To [1].	Creado por: Isabel Maniega Natural Language Processing (NLP)
	<pre>warnings.filterwarnings('ignore') import matplotlib.pyplot as plt import seaborn as sns # pip install nltk</pre>
out[5]:	<pre># nltk.download() nltk.download("gutenberg") [nltk_data] Downloading package gutenberg to [nltk_data] /home/isabelmaniega/nltk_data [nltk_data] Package gutenberg is already up-to-date!</pre>
Out[6]:	<pre>nltk.download("genesis") [nltk_data] Downloading package genesis to [nltk_data] /home/isabelmaniega/nltk_data [nltk_data] Package genesis is already up-to-date! True nltk.download("inaugural")</pre>
Out[7]:	<pre>[nltk_data] Downloading package inaugural to [nltk_data] /home/isabelmaniega/nltk_data [nltk_data] Package inaugural is already up-to-date! True nltk.download("nps_chat")</pre>
Out[8]: In [9]:	<pre>[nltk_data] Downloading package nps_chat to [nltk_data] /home/isabelmaniega/nltk_data [nltk_data] Package nps_chat is already up-to-date! True nltk.download('webtext') [nltk_data] Downloading package webtext to [nltk_data] /home/isabelmaniega/nltk_data</pre>
	<pre>[nltk_data] Package webtext is already up-to-date! True nltk.download('treebank') [nltk_data] Downloading package treebank to [nltk_data] /home/isabelmaniega/nltk_data [nltk_data] Package treebank is already up-to-date!</pre>
Out[10]:	<pre>1 Ejemplo from nltk.book import * *** Introductory Examples for the NLTK Book *** Loading text1,, text9 and sent1,, sent9</pre>
	Type the name of the text or sentence to view it. Type: 'texts()' or 'sents()' to list the materials. text1: Moby Dick by Herman Melville 1851 text2: Sense and Sensibility by Jane Austen 1811 text3: The Book of Genesis text4: Inaugural Address Corpus text5: Chat Corpus text6: Monty Python and the Holy Grail text7: Wall Street Journal
	text8: Personals Corpus text9: The Man Who Was Thursday by G . K . Chesterton 1908 print(text1) <text: 1851="" by="" dick="" herman="" melville="" moby=""> len(text1) 260819</text:>
	nltk.text.Text 1.1 Frecuencia de palabras en un texto
	Palabras con mayo frecuencia. Podría ayudar a distinguir autores por su forma de escribir. dist1 = FreqDist(text1) dist1 FreqDist({',': 18713, 'the': 13721, '.': 6862, 'of': 6536, 'and': 6024, 'a': 4569, 'to': 4542, ';': 4072, 'in': 3916, 'that': 2982,})
Out[16]:	<pre>dist1[","] 18713 dist1["whale"]</pre>
In [18]:	dist1.plot(25) plt.show() 17500 15000
	12500 10000 7500
	5000
In [19]:	
	14000 14000 12000
	10000
	Samples Vemos que hay palabras que no aportan mucha información 1.2 Stopwords
	<pre>from nltk.corpus import stopwords nltk.download("stopwords") [nltk_data] Downloading package stopwords to [nltk_data] /home/isabelmaniega/nltk_data [nltk_data] Package stopwords is already up-to-date! True</pre>
	<pre>sw = stopwords.words("spanish") sw</pre>
	'y', 'a', 'los', 'del', 'se', 'las', 'por', 'un', 'para',
	'con', 'no', 'una', 'su', 'al', 'lo', 'como', 'más', 'pero',
	<pre>'sus', 'le', 'ya', 'o', 'este', 'si', 'porque', 'esta', 'entre', 'cuando',</pre>
	<pre>'muy', 'sin', 'sobre', 'también', 'me', 'hasta', 'hay', 'donde', 'quien',</pre>
	<pre>'desde', 'todo', 'nos', 'durante', 'todos', 'uno', 'les', 'ni', 'contra', 'otros',</pre>
	'ese', 'eso', 'ante', 'ellos', 'e', 'esto', 'mí', 'antes', 'algunos',
	'qué', 'unos', 'yo', 'otro', 'otras', 'otra', 'él', 'tanto', 'esa',
	<pre>'estos', 'mucho', 'quienes', 'nada', 'muchos', 'cual', 'poco', 'ella', 'estar', 'estas',</pre>
	'algunas', 'algo', 'nosotros', 'mi', 'mis', 'tú', 'té', 'te', 'ti',
	<pre>'tus', 'ellas', 'nosotras', 'vosotros', 'vosotras', 'os', 'mío', 'mío', 'mía', 'míos',</pre>
	<pre>'mías', 'tuyo', 'tuya', 'tuyas', 'tuyas', 'suyo', 'suyo', 'suya', 'suyos', 'nuestro',</pre>
	<pre>'nuestra', 'nuestros', 'nuestras', 'vuestro', 'vuestra', 'vuestros', 'vuestras', 'esos', 'esos',</pre>
	<pre>'estoy', 'estás', 'está', 'estamos', 'estáis', 'están', 'esté', 'esté', 'estés', 'estés', 'estemos', 'estéis',</pre>
	<pre>'estén', 'estaré', 'estarás', 'estará', 'estaremos', 'estaréis', 'estarán', 'estaría', 'estaría',</pre>
	<pre>'estaríamos', 'estaríais', 'estarían', 'estaba', 'estabas', 'estabass', 'estabais', 'estabais', 'estaban', 'estaban',</pre>
	<pre>'estuviste', 'estuvo', 'estuvimos', 'estuvisteis', 'estuvieron', 'estuviera', 'estuvieras', 'estuviéramos', 'estuvierais', 'estuvierais',</pre>
	'estuviese', 'estuviésemos', 'estuvieseis', 'estuviesen', 'estando', 'estado', 'estado', 'estados',
	<pre>'estadas', 'estad', 'he', 'has', 'ha', 'hemos', 'habéis', 'habéis', 'han', 'haya',</pre>
	'hayanos', 'hayais', 'hayan', 'habré', 'habrás', 'habrás', 'habréis', 'habréis', 'habréis',
	'habría', 'habrías', 'habríamos', 'habríais', 'habríain', 'había', 'habías', 'habías', 'habíamos',
	<pre>'habían', 'hube', 'hubiste', 'hubo', 'hubimos', 'hubisteis', 'hubieron', 'hubiera', 'hubieras',</pre>
	'hubierais', 'hubieran', 'hubiese', 'hubieses', 'hubiésemos', 'hubiésemos', 'hubieseis', 'hubiesen', 'hubiesen', 'habiendo',
	<pre>'habida', 'habidos', 'habidas', 'soy', 'eres', 'es', 'somos', 'somos', 'sois', 'son',</pre>
	<pre>'sea', 'seas', 'seamos', 'seáis', 'sean', 'seré', 'serás', 'será', 'serés', 'serés', 'serési',</pre>
	'sería', 'serías', 'serías', 'seríamos', 'seríais', 'serían', 'era', 'eras', 'éramos',
	<pre>'erais', 'eran', 'fui', 'fuiste', 'fue', 'fuimos', 'fuisteis', 'fueron', 'fueras',</pre>
	'fuéramos', 'fuerais', 'fueran', 'fuese', 'fueses', 'fuésemos', 'fuéseis', 'fuesen', 'sintiendo', 'sentido',
	<pre>'sentido', 'sentidos', 'sentidos', 'sentidas', 'siente', 'siente', 'tengo', 'tienes', 'tienes', 'tienes',</pre>
	<pre>'tenéis', 'tienen', 'tenga', 'tengas', 'tengamos', 'tengáis', 'tengan', 'tendré', 'tendrás', 'tendrás',</pre>
	'tendremos', 'tendréis', 'tendrán', 'tendría', 'tendrías', 'tendríamos', 'tendríais', 'tendríais', 'tendrían', 'tendrían',
	<pre>'tenías', 'teníamos', 'teníais', 'teníais', 'tuve', 'tuve', 'tuviste', 'tuvo', 'tuvimos', 'tuvimos', 'tuvisteis', 'tuvieron',</pre>
	'tuviera', 'tuvieras', 'tuviéramos', 'tuvierais', 'tuvieran', 'tuviese', 'tuvieses', 'tuviésemos', 'tuviésemos',
In [23]:	<pre>'tuviesen', 'teniendo', 'tenido', 'tenida', 'tenidos', 'tenidas', 'tened']</pre>
Out[23]: In [24]: Out[24]:	<pre>stopwords = stopwords.words("english") stopwords</pre>
	<pre>'we', 'our', 'ours', 'ourselves', 'you', "you're", "you've", "you'd",</pre>
	<pre>'your', 'yourself', 'yourselves', 'he', 'him', 'his', 'himself', 'she', "she's",</pre>
	<pre>'her', 'hers', 'herself', 'it', "it's", 'its', 'itself', 'they', 'them', 'their'</pre>
	<pre>'their', 'theirs', 'themselves', 'what', 'which', 'who', 'whom', 'this', 'that', "that',"</pre>
	<pre>'these', 'those', 'am', 'is', 'are', 'was', 'be', 'been', 'being',</pre>
	<pre>'being', 'have', 'has', 'had', 'having', 'do', 'does', 'did', 'doing', 'a',</pre>
	<pre>'an', 'the', 'and', 'but', 'if', 'or', 'because', 'as', 'until', 'while',</pre>
	<pre>'of', 'at', 'by', 'for', 'with', 'about', 'against', 'between', 'into', 'through',</pre>
	<pre>'during', 'before', 'after', 'above', 'below', 'to', 'from', 'up', 'down', 'in',</pre>
	<pre>'in', 'out', 'on', 'off', 'over', 'under', 'again', 'further', 'then', 'once',</pre>
	<pre>'here', 'there', 'when', 'where', 'why', 'how', 'all', 'any', 'both', 'each',</pre>
	<pre>'few', 'more', 'most', 'other', 'some', 'such', 'no', 'nor', 'nor', 'not', 'only',</pre>
	<pre>'only', 'own', 'same', 'so', 'than', 'too', 'very', 's', 't',</pre>
	<pre>'will', 'just', 'don', "don't", 'should', "should've", 'now', 'd', 'd', 'll', 'm',</pre>
	<pre>'m', 'o', 're', 've', 'y', 'ain', 'aren't", 'couldn', "couldn't",</pre>
	<pre>'didn', "didn't", 'doesn', "doesn't", 'hadn', "hadn't", 'hasn', "hasn't", "hasn't", 'haven', "haven't",</pre>
	<pre>"haven't", 'isn', "isn't", 'ma', 'mightn', "mightn't", 'mustn', "mustn't", "needn', "needn't",</pre>
	<pre>"needn't", 'shan', "shan't", 'shouldn', "shouldn't", 'wasn', "wasn't", 'weren', "weren't", "weren't", 'won', "won't",</pre>
In [25]: Out[25]:	<pre>"won't", 'wouldn', "wouldn't"] len(stopwords)</pre>
In [28]: Out[28]:	<pre>import string string.punctuation '!"#\$%&\'()*+,/:;<=>?@[\\]^_`{ }~' type(string.punctuation)</pre>
Out[29]:	<pre>signos_puntuacion = list(string.punctuation) print(signos_puntuacion) ['!', '"', '#', '\$', '&', "'", '(', ')', '*', '+', ',', '-', '.', '/', ':', ';', '<', '=', '>', '?', '@', '[', '\\', ']', '^', '_', '\'', '\'', '\'', '\'']</pre>
In [31]:	1.4 Palabras a eliminar palabras_eliminar = stopwords + signos_puntuacion palabras_eliminar











Out[46]: Counter(<pre>{'Moby': 84, 'Dick': 84, 'Herman': 1, 'Melville': 1, '1851': 3, 'ETYMOLOGY': 1, 'Supplied': 2, 'Late': 1, 'Consumptive': 1,</pre>
	'Usher': 2, 'Grammar': 2, 'School': 1, 'The': 612, 'pale': 18, '': 1070, 'threadbare': 1, 'coat': 28, 'heart': 90, 'body': 110,
	'brain': 37, 'I': 2124, 'see': 253, 'He': 230, 'ever': 203, 'dusting': 2, 'old': 436, 'lexicons': 1, 'grammars': 2, 'queer': 44,
	'handkerchief': 5, 'mockingly': 1, 'embellished': 3, 'gay': 13, 'flags': 1, 'known': 80, 'nations': 12, 'world': 173, 'loved': 3, 'dust': 10,
	<pre>'somehow': 43, 'mildly': 10, 'reminded': 4, 'mortality': 1, 'While': 30, 'take': 118, 'hand': 205, 'school': 9, 'others': 37, 'teach': 5,</pre>
	<pre>'name': 69, 'whale': 906, 'fish': 133, 'called': 116, 'tongue': 11, 'leaving': 38, 'ignorance': 10, 'letter': 12, 'H': 2, 'almost': 186,</pre>
	<pre>'alone': 37, 'maketh': 5, 'signification': 1, 'word': 76, 'deliver': 3, 'true': 73, '."': 489, 'HACKLUYT': 1, 'WHALE': 38, '': 11,</pre>
	'Sw': 1, 'Dan': 4, 'HVAL': 1, 'This': 107, 'animal': 16, 'named': 11, 'roundness': 1, 'rolling': 35, 'HVALT': 2, 'arched': 10, 'vaulted': 2,
	'WEBSTER': 2, 'S': 63, 'DICTIONARY': 3, 'It': 310, 'immediately': 14, 'Dut': 1, 'Ger': 1, 'WALLEN': 1, 'A': 167, 'WALW': 1,
	'IAN': 1, 'roll': 16, 'wallow': 3, 'RICHARDSON': 1, 'KETOS': 1, 'GREEK': 1, 'CETUS': 1, 'LATIN': 1, 'WHOEL': 1, 'ANGLO': 1, 'SAXON': 1,
	'DANISH': 2, 'WAL': 1, 'DUTCH': 2, 'HWAL': 1, 'SWEDISH': 1, 'ICELANDIC': 1, 'ENGLISH': 3, 'BALEINE': 1, 'FRENCH': 5, 'BALLENA': 1,
	'SPANISH': 4, 'PEKEE': 2, 'NUEE': 4, 'FEGEE': 1, 'ERROMANGOAN': 1, 'EXTRACTS': 2, 'Sub': 7, 'Librarian': 1, ').': 15, 'seen': 161,
	<pre>'mere': 42, 'painstaking': 1, 'burrower': 1, 'grub': 1, 'worm': 5, 'poor': 106, 'devil': 51, 'appears': 7, 'gone': 58, 'long': 318, 'Vaticans': 1,</pre>
	<pre>'street': 10, 'stalls': 1, 'earth': 46, 'picking': 4, 'whatever': 44, 'random': 5, 'allusions': 7, 'whales': 237, 'could': 215, 'anyways': 2,</pre>
	'find': 55, 'book': 38, 'whatsoever': 5, 'sacred': 7, 'profane': 4, 'Therefore': 11, 'must': 282, 'every': 222, 'case': 69, 'least': 75, 'higgledy': 1,
	<pre>'piggledy': 1, 'statements': 4, 'however': 80, 'authentic': 3, 'extracts': 4, 'veritable': 4, 'gospel': 1, 'cetology': 4, 'Far': 4, 'As': 120,</pre>
	<pre>'touching': 45, 'ancient': 28, 'authors': 3, 'generally': 30, 'well': 174, 'poets': 3, 'appearing': 3, 'solely': 6, 'valuable': 11, 'entertaining': 2, 'affording': 2,</pre>
	'glancing': 12, 'bird': 17, 'eye': 88, 'view': 46, 'promiscuously': 1, 'said': 302, 'thought': 149, 'fancied': 9, 'sung': 2, 'Leviathan': 64,
	<pre>'many': 161, 'generations': 6, 'including': 8, 'So': 147, 'fare': 10, 'thee': 129, 'whose': 86, 'commentator': 1, 'Thou': 39, 'belongest': 1, 'hopeless': 8,</pre>
	<pre>'sallow': 1, 'tribe': 9, 'wine': 12, 'warm': 24, 'even': 181, 'Pale': 1, 'Sherry': 1, 'would': 421, 'rosy': 2, 'strong': 34, 'one': 889,</pre>
	'sometimes': 81, 'loves': 1, 'sit': 16, 'feel': 56, 'devilish': 9, 'grow': 24, 'convivial': 4, 'upon': 538, 'tears': 6, 'say': 237,
	'bluntly': 1, 'full': 121, 'eyes': 155, 'empty': 22, 'glasses': 10, 'altogether': 29, 'unpleasant': 3, 'sadness': 2, 'Give': 15, 'Subs': 1, 'For': 197,
	<pre>'much': 218, 'pains': 15, 'ye': 460, 'please': 12, 'shall': 84, 'go': 182, 'thankless': 1, 'Would': 11, 'clear': 39, 'Hampton': 1,</pre>
	'Court': 1, 'Tuileries': 3, 'But': 705, 'gulp': 2, 'hie': 1, 'aloft': 53, 'royal': 29, 'mast': 124, 'hearts': 29, 'friends': 16, 'clearing': 4,
	<pre>'seven': 15, 'storied': 4, 'heavens': 17, 'making': 45, 'refugees': 1, 'pampered': 1, 'Gabriel': 20, 'Michael': 1, 'Raphael': 1, 'coming': 50,</pre>
	'Here': 49, 'strike': 28, 'splintered': 7, 'together': 64, 'unsplinterable': 1, 'And': 369, 'God': 132, 'created': 6, 'great': 293, 'GENESIS': 1, 'path': 10,
	<pre>'shine': 3, 'One': 28, 'think': 111, 'deep': 69, 'hoary': 2, 'JOB': 1, 'Now': 139, 'Lord': 44, 'prepared': 12, 'swallow': 10, 'Jonah': 84,</pre>
	'JONAH': 1, 'There': 147, 'ships': 87, 'thou': 232, 'hast': 19, 'made': 178, 'play': 29, 'therein': 5, 'PSALMS': 2, 'In': 237,
	'day': 172, 'sore': 3, 'sword': 23, 'punish': 1, 'piercing': 4, 'serpent': 4, 'crooked': 8, 'slay': 4, 'dragon': 5, 'sea': 433, 'ISAIAH': 1,
	'thing': 188, 'soever': 1, 'besides': 20, 'cometh': 1, 'within': 78, 'chaos': 4, 'monster': 49, 'mouth': 60, 'beast': 4, 'boat': 330, 'stone': 18,
	'goes': 54, 'incontinently': 1, 'foul': 11, 'perisheth': 1, 'bottomless': 7, 'gulf': 4, 'paunch': 2, 'HOLLAND': 2, 'PLUTARCH': 1, 'MORALS': 1,
	'Indian': 57, 'Sea': 22, 'breedeth': 1, 'biggest': 2, 'fishes': 3, 'among': 158, 'Whales': 29, 'Whirlpooles': 1, 'Balaene': 1, 'length': 78, 'four': 74,
	<pre>'acres': 3, 'arpens': 1, 'land': 78, 'PLINY': 1, 'Scarcely': 1, 'proceeded': 6, 'two': 285, 'days': 80, 'sunrise': 11, 'monsters': 9,</pre>
	<pre>'appeared': 10, 'Among': 8, 'former': 20, 'monstrous': 10, 'size': 12, 'came': 130, 'towards': 107, 'us': 228, 'open': 75, 'mouthed': 4, 'raising': 4,</pre>
	<pre>'waves': 48, 'sides': 28, 'beating': 11, 'foam': 19, 'TOOKE': 1, 'LUCIAN': 1, 'THE': 98, 'TRUE': 1, 'HISTORY': 4, 'visited': 7, 'country': 25,</pre>
	<pre>'also': 88, 'catching': 7, 'horse': 26, 'bones': 48, 'value': 6, 'teeth': 47, 'brought': 38, 'king': 25, 'best': 62, 'catched': 1,</pre>
	'forty': 28, 'eight': 24, 'fifty': 33, 'yards': 27, 'six': 31, 'killed': 33, 'sixty': 17, 'OTHER': 2, 'OR': 5, 'OCTHER': 1, 'VERBAL': 1,
	'NARRATIVE': 4, 'TAKEN': 1, 'DOWN': 2, 'FROM': 10, 'HIS': 17, 'MOUTH': 1, 'BY': 15, 'KING': 2, 'ALFRED': 1, 'D': 23, '890': 1,
	<pre>'whereas': 14, 'things': 132, 'whether': 76, 'vessel': 54, 'enter': 9, 'dreadful': 6, 'lost': 45, 'swallowed': 13, 'gudgeon': 1, 'retires': 1, 'security': 3,</pre>
	'sleeps': 7, 'MONTAIGNE': 1, 'APOLOGY': 1, 'FOR': 6, 'RAIMOND': 1, 'SEBOND': 1, 'Let': 38, 'fly': 14, 'let': 120, 'Old': 11, 'Nick': 1,
	<pre>'described': 8, 'noble': 44, 'prophet': 13, 'Moses': 2, 'life': 162, 'patient': 6, 'Job': 5, 'RABELAIS': 1, 'liver': 4, 'cartloads': 1,</pre>
	'ANNALS': 1, 'seas': 80, 'seethe': 3, 'like': 624, 'boiling': 15, 'pan': 3, 'LORD': 1, 'BACON': 1, 'VERSION': 2, 'OF': 49, 'Touching': 1,
	'bulk': 32, 'ork': 1, 'received': 26, 'nothing': 102, 'certain': 87, 'They': 78, 'exceeding': 12, 'fat': 12, 'insomuch': 2, 'incredible': 9, 'quantity': 13,
	<pre>'oil': 90, 'extracted': 5, 'IBID': 3, 'LIFE': 2, 'AND': 37, 'DEATH': 1, 'sovereignest': 1, 'parmacetti': 3, 'inward': 4, 'bruise': 1, 'HENRY': 2,</pre>
	'Very': 11, 'HAMLET': 1, 'Which': 6, 'secure': 12, 'skill': 10, 'leach': 1, 'art': 39, 'Mote': 1, 'availle': 1, 'returne': 1, 'againe': 1,
	'To': 62, 'wound': 17, 'worker': 1, 'lowly': 8, 'dart': 24, 'Dinting': 1, 'breast': 7, 'bred': 9, 'restless': 4, 'paine': 1, 'Like': 23,
	<pre>'wounded': 9, 'shore': 23, 'flies': 7, 'thro': 1, 'maine': 1, 'FAERIE': 1, 'QUEEN': 2, 'Immense': 1, 'motion': 17, 'vast': 73, 'bodies': 9,</pre>
	<pre>'peaceful': 6, 'calm': 41, 'trouble': 11, 'ocean': 71, 'til': 1, 'boil': 2, 'SIR': 4, 'WILLIAM': 2, 'DAVENANT': 1, 'PREFACE': 1, 'TO': 21,</pre>
	'GONDIBERT': 1, 'What': 174, 'spermacetti': 1, 'men': 236, 'might': 183, 'justly': 3, 'doubt': 32, 'since': 63, 'learned': 25, 'Hosmannus': 1, 'work': 62,
	'thirty': 26, 'years': 94, 'saith': 5, 'plainly': 39, 'Nescio': 1, 'quid': 2, 'T': 8, 'BROWNE': 3, 'SPERMA': 2, 'CETI': 2,
	'VIDE': 1, 'V': 6, 'E': 5, 'Spencer': 1, 'Talus': 1, 'modern': 18, 'flail': 1, 'threatens': 1, 'ruin': 5, 'ponderous': 13, 'tail': 78,
	'Their': 6, 'fixed': 37, 'jav': 1, 'lins': 1, 'side': 208, 'wears': 7, 'back': 151, 'grove': 3, 'pikes': 2, 'WALLER': 1, 'BATTLE': 2,
	'SUMMER': 1, 'ISLANDS': 1, 'By': 52, 'Commonwealth': 1, 'State': 4, '(': 4, 'Latin': 4, 'Civitas': 1, 'artificial': 5, 'man': 508,
	'OPENING': 1, 'SENTENCE': 1, 'HOBBES': 1, 'LEVIATHAN': 1, 'Silly': 1, 'Mansoul': 1, 'without': 154, 'chewing': 1, 'sprat': 1, 'PILGRIM': 1, 'PROGRESS': 1,
	'That': 85, 'works': 30, 'Created': 1, 'hugest': 3, 'swim': 20, 'stream': 11, 'PARADISE': 1, 'LOST': 1, '"': 1, 'Hugest': 1, 'living': 73,
	<pre>'creatures': 34, 'Stretched': 1, 'promontory': 5, 'swims': 8, 'seems': 84, 'moving': 13, 'gills': 4, 'Draws': 1, 'breath': 18, 'spouts': 16, 'mighty': 47,</pre>
	<pre>'water': 185, 'swimming': 29, 'FULLLER': 1, 'PROFANE': 1, 'HOLY': 1, 'STATE': 1, 'close': 54, 'behind': 50, 'lie': 26, 'huge': 29, 'attend': 9,</pre>
	'prey': 14, 'give': 71, 'chance': 46, 'fry': 3, 'gaping': 3, 'jaws': 30, 'mistake': 9, 'way': 269, 'DRYDEN': 1, 'ANNUS': 1,
	'MIRABILIS': 1, 'floating': 24, 'stern': 48, 'ship': 507, 'cut': 50, 'head': 335, 'tow': 16, 'near': 69, 'come': 150, 'aground': 1, 'twelve': 16,
	'thirteen': 2, 'feet': 125, 'THOMAS': 3, 'EDGE': 1, 'TEN': 1, 'VOYAGES': 3, 'SPITZBERGEN': 1, 'IN': 19, 'PURCHAS': 1, 'saw': 115, 'sporting': 4,
	'wantonness': 1, 'fuzzing': 1, 'pipes': 8, 'vents': 1, 'nature': 34, 'placed': 36, 'shoulders': 13, 'HERBERT': 1, 'INTO': 1, 'ASIA': 1, 'AFRICA': 1,
	'HARRIS': 2, 'COLL': 2, 'troops': 2, 'forced': 18, 'proceed': 14, 'deal': 12, 'caution': 2, 'fear': 32, 'run': 51, 'SCHOUTEN': 1, 'SIXTH': 1,
	'CIRCUMNAVIGATION': 1, 'We': 47, 'set': 74, 'sail': 96, 'Elbe': 1, 'wind': 67, 'N': 2, 'Jonas': 3, 'Whale': 282, 'Some': 38, 'fable': 5,
	'frequently': 14, 'climb': 4, 'masts': 26, 'first': 210, 'discoverer': 2, 'ducat': 1, 'told': 45, 'taken': 50, 'Shetland': 2, 'barrel': 5,
	'herrings': 1, 'belly': 9, 'harpooneers': 55, 'caught': 44, 'Spitzbergen': 4, 'white': 191, 'VOYAGE': 7, 'GREENLAND': 1, '1671': 2, 'Several': 1,
	'coast': 40, 'Fife': 1, 'Anno': 1, '1652': 1, 'eighty': 4, 'bone': 50, 'kind': 27, 'informed': 4, '),': 62, 'afford': 9, '500': 2,
	<pre>'weight': 17, 'baleen': 11, 'stand': 83, 'gate': 4, 'garden': 6, 'Pitferren': 1, 'SIBBALD': 1, 'FIFE': 1, 'KINROSS': 1, 'Myself': 1, 'agreed': 3,</pre>
	'try': 49, 'master': 23, 'kill': 30, 'Sperma': 1, 'ceti': 1, 'never': 195, 'hear': 60, 'sort': 152, 'fierceness': 1, 'swiftness': 6, 'RICHARD': 1
	'RICHARD': 1, 'STRAFFORD': 1, 'LETTER': 1, 'BERMUDAS': 1, 'PHIL': 1, 'TRANS': 1, 'toice': 46, 'obey': 8, 'PRIMER': 1, 'abundance': 4, 'large': 77,
	'large': 77, 'southern': 5, 'may': 230, 'hundred': 50, 'northward': 4, 'CAPTAIN': 2, 'COWLEY': 1, 'ROUND': 2, 'GLOBE': 4, '1729': 1, '"': 1,
	'frequendy': 1, 'attended': 6, 'insupportable': 1, 'smell': 11, 'bring': 18, 'disorder': 1, 'ULLOA': 1, 'SOUTH': 1, 'AMERICA': 1, 'chosen': 3, 'sylphs': 1,
	<pre>'sylphs': 1, 'special': 15, 'note': 6, 'trust': 4, 'important': 19, 'charge': 10, 'petticoat': 1, '0ft': 1, 'fold': 6, 'fence': 2, 'fail': 17,</pre>
	'Tho': 1, 'stuffed': 5, 'hoops': 11, 'armed': 15, 'ribs': 20, 'RAPE': 1, 'LOCK': 1, 'If': 60, 'compare': 7, 'animals': 5, 'respect': 16,
	<pre>'respect': 16, 'magnitude': 21, 'abode': 2, 'appear': 6, 'contemptible': 4, 'comparison': 6, 'doubtless': 12, 'largest': 19, 'creation': 7, 'GOLDSMITH': 2, 'NAT': 1, 'HIST': 2,</pre>
	'HIST': 2, 'write': 5, 'little': 247, 'make': 109, 'speak': 47, 'wales': 1, 'JOHNSON': 1, 'afternoon': 10, 'supposed': 15, 'rock': 14, 'found': 115,
	'dead': 90, 'Asiatics': 2, 'towing': 14, 'ashore': 33, 'seemed': 283, 'endeavor': 4, 'conceal': 2, 'order': 64, 'avoid': 6, 'COOK': 1, 'larger': 9,
	<pre>'seldom': 24, 'venture': 5, 'attack': 8, 'dread': 12, 'afraid': 15, 'mention': 15, 'names': 12, 'carry': 25, 'dung': 1, 'lime': 1, 'juniper': 1, 'wood': 34,</pre>
	'BANKS': 1, 'SOLANDER': 1, 'ICELAND': 2, '1772': 1, 'Spermacetti': 2, 'Nantuckois': 1, 'active': 8, 'fierce': 11, 'requires': 7, 'address': 4, 'boldness': 2, 'fishermen': 28,



'Dick', 'Herman', 'Melville', '1851', 'ETYMOLOGY', 'Supplied', 'Late', 'Consumptive', 'Usher',
'Grammar', 'School', 'The', 'pale', '', 'threadbare', 'coat', 'heart', 'body', 'brain',
<pre>'I', 'see', 'He', 'ever', 'dusting', 'old', 'lexicons', 'grammars', 'queer', 'handkerchief',</pre>
<pre>'mockingly', 'embellished', 'gay', 'flags', 'known', 'nations', 'world', 'loved', 'dust', 'somehow',</pre>
<pre>'mildly', 'reminded', 'mortality', 'While', 'take', 'hand', 'school', 'others', 'teach', 'name',</pre>
<pre>'whale', 'fish', 'called', 'tongue', 'leaving', 'ignorance', 'letter', 'H', 'almost', 'alone', 'maketh',</pre>
<pre>"signification', 'word', 'deliver', 'true', '."', 'HACKLUYT', 'WHALE', '', 'Sw', 'Dan',</pre>
'HVAL', 'This', 'animal', 'named', 'roundness', 'rolling', 'HVALT', 'arched', 'vaulted', 'WEBSTER',
'S', 'DICTIONARY', 'It', 'immediately', 'Dut', 'Ger', 'WALLEN', 'A',
'IAN', 'roll', 'wallow', 'RICHARDSON', 'KETOS', 'GREEK', 'CETUS', 'LATIN', 'WHOEL', 'ANGLO',
'SAXON', 'DANISH', 'WAL', 'DUTCH', 'HWAL', 'SWEDISH', 'ICELANDIC', 'ENGLISH', 'BALEINE', 'BRALEINE',
'BALLENA', 'SPANISH', 'PEKEE', 'NUEE', 'FEGEE', 'ERROMANGOAN', 'EXTRACTS', 'Sub', 'Librarian', ').',
<pre>'seen', 'mere', 'painstaking', 'burrower', 'grub', 'worm', 'poor', 'devil', 'appears', 'gone',</pre>
'long', 'Vaticans', 'street', 'stalls', 'earth', 'picking', 'whatever', 'random', 'allusions',
<pre>'could', 'anyways', 'find', 'book', 'whatsoever', 'sacred', 'profane', 'Therefore', 'must', 'every',</pre>
<pre>'case', 'least', 'higgledy', 'piggledy', 'statements', 'however', 'authentic', 'extracts', 'veritable', 'gospel',</pre>
'cetology', 'Far', 'As', 'touching', 'ancient', 'authors', 'generally', 'well', 'poets', 'appearing',
<pre>'solely', 'valuable', 'entertaining', 'affording', 'glancing', 'bird', 'eye', 'view', 'promiscuously', 'said',</pre>
'thought', 'fancied', 'sung', 'Leviathan', 'many', 'generations', 'including', 'So', 'fare', 'thee',
<pre>'whose', 'commentator', 'Thou', 'belongest', 'hopeless', 'sallow', 'tribe', 'wine', 'warm', 'even',</pre>
'Pale', 'Sherry', 'would', 'rosy', 'strong', 'one', 'sometimes', 'loves', 'sit',
<pre>'feel', 'devilish', 'grow', 'convivial', 'upon', 'tears', 'say', 'bluntly', 'full', 'eyes',</pre>
<pre>'empty', 'glasses', 'altogether', 'unpleasant', 'sadness', 'Give', 'Subs', 'For', 'much', 'pains',</pre>
<pre>'ye', 'please', 'shall', 'go', 'thankless', 'Would', 'clear', 'Hampton', 'Court', 'Tuileries',</pre>
'But', 'gulp', 'hie', 'aloft', 'royal', 'mast', 'hearts', 'friends', 'clearing', 'seven',
<pre>'storied', 'heavens', 'making', 'refugees', 'pampered', 'Gabriel', 'Michael', 'Raphael', 'there',</pre>
<pre>'strike', 'splintered', 'together', 'unsplinterable', 'And', 'God', 'created', 'great', 'great', 'path',</pre>
<pre>'shine', 'One', 'think', 'deep', 'hoary', 'JOB', 'Now', 'Lord', 'prepared', 'swallow',</pre>
'JONAH', 'JONAH', 'There', 'ships', 'thou', 'hast', 'made', 'play', 'therein', 'PSALMS',
'In', 'day', 'sore', 'sword', 'punish', 'piercing', 'serpent', 'crooked', 'slay',
<pre>'dragon', 'sea', 'ISAIAH', 'thing', 'soever', 'besides', 'cometh', 'within', 'chaos', 'monster',</pre>
<pre>'mouth', 'beast', 'boat', 'stone', 'goes', 'incontinently', 'foul', 'perisheth', 'bottomless', 'gulf',</pre>
<pre>'paunch', 'HOLLAND', 'PLUTARCH', 'MORALS', 'Indian', 'Sea', 'breedeth', 'biggest', 'fishes', 'among',</pre>
<pre>'Whales', 'Whirlpooles', 'Balaene', 'length', 'four', 'acres', 'arpens', 'land', 'PLINY', 'Scarcely',</pre>
<pre>'proceeded', 'two', 'days', 'sunrise', 'monsters', 'appeared', 'Among', 'former', 'monstrous',</pre>
<pre>'came', 'towards', 'us', 'open', 'mouthed', 'raising', 'waves', 'sides', 'beating', 'foam',</pre>
'TOOKE', 'LUCIAN', 'THE', 'TRUE', 'HISTORY', 'visited', 'country', 'also', 'catching',
<pre>'bones', 'value', 'teeth', 'brought', 'king', 'best', 'catched', 'forty', 'eight',</pre>
'yards', 'six', 'killed', 'sixty', 'OTHER', 'OR', 'OCTHER', 'VERBAL', 'NARRATIVE', 'TAKEN',
'DOWN', 'FROM', 'HIS', 'MOUTH', 'BY', 'KING', 'ALFRED', 'D', '890', 'whereas',
<pre>'things', 'whether', 'vessel', 'enter', 'dreadful', 'lost', 'swallowed', 'gudgeon', 'retires', 'security',</pre>
'sleeps', 'MONTAIGNE', 'APOLOGY', 'FOR', 'RAIMOND', 'SEBOND', 'Let', 'fly', 'let',
'Old', 'Nick', 'described', 'noble', 'prophet', 'Moses', 'life', 'patient', 'Job', 'RABELAIS',
<pre>'liver', 'cartloads', 'STOWE', 'ANNALS', 'seas', 'seethe', 'like', 'boiling', 'pan', 'LORD',</pre>
'BACON', 'VERSION', 'OF', 'Touching', 'bulk', 'ork', 'received', 'nothing', 'certain', 'They',
<pre>'exceeding', 'fat', 'insomuch', 'incredible', 'quantity', 'oil', 'extracted', 'IBID', 'LIFE', 'AND',</pre>
'DEATH', 'sovereignest', 'parmacetti', 'inward', 'bruise', 'HENRY', 'Very', 'HAMLET', 'Which', 'secure',
<pre>'skill', 'leach', 'art', 'Mote', 'availle', 'returne', 'againe', 'To', 'wound',</pre>
<pre>'lowly', 'dart', 'Dinting', 'breast', 'bred', 'restless', 'paine', 'Like', 'wounded', 'shore',</pre>
<pre>'flies', 'thro', 'maine', 'FAERIE', 'QUEEN', 'Immense', 'motion', 'vast', 'bodies', 'peaceful',</pre>
'calm', 'trouble', 'ocean', 'til', 'boil', 'SIR', 'WILLIAM', 'DAVENANT', 'PREFACE', 'TO',
'GONDIBERT', 'What', 'spermacetti', 'men', 'might', 'justly', 'doubt', 'since', 'learned', 'Hosmannus',
'work', 'thirty', 'years', 'saith', 'plainly', 'Nescio', 'quid', 'T', 'BROWNE', 'SPERMA',
'CETI', 'VIDE', 'V', 'E', 'Spencer', 'Talus', 'modern', 'flail', 'threatens', 'ruin',
<pre>'ponderous', 'tail', 'Their', 'fixed', 'jav', 'lins', 'side', 'wears',</pre>
'grove', 'pikes', 'WALLER', 'BATTLE', 'SUMMER', 'ISLANDS', 'By', 'Commonwealth', 'State', '(',
'Latin', 'Civitas', 'artificial', 'man', 'OPENING', 'SENTENCE', 'HOBBES', 'LEVIATHAN', 'Silly', 'Mansoul',
<pre>'without', 'chewing', 'sprat', 'PILGRIM', 'PROGRESS', 'That', 'works', 'Created', 'hugest', 'swim',</pre>
<pre>'stream', 'PARADISE', 'LOST', '"', 'Hugest', 'living', 'creatures', 'Stretched', 'promontory', 'swims',</pre>
<pre>'seems', 'moving', 'gills', 'Draws', 'breath', 'spouts', 'mighty', 'water', 'swimming', 'FULLLER',</pre>
'PROFANE', 'HOLY', 'STATE', 'close', 'behind', 'lie', 'huge', 'attend', 'prey', 'give', 'chance',
'fry', 'gaping', 'jaws', 'mistake', 'way', 'DRYDEN', 'ANNUS', 'MIRABILIS', 'floating', 'stern',
'ship', 'cut', 'head', 'tow', 'near', 'come', 'aground', 'twelve', 'thirteen', 'feet',
'THOMAS', 'EDGE', 'TEN', 'VOYAGES', 'SPITZBERGEN', 'IN', 'PURCHAS', 'saw', 'sporting', 'wantonness',
'fuzzing', 'pipes', 'vents', 'nature', 'placed', 'shoulders', 'HERBERT', 'INTO', 'ASIA', 'AFRICA',
'HARRIS', 'COLL', 'troops', 'forced', 'proceed', 'deal', 'caution', 'fear', 'run',
'SCHOUTEN', 'SIXTH', 'CIRCUMNAVIGATION', 'We', 'set', 'sail', 'Elbe', 'wind', 'N', 'Jonas',
'Whale', 'Some', 'fable', 'frequently', 'climb', 'masts', 'first', 'discoverer', 'ducat', 'told',
<pre>'taken', 'Shetland', 'barrel', 'herrings', 'belly', 'harpooneers', 'caught', 'Spitzbergen', 'white', 'VOYAGE',</pre>
'GREENLAND', '1671', 'Several', 'coast', 'Fife', 'Anno', '1652', 'eighty', 'bone', 'kind',
<pre>'informed', '),', 'afford', '500', 'weight', 'baleen', 'stand', 'gate', 'garden',</pre>
'Pitferren', 'SIBBALD', 'FIFE', 'KINROSS', 'Myself', 'agreed', 'try', 'master', 'kill', 'Sperma', 'ceti', 'never',
<pre>'ceti', 'never', 'hear', 'sort', 'fierceness', 'swiftness', 'RICHARD', 'STRAFFORD', 'LETTER', 'BERMUDAS', 'PHIL',</pre>
'PHIL', 'TRANS', '1668', 'voice', 'obey', 'PRIMER', 'abundance', 'large', 'southern', 'may', 'hundred',
<pre>'hundred', 'northward', 'CAPTAIN', 'COWLEY', 'ROUND', 'GLOBE', '1729', '"', 'frequendy', 'attended', 'insupportable',</pre>
<pre>'insupportable', 'smell', 'bring', 'disorder', 'ULLOA', 'SOUTH', 'AMERICA', 'chosen', 'sylphs',</pre>
<pre>'special', 'note', 'trust', 'important', 'charge', 'petticoat', '0ft', 'fold', 'fence', 'fail', 'Tho',</pre>
'Tho', 'stuffed', 'hoops', 'armed', 'ribs', 'RAPE', 'LOCK', 'If', 'compare', 'animals', 'respect',
<pre>'respect', 'magnitude', 'abode', 'appear', 'contemptible', 'comparison', 'doubtless', 'largest', 'creation', 'GOLDSMITH', 'NAT',</pre>
'NAT', 'HIST', 'write', 'little', 'make', 'speak', 'wales', 'JOHNSON', 'afternoon', 'supposed',
<pre>'rock', 'found', 'dead', 'Asiatics', 'towing', 'ashore', 'seemed', 'endeavor', 'conceal', 'order',</pre>
<pre>'avoid', 'C00K', 'larger', 'seldom', 'venture', 'attack', 'dread', 'afraid', 'mention', 'names',</pre>
<pre>'carry', 'dung', 'lime', 'juniper', 'wood', 'articles', 'boats', 'terrify', 'prevent', 'approach',</pre>
<pre>'approach', 'UNO', 'VON', 'TROIL', 'LETTERS', 'ON', 'BANKS', 'SOLANDER', 'ICELAND', '1772', 'Spermacetti',</pre>
'Spermacetti', 'Nantuckois', 'active', 'fierce', 'requires', 'address', 'boldness', 'fishermen', 'JEFFERSON', 'MEMORIAL', 'MINISTER', '1778',



Out[48]:	184, 84, 1, 1, 3,		
Out[48]:	1, 2, 1, 1, 2, 2, 1, 612,		
	1070, 1, 28, 90, 110, 37, 2124, 253,		
	253, 230, 203, 2, 436, 1, 2, 44, 5, 1, 3, 13,		
	3, 13, 1, 80, 12, 173, 3,		
	43, 10, 4, 1, 30, 118, 205, 9,		
	12, 173, 3, 10, 43, 10, 4, 1, 30, 118, 205, 9, 37, 5, 69, 906, 133, 116, 11, 38, 10, 12, 2, 186, 37, 5, 1, 76, 3, 73, 489, 1, 38, 11, 1,		
	12, 2, 186, 37, 5, 1, 76,		
	489, 1, 38, 11, 1, 4, 1,		
	16, 11, 1, 35, 2, 10, 2,		
	63, 3, 310, 14, 1, 1, 1,		
	1, 16, 3, 1, 1, 1,		
	1, 1, 1, 2, 1, 2, 1,		
	1, 3, 1, 5, 1, 4, 2, 4,		
	1, 1, 2, 7, 1, 15, 161, 42,		
	4, 1, 107, 16, 11, 1, 35, 2, 10, 2, 63, 3, 310, 14, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,		
	1, 10, 1, 46, 4, 44, 5, 7, 237,		
	282, 222, 69, 75, 1, 1, 4,		
	4, 4, 1, 4, 4, 120, 45,		
	2, 55, 38, 5, 7, 4, 11, 282, 222, 69, 75, 1, 1, 1, 4, 80, 3, 4, 4, 11, 4, 4, 120, 45, 28, 3, 30, 174, 3, 3, 6, 11, 2, 2, 2, 12, 17, 88, 46, 1, 302, 149, 9, 2, 64, 161, 6, 6, 8, 147, 10, 129, 86, 1, 39, 1, 8, 1, 1, 1, 421, 2, 34, 889, 81, 1, 1, 421, 2, 34, 889, 81, 1, 1, 1, 421, 2, 34, 889, 81, 1, 1, 16, 56, 9, 24, 4, 45, 38, 66, 237, 1, 121, 125, 125, 120, 120, 120, 120, 120, 120, 120, 120		
	12, 17, 88, 46, 1, 302, 149,		
	64, 161, 6, 8, 147, 10, 129,		
	1, 39, 1, 8, 1, 9, 12, 24,		
	101, 1, 1, 421, 2, 34, 889, 81,		
	16, 56, 9, 24, 4, 538, 6,		
	1, 121, 155, 22, 10, 29, 3,		
	1, 197, 218, 15, 460, 12, 84, 182,		
	12, 84, 182, 1, 11, 39, 1, 1, 3, 705, 2, 1, 53, 29, 124,		
	16,		
	17, 45, 1, 1, 20, 1, 1,		
	28, 7, 64, 1, 369, 132, 6, 293,		
	4, 15, 4, 17, 45, 1, 1, 20, 1, 1, 50, 49, 28, 7, 64, 1, 369, 132, 6, 293, 1, 10, 3, 28, 111, 69, 2, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1		
	139, 44, 12, 10, 84, 1, 147, 87,		
	44, 12, 10, 84, 1, 147, 87, 232, 19, 178, 29, 5, 2, 237, 172, 3, 23, 1, 4, 4, 8, 4, 5, 433,		
	23, 1, 4, 4, 8, 4, 5, 433,		
	1, 188, 1, 20, 1, 78, 4, 49,		
	4, 330, 18, 54, 1, 11, 1, 7,		
	2, 2, 1, 1, 57, 22, 1, 2,		
	3, 158, 29, 1, 1, 78, 74, 3,		
	4, 330, 18, 54, 1, 11, 11, 1, 7, 4, 2, 2, 2, 1, 1, 57, 22, 1, 1, 78, 74, 3, 1, 78, 74, 3, 1, 1, 6, 285, 80, 11, 9, 10, 8, 20, 10, 12, 130, 107, 228, 75, 4, 4, 4, 48, 28, 11, 19, 1, 1, 98, 1, 1, 98, 1, 4, 7, 25, 88, 7, 26, 48, 6, 47, 38, 25, 62, 1, 1, 1, 98, 1, 4, 7, 25, 88, 7, 26, 48, 6, 47, 38, 25, 62, 1, 1, 1, 1, 98, 1, 4, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,		
	9, 10, 8, 20, 10, 12, 130, 107, 228,		
	75, 4, 4, 48, 28, 11, 19,		
	1, 98, 1, 4, 7, 25, 88,		
	26, 48, 6, 47, 38, 25, 62, 1,		
	20, 24, 33, 27, 31, 33, 17, 2,		
	2, 10, 17, 1, 15, 2, 1, 23, 1, 14, 132, 76, 54, 9, 6, 45, 13, 1, 1, 1, 3, 7, 1, 1, 1, 1, 8, 44, 13, 2,		
	9, 6, 45, 13, 1, 1, 1,		
	1, 1, 6, 1, 1, 38, 14, 120,		
	1, 8, 44, 13, 2, 162, 6,		
	1, 4, 1, 1, 1, 80, 3,		
	15, 3, 1, 1, 2, 49, 1, 32,		
	162, 6, 5, 1, 4, 1, 1, 1, 1, 80, 3, 624, 15, 3, 1, 1, 2, 49, 1, 32, 1, 26, 102, 87, 78, 12, 2, 9, 13, 90, 5, 3, 2, 37, 1, 1, 1, 1, 6, 12, 10, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,		
	13, 90, 5, 3, 2, 37, 1,		
	4, 1, 2, 11, 1, 6, 12,		
	39, 1, 1, 1, 1, 62, 17,		
	8, 24, 1, 7, 9, 4, 1, 23,		
	23, 7, 1, 1, 1, 2, 1,		
	73, 9, 6, 41, 11, 71, 1, 2,		
	1, 1, 21, 1, 174, 1, 236,		
	183, 3, 32, 63, 25, 1, 62, 26,		
	5, 39, 1, 2, 8, 3, 2,		
	1, 6, 5, 1, 1, 18, 1,		
	73, 9, 66, 41, 111, 71, 11, 2, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4,		
	3, 2, 1, 2, 1, 1, 52, 1, 4.		
	4, 4, 1, 5, 508, 1, 1,		
	1, 1, 1, 154, 1, 1, 1,		
	30, 1, 3, 20, 11, 1,		
	73, 34, 1, 5, 8, 84, 13,		
	1, 18, 16, 47, 185, 29, 1,		
	1, 54, 50, 26, 29, 9, 14,		
	3, 3, 30, 9, 269, 1, 1,		
	24, 48, 507, 50, 335, 16, 69, 150,		
	150, 1, 16, 2, 125, 3, 1, 1, 1, 19, 1, 115,		
	1, 1, 115, 4, 1, 1, 8, 1, 34,		
	4, 1, 1, 8, 1, 34, 36, 13, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,		
	18, 14, 12, 2, 32, 51, 1,		
	1, 47, 74, 96, 1, 67, 2, 3,		
	4, 26, 210, 2, 1, 45, 50, 2, 5, 1, 9, 55, 44,		
	191, 7, 1, 2, 1, 40, 1,		
	4, 50, 27, 4, 62, 9, 2,		
	11, 83, 4, 6, 1, 1, 1,		
	3, 49, 23, 30, 1, 1, 195,		
	4, 191, 7, 1, 2, 1, 40, 1, 1, 1, 1, 4, 50, 27, 4, 62, 9, 2, 17, 11, 13, 49, 23, 30, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,		
	1, 1, 46, 8, 1, 4, 77, 5, 230,		
	50, 4, 2, 1, 2, 4,		
	1, 6, 1, 11, 18, 1, 1,		
	1, 1, 1, 6, 1, 11, 18, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 5, 11, 15, 20, 1, 1, 1, 60, 7, 5, 16, 21,		
	1, 1, 6, 2, 17, 1, 5,		
	20, 1, 1, 60, 7, 5, 16, 21,		
	2, 6, 4, 6, 12, 19, 7, 2,		
	1, 2, 5, 247, 109, 47, 1,		
	15, 14, 115, 90, 2, 14, 33, 283,		
	4, 2, 64, 6, 1, 9, 24,		
	2, 6, 4, 6, 12, 19, 7, 2, 1, 2, 5, 247, 109, 47, 1, 1, 10, 115, 90, 2, 14, 33, 283, 4, 2, 64, 6, 1, 9, 24, 5, 8, 12, 15, 11, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1		
	1, 34, 10, 145, 3, 6, 5, 1,		
	1, 1, 11, 1, 2, 1, 2,		
	8, 11, 7, 4, 2, 28, 1,		
	2,		



'Moby': 84, 'Dick': 84, 'Herman': 1, 'Melville': 1, '1851': 3, 'ETYMOLOGY': 1, 'Supplied': 2, 'Late': 1, 'Consumptive': 1, 'Usher': 2, 'Grammar': 2,
'School': 1, 'The': 612, 'pale': 18, '': 1070, 'threadbare': 1, 'coat': 28, 'heart': 90, 'body': 110, 'brain': 37, 'I': 2124, 'see': 253,
'He': 230, 'ever': 203, 'dusting': 2, 'old': 436, 'lexicons': 1, 'grammars': 2, 'queer': 44, 'handkerchief': 5, 'mockingly': 1, 'embellished': 3, 'gay': 13, 'flags': 1,
'known': 80, 'nations': 12, 'world': 173, 'loved': 3, 'dust': 10, 'somehow': 43, 'mildly': 10, 'reminded': 4, 'mortality': 1, 'While': 30, 'take': 118,
'hand': 205, 'school': 9, 'others': 37, 'teach': 5, 'name': 69, 'whale': 906, 'fish': 133, 'called': 116, 'tongue': 11, 'leaving': 38, 'ignorance': 10, 'letter': 12,
'H': 2, 'almost': 186, 'alone': 37, 'maketh': 5, 'signification': 1, 'word': 76, 'deliver': 3, 'true': 73, '."': 489, 'HACKLUYT': 1, 'WHALE': 38,
'': 11, 'Sw': 1, 'Dan': 4, 'HVAL': 1, 'This': 107, 'animal': 16, 'named': 11, 'roundness': 1, 'rolling': 35, 'HVALT': 2, 'arched': 10, 'vaulted': 2,
<pre>'WEBSTER': 2, 'S': 63, 'DICTIONARY': 3, 'It': 310, 'immediately': 14, 'Dut': 1, 'Ger': 1, 'WALLEN': 1, 'A': 167, 'WALW': 1, 'IAN': 1, 'roll': 16,</pre>
<pre>'wallow': 3, 'RICHARDSON': 1, 'KETOS': 1, 'GREEK': 1, 'CETUS': 1, 'LATIN': 1, 'WHOEL': 1, 'ANGLO': 1, 'SAXON': 1, 'DANISH': 2, 'WAL': 1,</pre>
'DUTCH': 2, 'HWAL': 1, 'SWEDISH': 1, 'ICELANDIC': 1, 'ENGLISH': 3, 'BALEINE': 1, 'FRENCH': 5, 'BALLENA': 1, 'SPANISH': 4, 'PEKEE': 2, 'NUEE': 4, 'FEGEE': 1,
'ERROMANGOAN': 1, 'EXTRACTS': 2, 'Sub': 7, 'Librarian': 1, ').': 15, 'seen': 161, 'mere': 42, 'painstaking': 1, 'burrower': 1, 'grub': 1, 'worm': 5, 'poor': 106,
'devil': 51, 'appears': 7, 'gone': 58, 'long': 318, 'Vaticans': 1, 'street': 10, 'stalls': 1, 'earth': 46, 'picking': 4, 'whatever': 44, 'random': 5, 'allusions': 7,
<pre>'whales': 237, 'could': 215, 'anyways': 2, 'find': 55, 'book': 38, 'whatsoever': 5, 'sacred': 7, 'profane': 4, 'Therefore': 11, 'must': 282, 'every': 222, 'case': 69,</pre>
'least': 75, 'higgledy': 1, 'piggledy': 1, 'statements': 4, 'however': 80, 'authentic': 3, 'extracts': 4, 'veritable': 4, 'gospel': 1, 'cetology': 4, 'Far': 4, 'As': 120,
'touching': 45, 'ancient': 28, 'authors': 3, 'generally': 30, 'well': 174, 'poets': 3, 'appearing': 3, 'solely': 6, 'valuable': 11, 'entertaining': 2, 'affording': 2, 'glancing': 12,
'bird': 17, 'eye': 88, 'view': 46, 'promiscuously': 1, 'said': 302, 'thought': 149, 'fancied': 9, 'sung': 2, 'Leviathan': 64, 'many': 161, 'generations': 6, 'including': 8, 'So': 147,
'fare': 10, 'thee': 129, 'whose': 86, 'commentator': 1, 'Thou': 39, 'belongest': 1, 'hopeless': 8, 'sallow': 1, 'tribe': 9, 'wine': 12, 'warm': 24, 'even': 181,
'Pale': 1, 'Sherry': 1, 'would': 421, 'rosy': 2, 'strong': 34, 'one': 889, 'sometimes': 81, 'loves': 1, 'sit': 16, 'feel': 56, 'devilish': 9, 'grow': 24,
'convivial': 4, 'upon': 538, 'tears': 6, 'say': 237, 'bluntly': 1, 'full': 121, 'eyes': 155, 'empty': 22, 'glasses': 10, 'altogether': 29, 'unpleasant': 3, 'sadness': 2,
'Give': 15, 'Subs': 1, 'For': 197, 'much': 218, 'pains': 15, 'ye': 460, 'please': 12, 'shall': 84, 'go': 182, 'thankless': 1, 'Would': 11, 'clear': 39,
'Hampton': 1, 'Court': 1, 'Tuileries': 3, 'But': 705, 'gulp': 2, 'hie': 1, 'aloft': 53, 'royal': 29, 'mast': 124, 'hearts': 29, 'friends': 16, 'clearing': 4, 'seven': 15,
'storied': 4, 'heavens': 17, 'making': 45, 'refugees': 1, 'pampered': 1, 'Gabriel': 20, 'Michael': 1, 'Raphael': 1, 'coming': 50, 'Here': 49, 'strike': 28, 'splintered': 7,
'together': 64, 'unsplinterable': 1, 'And': 369, 'God': 132, 'created': 6, 'great': 293, 'GENESIS': 1, 'path': 10, 'shine': 3, 'One': 28, 'think': 111, 'deep': 69,
'hoary': 2, 'JOB': 1, 'Now': 139, 'Lord': 44, 'prepared': 12, 'swallow': 10, 'Jonah': 84, 'JONAH': 1, 'There': 147, 'ships': 87, 'thou': 232, 'hast': 19,
<pre>'made': 178, 'play': 29, 'therein': 5, 'PSALMS': 2, 'In': 237, 'day': 172, 'sore': 3, 'sword': 23, 'punish': 1, 'piercing': 4, 'serpent': 4, 'crooked': 8, 'slay': 4,</pre>
'dragon': 5, 'sea': 433, 'ISAIAH': 1, 'thing': 188, 'soever': 1, 'besides': 20, 'cometh': 1, 'within': 78, 'chaos': 4, 'monster': 49, 'mouth': 60, 'beast': 4,
'boat': 330, 'stone': 18, 'goes': 54, 'incontinently': 1, 'foul': 11, 'perisheth': 1, 'bottomless': 7, 'gulf': 4, 'paunch': 2, 'HOLLAND': 2, 'PLUTARCH': 1, 'MORALS': 1,
'Indian': 57, 'Sea': 22, 'breedeth': 1, 'biggest': 2, 'fishes': 3, 'among': 158, 'Whales': 29, 'Whirlpooles': 1, 'Balaene': 1, 'length': 78, 'four': 74, 'acres': 3, 'arpens': 1,
'land': 78, 'PLINY': 1, 'Scarcely': 1, 'proceeded': 6, 'two': 285, 'days': 80, 'sunrise': 11, 'monsters': 9, 'appeared': 10, 'Among': 8, 'former': 20, 'monstrous': 10,
<pre>'size': 12, 'came': 130, 'towards': 107, 'us': 228, 'open': 75, 'mouthed': 4, 'raising': 4, 'waves': 48, 'sides': 28, 'beating': 11, 'foam': 19, 'TOOKE': 1, 'LUCIAN': 1,</pre>
'THE': 98, 'TRUE': 1, 'HISTORY': 4, 'visited': 7, 'country': 25, 'also': 88, 'catching': 7, 'horse': 26, 'bones': 48, 'value': 6, 'teeth': 47, 'brought': 38,
<pre>'king': 25, 'best': 62, 'catched': 1, 'forty': 28, 'eight': 24, 'fifty': 33, 'yards': 27, 'six': 31, 'killed': 33, 'sixty': 17, 'OTHER': 2, 'OR': 5, 'OCTHER': 1,</pre>
'VERBAL': 1, 'NARRATIVE': 4, 'TAKEN': 1, 'DOWN': 2, 'FROM': 10, 'HIS': 17, 'MOUTH': 1, 'BY': 15, 'KING': 2, 'ALFRED': 1, 'D': 23, '890': 1,
<pre>'whereas': 14, 'things': 132, 'whether': 76, 'vessel': 54, 'enter': 9, 'dreadful': 6, 'lost': 45, 'swallowed': 13, 'gudgeon': 1, 'retires': 1, 'security': 3, 'sleeps': 7, 'MONTAIGNE': 1,</pre>
'APOLOGY': 1, 'FOR': 6, 'RAIMOND': 1, 'SEBOND': 1, 'Let': 38, 'fly': 14, 'let': 120, 'Old': 11, 'Nick': 1, 'described': 8, 'noble': 44, 'prophet': 13, 'Moses': 2,
'life': 162, 'patient': 6, 'Job': 5, 'RABELAIS': 1, 'liver': 4, 'cartloads': 1, 'STOWE': 1, 'ANNALS': 1, 'seas': 80, 'seethe': 3, 'like': 624, 'boiling': 15,
'pan': 3, 'LORD': 1, 'BACON': 1, 'VERSION': 2, 'OF': 49, 'Touching': 1, 'bulk': 32, 'ork': 1, 'received': 26, 'nothing': 102, 'certain': 87, 'They': 78, 'exceeding': 12,
'fat': 12, 'insomuch': 2, 'incredible': 9, 'quantity': 13, 'oil': 90, 'extracted': 5, 'IBID': 3, 'LIFE': 2, 'AND': 37, 'DEATH': 1, 'sovereignest': 1, 'parmacetti': 3, 'inward': 4,
'bruise': 1, 'HENRY': 2, 'Very': 11, 'HAMLET': 1, 'Which': 6, 'secure': 12, 'skill': 10, 'leach': 1, 'art': 39, 'Mote': 1, 'availle': 1, 'returne': 1, 'againe': 1,
'To': 62, 'wound': 17, 'worker': 1, 'lowly': 8, 'dart': 24, 'Dinting': 1, 'breast': 7, 'bred': 9, 'restless': 4, 'paine': 1, 'Like': 23, 'wounded': 9, 'shore': 23,
'flies': 7, 'thro': 1, 'maine': 1, 'FAERIE': 1, 'QUEEN': 2, 'Immense': 1, 'motion': 17, 'vast': 73, 'bodies': 9, 'peaceful': 6, 'calm': 41, 'trouble': 11, 'ocean': 71,
<pre>'til': 1, 'boil': 2, 'SIR': 4, 'WILLIAM': 2, 'DAVENANT': 1, 'PREFACE': 1, 'TO': 21, 'GONDIBERT': 1, 'What': 174, 'spermacetti': 1, 'men': 236, 'might': 183,</pre>
'justly': 3, 'doubt': 32, 'since': 63, 'learned': 25, 'Hosmannus': 1, 'work': 62, 'thirty': 26, 'years': 94, 'saith': 5, 'plainly': 39, 'Nescio': 1, 'quid': 2,
'T': 8, 'BROWNE': 3, 'SPERMA': 2, 'CETI': 2, 'VIDE': 1, 'V': 6, 'E': 5, 'Spencer': 1, 'Talus': 1, 'modern': 18, 'flail': 1, 'threatens': 1, 'ruin': 5,
'ponderous': 13, 'tail': 78, 'Their': 6, 'fixed': 37, 'jav': 1, 'lins': 1, 'side': 208, 'wears': 7, 'back': 151, 'grove': 3, 'pikes': 2, 'WALLER': 1,
'BATTLE': 2, 'SUMMER': 1, 'ISLANDS': 1, 'By': 52, 'Commonwealth': 1, 'State': 4, '(': 4, 'Latin': 4, 'Civitas': 1, 'artificial': 5, 'man': 508, 'OPENING': 1,
'SENTENCE': 1, 'HOBBES': 1, 'LEVIATHAN': 1, 'Silly': 1, 'Mansoul': 1, 'without': 154, 'chewing': 1, 'sprat': 1, 'PILGRIM': 1, 'PROGRESS': 1, 'That': 85, 'works': 30, 'Created': 1,
'hugest': 3, 'swim': 20, 'stream': 11, 'PARADISE': 1, 'LOST': 1, '"': 1, 'Hugest': 1, 'living': 73, 'creatures': 34, 'Stretched': 1, 'promontory': 5, 'swims': 8,
<pre>'seems': 84, 'moving': 13, 'gills': 4, 'Draws': 1, 'breath': 18, 'spouts': 16, 'mighty': 47, 'water': 185, 'swimming': 29, 'FULLER': 1, 'PROFANE': 1, 'HOLY': 1, 'STATE': 1,</pre>
'close': 54, 'behind': 50, 'lie': 26, 'huge': 29, 'attend': 9, 'prey': 14, 'give': 71, 'chance': 46, 'fry': 3, 'gaping': 3, 'jaws': 30, 'mistake': 9,
'way': 269, 'DRYDEN': 1, 'ANNUS': 1, 'MIRABILIS': 1, 'floating': 24, 'stern': 48, 'ship': 507, 'cut': 50, 'head': 335, 'tow': 16, 'near': 69, 'come': 150, 'aground': 1,
'twelve': 16, 'thirteen': 2, 'feet': 125, 'THOMAS': 3, 'EDGE': 1, 'TEN': 1, 'VOYAGES': 3, 'SPITZBERGEN': 1, 'IN': 19, 'PURCHAS': 1, 'saw': 115, 'sporting': 4,
<pre>'wantonness': 1, 'fuzzing': 1, 'pipes': 8, 'vents': 1, 'nature': 34, 'placed': 36, 'shoulders': 13, 'HERBERT': 1, 'INTO': 1, 'ASIA': 1, 'AFRICA': 1, 'HARRIS': 2, 'COLL': 2,</pre>
'troops': 2, 'forced': 18, 'proceed': 14, 'deal': 12, 'caution': 2, 'fear': 32, 'run': 51, 'SCHOUTEN': 1, 'SIXTH': 1, 'CIRCUMNAVIGATION': 1, 'We': 47, 'set': 74, 'sail': 96, 'Elbe': 1,
<pre>'wind': 67, 'N': 2, 'Jonas': 3, 'Whale': 282, 'Some': 38, 'fable': 5, 'frequently': 14, 'climb': 4, 'masts': 26, 'first': 210, 'discoverer': 2, 'ducat': 1, 'told': 45,</pre>
'taken': 50, 'Shetland': 2, 'barrel': 5, 'herrings': 1, 'belly': 9, 'harpooneers': 55, 'caught': 44, 'Spitzbergen': 4, 'white': 191, 'VOYAGE': 7, 'GREENLAND': 1, '1671': 2,
'Several': 1, 'coast': 40, 'Fife': 1, 'Anno': 1, '1652': 1, 'eighty': 4, 'bone': 50, 'kind': 27, 'informed': 4, '),': 62, 'afford': 9, '500': 2, 'weight': 17,
<pre>'baleen': 11, 'stand': 83, 'gate': 4, 'garden': 6, 'Pitferren': 1, 'SIBBALD': 1, 'FIFE': 1, 'KINROSS': 1, 'Myself': 1, 'agreed': 3, 'try': 49, 'master': 23, 'kill': 30,</pre>
<pre>'Sperma': 1, 'ceti': 1, 'never': 195, 'hear': 60, 'sort': 152, 'fierceness': 1, 'swiftness': 6, 'RICHARD': 1, 'STRAFFORD': 1, 'LETTER': 1, 'BERMUDAS': 1, 'PHIL': 1,</pre>
'TRANS': 1, '1668': 1, 'voice': 46, 'obey': 8, 'PRIMER': 1, 'abundance': 4, 'large': 77, 'southern': 5, 'may': 230, 'hundred': 50, 'northward': 4, 'CAPTAIN': 2, 'COWLEY': 1, 'ROUND': 2,
'ROUND': 2, 'GLOBE': 4, '1729': 1, '"': 1, 'frequendy': 1, 'attended': 6, 'insupportable': 1, 'smell': 11, 'bring': 18, 'disorder': 1, 'ULLOA': 1, 'SOUTH': 1, 'AMERICA': 1,
<pre>'chosen': 3, 'sylphs': 1, 'special': 15, 'note': 6, 'trust': 4, 'important': 19, 'charge': 10, 'petticoat': 1, 'oft': 1, 'fold': 6, 'fence': 2, 'fail': 17, 'Tho': 1,</pre>
<pre>'stuffed': 5, 'hoops': 11, 'armed': 15, 'ribs': 20, 'RAPE': 1, 'LOCK': 1, 'If': 60, 'compare': 7, 'animals': 5, 'respect': 16, 'magnitude': 21, 'abode': 2, 'appear': 6,</pre>
<pre>'contemptible': 4, 'comparison': 6, 'doubtless': 12, 'largest': 19, 'creation': 7, 'GOLDSMITH': 2, 'NAT': 1, 'HIST': 2, 'write': 5, 'little': 247, 'make': 109, 'speak': 47, 'wales': 1,</pre>
<pre>'wales': 1, 'JOHNSON': 1, 'afternoon': 10, 'supposed': 15, 'rock': 14, 'found': 115, 'dead': 90, 'Asiatics': 2, 'towing': 14, 'ashore': 33, 'seemed': 283, 'endeavor': 4, 'conceal': 2, 'order': 64,</pre>
<pre>'order': 64, 'avoid': 6, 'COOK': 1, 'larger': 9, 'seldom': 24, 'venture': 5, 'attack': 8, 'dread': 12, 'afraid': 15, 'mention': 15, 'names': 12, 'carry': 25, 'dung': 1, 'lime': 1,</pre>
'lime': 1, 'juniper': 1, 'wood': 34, 'articles': 10, 'boats': 145, 'terrify': 3, 'prevent': 6, 'approach': 5, 'UNO': 1, 'VON': 1, 'TROIL': 1, 'LETTERS': 1, 'ON': 11, 'BANKS': 1,
'SOLANDER': 1, 'ICELAND': 2, '1772': 1, 'Spermacetti': 2, 'Nantuckois': 1, 'active': 8, 'fierce': 11, 'requires': 7, 'address': 4, 'boldness': 2, 'fishermen': 28, 'JEFFERSON': 1, 'MEMORIAL': 1, 'MINISTER': 1,
'MINISTER': 1, '1778': 2, 'pray': 6,

