

AI Lab (ID5003W) - Assignment 1 - CH22M503

Author - Aloy Banerjee

Roll No. CH22M503

Google colab link : -

<https://colab.research.google.com/drive/10YSYUANd1ftJqbHd-ExCW4kFINIGN1ML?usp=sharing>

Submission Deadline: 31/05/2023 - 23:59

Question: 01

MAP REDUCE - DATASET: write map reduce program(s) using python-mr in colab notebook and complete the following exercises

- Find the files which contain the following phrases
 - "draping the folds of her velvet dress"
 - "grim clasp to the iron stanchions"
 - "Lo, I teach you the Superman!"
 - "Through that divine allegiance"
- Find the file which has maximum number of tokens
- Find the file which has the longest sentence (a sentence ends in a full stop and can span multiple lines)

Dataset:

- [pg1513.txt](#)
- [2160-0.txt](#)
- [pg11.txt](#)
- [pg2600.txt](#)
- [345.txt.utf-8](#)
- [64317.txt.utf-8](#)
- [70815.txt.utf-8](#)
- [1080-0.txt](#)
- [70817.txt.utf-8](#)
- [1998.txt.utf-8](#)

Answer: 01

Supporting URL

<https://mrjob.readthedocs.io/en/latest/>

Importing common library

```
In [1]: import os
```

```
import urllib.request
```

```
In [2]: basePath = '/content/sample_data/'
```

```
In [3]: os.chdir(basePath)
!mkdir 'Assignment 1'
```

```
In [4]: workingPath = '/content/sample_data/Assignment 1' #'C:/Users/Admin/Coding-M-Tech/AI Lab 2
```

```
In [5]: !pip install --quiet mrjob
```

6 MB/s eta 0:00:00 439.6/439.6 kB 8.

Changing the working directory

```
In [6]: os.chdir(workingPath)
```

Common Variable for downloading the files

```
In [7]: ## All the downloadable url
url_pg1513 = 'https://www.gutenberg.org/cache/epub/1513/pg1513.txt'
url_2160 = 'https://www.gutenberg.org/files/2160/2160-0.txt'
url_pg11 = 'https://www.gutenberg.org/cache/epub/11/pg11.txt'
url_pg2600 = 'https://www.gutenberg.org/cache/epub/2600/pg2600.txt'
url_pg345 = 'https://www.gutenberg.org/cache/epub/345/pg345.txt'
url_pg64317 = 'https://www.gutenberg.org/cache/epub/64317/pg64317.txt'
url_pg70815 = 'https://www.gutenberg.org/cache/epub/70815/pg70815.txt'
url_1080 = 'https://www.gutenberg.org/files/1080/1080-0.txt'
url_pg70817 = 'https://www.gutenberg.org/cache/epub/70817/pg70817.txt'
url_pg1998 = 'https://www.gutenberg.org/cache/epub/1998/pg1998.txt'
```

Map reduce Functionality

Phrase Searcher

```
In [8]: %%file phrase_search.py
import os
from mrjob.job import MRJob
from mrjob.step import MRStep

class FindPhrase(MRJob):

    def configure_args(self):
        super(FindPhrase, self).configure_args()
        self.add_passthru_arg('--phrase', type=str, help='Phrase to find in files')

    def mapper_init(self):
        self.phrase = self.options.phrase

    def mapper(self, _, line):
        full_filename = os.getenv('mapreduce_map_input_file')
        file_name = full_filename.split('/')[-1]
        if self.phrase in line:
            yield (self.phrase, file_name), 1

    def reducer(self, key, values):
        yield key, f"{sum(values)} occurrence of : {key[0]} phrase, is present in : {key[0]}"

    def steps(self):
        return [
```

```

        MRStep(mapper_init=self.mapper_init, mapper=self.mapper, reducer=self.reduce
    ]
if __name__ == '__main__':
    FindPhrase.run()

```

Writing phrase_search.py

Max Token Searcher

```

In [9]: %%file max_token_search.py
import os
from mrjob.job import MRJob
from mrjob.step import MRStep

class FindMaxToken(MRJob):

    def mapper(self, _, line):
        full_filename = os.getenv('mapreduce_map_input_file')
        tokens, file_name = line.strip().split(), full_filename.split('/')[-1]
        yield file_name, len(tokens)

    def reducer(self, key, values):
        yield key, f"file has token value of : {sum(values)}"

    def steps(self):
        return [
            MRStep(mapper=self.mapper, reducer=self.reducer)
        ]

if __name__ == '__main__':
    FindMaxToken.run()

```

Writing max_token_search.py

Maximum Length Sentence

```

In [10]: %%file max_length_search.py
import os
from mrjob.job import MRJob
from mrjob.step import MRStep

class MaxSentenceLength(MRJob):

    def mapper_init(self):
        self.current_sentence = ''

    def mapper(self, _, value):
        file_name, modified_line = os.getenv('mapreduce_map_input_file'), ' '.join(value)
        if '.' in modified_line:
            updated_sentence = self.current_sentence + ' ' if self.current_sentence else
            self.current_sentence = ' '.join(modified_line.split('.')[1:]) if len(modif
            yield file_name, max((updated_sentence + ' ' + modified_line.split('.')[0]).
        else:
            self.current_sentence += ' ' + modified_line if modified_line != '' else '

    def reducer(self, key, values):
        sentence = max(list(values), key=len)
        sentence_length = len(sentence)
        yield key, f"Maximum length of sentence is {sentence_length} from {key} file. Lo

    def steps(self):
        return [
            MRStep(mapper_init=self.mapper_init, mapper=self.mapper, reducer=self.reduce
        ]

```

```
if __name__ == '__main__':  
    MaxSentenceLength.run()
```

Writing max_length_search.py

Creating a dataset folder

```
In [11]: !mkdir Q1-Dataset
```

Downloading the file

```
In [12]: urllib.request.urlretrieve(url_pg1513, 'Q1-Dataset/pg1513.txt')  
urllib.request.urlretrieve(url_2160, 'Q1-Dataset/2160-0.txt')  
urllib.request.urlretrieve(url_pg11, 'Q1-Dataset/pg11.txt')  
urllib.request.urlretrieve(url_pg2600, 'Q1-Dataset/pg2600.txt')  
urllib.request.urlretrieve(url_pg345, 'Q1-Dataset/pg345.txt')  
urllib.request.urlretrieve(url_pg64317, 'Q1-Dataset/pg64317.txt')  
urllib.request.urlretrieve(url_pg70815, 'Q1-Dataset/pg70815.txt')  
urllib.request.urlretrieve(url_1080, 'Q1-Dataset/1080-0.txt')  
urllib.request.urlretrieve(url_pg70817, 'Q1-Dataset/pg70817.txt')  
urllib.request.urlretrieve(url_pg1998, 'Q1-Dataset/pg1998.txt')
```

```
Out[12]: ('Q1-Dataset/pg1998.txt', <http.client.HTTPMessage at 0x7fb6d46fd000>)
```

1a. Running Phrase Finder method

```
In [13]: !python phrase_search.py --phrase="draping the folds of her velvet dress" "Q1-Dataset/"  
  
!python phrase_search.py --phrase="grim clasp to the iron stanchions" "Q1-Dataset/"  
  
!python phrase_search.py --phrase="Lo, I teach you the Superman!" "Q1-Dataset/"  
  
!python phrase_search.py --phrase="Through that divine allegiance" "Q1-Dataset/"  
  
No configs found; falling back on auto-configuration  
No configs specified for inline runner  
Creating temp directory /tmp/phrase_search.root.20230612.164955.752218  
Running step 1 of 1...  
job output is in /tmp/phrase_search.root.20230612.164955.752218/output  
Streaming final output from /tmp/phrase_search.root.20230612.164955.752218/output...  
["draping the folds of her velvet dress", "pg2600.txt"] "1 occurrence of : draping the fo  
lds of her velvet dress phrase, is present in : pg2600.txt file"  
Removing temp directory /tmp/phrase_search.root.20230612.164955.752218...  
No configs found; falling back on auto-configuration  
No configs specified for inline runner  
Creating temp directory /tmp/phrase_search.root.20230612.164956.258601  
Running step 1 of 1...  
job output is in /tmp/phrase_search.root.20230612.164956.258601/output  
Streaming final output from /tmp/phrase_search.root.20230612.164956.258601/output...  
["grim clasp to the iron stanchions", "pg345.txt"] "1 occurrence of : grim clasp to  
the iron stanchions phrase, is present in : pg345.txt file"  
Removing temp directory /tmp/phrase_search.root.20230612.164956.258601...  
No configs found; falling back on auto-configuration  
No configs specified for inline runner  
Creating temp directory /tmp/phrase_search.root.20230612.164956.763957  
Running step 1 of 1...  
job output is in /tmp/phrase_search.root.20230612.164956.763957/output  
Streaming final output from /tmp/phrase_search.root.20230612.164956.763957/output...  
["Lo, I teach you the Superman!", "pg1998.txt"] "1 occurrence of : Lo, I teach you the Su  
perman! phrase, is present in : pg1998.txt file"  
Removing temp directory /tmp/phrase_search.root.20230612.164956.763957...  
No configs found; falling back on auto-configuration  
No configs specified for inline runner
```

```
Creating temp directory /tmp/phrase_search.root.20230612.164957.273619
Running step 1 of 1...
job output is in /tmp/phrase_search.root.20230612.164957.273619/output
Streaming final output from /tmp/phrase_search.root.20230612.164957.273619/output...
Removing temp directory /tmp/phrase_search.root.20230612.164957.273619...
```

Conclusion 1a .

"draping the folds of her velvet dress" phrase is available in **pg2600.txt** file

"grim clasp to the iron stanchions" phrase is available in **pg345.txt** file

"Lo, I teach you the Superman!" phrase is available in **pg1998.txt** file

"Through that divine allegiance" phrase is not available in any text file.

1b. Running Max Token Finder method

```
In [14]: !python max_token_search.py "Q1-Dataset/*"

No configs found; falling back on auto-configuration
No configs specified for inline runner
Creating temp directory /tmp/max_token_search.root.20230612.164957.776967
Running step 1 of 1...
job output is in /tmp/max_token_search.root.20230612.164957.776967/output
Streaming final output from /tmp/max_token_search.root.20230612.164957.776967/output...
"pg2600.txt"      " file has token value of : 566330"
"pg345.txt"       " file has token value of : 164459"
"pg64317.txt"     " file has token value of : 51257"
"pg70815.txt"     " file has token value of : 82782"
"pg70817.txt"     " file has token value of : 28184"
"1080-0.txt"      " file has token value of : 6512"
"2160-0.txt"      " file has token value of : 151379"
"pg11.txt"        " file has token value of : 29590"
"pg1513.txt"      " file has token value of : 29002"
"pg1998.txt"      " file has token value of : 113997"
Removing temp directory /tmp/max_token_search.root.20230612.164957.776967...
```

Conclusion 1b.

pg2600.txt has maximum number of token i.e. **566330**

1c. Running Find Max Sentence Length method

```
In [15]: !python max_length_search.py "Q1-Dataset/*"

No configs found; falling back on auto-configuration
No configs specified for inline runner
Creating temp directory /tmp/max_length_search.root.20230612.164959.198689
Running step 1 of 1...
job output is in /tmp/max_length_search.root.20230612.164959.198689/output
Streaming final output from /tmp/max_length_search.root.20230612.164959.198689/output...
"file://Q1-Dataset/pg2600.txt" "Maximum length of sentence is 1757 from file://Q1-Datas
et/pg2600.txt file. Longest Sentence of file://Q1-Dataset/pg2600.txt file is : But Count
Rostopchn, who now taunted those who left Moscow and now had the government offices remo
ved; now distributed quite useless weapons to the drunken rabble; now had processions di
splaying the icons, and now forbade Father Augustin to remove icons or the relics of sai
nts; now seized all the private carts in Moscow and on one hundred and thirty-six of the
m removed the balloon that was being constructed by Leppich; now hinted that he would bu
rn Moscow and related how he had set fire to his own house; now wrote a proclamation to
the French solemnly upbraiding them for having destroyed his Orphanage; now claimed the
glory of having hinted that he would burn Moscow and now repudiated the deed; now ordere
```

the people to catch all the spies and bring them to him, and now reproached them for doing so; now expelled all the French residents from Moscow, and now allowed Madame Aubert-Chalm (the center of the whole French colony in Moscow) to remain, but ordered the venerable old postmaster Klyucharov to be arrested and exiled for no particular offense; now assembled the people at the Three Hills to fight the French and now, to get rid of them, handed over to them a man to be killed and himself drove away by a back gate; now declared that he would not survive the fall of Moscow, and now wrote French verses in albums concerning his share in the affair this man did not understand the meaning of what was happening but merely wanted to do something himself that would astonish people, to perform some patriotically heroic feat; and like a child he made sport of the momentous, and unavoidable event the abandonment and burning of Moscow and tried with his puny hand now to speed and now to stay the enormous, popular tide that bore him along with it"

"file:///Q1-Dataset/pg345.txt" "Maximum length of sentence is 976 from file:///Q1-Dataset/pg345.txt file. Longest Sentence of file:///Q1-Dataset/pg345.txt file is : 'Who more gladly than we throughout the Four Nations received the 'bloody sword,' or at its warlike call flocked quicker to the standard of the King? When was redeemed that great shame of my nation, the shame of Cassova, when the flags of the Wallach and the Magyar went down beneath the Crescent? Who was it but one of my own race who as Voivode crossed the Danube and beat the Turk on his own ground? This was a Dracula indeed! Woe was it that his own unworthy brother, when he had fallen, sold his people to the Turk and brought the shame of slavery on them! Was it not this Dracula, indeed, who inspired that other of his race who in a later age again and again brought his forces over the great river into Turkey-land; who, when he was beaten back, came again, and again, and again, though he had to come alone from the bloody field where his troops were being slaughtered, since he knew that he alone could ultimately triumph! They said that he thought only of himself"

"file:///Q1-Dataset/pg64317.txt" "Maximum length of sentence is 417 from file:///Q1-Dataset/pg64317.txt file. Longest Sentence of file:///Q1-Dataset/pg64317.txt file is : Its vanished trees, the trees that had made way for Gatsby's house, had once pandered in whispers to the last and greatest of all human dreams; for a transitory enchanted moment man must have held his breath in the presence of this continent, compelled into an aesthetic contemplation he neither understood nor desired, face to face for the last time in history with something commensurate to his capacity for wonder"

"file:///Q1-Dataset/pg70815.txt" "Maximum length of sentence is 802 from file:///Q1-Dataset/pg70815.txt file. Longest Sentence of file:///Q1-Dataset/pg70815.txt file is : The sun, beating down upon his naked body, had no such effect upon his skin as would the sun of the outer world under like conditions, but it did impart to him a golden bronze color, which gave him a new confidence similar to that which he would have felt had he been able to retrieve his lost apparel, and in this fact he saw what he believed to be the real cause of his first embarrassment at his nakedness--it had been the whiteness of his skin that had made him seem so naked by contrast with other creatures, for this whiteness had suggested softness and weakness, arousing within him a disturbing sensation of inferiority; but now as he took on his heavy coat of tan and his feet became hardened and accustomed to the new conditions, he walked no longer in constant realization of his nakedness"

"file:///Q1-Dataset/pg70817.txt" "Maximum length of sentence is 402 from file:///Q1-Dataset/pg70817.txt file. Longest Sentence of file:///Q1-Dataset/pg70817.txt file is : Despite these efforts, Project Gutenberg electronic works, and the medium on which they may be stored, may contain Defects, such as, but not limited to, incomplete, inaccurate or corrupt data, transcription errors, a copyright or other intellectual property infringement, a defective or damaged disk or other medium, a computer virus, or computer codes that damage or cannot be read by your equipment"

"file:///Q1-Dataset/1080-0.txt" "Maximum length of sentence is 1045 from file:///Q1-Dataset/1080-0.txt file. Longest Sentence of file:///Q1-Dataset/1080-0.txt file is : And secondly, There being a round million of creatures in humane figure throughout this kingdom, whose whole subsistence put into a common stock, would leave them in debt two million of pounds sterling, adding those who are beggars by profession, to the bulk of farmers, cottagers and labourers, with their wives and children, who are beggars in effect; I desire those politicians who dislike my overture, and may perhaps be so bold to attempt an answer, that they will first ask the parents of these mortals, whether they would not at this day think it a great happiness to have been sold for food at a year old, in the manner I prescribe, and thereby have avoided such a perpetual scene of misfortunes, as they have since gone through, by the oppression of landlords, the impossibility of paying rent without money or trade, the want of common sustenance, with neither house nor clothes to cover them from the inclemencies of the weather, and the most inevitable prospect of intailing the like, or greater miseries, upon their breed for ever"

"file:///Q1-Dataset/2160-0.txt" "Maximum length of sentence is 2079 from file:///Q1-Datas

et/2160-0.txt file. Longest Sentence of file://Q1-Dataset/2160-0.txt file is :--The word s fiquidis perfusus odoribus, he translates van civet & moschata gestinken: that individuals differed toto coelo in their opinion of smells, which, indeed, was altogether as arbitrary as the opinion of beauty; that the French were pleased with the putrid effluvia of animal food; and so were the Hottentots in Africa, and the Savages in Greenland; and that the Negroes on the coast of Senegal would not touch fish till it was rotten; strong presumptions in favour of what is generally called stink, as those nations are in a state of nature, undebauched by luxury, unseduced by whim and caprice: that he had reason to believe the stercoreaceous flavour, condemned by prejudice as a stink, was, in fact, most agreeable to the organs of smelling; for, that every person who pretended to nauseate the smell of anothers excretions, snuffed up his own with particular complacency; for the truth of which he appealed to all the ladies and gentlemen then present: he said, the inhabitants of Madrid and Edinburgh found particular satisfaction in breathing their own atmosphere, which was always impregnated with stercoreaceous effluvia: that the learned Dr B--, in his treatise on the Four Digestions, explains in what manner the volatile effluvia from the intestines stimulate and promote the operations of the animal economy: he affirmed, the last Grand Duke of Tuscany, of the Medicis family, who refined upon sensuality with the spirit of a philosopher, was so delighted with that odour, that he caused the essence of ordure to be extracted, and used it as the most delicious perfume: that he himself (the doctor) when he happened to be low-spirited, or fatigued with business, found immediate relief and uncommon satisfaction from hanging over the stale contents of a close-stool, while his servant stirred it about under his nose; nor was this effect to be wondered at, when we consider that this substance abounds with the self-same volatile salts that are so greedily smelled to by the most delicate invalids, after they have been extracted and sublimed by the chemists"

"file://Q1-Dataset/pg11.txt" "Maximum length of sentence is 920 from file://Q1-Dataset/pg11.txt file. Longest Sentence of file://Q1-Dataset/pg11.txt file is : Hardly knowing what she did, she picked up a little bit of stick, and held it out to the puppy; whereupon the puppy jumped into the air off all its feet at once, with a yelp of delight, and rushed at the stick, and made believe to worry it; then Alice dodged behind a great thistle, to keep herself from being run over; and the moment she appeared on the other side, the puppy made another rush at the stick, and tumbled head over heels in its hurry to get hold of it; then Alice, thinking it was very like having a game of play with a cart-horse, and expecting every moment to be trampled under its feet, ran round the thistle again; then the puppy began a series of short charges at the stick, running a very little way forwards each time and a long way back, and barking hoarsely all the while, till at last it sat down a good way off, panting, with its tongue hanging out of its mouth, and its great eyes half shut"

"file://Q1-Dataset/pg1513.txt" "Maximum length of sentence is 845 from file://Q1-Dataset/pg1513.txt file. Longest Sentence of file://Q1-Dataset/pg1513.txt file is : And in this state she gallops night by night Through lovers brains, and then they dream of love; O'er courtiers knees, that dream on curtsies straight; O'er lawyers fingers, who straight dream on fees; O'er ladies lips, who straight on kisses dream, Which oft the angry Mab with blisters plagues, Because their breaths with sweetmeats tainted are: Sometime she gallops o'er a courtiers nose, And then dreams he of smelling out a suit; And sometime comes she with a tithe-pigs tail, Tickling a parsons nose as a lies asleep, Then dreams he of another benefice: Sometime she driveth o'er a soldiers neck, And then dreams he of cutting foreign throats, Of breaches, ambuscados, Spanish blades, Of healths five fathom deep; and then anon Drums in his ear, at which he starts and wakes; And, being thus frightened, swears a prayer or two, And sleeps again"

"file://Q1-Dataset/pg1998.txt" "Maximum length of sentence is 1084 from file://Q1-Dataset/pg1998.txt file. Longest Sentence of file://Q1-Dataset/pg1998.txt file is : Another ideal runs on before us, a strange, tempting ideal full of danger, to which we should not like to persuade any one, because we do not so readily acknowledge any ones RIGHT THERETO: the ideal of a spirit who plays naively (that is to say involuntarily and from overflowing abundance and power) with everything that has hitherto been called holy, good, intangible, or divine; to whom the loftiest conception which the people have reasonably made their measure of value, would already practically imply danger, ruin, abasement, or at least relaxation, blindness, or temporary self-forgetfulness; the ideal of a humanly superhuman welfare and benevolence, which will often enough appear INHUMAN, for example, when put alongside of all past seriousness on earth, and alongside of all past solemnities in bearing, word, tone, look, morality, and pursuit, as their truest involuntary parody and WITH which, nevertheless, perhaps THE GREAT SERIOUSNESS only commences, when the proper interrogative mark is set up, the fate of the soul changes, the hour-hand moves, and tragedy begins"

Removing temp directory /tmp/max_length_search.root.20230612.164959.198689...

Conclusion 1c.

Maximum length of sentence is **2079** from **2160-0.txt** file

Conclusion of Question 1:

Q. Find the files which contain the following phrases

A. **"draping the folds of her velvet dress"** phrase is available in **pg2600.txt** file

"grim clasp to the iron stanchions" phrase is available in **pg345.txt** file

"Lo, I teach you the Superman!" phrase is available in **pg1998.txt** file

Q. Find the file which has maximum number of tokens

A. **pg2600.txt** has maximum number of token i.e. **566330**

Q. Find the file which has the longest sentence (a sentence ends in a full stop and can span multiple lines)

A. Maximum length of sentence is **2079** from **2160-0.txt** file

Question: 02

Write a PySpark RDD program to compute:

- The list of common (overlapping / intersection) words across all the ten files listed above.
- List the word frequency of common words across files as a dict (tip: use join method in the rdd). The size of the dict equals the list of common words and each key in the top level dict contains the word count corresponding to all the ten files
 - {"word": {"pg1513.txt":10,"2160-0.txt":25,... } }

Answer: 02

```
In [16]: !pip install pyspark
```

```
Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-wheels/public/simple/
Collecting pyspark
  Downloading pyspark-3.4.0.tar.gz (310.8 MB)
    7 MB/s eta 0:00:00
  Preparing metadata (setup.py) ... done
Requirement already satisfied: py4j==0.10.9.7 in /usr/local/lib/python3.10/dist-packages (from pyspark) (0.10.9.7)
Building wheels for collected packages: pyspark
```



```
Building wheel for pyspark (setup.py) ... done
Created wheel for pyspark: filename=pyspark-3.4.0-py2.py3-none-any.whl size=311317130
sha256=c502492a3bfe97e14986b975f1d678029e4a9c969c4aab396077b213ca1ba59
Stored in directory: /root/.cache/pip/wheels/7b/1b/4b/3363a1d04368e7ff0d408e57ff57966f
cdf00583774e761327
Successfully built pyspark
Installing collected packages: pyspark
Successfully installed pyspark-3.4.0
```

Importing library

```
In [17]: import os
from pyspark import SparkConf, SparkContext
from operator import add
import pandas as pd
import re
```

```
In [18]: workingPath = '/content/sample_data/Assignment 1'
os.chdir(workingPath)
```

Common Variable Declaration

```
In [19]: # Initialize Spark
spark_config = SparkConf().setMaster('local').setAppName('CommonWordsFinderAppSparkConte
spark_context = SparkContext.getOrCreate(conf=spark_config)

# Define the list of files
file_paths = ['Q1-Dataset/pg1513.txt', 'Q1-Dataset/2160-0.txt', 'Q1-Dataset/pg11.txt',
              'Q1-Dataset/pg2600.txt', 'Q1-Dataset/pg345.txt', 'Q1-Dataset/pg64317.txt',
              'Q1-Dataset/pg70815.txt', 'Q1-Dataset/1080-0.txt', 'Q1-Dataset/pg70817.txt',
              'Q1-Dataset/pg1998.txt']

# Initialize the word frequency dictionary
word_frequency_dict = {}
```

Common Function Declation

```
In [20]: def clean_word(word):
          return re.sub(r'^\w\s', '', word).lower()
```

Perform the main operation of searching and formatting of data as requested

```
In [21]: # Calculate the word frequencies and word intersections for each file.
for file_index, file_path in enumerate(file_paths):
    word_rdd = spark_context.textFile(file_path).flatMap(lambda line: line.split(" ")).m
    word_frequency = word_rdd.countByValue()
    #Calculate intersections and frequencies
    word_frequency_dict = {word: {file_path: frequency} if word not in word_frequency_di
    intersecting_words_rdd = word_rdd if file_index == 0 else intersecting_words_rdd.int

#Collect the common words, removing the pure number, alpha numeric like 1e and 2f etc. a
common_words = [element for element in intersecting_words_rdd.distinct().collect() if no

# Filter the frequency dictionary to include only common words
common_words_frequency_dict = {word: word_frequency_dict[word] for word in common_words}

print(f"Common word count is {len(common_words)}\nCommon words are: {common_words}")
```

Common word count is 816

Common words are: ['rather', 'references', 'two', 'immediate', 'family', 'of', 'accessib
le', 'ago', 'away', 'official', 'right', 'before', 'ready', 'lieu', 'volunteers', 'perma
nent', 'than', 'computers', 'sent', 'united', 'law', 'damage', 'her', 'wwwgutenbergorgco

nact', 'gross', 'agent', 'would', 'putting', 'something', 'registered', 'anything', 'was', 'put', 'agreement', 'us', 'leave', 'think', 'particular', 'never', 'true', 'help', 'especially', 'format', 'files', 'links', 'language', 'allow', 'entity', 'have', 'expense', 'no', 'ein', 'four', 'die', 'charities', 'appears', 'thousand', 'equipment', 'demand', 'is', 'support', 'bring', 'though', 'opportunity', 'renamed', 'thus', 'till', 'making', 'derive', 'prohibition', 'solicit', 'manager', 'free', 'send', 'indicate', 'bear', 'viewing', 'strength', 'head', 'rules', 'phrase', 'obtain', 'share', 'displayed', 'synonymous', 'compliance', 'widest', 'indirect', 'looked', 'poor', 'countries', 'refund', 'freely', 'an', 'volunteer', 'said', 'charitable', 'at', 'tax', 'earth', 'requirements', 'archive', 'mission', 'following', 'imposed', 'almost', 'which', 'most', 'paperwork', 'bound', 'alone', 'copy', 'copies', 'and', 'glad', 'five', 'accordance', 'strict', 'goals', 'displaying', 'required', 'remain', 'gone', 'warranty', 'knew', 'available', 'any', 'follow', 'you', 'by', 'english', 'pg', 'reports', 'limitation', 'offer', 'disclaimer', 'editions', 'federal', 'preserve', 'harmless', 'received', 'ebooks', 'educational', 'within', 'care', 'concerning', 'parts', 'necessarily', 'processing', 'ut', 'indemnity', 'soon', 'taxes', 'distribute', 'compilation', 'web', 'email', 'how', 'return', 'wwwgutenbergorglicense', 'much', 'manner', 'treatment', 'codes', 'were', 'whatsoever', 'attempt', 'state', 'replace', 'several', 'both', 'checks', 'cease', 'govern', 'against', 'service', 'card', 'expend', 'abide', 'hart', 'use', 'awayyou', 'wwwgutenbergorgdonate', 'gratefully', 'claim', 'version', 'myself', 'created', 'internal', 'days', 'out', 'twenty', 'appearing', 'violates', 'financial', 'new', 'form', 'incomplete', 'cannot', 'derived', 'heart', 'donate', 'these', 'longer', 'ten', 'month', 'way', 'medium', 'nor', 'title', 'things', 'in', 'provision', 'long', 'inaccurate', 'additions', 'returns', 'proprietary', 'we', 'print', 'negligence', 'walk', 'change', 'array', 'holder', 'production', 'consequential', 'whose', 'carry', 'meet', 'actual', 'subscribe', 'include', 'pglaf', 'extent', 'visit', 'far', 'produce', 'this', 'rest', 'already', 'project', 'prominently', 'based', 'compressed', 'corrupt', 'search', 'modified', 'punitive', 'states', 'also', 'north', 'updated', 'it', 'forth', 'periodic', 'posted', 'agreed', 'not', 'contain', 'reading', 'trademarkcopyright', 'fell', 'sentence', 'data', 'fast', 'brought', 'pages', 'its', 'without', 'few', 'hypertext', 'destroy', 'every', 'my', 'profits', 'since', 'concept', 'except', 'defective', 'merchantability', 'maximum', 'physical', 'body', 'under', 'methods', 'wish', 'immediately', 'so', 'constant', 'outside', 'applicable', 'possessed', 'or', 'identification', 'onethe', 'previous', 'exporting', 'city', 'shall', 'small', 'found', 'whom', 'wife', 'general', 'facility', 'employees', 'easy', 'discontinue', 'variety', 'about', 'be', 'first', 'friend', 'original', 'seem', 'accept', 'main', 'apply', 'get', 'contact', 'alternate', 'section', 'notice', 'intellectual', 'direct', 's', 'common', 'there', 'turn', 'even', 'anywhere', 'told', 'perhaps', 'further', 'identify', 'produced', 'charge', 'sharing', 'modification', 'ebook', 'confirmed', 'stored', 'redistribution', 'copied', 'own', 'online', 'accessed', 'more', 'addresses', 'deletions', 'fine', 'ask', 'redistribute', 'implied', 'mean', 'very', 'charges', 'i', 'just', 'regulating', 'last', 'salt', 'alteration', 'past', 'fight', 'after', 'reuse', 'trouble', 'confirmation', 'appear', 'may', 'comply', 'depends', 'measure', 'transcribe', 'came', 'pay', 'are', 'offers', 'person', 'removed', 'make', 'easily', 'legal', 'good', 'provide', 'now', 'b', 'formats', 'houses', 'paid', 'world', 'neither', 'check', 'set', 'matter', 'remaining', 'public', 'different', 'disclaimers', 'forty', 'paragraph', 'providing', 'upon', 'discover', 'directly', 'require', 'receiving', 'granted', 'obtaining', 'distributing', 'license', 'donations', 'number', 'off', 'organized', 'office', 'one', 'release', 'learn', 'calculated', 'creation', 'their', 'written', 'warranties', 'foundation', 'place', 'second', 'hundred', 'on', 'located', 'generations', 'purpose', 'elect', 'performing', 'word', 'all', 'secure', 'whenever', 'has', 'does', 'agree', 'table', 'if', 'makes', 'hold', 'met', 'distributed', 'nothing', 'express', 'takes', 'come', 'lot', 'fitness', 'cost', 'protect', 'newsletter', 'status', 'invalidity', 'less', 'bent', 'answer', 'nearly', 'among', 'choose', 'because', 'west', 'disclaim', 'while', 'literary', 'virus', 'to', 'credit', 'grow', 'effort', 'problem', 'want', 'readable', 'such', 'viewed', 'people', 'should', 'lake', 'alive', 'plain', 'see', 'me', 'associated', 'reaching', 'staff', 'when', 'proofread', 'opportunities', 'revenue', 'owner', 'complying', 'copyright', 'sure', 'twelve', 'yet', 'michael', 'errors', 'creating', 'copying', 'figure', 'ascii', 'liability', 'work', 'he', 'library', 'calculate', 'beat', 'asis', 'home', 'active', 'page', 'c', 'above', 'disk', 'able', 'occur', 'gutenberg', 'let', 'start', 'distribution', 'particularly', 'donors', 'linked', 'three', 'certain', 'owns', 'transcription', 'maintaining', 'approach', 'group', 'accepted', 'kind', 'however', 'obsolete', 'subject', 'again', 'writing', 'stand', 'loose', 'accepting', 'cause', 'but', 'money', 'anyone', 'contract', 'house', 'am', 'take', 'exists', 'commercial', 'do', 'years', 'expenses', 'read', 'restrictions', 'like', 'edition', 'limited', 'promoting', 'prepare', 'child', 'young', 'using', 'unenforceability', 'individual', 'prevent', 'collection', 'other', 'vanilla', 'efforts', 'despite', 'others', 'country', 'move', 'uniform', 'another', 'children', 'for', 'distributor',

'up', 'permission', 'file', 'they', 'donation', 'property', 'licensed', 'includes', 'you
r', 'indicating', 'method', 'laws', 'present', 'future', 'age', 'girl', 'special', 'reas
on', 'each', 'types', 'specified', 'will', 'provisions', 'committed', 'ensuring', 'coul
d', 'swamp', 'unless', 'might', 'that', 'life', 'being', 'critical', 'been', 'little',
'payments', 'increasing', 'provided', 'wwwgutenbergorg', 'times', 'clearly', 'determin
e', 'considerable', 'beginning', 'death', 'receipt', 'please', 'international', 'unsolic
ited', 'half', 'made', 'irs', 'the', 'legally', 'user', 'same', 'survive', 'hope', 'perf
ormances', 'foundations', 'date', 'sending', 'employee', 'often', 'practically', 'explan
ation', 'replacement', 'author', 'taken', 'whole', 'some', 'electronic', 'then', 'grea
t', 'sort', 'redistributing', 'included', 'royalty', 'deductible', 'order', 'time', 'lia
ble', 'him', 'address', 'find', 'ever', 'full', 'sound', 'middleaged', 'must', 'addition
al', 'know', 'possibility', 'promotion', 'late', 'hear', 'below', 'known', 'performed',
'mississippi', 'his', 'believe', 'current', 'research', 'assistance', 'indemnify', 'doma
in', 'walks', 'exempt', 'means', 'as', 'possession', 'locations', 'into', 'where', 'fe
e', 'representations', 'hot', 'remedies', 'them', 'effect', 'reported', 'give', 'indirec
tly', 'business', 'downloading', 'used', 'incidental', 'nonproprietary', 'binary', 'aris
e', 'remove', 'cold', 'convert', 'end', 'only', 'best', 'electronically', 'perform', 'na
me', 'permitted', 'breach', 'interpreted', 'fees', 'understand', 'fix', 'network', 'wors
e', 'paying', 'sections', 'stay', 'old', 'outdated', 'costs', 'trademark', 'what', 'wel
l', 'part', 'from', 'shared', 'void', 'with', 'arms', 'marked', 'unlink', 'many', 'too',
'man', 'access', 'request', 'exclusion', 'discovered', 'display', 'defect', 'course', 'o
ur', 'either', 'reasonable', 'six', 'defects', 'originator', 'addition', 'those', 'impor
tant', 'including', 'owed', 'grown', 'infringement', 'keeping', 'damages', 'thing', 'hun
dreds', 'corporation', 'least', 'detach', 'having', 'information', 'can', 'terms', 'pape
r', 'a', 'printed', 'fair', 'better', 'solicitation', 'doubt', 'computer', 'royalties',
'contributions', 'derivative', 'ways', 'she', 'statements', 'keep', 'who', 'notifies',
'day', 'giving', 'works', 'containing', 'attached', 'need', 'paragraphs', 'through', 'de
scribed', 'professor', 'damaged', 'talk', 'receive', 'party', 'given']

Representation of the requested file wise common words frequency

```
In [22]: print(f"Frequencies of the words are :{common_words_frequency_dict}")
```

Frequencies of the words are :{'rather': {'Q1-Dataset/pg1513.txt': 4, 'Q1-Dataset/2160-0.txt': 43, 'Q1-Dataset/pg11.txt': 25, 'Q1-Dataset/pg2600.txt': 88, 'Q1-Dataset/pg345.txt': 25, 'Q1-Dataset/pg64317.txt': 25, 'Q1-Dataset/pg70815.txt': 15, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 10, 'Q1-Dataset/pg1998.txt': 40}, 'references': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 2, 'Q1-Dataset/pg345.txt': 2, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 2}, 'two': {'Q1-Dataset/pg1513.txt': 20, 'Q1-Dataset/2160-0.txt': 144, 'Q1-Dataset/pg11.txt': 37, 'Q1-Dataset/pg2600.txt': 587, 'Q1-Dataset/pg345.txt': 128, 'Q1-Dataset/pg64317.txt': 72, 'Q1-Dataset/pg70815.txt': 118, 'Q1-Dataset/1080-0.txt': 8, 'Q1-Dataset/pg70817.txt': 26, 'Q1-Dataset/pg1998.txt': 47}, 'immediate': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 12, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.txt': 10, 'Q1-Dataset/pg345.txt': 6, 'Q1-Dataset/pg64317.txt': 5, 'Q1-Dataset/pg70815.txt': 7, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 6, 'Q1-Dataset/pg1998.txt': 3}, 'family': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 113, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 135, 'Q1-Dataset/pg345.txt': 3, 'Q1-Dataset/pg64317.txt': 13, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 3, 'Q1-Dataset/pg1998.txt': 1}, 'of': {'Q1-Dataset/pg1513.txt': 518, 'Q1-Dataset/2160-0.txt': 5676, 'Q1-Dataset/pg11.txt': 632, 'Q1-Dataset/pg2600.txt': 14994, 'Q1-Dataset/pg345.txt': 3746, 'Q1-Dataset/pg64317.txt': 1233, 'Q1-Dataset/pg70815.txt': 3108, 'Q1-Dataset/1080-0.txt': 261, 'Q1-Dataset/pg70817.txt': 769, 'Q1-Dataset/pg1998.txt': 2895}, 'accessible': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 4, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'ago': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 45, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 74, 'Q1-Dataset/pg345.txt': 26, 'Q1-Dataset/pg64317.txt': 19, 'Q1-Dataset/pg70815.txt': 6, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 8, 'Q1-Dataset/pg1998.txt': 6}, 'away': {'Q1-Dataset/pg1513.txt': 25, 'Q1-Dataset/2160-0.txt': 51, 'Q1-Dataset/pg11.txt': 27, 'Q1-Dataset/pg2600.txt': 613, 'Q1-Dataset/pg345.txt': 187, 'Q1-Dataset/pg64317.txt': 76, 'Q1-Dataset/pg70815.txt': 71, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 35, 'Q1-Dataset/pg1998.txt': 116}, 'official': {'Q1-Dataset/pg1513.txt': 3, 'Q1-Dataset/2160-0.txt': 3, 'Q1-

'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.txt': 33, 'Q1-Dataset/pg345.txt': 6, 'Q1-Dataset/pg64317.txt': 4, 'Q1-Dataset/pg70815.txt': 4, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 3, 'Q1-Dataset/pg1998.txt': 3}, 'right': {'Q1-Dataset/pg1513.txt': 9, 'Q1-Dataset/2160-0.txt': 45, 'Q1-Dataset/pg11.txt': 35, 'Q1-Dataset/pg2600.txt': 449, 'Q1-Dataset/pg345.txt': 102, 'Q1-Dataset/pg64317.txt': 54, 'Q1-Dataset/pg70815.txt': 49, 'Q1-Dataset/1080-0.txt': 4, 'Q1-Dataset/pg70817.txt': 13, 'Q1-Dataset/pg1998.txt': 56}, 'before': {'Q1-Dataset/pg1513.txt': 24, 'Q1-Dataset/2160-0.txt': 118, 'Q1-Dataset/pg11.txt': 41, 'Q1-Dataset/pg2600.txt': 825, 'Q1-Dataset/pg345.txt': 278, 'Q1-Dataset/pg64317.txt': 91, 'Q1-Dataset/pg70815.txt': 119, 'Q1-Dataset/1080-0.txt': 7, 'Q1-Dataset/pg70817.txt': 61, 'Q1-Dataset/pg1998.txt': 113}, 'ready': {'Q1-Dataset/pg1513.txt': 5, 'Q1-Dataset/2160-0.txt': 25, 'Q1-Dataset/pg11.txt': 8, 'Q1-Dataset/pg2600.txt': 164, 'Q1-Dataset/pg345.txt': 71, 'Q1-Dataset/pg64317.txt': 3, 'Q1-Dataset/pg70815.txt': 20, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 4, 'Q1-Dataset/pg1998.txt': 26}, 'lieu': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 3, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 2, 'Q1-Dataset/pg345.txt': 2, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 3, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 2}, 'volunteers': {'Q1-Dataset/pg1513.txt': 6, 'Q1-Dataset/2160-0.txt': 6, 'Q1-Dataset/pg11.txt': 5, 'Q1-Dataset/pg2600.txt': 5, 'Q1-Dataset/pg345.txt': 5, 'Q1-Dataset/pg64317.txt': 5, 'Q1-Dataset/pg70815.txt': 5, 'Q1-Dataset/1080-0.txt': 5, 'Q1-Dataset/pg70817.txt': 5, 'Q1-Dataset/pg1998.txt': 5}, 'permanent': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 2, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 4, 'Q1-Dataset/pg1998.txt': 2}, 'than': {'Q1-Dataset/pg1513.txt': 50, 'Q1-Dataset/2160-0.txt': 280, 'Q1-Dataset/pg11.txt': 27, 'Q1-Dataset/pg2600.txt': 535, 'Q1-Dataset/pg345.txt': 222, 'Q1-Dataset/pg64317.txt': 37, 'Q1-Dataset/pg70815.txt': 151, 'Q1-Dataset/1080-0.txt': 10, 'Q1-Dataset/pg70817.txt': 37, 'Q1-Dataset/pg1998.txt': 205}, 'computers': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 2, 'Q1-Dataset/pg345.txt': 2, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 2}, 'sent': {'Q1-Dataset/pg1513.txt': 3, 'Q1-Dataset/2160-0.txt': 49, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.txt': 224, 'Q1-Dataset/pg345.txt': 39, 'Q1-Dataset/pg64317.txt': 17, 'Q1-Dataset/pg70815.txt': 14, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 4, 'Q1-Dataset/pg1998.txt': 4}, 'united': {'Q1-Dataset/pg1513.txt': 15, 'Q1-Dataset/2160-0.txt': 15, 'Q1-Dataset/pg11.txt': 15, 'Q1-Dataset/pg2600.txt': 31, 'Q1-Dataset/pg345.txt': 18, 'Q1-Dataset/pg64317.txt': 15, 'Q1-Dataset/pg70815.txt': 16, 'Q1-Dataset/1080-0.txt': 15, 'Q1-Dataset/pg70817.txt': 15, 'Q1-Dataset/pg1998.txt': 15}, 'law': {'Q1-Dataset/pg1513.txt': 18, 'Q1-Dataset/2160-0.txt': 23, 'Q1-Dataset/pg11.txt': 9, 'Q1-Dataset/pg2600.txt': 79, 'Q1-Dataset/pg345.txt': 20, 'Q1-Dataset/pg64317.txt': 7, 'Q1-Dataset/pg70815.txt': 9, 'Q1-Dataset/1080-0.txt': 8, 'Q1-Dataset/pg70817.txt': 7, 'Q1-Dataset/pg1998.txt': 29}, 'damage': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 10, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 3, 'Q1-Dataset/pg345.txt': 2, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 3, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 2}, 'her': {'Q1-Dataset/pg1513.txt': 153, 'Q1-Dataset/2160-0.txt': 750, 'Q1-Dataset/pg11.txt': 247, 'Q1-Dataset/pg2600.txt': 4697, 'Q1-Dataset/pg345.txt': 1054, 'Q1-Dataset/pg64317.txt': 392, 'Q1-Dataset/pg70815.txt': 388, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 156, 'Q1-Dataset/pg1998.txt': 36}, 'wwwgutenbergorgcontact': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'gross': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 7, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 3, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'agent': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 2, 'Q1-Dataset/pg345.txt': 13, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'would': {'Q1-Dataset/pg1513.txt': 54, 'Q1-Dataset/2160-0.txt': 445, 'Q1-Dataset/pg11.txt': 78, 'Q1-Dataset/pg2600.txt': 1362, 'Q1-Dataset/pg345.txt': 428, 'Q1-Dataset/pg64317.txt': 52, 'Q1-Dataset/pg70815.txt': 219, 'Q1-Dataset/1080-0.txt': 18, 'Q1-Dataset/pg70817.txt': 88, 'Q1-Dataset/pg1998.txt': 259}, 'putting': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 4, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.txt': 51, 'Q1-Dataset/pg345.txt': 21, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 3}, 'something': {'Q1-Dataset/pg1513.txt': 4, 'Q1-Dataset/2160-0.txt': 51, 'Q1-Dataset/pg11.txt': 17, 'Q1-Dataset/pg2600.txt': 597, 'Q1-Dataset/pg345.txt': 137, 'Q1-Dataset/pg64317.txt': 66, 'Q1-Dataset/pg7081

5.txt': 44, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 17, 'Q1-Dataset/pg1998.txt': 65}, 'registered': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 2, 'Q1-Dataset/pg345.txt': 2, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 2}, 'anything': {'Q1-Dataset/pg1513.txt': 6, 'Q1-Dataset/2160-0.txt': 5, 'Q1-Dataset/pg11.txt': 21, 'Q1-Dataset/pg2600.txt': 293, 'Q1-Dataset/pg345.txt': 99, 'Q1-Dataset/pg64317.txt': 29, 'Q1-Dataset/pg70815.txt': 20, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 9, 'Q1-Dataset/pg1998.txt': 18}, 'was': {'Q1-Dataset/pg1513.txt': 51, 'Q1-Dataset/2160-0.txt': 1399, 'Q1-Dataset/pg11.txt': 354, 'Q1-Dataset/pg2600.txt': 7332, 'Q1-Dataset/pg345.txt': 1874, 'Q1-Dataset/pg64317.txt': 766, 'Q1-Dataset/pg70815.txt': 1094, 'Q1-Dataset/1080-0.txt': 10, 'Q1-Dataset/pg70817.txt': 440, 'Q1-Dataset/pg1998.txt': 362}, 'put': {'Q1-Dataset/pg1513.txt': 17, 'Q1-Dataset/2160-0.txt': 86, 'Q1-Dataset/pg11.txt': 31, 'Q1-Dataset/pg2600.txt': 270, 'Q1-Dataset/pg345.txt': 106, 'Q1-Dataset/pg64317.txt': 21, 'Q1-Dataset/pg70815.txt': 10, 'Q1-Dataset/1080-0.txt': 4, 'Q1-Dataset/pg70817.txt': 6, 'Q1-Dataset/pg1998.txt': 59}, 'agreement': {'Q1-Dataset/pg1513.txt': 18, 'Q1-Dataset/2160-0.txt': 20, 'Q1-Dataset/pg11.txt': 18, 'Q1-Dataset/pg2600.txt': 25, 'Q1-Dataset/pg345.txt': 21, 'Q1-Dataset/pg64317.txt': 20, 'Q1-Dataset/pg70815.txt': 18, 'Q1-Dataset/1080-0.txt': 18, 'Q1-Dataset/pg70817.txt': 18, 'Q1-Dataset/pg1998.txt': 18}, 'us': {'Q1-Dataset/pg1513.txt': 38, 'Q1-Dataset/2160-0.txt': 152, 'Q1-Dataset/pg11.txt': 22, 'Q1-Dataset/pg2600.txt': 424, 'Q1-Dataset/pg345.txt': 463, 'Q1-Dataset/pg64317.txt': 71, 'Q1-Dataset/pg70815.txt': 123, 'Q1-Dataset/1080-0.txt': 12, 'Q1-Dataset/pg70817.txt': 36, 'Q1-Dataset/pg1998.txt': 156}, 'leave': {'Q1-Dataset/pg1513.txt': 13, 'Q1-Dataset/2160-0.txt': 59, 'Q1-Dataset/pg11.txt': 9, 'Q1-Dataset/pg2600.txt': 209, 'Q1-Dataset/pg345.txt': 56, 'Q1-Dataset/pg64317.txt': 12, 'Q1-Dataset/pg70815.txt': 10, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 11, 'Q1-Dataset/pg1998.txt': 17}, 'think': {'Q1-Dataset/pg1513.txt': 21, 'Q1-Dataset/2160-0.txt': 127, 'Q1-Dataset/pg11.txt': 49, 'Q1-Dataset/pg2600.txt': 355, 'Q1-Dataset/pg345.txt': 218, 'Q1-Dataset/pg64317.txt': 64, 'Q1-Dataset/pg70815.txt': 32, 'Q1-Dataset/1080-0.txt': 6, 'Q1-Dataset/pg70817.txt': 8, 'Q1-Dataset/pg1998.txt': 50}, 'particular': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 66, 'Q1-Dataset/pg11.txt': 5, 'Q1-Dataset/pg2600.txt': 44, 'Q1-Dataset/pg345.txt': 10, 'Q1-Dataset/pg64317.txt': 8, 'Q1-Dataset/pg70815.txt': 9, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 3, 'Q1-Dataset/pg1998.txt': 7}, 'never': {'Q1-Dataset/pg1513.txt': 18, 'Q1-Dataset/2160-0.txt': 139, 'Q1-Dataset/pg11.txt': 43, 'Q1-Dataset/pg2600.txt': 407, 'Q1-Dataset/pg345.txt': 128, 'Q1-Dataset/pg64317.txt': 64, 'Q1-Dataset/pg70815.txt': 94, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 29, 'Q1-Dataset/pg1998.txt': 62}, 'true': {'Q1-Dataset/pg1513.txt': 27, 'Q1-Dataset/2160-0.txt': 21, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.txt': 120, 'Q1-Dataset/pg345.txt': 56, 'Q1-Dataset/pg64317.txt': 9, 'Q1-Dataset/pg70815.txt': 14, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 5, 'Q1-Dataset/pg1998.txt': 31}, 'help': {'Q1-Dataset/pg1513.txt': 21, 'Q1-Dataset/2160-0.txt': 54, 'Q1-Dataset/pg11.txt': 12, 'Q1-Dataset/pg2600.txt': 144, 'Q1-Dataset/pg345.txt': 103, 'Q1-Dataset/pg64317.txt': 10, 'Q1-Dataset/pg70815.txt': 9, 'Q1-Dataset/1080-0.txt': 4, 'Q1-Dataset/pg70817.txt': 5, 'Q1-Dataset/pg1998.txt': 27}, 'especially': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 17, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 128, 'Q1-Dataset/pg345.txt': 10, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 7, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 4, 'Q1-Dataset/pg1998.txt': 32}, 'format': {'Q1-Dataset/pg1513.txt': 4, 'Q1-Dataset/2160-0.txt': 4, 'Q1-Dataset/pg11.txt': 4, 'Q1-Dataset/pg2600.txt': 4, 'Q1-Dataset/pg345.txt': 4, 'Q1-Dataset/pg64317.txt': 4, 'Q1-Dataset/pg70815.txt': 4, 'Q1-Dataset/1080-0.txt': 4, 'Q1-Dataset/pg70817.txt': 4, 'Q1-Dataset/pg1998.txt': 4}, 'files': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 2, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 2}, 'links': {'Q1-Dataset/pg1513.txt': 3, 'Q1-Dataset/2160-0.txt': 6, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.txt': 4, 'Q1-Dataset/pg345.txt': 4, 'Q1-Dataset/pg64317.txt': 3, 'Q1-Dataset/pg70815.txt': 3, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 4, 'Q1-Dataset/pg1998.txt': 3}, 'language': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 23, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 9, 'Q1-Dataset/pg345.txt': 9, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 31, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 18}, 'allow': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 11, 'Q1-Dataset/pg11.txt': 4, 'Q1-Dataset/pg2600.txt': 57, 'Q1-Dataset/pg345.txt': 7, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 7}, 'entity': {'Q1-Dataset/pg1513.txt': 3, 'Q1-Dataset/2160-0.txt': 3, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.txt': 3, 'Q1-Dataset/pg345.txt': 4, 'Q1-Dataset/pg64317.txt': 3, 'Q1-Dataset/pg70815.txt': 3, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 4, 'Q1-Dataset/pg1998.txt': 3}, 'have': {'Q1-Dataset/

t/pg1513.txt': 133, 'Q1-Dataset/2160-0.txt': 921, 'Q1-Dataset/pg11.txt': 83, 'Q1-Dataset/pg2600.txt': 1980, 'Q1-Dataset/pg345.txt': 1066, 'Q1-Dataset/pg64317.txt': 143, 'Q1-Dataset/pg70815.txt': 406, 'Q1-Dataset/1080-0.txt': 36, 'Q1-Dataset/pg70817.txt': 71, 'Q1-Dataset/pg1998.txt': 673}, 'expense': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 4, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 7, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 2}, 'no': {'Q1-Dataset/pg1513.txt': 111, 'Q1-Dataset/2160-0.txt': 469, 'Q1-Dataset/pg11.txt': 96, 'Q1-Dataset/pg2600.txt': 1402, 'Q1-Dataset/pg345.txt': 486, 'Q1-Dataset/pg64317.txt': 111, 'Q1-Dataset/pg70815.txt': 273, 'Q1-Dataset/1080-0.txt': 22, 'Q1-Dataset/pg70817.txt': 78, 'Q1-Dataset/pg1998.txt': 282}, 'ein': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 2, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'four': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 43, 'Q1-Dataset/pg11.txt': 6, 'Q1-Dataset/pg2600.txt': 103, 'Q1-Dataset/pg345.txt': 24, 'Q1-Dataset/pg64317.txt': 26, 'Q1-Dataset/pg70815.txt': 37, 'Q1-Dataset/1080-0.txt': 5, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 3}, 'die': {'Q1-Dataset/pg1513.txt': 23, 'Q1-Dataset/2160-0.txt': 8, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 67, 'Q1-Dataset/pg345.txt': 43, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 16, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 3, 'Q1-Dataset/pg1998.txt': 32}, 'charities': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 3, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'appears': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 16, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 24, 'Q1-Dataset/pg345.txt': 2, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 4}, 'thousand': {'Q1-Dataset/pg1513.txt': 7, 'Q1-Dataset/2160-0.txt': 52, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 168, 'Q1-Dataset/pg345.txt': 10, 'Q1-Dataset/pg64317.txt': 6, 'Q1-Dataset/pg70815.txt': 10, 'Q1-Dataset/1080-0.txt': 15, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 23}, 'equipment': {'Q1-Dataset/pg1513.txt': 3, 'Q1-Dataset/2160-0.txt': 3, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.txt': 11, 'Q1-Dataset/pg345.txt': 5, 'Q1-Dataset/pg64317.txt': 3, 'Q1-Dataset/pg70815.txt': 4, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 3, 'Q1-Dataset/pg1998.txt': 3}, 'demand': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 8, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 31, 'Q1-Dataset/pg345.txt': 2, 'Q1-Dataset/pg64317.txt': 4, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'is': {'Q1-Dataset/pg1513.txt': 366, 'Q1-Dataset/2160-0.txt': 1319, 'Q1-Dataset/pg11.txt': 116, 'Q1-Dataset/pg2600.txt': 3321, 'Q1-Dataset/pg345.txt': 1521, 'Q1-Dataset/pg64317.txt': 126, 'Q1-Dataset/pg70815.txt': 404, 'Q1-Dataset/1080-0.txt': 43, 'Q1-Dataset/pg70817.txt': 59, 'Q1-Dataset/pg1998.txt': 1816}, 'support': {'Q1-Dataset/pg1513.txt': 4, 'Q1-Dataset/2160-0.txt': 16, 'Q1-Dataset/pg11.txt': 4, 'Q1-Dataset/pg2600.txt': 20, 'Q1-Dataset/pg345.txt': 5, 'Q1-Dataset/pg64317.txt': 6, 'Q1-Dataset/pg70815.txt': 10, 'Q1-Dataset/1080-0.txt': 5, 'Q1-Dataset/pg70817.txt': 4, 'Q1-Dataset/pg1998.txt': 6}, 'bring': {'Q1-Dataset/pg1513.txt': 11, 'Q1-Dataset/2160-0.txt': 25, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.txt': 83, 'Q1-Dataset/pg345.txt': 32, 'Q1-Dataset/pg64317.txt': 4, 'Q1-Dataset/pg70815.txt': 3, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 4, 'Q1-Dataset/pg1998.txt': 13}, 'though': {'Q1-Dataset/pg1513.txt': 11, 'Q1-Dataset/2160-0.txt': 130, 'Q1-Dataset/pg11.txt': 11, 'Q1-Dataset/pg2600.txt': 443, 'Q1-Dataset/pg345.txt': 216, 'Q1-Dataset/pg64317.txt': 34, 'Q1-Dataset/pg70815.txt': 75, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 5, 'Q1-Dataset/pg1998.txt': 29}, 'opportunity': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 33, 'Q1-Dataset/pg11.txt': 9, 'Q1-Dataset/pg2600.txt': 33, 'Q1-Dataset/pg345.txt': 23, 'Q1-Dataset/pg64317.txt': 4, 'Q1-Dataset/pg70815.txt': 13, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 4}, 'renamed': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'thus': {'Q1-Dataset/pg1513.txt': 10, 'Q1-Dataset/2160-0.txt': 26, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 27, 'Q1-Dataset/pg345.txt': 27, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 37, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 386}, 'till': {'Q1-Dataset/pg1513.txt': 22, 'Q1-Dataset/2160-0.txt': 85, 'Q1-Dataset/pg11.txt': 21, 'Q1-Dataset/pg2600.txt': 173, 'Q1-Dataset/pg345.txt': 133, 'Q1-Dataset/pg64317.txt': 7, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 5, 'Q1-Dataset/pg1998.txt': 4}, 'making': {'Q1-Dataset/pg1513.txt': 3, 'Q1-Dataset/2160-0.txt': 41, 'Q1-Dataset/pg11.txt': 8, 'Q1-Dataset/pg2600.txt': 107, 'Q1-Dataset/pg345.txt': 30, 'Q1-

[illegible]

et/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 2, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 2}, 'looked': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 37, 'Q1-Dataset/pg11.txt': 45, 'Q1-Dataset/pg2600.txt': 647, 'Q1-Dataset/pg345.txt': 184, 'Q1-Dataset/pg64317.txt': 82, 'Q1-Dataset/pg70815.txt': 57, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 33, 'Q1-Dataset/pg1998.txt': 37}, 'poor': {'Q1-Dataset/pg1513.txt': 18, 'Q1-Dataset/2160-0.txt': 161, 'Q1-Dataset/pg11.txt': 27, 'Q1-Dataset/pg2600.txt': 75, 'Q1-Dataset/pg345.txt': 191, 'Q1-Dataset/pg64317.txt': 8, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 7, 'Q1-Dataset/pg70817.txt': 5, 'Q1-Dataset/pg1998.txt': 27}, 'countries': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 6, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 2, 'Q1-Dataset/pg345.txt': 3, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 4, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 4}, 'refund': {'Q1-Dataset/pg1513.txt': 10, 'Q1-Dataset/2160-0.txt': 10, 'Q1-Dataset/pg11.txt': 10, 'Q1-Dataset/pg2600.txt': 10, 'Q1-Dataset/pg345.txt': 10, 'Q1-Dataset/pg64317.txt': 10, 'Q1-Dataset/pg70815.txt': 10, 'Q1-Dataset/1080-0.txt': 10, 'Q1-Dataset/pg70817.txt': 10, 'Q1-Dataset/pg1998.txt': 10}, 'freely': {'Q1-Dataset/pg1513.txt': 4, 'Q1-Dataset/2160-0.txt': 15, 'Q1-Dataset/pg11.txt': 4, 'Q1-Dataset/pg2600.txt': 16, 'Q1-Dataset/pg345.txt': 17, 'Q1-Dataset/pg64317.txt': 5, 'Q1-Dataset/pg70815.txt': 5, 'Q1-Dataset/1080-0.txt': 5, 'Q1-Dataset/pg70817.txt': 5, 'Q1-Dataset/pg1998.txt': 5}, 'an': {'Q1-Dataset/pg1513.txt': 76, 'Q1-Dataset/2160-0.txt': 556, 'Q1-Dataset/pg11.txt': 62, 'Q1-Dataset/pg2600.txt': 1623, 'Q1-Dataset/pg345.txt': 360, 'Q1-Dataset/pg64317.txt': 191, 'Q1-Dataset/pg70815.txt': 248, 'Q1-Dataset/1080-0.txt': 14, 'Q1-Dataset/pg70817.txt': 104, 'Q1-Dataset/pg1998.txt': 243}, 'volunteer': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 3, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'said': {'Q1-Dataset/pg1513.txt': 14, 'Q1-Dataset/2160-0.txt': 278, 'Q1-Dataset/pg11.txt': 460, 'Q1-Dataset/pg2600.txt': 2834, 'Q1-Dataset/pg345.txt': 569, 'Q1-Dataset/pg64317.txt': 233, 'Q1-Dataset/pg70815.txt': 386, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 37, 'Q1-Dataset/pg1998.txt': 191}, 'charitable': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 5, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 6, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 1}, 'at': {'Q1-Dataset/pg1513.txt': 83, 'Q1-Dataset/2160-0.txt': 1090, 'Q1-Dataset/pg11.txt': 221, 'Q1-Dataset/pg2600.txt': 4533, 'Q1-Dataset/pg345.txt': 1091, 'Q1-Dataset/pg64317.txt': 417, 'Q1-Dataset/pg70815.txt': 408, 'Q1-Dataset/1080-0.txt': 42, 'Q1-Dataset/pg70817.txt': 211, 'Q1-Dataset/pg1998.txt': 506}, 'tax': {'Q1-Dataset/pg1513.txt': 6, 'Q1-Dataset/2160-0.txt': 8, 'Q1-Dataset/pg11.txt': 6, 'Q1-Dataset/pg2600.txt': 6, 'Q1-Dataset/pg345.txt': 9, 'Q1-Dataset/pg64317.txt': 6, 'Q1-Dataset/pg70815.txt': 6, 'Q1-Dataset/1080-0.txt': 6, 'Q1-Dataset/pg70817.txt': 6, 'Q1-Dataset/pg1998.txt': 6}, 'earth': {'Q1-Dataset/pg1513.txt': 19, 'Q1-Dataset/2160-0.txt': 25, 'Q1-Dataset/pg11.txt': 4, 'Q1-Dataset/pg2600.txt': 57, 'Q1-Dataset/pg345.txt': 36, 'Q1-Dataset/pg64317.txt': 3, 'Q1-Dataset/pg70815.txt': 26, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 7, 'Q1-Dataset/pg1998.txt': 117}, 'requirements': {'Q1-Dataset/pg1513.txt': 4, 'Q1-Dataset/2160-0.txt': 4, 'Q1-Dataset/pg11.txt': 4, 'Q1-Dataset/pg2600.txt': 6, 'Q1-Dataset/pg345.txt': 6, 'Q1-Dataset/pg64317.txt': 4, 'Q1-Dataset/pg70815.txt': 4, 'Q1-Dataset/1080-0.txt': 4, 'Q1-Dataset/pg70817.txt': 4, 'Q1-Dataset/pg1998.txt': 4}, 'archive': {'Q1-Dataset/pg1513.txt': 13, 'Q1-Dataset/2160-0.txt': 13, 'Q1-Dataset/pg11.txt': 13, 'Q1-Dataset/pg2600.txt': 13, 'Q1-Dataset/pg345.txt': 13, 'Q1-Dataset/pg64317.txt': 13, 'Q1-Dataset/pg70815.txt': 13, 'Q1-Dataset/1080-0.txt': 13, 'Q1-Dataset/pg70817.txt': 13, 'Q1-Dataset/pg1998.txt': 13}, 'mission': {'Q1-Dataset/pg1513.txt': 4, 'Q1-Dataset/2160-0.txt': 6, 'Q1-Dataset/pg11.txt': 4, 'Q1-Dataset/pg2600.txt': 15, 'Q1-Dataset/pg345.txt': 6, 'Q1-Dataset/pg64317.txt': 4, 'Q1-Dataset/pg70815.txt': 9, 'Q1-Dataset/1080-0.txt': 4, 'Q1-Dataset/pg70817.txt': 4, 'Q1-Dataset/pg1998.txt': 4}, 'following': {'Q1-Dataset/pg1513.txt': 5, 'Q1-Dataset/2160-0.txt': 21, 'Q1-Dataset/pg11.txt': 4, 'Q1-Dataset/pg2600.txt': 74, 'Q1-Dataset/pg345.txt': 17, 'Q1-Dataset/pg64317.txt': 9, 'Q1-Dataset/pg70815.txt': 25, 'Q1-Dataset/1080-0.txt': 4, 'Q1-Dataset/pg70817.txt': 6, 'Q1-Dataset/pg1998.txt': 11}, 'imposed': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 4, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 4}, 'almost': {'Q1-Dataset/pg1513.txt': 5, 'Q1-Dataset/2160-0.txt': 31, 'Q1-Dataset/pg11.txt': 8, 'Q1-Dataset/pg2600.txt': 154, 'Q1-Dataset/pg345.txt': 70, 'Q1-Dataset/pg64317.txt': 22, 'Q1-Dataset/pg70815.txt': 92, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 14, 'Q1-Dataset/pg1998.txt': 29}, 'which': {'Q1-Dataset/pg1513.txt': 69, 'Q1-Dataset/2160-0.txt': 1024, 'Q1-Dataset/pg11.txt': 56, 'Q1-Dataset/pg2600.txt': 2017, 'Q1-Dataset/pg345.txt': 656, 'Q1-Dataset/pg6

4317.txt': 72, 'Q1-Dataset/pg70815.txt': 400, 'Q1-Dataset/1080-0.txt': 30, 'Q1-Dataset/pg70817.txt': 132, 'Q1-Dataset/pg1998.txt': 468}, 'most': {'Q1-Dataset/pg1513.txt': 22, 'Q1-Dataset/2160-0.txt': 177, 'Q1-Dataset/pg11.txt': 14, 'Q1-Dataset/pg2600.txt': 328, 'Q1-Dataset/pg345.txt': 80, 'Q1-Dataset/pg64317.txt': 36, 'Q1-Dataset/pg70815.txt': 37, 'Q1-Dataset/1080-0.txt': 14, 'Q1-Dataset/pg70817.txt': 17, 'Q1-Dataset/pg1998.txt': 14 9}, 'paperwork': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'bound': {'Q1-Dataset/pg1513.txt': 10, 'Q1-Dataset/2160-0.txt': 14, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.txt': 44, 'Q1-Dataset/pg345.txt': 12, 'Q1-Dataset/pg64317.txt': 6, 'Q1-Dataset/pg70815.txt': 11, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 6, 'Q1-Dataset/pg1998.txt': 5}, 'alone': {'Q1-Dataset/pg1513.txt': 16, 'Q1-Dataset/2160-0.txt': 10, 'Q1-Dataset/pg11.txt': 5, 'Q1-Dataset/pg2600.txt': 237, 'Q1-Dataset/pg345.txt': 81, 'Q1-Dataset/pg64317.txt': 32, 'Q1-Dataset/pg70815.txt': 37, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 6, 'Q1-Dataset/pg1998.txt': 61}, 'copy': {'Q1-Dataset/pg1513.txt': 12, 'Q1-Dataset/2160-0.txt': 13, 'Q1-Dataset/pg11.txt': 12, 'Q1-Dataset/pg2600.txt': 16, 'Q1-Dataset/pg345.txt': 21, 'Q1-Dataset/pg64317.txt': 18, 'Q1-Dataset/pg70815.txt': 13, 'Q1-Dataset/1080-0.txt': 12, 'Q1-Dataset/pg70817.txt': 12, 'Q1-Dataset/pg1998.txt': 14}, 'copies': {'Q1-Dataset/pg1513.txt': 7, 'Q1-Dataset/2160-0.txt': 7, 'Q1-Dataset/pg11.txt': 7, 'Q1-Dataset/pg2600.txt': 9, 'Q1-Dataset/pg345.txt': 9, 'Q1-Dataset/pg64317.txt': 8, 'Q1-Dataset/pg70815.txt': 8, 'Q1-Dataset/1080-0.txt': 7, 'Q1-Dataset/pg70817.txt': 7, 'Q1-Dataset/pg1998.txt': 10}, 'and': {'Q1-Dataset/pg1513.txt': 803, 'Q1-Dataset/2160-0.txt': 5053, 'Q1-Dataset/pg11.txt': 913, 'Q1-Dataset/pg2600.txt': 22147, 'Q1-Dataset/pg345.txt': 5894, 'Q1-Dataset/pg64317.txt': 1594, 'Q1-Dataset/pg70815.txt': 2541, 'Q1-Dataset/1080-0.txt': 184, 'Q1-Dataset/pg70817.txt': 782, 'Q1-Dataset/pg1998.txt': 4590}, 'glad': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 23, 'Q1-Dataset/pg11.txt': 11, 'Q1-Dataset/pg2600.txt': 129, 'Q1-Dataset/pg345.txt': 35, 'Q1-Dataset/pg64317.txt': 11, 'Q1-Dataset/pg70815.txt': 9, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 3, 'Q1-Dataset/pg1998.txt': 3}, 'five': {'Q1-Dataset/pg1513.txt': 6, 'Q1-Dataset/2160-0.txt': 39, 'Q1-Dataset/pg11.txt': 8, 'Q1-Dataset/pg2600.txt': 85, 'Q1-Dataset/pg345.txt': 19, 'Q1-Dataset/pg64317.txt': 32, 'Q1-Dataset/pg70815.txt': 11, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 5, 'Q1-Dataset/pg1998.txt': 3}, 'accordance': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 7, 'Q1-Dataset/pg345.txt': 5, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 5, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 3}, 'strict': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 8, 'Q1-Dataset/pg345.txt': 5, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 2}, 'goal': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 5}, 'displaying': {'Q1-Dataset/pg1513.txt': 4, 'Q1-Dataset/2160-0.txt': 8, 'Q1-Dataset/pg11.txt': 4, 'Q1-Dataset/pg2600.txt': 9, 'Q1-Dataset/pg345.txt': 4, 'Q1-Dataset/pg64317.txt': 4, 'Q1-Dataset/pg70815.txt': 4, 'Q1-Dataset/1080-0.txt': 4, 'Q1-Dataset/pg70817.txt': 5, 'Q1-Dataset/pg1998.txt': 4}, 'required': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 9, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 17, 'Q1-Dataset/pg345.txt': 10, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 3, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 4, 'Q1-Dataset/pg1998.txt': 6}, 'remain': {'Q1-Dataset/pg1513.txt': 3, 'Q1-Dataset/2160-0.txt': 5, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 77, 'Q1-Dataset/pg345.txt': 23, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 23, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 4, 'Q1-Dataset/pg1998.txt': 13}, 'gone': {'Q1-Dataset/pg1513.txt': 33, 'Q1-Dataset/2160-0.txt': 30, 'Q1-Dataset/pg11.txt': 13, 'Q1-Dataset/pg2600.txt': 161, 'Q1-Dataset/pg345.txt': 72, 'Q1-Dataset/pg64317.txt': 23, 'Q1-Dataset/pg70815.txt': 36, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 18, 'Q1-Dataset/pg1998.txt': 33}, 'warranty': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 2, 'Q1-Dataset/pg345.txt': 2, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 2}, 'knew': {'Q1-Dataset/pg1513.txt': 3, 'Q1-Dataset/2160-0.txt': 35, 'Q1-Dataset/pg11.txt': 15, 'Q1-Dataset/pg2600.txt': 409, 'Q1-Dataset/pg345.txt': 126, 'Q1-Dataset/pg64317.txt': 54, 'Q1-Dataset/pg70815.txt': 89, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 21, 'Q1-Dataset/pg1998.txt': 16}, 'available': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 7, 'Q1-Dataset/pg345.txt': 8, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 2},

'any': {'Q1-Dataset/pg1513.txt': 53, 'Q1-Dataset/2160-0.txt': 336, 'Q1-Dataset/pg11.txt': 76, 'Q1-Dataset/pg2600.txt': 502, 'Q1-Dataset/pg345.txt': 322, 'Q1-Dataset/pg64317.txt': 104, 'Q1-Dataset/pg70815.txt': 162, 'Q1-Dataset/1080-0.txt': 45, 'Q1-Dataset/pg70817.txt': 67, 'Q1-Dataset/pg1998.txt': 113}, 'follow': {'Q1-Dataset/pg1513.txt': 8, 'Q1-Dataset/2160-0.txt': 15, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.txt': 40, 'Q1-Dataset/pg345.txt': 37, 'Q1-Dataset/pg64317.txt': 5, 'Q1-Dataset/pg70815.txt': 30, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 14}, 'you': {'Q1-Dataset/pg1513.txt': 362, 'Q1-Dataset/2160-0.txt': 742, 'Q1-Dataset/pg11.txt': 421, 'Q1-Dataset/pg2600.txt': 3647, 'Q1-Dataset/pg345.txt': 1453, 'Q1-Dataset/pg64317.txt': 528, 'Q1-Dataset/pg70815.txt': 569, 'Q1-Dataset/1080-0.txt': 74, 'Q1-Dataset/pg70817.txt': 134, 'Q1-Dataset/pg1998.txt': 364}, 'by': {'Q1-Dataset/pg1513.txt': 142, 'Q1-Dataset/2160-0.txt': 954, 'Q1-Dataset/pg11.txt': 84, 'Q1-Dataset/pg2600.txt': 2407, 'Q1-Dataset/pg345.txt': 530, 'Q1-Dataset/pg64317.txt': 150, 'Q1-Dataset/pg70815.txt': 382, 'Q1-Dataset/1080-0.txt': 61, 'Q1-Dataset/pg70817.txt': 105, 'Q1-Dataset/pg1998.txt': 313}, 'english': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 53, 'Q1-Dataset/pg11.txt': 7, 'Q1-Dataset/pg2600.txt': 30, 'Q1-Dataset/pg345.txt': 18, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 3, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 4}, 'pg': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 9, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 2}, 'reports': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 3, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 47, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 3, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 1}, 'limitation': {'Q1-Dataset/pg1513.txt': 3, 'Q1-Dataset/2160-0.txt': 3, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.txt': 4, 'Q1-Dataset/pg345.txt': 4, 'Q1-Dataset/pg64317.txt': 3, 'Q1-Dataset/pg70815.txt': 3, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 3, 'Q1-Dataset/pg1998.txt': 3}, 'offer': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 16, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 36, 'Q1-Dataset/pg345.txt': 2, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 4, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 4}, 'disclaimer': {'Q1-Dataset/pg1513.txt': 3, 'Q1-Dataset/2160-0.txt': 3, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.txt': 3, 'Q1-Dataset/pg345.txt': 3, 'Q1-Dataset/pg64317.txt': 3, 'Q1-Dataset/pg70815.txt': 3, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 3, 'Q1-Dataset/pg1998.txt': 3}, 'editions': {'Q1-Dataset/pg1513.txt': 4, 'Q1-Dataset/2160-0.txt': 4, 'Q1-Dataset/pg11.txt': 6, 'Q1-Dataset/pg2600.txt': 4, 'Q1-Dataset/pg345.txt': 4, 'Q1-Dataset/pg64317.txt': 4, 'Q1-Dataset/pg70815.txt': 4, 'Q1-Dataset/1080-0.txt': 4, 'Q1-Dataset/pg70817.txt': 4, 'Q1-Dataset/pg1998.txt': 4}, 'federal': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 2, 'Q1-Dataset/pg345.txt': 2, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 2}, 'preserve': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 5, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 12, 'Q1-Dataset/pg345.txt': 3, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 7}, 'harmless': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 4, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 3, 'Q1-Dataset/pg1998.txt': 1}, 'received': {'Q1-Dataset/pg1513.txt': 5, 'Q1-Dataset/2160-0.txt': 83, 'Q1-Dataset/pg11.txt': 6, 'Q1-Dataset/pg2600.txt': 204, 'Q1-Dataset/pg345.txt': 22, 'Q1-Dataset/pg64317.txt': 5, 'Q1-Dataset/pg70815.txt': 16, 'Q1-Dataset/1080-0.txt': 5, 'Q1-Dataset/pg70817.txt': 5, 'Q1-Dataset/pg1998.txt': 9}, 'ebooks': {'Q1-Dataset/pg1513.txt': 7, 'Q1-Dataset/2160-0.txt': 7, 'Q1-Dataset/pg11.txt': 7, 'Q1-Dataset/pg2600.txt': 7, 'Q1-Dataset/pg345.txt': 7, 'Q1-Dataset/pg64317.txt': 8, 'Q1-Dataset/pg70815.txt': 7, 'Q1-Dataset/1080-0.txt': 7, 'Q1-Dataset/pg70817.txt': 7, 'Q1-Dataset/pg1998.txt': 7}, 'educational': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 5, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'within': {'Q1-Dataset/pg1513.txt': 15, 'Q1-Dataset/2160-0.txt': 54, 'Q1-Dataset/pg11.txt': 5, 'Q1-Dataset/pg2600.txt': 86, 'Q1-Dataset/pg345.txt': 48, 'Q1-Dataset/pg64317.txt': 10, 'Q1-Dataset/pg70815.txt': 63, 'Q1-Dataset/1080-0.txt': 5, 'Q1-Dataset/pg70817.txt': 13, 'Q1-Dataset/pg1998.txt': 11}, 'care': {'Q1-Dataset/pg1513.txt': 6, 'Q1-Dataset/2160-0.txt': 47, 'Q1-Dataset/pg11.txt': 4, 'Q1-Dataset/pg2600.txt': 48, 'Q1-Dataset/pg345.txt': 52, 'Q1-Dataset/pg64317.txt': 14, 'Q1-Dataset/pg70815.txt': 10, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 10}, 'concerning': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 10, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 30, 'Q1-Dataset/

pg345.txt': 7, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 6, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 4, 'Q1-Dataset/pg1998.txt': 34}, 'parts': {'Q1-Dataset/pg1513.txt': 4, 'Q1-Dataset/2160-0.txt': 39, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.txt': 24, 'Q1-Dataset/pg345.txt': 6, 'Q1-Dataset/pg64317.txt': 3, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 4, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 7}, 'necessarily': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 3, 'Q1-Dataset/pg345.txt': 2, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 2}, 'processing': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'ut': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'indemnity': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'soon': {'Q1-Dataset/pg1513.txt': 13, 'Q1-Dataset/2160-0.txt': 40, 'Q1-Dataset/pg11.txt': 25, 'Q1-Dataset/pg2600.txt': 252, 'Q1-Dataset/pg345.txt': 68, 'Q1-Dataset/pg64317.txt': 6, 'Q1-Dataset/pg70815.txt': 27, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 8, 'Q1-Dataset/pg1998.txt': 19}, 'taxes': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'distribute': {'Q1-Dataset/pg1513.txt': 6, 'Q1-Dataset/2160-0.txt': 7, 'Q1-Dataset/pg11.txt': 6, 'Q1-Dataset/pg2600.txt': 7, 'Q1-Dataset/pg345.txt': 7, 'Q1-Dataset/pg64317.txt': 6, 'Q1-Dataset/pg70815.txt': 8, 'Q1-Dataset/1080-0.txt': 6, 'Q1-Dataset/pg70817.txt': 6, 'Q1-Dataset/pg1998.txt': 8}, 'compilation': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'web': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 5, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 5, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 4}, 'email': {'Q1-Dataset/pg1513.txt': 3, 'Q1-Dataset/2160-0.txt': 3, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.txt': 3, 'Q1-Dataset/pg345.txt': 3, 'Q1-Dataset/pg64317.txt': 3, 'Q1-Dataset/pg70815.txt': 3, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 3, 'Q1-Dataset/pg1998.txt': 3}, 'how': {'Q1-Dataset/pg1513.txt': 62, 'Q1-Dataset/2160-0.txt': 98, 'Q1-Dataset/pg11.txt': 67, 'Q1-Dataset/pg2600.txt': 995, 'Q1-Dataset/pg345.txt': 217, 'Q1-Dataset/pg64317.txt': 41, 'Q1-Dataset/pg70815.txt': 70, 'Q1-Dataset/1080-0.txt': 8, 'Q1-Dataset/pg70817.txt': 19, 'Q1-Dataset/pg1998.txt': 251}, 'return': {'Q1-Dataset/pg1513.txt': 8, 'Q1-Dataset/2160-0.txt': 49, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.txt': 98, 'Q1-Dataset/pg345.txt': 30, 'Q1-Dataset/pg64317.txt': 4, 'Q1-Dataset/pg70815.txt': 48, 'Q1-Dataset/1080-0.txt': 4, 'Q1-Dataset/pg70817.txt': 7, 'Q1-Dataset/pg1998.txt': 37}, 'wwwgutenbergorglicense': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'much': {'Q1-Dataset/pg1513.txt': 38, 'Q1-Dataset/2160-0.txt': 235, 'Q1-Dataset/pg11.txt': 52, 'Q1-Dataset/pg2600.txt': 394, 'Q1-Dataset/pg345.txt': 218, 'Q1-Dataset/pg64317.txt': 37, 'Q1-Dataset/pg70815.txt': 59, 'Q1-Dataset/1080-0.txt': 6, 'Q1-Dataset/pg70817.txt': 19, 'Q1-Dataset/pg1998.txt': 142}, 'manner': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 70, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 61, 'Q1-Dataset/pg345.txt': 33, 'Q1-Dataset/pg64317.txt': 4, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 4, 'Q1-Dataset/pg1998.txt': 5}, 'treatment': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 10, 'Q1-Dataset/pg345.txt': 4, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'codes': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 4, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 2}, 'were': {'Q1-Dataset/pg1513.txt': 37, 'Q1-Dataset/2160-0.txt': 341, 'Q1-Dataset/pg11.txt': 87, 'Q1-Dataset/pg2600.txt': 2433, 'Q1-Dataset/pg345.txt': 550, 'Q1-Dataset/pg64317.txt': 184, 'Q1-Dataset/pg70815.txt': 483, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 112, 'Q1-Dataset/pg1998.txt': 19}}

[illegible]

-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 5, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 4}, 'version': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 4, 'Q1-Dataset/pg1998.txt': 1}, 'mysql': {'Q1-Dataset/pg1513.txt': 15, 'Q1-Dataset/2160-0.txt': 88, 'Q1-Dataset/pg11.txt': 6, 'Q1-Dataset/pg2600.txt': 146, 'Q1-Dataset/pg345.txt': 132, 'Q1-Dataset/pg64317.txt': 17, 'Q1-Dataset/pg70815.txt': 4, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 12, 'Q1-Dataset/pg1998.txt': 93}, 'created': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 5, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 7, 'Q1-Dataset/pg345.txt': 3, 'Q1-Dataset/pg64317.txt': 3, 'Q1-Dataset/pg70815.txt': 5, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 28}, 'internal': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 3, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 5, 'Q1-Dataset/pg345.txt': 2, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 5}, 'days': {'Q1-Dataset/pg1513.txt': 16, 'Q1-Dataset/2160-0.txt': 55, 'Q1-Dataset/pg11.txt': 8, 'Q1-Dataset/pg2600.txt': 214, 'Q1-Dataset/pg345.txt': 63, 'Q1-Dataset/pg64317.txt': 22, 'Q1-Dataset/pg70815.txt': 8, 'Q1-Dataset/1080-0.txt': 4, 'Q1-Dataset/pg70817.txt': 17, 'Q1-Dataset/pg1998.txt': 22}, 'out': {'Q1-Dataset/pg1513.txt': 54, 'Q1-Dataset/2160-0.txt': 219, 'Q1-Dataset/pg11.txt': 114, 'Q1-Dataset/pg2600.txt': 1225, 'Q1-Dataset/pg345.txt': 429, 'Q1-Dataset/pg64317.txt': 211, 'Q1-Dataset/pg70815.txt': 116, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 107, 'Q1-Dataset/pg1998.txt': 246}, 'twenty': {'Q1-Dataset/pg1513.txt': 11, 'Q1-Dataset/2160-0.txt': 43, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 55, 'Q1-Dataset/pg345.txt': 8, 'Q1-Dataset/pg64317.txt': 4, 'Q1-Dataset/pg70815.txt': 8, 'Q1-Dataset/1080-0.txt': 5, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 2}, 'appearing': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 3, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 13, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 2}, 'violates': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'financial': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 3, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 3, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'new': {'Q1-Dataset/pg1513.txt': 16, 'Q1-Dataset/2160-0.txt': 62, 'Q1-Dataset/pg11.txt': 8, 'Q1-Dataset/pg2600.txt': 285, 'Q1-Dataset/pg345.txt': 80, 'Q1-Dataset/pg64317.txt': 73, 'Q1-Dataset/pg70815.txt': 45, 'Q1-Dataset/1080-0.txt': 5, 'Q1-Dataset/pg70817.txt': 24, 'Q1-Dataset/pg1998.txt': 155}, 'form': {'Q1-Dataset/pg1513.txt': 9, 'Q1-Dataset/2160-0.txt': 19, 'Q1-Dataset/pg11.txt': 5, 'Q1-Dataset/pg2600.txt': 56, 'Q1-Dataset/pg345.txt': 24, 'Q1-Dataset/pg64317.txt': 7, 'Q1-Dataset/pg70815.txt': 13, 'Q1-Dataset/1080-0.txt': 4, 'Q1-Dataset/pg70817.txt': 7, 'Q1-Dataset/pg1998.txt': 20}, 'incomplete': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 2, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 1}, 'cannot': {'Q1-Dataset/pg1513.txt': 25, 'Q1-Dataset/2160-0.txt': 102, 'Q1-Dataset/pg11.txt': 5, 'Q1-Dataset/pg2600.txt': 167, 'Q1-Dataset/pg345.txt': 88, 'Q1-Dataset/pg64317.txt': 6, 'Q1-Dataset/pg70815.txt': 36, 'Q1-Dataset/1080-0.txt': 10, 'Q1-Dataset/pg70817.txt': 5, 'Q1-Dataset/pg1998.txt': 35}, 'derived': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 7, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 5, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'heart': {'Q1-Dataset/pg1513.txt': 26, 'Q1-Dataset/2160-0.txt': 88, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 175, 'Q1-Dataset/pg345.txt': 129, 'Q1-Dataset/pg64317.txt': 19, 'Q1-Dataset/pg70815.txt': 18, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 15, 'Q1-Dataset/pg1998.txt': 171}, 'donate': {'Q1-Dataset/pg1513.txt': 3, 'Q1-Dataset/2160-0.txt': 3, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.txt': 3, 'Q1-Dataset/pg345.txt': 3, 'Q1-Dataset/pg64317.txt': 3, 'Q1-Dataset/pg70815.txt': 3, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 3, 'Q1-Dataset/pg1998.txt': 3}, 'these': {'Q1-Dataset/pg1513.txt': 45, 'Q1-Dataset/2160-0.txt': 194, 'Q1-Dataset/pg11.txt': 16, 'Q1-Dataset/pg2600.txt': 590, 'Q1-Dataset/pg345.txt': 117, 'Q1-Dataset/pg64317.txt': 28, 'Q1-Dataset/pg70815.txt': 102, 'Q1-Dataset/1080-0.txt': 10, 'Q1-Dataset/pg70817.txt': 17, 'Q1-Dataset/pg1998.txt': 127}, 'longer': {'Q1-Dataset/pg1513.txt': 5, 'Q1-Dataset/2160-0.txt': 34, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.txt': 176, 'Q1-Dataset/pg345.txt': 17, 'Q1-Dataset/pg64317.txt': 14, 'Q1-Dataset/pg70815.txt': 18, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 10, 'Q1-Dataset/pg1998.txt': 1}

89}, 'ten': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/pg2160-0.txt': 24, 'Q1-Dataset/pg11.txt': 6, 'Q1-Dataset/pg2600.txt': 98, 'Q1-Dataset/pg345.txt': 14, 'Q1-Dataset/pg64317.txt': 8, 'Q1-Dataset/pg70815.txt': 15, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 19}, 'month': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 7, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 32, 'Q1-Dataset/pg345.txt': 4, 'Q1-Dataset/pg64317.txt': 7, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 2}, 'way': {'Q1-Dataset/pg1513.txt': 11, 'Q1-Dataset/2160-0.txt': 134, 'Q1-Dataset/pg11.txt': 54, 'Q1-Dataset/pg2600.txt': 493, 'Q1-Dataset/pg345.txt': 226, 'Q1-Dataset/pg64317.txt': 75, 'Q1-Dataset/pg70815.txt': 110, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 18, 'Q1-Dataset/pg1998.txt': 110}, 'medium': {'Q1-Dataset/pg1513.txt': 5, 'Q1-Dataset/2160-0.txt': 7, 'Q1-Dataset/pg11.txt': 5, 'Q1-Dataset/pg2600.txt': 10, 'Q1-Dataset/pg345.txt': 9, 'Q1-Dataset/pg64317.txt': 7, 'Q1-Dataset/pg70815.txt': 5, 'Q1-Dataset/1080-0.txt': 6, 'Q1-Dataset/pg70817.txt': 5, 'Q1-Dataset/pg1998.txt': 6}, 'nor': {'Q1-Dataset/pg1513.txt': 20, 'Q1-Dataset/2160-0.txt': 89, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.txt': 175, 'Q1-Dataset/pg345.txt': 29, 'Q1-Dataset/pg64317.txt': 4, 'Q1-Dataset/pg70815.txt': 85, 'Q1-Dataset/1080-0.txt': 7, 'Q1-Dataset/pg70817.txt': 6, 'Q1-Dataset/pg1998.txt': 35}, 'title': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 11, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 10, 'Q1-Dataset/pg345.txt': 5, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 3}, 'things': {'Q1-Dataset/pg1513.txt': 10, 'Q1-Dataset/2160-0.txt': 26, 'Q1-Dataset/pg11.txt': 30, 'Q1-Dataset/pg2600.txt': 232, 'Q1-Dataset/pg345.txt': 173, 'Q1-Dataset/pg64317.txt': 32, 'Q1-Dataset/pg70815.txt': 15, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 11, 'Q1-Dataset/pg1998.txt': 156}, 'in': {'Q1-Dataset/pg1513.txt': 393, 'Q1-Dataset/2160-0.txt': 3155, 'Q1-Dataset/pg11.txt': 431, 'Q1-Dataset/pg2600.txt': 8930, 'Q1-Dataset/pg345.txt': 2559, 'Q1-Dataset/pg64317.txt': 861, 'Q1-Dataset/pg70815.txt': 1169, 'Q1-Dataset/1080-0.txt': 132, 'Q1-Dataset/pg70817.txt': 465, 'Q1-Dataset/pg1998.txt': 1534}, 'provision': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 9, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 4, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 4, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 3, 'Q1-Dataset/pg1998.txt': 1}, 'long': {'Q1-Dataset/pg1513.txt': 17, 'Q1-Dataset/2160-0.txt': 102, 'Q1-Dataset/pg11.txt': 34, 'Q1-Dataset/pg2600.txt': 512, 'Q1-Dataset/pg345.txt': 148, 'Q1-Dataset/pg64317.txt': 63, 'Q1-Dataset/pg70815.txt': 107, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 27, 'Q1-Dataset/pg1998.txt': 165}, 'inaccurate': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'additions': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 2, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'returns': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 6, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 8, 'Q1-Dataset/pg345.txt': 3, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 7, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 2}, 'proprietary': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'we': {'Q1-Dataset/pg1513.txt': 75, 'Q1-Dataset/2160-0.txt': 531, 'Q1-Dataset/pg11.txt': 37, 'Q1-Dataset/pg2600.txt': 956, 'Q1-Dataset/pg345.txt': 1534, 'Q1-Dataset/pg64317.txt': 201, 'Q1-Dataset/pg70815.txt': 361, 'Q1-Dataset/1080-0.txt': 18, 'Q1-Dataset/pg70817.txt': 86, 'Q1-Dataset/pg1998.txt': 283}, 'print': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 3, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'negligence': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 3, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'walk': {'Q1-Dataset/pg1513.txt': 4, 'Q1-Dataset/2160-0.txt': 17, 'Q1-Dataset/pg11.txt': 5, 'Q1-Dataset/pg2600.txt': 46, 'Q1-Dataset/pg345.txt': 18, 'Q1-Dataset/pg64317.txt': 5, 'Q1-Dataset/pg70815.txt': 6, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 8, 'Q1-Dataset/pg1998.txt': 16}, 'change': {'Q1-Dataset/pg1513.txt': 6, 'Q1-Dataset/2160-0.txt': 21, 'Q1-Dataset/pg11.txt': 15, 'Q1-Dataset/pg2600.txt': 83, 'Q1-Dataset/pg345.txt': 44, 'Q1-Dataset/pg64317.txt': 13, 'Q1-Dataset/pg70815.txt': 5, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 4, 'Q1-Dataset/pg1998.txt': 15}, 'array': {'Q1-Dataset/pg1513.txt': 3, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 2, 'Q1-D

```

Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/pg1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 1}, 'holder': {'Q1-Dataset/pg1513.txt': 4, 'Q1-Dataset/2160-0.txt': 7, 'Q1-Dataset/pg11.txt': 4, 'Q1-Dataset/pg2600.txt': 4, 'Q1-Dataset/pg345.txt': 4, 'Q1-Dataset/pg64317.txt': 4, 'Q1-Dataset/pg70815.txt': 4, 'Q1-Dataset/pg1080-0.txt': 4, 'Q1-Dataset/pg70817.txt': 4, 'Q1-Dataset/pg1998.txt': 4}, 'production': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 3, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 2, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/pg1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 3}, 'consequential': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/pg1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'whose': {'Q1-Dataset/pg1513.txt': 15, 'Q1-Dataset/2160-0.txt': 69, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 111, 'Q1-Dataset/pg345.txt': 30, 'Q1-Dataset/pg64317.txt': 9, 'Q1-Dataset/pg70815.txt': 29, 'Q1-Dataset/pg1080-0.txt': 4, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 43}, 'carry': {'Q1-Dataset/pg1513.txt': 3, 'Q1-Dataset/2160-0.txt': 17, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 53, 'Q1-Dataset/pg345.txt': 16, 'Q1-Dataset/pg64317.txt': 5, 'Q1-Dataset/pg70815.txt': 16, 'Q1-Dataset/pg1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 21}, 'meet': {'Q1-Dataset/pg1513.txt': 9, 'Q1-Dataset/2160-0.txt': 21, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.txt': 107, 'Q1-Dataset/pg345.txt': 42, 'Q1-Dataset/pg64317.txt': 9, 'Q1-Dataset/pg70815.txt': 17, 'Q1-Dataset/pg1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 3, 'Q1-Dataset/pg1998.txt': 20}, 'actual': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 16, 'Q1-Dataset/pg345.txt': 4, 'Q1-Dataset/pg64317.txt': 3, 'Q1-Dataset/pg70815.txt': 3, 'Q1-Dataset/pg1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 6}, 'subscribe': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 4, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 2, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/pg1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'include': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 3, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/pg1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'pglaf': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/pg1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'extent': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 7, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 13, 'Q1-Dataset/pg345.txt': 3, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 7, 'Q1-Dataset/pg1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 4}, 'visit': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 40, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.txt': 61, 'Q1-Dataset/pg345.txt': 17, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 4, 'Q1-Dataset/pg1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 7, 'Q1-Dataset/pg1998.txt': 2}, 'far': {'Q1-Dataset/pg1513.txt': 5, 'Q1-Dataset/2160-0.txt': 75, 'Q1-Dataset/pg11.txt': 13, 'Q1-Dataset/pg2600.txt': 169, 'Q1-Dataset/pg345.txt': 91, 'Q1-Dataset/pg64317.txt': 18, 'Q1-Dataset/pg70815.txt': 84, 'Q1-Dataset/pg1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 21, 'Q1-Dataset/pg1998.txt': 52}, 'produce': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 25, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 18, 'Q1-Dataset/pg345.txt': 2, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/pg1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 3}, 'this': {'Q1-Dataset/pg1513.txt': 265, 'Q1-Dataset/2160-0.txt': 974, 'Q1-Dataset/pg11.txt': 178, 'Q1-Dataset/pg2600.txt': 2101, 'Q1-Dataset/pg345.txt': 665, 'Q1-Dataset/pg64317.txt': 210, 'Q1-Dataset/pg70815.txt': 313, 'Q1-Dataset/pg1080-0.txt': 73, 'Q1-Dataset/pg70817.txt': 85, 'Q1-Dataset/pg1998.txt': 658}, 'rest': {'Q1-Dataset/pg1513.txt': 16, 'Q1-Dataset/2160-0.txt': 29, 'Q1-Dataset/pg11.txt': 10, 'Q1-Dataset/pg2600.txt': 107, 'Q1-Dataset/pg345.txt': 95, 'Q1-Dataset/pg64317.txt': 11, 'Q1-Dataset/pg70815.txt': 18, 'Q1-Dataset/pg1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 6, 'Q1-Dataset/pg1998.txt': 17}, 'already': {'Q1-Dataset/pg1513.txt': 6, 'Q1-Dataset/2160-0.txt': 39, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.txt': 326, 'Q1-Dataset/pg345.txt': 62, 'Q1-Dataset/pg64317.txt': 16, 'Q1-Dataset/pg70815.txt': 27, 'Q1-Dataset/pg1080-0.txt': 7, 'Q1-Dataset/pg70817.txt': 17, 'Q1-Dataset/pg1998.txt': 78}, 'project': {'Q1-Dataset/pg1513.txt': 89, 'Q1-Dataset/2160-0.txt': 94, 'Q1-Dataset/pg11.txt': 88, 'Q1-Dataset/pg2600.txt': 101, 'Q1-Dataset/pg345.txt': 89, 'Q1-Dataset/pg64317.txt': 90, 'Q1-Dataset/pg70815.txt': 90, 'Q1-Dataset/pg1080-0.txt': 89, 'Q1-Dataset/pg70817.txt': 89, 'Q1-Dataset/pg1998.txt': 88}, 'prominently': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 2, 'Q1-Dataset/pg345.txt': 3, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/pg1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 2}

```

txt': 2, 'Q1-Dataset/pg1998.txt': 2}, 'based': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 26, 'Q1-Dataset/pg345.txt': 3, 'Q1-Dataset/pg64317.txt': 3, 'Q1-Dataset/pg70815.txt': 6, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 2}, 'compressed': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 11, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 4}, 'corrupt': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 2, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 2}, 'search': {'Q1-Dataset/pg1513.txt': 5, 'Q1-Dataset/2160-0.txt': 6, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 21, 'Q1-Dataset/pg345.txt': 27, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 41, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 5, 'Q1-Dataset/pg1998.txt': 7}, 'modified': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'punitive': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 2}, 'states': {'Q1-Dataset/pg1513.txt': 20, 'Q1-Dataset/2160-0.txt': 16, 'Q1-Dataset/pg11.txt': 20, 'Q1-Dataset/pg2600.txt': 25, 'Q1-Dataset/pg345.txt': 23, 'Q1-Dataset/pg64317.txt': 21, 'Q1-Dataset/pg70815.txt': 20, 'Q1-Dataset/1080-0.txt': 20, 'Q1-Dataset/pg70817.txt': 20, 'Q1-Dataset/pg1998.txt': 21}, 'also': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 41, 'Q1-Dataset/pg11.txt': 4, 'Q1-Dataset/pg2600.txt': 268, 'Q1-Dataset/pg345.txt': 43, 'Q1-Dataset/pg64317.txt': 7, 'Q1-Dataset/pg70815.txt': 20, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 4, 'Q1-Dataset/pg1998.txt': 337}, 'north': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 18, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 2, 'Q1-Dataset/pg345.txt': 13, 'Q1-Dataset/pg64317.txt': 4, 'Q1-Dataset/pg70815.txt': 24, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 4, 'Q1-Dataset/pg1998.txt': 2}, 'updated': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 2, 'Q1-Dataset/pg345.txt': 2, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 2}, 'it': {'Q1-Dataset/pg1513.txt': 242, 'Q1-Dataset/2160-0.txt': 850, 'Q1-Dataset/pg11.txt': 538, 'Q1-Dataset/pg2600.txt': 5179, 'Q1-Dataset/pg345.txt': 2137, 'Q1-Dataset/pg64317.txt': 539, 'Q1-Dataset/pg70815.txt': 907, 'Q1-Dataset/1080-0.txt': 41, 'Q1-Dataset/pg70817.txt': 325, 'Q1-Dataset/pg1998.txt': 1477}, 'forth': {'Q1-Dataset/pg1513.txt': 20, 'Q1-Dataset/2160-0.txt': 34, 'Q1-Dataset/pg11.txt': 8, 'Q1-Dataset/pg2600.txt': 19, 'Q1-Dataset/pg345.txt': 18, 'Q1-Dataset/pg64317.txt': 11, 'Q1-Dataset/pg70815.txt': 16, 'Q1-Dataset/1080-0.txt': 8, 'Q1-Dataset/pg70817.txt': 9, 'Q1-Dataset/pg1998.txt': 37}, 'periodic': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'posted': {'Q1-Dataset/pg1513.txt': 4, 'Q1-Dataset/2160-0.txt': 5, 'Q1-Dataset/pg11.txt': 4, 'Q1-Dataset/pg2600.txt': 8, 'Q1-Dataset/pg345.txt': 6, 'Q1-Dataset/pg64317.txt': 4, 'Q1-Dataset/pg70815.txt': 4, 'Q1-Dataset/1080-0.txt': 4, 'Q1-Dataset/pg70817.txt': 4, 'Q1-Dataset/pg1998.txt': 4}, 'agreed': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 21, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 33, 'Q1-Dataset/pg345.txt': 13, 'Q1-Dataset/pg64317.txt': 10, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 6, 'Q1-Dataset/pg1998.txt': 2}, 'not': {'Q1-Dataset/pg1513.txt': 288, 'Q1-Dataset/2160-0.txt': 923, 'Q1-Dataset/pg11.txt': 165, 'Q1-Dataset/pg2600.txt': 4665, 'Q1-Dataset/pg345.txt': 1410, 'Q1-Dataset/pg64317.txt': 97, 'Q1-Dataset/pg70815.txt': 586, 'Q1-Dataset/1080-0.txt': 44, 'Q1-Dataset/pg70817.txt': 115, 'Q1-Dataset/pg1998.txt': 1003}, 'contain': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 7, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 5, 'Q1-Dataset/pg345.txt': 6, 'Q1-Dataset/pg64317.txt': 3, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 2}, 'reading': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 6, 'Q1-Dataset/pg11.txt': 4, 'Q1-Dataset/pg2600.txt': 83, 'Q1-Dataset/pg345.txt': 18, 'Q1-Dataset/pg64317.txt': 4, 'Q1-Dataset/pg70815.txt': 4, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 8}, 'trademarkcopyright': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'fell': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 15, 'Q1-Dataset/pg11.txt': 6, 'Q1-

-Dataset/pg2600.txt': 125, 'Q1-Dataset/pg345.txt': 59, 'Q1-Dataset/pg64317.txt': 17, 'Q1-Dataset/pg70815.txt': 25, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 11, 'Q1-Dataset/pg1998.txt': 12}, 'sentence': {'Q1-Dataset/pg1513.txt': 4, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 8, 'Q1-Dataset/pg2600.txt': 14, 'Q1-Dataset/pg345.txt': 3, 'Q1-Dataset/pg64317.txt': 4, 'Q1-Dataset/pg70815.txt': 3, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 3, 'Q1-Dataset/pg1998.txt': 2}, 'data': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 3, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'fast': {'Q1-Dataset/pg1513.txt': 3, 'Q1-Dataset/2160-0.txt': 8, 'Q1-Dataset/pg11.txt': 4, 'Q1-Dataset/pg2600.txt': 23, 'Q1-Dataset/pg345.txt': 15, 'Q1-Dataset/pg64317.txt': 3, 'Q1-Dataset/pg70815.txt': 6, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 3, 'Q1-Dataset/pg1998.txt': 13}, 'brought': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 62, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.txt': 210, 'Q1-Dataset/pg345.txt': 36, 'Q1-Dataset/pg64317.txt': 18, 'Q1-Dataset/pg70815.txt': 21, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 7, 'Q1-Dataset/pg1998.txt': 8}, 'pages': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 9, 'Q1-Dataset/pg345.txt': 3, 'Q1-Dataset/pg64317.txt': 4, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 4, 'Q1-Dataset/pg1998.txt': 4}, 'its': {'Q1-Dataset/pg1513.txt': 3, 'Q1-Dataset/2160-0.txt': 96, 'Q1-Dataset/pg11.txt': 114, 'Q1-Dataset/pg2600.txt': 945, 'Q1-Dataset/pg345.txt': 153, 'Q1-Dataset/pg64317.txt': 96, 'Q1-Dataset/pg70815.txt': 239, 'Q1-Dataset/1080-0.txt': 4, 'Q1-Dataset/pg70817.txt': 71, 'Q1-Dataset/pg1998.txt': 265}, 'without': {'Q1-Dataset/pg1513.txt': 21, 'Q1-Dataset/2160-0.txt': 305, 'Q1-Dataset/pg11.txt': 33, 'Q1-Dataset/pg2600.txt': 670, 'Q1-Dataset/pg345.txt': 140, 'Q1-Dataset/pg64317.txt': 35, 'Q1-Dataset/pg70815.txt': 57, 'Q1-Dataset/1080-0.txt': 12, 'Q1-Dataset/pg70817.txt': 31, 'Q1-Dataset/pg1998.txt': 73}, 'few': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 69, 'Q1-Dataset/pg11.txt': 10, 'Q1-Dataset/pg2600.txt': 171, 'Q1-Dataset/pg345.txt': 86, 'Q1-Dataset/pg64317.txt': 26, 'Q1-Dataset/pg70815.txt': 44, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 25, 'Q1-Dataset/pg1998.txt': 12}, 'hypertext': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'destroy': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 5, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 35, 'Q1-Dataset/pg345.txt': 20, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 13, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 3, 'Q1-Dataset/pg1998.txt': 8}, 'every': {'Q1-Dataset/pg1513.txt': 10, 'Q1-Dataset/2160-0.txt': 212, 'Q1-Dataset/pg11.txt': 12, 'Q1-Dataset/pg2600.txt': 334, 'Q1-Dataset/pg345.txt': 84, 'Q1-Dataset/pg64317.txt': 27, 'Q1-Dataset/pg70815.txt': 52, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 21, 'Q1-Dataset/pg1998.txt': 125}, 'my': {'Q1-Dataset/pg1513.txt': 356, 'Q1-Dataset/2160-0.txt': 1267, 'Q1-Dataset/pg11.txt': 58, 'Q1-Dataset/pg2600.txt': 1220, 'Q1-Dataset/pg345.txt': 1257, 'Q1-Dataset/pg64317.txt': 242, 'Q1-Dataset/pg70815.txt': 94, 'Q1-Dataset/1080-0.txt': 20, 'Q1-Dataset/pg70817.txt': 74, 'Q1-Dataset/pg1998.txt': 974}, 'profits': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 4, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'since': {'Q1-Dataset/pg1513.txt': 11, 'Q1-Dataset/2160-0.txt': 44, 'Q1-Dataset/pg11.txt': 4, 'Q1-Dataset/pg2600.txt': 188, 'Q1-Dataset/pg345.txt': 85, 'Q1-Dataset/pg64317.txt': 12, 'Q1-Dataset/pg70815.txt': 63, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 8, 'Q1-Dataset/pg1998.txt': 17}, 'concept': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 3, 'Q1-Dataset/pg345.txt': 2, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 3}, 'except': {'Q1-Dataset/pg1513.txt': 4, 'Q1-Dataset/2160-0.txt': 31, 'Q1-Dataset/pg11.txt': 8, 'Q1-Dataset/pg2600.txt': 71, 'Q1-Dataset/pg345.txt': 40, 'Q1-Dataset/pg64317.txt': 19, 'Q1-Dataset/pg70815.txt': 21, 'Q1-Dataset/1080-0.txt': 6, 'Q1-Dataset/pg70817.txt': 8, 'Q1-Dataset/pg1998.txt': 23}, 'defective': {'Q1-Dataset/pg1513.txt': 3, 'Q1-Dataset/2160-0.txt': 6, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.txt': 4, 'Q1-Dataset/pg345.txt': 3, 'Q1-Dataset/pg64317.txt': 3, 'Q1-Dataset/pg70815.txt': 3, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 3, 'Q1-Dataset/pg1998.txt': 3}, 'merchantability': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'maximum': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 4, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-

```

t.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'physical': {'Q1-Da
taset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/p
g2600.txt': 36, 'Q1-Dataset/pg345.txt': 10, 'Q1-Dataset/pg64317.txt': 10, 'Q1-Dataset/p
g70815.txt': 9, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 4, 'Q1-Dataset/pg1
998.txt': 4}, 'body': {'Q1-Dataset/pg1513.txt': 9, 'Q1-Dataset/2160-0.txt': 43, 'Q1-Da
set/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 99, 'Q1-Dataset/pg345.txt': 54, 'Q1-Dataset/p
g64317.txt': 19, 'Q1-Dataset/pg70815.txt': 54, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/p
g70817.txt': 4, 'Q1-Dataset/pg1998.txt': 70}, 'under': {'Q1-Dataset/pg1513.txt': 12, 'Q1
-Dataset/2160-0.txt': 112, 'Q1-Dataset/pg11.txt': 22, 'Q1-Dataset/pg2600.txt': 371, 'Q1-
Dataset/pg345.txt': 69, 'Q1-Dataset/pg64317.txt': 38, 'Q1-Dataset/pg70815.txt': 21, 'Q1-
Dataset/1080-0.txt': 10, 'Q1-Dataset/pg70817.txt': 19, 'Q1-Dataset/pg1998.txt': 34}, 'me
thods': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt':
1, 'Q1-Dataset/pg2600.txt': 10, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1,
'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 1,
'Q1-Dataset/pg1998.txt': 1}, 'wish': {'Q1-Dataset/pg1513.txt': 7, 'Q1-Dataset/2160-0.tx
t': 29, 'Q1-Dataset/pg11.txt': 22, 'Q1-Dataset/pg2600.txt': 185, 'Q1-Dataset/pg345.txt':
51, 'Q1-Dataset/pg64317.txt': 3, 'Q1-Dataset/pg70815.txt': 16, 'Q1-Dataset/1080-0.txt':
1, 'Q1-Dataset/pg70817.txt': 4, 'Q1-Dataset/pg1998.txt': 25}, 'immediately': {'Q1-Datas
et/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 56, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg
2600.txt': 130, 'Q1-Dataset/pg345.txt': 6, 'Q1-Dataset/pg64317.txt': 15, 'Q1-Dataset/pg7
0815.txt': 17, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 5, 'Q1-Dataset/pg19
98.txt': 15}, 'so': {'Q1-Dataset/pg1513.txt': 149, 'Q1-Dataset/2160-0.txt': 598, 'Q1-Da
taset/pg11.txt': 152, 'Q1-Dataset/pg2600.txt': 1853, 'Q1-Dataset/pg345.txt': 1094, 'Q1-Da
taset/pg64317.txt': 137, 'Q1-Dataset/pg70815.txt': 230, 'Q1-Dataset/1080-0.txt': 15, 'Q1
-Dataset/pg70817.txt': 52, 'Q1-Dataset/pg1998.txt': 306}, 'constant': {'Q1-Dataset/pg151
3.txt': 1, 'Q1-Dataset/2160-0.txt': 11, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.tx
t': 18, 'Q1-Dataset/pg345.txt': 4, 'Q1-Dataset/pg64317.txt': 5, 'Q1-Dataset/pg70815.tx
t': 9, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt':
1}, 'outside': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 5, 'Q1-Dataset/pg1
1.txt': 5, 'Q1-Dataset/pg2600.txt': 46, 'Q1-Dataset/pg345.txt': 37, 'Q1-Dataset/pg64317.
txt': 18, 'Q1-Dataset/pg70815.txt': 5, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.t
xt': 18, 'Q1-Dataset/pg1998.txt': 7}, 'applicable': {'Q1-Dataset/pg1513.txt': 3, 'Q1-Da
taset/2160-0.txt': 4, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.txt': 5, 'Q1-Dataset/p
g345.txt': 3, 'Q1-Dataset/pg64317.txt': 3, 'Q1-Dataset/pg70815.txt': 3, 'Q1-Dataset/1080
-0.txt': 3, 'Q1-Dataset/pg70817.txt': 3, 'Q1-Dataset/pg1998.txt': 3}, 'possessed': {'Q1-
Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 8, 'Q1-Dataset/pg11.txt': 1, 'Q1-Datase
t/pg2600.txt': 21, 'Q1-Dataset/pg345.txt': 2, 'Q1-Dataset/pg64317.txt': 4, 'Q1-Dataset/p
g70815.txt': 4, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 5, 'Q1-Dataset/pg1
998.txt': 7}, 'or': {'Q1-Dataset/pg1513.txt': 144, 'Q1-Dataset/2160-0.txt': 486, 'Q1-Da
taset/pg11.txt': 154, 'Q1-Dataset/pg2600.txt': 1589, 'Q1-Dataset/pg345.txt': 524, 'Q1-Da
taset/pg64317.txt': 173, 'Q1-Dataset/pg70815.txt': 297, 'Q1-Dataset/1080-0.txt': 113, 'Q1
-Dataset/pg70817.txt': 124, 'Q1-Dataset/pg1998.txt': 289}, 'identification': {'Q1-Datase
t/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg26
00.txt': 2, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.
txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.tx
t': 1}, 'onethe': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/p
g11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.
txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.tx
t': 1, 'Q1-Dataset/pg1998.txt': 3}, 'previous': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Datase
t/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 39, 'Q1-Dataset/pg3
45.txt': 9, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-
0.txt': 1, 'Q1-Dataset/pg70817.txt': 3, 'Q1-Dataset/pg1998.txt': 5}, 'exporting': {'Q1-D
ataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Datase
t/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg
70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg19
98.txt': 1}, 'city': {'Q1-Dataset/pg1513.txt': 5, 'Q1-Dataset/2160-0.txt': 25, 'Q1-Datas
et/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 67, 'Q1-Dataset/pg345.txt': 8, 'Q1-Dataset/pg6
4317.txt': 20, 'Q1-Dataset/pg70815.txt': 14, 'Q1-Dataset/1080-0.txt': 4, 'Q1-Dataset/pg7
0817.txt': 8, 'Q1-Dataset/pg1998.txt': 18}, 'shall': {'Q1-Dataset/pg1513.txt': 112, 'Q1-
Dataset/2160-0.txt': 221, 'Q1-Dataset/pg11.txt': 27, 'Q1-Dataset/pg2600.txt': 248, 'Q1-D
ataset/pg345.txt': 426, 'Q1-Dataset/pg64317.txt': 6, 'Q1-Dataset/pg70815.txt': 96, 'Q1-D
ataset/1080-0.txt': 9, 'Q1-Dataset/pg70817.txt': 6, 'Q1-Dataset/pg1998.txt': 202}, 'smal
l': {'Q1-Dataset/pg1513.txt': 4, 'Q1-Dataset/2160-0.txt': 54, 'Q1-Dataset/pg11.txt': 12,
'Q1-Dataset/pg2600.txt': 181, 'Q1-Dataset/pg345.txt': 41, 'Q1-Dataset/pg64317.txt': 34,
'Q1-Dataset/pg70815.txt': 34, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 16,
'Q1-Dataset/pg1998.txt': 80}, 'found': {'Q1-Dataset/pg1513.txt': 9, 'Q1-Dataset/2160-0.

```

txt': 120, 'Q1-Dataset/pg11.txt': 35, 'Q1-Dataset/pg2600.txt': 222, 'Q1-Dataset/pg345.txt': 155, 'Q1-Dataset/pg64317.txt': 38, 'Q1-Dataset/pg70815.txt': 74, 'Q1-Dataset/1080-0.txt': 5, 'Q1-Dataset/pg70817.txt': 25, 'Q1-Dataset/pg1998.txt': 83}, 'whom': {'Q1-Dataset/pg1513.txt': 5, 'Q1-Dataset/2160-0.txt': 136, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 400, 'Q1-Dataset/pg345.txt': 43, 'Q1-Dataset/pg64317.txt': 10, 'Q1-Dataset/pg70815.txt': 23, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 100}, 'wife': {'Q1-Dataset/pg1513.txt': 20, 'Q1-Dataset/2160-0.txt': 42, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 243, 'Q1-Dataset/pg345.txt': 30, 'Q1-Dataset/pg64317.txt': 22, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 2}, 'general': {'Q1-Dataset/pg1513.txt': 5, 'Q1-Dataset/2160-0.txt': 48, 'Q1-Dataset/pg11.txt': 6, 'Q1-Dataset/pg2600.txt': 484, 'Q1-Dataset/pg345.txt': 17, 'Q1-Dataset/pg64317.txt': 5, 'Q1-Dataset/pg70815.txt': 9, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 4, 'Q1-Dataset/pg1998.txt': 9}, 'facility': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 2, 'Q1-Dataset/pg345.txt': 2, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'employees': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'easy': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 19, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.txt': 49, 'Q1-Dataset/pg345.txt': 16, 'Q1-Dataset/pg64317.txt': 3, 'Q1-Dataset/pg70815.txt': 7, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 4, 'Q1-Dataset/pg1998.txt': 4}, 'discontinue': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 2}, 'variety': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 23, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 12, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 2}, 'about': {'Q1-Dataset/pg1513.txt': 20, 'Q1-Dataset/2160-0.txt': 146, 'Q1-Dataset/pg11.txt': 102, 'Q1-Dataset/pg2600.txt': 1016, 'Q1-Dataset/pg345.txt': 239, 'Q1-Dataset/pg64317.txt': 157, 'Q1-Dataset/pg70815.txt': 133, 'Q1-Dataset/1080-0.txt': 13, 'Q1-Dataset/pg70817.txt': 58, 'Q1-Dataset/pg1998.txt': 130}, 'be': {'Q1-Dataset/pg1513.txt': 227, 'Q1-Dataset/2160-0.txt': 985, 'Q1-Dataset/pg11.txt': 165, 'Q1-Dataset/pg2600.txt': 2448, 'Q1-Dataset/pg345.txt': 1125, 'Q1-Dataset/pg64317.txt': 140, 'Q1-Dataset/pg70815.txt': 325, 'Q1-Dataset/1080-0.txt': 70, 'Q1-Dataset/pg70817.txt': 103, 'Q1-Dataset/pg1998.txt': 782}, 'first': {'Q1-Dataset/pg1513.txt': 34, 'Q1-Dataset/2160-0.txt': 114, 'Q1-Dataset/pg11.txt': 49, 'Q1-Dataset/pg2600.txt': 618, 'Q1-Dataset/pg345.txt': 132, 'Q1-Dataset/pg64317.txt': 55, 'Q1-Dataset/pg70815.txt': 93, 'Q1-Dataset/1080-0.txt': 4, 'Q1-Dataset/pg70817.txt': 18, 'Q1-Dataset/pg1998.txt': 116}, 'friend': {'Q1-Dataset/pg1513.txt': 15, 'Q1-Dataset/2160-0.txt': 111, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.txt': 163, 'Q1-Dataset/pg345.txt': 182, 'Q1-Dataset/pg64317.txt': 17, 'Q1-Dataset/pg70815.txt': 30, 'Q1-Dataset/1080-0.txt': 4, 'Q1-Dataset/pg70817.txt': 3, 'Q1-Dataset/pg1998.txt': 57}, 'original': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 28, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 16, 'Q1-Dataset/pg345.txt': 5, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 10}, 'seem': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 27, 'Q1-Dataset/pg11.txt': 8, 'Q1-Dataset/pg2600.txt': 82, 'Q1-Dataset/pg345.txt': 65, 'Q1-Dataset/pg64317.txt': 4, 'Q1-Dataset/pg70815.txt': 14, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 4, 'Q1-Dataset/pg1998.txt': 32}, 'accept': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 8, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 24, 'Q1-Dataset/pg345.txt': 17, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 6, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 6}, 'main': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 3, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 14, 'Q1-Dataset/pg345.txt': 10, 'Q1-Dataset/pg64317.txt': 4, 'Q1-Dataset/pg70815.txt': 7, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 3}, 'apply': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 4, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 9, 'Q1-Dataset/pg345.txt': 2, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'get': {'Q1-Dataset/pg1513.txt': 12, 'Q1-Dataset/2160-0.txt': 25, 'Q1-Dataset/pg11.txt': 46, 'Q1-Dataset/pg2600.txt': 356, 'Q1-Dataset/pg345.txt': 172, 'Q1-Dataset/pg64317.txt': 74, 'Q1-Dataset/pg70815.txt': 19, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 11, 'Q1-Dataset/pg1998.txt': 19}, 'contact': {'Q1-Dataset/pg1513.txt': 3, 'Q1-Dataset/2160-0.txt': 5, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.txt': 8, 'Q1-Dataset/pg345.txt': 5, 'Q1-Dataset/pg64317.txt': 6, 'Q1-

-Dataset/pg70815.txt': 5, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 3, 'Q1-Dataset/pg1998.txt': 3}, 'alternate': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 3, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'section': {'Q1-Dataset/pg1513.txt': 7, 'Q1-Dataset/2160-0.txt': 7, 'Q1-Dataset/pg11.txt': 7, 'Q1-Dataset/pg2600.txt': 15, 'Q1-Dataset/pg345.txt': 9, 'Q1-Dataset/pg64317.txt': 7, 'Q1-Dataset/pg70815.txt': 7, 'Q1-Dataset/1080-0.txt': 7, 'Q1-Dataset/pg70817.txt': 8, 'Q1-Dataset/pg1998.txt': 7}, 'notice': {'Q1-Dataset/pg1513.txt': 4, 'Q1-Dataset/2160-0.txt': 25, 'Q1-Dataset/pg11.txt': 8, 'Q1-Dataset/pg2600.txt': 69, 'Q1-Dataset/pg345.txt': 29, 'Q1-Dataset/pg64317.txt': 6, 'Q1-Dataset/pg70815.txt': 8, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 3, 'Q1-Dataset/pg1998.txt': 5}, 'intellectual': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 24, 'Q1-Dataset/pg345.txt': 6, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 5}, 'direct': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 8, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 49, 'Q1-Dataset/pg345.txt': 5, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 4, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 2}, 's': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 18, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 10, 'Q1-Dataset/pg345.txt': 5, 'Q1-Dataset/pg64317.txt': 4, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'common': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 43, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 50, 'Q1-Dataset/pg345.txt': 23, 'Q1-Dataset/pg64317.txt': 5, 'Q1-Dataset/pg70815.txt': 11, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 15}, 'there': {'Q1-Dataset/pg1513.txt': 64, 'Q1-Dataset/2160-0.txt': 268, 'Q1-Dataset/pg11.txt': 76, 'Q1-Dataset/pg2600.txt': 1458, 'Q1-Dataset/pg345.txt': 768, 'Q1-Dataset/pg64317.txt': 191, 'Q1-Dataset/pg70815.txt': 297, 'Q1-Dataset/1080-0.txt': 10, 'Q1-Dataset/pg70817.txt': 81, 'Q1-Dataset/pg1998.txt': 489}, 'turn': {'Q1-Dataset/pg1513.txt': 12, 'Q1-Dataset/2160-0.txt': 39, 'Q1-Dataset/pg11.txt': 10, 'Q1-Dataset/pg2600.txt': 103, 'Q1-Dataset/pg345.txt': 33, 'Q1-Dataset/pg64317.txt': 10, 'Q1-Dataset/pg70815.txt': 22, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 9, 'Q1-Dataset/pg1998.txt': 17}, 'even': {'Q1-Dataset/pg1513.txt': 14, 'Q1-Dataset/2160-0.txt': 177, 'Q1-Dataset/pg11.txt': 20, 'Q1-Dataset/pg2600.txt': 503, 'Q1-Dataset/pg345.txt': 234, 'Q1-Dataset/pg64317.txt': 45, 'Q1-Dataset/pg70815.txt': 98, 'Q1-Dataset/1080-0.txt': 5, 'Q1-Dataset/pg70817.txt': 39, 'Q1-Dataset/pg1998.txt': 251}, 'anywhere': {'Q1-Dataset/pg1513.txt': 3, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.txt': 26, 'Q1-Dataset/pg345.txt': 18, 'Q1-Dataset/pg64317.txt': 5, 'Q1-Dataset/pg70815.txt': 4, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 3, 'Q1-Dataset/pg1998.txt': 3}, 'told': {'Q1-Dataset/pg1513.txt': 3, 'Q1-Dataset/2160-0.txt': 79, 'Q1-Dataset/pg11.txt': 6, 'Q1-Dataset/pg2600.txt': 409, 'Q1-Dataset/pg345.txt': 144, 'Q1-Dataset/pg64317.txt': 44, 'Q1-Dataset/pg70815.txt': 27, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 12, 'Q1-Dataset/pg1998.txt': 13}, 'perhaps': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 47, 'Q1-Dataset/pg11.txt': 17, 'Q1-Dataset/pg2600.txt': 115, 'Q1-Dataset/pg345.txt': 54, 'Q1-Dataset/pg64317.txt': 19, 'Q1-Dataset/pg70815.txt': 64, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 10, 'Q1-Dataset/pg1998.txt': 67}, 'further': {'Q1-Dataset/pg1513.txt': 4, 'Q1-Dataset/2160-0.txt': 40, 'Q1-Dataset/pg11.txt': 4, 'Q1-Dataset/pg2600.txt': 51, 'Q1-Dataset/pg345.txt': 34, 'Q1-Dataset/pg64317.txt': 6, 'Q1-Dataset/pg70815.txt': 21, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 11, 'Q1-Dataset/pg1998.txt': 21}, 'identify': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'produced': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 38, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.txt': 42, 'Q1-Dataset/pg345.txt': 10, 'Q1-Dataset/pg64317.txt': 5, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 3}, 'charge': {'Q1-Dataset/pg1513.txt': 9, 'Q1-Dataset/2160-0.txt': 16, 'Q1-Dataset/pg11.txt': 6, 'Q1-Dataset/pg2600.txt': 28, 'Q1-Dataset/pg345.txt': 17, 'Q1-Dataset/pg64317.txt': 6, 'Q1-Dataset/pg70815.txt': 21, 'Q1-Dataset/1080-0.txt': 11, 'Q1-Dataset/pg70817.txt': 6, 'Q1-Dataset/pg1998.txt': 7}, 'sharing': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 8, 'Q1-Dataset/pg345.txt': 3, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'modification': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'ebook': {'Q1-Dataset/pg15

13.txt': 13, 'Q1-Dataset/2160-0.txt': 10, 'Q1-Dataset/pg11.txt': 13, 'Q1-Dataset/pg2600.txt': 13, 'Q1-Dataset/pg345.txt': 13, 'Q1-Dataset/pg64317.txt': 13, 'Q1-Dataset/pg70815.txt': 13, 'Q1-Dataset/1080-0.txt': 13, 'Q1-Dataset/pg70817.txt': 13, 'Q1-Dataset/pg1998.txt': 13}, 'confirmed': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 11, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 20, 'Q1-Dataset/pg345.txt': 2, 'Q1-Dataset/pg64317.txt': 3, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'stored': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 2, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 1}, 'redistribution': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 2, 'Q1-Dataset/pg345.txt': 2, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 2}, 'copied': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 6, 'Q1-Dataset/pg345.txt': 6, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 2}, 'own': {'Q1-Dataset/pg1513.txt': 17, 'Q1-Dataset/2160-0.txt': 326, 'Q1-Dataset/pg11.txt': 10, 'Q1-Dataset/pg2600.txt': 498, 'Q1-Dataset/pg345.txt': 187, 'Q1-Dataset/pg64317.txt': 31, 'Q1-Dataset/pg70815.txt': 64, 'Q1-Dataset/1080-0.txt': 10, 'Q1-Dataset/pg70817.txt': 26, 'Q1-Dataset/pg1998.txt': 128}, 'online': {'Q1-Dataset/pg1513.txt': 4, 'Q1-Dataset/2160-0.txt': 4, 'Q1-Dataset/pg11.txt': 4, 'Q1-Dataset/pg2600.txt': 4, 'Q1-Dataset/pg345.txt': 5, 'Q1-Dataset/pg64317.txt': 4, 'Q1-Dataset/pg70815.txt': 5, 'Q1-Dataset/1080-0.txt': 4, 'Q1-Dataset/pg70817.txt': 5, 'Q1-Dataset/pg1998.txt': 4}, 'accessed': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'more': {'Q1-Dataset/pg1513.txt': 67, 'Q1-Dataset/2160-0.txt': 283, 'Q1-Dataset/pg11.txt': 48, 'Q1-Dataset/pg2600.txt': 1054, 'Q1-Dataset/pg345.txt': 373, 'Q1-Dataset/pg64317.txt': 75, 'Q1-Dataset/pg70815.txt': 121, 'Q1-Dataset/1080-0.txt': 7, 'Q1-Dataset/pg70817.txt': 40, 'Q1-Dataset/pg1998.txt': 240}, 'addresses': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 10, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 3, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 2}, 'deletions': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'fine': {'Q1-Dataset/pg1513.txt': 4, 'Q1-Dataset/2160-0.txt': 49, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 120, 'Q1-Dataset/pg345.txt': 21, 'Q1-Dataset/pg64317.txt': 10, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 4, 'Q1-Dataset/pg1998.txt': 17}, 'ask': {'Q1-Dataset/pg1513.txt': 5, 'Q1-Dataset/2160-0.txt': 8, 'Q1-Dataset/pg11.txt': 11, 'Q1-Dataset/pg2600.txt': 199, 'Q1-Dataset/pg345.txt': 68, 'Q1-Dataset/pg64317.txt': 15, 'Q1-Dataset/pg70815.txt': 8, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 26}, 'redistribute': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'implied': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 4, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 4, 'Q1-Dataset/pg345.txt': 3, 'Q1-Dataset/pg64317.txt': 3, 'Q1-Dataset/pg70815.txt': 3, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 2}, 'mean': {'Q1-Dataset/pg1513.txt': 7, 'Q1-Dataset/2160-0.txt': 44, 'Q1-Dataset/pg11.txt': 10, 'Q1-Dataset/pg2600.txt': 63, 'Q1-Dataset/pg345.txt': 42, 'Q1-Dataset/pg64317.txt': 18, 'Q1-Dataset/pg70815.txt': 8, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 11}, 'very': {'Q1-Dataset/pg1513.txt': 26, 'Q1-Dataset/2160-0.txt': 358, 'Q1-Dataset/pg11.txt': 131, 'Q1-Dataset/pg2600.txt': 801, 'Q1-Dataset/pg345.txt': 278, 'Q1-Dataset/pg64317.txt': 61, 'Q1-Dataset/pg70815.txt': 74, 'Q1-Dataset/1080-0.txt': 11, 'Q1-Dataset/pg70817.txt': 32, 'Q1-Dataset/pg1998.txt': 79}, 'charges': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 7, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'i': {'Q1-Dataset/pg1513.txt': 580, 'Q1-Dataset/2160-0.txt': 2442, 'Q1-Dataset/pg11.txt': 385, 'Q1-Dataset/pg2600.txt': 4106, 'Q1-Dataset/pg345.txt': 4711, 'Q1-Dataset/pg64317.txt': 1179, 'Q1-Dataset/pg70815.txt': 528, 'Q1-Dataset/1080-0.txt': 55, 'Q1-Dataset/pg70817.txt': 364, 'Q1-Dataset/pg1998.txt': 1767}, 'just': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 42, 'Q1-Dataset/pg11.txt': 51, 'Q1-Dataset/pg2600.txt': 563, 'Q1-Dataset/pg345.txt': 164,

'Q1-Dataset/pg64317.txt': 100, 'Q1-Dataset/pg70815.txt': 57, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 26, 'Q1-Dataset/pg1998.txt': 99}, 'regulating': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 2}, 'last': {'Q1-Dataset/pg1513.txt': 8, 'Q1-Dataset/2160-0.txt': 99, 'Q1-Dataset/pg11.txt': 33, 'Q1-Dataset/pg2600.txt': 345, 'Q1-Dataset/pg345.txt': 136, 'Q1-Dataset/pg64317.txt': 45, 'Q1-Dataset/pg70815.txt': 43, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 24, 'Q1-Dataset/pg1998.txt': 161}, 'salt': {'Q1-Dataset/pg1513.txt': 3, 'Q1-Dataset/2160-0.txt': 10, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.txt': 12, 'Q1-Dataset/pg345.txt': 2, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 5}, 'alteration': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 3, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 2}, 'past': {'Q1-Dataset/pg1513.txt': 7, 'Q1-Dataset/2160-0.txt': 8, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 172, 'Q1-Dataset/pg345.txt': 29, 'Q1-Dataset/pg64317.txt': 14, 'Q1-Dataset/pg70815.txt': 10, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 16, 'Q1-Dataset/pg1998.txt': 32}, 'fight': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 77, 'Q1-Dataset/pg345.txt': 15, 'Q1-Dataset/pg64317.txt': 3, 'Q1-Dataset/pg70815.txt': 12, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 4, 'Q1-Dataset/pg1998.txt': 7}, 'after': {'Q1-Dataset/pg1513.txt': 4, 'Q1-Dataset/2160-0.txt': 107, 'Q1-Dataset/pg11.txt': 42, 'Q1-Dataset/pg2600.txt': 834, 'Q1-Dataset/pg345.txt': 206, 'Q1-Dataset/pg64317.txt': 88, 'Q1-Dataset/pg70815.txt': 79, 'Q1-Dataset/1080-0.txt': 5, 'Q1-Dataset/pg70817.txt': 50, 'Q1-Dataset/pg1998.txt': 77}, 'reuse': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 2, 'Q1-Dataset/pg345.txt': 2, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 2}, 'trouble': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 26, 'Q1-Dataset/pg11.txt': 6, 'Q1-Dataset/pg2600.txt': 33, 'Q1-Dataset/pg345.txt': 52, 'Q1-Dataset/pg64317.txt': 9, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 5}, 'confirmation': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 9, 'Q1-Dataset/pg345.txt': 3, 'Q1-Dataset/pg64317.txt': 3, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'appear': {'Q1-Dataset/pg1513.txt': 4, 'Q1-Dataset/2160-0.txt': 28, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.txt': 53, 'Q1-Dataset/pg345.txt': 8, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 4, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 4, 'Q1-Dataset/pg1998.txt': 11}, 'may': {'Q1-Dataset/pg1513.txt': 64, 'Q1-Dataset/2160-0.txt': 233, 'Q1-Dataset/pg11.txt': 28, 'Q1-Dataset/pg2600.txt': 289, 'Q1-Dataset/pg345.txt': 412, 'Q1-Dataset/pg64317.txt': 21, 'Q1-Dataset/pg70815.txt': 87, 'Q1-Dataset/1080-0.txt': 26, 'Q1-Dataset/pg70817.txt': 24, 'Q1-Dataset/pg1998.txt': 174}, 'comply': {'Q1-Dataset/pg1513.txt': 6, 'Q1-Dataset/2160-0.txt': 7, 'Q1-Dataset/pg11.txt': 6, 'Q1-Dataset/pg2600.txt': 9, 'Q1-Dataset/pg345.txt': 6, 'Q1-Dataset/pg64317.txt': 6, 'Q1-Dataset/pg70815.txt': 6, 'Q1-Dataset/1080-0.txt': 6, 'Q1-Dataset/pg70817.txt': 6, 'Q1-Dataset/pg1998.txt': 6}, 'depends': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 18, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 3}, 'measure': {'Q1-Dataset/pg1513.txt': 7, 'Q1-Dataset/2160-0.txt': 24, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 9, 'Q1-Dataset/pg345.txt': 11, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 3, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 3, 'Q1-Dataset/pg1998.txt': 8}, 'transcribe': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 2}, 'came': {'Q1-Dataset/pg1513.txt': 13, 'Q1-Dataset/2160-0.txt': 72, 'Q1-Dataset/pg11.txt': 40, 'Q1-Dataset/pg2600.txt': 681, 'Q1-Dataset/pg345.txt': 307, 'Q1-Dataset/pg64317.txt': 108, 'Q1-Dataset/pg70815.txt': 92, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 30, 'Q1-Dataset/pg1998.txt': 87}, 'pay': {'Q1-Dataset/pg1513.txt': 5, 'Q1-Dataset/2160-0.txt': 38, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 48, 'Q1-Dataset/pg345.txt': 8, 'Q1-Dataset/pg64317.txt': 5, 'Q1-Dataset/pg70815.txt': 3, 'Q1-Dataset/1080-0.txt': 4, 'Q1-Dataset/pg70817.txt': 3, 'Q1-Dataset/pg1998.txt': 3}, 'are': {'Q1-Dataset/pg1513.txt': 89, 'Q1-Dataset/2160-0.txt': 560, 'Q1-Dataset/pg11.txt': 72, 'Q1-Dataset/pg2600.txt': 1281, 'Q1-Dataset/pg345.txt': 604, 'Q1-Dataset/pg64317.txt': 69, 'Q1-Dataset/pg70815.txt': 296, 'Q1-Dataset/1080-0.txt': 45, 'Q1-Dataset/pg70817.txt': 33, 'Q1-Dataset/pg1998.txt': 50

3}, 'offers': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 7, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 2}, 'person': {'Q1-Dataset/pg1513.txt': 6, 'Q1-Dataset/2160-0.txt': 108, 'Q1-Dataset/pg11.txt': 8, 'Q1-Dataset/pg2600.txt': 75, 'Q1-Dataset/pg345.txt': 15, 'Q1-Dataset/pg64317.txt': 17, 'Q1-Dataset/pg70815.txt': 7, 'Q1-Dataset/1080-0.txt': 6, 'Q1-Dataset/pg70817.txt': 10, 'Q1-Dataset/pg1998.txt': 16}, 'removed': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 21, 'Q1-Dataset/pg11.txt': 4, 'Q1-Dataset/pg2600.txt': 37, 'Q1-Dataset/pg345.txt': 12, 'Q1-Dataset/pg64317.txt': 5, 'Q1-Dataset/pg70815.txt': 14, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 3, 'Q1-Dataset/pg1998.txt': 3}, 'make': {'Q1-Dataset/pg1513.txt': 43, 'Q1-Dataset/2160-0.txt': 189, 'Q1-Dataset/pg11.txt': 29, 'Q1-Dataset/pg2600.txt': 276, 'Q1-Dataset/pg345.txt': 165, 'Q1-Dataset/pg64317.txt': 30, 'Q1-Dataset/pg70815.txt': 43, 'Q1-Dataset/1080-0.txt': 11, 'Q1-Dataset/pg70817.txt': 14, 'Q1-Dataset/pg1998.txt': 76}, 'easily': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 24, 'Q1-Dataset/pg11.txt': 4, 'Q1-Dataset/pg2600.txt': 43, 'Q1-Dataset/pg345.txt': 11, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 15, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 5}, 'legal': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 4, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 8, 'Q1-Dataset/pg345.txt': 6, 'Q1-Dataset/pg64317.txt': 3, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 2}, 'good': {'Q1-Dataset/pg1513.txt': 80, 'Q1-Dataset/2160-0.txt': 237, 'Q1-Dataset/pg11.txt': 24, 'Q1-Dataset/pg2600.txt': 380, 'Q1-Dataset/pg345.txt': 220, 'Q1-Dataset/pg64317.txt': 30, 'Q1-Dataset/pg70815.txt': 40, 'Q1-Dataset/1080-0.txt': 8, 'Q1-Dataset/pg70817.txt': 15, 'Q1-Dataset/pg1998.txt': 284}, 'provide': {'Q1-Dataset/pg1513.txt': 7, 'Q1-Dataset/2160-0.txt': 19, 'Q1-Dataset/pg11.txt': 7, 'Q1-Dataset/pg2600.txt': 10, 'Q1-Dataset/pg345.txt': 9, 'Q1-Dataset/pg64317.txt': 7, 'Q1-Dataset/pg70815.txt': 7, 'Q1-Dataset/1080-0.txt': 9, 'Q1-Dataset/pg70817.txt': 8, 'Q1-Dataset/pg1998.txt': 8}, 'now': {'Q1-Dataset/pg1513.txt': 81, 'Q1-Dataset/2160-0.txt': 191, 'Q1-Dataset/pg11.txt': 57, 'Q1-Dataset/pg2600.txt': 1303, 'Q1-Dataset/pg345.txt': 401, 'Q1-Dataset/pg64317.txt': 97, 'Q1-Dataset/pg70815.txt': 146, 'Q1-Dataset/1080-0.txt': 6, 'Q1-Dataset/pg70817.txt': 45, 'Q1-Dataset/pg1998.txt': 302}, 'b': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 4, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 9, 'Q1-Dataset/pg345.txt': 3, 'Q1-Dataset/pg64317.txt': 6, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 5}, 'formats': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 2}, 'houses': {'Q1-Dataset/pg1513.txt': 5, 'Q1-Dataset/2160-0.txt': 22, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 61, 'Q1-Dataset/pg345.txt': 11, 'Q1-Dataset/pg64317.txt': 7, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 7, 'Q1-Dataset/pg1998.txt': 3}, 'paid': {'Q1-Dataset/pg1513.txt': 6, 'Q1-Dataset/2160-0.txt': 17, 'Q1-Dataset/pg11.txt': 6, 'Q1-Dataset/pg2600.txt': 41, 'Q1-Dataset/pg345.txt': 10, 'Q1-Dataset/pg64317.txt': 8, 'Q1-Dataset/pg70815.txt': 14, 'Q1-Dataset/1080-0.txt': 6, 'Q1-Dataset/pg70817.txt': 9, 'Q1-Dataset/pg1998.txt': 9}, 'world': {'Q1-Dataset/pg1513.txt': 17, 'Q1-Dataset/2160-0.txt': 41, 'Q1-Dataset/pg11.txt': 9, 'Q1-Dataset/pg2600.txt': 158, 'Q1-Dataset/pg345.txt': 63, 'Q1-Dataset/pg64317.txt': 29, 'Q1-Dataset/pg70815.txt': 92, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 4, 'Q1-Dataset/pg1998.txt': 132}, 'neither': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 42, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 96, 'Q1-Dataset/pg345.txt': 15, 'Q1-Dataset/pg64317.txt': 8, 'Q1-Dataset/pg70815.txt': 16, 'Q1-Dataset/1080-0.txt': 5, 'Q1-Dataset/pg70817.txt': 4, 'Q1-Dataset/pg1998.txt': 26}, 'check': {'Q1-Dataset/pg1513.txt': 4, 'Q1-Dataset/2160-0.txt': 6, 'Q1-Dataset/pg11.txt': 4, 'Q1-Dataset/pg2600.txt': 11, 'Q1-Dataset/pg345.txt': 8, 'Q1-Dataset/pg64317.txt': 5, 'Q1-Dataset/pg70815.txt': 5, 'Q1-Dataset/1080-0.txt': 4, 'Q1-Dataset/pg70817.txt': 5, 'Q1-Dataset/pg1998.txt': 4}, 'set': {'Q1-Dataset/pg1513.txt': 18, 'Q1-Dataset/2160-0.txt': 77, 'Q1-Dataset/pg11.txt': 22, 'Q1-Dataset/pg2600.txt': 160, 'Q1-Dataset/pg345.txt': 71, 'Q1-Dataset/pg64317.txt': 24, 'Q1-Dataset/pg70815.txt': 28, 'Q1-Dataset/1080-0.txt': 10, 'Q1-Dataset/pg70817.txt': 13, 'Q1-Dataset/pg1998.txt': 33}, 'matter': {'Q1-Dataset/pg1513.txt': 6, 'Q1-Dataset/2160-0.txt': 46, 'Q1-Dataset/pg11.txt': 9, 'Q1-Dataset/pg2600.txt': 167, 'Q1-Dataset/pg345.txt': 53, 'Q1-Dataset/pg64317.txt': 20, 'Q1-Dataset/pg70815.txt': 17, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 6, 'Q1-Dataset/pg1998.txt': 40}, 'remaining': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 23, 'Q1-Dataset/pg345.txt': 6, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 14, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 4, 'Q1-Dataset/pg1998.txt': 1}, 'public': {'Q1-Dataset/pg1513.txt': 8, 'Q1-Dataset/2160-0.txt': 61, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.txt': 27, 'Q1-Dataset/pg345.txt':

7, 'Q1-Dataset/pg64317.txt': 3, 'Q1-Dataset/pg70815.txt': 4, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 3, 'Q1-Dataset/pg1998.txt': 11}, 'different': {'Q1-Dataset/pg1513.txt': 3, 'Q1-Dataset/2160-0.txt': 69, 'Q1-Dataset/pg11.txt': 10, 'Q1-Dataset/pg2600.txt': 144, 'Q1-Dataset/pg345.txt': 23, 'Q1-Dataset/pg64317.txt': 4, 'Q1-Dataset/pg70815.txt': 8, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 4, 'Q1-Dataset/pg1998.txt': 19}, 'disclaimers': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'forty': {'Q1-Dataset/pg1513.txt': 3, 'Q1-Dataset/2160-0.txt': 14, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 31, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 4, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 3, 'Q1-Dataset/pg1998.txt': 4}, 'paragraph': {'Q1-Dataset/pg1513.txt': 11, 'Q1-Dataset/2160-0.txt': 12, 'Q1-Dataset/pg11.txt': 11, 'Q1-Dataset/pg2600.txt': 13, 'Q1-Dataset/pg345.txt': 12, 'Q1-Dataset/pg64317.txt': 11, 'Q1-Dataset/pg70815.txt': 11, 'Q1-Dataset/1080-0.txt': 11, 'Q1-Dataset/pg70817.txt': 12, 'Q1-Dataset/pg1998.txt': 23}, 'providing': {'Q1-Dataset/pg1513.txt': 4, 'Q1-Dataset/2160-0.txt': 6, 'Q1-Dataset/pg11.txt': 4, 'Q1-Dataset/pg2600.txt': 5, 'Q1-Dataset/pg345.txt': 4, 'Q1-Dataset/pg64317.txt': 4, 'Q1-Dataset/pg70815.txt': 4, 'Q1-Dataset/1080-0.txt': 5, 'Q1-Dataset/pg70817.txt': 5, 'Q1-Dataset/pg1998.txt': 4}, 'upon': {'Q1-Dataset/pg1513.txt': 44, 'Q1-Dataset/2160-0.txt': 463, 'Q1-Dataset/pg11.txt': 28, 'Q1-Dataset/pg2600.txt': 97, 'Q1-Dataset/pg345.txt': 90, 'Q1-Dataset/pg64317.txt': 35, 'Q1-Dataset/pg70815.txt': 482, 'Q1-Dataset/1080-0.txt': 15, 'Q1-Dataset/pg70817.txt': 11, 'Q1-Dataset/pg1998.txt': 144}, 'discover': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 5, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 9, 'Q1-Dataset/pg345.txt': 5, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 11, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 5}, 'directly': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 10, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.txt': 36, 'Q1-Dataset/pg345.txt': 4, 'Q1-Dataset/pg64317.txt': 3, 'Q1-Dataset/pg70815.txt': 28, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 5, 'Q1-Dataset/pg1998.txt': 2}, 'require': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 7, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 6, 'Q1-Dataset/pg345.txt': 4, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 4, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 9}, 'receiving': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 4, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 40, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 4}, 'granted': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 9, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 10, 'Q1-Dataset/pg345.txt': 3, 'Q1-Dataset/pg64317.txt': 3, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 2}, 'obtaining': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 7, 'Q1-Dataset/pg345.txt': 2, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 3, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 2}, 'distributing': {'Q1-Dataset/pg1513.txt': 7, 'Q1-Dataset/2160-0.txt': 8, 'Q1-Dataset/pg11.txt': 7, 'Q1-Dataset/pg2600.txt': 9, 'Q1-Dataset/pg345.txt': 7, 'Q1-Dataset/pg64317.txt': 7, 'Q1-Dataset/pg70815.txt': 7, 'Q1-Dataset/1080-0.txt': 7, 'Q1-Dataset/pg70817.txt': 7, 'Q1-Dataset/pg1998.txt': 8}, 'license': {'Q1-Dataset/pg1513.txt': 17, 'Q1-Dataset/2160-0.txt': 15, 'Q1-Dataset/pg11.txt': 17, 'Q1-Dataset/pg2600.txt': 17, 'Q1-Dataset/pg345.txt': 17, 'Q1-Dataset/pg64317.txt': 17, 'Q1-Dataset/pg70815.txt': 17, 'Q1-Dataset/1080-0.txt': 17, 'Q1-Dataset/pg70817.txt': 17, 'Q1-Dataset/pg1998.txt': 17}, 'donations': {'Q1-Dataset/pg1513.txt': 15, 'Q1-Dataset/2160-0.txt': 15, 'Q1-Dataset/pg11.txt': 15, 'Q1-Dataset/pg2600.txt': 15, 'Q1-Dataset/pg345.txt': 15, 'Q1-Dataset/pg64317.txt': 15, 'Q1-Dataset/pg70815.txt': 15, 'Q1-Dataset/1080-0.txt': 15, 'Q1-Dataset/pg70817.txt': 15, 'Q1-Dataset/pg1998.txt': 15}, 'number': {'Q1-Dataset/pg1513.txt': 5, 'Q1-Dataset/2160-0.txt': 55, 'Q1-Dataset/pg11.txt': 8, 'Q1-Dataset/pg2600.txt': 78, 'Q1-Dataset/pg345.txt': 21, 'Q1-Dataset/pg64317.txt': 7, 'Q1-Dataset/pg70815.txt': 17, 'Q1-Dataset/1080-0.txt': 14, 'Q1-Dataset/pg70817.txt': 5, 'Q1-Dataset/pg1998.txt': 12}, 'off': {'Q1-Dataset/pg1513.txt': 10, 'Q1-Dataset/2160-0.txt': 73, 'Q1-Dataset/pg11.txt': 73, 'Q1-Dataset/pg2600.txt': 426, 'Q1-Dataset/pg345.txt': 149, 'Q1-Dataset/pg64317.txt': 45, 'Q1-Dataset/pg70815.txt': 38, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 16, 'Q1-Dataset/pg1998.txt': 16}, 'organized': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 4, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 3, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'office': {'Q1-Dataset/pg1513.txt': 3, 'Q1-Dataset/2160-0.txt': 15, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 9, 'Q1-Dataset/pg345.txt': 11, 'Q1-Dataset/pg64317.txt': 17, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 5, 'Q1-Dataset/pg1998.txt': 1}

aset/pg1998.txt': 1}, 'one': {'Q1-Dataset/pg1513.txt': 60, 'Q1-Dataset/2160-0.txt': 343, 'Q1-Dataset/pg11.txt': 99, 'Q1-Dataset/pg2600.txt': 2048, 'Q1-Dataset/pg345.txt': 490, 'Q1-Dataset/pg64317.txt': 146, 'Q1-Dataset/pg70815.txt': 232, 'Q1-Dataset/1080-0.txt': 16, 'Q1-Dataset/pg70817.txt': 69, 'Q1-Dataset/pg1998.txt': 711}, 'release': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 11, 'Q1-Dataset/pg345.txt': 3, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 3, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 1}, 'learn': {'Q1-Dataset/pg1513.txt': 4, 'Q1-Dataset/2160-0.txt': 6, 'Q1-Dataset/pg11.txt': 8, 'Q1-Dataset/pg2600.txt': 17, 'Q1-Dataset/pg345.txt': 36, 'Q1-Dataset/pg64317.txt': 3, 'Q1-Dataset/pg70815.txt': 6, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 82}, 'calculated': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 3, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 6, 'Q1-Dataset/pg345.txt': 3, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 1}, 'creation': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 5, 'Q1-Dataset/pg345.txt': 3, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 4, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 5}, 'their': {'Q1-Dataset/pg1513.txt': 47, 'Q1-Dataset/2160-0.txt': 475, 'Q1-Dataset/pg11.txt': 51, 'Q1-Dataset/pg2600.txt': 1437, 'Q1-Dataset/pg345.txt': 218, 'Q1-Dataset/pg64317.txt': 55, 'Q1-Dataset/pg70815.txt': 376, 'Q1-Dataset/1080-0.txt': 38, 'Q1-Dataset/pg70817.txt': 37, 'Q1-Dataset/pg1998.txt': 424}, 'written': {'Q1-Dataset/pg1513.txt': 8, 'Q1-Dataset/2160-0.txt': 17, 'Q1-Dataset/pg11.txt': 9, 'Q1-Dataset/pg2600.txt': 65, 'Q1-Dataset/pg345.txt': 32, 'Q1-Dataset/pg64317.txt': 4, 'Q1-Dataset/pg70815.txt': 4, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 6, 'Q1-Dataset/pg1998.txt': 21}, 'warranties': {'Q1-Dataset/pg1513.txt': 3, 'Q1-Dataset/2160-0.txt': 3, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.txt': 3, 'Q1-Dataset/pg345.txt': 3, 'Q1-Dataset/pg64317.txt': 3, 'Q1-Dataset/pg70815.txt': 3, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 3, 'Q1-Dataset/pg1998.txt': 3}, 'foundation': {'Q1-Dataset/pg1513.txt': 22, 'Q1-Dataset/2160-0.txt': 29, 'Q1-Dataset/pg11.txt': 22, 'Q1-Dataset/pg2600.txt': 26, 'Q1-Dataset/pg345.txt': 22, 'Q1-Dataset/pg64317.txt': 22, 'Q1-Dataset/pg70815.txt': 22, 'Q1-Dataset/1080-0.txt': 22, 'Q1-Dataset/pg70817.txt': 22, 'Q1-Dataset/pg1998.txt': 22}, 'place': {'Q1-Dataset/pg1513.txt': 17, 'Q1-Dataset/2160-0.txt': 147, 'Q1-Dataset/pg11.txt': 9, 'Q1-Dataset/pg2600.txt': 310, 'Q1-Dataset/pg345.txt': 133, 'Q1-Dataset/pg64317.txt': 21, 'Q1-Dataset/pg70815.txt': 29, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 14, 'Q1-Dataset/pg1998.txt': 35}, 'second': {'Q1-Dataset/pg1513.txt': 18, 'Q1-Dataset/2160-0.txt': 26, 'Q1-Dataset/pg11.txt': 6, 'Q1-Dataset/pg2600.txt': 118, 'Q1-Dataset/pg345.txt': 20, 'Q1-Dataset/pg64317.txt': 11, 'Q1-Dataset/pg70815.txt': 9, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 6, 'Q1-Dataset/pg1998.txt': 28}, 'hundred': {'Q1-Dataset/pg1513.txt': 3, 'Q1-Dataset/2160-0.txt': 57, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 122, 'Q1-Dataset/pg345.txt': 12, 'Q1-Dataset/pg64317.txt': 14, 'Q1-Dataset/pg70815.txt': 33, 'Q1-Dataset/1080-0.txt': 8, 'Q1-Dataset/pg70817.txt': 7, 'Q1-Dataset/pg1998.txt': 13}, 'on': {'Q1-Dataset/pg1513.txt': 92, 'Q1-Dataset/2160-0.txt': 527, 'Q1-Dataset/pg11.txt': 204, 'Q1-Dataset/pg2600.txt': 3996, 'Q1-Dataset/pg345.txt': 1081, 'Q1-Dataset/pg64317.txt': 371, 'Q1-Dataset/pg70815.txt': 150, 'Q1-Dataset/1080-0.txt': 19, 'Q1-Dataset/pg70817.txt': 197, 'Q1-Dataset/pg1998.txt': 553}, 'located': {'Q1-Dataset/pg1513.txt': 7, 'Q1-Dataset/2160-0.txt': 4, 'Q1-Dataset/pg11.txt': 7, 'Q1-Dataset/pg2600.txt': 8, 'Q1-Dataset/pg345.txt': 7, 'Q1-Dataset/pg64317.txt': 7, 'Q1-Dataset/pg70815.txt': 12, 'Q1-Dataset/1080-0.txt': 7, 'Q1-Dataset/pg70817.txt': 13, 'Q1-Dataset/pg1998.txt': 7}, 'generations': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 4, 'Q1-Dataset/pg345.txt': 3, 'Q1-Dataset/pg64317.txt': 3, 'Q1-Dataset/pg70815.txt': 3, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 3}, 'purpose': {'Q1-Dataset/pg1513.txt': 5, 'Q1-Dataset/2160-0.txt': 37, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.txt': 50, 'Q1-Dataset/pg345.txt': 36, 'Q1-Dataset/pg64317.txt': 3, 'Q1-Dataset/pg70815.txt': 17, 'Q1-Dataset/1080-0.txt': 4, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 26}, 'elect': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'performing': {'Q1-Dataset/pg1513.txt': 3, 'Q1-Dataset/2160-0.txt': 6, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.txt': 8, 'Q1-Dataset/pg345.txt': 3, 'Q1-Dataset/pg64317.txt': 3, 'Q1-Dataset/pg70815.txt': 3, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 3, 'Q1-Dataset/pg1998.txt': 3}, 'word': {'Q1-Dataset/pg1513.txt': 28, 'Q1-Dataset/2160-0.txt': 42, 'Q1-Dataset/pg11.txt': 11, 'Q1-Dataset/pg2600.txt': 233, 'Q1-Dataset/pg345.txt': 65, 'Q1-Dataset/pg64317.txt': 23, 'Q1-Dataset/pg70815.txt': 15, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 6, 'Q1-Dataset/pg1998.txt': 52}, 'all': {'Q1-Dataset/pg1513.txt': 112, 'Q1-Dataset/2160-0.txt': 697, 'Q1-Dataset/pg11.txt': 19

5, 'Q1-Dataset/pg2600.txt': 2771, 'Q1-Dataset/pg345.txt': 1164, 'Q1-Dataset/pg4317.txt': 233, 'Q1-Dataset/pg70815.txt': 186, 'Q1-Dataset/1080-0.txt': 27, 'Q1-Dataset/pg70817.txt': 86, 'Q1-Dataset/pg1998.txt': 884}, 'secure': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 6, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 16, 'Q1-Dataset/pg345.txt': 5, 'Q1-Dataset/pg64317.txt': 3, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 2}, 'whenever': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 6, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 13, 'Q1-Dataset/pg345.txt': 8, 'Q1-Dataset/pg64317.txt': 8, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 4}, 'has': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 399, 'Q1-Dataset/pg11.txt': 9, 'Q1-Dataset/pg2600.txt': 704, 'Q1-Dataset/pg345.txt': 347, 'Q1-Dataset/pg64317.txt': 21, 'Q1-Dataset/pg70815.txt': 68, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 11, 'Q1-Dataset/pg1998.txt': 47}, 'does': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 32, 'Q1-Dataset/pg11.txt': 11, 'Q1-Dataset/pg2600.txt': 208, 'Q1-Dataset/pg345.txt': 74, 'Q1-Dataset/pg64317.txt': 12, 'Q1-Dataset/pg70815.txt': 17, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 4, 'Q1-Dataset/pg1998.txt': 19}, 'agree': {'Q1-Dataset/pg1513.txt': 10, 'Q1-Dataset/2160-0.txt': 11, 'Q1-Dataset/pg11.txt': 11, 'Q1-Dataset/pg2600.txt': 49, 'Q1-Dataset/pg345.txt': 14, 'Q1-Dataset/pg64317.txt': 9, 'Q1-Dataset/pg70815.txt': 14, 'Q1-Dataset/1080-0.txt': 9, 'Q1-Dataset/pg70817.txt': 10, 'Q1-Dataset/pg1998.txt': 9}, 'table': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 34, 'Q1-Dataset/pg11.txt': 18, 'Q1-Dataset/pg2600.txt': 222, 'Q1-Dataset/pg345.txt': 32, 'Q1-Dataset/pg64317.txt': 31, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 6, 'Q1-Dataset/pg1998.txt': 19}, 'if': {'Q1-Dataset/pg1513.txt': 106, 'Q1-Dataset/2160-0.txt': 406, 'Q1-Dataset/pg11.txt': 116, 'Q1-Dataset/pg2600.txt': 1303, 'Q1-Dataset/pg345.txt': 649, 'Q1-Dataset/pg64317.txt': 165, 'Q1-Dataset/pg70815.txt': 209, 'Q1-Dataset/1080-0.txt': 27, 'Q1-Dataset/pg70817.txt': 65, 'Q1-Dataset/pg1998.txt': 240}, 'makes': {'Q1-Dataset/pg1513.txt': 13, 'Q1-Dataset/2160-0.txt': 19, 'Q1-Dataset/pg11.txt': 12, 'Q1-Dataset/pg2600.txt': 27, 'Q1-Dataset/pg345.txt': 20, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 4, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 6}, 'hold': {'Q1-Dataset/pg1513.txt': 18, 'Q1-Dataset/2160-0.txt': 33, 'Q1-Dataset/pg11.txt': 10, 'Q1-Dataset/pg2600.txt': 64, 'Q1-Dataset/pg345.txt': 32, 'Q1-Dataset/pg64317.txt': 6, 'Q1-Dataset/pg70815.txt': 23, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 5, 'Q1-Dataset/pg1998.txt': 18}, 'met': {'Q1-Dataset/pg1513.txt': 5, 'Q1-Dataset/2160-0.txt': 39, 'Q1-Dataset/pg11.txt': 4, 'Q1-Dataset/pg2600.txt': 155, 'Q1-Dataset/pg345.txt': 30, 'Q1-Dataset/pg64317.txt': 21, 'Q1-Dataset/pg70815.txt': 10, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 5, 'Q1-Dataset/pg1998.txt': 8}, 'distributed': {'Q1-Dataset/pg1513.txt': 4, 'Q1-Dataset/2160-0.txt': 9, 'Q1-Dataset/pg11.txt': 4, 'Q1-Dataset/pg2600.txt': 13, 'Q1-Dataset/pg345.txt': 6, 'Q1-Dataset/pg64317.txt': 4, 'Q1-Dataset/pg70815.txt': 5, 'Q1-Dataset/1080-0.txt': 4, 'Q1-Dataset/pg70817.txt': 5, 'Q1-Dataset/pg1998.txt': 4}, 'nothing': {'Q1-Dataset/pg1513.txt': 14, 'Q1-Dataset/2160-0.txt': 134, 'Q1-Dataset/pg11.txt': 34, 'Q1-Dataset/pg2600.txt': 461, 'Q1-Dataset/pg345.txt': 112, 'Q1-Dataset/pg64317.txt': 32, 'Q1-Dataset/pg70815.txt': 42, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 21, 'Q1-Dataset/pg1998.txt': 64}, 'express': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 12, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 62, 'Q1-Dataset/pg345.txt': 4, 'Q1-Dataset/pg64317.txt': 3, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 5}, 'takes': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 11, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.txt': 25, 'Q1-Dataset/pg345.txt': 8, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 4}, 'come': {'Q1-Dataset/pg1513.txt': 96, 'Q1-Dataset/2160-0.txt': 75, 'Q1-Dataset/pg11.txt': 46, 'Q1-Dataset/pg2600.txt': 679, 'Q1-Dataset/pg345.txt': 338, 'Q1-Dataset/pg64317.txt': 85, 'Q1-Dataset/pg70815.txt': 112, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 11, 'Q1-Dataset/pg1998.txt': 139}, 'lot': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 21, 'Q1-Dataset/pg345.txt': 17, 'Q1-Dataset/pg64317.txt': 10, 'Q1-Dataset/pg70815.txt': 3, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 8, 'Q1-Dataset/pg1998.txt': 7}, 'fitness': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 2, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 3}, 'cost': {'Q1-Dataset/pg1513.txt': 4, 'Q1-Dataset/2160-0.txt': 12, 'Q1-Dataset/pg11.txt': 4, 'Q1-Dataset/pg2600.txt': 26, 'Q1-Dataset/pg345.txt': 3, 'Q1-Dataset/pg64317.txt': 4, 'Q1-Dataset/pg70815.txt': 3, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 4, 'Q1-Dataset/pg1998.txt': 4}, 'protect': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 7, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 10, 'Q1-Dataset/pg345.txt': 13, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 9, 'Q1-Dataset/

1080-0.txt': 2, 'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/pg1998.txt': 5}, 'newletter': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'status': {'Q1-Dataset/pg1513.txt': 4, 'Q1-Dataset/2160-0.txt': 4, 'Q1-Dataset/pg11.txt': 4, 'Q1-Dataset/pg2600.txt': 6, 'Q1-Dataset/pg345.txt': 4, 'Q1-Dataset/pg64317.txt': 4, 'Q1-Dataset/pg70815.txt': 4, 'Q1-Dataset/1080-0.txt': 4, 'Q1-Dataset/pg70817.txt': 4, 'Q1-Dataset/pg1998.txt': 4}, 'invalidity': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 2, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'less': {'Q1-Dataset/pg1513.txt': 6, 'Q1-Dataset/2160-0.txt': 46, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.txt': 142, 'Q1-Dataset/pg345.txt': 31, 'Q1-Dataset/pg64317.txt': 12, 'Q1-Dataset/pg70815.txt': 20, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 6, 'Q1-Dataset/pg1998.txt': 15}, 'bent': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 62, 'Q1-Dataset/pg345.txt': 26, 'Q1-Dataset/pg64317.txt': 4, 'Q1-Dataset/pg70815.txt': 3, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 5, 'Q1-Dataset/pg1998.txt': 3}, 'answer': {'Q1-Dataset/pg1513.txt': 10, 'Q1-Dataset/2160-0.txt': 21, 'Q1-Dataset/pg11.txt': 9, 'Q1-Dataset/pg2600.txt': 142, 'Q1-Dataset/pg345.txt': 44, 'Q1-Dataset/pg64317.txt': 9, 'Q1-Dataset/pg70815.txt': 8, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 3, 'Q1-Dataset/pg1998.txt': 14}, 'nearly': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 9, 'Q1-Dataset/pg11.txt': 13, 'Q1-Dataset/pg2600.txt': 57, 'Q1-Dataset/pg345.txt': 22, 'Q1-Dataset/pg64317.txt': 4, 'Q1-Dataset/pg70815.txt': 3, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 18, 'Q1-Dataset/pg1998.txt': 6}, 'among': {'Q1-Dataset/pg1513.txt': 6, 'Q1-Dataset/2160-0.txt': 93, 'Q1-Dataset/pg11.txt': 12, 'Q1-Dataset/pg2600.txt': 216, 'Q1-Dataset/pg345.txt': 8, 'Q1-Dataset/pg64317.txt': 21, 'Q1-Dataset/pg70815.txt': 57, 'Q1-Dataset/1080-0.txt': 6, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 69}, 'choose': {'Q1-Dataset/pg1513.txt': 4, 'Q1-Dataset/2160-0.txt': 9, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 26, 'Q1-Dataset/pg345.txt': 8, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 4, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 6}, 'because': {'Q1-Dataset/pg1513.txt': 11, 'Q1-Dataset/2160-0.txt': 43, 'Q1-Dataset/pg11.txt': 16, 'Q1-Dataset/pg2600.txt': 487, 'Q1-Dataset/pg345.txt': 55, 'Q1-Dataset/pg64317.txt': 46, 'Q1-Dataset/pg70815.txt': 43, 'Q1-Dataset/1080-0.txt': 4, 'Q1-Dataset/pg70817.txt': 11, 'Q1-Dataset/pg1998.txt': 97}, 'west': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 9, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 28, 'Q1-Dataset/pg345.txt': 12, 'Q1-Dataset/pg64317.txt': 38, 'Q1-Dataset/pg70815.txt': 8, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'disclaim': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 2}, 'while': {'Q1-Dataset/pg1513.txt': 8, 'Q1-Dataset/2160-0.txt': 94, 'Q1-Dataset/pg11.txt': 25, 'Q1-Dataset/pg2600.txt': 397, 'Q1-Dataset/pg345.txt': 49, 'Q1-Dataset/pg64317.txt': 48, 'Q1-Dataset/pg70815.txt': 126, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 15, 'Q1-Dataset/pg1998.txt': 41}, 'literary': {'Q1-Dataset/pg1513.txt': 13, 'Q1-Dataset/2160-0.txt': 17, 'Q1-Dataset/pg11.txt': 13, 'Q1-Dataset/pg2600.txt': 13, 'Q1-Dataset/pg345.txt': 13, 'Q1-Dataset/pg64317.txt': 14, 'Q1-Dataset/pg70815.txt': 13, 'Q1-Dataset/1080-0.txt': 13, 'Q1-Dataset/pg70817.txt': 13, 'Q1-Dataset/pg1998.txt': 13}, 'virus': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 3, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'to': {'Q1-Dataset/pg1513.txt': 622, 'Q1-Dataset/2160-0.txt': 4412, 'Q1-Dataset/pg11.txt': 803, 'Q1-Dataset/pg2600.txt': 16713, 'Q1-Dataset/pg345.txt': 4543, 'Q1-Dataset/pg64317.txt': 1209, 'Q1-Dataset/pg70815.txt': 2032, 'Q1-Dataset/1080-0.txt': 190, 'Q1-Dataset/pg70817.txt': 578, 'Q1-Dataset/pg1998.txt': 2681}, 'credit': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 18, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 9, 'Q1-Dataset/pg345.txt': 4, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 3, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 2}, 'grow': {'Q1-Dataset/pg1513.txt': 3, 'Q1-Dataset/2160-0.txt': 14, 'Q1-Dataset/pg11.txt': 13, 'Q1-Dataset/pg2600.txt': 28, 'Q1-Dataset/pg345.txt': 21, 'Q1-Dataset/pg64317.txt': 4, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 5}, 'effort': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 13, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 78, 'Q1-Dataset/pg345.txt': 38, 'Q1-Dataset/pg64317.txt': 15, 'Q1-Dataset/pg70815.txt': 32, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 5, 'Q1-Dataset/pg1998.txt': 3}, 'problem': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1,

'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 20, 'Q1-Dataset/pg345.txt': 4, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 3, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 3, 'Q1-Dataset/pg1998.txt': 1}, 'want': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 54, 'Q1-Dataset/pg11.txt': 9, 'Q1-Dataset/pg2600.txt': 249, 'Q1-Dataset/pg345.txt': 114, 'Q1-Dataset/pg64317.txt': 59, 'Q1-Dataset/pg70815.txt': 17, 'Q1-Dataset/1080-0.txt': 5, 'Q1-Dataset/pg70817.txt': 6, 'Q1-Dataset/pg1998.txt': 86}, 'readable': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'such': {'Q1-Dataset/pg1513.txt': 52, 'Q1-Dataset/2160-0.txt': 397, 'Q1-Dataset/pg11.txt': 47, 'Q1-Dataset/pg2600.txt': 590, 'Q1-Dataset/pg345.txt': 229, 'Q1-Dataset/pg64317.txt': 25, 'Q1-Dataset/pg70815.txt': 94, 'Q1-Dataset/1080-0.txt': 9, 'Q1-Dataset/pg70817.txt': 16, 'Q1-Dataset/pg1998.txt': 178}, 'viewed': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 3, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 2, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'people': {'Q1-Dataset/pg1513.txt': 4, 'Q1-Dataset/2160-0.txt': 106, 'Q1-Dataset/pg11.txt': 16, 'Q1-Dataset/pg2600.txt': 524, 'Q1-Dataset/pg345.txt': 46, 'Q1-Dataset/pg64317.txt': 63, 'Q1-Dataset/pg70815.txt': 60, 'Q1-Dataset/1080-0.txt': 8, 'Q1-Dataset/pg70817.txt': 12, 'Q1-Dataset/pg1998.txt': 132}, 'should': {'Q1-Dataset/pg1513.txt': 51, 'Q1-Dataset/2160-0.txt': 259, 'Q1-Dataset/pg11.txt': 29, 'Q1-Dataset/pg2600.txt': 519, 'Q1-Dataset/pg345.txt': 197, 'Q1-Dataset/pg64317.txt': 11, 'Q1-Dataset/pg70815.txt': 59, 'Q1-Dataset/1080-0.txt': 4, 'Q1-Dataset/pg70817.txt': 6, 'Q1-Dataset/pg1998.txt': 136}, 'lake': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 9, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 3, 'Q1-Dataset/pg64317.txt': 5, 'Q1-Dataset/pg70815.txt': 11, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 7}, 'alive': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.txt': 45, 'Q1-Dataset/pg345.txt': 12, 'Q1-Dataset/pg64317.txt': 5, 'Q1-Dataset/pg70815.txt': 13, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 4, 'Q1-Dataset/pg1998.txt': 4}, 'plain': {'Q1-Dataset/pg1513.txt': 3, 'Q1-Dataset/2160-0.txt': 23, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 67, 'Q1-Dataset/pg345.txt': 6, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 29, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 3, 'Q1-Dataset/pg1998.txt': 5}, 'see': {'Q1-Dataset/pg1513.txt': 39, 'Q1-Dataset/2160-0.txt': 112, 'Q1-Dataset/pg11.txt': 69, 'Q1-Dataset/pg2600.txt': 730, 'Q1-Dataset/pg345.txt': 394, 'Q1-Dataset/pg64317.txt': 84, 'Q1-Dataset/pg70815.txt': 69, 'Q1-Dataset/1080-0.txt': 5, 'Q1-Dataset/pg70817.txt': 30, 'Q1-Dataset/pg1998.txt': 151}, 'me': {'Q1-Dataset/pg1513.txt': 265, 'Q1-Dataset/2160-0.txt': 510, 'Q1-Dataset/pg11.txt': 64, 'Q1-Dataset/pg2600.txt': 1263, 'Q1-Dataset/pg345.txt': 1444, 'Q1-Dataset/pg64317.txt': 325, 'Q1-Dataset/pg70815.txt': 105, 'Q1-Dataset/1080-0.txt': 5, 'Q1-Dataset/pg70817.txt': 61, 'Q1-Dataset/pg1998.txt': 783}, 'associated': {'Q1-Dataset/pg1513.txt': 7, 'Q1-Dataset/2160-0.txt': 8, 'Q1-Dataset/pg11.txt': 8, 'Q1-Dataset/pg2600.txt': 14, 'Q1-Dataset/pg345.txt': 7, 'Q1-Dataset/pg64317.txt': 9, 'Q1-Dataset/pg70815.txt': 8, 'Q1-Dataset/1080-0.txt': 8, 'Q1-Dataset/pg70817.txt': 7, 'Q1-Dataset/pg1998.txt': 9}, 'reaching': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 5, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 32, 'Q1-Dataset/pg345.txt': 3, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 3, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 4, 'Q1-Dataset/pg1998.txt': 1}, 'staff': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 135, 'Q1-Dataset/pg345.txt': 2, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 6}, 'when': {'Q1-Dataset/pg1513.txt': 54, 'Q1-Dataset/2160-0.txt': 327, 'Q1-Dataset/pg11.txt': 80, 'Q1-Dataset/pg2600.txt': 1486, 'Q1-Dataset/pg345.txt': 771, 'Q1-Dataset/pg64317.txt': 144, 'Q1-Dataset/pg70815.txt': 220, 'Q1-Dataset/1080-0.txt': 8, 'Q1-Dataset/pg70817.txt': 78, 'Q1-Dataset/pg1998.txt': 393}, 'proofread': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'opportunities': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 5, 'Q1-Dataset/pg345.txt': 5, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 2}, 'revenue': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 4, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 2, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'owner': {'Q1-Dataset/pg1513.txt': 4, 'Q1-Dataset/2160-0.txt': 6, 'Q1-Dataset/pg11.txt': 4, 'Q1-Dataset/pg2600.txt': 10, 'Q1-Dataset/pg345.txt': 10, 'Q1-Dataset/pg64317.txt': 4, 'Q1-Dataset/pg70815.txt': 10, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}

xt': 4, 'Q1-Dataset/1080-0.txt': 4, 'Q1-Dataset/pg10817.txt': 5, 'Q1-Dataset/pg1998.txt': 4}, 'complying': {'Q1-Dataset/pg1513.txt': 3, 'Q1-Dataset/2160-0.txt': 3, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.txt': 4, 'Q1-Dataset/pg345.txt': 3, 'Q1-Dataset/pg64317.txt': 3, 'Q1-Dataset/pg70815.txt': 3, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 3, 'Q1-Dataset/pg1998.txt': 3}, 'copyright': {'Q1-Dataset/pg1513.txt': 19, 'Q1-Dataset/2160-0.txt': 13, 'Q1-Dataset/pg11.txt': 19, 'Q1-Dataset/pg2600.txt': 19, 'Q1-Dataset/pg345.txt': 20, 'Q1-Dataset/pg64317.txt': 19, 'Q1-Dataset/pg70815.txt': 19, 'Q1-Dataset/1080-0.txt': 19, 'Q1-Dataset/pg70817.txt': 21, 'Q1-Dataset/pg1998.txt': 19}, 'sure': {'Q1-Dataset/pg1513.txt': 5, 'Q1-Dataset/2160-0.txt': 63, 'Q1-Dataset/pg11.txt': 24, 'Q1-Dataset/pg2600.txt': 68, 'Q1-Dataset/pg345.txt': 47, 'Q1-Dataset/pg64317.txt': 17, 'Q1-Dataset/pg70815.txt': 21, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 18, 'Q1-Dataset/pg1998.txt': 24}, 'twelve': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 10, 'Q1-Dataset/pg11.txt': 4, 'Q1-Dataset/pg2600.txt': 14, 'Q1-Dataset/pg345.txt': 6, 'Q1-Dataset/pg64317.txt': 5, 'Q1-Dataset/pg70815.txt': 3, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 2}, 'yet': {'Q1-Dataset/pg1513.txt': 47, 'Q1-Dataset/2160-0.txt': 81, 'Q1-Dataset/pg11.txt': 21, 'Q1-Dataset/pg2600.txt': 339, 'Q1-Dataset/pg345.txt': 152, 'Q1-Dataset/pg64317.txt': 9, 'Q1-Dataset/pg70815.txt': 75, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 20, 'Q1-Dataset/pg1998.txt': 129}, 'michael': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 52, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'errors': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 3, 'Q1-Dataset/pg345.txt': 3, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 4}, 'creating': {'Q1-Dataset/pg1513.txt': 4, 'Q1-Dataset/2160-0.txt': 4, 'Q1-Dataset/pg11.txt': 4, 'Q1-Dataset/pg2600.txt': 4, 'Q1-Dataset/pg345.txt': 4, 'Q1-Dataset/pg64317.txt': 4, 'Q1-Dataset/pg70815.txt': 4, 'Q1-Dataset/1080-0.txt': 4, 'Q1-Dataset/pg70817.txt': 4, 'Q1-Dataset/pg1998.txt': 43}, 'copying': {'Q1-Dataset/pg1513.txt': 4, 'Q1-Dataset/2160-0.txt': 4, 'Q1-Dataset/pg11.txt': 4, 'Q1-Dataset/pg2600.txt': 7, 'Q1-Dataset/pg345.txt': 4, 'Q1-Dataset/pg64317.txt': 4, 'Q1-Dataset/pg70815.txt': 4, 'Q1-Dataset/1080-0.txt': 4, 'Q1-Dataset/pg70817.txt': 4, 'Q1-Dataset/pg1998.txt': 4}, 'figure': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 23, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.txt': 69, 'Q1-Dataset/pg345.txt': 22, 'Q1-Dataset/pg64317.txt': 6, 'Q1-Dataset/pg70815.txt': 6, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 4, 'Q1-Dataset/pg1998.txt': 9}, 'ascii': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 2, 'Q1-Dataset/pg345.txt': 2, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 2}, 'liability': {'Q1-Dataset/pg1513.txt': 3, 'Q1-Dataset/2160-0.txt': 3, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.txt': 3, 'Q1-Dataset/pg345.txt': 3, 'Q1-Dataset/pg64317.txt': 3, 'Q1-Dataset/pg70815.txt': 3, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 3, 'Q1-Dataset/pg1998.txt': 3}, 'work': {'Q1-Dataset/pg1513.txt': 49, 'Q1-Dataset/2160-0.txt': 61, 'Q1-Dataset/pg11.txt': 53, 'Q1-Dataset/pg2600.txt': 166, 'Q1-Dataset/pg345.txt': 183, 'Q1-Dataset/pg64317.txt': 56, 'Q1-Dataset/pg70815.txt': 56, 'Q1-Dataset/1080-0.txt': 51, 'Q1-Dataset/pg70817.txt': 54, 'Q1-Dataset/pg1998.txt': 77}, 'he': {'Q1-Dataset/pg1513.txt': 122, 'Q1-Dataset/2160-0.txt': 2012, 'Q1-Dataset/pg11.txt': 119, 'Q1-Dataset/pg2600.txt': 9811, 'Q1-Dataset/pg345.txt': 2541, 'Q1-Dataset/pg64317.txt': 797, 'Q1-Dataset/pg70815.txt': 1472, 'Q1-Dataset/1080-0.txt': 8, 'Q1-Dataset/pg70817.txt': 551, 'Q1-Dataset/pg1998.txt': 1077}, 'library': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 3, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 6, 'Q1-Dataset/pg345.txt': 9, 'Q1-Dataset/pg64317.txt': 11, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 5, 'Q1-Dataset/pg1998.txt': 1}, 'calculate': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 2}, 'beat': {'Q1-Dataset/pg1513.txt': 4, 'Q1-Dataset/2160-0.txt': 8, 'Q1-Dataset/pg11.txt': 4, 'Q1-Dataset/pg2600.txt': 33, 'Q1-Dataset/pg345.txt': 17, 'Q1-Dataset/pg64317.txt': 9, 'Q1-Dataset/pg70815.txt': 7, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 4}, 'asis': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'home': {'Q1-Dataset/pg1513.txt': 6, 'Q1-Dataset/2160-0.txt': 30, 'Q1-Dataset/pg11.txt': 5, 'Q1-Dataset/pg2600.txt': 172, 'Q1-Dataset/pg345.txt': 75, 'Q1-Dataset/pg64317.txt': 27, 'Q1-Dataset/pg70815.txt': 8, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 9, 'Q1-Dataset/pg1998.txt': 44}, 'active': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 3, 'Q1-

Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 24, 'Q1-Dataset/pg345.txt': 9, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 4, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 6}, 'page': {'Q1-Dataset/pg1513.txt': 13, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 14, 'Q1-Dataset/pg345.txt': 4, 'Q1-Dataset/pg64317.txt': 3, 'Q1-Dataset/pg70815.txt': 3, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 4, 'Q1-Dataset/pg1998.txt': 5}, 'c': {'Q1-Dataset/pg1513.txt': 3, 'Q1-Dataset/2160-0.txt': 28, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 5, 'Q1-Dataset/pg345.txt': 4, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 2}, 'above': {'Q1-Dataset/pg1513.txt': 10, 'Q1-Dataset/2160-0.txt': 79, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.txt': 146, 'Q1-Dataset/pg345.txt': 17, 'Q1-Dataset/pg64317.txt': 7, 'Q1-Dataset/pg70815.txt': 99, 'Q1-Dataset/1080-0.txt': 4, 'Q1-Dataset/pg70817.txt': 7, 'Q1-Dataset/pg1998.txt': 58}, 'disk': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'able': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 31, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 107, 'Q1-Dataset/pg345.txt': 47, 'Q1-Dataset/pg64317.txt': 6, 'Q1-Dataset/pg70815.txt': 18, 'Q1-Dataset/1080-0.txt': 4, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 19}, 'occur': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 3, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 22, 'Q1-Dataset/pg345.txt': 9, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 3}, 'guttenberg': {'Q1-Dataset/pg1513.txt': 32, 'Q1-Dataset/2160-0.txt': 29, 'Q1-Dataset/pg11.txt': 31, 'Q1-Dataset/pg2600.txt': 31, 'Q1-Dataset/pg345.txt': 31, 'Q1-Dataset/pg64317.txt': 32, 'Q1-Dataset/pg70815.txt': 87, 'Q1-Dataset/1080-0.txt': 31, 'Q1-Dataset/pg70817.txt': 87, 'Q1-Dataset/pg1998.txt': 87}, 'let': {'Q1-Dataset/pg1513.txt': 44, 'Q1-Dataset/2160-0.txt': 46, 'Q1-Dataset/pg11.txt': 17, 'Q1-Dataset/pg2600.txt': 376, 'Q1-Dataset/pg345.txt': 144, 'Q1-Dataset/pg64317.txt': 26, 'Q1-Dataset/pg70815.txt': 40, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 4, 'Q1-Dataset/pg1998.txt': 102}, 'start': {'Q1-Dataset/pg1513.txt': 3, 'Q1-Dataset/2160-0.txt': 7, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.txt': 33, 'Q1-Dataset/pg345.txt': 29, 'Q1-Dataset/pg64317.txt': 12, 'Q1-Dataset/pg70815.txt': 13, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 7, 'Q1-Dataset/pg1998.txt': 6}, 'distribution': {'Q1-Dataset/pg1513.txt': 6, 'Q1-Dataset/2160-0.txt': 8, 'Q1-Dataset/pg11.txt': 6, 'Q1-Dataset/pg2600.txt': 10, 'Q1-Dataset/pg345.txt': 8, 'Q1-Dataset/pg64317.txt': 6, 'Q1-Dataset/pg70815.txt': 6, 'Q1-Dataset/1080-0.txt': 6, 'Q1-Dataset/pg70817.txt': 6, 'Q1-Dataset/pg1998.txt': 6}, 'particularly': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 7, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 87, 'Q1-Dataset/pg345.txt': 3, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 10, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 5}, 'donors': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 2, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'linked': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 5, 'Q1-Dataset/pg345.txt': 2, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 1}, 'three': {'Q1-Dataset/pg1513.txt': 8, 'Q1-Dataset/2160-0.txt': 77, 'Q1-Dataset/pg11.txt': 26, 'Q1-Dataset/pg2600.txt': 255, 'Q1-Dataset/pg345.txt': 58, 'Q1-Dataset/pg64317.txt': 35, 'Q1-Dataset/pg70815.txt': 72, 'Q1-Dataset/1080-0.txt': 4, 'Q1-Dataset/pg70817.txt': 13, 'Q1-Dataset/pg1998.txt': 33}, 'certain': {'Q1-Dataset/pg1513.txt': 3, 'Q1-Dataset/2160-0.txt': 46, 'Q1-Dataset/pg11.txt': 5, 'Q1-Dataset/pg2600.txt': 138, 'Q1-Dataset/pg345.txt': 41, 'Q1-Dataset/pg64317.txt': 15, 'Q1-Dataset/pg70815.txt': 21, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 6, 'Q1-Dataset/pg1998.txt': 20}, 'owns': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 3, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 2, 'Q1-Dataset/pg345.txt': 2, 'Q1-Dataset/pg64317.txt': 3, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 2}, 'transcription': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'maintaining': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 4, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 6, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'approach': {'Q1-Dataset/pg1513.txt': 4, 'Q1-Dataset/2160-0.txt': 3, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 28, 'Q1-Dataset/pg345.txt': 7, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 11, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}

t': 1, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 5}, 'group': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 69, 'Q1-Dataset/pg345.txt': 6, 'Q1-Dataset/pg64317.txt': 5, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 4, 'Q1-Dataset/pg1998.txt': 1}, 'accepted': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 4, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 36, 'Q1-Dataset/pg345.txt': 4, 'Q1-Dataset/pg64317.txt': 9, 'Q1-Dataset/pg70815.txt': 11, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 4, 'Q1-Dataset/pg1998.txt': 3}, 'kind': {'Q1-Dataset/pg1513.txt': 7, 'Q1-Dataset/2160-0.txt': 68, 'Q1-Dataset/pg11.txt': 8, 'Q1-Dataset/pg2600.txt': 141, 'Q1-Dataset/pg345.txt': 58, 'Q1-Dataset/pg64317.txt': 11, 'Q1-Dataset/pg70815.txt': 12, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 13, 'Q1-Dataset/pg1998.txt': 16}, 'however': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 97, 'Q1-Dataset/pg11.txt': 21, 'Q1-Dataset/pg2600.txt': 100, 'Q1-Dataset/pg345.txt': 103, 'Q1-Dataset/pg64317.txt': 8, 'Q1-Dataset/pg70815.txt': 23, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 19, 'Q1-Dataset/pg1998.txt': 362}, 'obsolete': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'subject': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 53, 'Q1-Dataset/pg11.txt': 7, 'Q1-Dataset/pg2600.txt': 84, 'Q1-Dataset/pg345.txt': 24, 'Q1-Dataset/pg64317.txt': 8, 'Q1-Dataset/pg70815.txt': 6, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 3, 'Q1-Dataset/pg1998.txt': 9}, 'again': {'Q1-Dataset/pg1513.txt': 27, 'Q1-Dataset/2160-0.txt': 48, 'Q1-Dataset/pg11.txt': 82, 'Q1-Dataset/pg2600.txt': 703, 'Q1-Dataset/pg345.txt': 246, 'Q1-Dataset/pg64317.txt': 56, 'Q1-Dataset/pg70815.txt': 82, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 35, 'Q1-Dataset/pg1998.txt': 157}, 'writing': {'Q1-Dataset/pg1513.txt': 4, 'Q1-Dataset/2160-0.txt': 9, 'Q1-Dataset/pg11.txt': 7, 'Q1-Dataset/pg2600.txt': 42, 'Q1-Dataset/pg345.txt': 22, 'Q1-Dataset/pg64317.txt': 4, 'Q1-Dataset/pg70815.txt': 4, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 5, 'Q1-Dataset/pg1998.txt': 11}, 'stand': {'Q1-Dataset/pg1513.txt': 25, 'Q1-Dataset/2160-0.txt': 28, 'Q1-Dataset/pg11.txt': 6, 'Q1-Dataset/pg2600.txt': 51, 'Q1-Dataset/pg345.txt': 13, 'Q1-Dataset/pg64317.txt': 9, 'Q1-Dataset/pg70815.txt': 9, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 4, 'Q1-Dataset/pg1998.txt': 52}, 'loose': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 10, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 17, 'Q1-Dataset/pg345.txt': 7, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 5, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'accepting': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 3, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 4, 'Q1-Dataset/pg345.txt': 4, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 2}, 'cause': {'Q1-Dataset/pg1513.txt': 7, 'Q1-Dataset/2160-0.txt': 31, 'Q1-Dataset/pg11.txt': 5, 'Q1-Dataset/pg2600.txt': 146, 'Q1-Dataset/pg345.txt': 51, 'Q1-Dataset/pg64317.txt': 3, 'Q1-Dataset/pg70815.txt': 12, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 21}, 'but': {'Q1-Dataset/pg1513.txt': 185, 'Q1-Dataset/2160-0.txt': 1020, 'Q1-Dataset/pg11.txt': 168, 'Q1-Dataset/pg2600.txt': 4011, 'Q1-Dataset/pg345.txt': 1069, 'Q1-Dataset/pg64317.txt': 243, 'Q1-Dataset/pg70815.txt': 555, 'Q1-Dataset/1080-0.txt': 13, 'Q1-Dataset/pg70817.txt': 151, 'Q1-Dataset/pg1998.txt': 680}, 'money': {'Q1-Dataset/pg1513.txt': 4, 'Q1-Dataset/2160-0.txt': 34, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.txt': 127, 'Q1-Dataset/pg345.txt': 19, 'Q1-Dataset/pg64317.txt': 24, 'Q1-Dataset/pg70815.txt': 9, 'Q1-Dataset/1080-0.txt': 6, 'Q1-Dataset/pg70817.txt': 5, 'Q1-Dataset/pg1998.txt': 3}, 'anyone': {'Q1-Dataset/pg1513.txt': 5, 'Q1-Dataset/2160-0.txt': 5, 'Q1-Dataset/pg11.txt': 5, 'Q1-Dataset/pg2600.txt': 182, 'Q1-Dataset/pg345.txt': 6, 'Q1-Dataset/pg64317.txt': 15, 'Q1-Dataset/pg70815.txt': 16, 'Q1-Dataset/1080-0.txt': 5, 'Q1-Dataset/pg70817.txt': 7, 'Q1-Dataset/pg1998.txt': 5}, 'contract': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 3, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 2, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 2}, 'house': {'Q1-Dataset/pg1513.txt': 32, 'Q1-Dataset/2160-0.txt': 170, 'Q1-Dataset/pg11.txt': 18, 'Q1-Dataset/pg2600.txt': 389, 'Q1-Dataset/pg345.txt': 129, 'Q1-Dataset/pg64317.txt': 91, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 61, 'Q1-Dataset/pg1998.txt': 14}, 'am': {'Q1-Dataset/pg1513.txt': 60, 'Q1-Dataset/2160-0.txt': 218, 'Q1-Dataset/pg11.txt': 15, 'Q1-Dataset/pg2600.txt': 545, 'Q1-Dataset/pg345.txt': 336, 'Q1-Dataset/pg64317.txt': 17, 'Q1-Dataset/pg70815.txt': 87, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 8, 'Q1-Dataset/pg1998.txt': 186}, 'take': {'Q1-Dataset/pg1513.txt': 40, 'Q1-Dataset/2160-0.txt': 143, 'Q1-Dataset/pg11.txt': 21, 'Q1-Dataset/pg2600.txt': 349, 'Q1-Dataset/pg345.txt': 148, 'Q1-Dataset/pg64317.txt': 28, 'Q1-Dataset/pg70815.txt': 61, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 8, 'Q1-Dataset/pg1998.txt': 66}, 'exists': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 2,

'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 18, 'Q1-Dataset/pg345.txt': 2, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 3, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'commercial': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 4, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 2, 'Q1-Dataset/pg345.txt': 2, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'do': {'Q1-Dataset/pg1513.txt': 105, 'Q1-Dataset/2160-0.txt': 108, 'Q1-Dataset/pg11.txt': 97, 'Q1-Dataset/pg2600.txt': 1048, 'Q1-Dataset/pg345.txt': 487, 'Q1-Dataset/pg64317.txt': 100, 'Q1-Dataset/pg70815.txt': 165, 'Q1-Dataset/1080-0.txt': 19, 'Q1-Dataset/pg70817.txt': 26, 'Q1-Dataset/pg1998.txt': 518}, 'years': {'Q1-Dataset/pg1513.txt': 13, 'Q1-Dataset/2160-0.txt': 78, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 124, 'Q1-Dataset/pg345.txt': 28, 'Q1-Dataset/pg64317.txt': 39, 'Q1-Dataset/pg70815.txt': 11, 'Q1-Dataset/1080-0.txt': 9, 'Q1-Dataset/pg70817.txt': 16, 'Q1-Dataset/pg1998.txt': 25}, 'expenses': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 3, 'Q1-Dataset/pg345.txt': 2, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 2}, 'read': {'Q1-Dataset/pg1513.txt': 9, 'Q1-Dataset/2160-0.txt': 26, 'Q1-Dataset/pg11.txt': 14, 'Q1-Dataset/pg2600.txt': 132, 'Q1-Dataset/pg345.txt': 84, 'Q1-Dataset/pg64317.txt': 14, 'Q1-Dataset/pg70815.txt': 5, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 13, 'Q1-Dataset/pg1998.txt': 26}, 'restrictions': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 4, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 3, 'Q1-Dataset/pg345.txt': 2, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 2}, 'like': {'Q1-Dataset/pg1513.txt': 38, 'Q1-Dataset/2160-0.txt': 162, 'Q1-Dataset/pg11.txt': 85, 'Q1-Dataset/pg2600.txt': 733, 'Q1-Dataset/pg345.txt': 275, 'Q1-Dataset/pg64317.txt': 120, 'Q1-Dataset/pg70815.txt': 50, 'Q1-Dataset/1080-0.txt': 4, 'Q1-Dataset/pg70817.txt': 39, 'Q1-Dataset/pg1998.txt': 296}, 'edition': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 3, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 3}, 'limited': {'Q1-Dataset/pg1513.txt': 5, 'Q1-Dataset/2160-0.txt': 5, 'Q1-Dataset/pg11.txt': 5, 'Q1-Dataset/pg2600.txt': 18, 'Q1-Dataset/pg345.txt': 8, 'Q1-Dataset/pg64317.txt': 8, 'Q1-Dataset/pg70815.txt': 6, 'Q1-Dataset/1080-0.txt': 5, 'Q1-Dataset/pg70817.txt': 6, 'Q1-Dataset/pg1998.txt': 6}, 'promoting': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 3, 'Q1-Dataset/pg345.txt': 2, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 2}, 'prepare': {'Q1-Dataset/pg1513.txt': 7, 'Q1-Dataset/2160-0.txt': 11, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 28, 'Q1-Dataset/pg345.txt': 9, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 4, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 5}, 'child': {'Q1-Dataset/pg1513.txt': 15, 'Q1-Dataset/2160-0.txt': 14, 'Q1-Dataset/pg11.txt': 10, 'Q1-Dataset/pg2600.txt': 100, 'Q1-Dataset/pg345.txt': 61, 'Q1-Dataset/pg64317.txt': 9, 'Q1-Dataset/pg70815.txt': 3, 'Q1-Dataset/1080-0.txt': 9, 'Q1-Dataset/pg70817.txt': 10, 'Q1-Dataset/pg1998.txt': 48}, 'young': {'Q1-Dataset/pg1513.txt': 24, 'Q1-Dataset/2160-0.txt': 81, 'Q1-Dataset/pg11.txt': 5, 'Q1-Dataset/pg2600.txt': 442, 'Q1-Dataset/pg345.txt': 52, 'Q1-Dataset/pg64317.txt': 44, 'Q1-Dataset/pg70815.txt': 20, 'Q1-Dataset/1080-0.txt': 5, 'Q1-Dataset/pg70817.txt': 18, 'Q1-Dataset/pg1998.txt': 25}, 'using': {'Q1-Dataset/pg1513.txt': 6, 'Q1-Dataset/2160-0.txt': 5, 'Q1-Dataset/pg11.txt': 8, 'Q1-Dataset/pg2600.txt': 18, 'Q1-Dataset/pg345.txt': 17, 'Q1-Dataset/pg64317.txt': 6, 'Q1-Dataset/pg70815.txt': 12, 'Q1-Dataset/1080-0.txt': 7, 'Q1-Dataset/pg70817.txt': 7, 'Q1-Dataset/pg1998.txt': 6}, 'unenforceability': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'individual': {'Q1-Dataset/pg1513.txt': 4, 'Q1-Dataset/2160-0.txt': 24, 'Q1-Dataset/pg11.txt': 4, 'Q1-Dataset/pg2600.txt': 18, 'Q1-Dataset/pg345.txt': 7, 'Q1-Dataset/pg64317.txt': 5, 'Q1-Dataset/pg70815.txt': 7, 'Q1-Dataset/1080-0.txt': 5, 'Q1-Dataset/pg70817.txt': 5, 'Q1-Dataset/pg1998.txt': 12}, 'prevent': {'Q1-Dataset/pg1513.txt': 3, 'Q1-Dataset/2160-0.txt': 11, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 26, 'Q1-Dataset/pg345.txt': 8, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 3, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 3}, 'collection': {'Q1-Dataset/pg1513.txt': 4, 'Q1-Dataset/2160-0.txt': 9, 'Q1-Dataset/pg11.txt': 4, 'Q1-Dataset/pg2600.txt': 8, 'Q1-Dataset/pg345.txt': 4, 'Q1-Dataset/pg64317.txt': 4, 'Q1-Dataset/pg70815.txt': 4, 'Q1-Dataset/1080-0.txt': 4, 'Q1-Dataset/pg70817.txt': 6, 'Q1-Dataset/pg1998.txt': 4}, 'other': {'Q1-Dataset/pg1513.txt': 33, 'Q1-Dataset/2160-0.txt': 256, 'Q1-Dataset/pg11.txt': 56, 'Q1-Dataset/pg2600.txt': 584, 'Q1-Dataset/pg345.txt': 190, 'Q1-Dataset/pg64317.txt': 81, 'Q1-Dataset/pg70815.txt': 16, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}

et/pg70815.txt': 154, 'Q1-Dataset/1080-0.txt': 26, 'Q1-Dataset/pg70817.txt': 34, 'Q1-Dataset/pg1998.txt': 113}, 'vanilla': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 2, 'Q1-Dataset/pg345.txt': 2, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 2}, 'efforts': {'Q1-Dataset/pg1513.txt': 3, 'Q1-Dataset/2160-0.txt': 12, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.txt': 4, 'Q1-Dataset/pg345.txt': 12, 'Q1-Dataset/pg64317.txt': 4, 'Q1-Dataset/pg70815.txt': 9, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 5, 'Q1-Dataset/pg1998.txt': 4}, 'despite': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 81, 'Q1-Dataset/pg345.txt': 3, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 3}, 'others': {'Q1-Dataset/pg1513.txt': 10, 'Q1-Dataset/2160-0.txt': 24, 'Q1-Dataset/pg11.txt': 8, 'Q1-Dataset/pg2600.txt': 307, 'Q1-Dataset/pg345.txt': 81, 'Q1-Dataset/pg64317.txt': 5, 'Q1-Dataset/pg70815.txt': 51, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 7, 'Q1-Dataset/pg1998.txt': 35}, 'country': {'Q1-Dataset/pg1513.txt': 5, 'Q1-Dataset/2160-0.txt': 142, 'Q1-Dataset/pg11.txt': 5, 'Q1-Dataset/pg2600.txt': 102, 'Q1-Dataset/pg345.txt': 28, 'Q1-Dataset/pg64317.txt': 9, 'Q1-Dataset/pg70815.txt': 67, 'Q1-Dataset/1080-0.txt': 17, 'Q1-Dataset/pg70817.txt': 5, 'Q1-Dataset/pg1998.txt': 10}, 'move': {'Q1-Dataset/pg1513.txt': 9, 'Q1-Dataset/2160-0.txt': 6, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.txt': 84, 'Q1-Dataset/pg345.txt': 27, 'Q1-Dataset/pg64317.txt': 10, 'Q1-Dataset/pg70815.txt': 12, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 10, 'Q1-Dataset/pg1998.txt': 3}, 'uniform': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 95, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 5, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 4}, 'another': {'Q1-Dataset/pg1513.txt': 10, 'Q1-Dataset/2160-0.txt': 137, 'Q1-Dataset/pg11.txt': 22, 'Q1-Dataset/pg2600.txt': 601, 'Q1-Dataset/pg345.txt': 79, 'Q1-Dataset/pg64317.txt': 36, 'Q1-Dataset/pg70815.txt': 79, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 17, 'Q1-Dataset/pg1998.txt': 71}, 'children': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 13, 'Q1-Dataset/pg11.txt': 10, 'Q1-Dataset/pg2600.txt': 103, 'Q1-Dataset/pg345.txt': 21, 'Q1-Dataset/pg64317.txt': 6, 'Q1-Dataset/pg70815.txt': 11, 'Q1-Dataset/1080-0.txt': 20, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 49}, 'for': {'Q1-Dataset/pg1513.txt': 249, 'Q1-Dataset/2160-0.txt': 1262, 'Q1-Dataset/pg11.txt': 179, 'Q1-Dataset/pg2600.txt': 3521, 'Q1-Dataset/pg345.txt': 1535, 'Q1-Dataset/pg64317.txt': 336, 'Q1-Dataset/pg70815.txt': 607, 'Q1-Dataset/1080-0.txt': 74, 'Q1-Dataset/pg70817.txt': 167, 'Q1-Dataset/pg1998.txt': 1136}, 'distributor': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'up': {'Q1-Dataset/pg1513.txt': 59, 'Q1-Dataset/2160-0.txt': 238, 'Q1-Dataset/pg11.txt': 101, 'Q1-Dataset/pg2600.txt': 1553, 'Q1-Dataset/pg345.txt': 445, 'Q1-Dataset/pg64317.txt': 194, 'Q1-Dataset/pg70815.txt': 96, 'Q1-Dataset/1080-0.txt': 7, 'Q1-Dataset/pg70817.txt': 78, 'Q1-Dataset/pg1998.txt': 179}, 'permission': {'Q1-Dataset/pg1513.txt': 6, 'Q1-Dataset/2160-0.txt': 13, 'Q1-Dataset/pg11.txt': 6, 'Q1-Dataset/pg2600.txt': 29, 'Q1-Dataset/pg345.txt': 9, 'Q1-Dataset/pg64317.txt': 6, 'Q1-Dataset/pg70815.txt': 7, 'Q1-Dataset/1080-0.txt': 6, 'Q1-Dataset/pg70817.txt': 6, 'Q1-Dataset/pg1998.txt': 6}, 'file': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 3, 'Q1-Dataset/pg345.txt': 2, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 2}, 'they': {'Q1-Dataset/pg1513.txt': 55, 'Q1-Dataset/2160-0.txt': 472, 'Q1-Dataset/pg11.txt': 128, 'Q1-Dataset/pg2600.txt': 2165, 'Q1-Dataset/pg345.txt': 464, 'Q1-Dataset/pg64317.txt': 146, 'Q1-Dataset/pg70815.txt': 678, 'Q1-Dataset/1080-0.txt': 30, 'Q1-Dataset/pg70817.txt': 72, 'Q1-Dataset/pg1998.txt': 635}, 'donation': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'property': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 12, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 36, 'Q1-Dataset/pg345.txt': 9, 'Q1-Dataset/pg64317.txt': 3, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 3, 'Q1-Dataset/pg1998.txt': 3}, 'licensed': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'includes': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 2, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'your': {'Q1-D

pg1513.txt': 113, 'Q1-Dataset/2160-0.txt': 404, 'Q1-Dataset/pg11.txt': 64, 'Q1-Dataset/pg2600.txt': 816, 'Q1-Dataset/pg345.txt': 312, 'Q1-Dataset/pg64317.txt': 78, 'Q1-Dataset/pg70815.txt': 75, 'Q1-Dataset/1080-0.txt': 9, 'Q1-Dataset/pg70817.txt': 17, 'Q1-Dataset/pg1998.txt': 363}, 'indicating': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 24, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 3, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'method': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 12, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 12, 'Q1-Dataset/pg345.txt': 7, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 3, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 4}, 'laws': {'Q1-Dataset/pg1513.txt': 10, 'Q1-Dataset/2160-0.txt': 8, 'Q1-Dataset/pg11.txt': 10, 'Q1-Dataset/pg2600.txt': 97, 'Q1-Dataset/pg345.txt': 10, 'Q1-Dataset/pg64317.txt': 11, 'Q1-Dataset/pg70815.txt': 14, 'Q1-Dataset/1080-0.txt': 11, 'Q1-Dataset/pg70817.txt': 11, 'Q1-Dataset/pg1998.txt': 13}, 'present': {'Q1-Dataset/pg1513.txt': 3, 'Q1-Dataset/2160-0.txt': 92, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 124, 'Q1-Dataset/pg345.txt': 57, 'Q1-Dataset/pg64317.txt': 6, 'Q1-Dataset/pg70815.txt': 6, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 55}, 'future': {'Q1-Dataset/pg1513.txt': 3, 'Q1-Dataset/2160-0.txt': 12, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.txt': 76, 'Q1-Dataset/pg345.txt': 11, 'Q1-Dataset/pg64317.txt': 8, 'Q1-Dataset/pg70815.txt': 15, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 7, 'Q1-Dataset/pg1998.txt': 44}, 'age': {'Q1-Dataset/pg1513.txt': 6, 'Q1-Dataset/2160-0.txt': 27, 'Q1-Dataset/pg11.txt': 4, 'Q1-Dataset/pg2600.txt': 42, 'Q1-Dataset/pg345.txt': 13, 'Q1-Dataset/pg64317.txt': 5, 'Q1-Dataset/pg70815.txt': 18, 'Q1-Dataset/1080-0.txt': 4, 'Q1-Dataset/pg70817.txt': 6, 'Q1-Dataset/pg1998.txt': 10}, 'girl': {'Q1-Dataset/pg1513.txt': 6, 'Q1-Dataset/2160-0.txt': 20, 'Q1-Dataset/pg11.txt': 4, 'Q1-Dataset/pg2600.txt': 129, 'Q1-Dataset/pg345.txt': 33, 'Q1-Dataset/pg64317.txt': 47, 'Q1-Dataset/pg70815.txt': 66, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 2}, 'special': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 73, 'Q1-Dataset/pg345.txt': 12, 'Q1-Dataset/pg64317.txt': 3, 'Q1-Dataset/pg70815.txt': 3, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 3, 'Q1-Dataset/pg1998.txt': 2}, 'reason': {'Q1-Dataset/pg1513.txt': 5, 'Q1-Dataset/2160-0.txt': 42, 'Q1-Dataset/pg11.txt': 9, 'Q1-Dataset/pg2600.txt': 126, 'Q1-Dataset/pg345.txt': 28, 'Q1-Dataset/pg64317.txt': 6, 'Q1-Dataset/pg70815.txt': 20, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 9, 'Q1-Dataset/pg1998.txt': 23}, 'each': {'Q1-Dataset/pg1513.txt': 3, 'Q1-Dataset/2160-0.txt': 30, 'Q1-Dataset/pg11.txt': 9, 'Q1-Dataset/pg2600.txt': 183, 'Q1-Dataset/pg345.txt': 87, 'Q1-Dataset/pg64317.txt': 34, 'Q1-Dataset/pg70815.txt': 45, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 8, 'Q1-Dataset/pg1998.txt': 42}, 'types': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 2, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 3}, 'specified': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 4, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 2, 'Q1-Dataset/pg345.txt': 2, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 2}, 'will': {'Q1-Dataset/pg1513.txt': 156, 'Q1-Dataset/2160-0.txt': 390, 'Q1-Dataset/pg11.txt': 37, 'Q1-Dataset/pg2600.txt': 1064, 'Q1-Dataset/pg345.txt': 458, 'Q1-Dataset/pg64317.txt': 28, 'Q1-Dataset/pg70815.txt': 132, 'Q1-Dataset/1080-0.txt': 45, 'Q1-Dataset/pg70817.txt': 22, 'Q1-Dataset/pg1998.txt': 404}, 'provisions': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 26, 'Q1-Dataset/pg345.txt': 4, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 4, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'committed': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 10, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 17, 'Q1-Dataset/pg345.txt': 3, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 2}, 'ensuring': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 2, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'could': {'Q1-Dataset/pg1513.txt': 18, 'Q1-Dataset/2160-0.txt': 247, 'Q1-Dataset/pg11.txt': 75, 'Q1-Dataset/pg2600.txt': 1111, 'Q1-Dataset/pg345.txt': 492, 'Q1-Dataset/pg64317.txt': 62, 'Q1-Dataset/pg70815.txt': 247, 'Q1-Dataset/1080-0.txt': 6, 'Q1-Dataset/pg70817.txt': 64, 'Q1-Dataset/pg1998.txt': 117}, 'swamp': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 3, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 10, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 14, 'Q1-Dataset/pg1998.txt': 12}, 'unless': {'Q1-Dataset/pg1513.txt': 9, 'Q1-Dataset/2160-0.txt': 13, 'Q1-Dataset/pg11.txt': 5, 'Q1-Dataset/pg2600.txt': 15, 'Q1-Dataset/pg345.txt': 2

2, 'Q1-Dataset/pg64317.txt': 4, 'Q1-Dataset/pg70815.txt': 15, 'Q1-Dataset/1080-0.txt': 4, 'Q1-Dataset/pg70817.txt': 5, 'Q1-Dataset/pg1998.txt': 20}, 'might': {'Q1-Dataset/pg1513.txt': 6, 'Q1-Dataset/2160-0.txt': 137, 'Q1-Dataset/pg11.txt': 28, 'Q1-Dataset/pg2600.txt': 314, 'Q1-Dataset/pg345.txt': 158, 'Q1-Dataset/pg64317.txt': 26, 'Q1-Dataset/pg70815.txt': 163, 'Q1-Dataset/1080-0.txt': 4, 'Q1-Dataset/pg70817.txt': 22, 'Q1-Dataset/pg1998.txt': 44}, 'that': {'Q1-Dataset/pg1513.txt': 361, 'Q1-Dataset/2160-0.txt': 1878, 'Q1-Dataset/pg11.txt': 280, 'Q1-Dataset/pg2600.txt': 7806, 'Q1-Dataset/pg345.txt': 2453, 'Q1-Dataset/pg64317.txt': 596, 'Q1-Dataset/pg70815.txt': 1441, 'Q1-Dataset/1080-0.txt': 54, 'Q1-Dataset/pg70817.txt': 270, 'Q1-Dataset/pg1998.txt': 1262}, 'life': {'Q1-Dataset/pg1513.txt': 24, 'Q1-Dataset/2160-0.txt': 118, 'Q1-Dataset/pg11.txt': 12, 'Q1-Dataset/pg2600.txt': 589, 'Q1-Dataset/pg345.txt': 143, 'Q1-Dataset/pg64317.txt': 30, 'Q1-Dataset/pg70815.txt': 63, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 20, 'Q1-Dataset/pg1998.txt': 180}, 'being': {'Q1-Dataset/pg1513.txt': 26, 'Q1-Dataset/2160-0.txt': 221, 'Q1-Dataset/pg11.txt': 19, 'Q1-Dataset/pg2600.txt': 460, 'Q1-Dataset/pg345.txt': 82, 'Q1-Dataset/pg64317.txt': 15, 'Q1-Dataset/pg70815.txt': 60, 'Q1-Dataset/1080-0.txt': 17, 'Q1-Dataset/pg70817.txt': 11, 'Q1-Dataset/pg1998.txt': 47}, 'critical': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 6, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'been': {'Q1-Dataset/pg1513.txt': 20, 'Q1-Dataset/2160-0.txt': 394, 'Q1-Dataset/pg11.txt': 38, 'Q1-Dataset/pg2600.txt': 1475, 'Q1-Dataset/pg345.txt': 393, 'Q1-Dataset/pg64317.txt': 107, 'Q1-Dataset/pg70815.txt': 269, 'Q1-Dataset/1080-0.txt': 7, 'Q1-Dataset/pg70817.txt': 68, 'Q1-Dataset/pg1998.txt': 149}, 'little': {'Q1-Dataset/pg1513.txt': 13, 'Q1-Dataset/2160-0.txt': 148, 'Q1-Dataset/pg11.txt': 127, 'Q1-Dataset/pg2600.txt': 554, 'Q1-Dataset/pg345.txt': 168, 'Q1-Dataset/pg64317.txt': 103, 'Q1-Dataset/pg70815.txt': 65, 'Q1-Dataset/1080-0.txt': 7, 'Q1-Dataset/pg70817.txt': 26, 'Q1-Dataset/pg1998.txt': 109}, 'payments': {'Q1-Dataset/pg1513.txt': 3, 'Q1-Dataset/2160-0.txt': 4, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.txt': 6, 'Q1-Dataset/pg345.txt': 3, 'Q1-Dataset/pg64317.txt': 3, 'Q1-Dataset/pg70815.txt': 3, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 3, 'Q1-Dataset/pg1998.txt': 3}, 'increasing': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 14, 'Q1-Dataset/pg345.txt': 5, 'Q1-Dataset/pg64317.txt': 3, 'Q1-Dataset/pg70815.txt': 5, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 3, 'Q1-Dataset/pg1998.txt': 2}, 'provided': {'Q1-Dataset/pg1513.txt': 5, 'Q1-Dataset/2160-0.txt': 30, 'Q1-Dataset/pg11.txt': 4, 'Q1-Dataset/pg2600.txt': 12, 'Q1-Dataset/pg345.txt': 4, 'Q1-Dataset/pg64317.txt': 6, 'Q1-Dataset/pg70815.txt': 4, 'Q1-Dataset/1080-0.txt': 6, 'Q1-Dataset/pg70817.txt': 4, 'Q1-Dataset/pg1998.txt': 8}, 'wwwgutenbergorg': {'Q1-Dataset/pg1513.txt': 5, 'Q1-Dataset/2160-0.txt': 5, 'Q1-Dataset/pg11.txt': 5, 'Q1-Dataset/pg2600.txt': 5, 'Q1-Dataset/pg345.txt': 5, 'Q1-Dataset/pg64317.txt': 5, 'Q1-Dataset/pg70815.txt': 5, 'Q1-Dataset/1080-0.txt': 5, 'Q1-Dataset/pg70817.txt': 5, 'Q1-Dataset/pg1998.txt': 5}, 'times': {'Q1-Dataset/pg1513.txt': 8, 'Q1-Dataset/2160-0.txt': 43, 'Q1-Dataset/pg11.txt': 6, 'Q1-Dataset/pg2600.txt': 151, 'Q1-Dataset/pg345.txt': 55, 'Q1-Dataset/pg64317.txt': 17, 'Q1-Dataset/pg70815.txt': 22, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 6, 'Q1-Dataset/pg1998.txt': 23}, 'clearly': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 91, 'Q1-Dataset/pg345.txt': 4, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 8}, 'determine': {'Q1-Dataset/pg1513.txt': 3, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 8, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 3, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 4}, 'considerable': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 30, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 7, 'Q1-Dataset/pg345.txt': 12, 'Q1-Dataset/pg64317.txt': 3, 'Q1-Dataset/pg70815.txt': 32, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 3}, 'beginning': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 10, 'Q1-Dataset/pg11.txt': 14, 'Q1-Dataset/pg2600.txt': 89, 'Q1-Dataset/pg345.txt': 32, 'Q1-Dataset/pg64317.txt': 6, 'Q1-Dataset/pg70815.txt': 4, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 5, 'Q1-Dataset/pg1998.txt': 14}, 'death': {'Q1-Dataset/pg1513.txt': 69, 'Q1-Dataset/2160-0.txt': 20, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 178, 'Q1-Dataset/pg345.txt': 76, 'Q1-Dataset/pg64317.txt': 5, 'Q1-Dataset/pg70815.txt': 33, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 24, 'Q1-Dataset/pg1998.txt': 61}, 'receipt': {'Q1-Dataset/pg1513.txt': 3, 'Q1-Dataset/2160-0.txt': 5, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 6, 'Q1-Dataset/pg345.txt': 5, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 2}, 'please': {'Q1-Dataset/pg1513.txt': 6, 'Q1-Dataset/2160-0.txt': 23, 'Q1-Dataset/pg11.txt': 21, 'Q1-Dataset/pg2600.txt': 134, 'Q1-Dataset/pg345.txt': 24, 'Q1-Dataset/pg64317.txt': 13, 'Q1-Dataset/pg70815.txt': 5, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}

txt': 4, 'Q1-Dataset/pg70817.txt': 7, 'Q1-Dataset/pg1998.txt': 11}, 'international': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 3, 'Q1-Dataset/pg345.txt': 2, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'unsolicited': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 2, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'half': {'Q1-Dataset/pg1513.txt': 5, 'Q1-Dataset/2160-0.txt': 64, 'Q1-Dataset/pg11.txt': 21, 'Q1-Dataset/pg2600.txt': 136, 'Q1-Dataset/pg345.txt': 43, 'Q1-Dataset/pg64317.txt': 36, 'Q1-Dataset/pg70815.txt': 22, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 13, 'Q1-Dataset/pg1998.txt': 4}, 'made': {'Q1-Dataset/pg1513.txt': 22, 'Q1-Dataset/2160-0.txt': 227, 'Q1-Dataset/pg11.txt': 30, 'Q1-Dataset/pg2600.txt': 458, 'Q1-Dataset/pg345.txt': 195, 'Q1-Dataset/pg64317.txt': 63, 'Q1-Dataset/pg70815.txt': 75, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 35, 'Q1-Dataset/pg1998.txt': 44}, 'irs': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'the': {'Q1-Dataset/pg1513.txt': 875, 'Q1-Dataset/2160-0.txt': 8521, 'Q1-Dataset/pg11.txt': 1825, 'Q1-Dataset/pg2600.txt': 34575, 'Q1-Dataset/pg345.txt': 8045, 'Q1-Dataset/pg64317.txt': 2573, 'Q1-Dataset/pg70815.txt': 6202, 'Q1-Dataset/1080-0.txt': 358, 'Q1-Dataset/pg70817.txt': 1904, 'Q1-Dataset/pg1998.txt': 6041}, 'legally': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'user': {'Q1-Dataset/pg1513.txt': 3, 'Q1-Dataset/2160-0.txt': 3, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.txt': 3, 'Q1-Dataset/pg345.txt': 3, 'Q1-Dataset/pg64317.txt': 3, 'Q1-Dataset/pg70815.txt': 3, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 3, 'Q1-Dataset/pg1998.txt': 3}, 'same': {'Q1-Dataset/pg1513.txt': 16, 'Q1-Dataset/2160-0.txt': 142, 'Q1-Dataset/pg11.txt': 25, 'Q1-Dataset/pg2600.txt': 663, 'Q1-Dataset/pg345.txt': 111, 'Q1-Dataset/pg64317.txt': 17, 'Q1-Dataset/pg70815.txt': 54, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 4, 'Q1-Dataset/pg1998.txt': 40}, 'survive': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 5, 'Q1-Dataset/pg345.txt': 3, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 3}, 'hope': {'Q1-Dataset/pg1513.txt': 5, 'Q1-Dataset/2160-0.txt': 67, 'Q1-Dataset/pg11.txt': 4, 'Q1-Dataset/pg2600.txt': 90, 'Q1-Dataset/pg345.txt': 65, 'Q1-Dataset/pg64317.txt': 8, 'Q1-Dataset/pg70815.txt': 42, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 3, 'Q1-Dataset/pg1998.txt': 37}, 'performances': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 8, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 2, 'Q1-Dataset/pg345.txt': 2, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'foundations': {'Q1-Dataset/pg1513.txt': 3, 'Q1-Dataset/2160-0.txt': 3, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.txt': 6, 'Q1-Dataset/pg345.txt': 3, 'Q1-Dataset/pg64317.txt': 3, 'Q1-Dataset/pg70815.txt': 3, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 4, 'Q1-Dataset/pg1998.txt': 3}, 'date': {'Q1-Dataset/pg1513.txt': 6, 'Q1-Dataset/2160-0.txt': 3, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.txt': 6, 'Q1-Dataset/pg345.txt': 6, 'Q1-Dataset/pg64317.txt': 4, 'Q1-Dataset/pg70815.txt': 3, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 6, 'Q1-Dataset/pg1998.txt': 4}, 'sending': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 5, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.txt': 19, 'Q1-Dataset/pg345.txt': 5, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 1}, 'employee': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'often': {'Q1-Dataset/pg1513.txt': 3, 'Q1-Dataset/2160-0.txt': 40, 'Q1-Dataset/pg11.txt': 6, 'Q1-Dataset/pg2600.txt': 140, 'Q1-Dataset/pg345.txt': 19, 'Q1-Dataset/pg64317.txt': 8, 'Q1-Dataset/pg70815.txt': 23, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 6, 'Q1-Dataset/pg1998.txt': 45}, 'practically': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 2, 'Q1-Dataset/pg345.txt': 2, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 10, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 7}, 'explanation': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 11, 'Q1-Dataset/pg11.txt': 4, 'Q1-Dataset/pg2600.txt': 40, 'Q1-Dataset/pg345.txt': 6, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 6, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 6, 'Q1-Dataset/pg1998.txt': 4}, 'replacement': {'Q1-Dataset/pg1513.txt': 5, 'Q1-Dataset/2160-0.txt': 5, 'Q1-

'Q1-Dataset/pg11.txt': 5, 'Q1-Dataset/pg2600.txt': 5, 'Q1-Dataset/pg345.txt': 5, 'Q1-Dataset/pg64317.txt': 5, 'Q1-Dataset/pg70815.txt': 5, 'Q1-Dataset/1080-0.txt': 5, 'Q1-Dataset/pg70817.txt': 5, 'Q1-Dataset/pg1998.txt': 5}, 'author': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 6, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 9, 'Q1-Dataset/pg345.txt': 3, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 5, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 6}, 'taken': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 85, 'Q1-Dataset/pg11.txt': 4, 'Q1-Dataset/pg2600.txt': 242, 'Q1-Dataset/pg345.txt': 47, 'Q1-Dataset/pg64317.txt': 9, 'Q1-Dataset/pg70815.txt': 20, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 4, 'Q1-Dataset/pg1998.txt': 20}, 'whole': {'Q1-Dataset/pg1513.txt': 3, 'Q1-Dataset/2160-0.txt': 102, 'Q1-Dataset/pg11.txt': 13, 'Q1-Dataset/pg2600.txt': 531, 'Q1-Dataset/pg345.txt': 73, 'Q1-Dataset/pg64317.txt': 17, 'Q1-Dataset/pg70815.txt': 15, 'Q1-Dataset/1080-0.txt': 4, 'Q1-Dataset/pg70817.txt': 4, 'Q1-Dataset/pg1998.txt': 29}, 'some': {'Q1-Dataset/pg1513.txt': 58, 'Q1-Dataset/2160-0.txt': 345, 'Q1-Dataset/pg11.txt': 51, 'Q1-Dataset/pg2600.txt': 843, 'Q1-Dataset/pg345.txt': 440, 'Q1-Dataset/pg64317.txt': 104, 'Q1-Dataset/pg70815.txt': 103, 'Q1-Dataset/1080-0.txt': 9, 'Q1-Dataset/pg70817.txt': 76, 'Q1-Dataset/pg1998.txt': 73}, 'electronic': {'Q1-Dataset/pg1513.txt': 27, 'Q1-Dataset/2160-0.txt': 27, 'Q1-Dataset/pg11.txt': 27, 'Q1-Dataset/pg2600.txt': 27, 'Q1-Dataset/pg345.txt': 27, 'Q1-Dataset/pg64317.txt': 27, 'Q1-Dataset/pg70815.txt': 27, 'Q1-Dataset/1080-0.txt': 27, 'Q1-Dataset/pg70817.txt': 27, 'Q1-Dataset/pg1998.txt': 27}, 'then': {'Q1-Dataset/pg1513.txt': 84, 'Q1-Dataset/2160-0.txt': 143, 'Q1-Dataset/pg11.txt': 90, 'Q1-Dataset/pg2600.txt': 934, 'Q1-Dataset/pg345.txt': 543, 'Q1-Dataset/pg64317.txt': 137, 'Q1-Dataset/pg70815.txt': 196, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 55, 'Q1-Dataset/pg1998.txt': 336}, 'great': {'Q1-Dataset/pg1513.txt': 9, 'Q1-Dataset/2160-0.txt': 252, 'Q1-Dataset/pg11.txt': 39, 'Q1-Dataset/pg2600.txt': 262, 'Q1-Dataset/pg345.txt': 181, 'Q1-Dataset/pg64317.txt': 30, 'Q1-Dataset/pg70815.txt': 164, 'Q1-Dataset/1080-0.txt': 10, 'Q1-Dataset/pg70817.txt': 16, 'Q1-Dataset/pg1998.txt': 236}, 'sort': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 28, 'Q1-Dataset/pg11.txt': 20, 'Q1-Dataset/pg2600.txt': 57, 'Q1-Dataset/pg345.txt': 80, 'Q1-Dataset/pg64317.txt': 23, 'Q1-Dataset/pg70815.txt': 7, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 4, 'Q1-Dataset/pg1998.txt': 5}, 'redistributing': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 2, 'Q1-Dataset/pg345.txt': 2, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 2}, 'included': {'Q1-Dataset/pg1513.txt': 3, 'Q1-Dataset/2160-0.txt': 3, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.txt': 11, 'Q1-Dataset/pg345.txt': 5, 'Q1-Dataset/pg64317.txt': 9, 'Q1-Dataset/pg70815.txt': 6, 'Q1-Dataset/1080-0.txt': 4, 'Q1-Dataset/pg70817.txt': 4, 'Q1-Dataset/pg1998.txt': 3}, 'royalty': {'Q1-Dataset/pg1513.txt': 3, 'Q1-Dataset/2160-0.txt': 3, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.txt': 4, 'Q1-Dataset/pg345.txt': 3, 'Q1-Dataset/pg64317.txt': 3, 'Q1-Dataset/pg70815.txt': 3, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 3, 'Q1-Dataset/pg1998.txt': 3}, 'deductible': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'order': {'Q1-Dataset/pg1513.txt': 3, 'Q1-Dataset/2160-0.txt': 59, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.txt': 295, 'Q1-Dataset/pg345.txt': 32, 'Q1-Dataset/pg64317.txt': 7, 'Q1-Dataset/pg70815.txt': 5, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 7, 'Q1-Dataset/pg1998.txt': 40}, 'time': {'Q1-Dataset/pg1513.txt': 32, 'Q1-Dataset/2160-0.txt': 210, 'Q1-Dataset/pg11.txt': 68, 'Q1-Dataset/pg2600.txt': 921, 'Q1-Dataset/pg345.txt': 369, 'Q1-Dataset/pg64317.txt': 81, 'Q1-Dataset/pg70815.txt': 141, 'Q1-Dataset/1080-0.txt': 5, 'Q1-Dataset/pg70817.txt': 36, 'Q1-Dataset/pg1998.txt': 152}, 'liable': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 2, 'Q1-Dataset/pg345.txt': 3, 'Q1-Dataset/pg64317.txt': 3, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'him': {'Q1-Dataset/pg1513.txt': 96, 'Q1-Dataset/2160-0.txt': 679, 'Q1-Dataset/pg11.txt': 40, 'Q1-Dataset/pg2600.txt': 4571, 'Q1-Dataset/pg345.txt': 941, 'Q1-Dataset/pg64317.txt': 273, 'Q1-Dataset/pg70815.txt': 496, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 108, 'Q1-Dataset/pg1998.txt': 395}, 'address': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 16, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 30, 'Q1-Dataset/pg345.txt': 9, 'Q1-Dataset/pg64317.txt': 5, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 3}, 'find': {'Q1-Dataset/pg1513.txt': 24, 'Q1-Dataset/2160-0.txt': 113, 'Q1-Dataset/pg11.txt': 21, 'Q1-Dataset/pg2600.txt': 161, 'Q1-Dataset/pg345.txt': 114, 'Q1-Dataset/pg64317.txt': 17, 'Q1-Dataset/pg70815.txt': 57, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 7, 'Q1-Dataset/pg1998.txt': 74}, 'ever': {'Q1-Dataset/pg1513.txt': 12, 'Q1-Dataset/2160-0.txt': 129, 'Q1-Dataset/pg11.txt': 20, 'Q1-Dataset/pg2600.txt': 128, 'Q1-Dataset/pg345.txt': 130, 'Q1-Dataset/pg64317.txt': 2

8, 'Q1-Dataset/pg70815.txt': 53, 'Q1-Dataset/1080-0.txt': 5, 'Q1-Dataset/pg70817.txt': 14, 'Q1-Dataset/pg1998.txt': 144}, 'full': {'Q1-Dataset/pg1513.txt': 24, 'Q1-Dataset/2160-0.txt': 48, 'Q1-Dataset/pg11.txt': 19, 'Q1-Dataset/pg2600.txt': 160, 'Q1-Dataset/pg345.txt': 100, 'Q1-Dataset/pg64317.txt': 34, 'Q1-Dataset/pg70815.txt': 28, 'Q1-Dataset/1080-0.txt': 14, 'Q1-Dataset/pg70817.txt': 21, 'Q1-Dataset/pg1998.txt': 48}, 'sound': {'Q1-Dataset/pg1513.txt': 13, 'Q1-Dataset/2160-0.txt': 18, 'Q1-Dataset/pg11.txt': 4, 'Q1-Dataset/pg2600.txt': 160, 'Q1-Dataset/pg345.txt': 43, 'Q1-Dataset/pg64317.txt': 30, 'Q1-Dataset/pg70815.txt': 34, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 6, 'Q1-Dataset/pg1998.txt': 13}, 'middleaged': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 6, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 3, 'Q1-Dataset/pg1998.txt': 1}, 'must': {'Q1-Dataset/pg1513.txt': 50, 'Q1-Dataset/2160-0.txt': 183, 'Q1-Dataset/pg11.txt': 51, 'Q1-Dataset/pg2600.txt': 482, 'Q1-Dataset/pg345.txt': 450, 'Q1-Dataset/pg64317.txt': 42, 'Q1-Dataset/pg70815.txt': 110, 'Q1-Dataset/1080-0.txt': 11, 'Q1-Dataset/pg70817.txt': 25, 'Q1-Dataset/pg1998.txt': 194}, 'additional': {'Q1-Dataset/pg1513.txt': 3, 'Q1-Dataset/2160-0.txt': 6, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.txt': 6, 'Q1-Dataset/pg345.txt': 5, 'Q1-Dataset/pg64317.txt': 3, 'Q1-Dataset/pg70815.txt': 5, 'Q1-Dataset/1080-0.txt': 4, 'Q1-Dataset/pg70817.txt': 3, 'Q1-Dataset/pg1998.txt': 4}, 'know': {'Q1-Dataset/pg1513.txt': 36, 'Q1-Dataset/2160-0.txt': 180, 'Q1-Dataset/pg11.txt': 86, 'Q1-Dataset/pg2600.txt': 841, 'Q1-Dataset/pg345.txt': 385, 'Q1-Dataset/pg64317.txt': 98, 'Q1-Dataset/pg70815.txt': 100, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 10, 'Q1-Dataset/pg1998.txt': 129}, 'possibility': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 60, 'Q1-Dataset/pg345.txt': 8, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 11, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 3, 'Q1-Dataset/pg1998.txt': 4}, 'promotion': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 8, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'late': {'Q1-Dataset/pg1513.txt': 12, 'Q1-Dataset/2160-0.txt': 13, 'Q1-Dataset/pg11.txt': 6, 'Q1-Dataset/pg2600.txt': 105, 'Q1-Dataset/pg345.txt': 57, 'Q1-Dataset/pg64317.txt': 13, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 9, 'Q1-Dataset/pg1998.txt': 13}, 'hear': {'Q1-Dataset/pg1513.txt': 23, 'Q1-Dataset/2160-0.txt': 39, 'Q1-Dataset/pg11.txt': 15, 'Q1-Dataset/pg2600.txt': 141, 'Q1-Dataset/pg345.txt': 102, 'Q1-Dataset/pg64317.txt': 24, 'Q1-Dataset/pg70815.txt': 8, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 8, 'Q1-Dataset/pg1998.txt': 59}, 'below': {'Q1-Dataset/pg1513.txt': 3, 'Q1-Dataset/2160-0.txt': 17, 'Q1-Dataset/pg11.txt': 6, 'Q1-Dataset/pg2600.txt': 46, 'Q1-Dataset/pg345.txt': 25, 'Q1-Dataset/pg64317.txt': 7, 'Q1-Dataset/pg70815.txt': 62, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 5, 'Q1-Dataset/pg1998.txt': 12}, 'known': {'Q1-Dataset/pg1513.txt': 4, 'Q1-Dataset/2160-0.txt': 28, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 127, 'Q1-Dataset/pg345.txt': 24, 'Q1-Dataset/pg64317.txt': 12, 'Q1-Dataset/pg70815.txt': 24, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 3, 'Q1-Dataset/pg1998.txt': 16}, 'performed': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 12, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 24, 'Q1-Dataset/pg345.txt': 3, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 3}, 'mississippi': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'his': {'Q1-Dataset/pg1513.txt': 134, 'Q1-Dataset/2160-0.txt': 1996, 'Q1-Dataset/pg11.txt': 95, 'Q1-Dataset/pg2600.txt': 7965, 'Q1-Dataset/pg345.txt': 1466, 'Q1-Dataset/pg64317.txt': 490, 'Q1-Dataset/pg70815.txt': 925, 'Q1-Dataset/1080-0.txt': 10, 'Q1-Dataset/pg70817.txt': 280, 'Q1-Dataset/pg1998.txt': 892}, 'believe': {'Q1-Dataset/pg1513.txt': 4, 'Q1-Dataset/2160-0.txt': 74, 'Q1-Dataset/pg11.txt': 9, 'Q1-Dataset/pg2600.txt': 124, 'Q1-Dataset/pg345.txt': 75, 'Q1-Dataset/pg64317.txt': 16, 'Q1-Dataset/pg70815.txt': 27, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 4, 'Q1-Dataset/pg1998.txt': 51}, 'current': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 4, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 19, 'Q1-Dataset/pg345.txt': 2, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 7, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 4}, 'research': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 3, 'Q1-Dataset/pg345.txt': 2, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 4, 'Q1-Dataset/pg1998.txt': 2}, 'assistance': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 22, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 8, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 7, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg19

98.txt': 2}, 'indemnify': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'domain': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 8, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 7, 'Q1-Dataset/pg345.txt': 2, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 4, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 3, 'Q1-Dataset/pg1998.txt': 13}, 'walks': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 12, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 3, 'Q1-Dataset/pg345.txt': 7, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 3}, 'exempt': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 3, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 4, 'Q1-Dataset/pg345.txt': 2, 'Q1-Dataset/pg64317.txt': 3, 'Q1-Dataset/pg70815.txt': 3, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 3}, 'means': {'Q1-Dataset/pg1513.txt': 12, 'Q1-Dataset/2160-0.txt': 63, 'Q1-Dataset/pg11.txt': 7, 'Q1-Dataset/pg2600.txt': 73, 'Q1-Dataset/pg345.txt': 27, 'Q1-Dataset/pg64317.txt': 5, 'Q1-Dataset/pg70815.txt': 35, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 6, 'Q1-Dataset/pg1998.txt': 32}, 'as': {'Q1-Dataset/pg1513.txt': 166, 'Q1-Dataset/2160-0.txt': 1130, 'Q1-Dataset/pg11.txt': 272, 'Q1-Dataset/pg2600.txt': 3986, 'Q1-Dataset/pg345.txt': 1582, 'Q1-Dataset/pg64317.txt': 277, 'Q1-Dataset/pg70815.txt': 688, 'Q1-Dataset/1080-0.txt': 49, 'Q1-Dataset/pg70817.txt': 132, 'Q1-Dataset/pg1998.txt': 541}, 'possession': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 16, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 16, 'Q1-Dataset/pg345.txt': 8, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 5, 'Q1-Dataset/pg1998.txt': 6}, 'locations': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'into': {'Q1-Dataset/pg1513.txt': 8, 'Q1-Dataset/2160-0.txt': 298, 'Q1-Dataset/pg11.txt': 67, 'Q1-Dataset/pg2600.txt': 1003, 'Q1-Dataset/pg345.txt': 287, 'Q1-Dataset/pg64317.txt': 168, 'Q1-Dataset/pg70815.txt': 182, 'Q1-Dataset/1080-0.txt': 4, 'Q1-Dataset/pg70817.txt': 80, 'Q1-Dataset/pg1998.txt': 251}, 'where': {'Q1-Dataset/pg1513.txt': 56, 'Q1-Dataset/2160-0.txt': 221, 'Q1-Dataset/pg11.txt': 18, 'Q1-Dataset/pg2600.txt': 616, 'Q1-Dataset/pg345.txt': 219, 'Q1-Dataset/pg64317.txt': 59, 'Q1-Dataset/pg70815.txt': 133, 'Q1-Dataset/1080-0.txt': 8, 'Q1-Dataset/pg70817.txt': 17, 'Q1-Dataset/pg1998.txt': 142}, 'fee': {'Q1-Dataset/pg1513.txt': 10, 'Q1-Dataset/2160-0.txt': 9, 'Q1-Dataset/pg11.txt': 8, 'Q1-Dataset/pg2600.txt': 8, 'Q1-Dataset/pg345.txt': 8, 'Q1-Dataset/pg64317.txt': 8, 'Q1-Dataset/pg70815.txt': 8, 'Q1-Dataset/1080-0.txt': 8, 'Q1-Dataset/pg70817.txt': 8, 'Q1-Dataset/pg1998.txt': 8}, 'representations': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'hot': {'Q1-Dataset/pg1513.txt': 7, 'Q1-Dataset/2160-0.txt': 34, 'Q1-Dataset/pg11.txt': 5, 'Q1-Dataset/pg2600.txt': 49, 'Q1-Dataset/pg345.txt': 9, 'Q1-Dataset/pg64317.txt': 25, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 4, 'Q1-Dataset/pg1998.txt': 13}, 'remedies': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 3, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 3, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'them': {'Q1-Dataset/pg1513.txt': 43, 'Q1-Dataset/2160-0.txt': 210, 'Q1-Dataset/pg11.txt': 84, 'Q1-Dataset/pg2600.txt': 1512, 'Q1-Dataset/pg345.txt': 467, 'Q1-Dataset/pg64317.txt': 81, 'Q1-Dataset/pg70815.txt': 395, 'Q1-Dataset/1080-0.txt': 16, 'Q1-Dataset/pg70817.txt': 36, 'Q1-Dataset/pg1998.txt': 326}, 'effect': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 47, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.txt': 65, 'Q1-Dataset/pg345.txt': 28, 'Q1-Dataset/pg64317.txt': 3, 'Q1-Dataset/pg70815.txt': 19, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 5}, 'reported': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 36, 'Q1-Dataset/pg345.txt': 11, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 4, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 2}, 'give': {'Q1-Dataset/pg1513.txt': 51, 'Q1-Dataset/2160-0.txt': 122, 'Q1-Dataset/pg11.txt': 16, 'Q1-Dataset/pg2600.txt': 280, 'Q1-Dataset/pg345.txt': 82, 'Q1-Dataset/pg64317.txt': 22, 'Q1-Dataset/pg70815.txt': 20, 'Q1-Dataset/1080-0.txt': 5, 'Q1-Dataset/pg70817.txt': 8, 'Q1-Dataset/pg1998.txt': 75}, 'indirectly': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 2, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'business': {'Q1-Dataset/pg1513.txt': 4, 'Q1-Dataset/2160-0.txt': 40, 'Q1-Dataset/pg11.txt': 8, 'Q1-Dataset/pg2600.txt': 99, 'Q1-

-Dataset/pg345.txt': 22, 'Q1-Dataset/pg64317.txt': 19, 'Q1-Dataset/pg70815.txt': 3, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 3, 'Q1-Dataset/pg1998.txt': 3}, 'downloading': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'used': {'Q1-Dataset/pg1513.txt': 3, 'Q1-Dataset/2160-0.txt': 38, 'Q1-Dataset/pg11.txt': 15, 'Q1-Dataset/pg2600.txt': 135, 'Q1-Dataset/pg345.txt': 46, 'Q1-Dataset/pg64317.txt': 20, 'Q1-Dataset/pg70815.txt': 13, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 8, 'Q1-Dataset/pg1998.txt': 7}, 'incidental': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'nonproprietary': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'binary': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'arise': {'Q1-Dataset/pg1513.txt': 4, 'Q1-Dataset/2160-0.txt': 4, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 9, 'Q1-Dataset/pg345.txt': 2, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 14}, 'remove': {'Q1-Dataset/pg1513.txt': 5, 'Q1-Dataset/2160-0.txt': 8, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 6, 'Q1-Dataset/pg345.txt': 5, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 4, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 3}, 'cold': {'Q1-Dataset/pg1513.txt': 7, 'Q1-Dataset/2160-0.txt': 33, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 137, 'Q1-Dataset/pg345.txt': 43, 'Q1-Dataset/pg64317.txt': 14, 'Q1-Dataset/pg70815.txt': 12, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 22, 'Q1-Dataset/pg1998.txt': 28}, 'convert': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 8, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 2, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'end': {'Q1-Dataset/pg1513.txt': 10, 'Q1-Dataset/2160-0.txt': 35, 'Q1-Dataset/pg11.txt': 19, 'Q1-Dataset/pg2600.txt': 184, 'Q1-Dataset/pg345.txt': 73, 'Q1-Dataset/pg64317.txt': 21, 'Q1-Dataset/pg70815.txt': 38, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 13, 'Q1-Dataset/pg1998.txt': 43}, 'only': {'Q1-Dataset/pg1513.txt': 11, 'Q1-Dataset/2160-0.txt': 114, 'Q1-Dataset/pg11.txt': 51, 'Q1-Dataset/pg2600.txt': 1292, 'Q1-Dataset/pg345.txt': 285, 'Q1-Dataset/pg64317.txt': 53, 'Q1-Dataset/pg70815.txt': 127, 'Q1-Dataset/1080-0.txt': 8, 'Q1-Dataset/pg70817.txt': 34, 'Q1-Dataset/pg1998.txt': 244}, 'best': {'Q1-Dataset/pg1513.txt': 12, 'Q1-Dataset/2160-0.txt': 60, 'Q1-Dataset/pg11.txt': 10, 'Q1-Dataset/pg2600.txt': 124, 'Q1-Dataset/pg345.txt': 51, 'Q1-Dataset/pg64317.txt': 7, 'Q1-Dataset/pg70815.txt': 17, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 5, 'Q1-Dataset/pg1998.txt': 97}, 'electronically': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 2, 'Q1-Dataset/pg345.txt': 2, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 2}, 'perform': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 7, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 16, 'Q1-Dataset/pg345.txt': 4, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'name': {'Q1-Dataset/pg1513.txt': 25, 'Q1-Dataset/2160-0.txt': 81, 'Q1-Dataset/pg11.txt': 11, 'Q1-Dataset/pg2600.txt': 117, 'Q1-Dataset/pg345.txt': 21, 'Q1-Dataset/pg64317.txt': 42, 'Q1-Dataset/pg70815.txt': 13, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 9, 'Q1-Dataset/pg1998.txt': 25}, 'permitted': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 8, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.txt': 5, 'Q1-Dataset/pg345.txt': 5, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 8, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 6, 'Q1-Dataset/pg1998.txt': 5}, 'breach': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 4, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 3, 'Q1-Dataset/pg345.txt': 2, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 3}, 'interpreted': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 3, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 3, 'Q1-Dataset/pg345.txt': 2, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 3, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 3}, 'fees': {'Q1-Dataset/pg1513.txt': 5, 'Q1-Dataset/2160-0.txt': 4, 'Q1-Dataset/pg11.txt': 4, 'Q1-Dataset/pg2600.txt': 4, 'Q1-Dataset/pg345.txt': 4, 'Q1-Dataset/pg64317.txt': 4, 'Q1-Dataset/pg70815.txt': 4, 'Q1-Dataset/1080-0.txt': 4, 'Q1-Dataset/pg70817.txt': 4, 'Q1-Dataset/pg1998.txt': 4}


```
t': 4}, 'understand': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 37, 'Q1-Dataset/pg11.txt': 7, 'Q1-Dataset/pg2600.txt': 347, 'Q1-Dataset/pg345.txt': 56, 'Q1-Dataset/pg64317.txt': 15, 'Q1-Dataset/pg70815.txt': 29, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 31}, 'fix': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 10, 'Q1-Dataset/pg345.txt': 3, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 3}, 'network': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 2, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'worse': {'Q1-Dataset/pg1513.txt': 3, 'Q1-Dataset/2160-0.txt': 15, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.txt': 48, 'Q1-Dataset/pg345.txt': 18, 'Q1-Dataset/pg64317.txt': 3, 'Q1-Dataset/pg70815.txt': 6, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 8, 'Q1-Dataset/pg1998.txt': 12}, 'paying': {'Q1-Dataset/pg1513.txt': 3, 'Q1-Dataset/2160-0.txt': 7, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.txt': 17, 'Q1-Dataset/pg345.txt': 4, 'Q1-Dataset/pg64317.txt': 4, 'Q1-Dataset/pg70815.txt': 4, 'Q1-Dataset/1080-0.txt': 4, 'Q1-Dataset/pg70817.txt': 4, 'Q1-Dataset/pg1998.txt': 3}, 'sections': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 2, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'stay': {'Q1-Dataset/pg1513.txt': 26, 'Q1-Dataset/2160-0.txt': 25, 'Q1-Dataset/pg11.txt': 5, 'Q1-Dataset/pg2600.txt': 57, 'Q1-Dataset/pg345.txt': 43, 'Q1-Dataset/pg64317.txt': 6, 'Q1-Dataset/pg70815.txt': 7, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 8}, 'old': {'Q1-Dataset/pg1513.txt': 30, 'Q1-Dataset/2160-0.txt': 134, 'Q1-Dataset/pg11.txt': 21, 'Q1-Dataset/pg2600.txt': 803, 'Q1-Dataset/pg345.txt': 188, 'Q1-Dataset/pg64317.txt': 83, 'Q1-Dataset/pg70815.txt': 10, 'Q1-Dataset/1080-0.txt': 10, 'Q1-Dataset/pg70817.txt': 28, 'Q1-Dataset/pg1998.txt': 170}, 'outdated': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'costs': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 4, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 2, 'Q1-Dataset/pg345.txt': 2, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 2}, 'trademark': {'Q1-Dataset/pg1513.txt': 13, 'Q1-Dataset/2160-0.txt': 10, 'Q1-Dataset/pg11.txt': 13, 'Q1-Dataset/pg2600.txt': 13, 'Q1-Dataset/pg345.txt': 13, 'Q1-Dataset/pg64317.txt': 13, 'Q1-Dataset/pg70815.txt': 13, 'Q1-Dataset/1080-0.txt': 13, 'Q1-Dataset/pg70817.txt': 13, 'Q1-Dataset/pg1998.txt': 13}, 'what': {'Q1-Dataset/pg1513.txt': 162, 'Q1-Dataset/2160-0.txt': 263, 'Q1-Dataset/pg11.txt': 130, 'Q1-Dataset/pg2600.txt': 2315, 'Q1-Dataset/pg345.txt': 460, 'Q1-Dataset/pg64317.txt': 128, 'Q1-Dataset/pg70815.txt': 175, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 35, 'Q1-Dataset/pg1998.txt': 429}, 'well': {'Q1-Dataset/pg1513.txt': 68, 'Q1-Dataset/2160-0.txt': 175, 'Q1-Dataset/pg11.txt': 60, 'Q1-Dataset/pg2600.txt': 735, 'Q1-Dataset/pg345.txt': 236, 'Q1-Dataset/pg64317.txt': 66, 'Q1-Dataset/pg70815.txt': 75, 'Q1-Dataset/1080-0.txt': 8, 'Q1-Dataset/pg70817.txt': 26, 'Q1-Dataset/pg1998.txt': 170}, 'part': {'Q1-Dataset/pg1513.txt': 28, 'Q1-Dataset/2160-0.txt': 136, 'Q1-Dataset/pg11.txt': 6, 'Q1-Dataset/pg2600.txt': 236, 'Q1-Dataset/pg345.txt': 62, 'Q1-Dataset/pg64317.txt': 13, 'Q1-Dataset/pg70815.txt': 23, 'Q1-Dataset/1080-0.txt': 7, 'Q1-Dataset/pg70817.txt': 9, 'Q1-Dataset/pg1998.txt': 50}, 'from': {'Q1-Dataset/pg1513.txt': 101, 'Q1-Dataset/2160-0.txt': 661, 'Q1-Dataset/pg11.txt': 52, 'Q1-Dataset/pg2600.txt': 2700, 'Q1-Dataset/pg345.txt': 637, 'Q1-Dataset/pg64317.txt': 247, 'Q1-Dataset/pg70815.txt': 501, 'Q1-Dataset/1080-0.txt': 29, 'Q1-Dataset/pg70817.txt': 92, 'Q1-Dataset/pg1998.txt': 45}, 'shared': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 13, 'Q1-Dataset/pg345.txt': 3, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 3, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 2}, 'void': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 5, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 2, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'with': {'Q1-Dataset/pg1513.txt': 299, 'Q1-Dataset/2160-0.txt': 1702, 'Q1-Dataset/pg11.txt': 227, 'Q1-Dataset/pg2600.txt': 5694, 'Q1-Dataset/pg345.txt': 1321, 'Q1-Dataset/pg64317.txt': 465, 'Q1-Dataset/pg70815.txt': 612, 'Q1-Dataset/1080-0.txt': 59, 'Q1-Dataset/pg70817.txt': 222, 'Q1-Dataset/pg1998.txt': 866}, 'arms': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 26, 'Q1-Dataset/pg11.txt': 6, 'Q1-Dataset/pg2600.txt': 177, 'Q1-Dataset/pg345.txt': 33, 'Q1-Dataset/pg64317.txt': 13, 'Q1-Dataset/pg70815.txt': 11, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 6, 'Q1-Dataset/pg1998.txt': 7}, 'marked': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 9, 'Q1-Dataset/pg11.txt': 8, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}}
```

t': 7, 'Q1-Dataset/pg345.txt': 15, 'Q1-Dataset/pg64317.txt': 3, 'Q1-Dataset/pg70815.txt': 12, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 3, 'Q1-Dataset/pg1998.txt': 2}, 'unlink': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'many': {'Q1-Dataset/pg1513.txt': 16, 'Q1-Dataset/2160-0.txt': 98, 'Q1-Dataset/pg11.txt': 14, 'Q1-Dataset/pg2600.txt': 233, 'Q1-Dataset/pg345.txt': 104, 'Q1-Dataset/pg64317.txt': 17, 'Q1-Dataset/pg70815.txt': 66, 'Q1-Dataset/1080-0.txt': 11, 'Q1-Dataset/pg70817.txt': 11, 'Q1-Dataset/pg1998.txt': 168}, 'too': {'Q1-Dataset/pg1513.txt': 56, 'Q1-Dataset/2160-0.txt': 70, 'Q1-Dataset/pg11.txt': 26, 'Q1-Dataset/pg2600.txt': 389, 'Q1-Dataset/pg345.txt': 209, 'Q1-Dataset/pg64317.txt': 60, 'Q1-Dataset/pg70815.txt': 42, 'Q1-Dataset/1080-0.txt': 5, 'Q1-Dataset/pg70817.txt': 16, 'Q1-Dataset/pg1998.txt': 173}, 'man': {'Q1-Dataset/pg1513.txt': 65, 'Q1-Dataset/2160-0.txt': 170, 'Q1-Dataset/pg11.txt': 5, 'Q1-Dataset/pg2600.txt': 1064, 'Q1-Dataset/pg345.txt': 216, 'Q1-Dataset/pg64317.txt': 106, 'Q1-Dataset/pg70815.txt': 168, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 34, 'Q1-Dataset/pg1998.txt': 377}, 'access': {'Q1-Dataset/pg1513.txt': 11, 'Q1-Dataset/2160-0.txt': 11, 'Q1-Dataset/pg11.txt': 10, 'Q1-Dataset/pg2600.txt': 11, 'Q1-Dataset/pg345.txt': 11, 'Q1-Dataset/pg64317.txt': 10, 'Q1-Dataset/pg70815.txt': 10, 'Q1-Dataset/1080-0.txt': 10, 'Q1-Dataset/pg70817.txt': 11, 'Q1-Dataset/pg1998.txt': 11}, 'request': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 15, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 20, 'Q1-Dataset/pg345.txt': 8, 'Q1-Dataset/pg64317.txt': 6, 'Q1-Dataset/pg70815.txt': 3, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'exclusion': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 2, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'discovered': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 22, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 14, 'Q1-Dataset/pg345.txt': 7, 'Q1-Dataset/pg64317.txt': 6, 'Q1-Dataset/pg70815.txt': 31, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 5, 'Q1-Dataset/pg1998.txt': 11}, 'display': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 4, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 10, 'Q1-Dataset/pg345.txt': 2, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 2}, 'defect': {'Q1-Dataset/pg1513.txt': 3, 'Q1-Dataset/2160-0.txt': 10, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.txt': 7, 'Q1-Dataset/pg345.txt': 4, 'Q1-Dataset/pg64317.txt': 3, 'Q1-Dataset/pg70815.txt': 3, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 3, 'Q1-Dataset/pg1998.txt': 4}, 'course': {'Q1-Dataset/pg1513.txt': 5, 'Q1-Dataset/2160-0.txt': 44, 'Q1-Dataset/pg11.txt': 25, 'Q1-Dataset/pg2600.txt': 123, 'Q1-Dataset/pg345.txt': 41, 'Q1-Dataset/pg64317.txt': 23, 'Q1-Dataset/pg70815.txt': 35, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 10, 'Q1-Dataset/pg1998.txt': 23}, 'our': {'Q1-Dataset/pg1513.txt': 64, 'Q1-Dataset/2160-0.txt': 409, 'Q1-Dataset/pg11.txt': 12, 'Q1-Dataset/pg2600.txt': 624, 'Q1-Dataset/pg345.txt': 409, 'Q1-Dataset/pg64317.txt': 28, 'Q1-Dataset/pg70815.txt': 80, 'Q1-Dataset/1080-0.txt': 26, 'Q1-Dataset/pg70817.txt': 18, 'Q1-Dataset/pg1998.txt': 115}, 'either': {'Q1-Dataset/pg1513.txt': 10, 'Q1-Dataset/2160-0.txt': 52, 'Q1-Dataset/pg11.txt': 11, 'Q1-Dataset/pg2600.txt': 127, 'Q1-Dataset/pg345.txt': 39, 'Q1-Dataset/pg64317.txt': 12, 'Q1-Dataset/pg70815.txt': 40, 'Q1-Dataset/1080-0.txt': 4, 'Q1-Dataset/pg70817.txt': 4, 'Q1-Dataset/pg1998.txt': 17}, 'reasonable': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 5, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 13, 'Q1-Dataset/pg345.txt': 4, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 8, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 3, 'Q1-Dataset/pg1998.txt': 1}, 'six': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 22, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 56, 'Q1-Dataset/pg345.txt': 12, 'Q1-Dataset/pg64317.txt': 6, 'Q1-Dataset/pg70815.txt': 14, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 5, 'Q1-Dataset/pg1998.txt': 1}, 'defects': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 3, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 4, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 2}, 'originator': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 3, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'addition': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 6, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 5, 'Q1-Dataset/pg345.txt': 3, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 10, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 4, 'Q1-Dataset/pg1998.txt': 2}, 'those': {'Q1-Dataset/pg1513.txt': 16, 'Q1-Dataset/2160-0.txt': 175, 'Q1-Dataset/pg11.txt': 11, 'Q1-Dataset/pg2600.txt': 651, 'Q1-Dataset/pg345.txt': 83, 'Q1-Dataset/pg64317.txt': 28, 'Q1-Dataset/pg70815.txt': 56, 'Q1-Dataset/1080-0.txt': 8, 'Q1-Da

taset/pg70817.txt': 20, 'Q1-Dataset/pg1998.txt': 200}, 'important': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 3, 'Q1-Dataset/pg11.txt': 6, 'Q1-Dataset/pg2600.txt': 129, 'Q1-Dataset/pg345.txt': 12, 'Q1-Dataset/pg64317.txt': 4, 'Q1-Dataset/pg70815.txt': 5, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 19}, 'including': {'Q1-Dataset/pg1513.txt': 9, 'Q1-Dataset/2160-0.txt': 17, 'Q1-Dataset/pg11.txt': 9, 'Q1-Dataset/pg2600.txt': 16, 'Q1-Dataset/pg345.txt': 10, 'Q1-Dataset/pg64317.txt': 12, 'Q1-Dataset/pg70815.txt': 11, 'Q1-Dataset/1080-0.txt': 9, 'Q1-Dataset/pg70817.txt': 9, 'Q1-Dataset/pg1998.txt': 10}, 'owed': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 4, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 4, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 2}, 'grown': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 5, 'Q1-Dataset/pg11.txt': 7, 'Q1-Dataset/pg2600.txt': 65, 'Q1-Dataset/pg345.txt': 11, 'Q1-Dataset/pg64317.txt': 3, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 8}, 'infringement': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 2, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'keeping': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 9, 'Q1-Dataset/pg11.txt': 4, 'Q1-Dataset/pg2600.txt': 35, 'Q1-Dataset/pg345.txt': 18, 'Q1-Dataset/pg64317.txt': 5, 'Q1-Dataset/pg70815.txt': 8, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 3, 'Q1-Dataset/pg1998.txt': 3}, 'damages': {'Q1-Dataset/pg1513.txt': 4, 'Q1-Dataset/2160-0.txt': 6, 'Q1-Dataset/pg11.txt': 4, 'Q1-Dataset/pg2600.txt': 4, 'Q1-Dataset/pg345.txt': 5, 'Q1-Dataset/pg64317.txt': 4, 'Q1-Dataset/pg70815.txt': 4, 'Q1-Dataset/1080-0.txt': 4, 'Q1-Dataset/pg70817.txt': 4, 'Q1-Dataset/pg1998.txt': 4}, 'thing': {'Q1-Dataset/pg1513.txt': 12, 'Q1-Dataset/2160-0.txt': 79, 'Q1-Dataset/pg11.txt': 49, 'Q1-Dataset/pg2600.txt': 224, 'Q1-Dataset/pg345.txt': 97, 'Q1-Dataset/pg64317.txt': 35, 'Q1-Dataset/pg70815.txt': 31, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 15, 'Q1-Dataset/pg1998.txt': 86}, 'hundreds': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 4, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 25, 'Q1-Dataset/pg345.txt': 5, 'Q1-Dataset/pg64317.txt': 5, 'Q1-Dataset/pg70815.txt': 9, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 3, 'Q1-Dataset/pg1998.txt': 1}, 'corporation': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 3, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'least': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 66, 'Q1-Dataset/pg11.txt': 8, 'Q1-Dataset/pg2600.txt': 84, 'Q1-Dataset/pg345.txt': 39, 'Q1-Dataset/pg64317.txt': 15, 'Q1-Dataset/pg70815.txt': 21, 'Q1-Dataset/1080-0.txt': 7, 'Q1-Dataset/pg70817.txt': 10, 'Q1-Dataset/pg1998.txt': 22}, 'detach': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 2, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'having': {'Q1-Dataset/pg1513.txt': 8, 'Q1-Dataset/2160-0.txt': 143, 'Q1-Dataset/pg11.txt': 10, 'Q1-Dataset/pg2600.txt': 487, 'Q1-Dataset/pg345.txt': 32, 'Q1-Dataset/pg64317.txt': 11, 'Q1-Dataset/pg70815.txt': 26, 'Q1-Dataset/1080-0.txt': 5, 'Q1-Dataset/pg70817.txt': 4, 'Q1-Dataset/pg1998.txt': 8}, 'information': {'Q1-Dataset/pg1513.txt': 8, 'Q1-Dataset/2160-0.txt': 25, 'Q1-Dataset/pg11.txt': 8, 'Q1-Dataset/pg2600.txt': 29, 'Q1-Dataset/pg345.txt': 18, 'Q1-Dataset/pg64317.txt': 13, 'Q1-Dataset/pg70815.txt': 10, 'Q1-Dataset/1080-0.txt': 8, 'Q1-Dataset/pg70817.txt': 8, 'Q1-Dataset/pg1998.txt': 8}, 'can': {'Q1-Dataset/pg1513.txt': 49, 'Q1-Dataset/2160-0.txt': 139, 'Q1-Dataset/pg11.txt': 40, 'Q1-Dataset/pg2600.txt': 473, 'Q1-Dataset/pg345.txt': 384, 'Q1-Dataset/pg64317.txt': 40, 'Q1-Dataset/pg70815.txt': 75, 'Q1-Dataset/1080-0.txt': 19, 'Q1-Dataset/pg70817.txt': 19, 'Q1-Dataset/pg1998.txt': 112}, 'terms': {'Q1-Dataset/pg1513.txt': 23, 'Q1-Dataset/2160-0.txt': 38, 'Q1-Dataset/pg11.txt': 23, 'Q1-Dataset/pg2600.txt': 38, 'Q1-Dataset/pg345.txt': 24, 'Q1-Dataset/pg64317.txt': 24, 'Q1-Dataset/pg70815.txt': 22, 'Q1-Dataset/1080-0.txt': 22, 'Q1-Dataset/pg70817.txt': 22, 'Q1-Dataset/pg1998.txt': 25}, 'paper': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 15, 'Q1-Dataset/pg11.txt': 5, 'Q1-Dataset/pg2600.txt': 58, 'Q1-Dataset/pg345.txt': 24, 'Q1-Dataset/pg64317.txt': 4, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 5, 'Q1-Dataset/pg1998.txt': 2}, 'a': {'Q1-Dataset/pg1513.txt': 541, 'Q1-Dataset/2160-0.txt': 3649, 'Q1-Dataset/pg11.txt': 689, 'Q1-Dataset/pg2600.txt': 10493, 'Q1-Dataset/pg345.txt': 2958, 'Q1-Dataset/pg64317.txt': 1451, 'Q1-Dataset/pg70815.txt': 1555, 'Q1-Dataset/1080-0.txt': 152, 'Q1-Dataset/pg70817.txt': 722, 'Q1-Dataset/pg1998.txt': 1797}, 'printed': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 5, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.txt': 9, 'Q1-Dataset/pg345.txt': 3, 'Q1-Dataset/pg64317.txt': 4, 'Q1-Dataset/pg70815.txt': 3, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 4}, 'fair': {'Q1-Dataset/pg1513.txt': 44, 'Q1-Dataset/2160-0.txt': 21, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}

set/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 15, 'Q1-Dataset/pg345.txt': 23, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 3, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 5}, 'better': {'Q1-Dataset/pg1513.txt': 12, 'Q1-Dataset/2160-0.txt': 46, 'Q1-Dataset/pg11.txt': 14, 'Q1-Dataset/pg2600.txt': 162, 'Q1-Dataset/pg345.txt': 77, 'Q1-Dataset/pg64317.txt': 18, 'Q1-Dataset/pg70815.txt': 29, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 11, 'Q1-Dataset/pg1998.txt': 60}, 'solicitation': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'doubt': {'Q1-Dataset/pg1513.txt': 4, 'Q1-Dataset/2160-0.txt': 56, 'Q1-Dataset/pg11.txt': 4, 'Q1-Dataset/pg2600.txt': 60, 'Q1-Dataset/pg345.txt': 41, 'Q1-Dataset/pg64317.txt': 5, 'Q1-Dataset/pg70815.txt': 7, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 5, 'Q1-Dataset/pg1998.txt': 14}, 'computer': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 2, 'Q1-Dataset/pg345.txt': 2, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 2}, 'royalties': {'Q1-Dataset/pg1513.txt': 3, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.txt': 3, 'Q1-Dataset/pg345.txt': 3, 'Q1-Dataset/pg64317.txt': 3, 'Q1-Dataset/pg70815.txt': 3, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 3, 'Q1-Dataset/pg1998.txt': 3}, 'contributions': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 2, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 5, 'Q1-Dataset/pg345.txt': 2, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 2}, 'derivative': {'Q1-Dataset/pg1513.txt': 3, 'Q1-Dataset/2160-0.txt': 3, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.txt': 3, 'Q1-Dataset/pg345.txt': 3, 'Q1-Dataset/pg64317.txt': 3, 'Q1-Dataset/pg70815.txt': 3, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 3, 'Q1-Dataset/pg1998.txt': 3}, 'ways': {'Q1-Dataset/pg1513.txt': 3, 'Q1-Dataset/2160-0.txt': 7, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 16, 'Q1-Dataset/pg345.txt': 20, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 5, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 12}, 'she': {'Q1-Dataset/pg1513.txt': 110, 'Q1-Dataset/2160-0.txt': 492, 'Q1-Dataset/pg11.txt': 535, 'Q1-Dataset/pg2600.txt': 3403, 'Q1-Dataset/pg345.txt': 811, 'Q1-Dataset/pg64317.txt': 371, 'Q1-Dataset/pg70815.txt': 333, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 286, 'Q1-Dataset/pg1998.txt': 47}, 'statements': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 3, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'keep': {'Q1-Dataset/pg1513.txt': 18, 'Q1-Dataset/2160-0.txt': 59, 'Q1-Dataset/pg11.txt': 13, 'Q1-Dataset/pg2600.txt': 83, 'Q1-Dataset/pg345.txt': 98, 'Q1-Dataset/pg64317.txt': 13, 'Q1-Dataset/pg70815.txt': 12, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 11, 'Q1-Dataset/pg1998.txt': 29}, 'who': {'Q1-Dataset/pg1513.txt': 36, 'Q1-Dataset/2160-0.txt': 611, 'Q1-Dataset/pg11.txt': 65, 'Q1-Dataset/pg2600.txt': 2116, 'Q1-Dataset/pg345.txt': 285, 'Q1-Dataset/pg64317.txt': 113, 'Q1-Dataset/pg70815.txt': 166, 'Q1-Dataset/1080-0.txt': 27, 'Q1-Dataset/pg70817.txt': 20, 'Q1-Dataset/pg1998.txt': 538}, 'notifies': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 1, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 1, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'day': {'Q1-Dataset/pg1513.txt': 45, 'Q1-Dataset/2160-0.txt': 125, 'Q1-Dataset/pg11.txt': 25, 'Q1-Dataset/pg2600.txt': 543, 'Q1-Dataset/pg345.txt': 156, 'Q1-Dataset/pg64317.txt': 43, 'Q1-Dataset/pg70815.txt': 14, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 21, 'Q1-Dataset/pg1998.txt': 106}, 'giving': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 22, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 107, 'Q1-Dataset/pg345.txt': 5, 'Q1-Dataset/pg64317.txt': 3, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 11}, 'works': {'Q1-Dataset/pg1513.txt': 32, 'Q1-Dataset/2160-0.txt': 47, 'Q1-Dataset/pg11.txt': 33, 'Q1-Dataset/pg2600.txt': 37, 'Q1-Dataset/pg345.txt': 34, 'Q1-Dataset/pg64317.txt': 33, 'Q1-Dataset/pg70815.txt': 32, 'Q1-Dataset/1080-0.txt': 32, 'Q1-Dataset/pg70817.txt': 32, 'Q1-Dataset/pg1998.txt': 44}, 'containing': {'Q1-Dataset/pg1513.txt': 2, 'Q1-Dataset/2160-0.txt': 6, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 14, 'Q1-Dataset/pg345.txt': 2, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 3, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'attached': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 3, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 23, 'Q1-Dataset/pg345.txt': 2, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 4, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 4}, 'need': {'Q1-Dataset/pg1513.txt': 11, 'Q1-Dataset/2160-0.txt': 12, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 108, 'Q1-Dataset/pg345.txt': 69, 'Q1-Dataset/pg64317.txt': 4, 'Q1-Dataset/pg7081

5.txt': 14, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 22}, 'paragraphs': {'Q1-Dataset/pg1513.txt': 3, 'Q1-Dataset/2160-0.txt': 3, 'Q1-Dataset/pg11.txt': 3, 'Q1-Dataset/pg2600.txt': 5, 'Q1-Dataset/pg345.txt': 3, 'Q1-Dataset/pg64317.txt': 3, 'Q1-Dataset/pg70815.txt': 3, 'Q1-Dataset/1080-0.txt': 3, 'Q1-Dataset/pg70817.txt': 4, 'Q1-Dataset/pg1998.txt': 8}, 'through': {'Q1-Dataset/pg1513.txt': 10, 'Q1-Dataset/2160-0.txt': 75, 'Q1-Dataset/pg11.txt': 15, 'Q1-Dataset/pg2600.txt': 378, 'Q1-Dataset/pg345.txt': 214, 'Q1-Dataset/pg64317.txt': 68, 'Q1-Dataset/pg70815.txt': 107, 'Q1-Dataset/1080-0.txt': 5, 'Q1-Dataset/pg70817.txt': 43, 'Q1-Dataset/pg1998.txt': 76}, 'described': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 5, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 31, 'Q1-Dataset/pg345.txt': 7, 'Q1-Dataset/pg64317.txt': 4, 'Q1-Dataset/pg70815.txt': 2, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 3}, 'professor': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 4, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 2, 'Q1-Dataset/pg345.txt': 156, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 1, 'Q1-Dataset/pg1998.txt': 1}, 'damaged': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 5, 'Q1-Dataset/pg11.txt': 1, 'Q1-Dataset/pg2600.txt': 3, 'Q1-Dataset/pg345.txt': 1, 'Q1-Dataset/pg64317.txt': 1, 'Q1-Dataset/pg70815.txt': 1, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 1}, 'talk': {'Q1-Dataset/pg1513.txt': 12, 'Q1-Dataset/2160-0.txt': 11, 'Q1-Dataset/pg11.txt': 14, 'Q1-Dataset/pg2600.txt': 251, 'Q1-Dataset/pg345.txt': 34, 'Q1-Dataset/pg64317.txt': 12, 'Q1-Dataset/pg70815.txt': 5, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 3, 'Q1-Dataset/pg1998.txt': 34}, 'receive': {'Q1-Dataset/pg1513.txt': 3, 'Q1-Dataset/2160-0.txt': 24, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 43, 'Q1-Dataset/pg345.txt': 6, 'Q1-Dataset/pg64317.txt': 2, 'Q1-Dataset/pg70815.txt': 5, 'Q1-Dataset/1080-0.txt': 2, 'Q1-Dataset/pg70817.txt': 2, 'Q1-Dataset/pg1998.txt': 5}, 'party': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 14, 'Q1-Dataset/pg11.txt': 9, 'Q1-Dataset/pg2600.txt': 70, 'Q1-Dataset/pg345.txt': 12, 'Q1-Dataset/pg64317.txt': 19, 'Q1-Dataset/pg70815.txt': 43, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 7, 'Q1-Dataset/pg1998.txt': 3}, 'given': {'Q1-Dataset/pg1513.txt': 1, 'Q1-Dataset/2160-0.txt': 54, 'Q1-Dataset/pg11.txt': 2, 'Q1-Dataset/pg2600.txt': 179, 'Q1-Dataset/pg345.txt': 43, 'Q1-Dataset/pg64317.txt': 8, 'Q1-Dataset/pg70815.txt': 15, 'Q1-Dataset/1080-0.txt': 1, 'Q1-Dataset/pg70817.txt': 4, 'Q1-Dataset/pg1998.txt': 37}}

```
In [23]: print('Frequencies of the words in tabuler format')
words = pd.DataFrame.from_dict(common_words_frequency_dict, orient = 'columns')
display(words)
words.to_excel('words.xlsx')
```

Frequencies of the words in tabuler format

	rather	references	two	immediate	family	of	accessible	ago	away	official	...	need
Q1-Dataset/pg1513.txt	4	2	20	2	2	518	1	2	25	3	...	11
Q1-Dataset/2160-0.txt	43	2	144	12	113	5676	2	45	51	3	...	12
Q1-Dataset/pg11.txt	25	2	37	3	1	632	1	2	27	3	...	1
Q1-Dataset/pg2600.txt	88	2	587	10	135	14994	4	74	613	33	...	108
Q1-Dataset/pg345.txt	25	2	128	6	3	3746	1	26	187	6	...	69
Q1-Dataset/pg64317.txt	25	2	72	5	13	1233	1	19	76	4	...	4
Q1-Dataset/pg70815.txt	15	2	118	7	1	3108	1	6	71	4	...	14
Q1-Dataset/1080-0.txt	2	2	8	2	2	261	1	1	3	3	...	1
Q1-Dataset/pg70817.txt	10	2	26	6	3	769	1	8	35	3	...	2

10 rows × 816 columns

```
In [24]: spark_context.stop()
```

Conclusion of Question 2:

To consider as a valid word we are not considering any special character and number, hence we are considering only the alphabet because as per definition of word it is defined as, a sound or letter or group of sounds or letters that expresses a particular meaning.

reference:<https://www.google.com/search?client=firefox-b-d&q=words+definition+>

There are 816 common words available and their corresponding frequency of occurrence in different files are displayed in tabular format.

Question: 03

Perform an exploratory data analysis on the primary_data.csv of the Mushroom dataset [Secondary Mushroom Dataset](#) using pandas and answer the following questions:

- How many mushrooms are poisonous?
- Which family contains most poisonous mushrooms?
- Do all poisonous mushrooms have a ring?
- Which seasons do poisonous mushrooms grow typically?

Answer: 03

Importing Libraries

```
In [25]: import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import os
import urllib.request
```

Common Variables

```
In [26]: working_dir_path = '/content/sample_data/Assignment 1/Q3-Dataset/MushroomDataset' #'C:/Us
primaryFileName = 'primary_data.csv'
Q3basePath = '/content/sample_data/Assignment 1'
```

File and folder operation

```
In [27]: os.chdir(Q3basePath)
!mkdir Q3-Dataset
os.chdir(Q3basePath+'/Q3-Dataset/')
```

Uploading the necessary file inside the Q3-Dataset/MushroomDataset folder before proceeding for further operation

```
In [28]: url_mushroomdataset = 'https://archive.ics.uci.edu/ml/machine-learning-databases/00615/MushroomDataset.zip'
urllib.request.urlretrieve(url_mushroomdataset, 'MushroomDataset.zip')
!unzip MushroomDataset.zip
```

```
Archive:  MushroomDataset.zip
  creating: MushroomDataset/
  inflating: MushroomDataset/primary_data.csv
  inflating: MushroomDataset/primary_data_meta.txt
  inflating: MushroomDataset/secondary_data.csv
  inflating: MushroomDataset/secondary_data_meta.txt
```

Change the folder location

```
In [29]: os.chdir(working_dir_path)
```

Loading primary data

```
In [30]: df_primary_data = pd.read_csv(primaryFileName, sep=';')
```

```
In [31]: df_primary_data
```

Out[31]:

	family	name	class	cap-diameter	cap-shape	Cap-surface	cap-color	does-bruise-or-bleed	gill-attachment	gill-spacing	...	stem-root	stem-surface
0	Amanita Family	Fly Agaric	p	[10, 20]	[x, f]	[g, h]	[e, o]	[f]	[e]	NaN	...	[s]	[y]
1	Amanita Family	Panther Cap	p	[5, 10]	[p, x]	[g]	[n]	[f]	[e]	NaN	...	NaN	[y]
2	Amanita Family	False Panther Cap	p	[10, 15]	[x, f]	NaN	[g, n]	[f]	[e]	NaN	...	NaN	NaN
3	Amanita Family	The Blusher	e	[5, 15]	[x, f]	NaN	[n]	[t]	NaN	NaN	...	[b]	NaN
4	Amanita Family	Death Cap	p	[5, 12]	[x, f]	[h]	[r]	[f]	NaN	[c]	...	NaN	NaN
...
168	Saddle-Cup Family	White Saddle	p	[2, 6]	[o]	NaN	[w]	[f]	[f]	[f]	...	NaN	[g]
169	Bracket Fungi	Elfin's Saddle	p	[2, 4]	[o]	NaN	[g, k]	[f]	[f]	[f]	...	NaN	[g]
170	Bracket Fungi	Turban Fungus	p	[5, 15]	[o]	NaN	[n, e]	[f]	[f]	[f]	...	NaN	[g]
171	Morel Family	Common Morel	e	[3, 8]	[p, c, o]	NaN	[n, g]	[f]	[f]	[f]	...	NaN	NaN
172	Jelly Discs Family	Jelly Babies	p	[1, 1.5]	[x, f, s]	[s]	[y]	[f]	[f]	[f]	...	NaN	NaN

173 rows × 23 columns

Perform basic statistical analysis and missing value check

```
In [32]: df_primary_data.describe(include='object')
```

Out[32]:

	family	name	class	cap-diameter	cap-shape	Cap-surface	cap-color	does-bruise-or-bleed	gill-attachment	gill-spacing	...	stem-root	stem-surface
count	173	173	173	173	173	133	173	173	145	102	...	27	
unique	23	173	2	51	27	40	67	2	8	3	...	5	
top	Tricholoma Family	Fly Agaric	p	[2, 5]	[x]	[y]	[n]	[f]	[a]	[c]	...	[s]	
freq	43	1	96	16	48	14	38	143	32	70	...	9	

4 rows × 23 columns

```
In [33]: df_primary_data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 173 entries, 0 to 172
Data columns (total 23 columns):
#   Column                                Non-Null Count  Dtype
---  -
0   family                                173 non-null    object
1   name                                  173 non-null    object
2   class                                 173 non-null    object
3   cap-diameter                          173 non-null    object
4   cap-shape                             173 non-null    object
5   Cap-surface                           133 non-null    object
6   cap-color                             173 non-null    object
7   does-bruise-or-bleed                  173 non-null    object
8   gill-attachment                       145 non-null    object
9   gill-spacing                           102 non-null    object
10  gill-color                             173 non-null    object
11  stem-height                           173 non-null    object
12  stem-width                             173 non-null    object
13  stem-root                             27 non-null     object
14  stem-surface                           65 non-null     object
15  stem-color                             173 non-null    object
16  veil-type                              9 non-null      object
17  veil-color                             21 non-null     object
18  has-ring                               173 non-null    object
19  ring-type                              166 non-null    object
20  Spore-print-color                      18 non-null     object
21  habitat                                173 non-null    object
22  season                                 173 non-null    object
dtypes: object(23)
memory usage: 31.2+ KB
```

```
In [34]: df_primary_data.isna().sum().sort_values(ascending=False)
```

Out[34]:

veil-type	164
Spore-print-color	155
veil-color	152
stem-root	146
stem-surface	108
gill-spacing	71


```

Cap-surface          40
gill-attachment      28
ring-type            7
family               0
habitat              0
has-ring             0
stem-color           0
stem-height          0
stem-width           0
name                 0
gill-color           0
does-bruise-or-bleed 0
cap-color            0
cap-shape            0
cap-diameter         0
class                0
season               0
dtype: int64

```

```

In [35]: poisonous_mushroom = df_primary_data[df_primary_data['class'] == 'p']
         edible_mushroom = df_primary_data[df_primary_data['class'] == 'e']

```

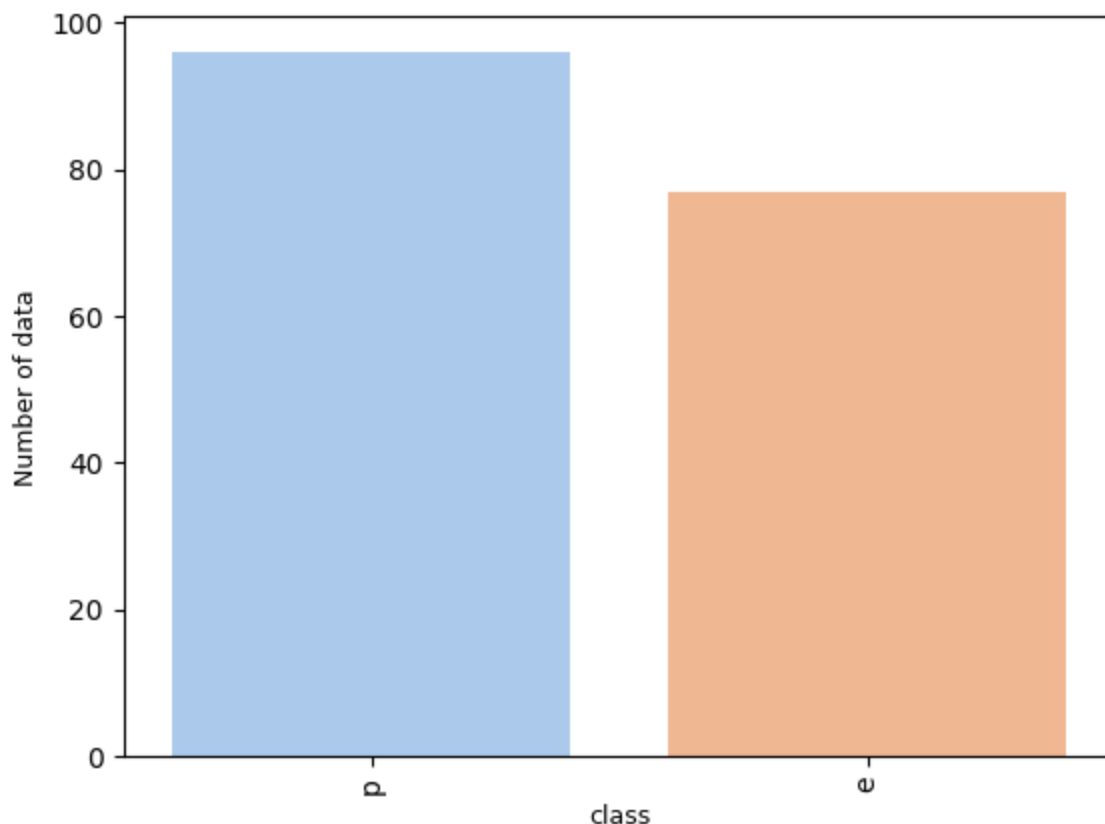
Graphical representation of the distribution of poisonous and edible mushroom

```

In [36]: print('====Distribution of poisonous and edible mushroom====')
         mushroom_type_count= df_primary_data['class'].value_counts()
         sns.barplot(x=mushroom_type_count.index, y=mushroom_type_count.values,palette='pastel')
         plt.ylabel('Number of data', fontsize=9)
         plt.xlabel('class', fontsize=9)
         plt.xticks(rotation=90)
         plt.show();
         df_primary_data['class'].value_counts()

```

====Distribution of poisonous and edible mushroom=====



```

Out[36]: p      96
         e      77
         Name: class, dtype: int64

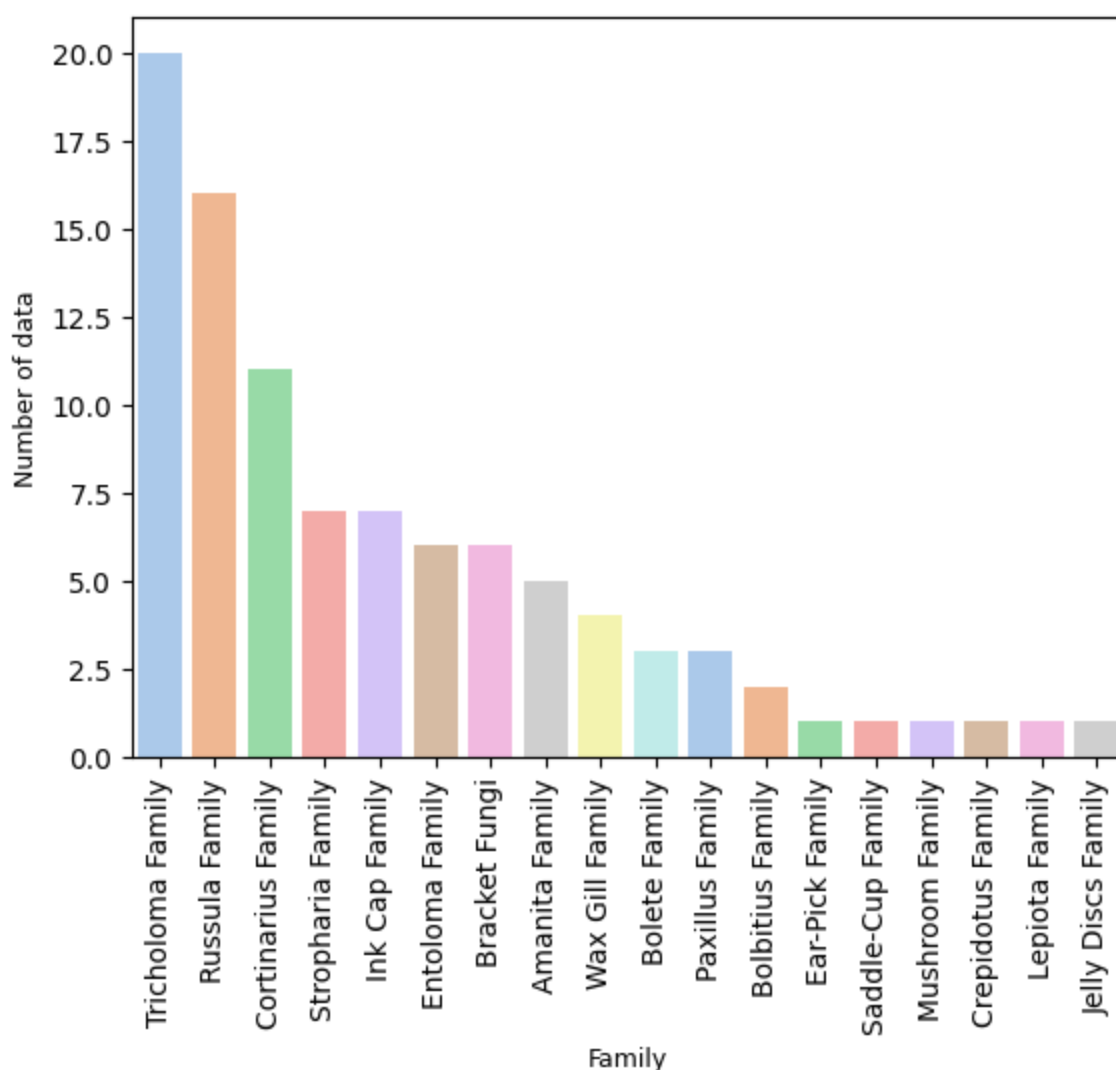
```

Q. How many mushrooms are poisonous?

A. 96 mushrooms are poisonous and 77 are edible

Graphical representation of the distribution of poisonous mushroom among different family

```
In [37]: poisonous_mushroom_type_count= poisonous_mushroom['family'].value_counts()
sns.barplot(x=poisonous_mushroom_type_count.index, y=poisonous_mushroom_type_count.value
plt.ylabel('Number of data', fontsize=9)
plt.xlabel('Family', fontsize=9)
plt.xticks(rotation=90)
plt.show();
df_primary_data['family'].value_counts()
```



```
Out[37]: Tricholoma Family      43
Russula Family      27
Bolete Family      14
Ink Cap Family      13
Cortinarius Family  11
Stropharia Family   8
Amanita Family       8
Wax Gill Family      8
Entoloma Family      7
Bracket Fungi        7
Mushroom Family      5
Bolbitius Family     3
Lepiota Family       3
Paxillus Family      3
Chanterelle Family   3
Pluteus Family       2
```

```
Oyster Mushroom Family      2
Crepidotus Family            1
Hydnum Family                 1
Ear-Pick Family               1
Saddle-Cup Family            1
Morel Family                  1
Jelly Discs Family            1
Name: family, dtype: int64
```

Q. Which family contains most poisonous mushrooms?

A. Tricholoma, mushroom family has highest number of poisonous mushrooms.

Analysis of poisonous mushroom based on its ring characteristics

```
In [38]: np.unique(poisonous_mushroom['has-ring'])
```

```
Out[38]: array(['f'], 't'], dtype=object)
```

It has been observed that the unique value of mushroom has-ring column is [f] & [t], which clearly indicate that only ring exists among the mushroom which has 'has-ring' column value as [t], so we will filter that out and get the desired result.

```
In [39]: ring_poisonous_mushroom = poisonous_mushroom[poisonous_mushroom['has-ring'] == '[t]']
```

```
In [40]: ring_poisonous_mushroom
```

```
Out[40]:
```

	family	name	class	cap-diameter	cap-shape	Cap-surface	cap-color	does-bruise-or-bleed	gill-attachment	gill-spacing	...	stem-root	st
0	Amanita Family	Fly Agaric	p	[10, 20]	[x, f]	[g, h]	[e, o]	[f]	[e]	NaN	...	[s]	
1	Amanita Family	Panther Cap	p	[5, 10]	[p, x]	[g]	[n]	[f]	[e]	NaN	...	NaN	
2	Amanita Family	False Panther Cap	p	[10, 15]	[x, f]	NaN	[g, n]	[f]	[e]	NaN	...	NaN	I
4	Amanita Family	Death Cap	p	[5, 12]	[x, f]	[h]	[r]	[f]	NaN	[c]	...	NaN	I
6	Amanita Family	Destroying Angel	p	[5, 10]	[b]	[t]	[w]	[f]	[e]	[c]	...	NaN	
10	Lepiota Family	Stinking Parasol	p	[2, 5]	[b, f]	[y]	[e, n, p, w]	[f]	[e]	[c]	...	NaN	
11	Tricholoma Family	Saffron Parasol	p	[2, 5]	[x]	[y]	[y, n]	[f]	[a]	NaN	...	NaN	
85	Russula Family	Woolly Milk Cap	p	[6, 12]	[s]	[t]	[p]	[t]	[d]	[c]	...	NaN	I
87	Russula Family	Spruce Milk Cap	p	[5, 12]	[s]	NaN	[o, y, r]	[t]	[d]	NaN	...	NaN	I
100	Stropharia Family	Shaggy Pholiota	p	[4, 12]	[x]	[y]	[y, n]	[f]	[a]	[c]	...	NaN	

101	Stropharia Family	Charcoal Pholiota	p	[2, 6]	[x]	[h, t]	[n, o]	[f]	[a]	[c]	...	NaN	I
103	Stropharia Family	Sulphur Tuft	p	[2, 6]	[b, x]	[s]	[y]	[f]	NaN	[c]	...	NaN	I
104	Stropharia Family	Brick Caps	p	[4, 8]	[x]	[e]	[e]	[f]	[a]	NaN	...	NaN	I
105	Stropharia Family	Verdigris Toadstool	p	[2, 6]	[b, x]	[h, t, y]	[r, l]	[f]	[s]	[c]	...	NaN	I
106	Stropharia Family	Dung Roundhead	p	[1, 3]	[p]	[t]	[n]	[f]	[a]	[c]	...	NaN	I
107	Stropharia Family	Magic Mushroom	p	[0.5, 1.5]	[b]	[t]	[n, y]	[f]	NaN	[c]	...	NaN	I
115	Cortinarius Family	Dark-centred Hebeloma	p	[2, 5]	[x]	[t]	[n]	[f]	[s]	[c]	...	NaN	I
116	Cortinarius Family	Rusty Wood Rotter	p	[4, 8]	[b, x, f]	[d, k]	[n]	[f]	[a, d]	[c]	...	NaN	I
117	Cortinarius Family	Spectacular Gymnopile	p	[6, 15]	[x]	[d, e, y, i]	[y, o]	[f]	[a]	NaN	...	[r]	I
122	Mushroom Family	Yellow-staining Mushroom	p	[5, 10]	[x]	[y]	[w]	[t]	NaN	NaN	...	[s]	I
125	Ink Cap Family	Common Ink Cap	p	[4, 8]	[p, b]	[s, y]	[g, n]	[f]	NaN	NaN	...	NaN	I
135	Ink Cap Family	Egg-shell Toadstool	p	[2, 6]	[b]	[t, h]	[n]	[f]	NaN	NaN	...	NaN	I
136	Ink Cap Family	Weeping Widow	p	[5, 10]	[x]	[i]	[n, y]	[f]	[a]	[c]	...	NaN	I
163	Bracket Fungi	Hairy Stereum	p	[2, 6]	[o]	[s]	[y, o, g, n, r]	[f]	[f]	[f]	...	[f]	I
167	Bracket Fungi	Brown Goblet	p	[3, 7]	[f, s]	[k]	[n, y]	[f]	[p]	NaN	...	NaN	I
168	Saddle-Cup Family	White Saddle	p	[2, 6]	[o]	NaN	[w]	[f]	[f]	[f]	...	NaN	I

26 rows × 23 columns

Q. Do all poisonous mushrooms have a ring?

A. No, all the poisonous mushroom does not have ring, only 70 out of 96 poisonous mushroom have no ring.

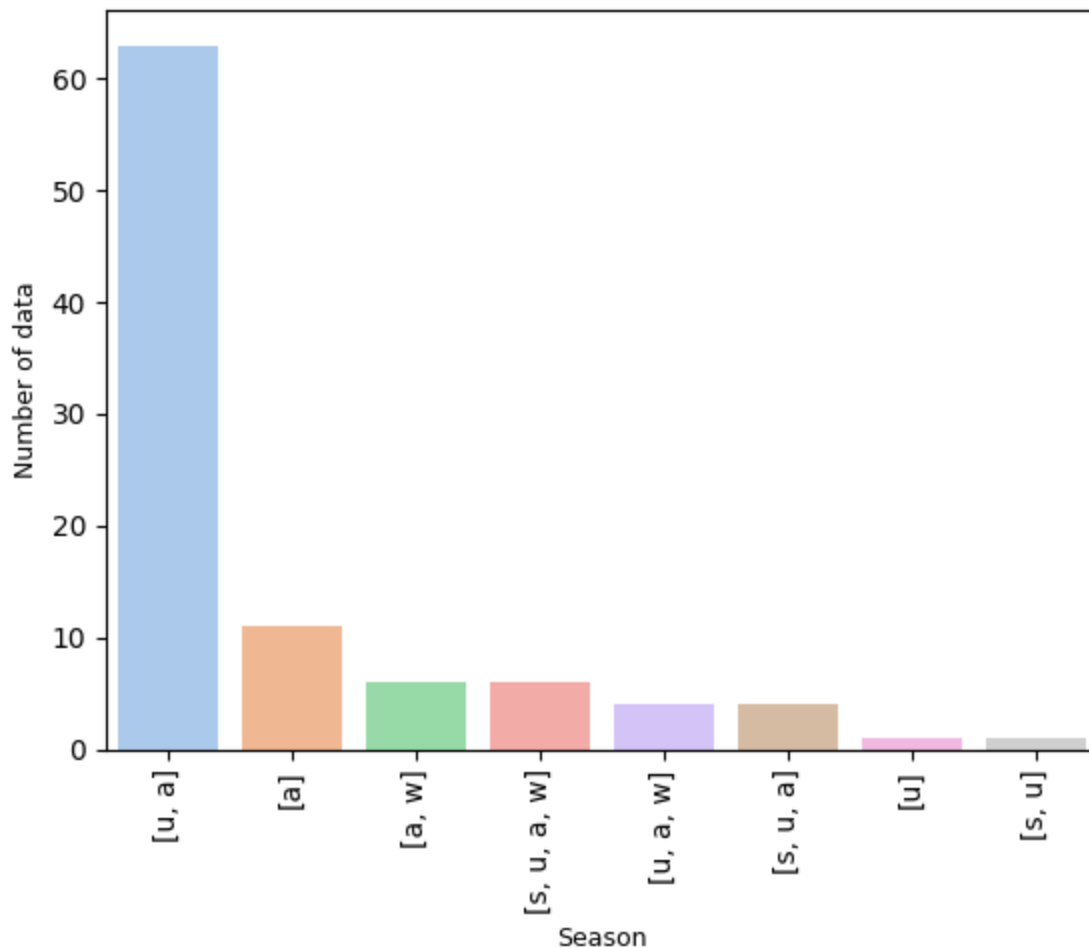
Analysis of poisonous mushroom based on the season

```
In [41]: np.unique(poisonous_mushroom['season'])

Out[41]: array(['[a, w]', '[a]', '[s, u, a, w]', '[s, u, a]', '[s, u]',
              '[u, a, w]', '[u, a]', '[u]'], dtype=object)
```

Graphical representation of the distribution of poisonous mushroom based on it season of growth

```
In [42]: poisonous_mushroom_type_count= poisonous_mushroom['season'].value_counts()
sns.barplot(x=poisonous_mushroom_type_count.index, y=poisonous_mushroom_type_count.value
plt.ylabel('Number of data', fontsize=9)
plt.xlabel('Season', fontsize=9)
plt.xticks(rotation=90)
plt.show();
df_primary_data['season'].value_counts()
```



```
Out[42]: [u, a]      106
[a]        16
[a, w]     15
[s, u, a, w] 13
[u, a, w]   12
[s, u, a]    5
[s, u]       3
[s, a, w]    1
[u]          1
[s]          1
Name: season, dtype: int64
```

Q. Which seasons do poisonous mushrooms grow typically?

A. Over Summer and Autumn most of the poisonous mushrooms typically grows.

Conclusion of Question 3:

Q. How many mushrooms are poisonous?

A. 96 mushrooms are poisonous and 77 are edible

Q. Which family contains most poisonous mushrooms?

A. Tricholoma, mushroom family has highest number of poisonous mushrooms.

Q. Do all poisonous mushrooms have a ring?

A. No, all the poisonous mushroom does not have ring, only 70 out of 96 poisonous mushroom have no ring.

Q. Which seasons do poisonous mushrooms grow typically?

A. Over Summer and Autumn most of the poisonous mushrooms typically grows.