (('At', 'first'), 8), (('was', 'at'), 8), (('it', 'had'), 8), (('my', 'eyes'), 8), (('the', 'white'), 8), (('and', 'was'), 8), (('come', 'to'), 8), (('then', 'I'), 8), (('a', 'long'), 8), (('I', 'fancied'), 8), (('I', 'judged'), 8), (('the', 'most'), 8), (('I', 'determined'), 8), (('I', 'got'), 8), (('sat', 'down'), 8), (('as', 'well'), 8), (('they', 'had'), 8), (('where', 'I'), 8), (('in', 'that'), 8), (('went', 'on'), 8), (('to', 'me'), 8), (('but', 'it'), 8), (('the', 'Morlocks'), 8), (('set', 'forth'), 8), (('forth', 'in'), 8), (('of', 'Project'), 8), (('access', 'to'), 8), (('electronic', 'work'), 8), (('with', 'this'), 7), (('will', 'be'), 7), (('have', 'to'), 7), (('the', 'Psychologist'), 7), (('have', 'a'), 7), (('I', 'will'), 7), (('in', 'one'), 7), (('side', 'of'), 7), (('a', 'man'), 7), (('it', 'and'), 7), (('above', 'the'), 7), (('if', 'I'), 7), (('so', 'I'), 7), (('with', 'his'), 7), (('to', 'his')

), 7), (('a', 'small'), 7), (('must', 'be'), 7), (('Then', 'he'), 7), (('so', 'that'), 7), (('I', 'sat'), 7), (('The', 'Medical'), 7), (('Medical', 'Man'), 7), (('look', 'at'), 7), (('the', 'lamp'), 7), (('to', 'have'), 7), (('on', 'my'), 7), (('would', 'have'), 7), (('a', 'minute'), 7), (('of', 'what'), 7), (('the', 'laboratory'), 7), (('how', 'to'), 7), (('it.', 'I'), 7), (('Time', 'Machine.'), 7), (('But', 'the'), 7), (('to', 'do'), 7), (('looked', 'round'), 7), (('in', 'an'), 7), (('not', 'a'), 7), (('made', 'a'), 7), (('turned', 'to'), 7), (('you', 'the'), 7), (('had', 'not'), 7), (('of', 'all'), 7), (('the', 'moon'), 7), (('and', 'had'), 7), (('the', 'sky'), 7), (('the', 'earth'), 7), (('the', 'ground'), 7), (('appeared', 'to'), 7), (('I', 'might'), 7), (('I', 'turned'), 7), (('and', 'their'), 7), (('Hundred', 'and'), 7), (('number', 'of'), 7), (('for', 'me'), 7), (('as', 'they'), 7), (('time', 'I'), 7), (('soon', 'as'), 7), (('my', 'first'), 7), (('were', 'no'), 7), (('It', 'may'), 7), (('the', 'light'), 7), (('time', 'in'), 7), (('the', 'black'), 7), (('might', 'be'), 7), (('It', 'had'), 7), (('I', 'lit'), 7), (('made', 'me'), 7), (('protect ed', 'by'), 7), (('copies', 'of'), 7), (('agree', 'to'), 7), (('this', 'agreement'), 7), (('use', 'of'), 6), (('You', 'may'), 6), (('Green', 'Porcelain'), 6), (('White', 'Sphinx'), 6), (('for', 'instance,'), 6), (('between', 'the'), 6), (('the', 'end'), 6), (('It', 'is'), 6), (('is', 'no'), 6), (('the', 'other'), 6), (('or', 'so'), 6), (('think', 'that'), 6), (('think', 'I'), 6), (('in', 'any'), 6), (('Medical', 'Man.'), 6), (('before', 'the'), 6), (('in', 'all'), 6), (('than', 'a'), 6), (('you', 'will'), 6), (('they', 'would'), 6), (('I', 'never'), 6), (('now', 'I'), 6), (('light', 'of'), 6), (('of', 'which'), 6), (('perhaps', 'a'), 6), (('as', 'to'), 6), (('that', 'there'), 6), (('as', 'though'), 6), (('in', 'some'), 6), (('off', 'the'), 6), (('he', 'was'), 6), (('to', 'look'), 6), (('the', 'future'), 6), (('or', 'the'), 6), (('if', 'it'), 6), (('of', 'it.'), 6), (('or', 'a'), 6), (('and', 'how'), 6), (('of', 'us'), 6), (('The', 'Editor'), 6), (('round', 'the'), 6), (('was', 'all'), 6), (('sound', 'of'), 6), (('the', 'new'), 6), (('the', 'place'), 6), (('if', 'you'), 6), (('now', 'and'), 6), (('and', 'again'), 6), (('I', 'seemed'), 6), (('felt', 'a'), 6), (('the', 'machine,'), 6), (('was', 'too'), 6), (('while', 'I'), 6), (('it', 'as'), 6), (('no', 'longer'), 6), (('instead', 'of'), 6), (('It', 'seemed'), 6), (('and', 'to'), 6), (('saw', 'that'), 6), (('the', 'machine.'), 6), (('me', 'as'), 6), (('took', 'my'), 6), (('my', 'hands'), 6), (('up', 'to'), 6), (('upon', 'my'), 6), (('wanted', 'to'), 6), (('a', 'sudden'), 6), (('might', 'have'), 6), (('to', 'and'), 6), (('that', 'they'), 6), (('a', 'great'), 6), (('me', 'with'), 6), (('found', 'the'), 6), (('found', 'a'), 6), (('came', 'the'), 6), (('signs', 'of'), 6), (('the', 'last'), 6), (('down', 'into'), 6), (('I', 'knew'), 6), (('went', 'down'), 6), (('and', 'fell'), 6), (('my', 'Time'), 6), (('my', 'head'), 6), (('trying', 'to'), 6), (('and', 'when'), 6), (('enough', 'to'), 6), (('At', 'last,'), 6), (('after', 'a'), 6), (('you', 'may'), 6), (('a', 'few'), 6), (('I', 'left'), 6), (('and', 'she'), 6), (('I', 'would'), 6), (('the', 'night'), 6), (('to', 'find'), 6), (('I', 'struck'), 6), (('to', 'go'), 6), (('the', 'human'), 6), (('long', 'since'), 6), (('full', 'of'), 6), (('the', 'red'), 6), (('my', 'iron'), 6), (('United', 'States'), 6), (('this', 'work'), 6), (('Gutenberg-tm', 'License'), 6), (('this', 'agreement,'), 6), (('in', 'our'), 5), (('You', 'know'), 5), (('it', 'must'), 5), (('through', 'a'), 5), (('to', 'you'), 5), (('know', 'how'), 5), (('they', 'could'), 5), (('I', 'think'), 5), (('telling', 'you'), 5), (('after', 'the'), 5), (('again,', 'and'), 5), (('move', 'about'), 5), (('that', 'is'), 5), (('more', 'than'), 5), (('the', 'Psychologist'), 5), (('going', 'to'), 5), (('The', 'Psychologist'), 5), (('to', 'tell'), 5), (('The', 'thing'), 5), (('and', 'very'), 5), (('the', 'table'), 5), (('pointed', 'to'), 5), (('got', 'up'), 5), (('out', 'his'), 5), (('At', 'that'), 5), (('he', 'said,'), 5), (('travelled', 'into'), 5), (('what', 'it'), 5), (('were', 'not'), 5), (('thing', 'was'), 5), (('at', 'me'), 5), (('felt', 'that'), 5), (('to', 'let'), 5), (('far', 'as'), 5), (('He', 'was'), 5), (('when', 'the'), 5), (('as', 'if'), 5), (('up', 'in'), 5), (('and', 'looked'), 5), (('a', 'most'), 5), (('a', 'mere'), 5), (('the', 'face'), 5), (('face', 'of'), 5), (('myself', 'in'), 5), (('another', 'moment'), 5), (('I', 'cannot'), 5), (('was', 'already'), 5), (('under', 'my'), 5), (('first', 'I'), 5), (('about', 'me,'), 5), (('time', 'and'), 5), (('I', 'noticed'), 5), (('my', 'eyes'), 5), (('a', 'strange'), 5), (('They', 'had'), 5), (('towards', 'me.'), 5), (('the', 'bushes'), 5), (('were', 'the'), 5), (('that,', 'I'), 5), (('the', 'air'), 5), (('struck', 'me'), 5), (('sight', 'of'), 5), (('them', 'in'), 5), (('and', 'presently'), 5), (('round', 'me'), 5), (('Eight', 'Hundred'), 5), (('of', 'one'), 5), (('towards', 'me,'), 5), (('about', 'my'), 5), (('the', 'sphinx'), 5), (('memory', 'of'), 5), (('the', 'big'), 5), (('a', 'number'), 5), (('I', 'say,'), 5), (('Machine', 'was'), 5), (('fancied', 'I'), 5), (('were', 'very'), 5), (('for', 'I'), 5), (('in', 'spite'), 5), (('spite', 'of'), 5), (('thing', 'I'), 5), (('cries', 'of'), 5), (('“As', 'I'), 5), (('Here', 'and'), 5), (('people', 'were'), 5), (('was', 'my'), 5), (('sun', 'had'), 5), (('there', 'came'), 5), (('the', 'conditions'), 5), (('and', 'more'), 5), (('now', 'a'), 5), (('the', 'end,'), 5), (('hither', 'and'), 5), (('all', 'my'), 5), (('And', 'I'), 5), (('well', 'as'), 5), (('pedestal', 'of'), 5), (('All', 'the'), 5), (('near', 'the'), 5), (('was', 'full'), 5), (('my', 'feet'), 5), (('on', 'either'), 5), (('don't', 'know'), 5), (('and', 'all'), 5), (('you', 'can'), 5), (('and', 'now'), 5), (('a', 'queer'), 5), (('dark,', 'and'), 5), (('and', 'its'), 5), (('a', 'peculiar'), 5), (('with', 'its'), 5), (('and', 'they'), 5), (('did', 'so,'), 5), (('had', 'long'), 5), (('in', 'part'), 5), (('to', 'think'), 5), (('the', 'gallery'), 5), (('the', 'huge'), 5), (('of', 'camphor'), 5), (('for', 'my'), 5), (('not', 'protected'), 5), (('by', 'U.S.'), 5), (('distribution', 'of'), 5), (('comply', 'with'), 5), (('works', 'in'), 5), (('a', 'Project'), 5), (('in', 'paragraph'), 5), (('work', 'is'), 5), (('work', 'in'), 5), (('full', 'Project'), 5), (('or', 'any'), 5), (('the', 'copyright'), 5), (('or', 'other'), 5), (('The', 'Project'), 4), (('H.', 'G.'), 4), (('G.', 'Wells'), 4), (('the', 'use'), 4), (('it', 'will'), 4), (('and', 'his'), 4), (('put', 'it'), 4), (('thing', 'to'), 4), (('mean', 'to'), 4), (('as', 'much'), 4), (('know', 'of'), 4), (('that', 'a'), 4), (('line', 'of'), 4), (('a', 'real'), 4), (('said', 'Filby.'), 4), (('does', 'not'), 4), (('for', 'any'), 4), (('which', 'we'), 4), (('that', 'this'), 4), (('a', 'slight'), 4), (('way', 'of'), 4), (('looking', 'at'), 4), (('and', 'any'), 4), (('any', 'of'), 4), (('to', 'say'), 4), (('the', 'Provincial'), 4), (('we', 'can'), 4), (('for', 'some'), 4), (('the', 'fire,'), 4), (('about', 'in'), 4), (('save', 'for'), 4), (('had', 'no'), 4), (('you', 'cannot'), 4), (('where', 'you'), 4), (('where', 'the'), 4), (('have', 'no'), 4), (('we', 'should'), 4), (('as', 'you'), 4), (('than', 'the'), 4), (('Very', 'Young'), 4), (('said', 'I,'), 4), (('the', 'Psychologist,'), 4), (('and', 'we'), 4), (('he', 'had'), 4), (('And', 'now'), 4), (('in', 'front'), 4), (('front', 'of'), 4), (('up', 'a'), 4), (('the', 'bright'), 4), (('us,', 'and'), 4), (('one', 'little'), 4), (('lever,', 'and'), 4), (('when', 'we'), 4), (('a', 'time'), 4), (('Presently', 'I'), 4), (('the', 'lever,'), 4), (('to', 'put'), 4), (('put', 'out'), 4), (('the', 'lever'), 4), (('One', 'of'), 4), (('stared', 'at'), 4), (('each', 'other.'), 4), (('his', 'pipe,'), 4), (('it', 'would'), 4), (('all', 'this'), 4), (('we', 'came'), 4), (('with', 'an'), 4), (('the', 'air.'), 4), (('it', 'were'), 4), (('the', 'space'), 4), (('minute', 'or'), 4), (('or', 'so.'), 4), (('the', 'way'), 4), (('there', 'in'), 4), (('were', 'of'), 4), (('sheets', 'of'), 4), (('to', 'take'), 4), (('I', 'caught'), 4), (('at', 'that'), 4), (('never', 'felt'), 4), (('that', 'you'), 4), (('that', 'would'), 4), (('very', 'much'), 4), (('most', 'of'), 4), (('again', 'to'), 4), (('the', 'fire'), 4), (('one', 'hand'), 4), (('but', 'a'), 4), (('I', 'remembered'), 4), (('mind', 'was'), 4), (('the', 'Silent'), 4), (('Silent', 'Man'), 4), (('took', 'up'), 4), (('home', 'to'), 4), (('would', 'not'), 4), (('you', 'have'), 4), (('the', 'Editor.'), 4), (('rest', 'of'), 4), (('At', 'last'), 4), (('come', 'into'), 4), (('have', 'told'), 4), (('not', 'be en'), 4), (('only', 'the'), 4), (('ceased', 'to'), 4), (('I', 'suppose'), 4), (('There', 'is'), 4), (('upon', 'a'), 4), (('felt', 'the'), 4), (('suggestion', 'of'), 4), (('the', 'sky,'), 4), (('impression', 'of'), 4), (('me.', 'The'), 4), (('glimpse', 'of'), 4), (('night', 'and'), 4), (('the', 'stars,'), 4), (('eyes.', 'The'), 4), (('me,', 'I'), 4), (('I', 'thought,'), 4), (('the', 'dim'), 4), (('myself', 'and'), 4), (('occurred', 'to'), 4), (('feeling', 'of'), 4), (('I', 'resolved'), 4), (('resolved', 'to'), 4), (('I', 'may'), 4), (('may', 'have'), 4), (('was', 'gone.'), 4), (('noticed', 'that'), 4), (('In', 'a'), 4), (('of', 'white'), 4), (('me,', 'was'), 4), (('to', 'watch'), 4), (('last', 'I'), 4), (('sky', 'was'), 4), (('the', 'sun.'), 4), (('the', 'more'), 4), (('about', 'me'), 4), (('felt', 'as'), 4), (('and', 'turned'), 4), (('at', 'this'), 4), (('first', 'time'), 4), (('reminded', 'me'), 4), (('me', 'of'), 4), (('At', 'the'), 4), (('sweet', 'and'), 4), (('me.', 'It'), 4), (('my', 'head,'), 4), (('my', 'pocket.'), 4), (('which', 'was'), 4), (('not', 'the'), 4), (('face,', 'and'), 4), (('to', 'each'), 4), (('each', 'other,'), 4), (('I', 'pointed'), 4), (('Machine', 'and'), 4), (('how', 'it'), 4), (('and', 'Two'), 4), (('Two', 'Thousand'), 4), (('a', 'question'), 4), (('the', 'sun,'), 4), (('altogether', 'new'), 4), (('was', 'almost'), 4), (('should', 'be'), 4), (('a', 'huge'), 4), (('crowd', 'of'), 4), (('by', 'an'), 4), (('mass', 'of'), 4), (('perhaps,', 'a'), 4), (('heaps', 'of'), 4), (('With', 'a'), 4), (('thing', 'that'), 4), (('the', 'lower'), 4), (('corner', 'of'), 4), (('the', 'time'), 4), (('thing', 'in'), 4), (('I', 'saw,'), 4), (('make', 'a'), 4), (('the

', 'next'), 4), (('make', 'the'), 4), (('at', 'least'), 4), (('even', 'the'), 4), (('very', 'little'), 4), (('my', 'little'), 4), (('as', 'soon'), 4), (('had', 'left'), 4), (('a', 'mile'), 4), (('from', 'which'), 4), (('could', 'get'), 4), (('I', 'walked'), 4), (('heap', 'of'), 4), (('masses', 'of'), 4), (('to', 'what'), 4), (('on', 'which'), 4), (('there', 'were'), 4), (('down', 'on'), 4), (('itself', 'to'), 4), (('conditions', 'of'), 4), (('triumph', 'of'), 4), (('of', 'human'), 4), (('leaving', 'the'), 4), (('and', 'still'), 4), (('tell', 'you'), 4), (('state', 'of'), 4), (('No', 'doubt'), 4), (('and', 'no'), 4), (('the', 'pedestal'), 4), (('the', 'pale'), 4), (('to', 'myself'), 4), (('this', 'strange'), 4), (('was', 'an'), 4), (('I', 'ran'), 4), (('trace', 'of'), 4), (('assured', 'of'), 4), (('a', 'match.'), 4), (('When', 'I'), 4), (('a', 'thing'), 4), (('the', 'sky.'), 4), (('at', 'last'), 4), (('had', 'only'), 4), (('found', 'myself'), 4), (('some', 'time'), 4), (('if', 'they'), 4), (('them', 'and'), 4), (('received', 'the'), 4), (('and', 'began'), 4), (('thought', 'I'), 4), (('heard', 'a'), 4), (('a', 'time'), 4), (('the', 'thought'), 4), (('do', 'with'), 4), (('Yet', 'I'), 4), (('the', 'mystery'), 4), (('presence', 'of'), 4), (('the', 'side'), 4), (('was', 'absolutely'), 4), (('and', 'these'), 4), (('had', 'taken'), 4), (('White', 'Sphinx.'), 4), (('she', 'was'), 4), (('before', 'I'), 4), (('of', 'her'), 4), (('it', 'was'), 4), (('down', 'and'), 4), (('I', 'think'), 4), (('“It', 'was'), 4), (('the', 'dark'), 4), (('and', 'for'), 4), (('over', 'my'), 4), (('the', 'day'), 4), (('Weena', 'was'), 4), (('be', 'that'), 4), (('before', 'me.'), 4), (('came', 'upon'), 4), (('And', 'then'), 4), (('down', 'in'), 4), (('but', 'that'), 4), (('the', 'beautiful'), 4), (('Weena', 'and'), 4), (('there', 'are'), 4), (('night', 'I'), 4), (('my', 'matches'), 4), (('waiting', 'for'), 4), (('of', 'light'), 4), (('came', 'back'), 4), (('to', 'strike'), 4), (('my', 'pocket'), 4), (('had', 'my'), 4), (('but', 'there'), 4), (('an', 'altogether'), 4), (('once', 'have'), 4), (('was', 'creeping'), 4), (('thousands', 'of'), 4), (('The', 'sky'), 4), (('and', 'thither'), 4), (('we', 'went'), 4), (('down', 'upon'), 4), (('Morlocks', 'had'), 4), (('and', 'from'), 4), (('the', 'Morlocks'), 4), (('that', 'if'), 4), (('array', 'of'), 4), (('old', 'familiar'), 4), (('I', 'stared'), 4), (('days', 'and'), 4), (('there', 'I'), 4), (('sky', 'and'), 4), (('days', 'of'), 4), (('U.S.', 'copyright'), 4), (('copyright', 'law'), 4), (('you', 'do'), 4), (('any', 'other'), 4), (('the', 'phrase'), 4), (('phrase', 'Project'), 4), (('Project', 'Gutenberg'), 4), (('United', 'States.'), 4), (('associated', 'with'), 4), (('laws', 'of'), 4), (('the', 'laws'), 4), (('work', 'or'), 4), (('provide', 'a'), 4), (('*, 'You'), 4), (('by', 'H.'), 3), (('at', 'no'), 3), (('give', 'it'), 3), (('it', 'away'), 3), (('it', 'under'), 3), (('this', 'eBook'), 3), (('or', 'online'), 3), (('online', 'at'), 3), (('PROJECT', 'GUTENBERG'), 3), (('the', 'Golden'), 3), (('The', 'Morlocks'), 3), (('face', 'was'), 3), (('caught', 'the'), 3), (('was', 'that'), 3), (('follow', 'me'), 3), (('I', 'shall'), 3), (('a', 'large'), 3), (('to', 'begin'), 3), (('“I', 'do'), 3), (('for', 'it.'), 3), (('much', 'as'), 3), (('“There', 'I'), 3), (('a', 'moment.'), 3), (('you', 'in'), 3), (('There', 'are'), 3), (('the', 'three'), 3), (('of', 'Space'), 3), (('three', 'dimensions'), 3), (('the', 'beginning'), 3), (('end', 'of'), 3), (('is', 'very'), 3), (('people', 'who'), 3), (('not', 'know'), 3), (('it.', 'It'), 3), (('dimensions', 'of'), 3), (('of', 'Space'), 3), (('hold', 'of'), 3), (('about', 'this'), 3), (('figure', 'of'), 3), (('the', 'thing.'), 3), (('“I', 'think'), 3), (('as', 'one'), 3), (('one', 'who'), 3), (('he', 'said'), 3), (('I', 'do'), 3), (('here', 'is'), 3), (('another', 'at'), 3), (('the', 'proper'), 3), (('and', 'why'), 3), (('And', 'why'), 3), (('we', 'move'), 3), (('as', 'we'), 3), (('But', 'how'), 3), (('man', 'had'), 3), (('cannot', 'move'), 3), (('at', 'all'), 3), (('the', 'present'), 3), (('That', 'is'), 3), (('whole', 'world'), 3), (('but', 'you'), 3), (('that', 'we'), 3), (('we', 'cannot'), 3), (('we', 'have'), 3), (('travel', 'through'), 3), (('the', 'Very'), 3), (('account', 'of'), 3), (('no', 'great'), 3), (('began', 'the'), 3), (('who', 'was'), 3), (('the', 'room'), 3), (('Traveller', 'came'), 3), (('hand', 'was'), 3), (('He', 'took'), 3), (('set', 'it'), 3), (('sat', 'down.'), 3), (('so', 'as'), 3), (('the', 'left.'), 3), (('could', 'have'), 3), (('a', 'machine'), 3), (('machine', 'to'), 3), (('an', 'odd'), 3), (('we', 'had'), 3), (('to', 'understand'), 3), (('and', 'this'), 3), (('seat', 'of'), 3), (('want', 'to'), 3), (('me', 'but'), 3), (('Traveller', 'put'), 3), (('him', 'to'), 3), (('the', 'model'), 3), (('a', 'second'), 3), (('as', 'an'), 3), (('he', 'went'), 3), (('at', 'each'), 3), (('believe', 'that'), 3), (('has', 'travelled'), 3), (('to', 'light'), 3), (('that', 'he'), 3), (('a', 'big'), 3)]

In [55]:

```
unique = set()
for i in range(500):
    unique.add(common[i][0][1])
    unique.add(common[i][0][0])
vocab_unique = { word:i for i, word in enumerate(unique) }
vocab_unique
```

Out[55]:

{'all': 0,
'which': 1,
'table': 2,
'They': 3,
'be': 4,
'part': 5,
'forth': 6,
'their': 7,
'into': 8,
'longer': 9,
'made': 10,
'do': 11,
'like': 12,
'stood': 13,
'may': 14,
'remember': 15,
'Foundation': 16,
'laboratory': 17,
'here': 18,
'very': 19,
'wanted': 20,
'should': 21,
'Sphinx': 22,
'But': 23,
'eyes.': 24,
'Project': 25,
'Then': 26,
'find': 27,
'fancied': 28,
'white': 29,
'pointed': 30,
'sense': 31,
'what': 32,
'with': 33,
'And': 34,
'little': 35,
'heard': 36,
'works': 37,
'it.': 38,
'along': 39,
'you': 40,

'never': 41,
'I': 42,
'States': 43,
'have': 44,
'head': 45,
'instance,': 46,
'am': 47,
'soon': 48,
'might': 49,
'iron': 50,
'Morlocks,': 51,
'so': 52,
'but': 53,
'it,': 54,
'me,': 55,
'great': 56,
'make': 57,
'he': 58,
'other': 59,
'United': 60,
'tell': 61,
'Traveller's': 62,
'Editor': 63,
'copies': 64,
'turned': 65,
'move': 66,
'small': 67,
'an': 68,
'out': 69,
'not': 70,
'under': 71,
'world': 72,
'thing': 73,
'Gutenberg-tm': 74,
'said.': 75,
'door': 76,
'Man,': 77,
'moon': 78,
'got': 79,
'get': 80,
'protected': 81,
'hands': 82,
'saw': 83,
'down': 84,
'me.': 85,
'Man': 86,
'against': 87,
'a': 88,
'upon': 89,
'moment': 90,
'there': 91,
'You': 92,
'whole': 93,
'now': 94,
'while': 95,
'mind': 96,
'his': 97,
'to': 98,
'my': 99,
'people': 100,
'felt': 101,
'machine.': 102,
'know': 103,
'come': 104,
'where': 105,
'then': 106,
'Machine,': 107,
'agreement,': 108,
'appeared': 109,
'Traveller.': 110,
'from': 111,
'such': 112,
'at': 113,
'Morlocks': 114,
'perhaps': 115,
'looked': 116,
'must': 117,
'don't': 118,
'will': 119,
'number': 120,
'time': 121,
'most': 122,
'certain': 123,
'before': 124,
'new': 125,
'human': 126,
'too': 127,
'Gutenberg': 128,
'between': 129,
'round': 130,
'us': 131,
'sudden': 132,
'If': 133,
'Medical': 134,
'instead': 135,
'Green': 136,
'left': 137,
'first': 138,
'or': 139,

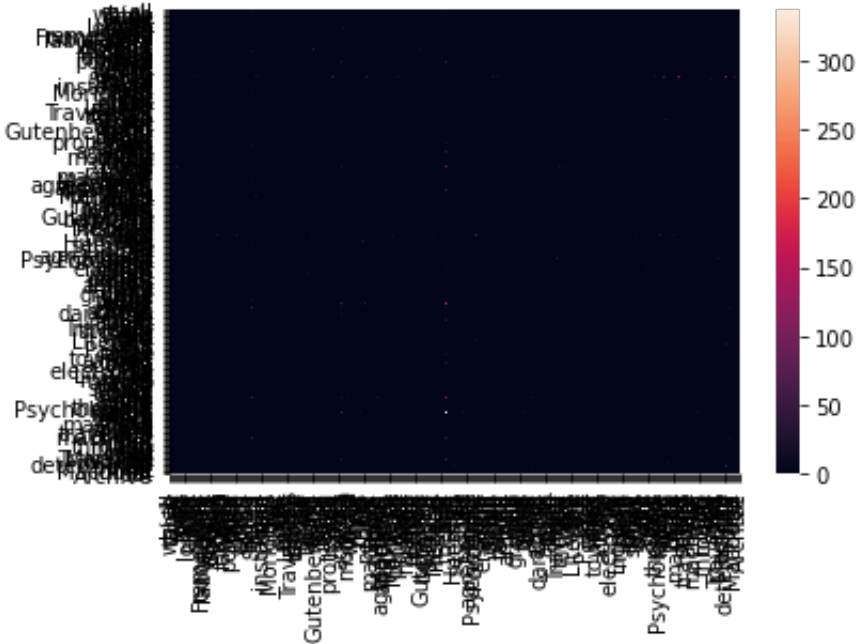
'means': 140,
'the': 141,
'seen': 142,
'one': 143,
'Hundred': 144,
'seemed': 145,
'As': 146,
'on': 147,
'let': 148,
'our': 149,
'In': 150,
'judged': 151,
'agreement': 152,
'knew': 153,
'Psychologist': 154,
'Porcelain': 155,
'Time': 156,
'came': 157,
'side': 158,
'again': 159,
'enough': 160,
'night': 161,
'up': 162,
'would': 163,
'full': 164,
'did': 165,
'going': 166,
'began': 167,
'work': 168,
'telling': 169,
'long': 170,
'among': 171,
'well': 172,
'its': 173,
'man': 174,
'if': 175,
'ground': 176,
'think': 177,
'this': 178,
'signs': 179,
'these': 180,
'when': 181,
'back': 182,
'tried': 183,
'in': 184,
'mind.': 185,
'as': 186,
'across': 187,
'darkness': 188,
'fell': 189,
'is': 190,
'again,': 191,
'than': 192,
'about': 193,
'said': 194,
'Man.': 195,
'end': 196,
'Traveller': 197,
'me': 198,
'though': 199,
'more': 200,
'see': 201,
'old': 202,
'went': 203,
'she': 204,
'Literary': 205,
'The': 206,
'said,': 207,
'own': 208,
'It': 209,
'Palace': 210,
'hill': 211,
'place': 212,
'after': 213,
'last': 214,
'over': 215,
'towards': 216,
'red': 217,
'light': 218,
'them': 219,
'access': 220,
'by': 221,
'go': 222,
'only': 223,
'electronic': 224,
'put': 225,
'look': 226,
'License': 227,
'At': 228,
'minute': 229,
'hand': 230,
'lamp': 231,
'earth': 232,
'kind': 233,
'sky': 234,
'struck': 235,
'set': 236,
'terms': 237,
'away': 238,

```
'some': 239,
'are': 240,
'who': 241,
'and': 242,
'agree': 243,
'since': 244,
'took': 245,
'any': 246,
'thought': 247,
'Psychologist.': 248,
'same': 249,
'future': 250,
'could': 251,
'of': 252,
'were': 253,
'black': 254,
'White': 255,
'no': 256,
'machine': 257,
'how': 258,
'had': 259,
'sound': 260,
'vast': 261,
'travelled': 262,
'There': 263,
'above': 264,
'machine,': 265,
'still': 266,
'lit': 267,
'that': 268,
'sun': 269,
'last,': 270,
'through': 271,
'trying': 272,
'use': 273,
'off': 274,
'been': 275,
'For': 276,
'few': 277,
'Traveller,': 278,
'Machine': 279,
'bronze': 280,
'for': 281,
'determined': 282,
'was': 283,
'they': 284,
'it': 285,
'Machine.': 286,
'found': 287,
'sat': 288,
'Archive': 289}
```

```
In [56]: matrix = np.zeros((len(unique), len(unique)))
for i in range(500):
    current = common[i][0][0]
    previous = common[i][0][1]
    count = common[i][1]
    pos_curr = vocab_unique[current]
    pos_prev = vocab_unique[previous]
    matrix[pos_curr][pos_prev] = count
```

```
In [57]: vmin = 0
vmax = 1000
cmap = 'blues'
xticklabels = vocab_unique.keys()
yticklabels = vocab_unique.keys()
sns.heatmap(data = matrix, xticklabels=xticklabels, yticklabels=yticklabels)
```

Out[57]: <AxesSubplot:>



In []: