

```
In [2]: import pandas as pd
import numpy as np
import textdistance
import re
from collections import Counter
```

```
In [5]: word=[]

with open('Next-World Suggestion.py.txt', mode='r', encoding='utf-8') as f:
    data=f.read()
    data=data.lower()
    words=re.findall('\w+',data)
    words+=words

print(words[0:10])
```

```
['the', 'project', 'gutenberg', 'ebook', 'of', 'moby', 'dick', 'or', 'the', 'whal
e']
```

```
<>:6: SyntaxWarning: invalid escape sequence '\w'
<>:6: SyntaxWarning: invalid escape sequence '\w'
C:\Users\91789\AppData\Local\Temp\ipykernel_13136\1984690620.py:6: SyntaxWarning:
invalid escape sequence '\w'
    words=re.findall('\w+',data)
```

```
In [6]: len(words)
```

```
Out[6]: 445326
```

```
In [7]: V=set(words)
V
```

```
Out[7]: {'infinite',
        'affronted',
        'fluke',
        'torrents',
        'consequence',
        'gnarled',
        'indefatigable',
        'credentials',
        'verse',
        'walruses',
        'harmattans',
        'efforts',
        'prescient',
        'emigrant',
        'occurs',
        'promiscuously',
        'bridal',
        'provincials',
        'exceeded',
        'paternal',
        'leaks',
        'loon',
        'sultanism',
        'carved',
        'twist',
        'simple',
        'frosty',
        'racing',
        '_cruise',
        'habitually',
        'villain',
        'million',
        'fastenings',
        'us',
        'brevet',
        'joe',
        'rented',
        'true',
        'mutiny',
        'arrives',
        'banister',
        'suffered',
        'elaboration',
        'emptying',
        'saco',
        'www',
        'literature',
        'cerebellum',
        'indiamen',
        'appropriate',
        'sultan',
        'disasters',
        'gardening',
        'ambergriese',
        'thee',
        'lowering',
        'steady',
        'likes',
        'roofs',
        'inshore',
```

'upbubble',
'powders',
'capricornus',
'rapidity',
'morass',
'undecided',
'making',
'fares',
'renders',
'cautious',
'settling',
'flailing',
'eminent',
'standpoint',
'endurance',
'preserved',
'community',
'varieties',
'waggish',
'butchers',
'unavoidably',
'110',
'closely',
'observed',
'fins',
'jollity',
'ails',
'habituated',
'1729',
'incumbrance',
'shop',
'flattening',
'angrily',
'east',
'squilgee',
'stabs',
'piteous',
'continent',
'couldst',
'particulars',
'bitterer',
'booting',
'petulance',
'chronicler',
'doltish',
'stigma',
'bone',
'dere',
'australian',
'staggering',
'irrational',
'savannas',
'uneven',
'beggar',
'clotted',
'web',
'stopping',
'connexions',
'grievous',
'alarmed',

'ramblings',
'deer',
'mess',
'surgeons',
'mutely',
'unexampled',
'duskier',
'puissant',
'crouch',
'commander',
'unbutton',
'captured',
'tenth',
'struggle',
'reigneth',
'tangles',
'elijah',
'slices',
'ambitious',
'vapors',
'perpetually',
'_thrasher_',
'copies',
'outburst',
'antwerp',
'roaring',
'discourseth',
'risk',
'inward',
'share',
'cocks',
'declares',
'evaporate',
'fast',
'landscape',
'manikin',
'heart',
'match',
'accidents',
'popping',
'outblown',
'americans',
'commended',
'rugged',
'pedro',
'strike',
'bowing',
'compunctions',
'fated',
'dollars',
'shuffling',
'prove',
'bandages',
'unhesitatingly',
'provoke',
'porpoises',
'fork',
'monadnock',
'apex',
'stray',

'soaked',
'rear',
'instance',
'valour',
'unspoken',
'blazed',
'covenant',
'overdoing',
'fresher',
'application',
'grenadier',
'constitution',
'loo',
'spotted',
'creeps',
'obtains',
'pikes',
'merry',
'envious',
'refuses',
'contradicted',
'ramparts',
'thousands',
'bending',
'49',
'derisive',
'matrimonial',
'notion',
'giddy',
'libertines',
'essences',
'constructed',
'hopelessly',
'insinuating',
'unrecking',
'adroit',
'congenial',
'latently',
'narrated',
'golconda',
'eastward',
'brawlers',
'swathed',
'presumed',
'prefecture',
'gazing',
'cheever',
'cobblestones',
'newspaper',
'c',
'prescription',
'flourishin',
'retires',
'solicited',
'endeavor',
'acts',
'throb',
'unsolved',
'tribulation',
'mobbing',

'detach',
'way',
'scars',
'splash',
'drench',
'operations',
'overawing',
'niger',
'resumed',
'therein',
'yoke',
'expedition',
'draggingly',
'joyously',
'policemen',
'frankfort',
'unwearied',
'scuttling',
'tribute',
'bravest',
'opposite',
'mss',
'lips',
'differs',
'south',
'tartar',
'procure',
'rulers',
'leakage',
'sailors',
'genteel',
'distress',
'literary',
'crossed',
'stake',
'buckets',
'tauntings',
'hired',
'volunteered',
'melted',
'swayed',
'pipe',
'elaborate',
'dismasted',
'centaurs',
'bandied',
'manufacturing',
'subs',
'sunk',
'abundantly',
'gainst',
'umbrellas',
'toss',
'cooled',
'cascading',
'applies',
'arrival',
'catacombs',
'conceal',
'savages',

'leagues',
'bewildered',
'heed',
'libraries',
'manifesto',
'liv',
'plaid',
'impersonal',
'saplings',
'ensued',
'shooks',
'exhumed',
'drills',
'beseeching',
'becomes',
'cervantes',
'pursuing',
'thick',
'lady',
'prepares',
'wakened',
'lilies',
'circumambulate',
'inwreathing',
'miniature',
'shave',
'grievances',
'clearest',
'uncomfortable',
'crime',
'compact',
'springs',
'eastern',
'soul',
'organs',
'pangs',
'reference',
'seasons',
'pursuits',
'platformed',
'papers',
'noting',
'wan',
'costliest',
'requirements',
'confidently',
'extorting',
'narrative_',
'cordially',
'careless',
'ubiquitous',
'tow',
'downward',
'speechless',
'pageant',
'maimed',
'humbug',
'bridegroom',
'arrangement',
'on',

'profit',
'politeness',
'bales',
'swum',
'reinforced',
'arisen',
'materially',
'trough',
'103',
'incantation',
'reward',
'bladder',
'hurried',
'roughly',
'blindest',
'ruin',
'soapstone',
'disembowelled',
'banks',
'wool',
'thousandth',
'futures',
'thrown',
'mobbed',
'colour',
'weighing',
'tasted',
'clouds',
'zone',
'moor',
'insane',
'pell',
'perceive',
'blowing',
'grateful',
'beginning',
'traditions',
'glim',
'premeditated',
'rowed',
'overscorning',
'escaped',
'legatee',
'lightens',
'gnashing',
'miles',
'sonorous',
'glee',
'zodiac',
'tendons',
'idiotic',
'off',
'innocently',
'scholar',
'toasting',
'federated',
'recent',
'unduly',
'labored',
'embrace',

'sighted',
'andromeda',
'obedience',
'mysteriously',
'cibil',
'performing',
'soggy',
'reached',
'palpably',
'mantle',
'wearer',
'spite',
'filaments',
'receptive',
'soap',
'horribles',
'indomitable',
'bobbed',
'jump',
'negligence',
'ka',
'today',
'regina',
'cleanse',
'hazards',
'apprehensions',
'crash',
'transient',
'decently',
'verbally',
'paul',
'lirra',
'underneath',
'captive',
'omnipotent',
'modifications',
'settlement',
'tawny',
'brit',
'loosening',
'transom',
'atmospheres',
'harder',
'learned',
'unsullied',
'surely',
'oh',
'dictator',
'czarship',
'glorying',
'told',
'shout',
'coverlid',
'context',
'_nudging_',
'unmistakable',
'ridiculous',
'germs',
'breast',
'leviathan_',

'jollies',
'illuminated',
'creamy',
'giants',
'davis',
'lifting',
'disposition',
'lock',
'overlays',
'unwarped',
'perished',
'bawl',
'foster',
'turkeys',
'belted',
'wherever',
'farces',
'unflinching',
'sports',
'adjust',
'visiting',
'pencil',
'warrant',
'estate',
'abominate',
'stoneless',
'utterly',
'undulated',
'size',
'pony',
'stem',
'truthfully',
'bowled',
'improving',
'engage',
'physical',
'_thrusts',
'separate',
'jeroboam',
'handspikes',
'bigotry',
'isolated',
'gurgling',
'grip',
'circling',
'fears',
'paregoric',
'uniqueness',
'emetic',
'manner',
'rome',
'nevertheless',
'secrecy',
'tandem',
'unmoor',
'speak',
'judging',
'martial',
'smothered',
'intention',

'shudderings',
'played',
'yarns',
'display',
'accidentally',
'betake',
'peering',
'unmomentous',
'disastrous',
'pulling',
'scorched',
'widely',
'complain',
'ensue',
'insure',
'inseparable',
'unmindful',
'reasons',
'shakespeare',
'quicksand',
'fainter',
'planed',
'indicating',
'din',
'judith',
'electricity',
'railway',
'troubledly',
'immutable',
'port',
'annals_',
'superhuman',
'busks',
'caw',
'grappled',
'leech',
'lightly',
'dromedary',
'dismasting',
'bound',
'lave',
'largely',
'hulls',
'winking',
'quohogs',
'inexpressive',
'wal',
'wonderfulness',
'retiring',
'gall',
'seekest',
'parti',
'limped',
'cabinets',
'endures',
'highwaymen',
'vitality',
'befriended',
'apprised',
'parmacetti',

'explain',
'doleful',
'ahab_',
'azure',
'headsman',
'ghent',
'bitingly',
'seaport',
'curved',
'supper',
'brokenly',
'deliverance',
'_edmund',
'prodigies',
'excellently',
'imputed',
'leeward',
'eating',
'manipulated',
'16th',
'churches',
'rafts',
'pantry',
'khan',
'broiled',
'dusky',
'league',
'heathenish',
'tender',
'hours',
'huff',
'answer',
'indies',
'mow',
'alfred',
'flooded',
'delusion',
'atmospheric',
'as',
'awry',
'marquee',
'flasks',
'stands',
'soldier',
'visitants',
'shocks',
'shaster',
'gudgeons',
'yeast',
'scamp',
'inquiringly',
'heels',
'psalmody',
'welcome',
'supremacy',
'cough',
'1846',
'route',
'mary',
'lash',

'entreaties',
'heavers',
'counterpoise',
'shortness',
'varying',
'circumstances',
'distrustful',
'rituals',
'occasions',
'blind',
'partake',
'witty',
'tarred',
'sheaths',
'supperless',
'cupbearers',
'district',
'meritoque',
'langsдорff',
'draughts',
'embayed',
'unlighted',
'civilization',
'flinging',
'overdone',
'sensible',
'morgana',
'pauses',
'inferred',
'buckler',
'appellations',
'ivory',
'hazel',
'circumspectly',
'maidenly',
'dandy',
'journey',
'hard',
'severed',
'discernment',
'complied',
'maternal',
'deadly',
'broke',
'hovered',
'baleen',
'gnawing',
'_report',
'jeer',
'augmented',
'heedfully',
'described',
'shouts',
'depicting',
'involved',
'fragmentary',
'feminam',
'lurching',
'array',
'weeks',

'upcast',
'chassee',
'oakum',
'afric',
'kin',
'afford',
'ideal',
'colonnades',
'render',
'eminence',
'avers',
'protuberance',
'historical',
'cords',
'withstand',
'covering',
'disagreeable',
'propelled',
'skrimshander',
'bethinking',
'abbreviation',
'weedy',
'towers',
'strawberries',
'bounteous',
'alternate',
'keeled',
'manfully',
'aunt',
'fit',
'dozen',
'deeply',
'selectest',
'princely',
'96',
'battalions',
'clearer',
'ontario',
'voyages_',
'hypos',
'æsthetically',
'sides',
'nations',
'change',
'canaris',
'cousin',
'execrations',
'bashee',
'irregularity',
'robed',
'disinfecting',
'frequently',
'cabaco',
'mutineer',
'repelling',
'commissioned',
'lends',
'murray',
'move',
'chestnuts',

'beefsteaks',
'fishing',
'involve',
'magnet',
'furiously',
'reverie',
'absurdly',
'duff',
'proved',
'controllable',
'quits',
'hundred',
'tangle',
'adhering',
'fickleness',
'formidable',
'fragments',
'worryings',
'delay',
'oarsman',
'firm',
'clearing',
'vehicle',
'behooves',
'between',
'jonas',
'spying',
'graze',
'bubbling',
'plump',
'false',
'greenly',
'maker',
'hereafters',
'obscure',
'tally',
'patronising',
'loosened',
'daybreak',
'defray',
'mightest',
'flowed',
'repaired',
'financial',
'drag',
'teetering',
'affect',
'button',
'circumstanced',
'crim',
'scolds',
'stakes',
'131',
'trio',
'struck',
'chiseled',
'digressively',
'pertaining',
'indian',
'voyagings',

'disguises',
'feegeeeans',
'corroborated',
'choking',
'amazing',
'bunch',
'fleecy',
'cretan',
'crows',
'yielded',
'jollily',
'thoughtfully',
'sich',
'equipped',
'artistic',
'eaves',
'king',
'sobbing',
'd',
'analogies',
'catcher',
'separately',
'trudging',
'lost',
'anew',
'shinbones',
'snows',
'yesterday',
'booble',
'restores',
'mill',
'solicit',
'pursuers',
'domineer',
'utmost',
'serenest',
'abbreviate',
'_what_',
'quote',
'pecks',
'tipping',
'recede',
'batten',
'examined',
'entrapped',
'schoolmaster',
'borrow',
'topic',
'base',
'weakness',
'economical',
'englander',
'catching',
'lanes',
'precedence',
'seated',
'homewards',
'verdigris',
'northward',
'alleged',

'proceeding',
'exhibited',
'black',
'shoemaker',
'epicures',
'bower',
'some',
'contract',
'car',
'imperiously',
'glue',
'boys',
'distended',
'rivals',
'butt',
'delightfully',
'approach',
'sultanically',
'commandment',
'beholds',
'deserts',
'evangelist',
'writhed',
'bought',
'hastily',
'crucified',
'slights',
'permitted',
'delineations',
'uttered',
'inhabitant',
'suburban',
'scrolls',
'unintelligent',
'bounce',
'wheel',
'extremities',
'trackless',
'inscrutably',
'unyielding',
'painter',
'webbed',
'helping',
'quite',
'whispering',
'peels',
'collection',
'tubs',
'pondered',
'visibly',
'mayest',
'critters',
'quoggy',
'comstock_',
'mails',
'catch',
'infernal',
'tongs',
'formally',
'forsake',

```
'awakened',  
'manliest',  
'dollar',  
'glad',  
'boomer',  
'ixion',  
'misdoubt',  
'constituting',  
'wrapping',  
'rejecting',  
'kitchen',  
'snuff',  
'hails',  
'unnamable',  
'decanter',  
'shifting',  
'unsuspected',  
'avenues',  
'forth',  
'dubiously',  
'tornado',  
'lounging',  
'attenuated',  
'anacharsis',  
'tables',  
'scorpio',  
'moderately',  
'driven',  
'mapped',  
'callao',  
'accompanies',  
'capable',  
'folks',  
'subdued',  
'raise',  
'pig',  
'unabated',  
'snapped',  
'hereabouts',  
'bawling',  
...}
```

```
In [8]: len(V)
```

```
Out[8]: 17647
```

```
In [9]: word_freq={}  
word_freq=Counter(words)
```

```
In [10]: word_freq.most_common(10)
```

```
Out[10]: [('the', 29406),
          ('of', 13484),
          ('and', 13034),
          ('a', 9598),
          ('to', 9414),
          ('in', 8476),
          ('that', 6162),
          ('it', 5068),
          ('his', 5060),
          ('i', 4240)]
```

```
In [11]: Total_words=sum(word_freq.values())

probs={}
for k in word_freq.keys():
    probs[k]=word_freq[k]/Total_words
```

```
In [12]: probs
```

```
Out[12]: {'the': 0.06603252448767868,
'project': 0.0004086893646452262,
'gutenberg': 0.0004221626404027611,
'ebook': 4.4910919191783095e-05,
'of': 0.030278941719100165,
'moby': 0.0004041982727260479,
'dick': 0.0004041982727260479,
'or': 0.003579400259585113,
'whale': 0.005524043060589321,
'by': 0.005488114325235894,
'herman': 1.796436767671324e-05,
'melville': 1.796436767671324e-05,
'this': 0.006462681271697588,
'is': 0.007863901950481221,
'for': 0.007383355115129142,
'use': 0.0002200635040397372,
'anyone': 2.694655151506986e-05,
'anywhere': 7.185747070685296e-05,
'at': 0.005995607712103043,
'no': 0.002667708599991916,
'cost': 1.796436767671324e-05,
'and': 0.029268446037285047,
'with': 0.00794474160502643,
'almost': 0.000884745108078127,
'restrictions': 8.98218383835662e-06,
'whatsoever': 3.143764343424817e-05,
'you': 0.004302466058572821,
'may': 0.001145228439390469,
'copy': 8.533074646438789e-05,
'it': 0.011380426923197837,
'give': 0.0004041982727260479,
'away': 0.0008353430969671656,
're': 8.083965454520958e-05,
'under': 0.0005658775818164671,
'terms': 0.00014820603333288423,
'license': 8.083965454520958e-05,
'included': 6.287528686849633e-05,
'online': 1.796436767671324e-05,
'www': 2.694655151506986e-05,
'org': 5.838419494931803e-05,
'title': 3.592873535342648e-05,
'author': 4.041982727260479e-05,
'release': 4.49109191917831e-06,
'date': 1.796436767671324e-05,
'december': 2.2455459595891548e-05,
'25': 1.796436767671324e-05,
'2008': 4.49109191917831e-06,
'2701': 1.796436767671324e-05,
'last': 0.00124852355353157,
'updated': 8.98218383835662e-06,
'3': 7.185747070685296e-05,
'2017': 4.49109191917831e-06,
'language': 3.143764343424817e-05,
'english': 0.0002200635040397372,
'character': 6.736637878767465e-05,
'set': 0.0003817428131301563,
'encoding': 4.49109191917831e-06,
'utf': 4.49109191917831e-06,
'8': 3.143764343424817e-05,
'start': 0.00016617040100959745,
```

'produced': 4.041982727260479e-05,
'daniel': 1.796436767671324e-05,
'lazarus': 3.143764343424817e-05,
'jonesey': 8.98218383835662e-06,
'david': 8.98218383835662e-06,
'widger': 8.98218383835662e-06,
'contents': 3.592873535342648e-05,
'etymology': 8.98218383835662e-06,
'extracts': 3.143764343424817e-05,
'supplied': 5.838419494931803e-05,
'a': 0.02155275012013671,
'sub': 4.041982727260479e-05,
'librarian': 8.98218383835662e-06,
'chapter': 0.0013832563111069193,
'1': 0.0002200635040397372,
'loomings': 8.98218383835662e-06,
'2': 3.143764343424817e-05,
'carpet': 4.041982727260479e-05,
'bag': 0.00010329511414110113,
'spouter': 3.592873535342648e-05,
'inn': 7.634856262603126e-05,
'4': 4.041982727260479e-05,
'counterpane': 3.592873535342648e-05,
'5': 3.143764343424817e-05,
'breakfast': 6.736637878767465e-05,
'6': 1.796436767671324e-05,
'street': 6.287528686849633e-05,
'7': 3.143764343424817e-05,
'chapel': 4.940201111096141e-05,
'pulpit': 7.634856262603126e-05,
'9': 2.2455459595891548e-05,
'sermon': 2.2455459595891548e-05,
'10': 2.2455459595891548e-05,
'bosom': 4.041982727260479e-05,
'friend': 0.00012125948181781436,
'11': 8.98218383835662e-06,
'nightgown': 8.98218383835662e-06,
'12': 8.98218383835662e-06,
'biographical': 8.98218383835662e-06,
'13': 1.796436767671324e-05,
'wheelbarrow': 1.796436767671324e-05,
'14': 8.98218383835662e-06,
'nantucket': 0.0004221626404027611,
'15': 8.98218383835662e-06,
'chowder': 4.4910919191783095e-05,
'16': 8.98218383835662e-06,
'ship': 0.002330876706053543,
'17': 8.98218383835662e-06,
'ramadan': 5.389310303013972e-05,
'18': 8.98218383835662e-06,
'his': 0.011362462555521124,
'mark': 0.00025150114747398534,
'19': 8.98218383835662e-06,
'prophet': 7.634856262603126e-05,
'20': 2.2455459595891548e-05,
'all': 0.006934245923211311,
'astir': 1.796436767671324e-05,
'21': 8.98218383835662e-06,
'going': 0.0004041982727260479,
'aboard': 9.880402222192281e-05,

'22': 8.98218383835662e-06,
'merry': 6.287528686849633e-05,
'christmas': 2.694655151506986e-05,
'23': 8.98218383835662e-06,
'lee': 9.431293030274451e-05,
'shore': 0.00011676838989863606,
'24': 8.98218383835662e-06,
'advocate': 1.796436767671324e-05,
'postscript': 8.98218383835662e-06,
'26': 8.98218383835662e-06,
'knights': 2.694655151506986e-05,
'squires': 1.796436767671324e-05,
'27': 8.98218383835662e-06,
'28': 8.98218383835662e-06,
'ahab': 0.0022994390626192946,
'29': 8.98218383835662e-06,
'enter': 4.940201111096141e-05,
'to': 0.021139569663572305,
'him': 0.0047830128939249,
'stubb': 0.0011631928070671821,
'30': 1.796436767671324e-05,
'pipe': 0.00018862586060548902,
'31': 8.98218383835662e-06,
'queen': 8.083965454520958e-05,
'mab': 8.98218383835662e-06,
'32': 8.98218383835662e-06,
'cetology': 4.940201111096141e-05,
'33': 8.98218383835662e-06,
'specksnyder': 1.796436767671324e-05,
'34': 8.98218383835662e-06,
'cabin': 0.00039072499696851296,
'table': 0.00017066149292877577,
'35': 8.98218383835662e-06,
'mast': 0.0005838419494931802,
'head': 0.0015628999878740518,
'36': 8.98218383835662e-06,
'quarter': 0.00021557241212055887,
'deck': 0.0008937272919164837,
'37': 8.98218383835662e-06,
'sunset': 4.4910919191783095e-05,
'38': 8.98218383835662e-06,
'dusk': 1.347327575753493e-05,
'39': 8.98218383835662e-06,
'first': 0.001073370968683616,
'night': 0.0006826459717151031,
'watch': 0.0002739566070698769,
'40': 1.347327575753493e-05,
'midnight': 0.0001347327575753493,
'forecastle': 0.00017066149292877577,
'41': 8.98218383835662e-06,
'42': 8.98218383835662e-06,
'whiteness': 0.00012575057373699267,
'43': 8.98218383835662e-06,
'hark': 7.634856262603126e-05,
'44': 8.98218383835662e-06,
'chart': 3.592873535342648e-05,
'45': 1.347327575753493e-05,
'affidavit': 1.347327575753493e-05,
'46': 8.98218383835662e-06,
'surmises': 1.796436767671324e-05,

'47': 8.98218383835662e-06,
'mat': 5.838419494931803e-05,
'maker': 2.2455459595891548e-05,
'48': 8.98218383835662e-06,
'lowering': 8.533074646438789e-05,
'49': 8.98218383835662e-06,
'hyena': 1.796436767671324e-05,
'50': 1.347327575753493e-05,
's': 0.008169296200985346,
'boat': 0.0015000247010055554,
'crew': 0.0006332439606041417,
'fedallah': 0.00012575057373699267,
'51': 8.98218383835662e-06,
'spirit': 8.533074646438789e-05,
'spout': 0.0002739566070698769,
'52': 8.98218383835662e-06,
'albatross': 4.4910919191783095e-05,
'53': 8.98218383835662e-06,
'gam': 2.694655151506986e-05,
'54': 8.98218383835662e-06,
'town': 0.00016167930909041916,
'ho': 0.00014371494141370592,
'story': 0.0002290456878780938,
'55': 8.98218383835662e-06,
'monstrous': 5.389310303013972e-05,
'pictures': 8.533074646438789e-05,
'whales': 0.001217085910097322,
'56': 8.98218383835662e-06,
'less': 0.00024251896363562873,
'erroneous': 8.98218383835662e-06,
'true': 0.00039521608888769125,
'whaling': 0.0005928241333315369,
'scenes': 5.389310303013972e-05,
'57': 8.98218383835662e-06,
'in': 0.019033247553477677,
'paint': 3.143764343424817e-05,
'teeth': 0.00022455459595891548,
'wood': 0.00016617040100959745,
'sheet': 4.041982727260479e-05,
'iron': 0.00039521608888769125,
'stone': 8.982183838356619e-05,
'mountains': 6.287528686849633e-05,
'stars': 6.287528686849633e-05,
'58': 8.98218383835662e-06,
'brit': 4.041982727260479e-05,
'59': 8.98218383835662e-06,
'squid': 3.592873535342648e-05,
'60': 1.796436767671324e-05,
'line': 0.0007185747070685295,
'61': 8.98218383835662e-06,
'kills': 1.347327575753493e-05,
'62': 8.98218383835662e-06,
'dart': 0.00011676838989863606,
'63': 8.98218383835662e-06,
'crotch': 4.940201111096141e-05,
'64': 1.347327575753493e-05,
'supper': 0.0001392238494945276,
'65': 8.98218383835662e-06,
'as': 0.007868393042400399,
'dish': 4.940201111096141e-05,

'66': 8.98218383835662e-06,
'shark': 0.0001392238494945276,
'massacre': 1.347327575753493e-05,
'67': 8.98218383835662e-06,
'cutting': 0.00017066149292877577,
'68': 8.98218383835662e-06,
'blanket': 6.287528686849633e-05,
'69': 8.98218383835662e-06,
'funeral': 2.694655151506986e-05,
'70': 8.98218383835662e-06,
'sphinx': 1.347327575753493e-05,
'71': 8.98218383835662e-06,
'jeroboam': 4.940201111096141e-05,
'72': 1.347327575753493e-05,
'monkey': 7.634856262603126e-05,
'rope': 0.0002694655151506986,
'73': 8.98218383835662e-06,
'flask': 0.0004895290191904357,
'kill': 0.00015269712525206252,
'right': 0.000696119247472638,
'then': 0.002829387909082335,
'have': 0.0034761051454440116,
'talk': 0.00011676838989863606,
'over': 0.001841347686863107,
'74': 8.98218383835662e-06,
'sperm': 0.001086844244441151,
'contrasted': 1.796436767671324e-05,
'view': 0.00022455459595891548,
'75': 8.98218383835662e-06,
'76': 8.98218383835662e-06,
'battering': 3.143764343424817e-05,
'ram': 3.592873535342648e-05,
'77': 8.98218383835662e-06,
'great': 0.0013787652191877412,
'heidelberg': 4.041982727260479e-05,
'tun': 6.736637878767465e-05,
'78': 8.98218383835662e-06,
'cistern': 2.2455459595891548e-05,
'buckets': 2.694655151506986e-05,
'79': 8.98218383835662e-06,
'prairie': 5.838419494931803e-05,
'80': 8.98218383835662e-06,
'nut': 2.2455459595891548e-05,
'81': 8.98218383835662e-06,
'pequod': 0.0007949232696945609,
'meets': 6.287528686849633e-05,
'virgin': 5.389310303013972e-05,
'82': 8.98218383835662e-06,
'honor': 0.00011227729797945774,
'glory': 8.533074646438789e-05,
'83': 8.98218383835662e-06,
'jonah': 0.0003817428131301563,
'historically': 8.98218383835662e-06,
'regarded': 0.00011227729797945774,
'84': 8.98218383835662e-06,
'pitchpoling': 3.143764343424817e-05,
'85': 8.98218383835662e-06,
'fountain': 4.4910919191783095e-05,
'86': 8.98218383835662e-06,
'tail': 0.0003637784454534431,

'87': 8.98218383835662e-06,
'grand': 0.00024251896363562873,
'armada': 8.98218383835662e-06,
'88': 8.98218383835662e-06,
'schools': 4.041982727260479e-05,
'schoolmasters': 1.347327575753493e-05,
'89': 8.98218383835662e-06,
'fast': 0.00028742988282741183,
'fish': 0.0007589945343411344,
'loose': 0.00016617040100959745,
'90': 1.796436767671324e-05,
'heads': 0.00039072499696851296,
'tails': 3.592873535342648e-05,
'91': 8.98218383835662e-06,
'rose': 0.00015718821717124084,
'bud': 2.694655151506986e-05,
'92': 8.98218383835662e-06,
'ambergris': 6.287528686849633e-05,
'93': 8.98218383835662e-06,
'castaway': 2.2455459595891548e-05,
'94': 8.98218383835662e-06,
'squeeze': 4.4910919191783095e-05,
'hand': 0.0009655847626233366,
'95': 8.98218383835662e-06,
'cassock': 1.347327575753493e-05,
'96': 8.98218383835662e-06,
'try': 0.00024701005555480705,
'works': 0.00028742988282741183,
'97': 8.98218383835662e-06,
'lamp': 0.0001347327575753493,
'98': 8.98218383835662e-06,
'stowing': 1.347327575753493e-05,
'down': 0.0017021238373685795,
'clearing': 2.694655151506986e-05,
'up': 0.002362314349487791,
'99': 8.98218383835662e-06,
'doubloon': 9.431293030274451e-05,
'100': 1.796436767671324e-05,
'leg': 0.0004086893646452262,
'arm': 0.0004041982727260479,
'101': 8.98218383835662e-06,
'decanter': 1.347327575753493e-05,
'102': 8.98218383835662e-06,
'bower': 1.796436767671324e-05,
'arsacides': 1.796436767671324e-05,
'103': 8.98218383835662e-06,
'measurement': 1.796436767671324e-05,
'skeleton': 0.00015718821717124084,
'104': 8.98218383835662e-06,
'fossil': 2.2455459595891548e-05,
'105': 8.98218383835662e-06,
'does': 0.0004401270080794744,
'magnitude': 0.00010329511414110113,
'diminish': 2.2455459595891548e-05,
'will': 0.0017919456757521455,
'he': 0.008515110278762075,
'perish': 2.2455459595891548e-05,
'106': 8.98218383835662e-06,
'107': 8.98218383835662e-06,
'carpenter': 0.00022455459595891548,

'108': 8.98218383835662e-06,
'109': 8.98218383835662e-06,
'starbuck': 0.0008937272919164837,
'110': 8.98218383835662e-06,
'queequag': 0.0011362462555521123,
'coffin': 0.0002380278717164504,
'111': 8.98218383835662e-06,
'pacific': 0.00015718821717124084,
'112': 8.98218383835662e-06,
'blacksmith': 9.880402222192281e-05,
'113': 8.98218383835662e-06,
'forge': 5.838419494931803e-05,
'114': 8.98218383835662e-06,
'gilder': 8.98218383835662e-06,
'115': 8.98218383835662e-06,
'bachelor': 4.041982727260479e-05,
'116': 8.98218383835662e-06,
'dying': 6.736637878767465e-05,
'117': 8.98218383835662e-06,
'118': 8.98218383835662e-06,
'quadrant': 4.041982727260479e-05,
'119': 8.98218383835662e-06,
'candles': 4.4910919191783095e-05,
'120': 8.98218383835662e-06,
'towards': 0.0005164755707055056,
'end': 0.0004625824676753659,
'121': 8.98218383835662e-06,
'bulwarks': 0.00017066149292877577,
'122': 8.98218383835662e-06,
'aloft': 0.00028293879090823354,
'thunder': 0.00014371494141370592,
'lightning': 0.00012575057373699267,
'123': 8.98218383835662e-06,
'musket': 4.041982727260479e-05,
'124': 8.98218383835662e-06,
'needle': 5.389310303013972e-05,
'125': 8.98218383835662e-06,
'log': 0.00010329511414110113,
'126': 8.98218383835662e-06,
'life': 0.0007859410858562043,
'buoy': 6.287528686849633e-05,
'127': 8.98218383835662e-06,
'128': 8.98218383835662e-06,
'rachel': 3.143764343424817e-05,
'129': 8.98218383835662e-06,
'130': 8.98218383835662e-06,
'hat': 0.0001841347686863107,
'131': 8.98218383835662e-06,
'delight': 8.533074646438789e-05,
'132': 8.98218383835662e-06,
'symphony': 8.98218383835662e-06,
'133': 8.98218383835662e-06,
'chase': 0.0002784476989890552,
'day': 0.0008039054535329174,
'134': 8.98218383835662e-06,
'second': 0.0002784476989890552,
'135': 8.98218383835662e-06,
'third': 0.00015269712525206252,
'epilogue': 8.98218383835662e-06,
'original': 0.00010329511414110113,

'transcriber': 4.49109191917831e-06,
'notes': 1.796436767671324e-05,
'text': 2.694655151506986e-05,
'combination': 8.98218383835662e-06,
'etexts': 4.49109191917831e-06,
'one': 0.004149768933320758,
'from': 0.004958165478772854,
'now': 0.0035299982484741513,
'defunct': 8.98218383835662e-06,
'eris': 4.49109191917831e-06,
'virginia': 2.694655151506986e-05,
'tech': 8.98218383835662e-06,
'archives': 4.49109191917831e-06,
'proofreaders': 4.49109191917831e-06,
'version': 2.694655151506986e-05,
'are': 0.0027754948060521954,
'indebted': 8.98218383835662e-06,
'university': 4.49109191917831e-06,
'adelaide': 4.49109191917831e-06,
'library': 1.347327575753493e-05,
'preserving': 1.796436767671324e-05,
'resulting': 1.347327575753493e-05,
'etext': 4.49109191917831e-06,
'was': 0.007392337298967498,
'compared': 4.041982727260479e-05,
'public': 6.736637878767465e-05,
'domain': 4.041982727260479e-05,
'hard': 0.00037725172121097803,
'late': 0.0001347327575753493,
'consumptive': 4.49109191917831e-06,
'usher': 8.98218383835662e-06,
'grammar': 8.98218383835662e-06,
'school': 4.4910919191783095e-05,
'pale': 8.533074646438789e-05,
'threadbare': 4.49109191917831e-06,
'coat': 0.00012575057373699267,
'heart': 0.0004086893646452262,
'body': 0.0004940201111096141,
'brain': 0.00016617040100959745,
'i': 0.009521114868658016,
'see': 0.0012350502777740352,
'ever': 0.0009251649353507318,
'dusting': 8.98218383835662e-06,
'old': 0.002029973547468596,
'lexicons': 4.49109191917831e-06,
'grammars': 8.98218383835662e-06,
'queer': 0.00019760804444384563,
'handkerchief': 2.2455459595891548e-05,
'mockingly': 4.49109191917831e-06,
'embellished': 1.347327575753493e-05,
'gay': 9.431293030274451e-05,
'flags': 4.49109191917831e-06,
'known': 0.00035928735353426476,
'nations': 5.389310303013972e-05,
'world': 0.0007904321777753825,
'loved': 1.347327575753493e-05,
'dust': 4.4910919191783095e-05,
'somehow': 0.00019760804444384563,
'mildly': 4.4910919191783095e-05,
'reminded': 1.796436767671324e-05,

'mortality': 4.49109191917831e-06,
'while': 0.0011048086121178643,
'take': 0.0006152795929274284,
'others': 0.00017964367676713238,
'teach': 2.2455459595891548e-05,
'them': 0.002128777569690519,
'what': 0.0027710037141330172,
'name': 0.0003143764343424817,
'be': 0.004774030710086544,
'called': 0.000520966662624684,
'our': 0.0009476203949466234,
'tongue': 4.940201111096141e-05,
'leaving': 0.00017066149292877577,
'out': 0.002411716360598752,
'through': 0.0010554066010069027,
'ignorance': 4.940201111096141e-05,
'letter': 7.185747070685296e-05,
'h': 1.347327575753493e-05,
'which': 0.002941665207061793,
'alone': 0.00018862586060548902,
'maketh': 2.2455459595891548e-05,
'signification': 4.49109191917831e-06,
'word': 0.00034581407777672983,
'deliver': 1.347327575753493e-05,
'that': 0.013837054202988372,
'not': 0.005268050821196158,
'_hackluyt': 4.49109191917831e-06,
'_': 0.0002200635040397372,
'sw': 4.49109191917831e-06,
'dan': 3.143764343424817e-05,
'_hval_': 4.49109191917831e-06,
'animal': 7.185747070685296e-05,
'named': 5.389310303013972e-05,
'roundness': 4.49109191917831e-06,
'rolling': 0.00015718821717124084,
'_hvalt_': 4.49109191917831e-06,
'arched': 4.940201111096141e-05,
'vaulted': 8.98218383835662e-06,
'_webster': 4.49109191917831e-06,
'dictionary': 1.796436767671324e-05,
'more': 0.002276983603023403,
'immediately': 7.634856262603126e-05,
'dut': 4.49109191917831e-06,
'ger': 4.49109191917831e-06,
'_wallen_': 4.49109191917831e-06,
'_walw': 4.49109191917831e-06,
'ian_': 4.49109191917831e-06,
'roll': 8.083965454520958e-05,
'wallow': 1.347327575753493e-05,
'_richardson': 4.49109191917831e-06,
'4.49109191917831 : 'me-06,
'_hebrew_': 4.49109191917831e-06,
'κητος': 4.49109191917831e-06,
'_greek_': 4.49109191917831e-06,
'cetus': 1.347327575753493e-05,
'_latin_': 4.49109191917831e-06,
'whæl': 4.49109191917831e-06,
'_anglo': 4.49109191917831e-06,
'saxon_': 4.49109191917831e-06,
'hvalt': 4.49109191917831e-06,

'_danish_': 4.49109191917831e-06,
'wal': 4.49109191917831e-06,
'_dutch_': 4.49109191917831e-06,
'hwal': 4.49109191917831e-06,
'_swedish_': 4.49109191917831e-06,
'_icelandic_': 4.49109191917831e-06,
'_english_': 4.49109191917831e-06,
'ballena': 4.49109191917831e-06,
'_spanish_': 4.49109191917831e-06,
'pekee': 4.49109191917831e-06,
'nuee': 1.796436767671324e-05,
'_fegee_': 4.49109191917831e-06,
'pehee': 4.49109191917831e-06,
'_erromangoan_': 4.49109191917831e-06,
'seen': 0.0007410301666644211,
'mere': 0.00018862586060548902,
'painstaking': 4.49109191917831e-06,
'burrower': 4.49109191917831e-06,
'grub': 8.98218383835662e-06,
'worm': 2.2455459595891548e-05,
'poor': 0.0005119844787863274,
'devil': 0.00024701005555480705,
'appears': 3.592873535342648e-05,
'gone': 0.00026497442323152027,
'long': 0.0014910425171671988,
'vaticans': 4.49109191917831e-06,
'stalls': 4.49109191917831e-06,
'earth': 0.00020659022828220226,
'picking': 1.796436767671324e-05,
'whatever': 0.00020659022828220226,
'random': 2.2455459595891548e-05,
'allusions': 3.143764343424817e-05,
'could': 0.0009745669464616933,
'anyways': 8.98218383835662e-06,
'find': 0.00025150114747398534,
'any': 0.0016347574585809047,
'book': 0.0002694655151506986,
'sacred': 4.4910919191783095e-05,
'profane': 2.2455459595891548e-05,
'therefore': 0.00030090315858494676,
'must': 0.0013158899323192448,
'every': 0.001041933325249368,
'case': 0.0003233586181808383,
'least': 0.00034132298585755154,
'higgledy': 4.49109191917831e-06,
'piggledy': 4.49109191917831e-06,
'statements': 2.2455459595891548e-05,
'however': 0.00043114482424111775,
'authentic': 1.347327575753493e-05,
'these': 0.0018188922272672156,
'veritable': 1.796436767671324e-05,
'gospel': 8.98218383835662e-06,
'far': 0.0007410301666644211,
'touching': 0.00020659022828220226,
'ancient': 0.00012575057373699267,
'authors': 1.796436767671324e-05,
'generally': 0.0001392238494945276,
'well': 0.0010329511414110112,
'poets': 1.347327575753493e-05,
'here': 0.001437149414137059,

'appearing': 1.796436767671324e-05,
'solely': 2.694655151506986e-05,
'valuable': 4.940201111096141e-05,
'entertaining': 8.98218383835662e-06,
'affording': 8.98218383835662e-06,
'glancing': 6.287528686849633e-05,
'bird': 7.634856262603126e-05,
'eye': 0.00039521608888769125,
'has': 0.0013158899323192448,
'been': 0.0018638031464589986,
'promiscuously': 4.49109191917831e-06,
'said': 0.0013697830353493845,
'thought': 0.0006736637878767465,
'fancied': 4.041982727260479e-05,
'sung': 8.98218383835662e-06,
'leviathan': 0.00042665373232193945,
'many': 0.0007545034424219561,
'generations': 3.592873535342648e-05,
'including': 7.185747070685296e-05,
'own': 0.0009206738434315535,
'so': 0.0047875039858440785,
'fare': 4.4910919191783095e-05,
'thee': 0.0005883330414123586,
'whose': 0.00039072499696851296,
'commentator': 4.49109191917831e-06,
'am': 0.0004041982727260479,
'thou': 0.001217085910097322,
'belongest': 4.49109191917831e-06,
'hopeless': 3.592873535342648e-05,
'sallow': 4.49109191917831e-06,
'tribe': 4.041982727260479e-05,
'wine': 5.389310303013972e-05,
'warm': 0.00010778620606027944,
'whom': 0.00020209913636302394,
'even': 0.0008667807404014138,
'sherry': 4.49109191917831e-06,
'would': 0.0019401517090850298,
'too': 0.0008308520050479873,
'rosy': 8.98218383835662e-06,
'strong': 0.00016167930909041916,
'but': 0.008187260568662058,
'sometimes': 0.00039072499696851296,
'loves': 4.49109191917831e-06,
'sit': 6.736637878767465e-05,
'feel': 0.0002559922393931637,
'devilish': 4.041982727260479e-05,
'grow': 0.00010778620606027944,
'convivial': 1.796436767671324e-05,
'upon': 0.00255094021009328,
'tears': 2.694655151506986e-05,
'say': 0.0010958264282795076,
'bluntly': 4.49109191917831e-06,
'full': 0.0006107885010082501,
'eyes': 0.0007006103393918164,
'empty': 9.880402222192281e-05,
'glasses': 4.4910919191783095e-05,
'altogether': 0.000130241665656171,
'unpleasant': 1.347327575753493e-05,
'sadness': 8.98218383835662e-06,
'subs': 4.49109191917831e-06,

'how': 0.001158701715148004,
'much': 0.0010060045898959414,
'pains': 6.736637878767465e-05,
'ye': 0.002119795385852162,
'please': 7.185747070685296e-05,
'shall': 0.00043563591616029604,
'go': 0.0008667807404014138,
'thankless': 4.49109191917831e-06,
'clear': 0.0001841347686863107,
'hampton': 4.49109191917831e-06,
'court': 8.98218383835662e-06,
'tuileries': 1.347327575753493e-05,
'gulp': 8.98218383835662e-06,
'your': 0.0011542106232288257,
'hie': 4.49109191917831e-06,
'royal': 0.00012575057373699267,
'hearts': 0.000130241665656171,
'friends': 7.634856262603126e-05,
'who': 0.0016437396424192614,
'before': 0.0013518186676726713,
'seven': 7.185747070685296e-05,
'storied': 1.796436767671324e-05,
'heavens': 7.634856262603126e-05,
'making': 0.00020659022828220226,
'refugees': 4.49109191917831e-06,
'pampered': 4.49109191917831e-06,
'gabriel': 8.982183838356619e-05,
'michael': 1.347327575753493e-05,
'raphael': 4.49109191917831e-06,
'against': 0.0006062974090890719,
'coming': 0.00024251896363562873,
'strike': 0.00014371494141370592,
'splintered': 3.143764343424817e-05,
'together': 0.00028742988282741183,
'there': 0.0038847945100892378,
'unsplinterable': 4.49109191917831e-06,
'god': 0.0006871370636342814,
'created': 4.041982727260479e-05,
'_genesis_': 4.49109191917831e-06,
'path': 4.4910919191783095e-05,
'shine': 1.347327575753493e-05,
'after': 0.0012125948181781437,
'think': 0.0005479132141397538,
'deep': 0.0003143764343424817,
'hoary': 8.98218383835662e-06,
'_job_': 4.49109191917831e-06,
'lord': 0.0002919209747465901,
'had': 0.0034985606050399035,
'prepared': 5.389310303013972e-05,
'swallow': 4.940201111096141e-05,
'_jonah_': 4.49109191917831e-06,
'ships': 0.00039521608888769125,
'hast': 0.000130241665656171,
'made': 0.0007994143616137392,
'play': 0.000130241665656171,
'therein': 2.694655151506986e-05,
'_psalms_': 4.49109191917831e-06,
'sore': 1.347327575753493e-05,
'sword': 0.00011227729797945774,
'punish': 4.49109191917831e-06,

'piercing': 1.796436767671324e-05,
'serpent': 1.796436767671324e-05,
'crooked': 3.592873535342648e-05,
'slay': 1.796436767671324e-05,
'dragon': 2.694655151506986e-05,
'sea': 0.002043446823226131,
'_isaiah_': 4.49109191917831e-06,
'thing': 0.0008443252808055222,
'soever': 4.49109191917831e-06,
'besides': 0.00022455459595891548,
'cometh': 4.49109191917831e-06,
'within': 0.0003727606292917997,
'chaos': 1.796436767671324e-05,
'monster': 0.0002200635040397372,
'mouth': 0.0002739566070698769,
'beast': 1.796436767671324e-05,
'goes': 0.00024701005555480705,
'incontinently': 4.49109191917831e-06,
'foul': 4.940201111096141e-05,
'perisheth': 4.49109191917831e-06,
'bottomless': 3.143764343424817e-05,
'gulf': 2.694655151506986e-05,
'paunch': 8.98218383835662e-06,
'_holland': 8.98218383835662e-06,
'plutarch': 4.49109191917831e-06,
'morals_': 4.49109191917831e-06,
'indian': 0.0002559922393931637,
'breedeth': 4.49109191917831e-06,
'most': 0.0012799611969658183,
'biggest': 8.98218383835662e-06,
'fishes': 2.2455459595891548e-05,
'among': 0.0007500123505027777,
'whirlpooles': 4.49109191917831e-06,
'balaene': 4.49109191917831e-06,
'length': 0.0003503051696959082,
'four': 0.0003323408020191949,
'acres': 1.347327575753493e-05,
'arpens': 4.49109191917831e-06,
'land': 0.00035928735353426476,
'pliny_': 4.49109191917831e-06,
'scarcely': 4.041982727260479e-05,
'we': 0.0021063221100946273,
'proceeded': 2.694655151506986e-05,
'two': 0.0013293632080767796,
'days': 0.00038623390504933467,
'on': 0.004818941629278327,
'when': 0.002726092794941234,
'about': 0.0014281672302987026,
'sunrise': 4.940201111096141e-05,
'other': 0.0019311695252466731,
'monsters': 4.041982727260479e-05,
'appeared': 4.4910919191783095e-05,
'former': 8.982183838356619e-05,
'size': 5.389310303013972e-05,
'came': 0.0005838419494931802,
'us': 0.0010284600494918329,
'open': 0.00034132298585755154,
'mouthed': 2.2455459595891548e-05,
'raising': 2.2455459595891548e-05,
'waves': 0.00021557241212055887,

'sides': 0.00012575057373699267,
'beating': 4.940201111096141e-05,
'into': 0.0023488410737302562,
'foam': 8.533074646438789e-05,
'_tooke': 4.49109191917831e-06,
'lucian_': 4.49109191917831e-06,
'_the': 5.389310303013972e-05,
'history_': 4.49109191917831e-06,
'visited': 3.143764343424817e-05,
'country': 0.00012125948181781436,
'also': 0.0004176715484835828,
'catching': 3.143764343424817e-05,
'horse': 0.00012125948181781436,
'bones': 0.0002200635040397372,
'very': 0.001450622689894594,
'value': 2.694655151506986e-05,
'their': 0.002784476989890552,
'brought': 0.00017066149292877577,
'some': 0.002779985897971374,
'king': 0.00028293879090823354,
'best': 0.00029641206666576847,
'were': 0.0030674157807987854,
'caught': 4.49109191917831e-06,
'forty': 0.00014820603333288423,
'eight': 0.00012125948181781436,
'fifty': 0.00014820603333288423,
'yards': 0.00012125948181781436,
'six': 0.0001392238494945276,
'killed': 0.00015718821717124084,
'sixty': 7.634856262603126e-05,
'_other': 4.49109191917831e-06,
'verbal': 8.98218383835662e-06,
'narrative': 4.041982727260479e-05,
'taken': 0.0002290456878780938,
'alfred': 8.98218383835662e-06,
'd': 0.00033683189393837325,
'890': 4.49109191917831e-06,
'whereas': 8.083965454520958e-05,
'things': 0.0006018063171698935,
'whether': 0.0004086893646452262,
'vessel': 0.00024701005555480705,
'dreadful': 2.694655151506986e-05,
'lost': 0.00021108132020138055,
'swallowed': 5.838419494931803e-05,
'gudgeon': 4.49109191917831e-06,
'retires': 4.49109191917831e-06,
'security': 1.347327575753493e-05,
'sleeps': 3.143764343424817e-05,
'montaigne': 4.49109191917831e-06,
'_apology': 4.49109191917831e-06,
'raimond': 4.49109191917831e-06,
'sebond_': 4.49109191917831e-06,
'let': 0.0007095925232301729,
'fly': 6.287528686849633e-05,
'nick': 8.98218383835662e-06,
'me': 0.002824896817163157,
'if': 0.002250037051508333,
'described': 4.041982727260479e-05,
'noble': 0.00019760804444384563,
'moses': 8.98218383835662e-06,

'patient': 2.694655151506986e-05,
'job': 6.287528686849633e-05,
'_rabelais_': 4.49109191917831e-06,
'liver': 1.796436767671324e-05,
'cartloads': 4.49109191917831e-06,
'_stowe': 4.49109191917831e-06,
'annals_': 4.49109191917831e-06,
'seas': 0.00039072499696851296,
'seethe': 1.347327575753493e-05,
'like': 0.0029057364717083663,
'boiling': 6.736637878767465e-05,
'pan': 2.2455459595891548e-05,
'_lord': 4.49109191917831e-06,
'bacon': 4.49109191917831e-06,
'psalms_': 4.49109191917831e-06,
'bulk': 0.00014371494141370592,
'ork': 4.49109191917831e-06,
'received': 0.0001392238494945276,
'nothing': 0.0005164755707055056,
'certain': 0.0004086893646452262,
'they': 0.003000049402011111,
'exceeding': 5.389310303013972e-05,
'fat': 6.287528686849633e-05,
'insomuch': 8.98218383835662e-06,
'an': 0.002694655151506986,
'incredible': 4.041982727260479e-05,
'quantity': 5.838419494931803e-05,
'oil': 0.0004041982727260479,
'extracted': 2.2455459595891548e-05,
'_ibid_': 1.347327575753493e-05,
'_history': 4.49109191917831e-06,
'death_': 4.49109191917831e-06,
'sovereignest': 4.49109191917831e-06,
'parmacetti': 1.347327575753493e-05,
'inward': 2.2455459595891548e-05,
'bruise': 4.49109191917831e-06,
'_king': 4.49109191917831e-06,
'henry_': 4.49109191917831e-06,
'_hamlet_': 4.49109191917831e-06,
'secure': 5.838419494931803e-05,
'skill': 4.4910919191783095e-05,
'leach': 4.49109191917831e-06,
'art': 0.0001841347686863107,
'mote': 4.49109191917831e-06,
'availle': 4.49109191917831e-06,
'returne': 4.49109191917831e-06,
'again': 4.49109191917831e-06,
'wound': 7.634856262603126e-05,
'worker': 4.49109191917831e-06,
'lowly': 3.592873535342648e-05,
'dinting': 4.49109191917831e-06,
'breast': 3.143764343424817e-05,
'bred': 4.041982727260479e-05,
'restless': 1.796436767671324e-05,
'paine': 4.49109191917831e-06,
'wounded': 4.041982727260479e-05,
'flies': 3.143764343424817e-05,
'thro': 4.49109191917831e-06,
'maine': 4.49109191917831e-06,
'faerie': 4.49109191917831e-06,

'queen_': 4.49109191917831e-06,
'immense': 7.185747070685296e-05,
'motion': 7.634856262603126e-05,
'vast': 0.0003278497101000166,
'bodies': 4.041982727260479e-05,
'can': 0.0011092997040370424,
'peaceful': 2.694655151506986e-05,
'calm': 0.0001841347686863107,
'trouble': 4.940201111096141e-05,
'ocean': 0.00035479626161508647,
'till': 0.0005524043060589322,
'boil': 8.98218383835662e-06,
'_sir': 1.796436767671324e-05,
'william': 1.796436767671324e-05,
'davenant': 4.49109191917831e-06,
'preface': 4.49109191917831e-06,
'gondibert_': 4.49109191917831e-06,
'spermacetti': 1.347327575753493e-05,
'men': 0.0010958264282795076,
'might': 0.0008218698212096307,
'justly': 1.347327575753493e-05,
'doubt': 0.00014371494141370592,
'since': 0.0002919209747465901,
'learned': 0.00011227729797945774,
'hosmannus': 4.49109191917831e-06,
'work': 0.0004895290191904357,
'thirty': 0.00012575057373699267,
'years': 0.00043563591616029604,
'saith': 2.2455459595891548e-05,
'plainly': 0.0001751525848479541,
'_nescio': 4.49109191917831e-06,
'quid': 8.98218383835662e-06,
'sit_': 4.49109191917831e-06,
't': 0.0013024166565617098,
'browne': 2.694655151506986e-05,
'sperma': 1.347327575753493e-05,
'ceti': 1.347327575753493e-05,
'vide': 4.49109191917831e-06,
'v': 3.143764343424817e-05,
'e': 0.00023353677979727212,
'spencer': 4.49109191917831e-06,
'talus': 4.49109191917831e-06,
'modern': 8.083965454520958e-05,
'flail': 4.49109191917831e-06,
'threatens': 4.49109191917831e-06,
'ruin': 2.2455459595891548e-05,
'ponderous': 5.838419494931803e-05,
'fixed': 0.00017066149292877577,
'jav': 4.49109191917831e-06,
'lins': 4.49109191917831e-06,
'side': 0.0009341471191890885,
'wears': 3.143764343424817e-05,
'back': 0.0007230657989877079,
'grove': 1.347327575753493e-05,
'pikes': 8.98218383835662e-06,
'_waller': 4.49109191917831e-06,
'battle': 0.0001392238494945276,
'summer': 7.185747070685296e-05,
'islands_': 4.49109191917831e-06,
'commonwealth': 4.49109191917831e-06,

```

'state': 0.00014820603333288423,
'latin': 1.796436767671324e-05,
'civitas': 4.49109191917831e-06,
'artificial': 2.2455459595891548e-05,
'man': 0.002366805441406969,
'_opening': 4.49109191917831e-06,
'sentence': 1.347327575753493e-05,
'hobbes': 4.49109191917831e-06,
'leviathan_': 4.49109191917831e-06,
'silly': 2.2455459595891548e-05,
'mansoul': 4.49109191917831e-06,
'without': 0.0007365390747452428,
'chewing': 4.49109191917831e-06,
'sprat': 4.49109191917831e-06,
'_pilgrim': 4.49109191917831e-06,
'progress_': 4.49109191917831e-06,
'hugest': 1.796436767671324e-05,
'swim': 9.431293030274451e-05,
'stream': 4.940201111096141e-05,
'_paradise': 4.49109191917831e-06,
'lost_': 4.49109191917831e-06,
'living': 0.0003278497101000166,
'creatures': 0.00015269712525206252,
'stretched': 4.4910919191783095e-05,
'promontory': 2.2455459595891548e-05,
'swims': 3.592873535342648e-05,
'seems': 0.00039072499696851296,
'moving': 6.736637878767465e-05,
'gills': 1.796436767671324e-05,
'draws': 4.49109191917831e-06,
'breath': 8.083965454520958e-05,
'spouts': 7.185747070685296e-05,
'mighty': 0.00021108132020138055,
'water': 0.0008533074646438789,
'swimming': 0.0001347327575753493,
'_fuller': 4.49109191917831e-06,
'holy': 9.431293030274451e-05,
'state_': 4.49109191917831e-06,
'close': 0.00025150114747398534,
'behind': 0.00022455459595891548,
...}

```

```

In [13]: def autocorrect(word):
          word=word.lower()
          if word in V:
              return ('Your word seems to be correct',word)
          else:
              similarities = [1-(textdistance.Jaccard(qval=2).distance(v,word)) for v in w
              df=pd.DataFrame.from_dict(probs, orient='index').reset_index()
              df=df.rename(columns={'index':'word',0:'Prob'})
              df['Similarity']=similarities
              output=df.sort_values(['Similarity','Prob'],ascending=False).head(3)
              return(output)
          autocorrect('pankaj')

```

Out[13]:

	word	Prob	Similarity
851	pan	0.000022	0.400000
5324	sank	0.000067	0.333333
5229	pans	0.000022	0.333333

In []: