Perrault vs. Grimm Fairy Tales

November 22, 2017

```
In [ ]: import nltk
        nltk.download()

In [1]: from datascience import *
        import numpy as np
        from collections import Counter
        from string import punctuation, digits
        from sklearn.feature_extraction.stop_words import ENGLISH_STOP_WORDS
        from sklearn.feature_extraction.text import CountVectorizer
        !pip install --no-cache-dir textblob
        from textblob import TextBlob
        import nltk
        from sklearn.feature_extraction.text import TfidfTransformer
        import matplotlib.pyplot as plt
        %matplotlib inline
        from sklearn.manifold import MDS
        from nltk.corpus import stopwords

Collecting textblob
  Downloading textblob-0.14.0-py2.py3-none-any.whl (631kB)
    100% || 634kB 15.6MB/s ta 0:00:01  4% |                          | 30kB 4.7MB/s eta 0:0
Requirement already satisfied: nltk>=3.1 in /srv/app/venv/lib/python3.6/site-packages (from te
Requirement already satisfied: six in /srv/app/venv/lib/python3.6/site-packages (from nltk>=3.
Installing collected packages: textblob
Successfully installed textblob-0.14.0


In [2]: from sklearn.cluster import KMeans
        from sklearn.metrics import pairwise

In [3]: stop_words = (list(ENGLISH_STOP_WORDS))

In [4]: with open('Perrault.csv', 'r') as f:
            perrault = f.read()
        perrault2 = perrault
        perrault

Out[4]: 'Title,Text\nThe Sleeping Beauty in the Wood,"\nOnce upon a time there lived a king an
```

```
In [5]: len_raw_p = len(perrault)
        len_raw_p

Out[5]: 96209

In [6]: perrault = perrault.replace("Title,Text", ""). strip()
        perrault

Out[6]: 'The Sleeping Beauty in the Wood,"\nOnce upon a time there lived a king and queen who u

In [7]: perrault1 = perrault.lower()
        perrault1 = ''.join([char for char in perrault1 if char not in punctuation + digits])
        p_tokens = perrault1.split()
        p_tokens

Out[7]: ['the',
         'sleeping',
         'beauty',
         'in',
         'the',
         'wood',
         'once',
         'upon',
         'a',
         'time',
         'there',
         'lived',
         'a',
         'king',
         'and',
         'queen',
         'who',
         'were',
         'grieved',
         'more',
         'grieved',
         'than',
         'words',
         'can',
         'tell',
         'because',
         'they',
         'had',
         'no',
         'children',
         'they',
         'tried',
         'the',
         'waters',
```

'of',
'every',
'country',
'made',
'vows',
'and',
'pilgrimages',
'and',
'did',
'everything',
'that',
'could',
'be',
'done',
'but',
'without',
'result',
'at',
'last',
'however',
'the',
'queen',
'found',
'that',
'her',
'wishes',
'were',
'fulfilled',
'and',
'in',
'due',
'course',
'she',
'gave',
'birth',
'to',
'a',
'daughter',
'a',
'grand',
'christening',
'was',
'held',
'and',
'all',
'the',
'fairies',
'that',

'could',
'be',
'found',
'in',
'the',
'realm',
'they',
'numbered',
'seven',
'in',
'all',
'were',
'invited',
'to',
'be',
'godmothers',
'to',
'the',
'little',
'princess',
'this',
'was',
'done',
'so',
'that',
'by',
'means',
'of',
'the',
'gifts',
'which',
'each',
'in',
'turn',
'would',
'bestow',
'upon',
'her',
'in',
'accordance',
'with',
'the',
'fairy',
'custom',
'of',
'those',
'days',
'the',

'princess',
'might',
'be',
'endowed',
'with',
'every',
'imaginable',
'perfection',
'when',
'the',
'christening',
'ceremony',
'was',
'over',
'all',
'the',
'company',
'returned',
'to',
'the',
'kings',
'palace',
'where',
'a',
'great',
'banquet',
'was',
'held',
'in',
'honor',
'of',
'the',
'fairies',
'places',
'were',
'laid',
'for',
'them',
'in',
'magnificent',
'style',
'and',
'before',
'each',
'was',
'placed',
'a',
'solid',

'gold',
'casket',
'containing',
'a',
'spoon',
'fork',
'and',
'knife',
'of',
'fine',
'gold',
'set',
'with',
'diamonds',
'and',
'rubies',
'but',
'just',
'as',
'all',
'were',
'sitting',
'down',
'to',
'table',
'an',
'aged',
'fairy',
'was',
'seen',
'to',
'enter',
'whom',
'no',
'one',
'had',
'thought',
'to',
'invite',
'the',
'reason',
'being',
'that',
'for',
'more',
'than',
'fifty',
'years',

'she',
'had',
'never',
'quitted',
'the',
'tower',
'in',
'which',
'she',
'lived',
'and',
'people',
'had',
'supposed',
'her',
'to',
'be',
'dead',
'or',
'bewitched',
'by',
'the',
'kings',
'orders',
'a',
'place',
'was',
'laid',
'for',
'her',
'but',
'it',
'was',
'impossible',
'to',
'give',
'her',
'a',
'golden',
'casket',
'like',
'the',
'others',
'for',
'only',
'seven',
'had',
'been',

'made',
'for',
'the',
'seven',
'fairies',
'the',
'old',
'creature',
'believed',
'that',
'she',
'was',
'intentionally',
'slighted',
'and',
'muttered',
'threats',
'between',
'her',
'teeth',
'she',
'was',
'overheard',
'by',
'one',
'of',
'the',
'young',
'fairies',
'who',
'was',
'seated',
'nearby',
'the',
'latter',
'guessing',
'that',
'some',
'mischievous',
'gift',
'might',
'be',
'bestowed',
'upon',
'the',
'little',
'princess',
'hid',

'behind',
'the',
'tapestry',
'as',
'soon',
'as',
'the',
'company',
'left',
'the',
'table',
'her',
'intention',
'was',
'to',
'be',
'the',
'last',
'to',
'speak',
'and',
'so',
'to',
'have',
'the',
'power',
'of',
'counteracting',
'as',
'far',
'as',
'possible',
'any',
'evil',
'which',
'the',
'old',
'fairy',
'might',
'do',
'presently',
'the',
'fairies',
'began',
'to',
'bestow',
'their',
'gifts',

'upon',
'the',
'princess',
'the',
'youngest',
'ordained',
'that',
'she',
'should',
'be',
'the',
'most',
'beautiful',
'person',
'in',
'the',
'world',
'the',
'next',
'that',
'she',
'should',
'have',
'the',
'temper',
'of',
'an',
'angel',
'the',
'third',
'that',
'she',
'should',
'do',
'everything',
'with',
'wonderful',
'grace',
'the',
'fourth',
'that',
'she',
'should',
'dance',
'to',
'perfection',
'the',
'fifth',

'that',
'she',
'should',
'sing',
'like',
'a',
'nightingale',
'and',
'the',
'sixth',
'that',
'she',
'should',
'play',
'every',
'kind',
'of',
'music',
'with',
'the',
'utmost',
'skill',
'it',
'was',
'now',
'the',
'turn',
'of',
'the',
'aged',
'fairy',
'shaking',
'her',
'head',
'in',
'token',
'of',
'spite',
'rather',
'than',
'of',
'infirmity',
'she',
'declared',
'that',
'the',
'princess',
'should',

'prick',
'her',
'hand',
'with',
'a',
'spindle',
'and',
'die',
'of',
'it',
'a',
'shudder',
'ran',
'through',
'the',
'company',
'at',
'this',
'terrible',
'gift',
'all',
'eyes',
'were',
'filled',
'with',
'tears',
'but',
'at',
'this',
'moment',
'the',
'young',
'fairy',
'stepped',
'forth',
'from',
'behind',
'the',
'tapestry',
'take',
'comfort',
'your',
'majesties',
'she',
'cried',
'in',
'a',
'loud',

'voice',
'your',
'daughter',
'shall',
'not',
'die',
'my',
'power',
'it',
'is',
'true',
'is',
'not',
'enough',
'to',
'undo',
'all',
'that',
'my',
'aged',
'kinswoman',
'has',
'decreed',
'the',
'princess',
'will',
'indeed',
'prick',
'her',
'hand',
'with',
'a',
'spindle',
'but',
'instead',
'of',
'dying',
'she',
'shall',
'merely',
'fall',
'into',
'a',
'profound',
'slumber',
'that',
'will',
'last',

'a',
'hundred',
'years',
'at',
'the',
'end',
'of',
'that',
'time',
'a',
'kings',
'son',
'shall',
'come',
'to',
'awaken',
'her',
'the',
'king',
'in',
'an',
'attempt',
'to',
'avert',
'the',
'unhappy',
'doom',
'pronounced',
'by',
'the',
'old',
'fairy',
'at',
'once',
'published',
'an',
'edict',
'forbidding',
'all',
'persons',
'under',
'pain',
'of',
'death',
'to',
'use',
'a',
'spinning',

'wheel',
'or',
'keep',
'a',
'spindle',
'in',
'the',
'house',
'at',
'the',
'end',
'of',
'fifteen',
'or',
'sixteen',
'years',
'the',
'king',
'and',
'queen',
'happened',
'one',
'day',
'to',
'be',
'away',
'on',
'pleasure',
'bent',
'the',
'princess',
'was',
'running',
'about',
'the',
'castle',
'and',
'going',
'upstairs',
'from',
'room',
'to',
'room',
'she',
'came',
'at',
'length',
'to',

'a',
'garret',
'at',
'the',
'top',
'of',
'a',
'tower',
'where',
'an',
'old',
'serving',
'woman',
'sat',
'alone',
'with',
'her',
'distaff',
'spinning',
'this',
'good',
'woman',
'had',
'never',
'heard',
'speak',
'of',
'the',
'kings',
'proclamation',
'forbidding',
'the',
'use',
'of',
'spinning',
'wheels',
'what',
'are',
'you',
'doing',
'my',
'good',
'woman',
'asked',
'the',
'princess',
'i',
'am',

'spinning',
'my',
'pretty',
'child',
'replied',
'the',
'dame',
'not',
'knowing',
'who',
'she',
'was',
'oh',
'what',
'fun',
'rejoined',
'the',
'princess',
'how',
'do',
'you',
'do',
'it',
'let',
'me',
'try',
'and',
'see',
'if',
'i',
'can',
'do',
'it',
'equally',
'well',
'partly',
'because',
'she',
'was',
'too',
'hasty',
'partly',
'because',
'she',
'was',
'a',
'little',
'heedless',

'but',
'also',
'because',
'the',
'fairy',
'decree',
'had',
'ordained',
'it',
'no',
'sooner',
'had',
'she',
'seized',
'the',
'spindle',
'than',
'she',
'pricked',
'her',
'hand',
'and',
'fell',
'down',
'in',
'a',
'swoon',
'in',
'great',
'alarm',
'the',
'good',
'dame',
'cried',
'out',
'for',
'help',
'people',
'came',
'running',
'from',
'every',
'quarter',
'to',
'the',
'princess',
'they',
'threw',

'water',
'on',
'her',
'face',
'chafed',
'her',
'with',
'their',
'hands',
'and',
'rubbed',
'her',
'temples',
'with',
'the',
'royal',
'essence',
'of',
'hungary',
'but',
'nothing',
'would',
'restore',
'her',
'then',
'the',
'king',
'who',
'had',
'been',
'brought',
'upstairs',
'by',
'the',
'commotion',
'remembered',
'the',
'fairy',
'prophecy',
'feeling',
'certain',
'that',
'what',
'had',
'happened',
'was',
'inevitable',
'since',

'the',
'fairies',
'had',
'decreed',
'it',
'he',
'gave',
'orders',
'that',
'the',
'princess',
'should',
'be',
'placed',
'in',
'the',
'finest',
'apartment',
'in',
'the',
'palace',
'upon',
'a',
'bed',
'embroidered',
'in',
'gold',
'and',
'silver',
'you',
'would',
'have',
'thought',
'her',
'an',
'angel',
'so',
'fair',
'was',
'she',
'to',
'behold',
'the',
'trance',
'had',
'not',
'taken',
'away',

'the',
'lovely',
'color',
'of',
'her',
'complexion',
'her',
'cheeks',
'were',
'delicately',
'flushed',
'her',
'lips',
'like',
'coral',
'her',
'eyes',
'indeed',
'were',
'closed',
'but',
'her',
'gentle',
'breathing',
'could',
'be',
'heard',
'and',
'it',
'was',
'therefore',
'plain',
'that',
'she',
'was',
'not',
'dead',
'the',
'king',
'commanded',
'that',
'she',
'should',
'be',
'left',
'to',
'sleep',
'in',

'peace',
'until',
'the',
'hour',
'of',
'her',
'awakening',
'should',
'come',
'when',
'the',
'accident',
'happened',
'to',
'the',
'princess',
'the',
'good',
'fairy',
'who',
'had',
'saved',
'her',
'life',
'by',
'condemning',
'her',
'to',
'sleep',
'a',
'hundred',
'years',
'was',
'in',
'the',
'kingdom',
'of',
'mataquin',
'twelve',
'thousand',
'leagues',
'away',
'she',
'was',
'instantly',
'warned',
'of',
'it',

```
        'however',
        'by',
        'a',
        'little',
        'dwarf',
        'who',
        ...]

In [8]: perrault = perrault.split("\n,\n")
        perrault

Out[8]: ['The Sleeping Beauty in the Wood,"\nOnce upon a time there lived a king and queen who




In [9]: len(perrault)

Out[9]: 8

In [10]: with open('Grimm.csv', 'r') as e:
             grimm = e.read()
         grimm2 = grimm
         grimm

Out[10]: 'All-Kinds-of-Fur,"Once upon a time there was a king, who had a wife with golden hair

In [11]: len_raw_g = len(grimm)
         len_raw_g

Out[11]: 69475

In [12]: grimm = grimm.replace(",,,,,,,,", "").strip()
         grimm

Out[12]: 'All-Kinds-of-Fur,"Once upon a time there was a king, who had a wife with golden hair

In [13]: grimm1 = ''.join([char for char in grimm if char not in punctuation + digits + '\n'])
         grimm1 = grimm1.lower()
         g_tokens = grimm1.split()
         g_tokens

Out[13]: ['allkindsoffuronce',
          'upon',
          'a',
          'time',
```

23

'there',
'was',
'a',
'king',
'who',
'had',
'a',
'wife',
'with',
'golden',
'hair',
'and',
'she',
'was',
'so',
'beautiful',
'that',
'her',
'equal',
'was',
'not',
'to',
'be',
'found',
'anywhere',
'on',
'earth',
'it',
'happened',
'that',
'she',
'fell',
'ill',
'and',
'when',
'she',
'felt',
'that',
'she',
'was',
'about',
'to',
'die',
'she',
'called',
'the',
'king',
'to',

```
'her',
'and',
'said',
'if',
'you',
'want',
'to',
'remarry',
'after',
'my',
'death',
'then',
'do',
'not',
'take',
'anyone',
'who',
'is',
'not',
'as',
'beautiful',
'as',
'i',
'am',
'and',
'who',
'does',
'not',
'have',
'such',
'golden',
'hair',
'as',
'mine',
'that',
'you',
'must',
'promise',
'me',
'after',
'the',
'king',
'had',
'promised',
'her',
'this',
'she',
'closed',
```

'her',
'eyes',
'and',
'died',
'for',
'a',
'long',
'time',
'the',
'king',
'could',
'not',
'be',
'comforted',
'and',
'he',
'did',
'not',
'consider',
'taking',
'a',
'second',
'wife',
'finally',
'his',
'councilors',
'said',
'there',
'is',
'no',
'other',
'way',
'the',
'king',
'must',
'remarry',
'so',
'that',
'we',
'will',
'have',
'a',
'queen',
'then',
'messengers',
'were',
'sent',
'out',

'far',
'and',
'wide',
'to',
'seek',
'a',
'bride',
'who',
'in',
'beauty',
'was',
'entirely',
'the',
'equal',
'of',
'the',
'deceased',
'queen',
'however',
'no',
'such',
'bride',
'could',
'be',
'found',
'in',
'the',
'whole',
'world',
'and',
'even',
'if',
'a',
'bride',
'this',
'beautiful',
'had',
'been',
'found',
'she',
'would',
'not',
'have',
'had',
'such',
'golden',
'hair',
'thus',

```
'the',
'messengers',
'returned',
'home',
'without',
'success',
'now',
'the',
'king',
'had',
'a',
'daughter',
'who',
'was',
'just',
'as',
'beautiful',
'as',
'her',
'deceased',
'mother',
'and',
'who',
'also',
'had',
'such',
'golden',
'hair',
'after',
'she',
'had',
'come',
'of',
'age',
'the',
'king',
'looked',
'at',
'her',
'one',
'day',
'that',
'she',
'resembled',
'his',
'deceased',
'wife',
'in',
```

'every',
'way',
'and',
'he',
'suddenly',
'felt',
'a',
'great',
'love',
'for',
'her',
'then',
'he',
'said',
'to',
'his',
'councilors',
'i',
'will',
'marry',
'my',
'daughter',
'for',
'she',
'is',
'the',
'image',
'of',
'my',
'deceased',
'wife',
'and',
'nowhere',
'else',
'can',
'i',
'find',
'a',
'bride',
'who',
'is',
'her',
'equal',
'when',
'the',
'councilors',
'heard',
'this',

'they',
'were',
'horrified',
'and',
'said',
'god',
'has',
'forbidden',
'that',
'a',
'father',
'should',
'marry',
'his',
'daughter',
'nothing',
'good',
'can',
'come',
'from',
'sin',
'and',
'the',
'kingdom',
'will',
'be',
'pulled',
'along',
'into',
'ruin',
'the',
'daughter',
'was',
'even',
'more',
'horrified',
'when',
'she',
'heard',
'her',
'fathers',
'decision',
'hoping',
'to',
'dissuade',
'him',
'from',
'his',

'intentions',
'she',
'said',
'to',
'him',
'before',
'i',
'fulfill',
'your',
'wish',
'i',
'must',
'have',
'three',
'dresses',
'one',
'as',
'golden',
'as',
'the',
'sun',
'one',
'as',
'silver',
'as',
'the',
'moon',
'and',
'one',
'that',
'glistens',
'like',
'the',
'stars',
'further',
'i',
'must',
'have',
'a',
'cloak',
'put',
'together',
'from',
'a',
'thousand',
'kinds',
'of',
'pelts',

```
'and',
'fur',
'every',
'animal',
'in',
'your',
'kingdom',
'must',
'contribute',
'a',
'piece',
'of',
'its',
'skin',
'for',
'it',
'now',
'she',
'thought',
'that',
'will',
'be',
'entirely',
'impossible',
'for',
'him',
'to',
'achieve',
'and',
'in',
'this',
'way',
'i',
'will',
'divert',
'my',
'father',
'from',
'his',
'wicked',
'thoughts',
'but',
'the',
'king',
'did',
'not',
'give',
'in',
```

```
'and',
'the',
'most',
'skilled',
'maidens',
'in',
'his',
'kingdom',
'had',
'to',
'weave',
'the',
'three',
'dresses',
'one',
'as',
'golden',
'as',
'the',
'sun',
'one',
'as',
'silver',
'as',
'the',
'moon',
'and',
'one',
'that',
'glistened',
'like',
'the',
'stars',
'and',
'his',
'huntsmen',
'had',
'to',
'capture',
'all',
'the',
'animals',
'in',
'his',
'entire',
'kingdom',
'and',
'take',
```

```
'a',
'piece',
'of',
'skin',
'from',
'each',
'one',
'from',
'these',
'a',
'cloak',
'of',
'a',
'thousand',
'kinds',
'of',
'fur',
'was',
'made',
'finally',
'when',
'everything',
'was',
'finished',
'the',
'king',
'had',
'the',
'cloak',
'brought',
'to',
'him',
'spreading',
'it',
'out',
'he',
'said',
'tomorrow',
'is',
'our',
'wedding',
'day',
'when',
'the',
'kings',
'daughter',
'saw',
'that',
```

'there',
'was',
'no',
'more',
'hope',
'to',
'change',
'her',
'fathers',
'mind',
'she',
'made',
'the',
'decision',
'to',
'run',
'away',
'during',
'the',
'night',
'while',
'everyone',
'was',
'asleep',
'she',
'got',
'up',
'and',
'took',
'three',
'things',
'from',
'among',
'her',
'valuables',
'a',
'golden',
'ring',
'a',
'little',
'golden',
'spinning',
'wheel',
'and',
'a',
'little',
'golden',
'reel',

```
'she',
'put',
'the',
'three',
'dresses',
'from',
'the',
'sun',
'moon',
'and',
'stars',
'into',
'a',
'nutshell',
'put',
'on',
'the',
'cloak',
'of',
'all',
'kinds',
'of',
'fur',
'blackened',
'her',
'hands',
'and',
'face',
'with',
'soot',
'then',
'surrendering',
'herself',
'to',
'god',
'she',
'set',
'forth',
'she',
'walked',
'the',
'entire',
'night',
'until',
'she',
'came',
'to',
'a',
```

'great',
'forest',
'being',
'tired',
'she',
'sat',
'down',
'in',
'a',
'hollow',
'tree',
'and',
'fell',
'asleep',
'the',
'sun',
'came',
'up',
'and',
'she',
'continued',
'to',
'sleep',
'and',
'she',
'was',
'still',
'asleep',
'by',
'broad',
'daylight',
'now',
'it',
'came',
'to',
'pass',
'that',
'the',
'king',
'who',
'owned',
'these',
'woods',
'was',
'hunting',
'in',
'them',
'when',

```
'his',
'dogs',
'approached',
'the',
'tree',
'they',
'sniffed',
'then',
'ran',
'around',
'it',
'barking',
'the',
'king',
'said',
'to',
'the',
'huntsmen',
'see',
'what',
'kind',
'of',
'wild',
'animal',
'is',
'hiding',
'there',
'the',
'huntsmen',
'followed',
'his',
'command',
'and',
'when',
'they',
'returned',
'they',
'said',
'a',
'strange',
'animal',
'like',
'none',
'we',
'have',
'ever',
'seen',
'before',
```

'is',
'lying',
'in',
'the',
'hollow',
'tree',
'there',
'are',
'a',
'thousand',
'kinds',
'of',
'fur',
'on',
'its',
'skin',
'it',
'is',
'just',
'lying',
'there',
'asleep',
'the',
'king',
'said',
'see',
'if',
'you',
'can',
'capture',
'it',
'alive',
'then',
'tie',
'it',
'onto',
'the',
'cart',
'and',
'bring',
'it',
'along',
'when',
'the',
'huntsmen',
'took',
'hold',
'of',

```
'the',
'girl',
'she',
'awoke',
'filled',
'with',
'fear',
'she',
'cried',
'out',
'i',
'am',
'a',
'poor',
'child',
'who',
'has',
'been',
'abandoned',
'by',
'her',
'father',
'and',
'mother',
'have',
'pity',
'on',
'me',
'and',
'take',
'me',
'with',
'you',
'then',
'they',
'said',
'allkindsoffur',
'you',
'are',
'good',
'for',
'the',
'kitchen',
'come',
'with',
'us',
'you',
'can',
```

'sweep',
'up',
'the',
'ashes',
'thus',
'they',
'set',
'her',
'on',
'the',
'cart',
'and',
'drove',
'her',
'home',
'to',
'the',
'royal',
'castle',
'there',
'they',
'showed',
'her',
'a',
'little',
'cubbyhole',
'under',
'the',
'stairs',
'where',
'the',
'light',
'of',
'day',
'never',
'entered',
'and',
'said',
'this',
'is',
'where',
'you',
'can',
'live',
'and',
'sleep',
'you',
'furry',

'little',
'animal',
'then',
'she',
'was',
'sent',
'to',
'the',
'kitchen',
'where',
'she',
'carried',
'wood',
'and',
'water',
'tended',
'the',
'fire',
'plucked',
'the',
'poultry',
'sorted',
'vegetables',
'swept',
'up',
'the',
'ashes',
'and',
'did',
'all',
'the',
'dirty',
'work',
'allkindsoffur',
'lived',
'there',
'miserable',
'for',
'a',
'long',
'time',
'oh',
'you',
'beautiful',
'princess',
'what',
'will',
'become',

```
'of',
'you',
'now',
'one',
'day',
'it',
'happened',
'that',
'a',
'banquet',
'was',
'being',
'held',
'in',
'the',
'castle',
'and',
'she',
'said',
'to',
'the',
'cook',
'may',
'i',
'go',
'up',
'for',
'a',
'little',
'while',
'and',
'take',
'a',
'look',
'i',
'will',
'stand',
'outside',
'the',
'door',
'the',
'cook',
'answered',
'yes',
'go',
'ahead',
'but',
'you',
```

```
                    'have',
                    'to',
                    'be',
                    'back',
                    'here',
                    'in',
                    'a',
                    'half',
                    'hour',
                    'and',
                    'carry',
                    'out',
                    'the',
                    'ashes',
                    'then',
                    'she',
                    'took',
                    'her',
                    'oil',
                    'lamp',
                    'and',
                    'went',
                    'into',
                    'her',
                    'cubbyhole',
                    'taking',
                    'off',
                    'her',
                    'fur',
                    'cloak',
                    'she',
                    'washed',
                    'the',
                    'soot',
                    'from',
                    'her',
                    ...]

In [14]: grimm = grimm.split("\n,\n")
         grimm

Out[14]: ['All-Kinds-of-Fur,"Once upon a time there was a king, who had a wife with golden hai
```

```
In [15]: len(grimm)
```

```
Out[15]: 8
```

```
In [16]: p_dict = Counter(p_tokens)
         g_dict = Counter(g_tokens)
```

```
In [17]: p_common_words = p_dict.most_common()
         p_common_35 = p_common_words[:35]
         p_common_35
```

```
Out[17]: [('the', 1060),
          ('and', 569),
          ('to', 519),
          ('her', 363),
          ('a', 360),
          ('of', 325),
          ('he', 288),
          ('she', 274),
          ('was', 272),
          ('that', 260),
          ('in', 224),
          ('his', 223),
          ('had', 200),
          ('they', 184),
          ('it', 158),
          ('with', 157),
          ('you', 153),
          ('for', 145),
          ('but', 132),
          ('all', 130),
          ('as', 127),
          ('not', 122),
          ('at', 117),
          ('them', 116),
          ('him', 106),
          ('so', 102),
          ('i', 102),
          ('this', 100),
          ('little', 92),
          ('who', 90),
          ('were', 87),
          ('be', 82),
          ('said', 82),
          ('very', 79),
          ('which', 75)]
```

```
In [18]: g_common_words = g_dict.most_common()
         g_common_35 = g_common_words[:35]
         g_common_35
```

```
Out[18]: [('the', 827),
         ('and', 601),
         ('to', 345),
         ('a', 228),
         ('she', 225),
         ('her', 218),
         ('he', 203),
         ('was', 173),
         ('of', 160),
         ('in', 159),
         ('had', 144),
         ('that', 142),
         ('it', 135),
         ('you', 119),
         ('with', 112),
         ('i', 103),
         ('his', 95),
         ('on', 92),
         ('as', 92),
         ('said', 91),
         ('him', 91),
         ('when', 86),
         ('for', 86),
         ('into', 75),
         ('out', 74),
         ('but', 74),
         ('then', 71),
         ('they', 71),
         ('not', 70),
         ('one', 67),
         ('so', 65),
         ('is', 63),
         ('have', 60),
         ('me', 60),
         ('at', 58)]

In [19]: p_word_35 = make_array()
         p_num_35 = make_array()
         g_word_35 = make_array()
         g_num_35 = make_array()

         for i in range(len(p_common_35)-1):
             p_word_35 = np.append(p_word_35, p_common_35[i][0])
             p_num_35 = np.append(p_num_35, p_common_35[i][1])

         for h in range(len(g_common_35)-1):
             g_word_35 = np.append(g_word_35, g_common_35[h][0])
             g_num_35 = np.append(g_num_35, g_common_35[h][1])
```
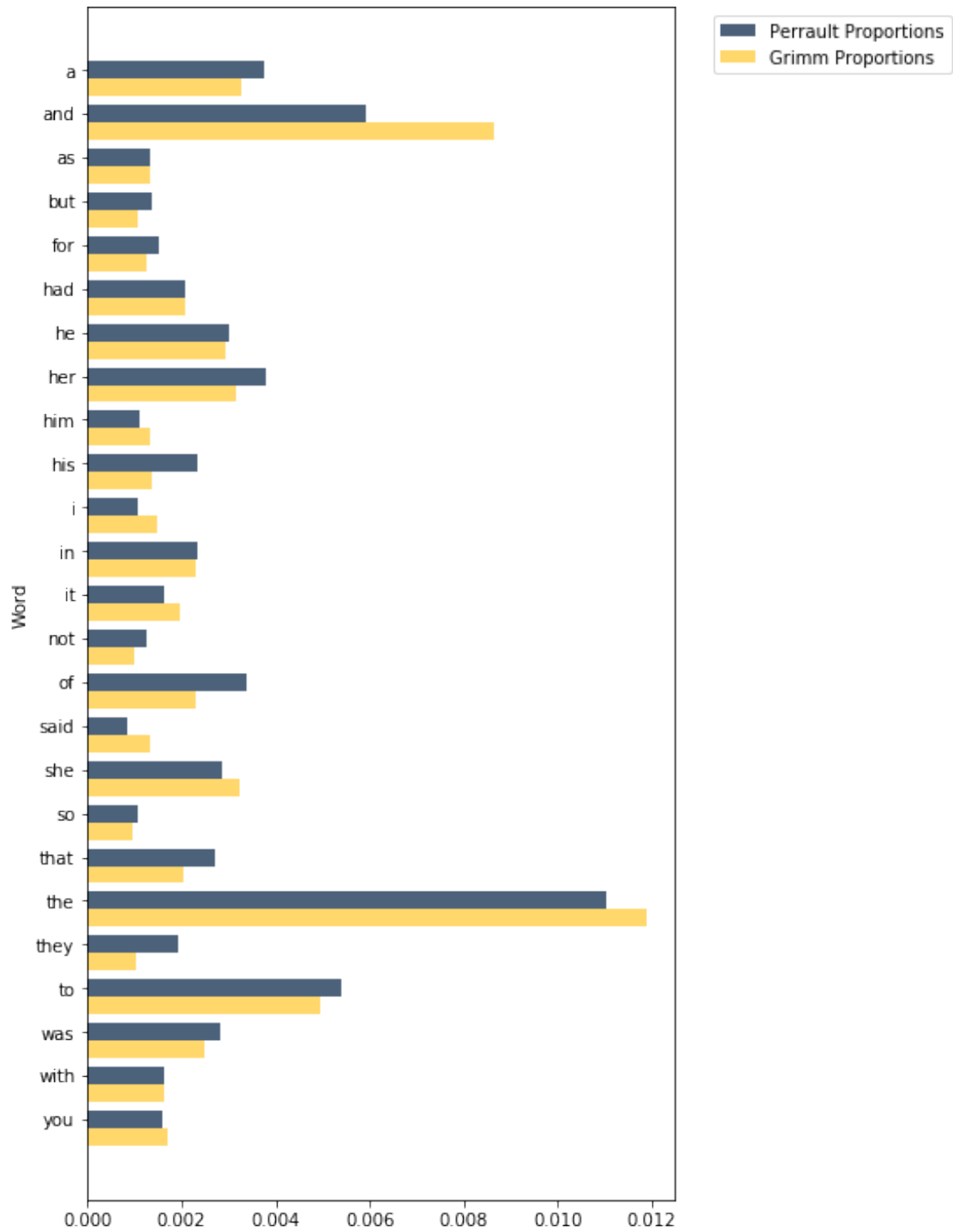
```python
p_prop = p_num_35 / len_raw_p
g_prop = g_num_35 /len_raw_g

p_table = Table().with_columns("Perrault", p_word_35, "Perrault Proportions", p_prop)
g_table = Table().with_columns("Grimm", g_word_35, "Grimm Proportions", g_prop)
common_table = p_table.join("Perrault", g_table, "Grimm")
common_table = common_table.relabeled(0,"Word")
common_table.barh("Word")
```

```
In [20]: cv= CountVectorizer()
         dtm_perrault = cv.fit_transform(perrault)
         dtm_perrault
```

```
Out[20]: <8x2502 sparse matrix of type '<class 'numpy.int64'>'
            with 4968 stored elements in Compressed Sparse Row format>

In [21]: p_desparse = dtm_perrault.toarray()

         p_word_list = cv.get_feature_names()

         p_dtm_tb = Table(p_word_list).with_rows(p_desparse).drop(stop_words)
         p_dtm_tb

Out[21]: abandon | abandoned | able | absence | abundance | accepted | accident | accordance |
         0       | 0         | 4    | 1       | 0         | 0        | 1        | 1          |

In [ ]:

In [22]: p_toks_tab = Table()
         p_toks_tab.append_column(label="Word List", values=p_word_list)
         p_toks_tab.append_column(label="Frequency", values=sum(p_desparse))
         p_row_sums = np.sum(p_desparse, axis=1)
         p_normed = p_desparse/p_row_sums[:,None]
         p_dtm_tb = Table(p_word_list).with_rows(p_normed)
         p_dtm_tb

Out[22]: abandon     | abandoned   | able        | about       | above       | absence     | abu

In [23]: cv = CountVectorizer(stop_words = stop_words)
         p_dtm = cv.fit_transform(perrault)
         pt = TfidfTransformer(norm='l1',use_idf=False)
         p_dtm_tf = pt.fit_transform(p_dtm)

         p_word_list = cv.get_feature_names()
         p_dtm_array = p_dtm_tf.toarray()
         Table(p_word_list).with_rows(p_dtm_array)
```

```
Out[23]: abandon     | abandoned   | able       | absence     | abundance | accepted   | accide
```

```
In [24]: p_married = p_dtm_tb.column("married")
         p_dtm_tb.column("moral")
```

```
Out[24]: array([ 0.00027601,  0.00149701,  0.00106383,  0.00234192,  0.00233918,
                 0.00042159,  0.00056117,  0.00041876])
```

```
In [25]: dtm_grimm = cv.fit_transform(grimm)
         dtm_grimm
```

```
Out[25]: <8x1520 sparse matrix of type '<class 'numpy.int64'>'
             with 2599 stored elements in Compressed Sparse Row format>
```

```
In [26]: g_desparse = dtm_grimm.toarray()

         g_word_list = cv.get_feature_names()

         g_dtm_tb = Table(g_word_list).with_rows(g_desparse).drop(stop_words)
         g_dtm_tb
```

```
Out[26]: abandoned | able | absolutely | accompanied | accomplish | accordance | account | achi
         1         | 1    | 0          | 0           | 0          | 0          | 0       | 1
```

```
In [27]: g_toks_tab = Table()
         g_toks_tab.append_column(label="Word List", values=g_word_list)
         g_toks_tab.append_column(label="Frequency", values=sum(g_desparse))
         g_row_sums = np.sum(g_desparse, axis=1)
         g_normed = g_desparse/g_row_sums[:,None]
         g_dtm_tb = Table(g_word_list).with_rows(g_normed)
         g_dtm_tb
```

```
Out[27]: abandoned  | able       | absolutely  | accompanied | accomplish | accordance | accou
```

```
   0              | 0.00170068 | 0            | 0            | 0            | 0            | 0
```

```
In [28]: cv = CountVectorizer(stop_words = stop_words)
         g_dtm = cv.fit_transform(grimm)
         gt = TfidfTransformer(norm='l1',use_idf=False)
         g_dtm_tf = gt.fit_transform(g_dtm)

         g_word_list = cv.get_feature_names()
         g_dtm_array = g_dtm_tf.toarray()
         Table(g_word_list).with_rows(g_dtm_array)
```

```
Out[28]: abandoned  | able       | absolutely  | accompanied | accomplish | accordance | accou
```

```
In [29]: g_married = g_dtm_tb.column("married")
```

```
In [30]: p_married , g_married
```

```
Out[30]: (array([ 0.00055203,  0.         ,  0.00159574,  0.00058548,  0.00116959,
                  0.00042159,  0.         ,  0.         ]),
           array([ 0.         ,  0.         ,  0.         ,  0.         ,  0.         ,
                  0.00208768,  0.         ,  0.         ]))
```

```
In [31]: tales = list(set(perrault + grimm))
         tales
```

```
Out[31]: ['The Fairies,"Once upon a time there was a widow who had two daughters. The elder was
          'The Sleeping Beauty in the Wood,"\nOnce upon a time there lived a king and queen who
```

'Frau Holle,"A widow had two daughters, the one was beautiful and industrious, the ot
ened my eyes to-day I saw you Glass Coffin, her dream fulfilled. the Help me to accomplish the great things which
'Donkey Skin,"\nOnce upon a time there was a king who was the most powerful ruler in

```
In [32]: tales1 = perrault1+grimm1
         tales_tokens = tales1.split()
         tales_tokens

Out[32]: ['the',
          'sleeping',
          'beauty',
          'in',
          'the',
          'wood',
          'once',
          'upon',
          'a',
          'time',
          'there',
          'lived',
          'a',
          'king',
          'and',
          'queen',
          'who',
          'were',
          'grieved',
          'more',
          'grieved',
          'than',
          'words',
          'can',
          'tell',
          'because',
          'they',
          'had',
          'no',
          'children',
          'they',
          'tried',
          'the',
          'waters',
          'of',
          'every',
          'country',
```

```
'made',
'vows',
'and',
'pilgrimages',
'and',
'did',
'everything',
'that',
'could',
'be',
'done',
'but',
'without',
'result',
'at',
'last',
'however',
'the',
'queen',
'found',
'that',
'her',
'wishes',
'were',
'fulfilled',
'and',
'in',
'due',
'course',
'she',
'gave',
'birth',
'to',
'a',
'daughter',
'a',
'grand',
'christening',
'was',
'held',
'and',
'all',
'the',
'fairies',
'that',
'could',
'be',
'found',
```

'in',
'the',
'realm',
'they',
'numbered',
'seven',
'in',
'all',
'were',
'invited',
'to',
'be',
'godmothers',
'to',
'the',
'little',
'princess',
'this',
'was',
'done',
'so',
'that',
'by',
'means',
'of',
'the',
'gifts',
'which',
'each',
'in',
'turn',
'would',
'bestow',
'upon',
'her',
'in',
'accordance',
'with',
'the',
'fairy',
'custom',
'of',
'those',
'days',
'the',
'princess',
'might',
'be',

'endowed',
'with',
'every',
'imaginable',
'perfection',
'when',
'the',
'christening',
'ceremony',
'was',
'over',
'all',
'the',
'company',
'returned',
'to',
'the',
'kings',
'palace',
'where',
'a',
'great',
'banquet',
'was',
'held',
'in',
'honor',
'of',
'the',
'fairies',
'places',
'were',
'laid',
'for',
'them',
'in',
'magnificent',
'style',
'and',
'before',
'each',
'was',
'placed',
'a',
'solid',
'gold',
'casket',
'containing',

'a',
'spoon',
'fork',
'and',
'knife',
'of',
'fine',
'gold',
'set',
'with',
'diamonds',
'and',
'rubies',
'but',
'just',
'as',
'all',
'were',
'sitting',
'down',
'to',
'table',
'an',
'aged',
'fairy',
'was',
'seen',
'to',
'enter',
'whom',
'no',
'one',
'had',
'thought',
'to',
'invite',
'the',
'reason',
'being',
'that',
'for',
'more',
'than',
'fifty',
'years',
'she',
'had',
'never',

'quitted',
'the',
'tower',
'in',
'which',
'she',
'lived',
'and',
'people',
'had',
'supposed',
'her',
'to',
'be',
'dead',
'or',
'bewitched',
'by',
'the',
'kings',
'orders',
'a',
'place',
'was',
'laid',
'for',
'her',
'but',
'it',
'was',
'impossible',
'to',
'give',
'her',
'a',
'golden',
'casket',
'like',
'the',
'others',
'for',
'only',
'seven',
'had',
'been',
'made',
'for',
'the',

```
'seven',
'fairies',
'the',
'old',
'creature',
'believed',
'that',
'she',
'was',
'intentionally',
'slighted',
'and',
'muttered',
'threats',
'between',
'her',
'teeth',
'she',
'was',
'overheard',
'by',
'one',
'of',
'the',
'young',
'fairies',
'who',
'was',
'seated',
'nearby',
'the',
'latter',
'guessing',
'that',
'some',
'mischievous',
'gift',
'might',
'be',
'bestowed',
'upon',
'the',
'little',
'princess',
'hid',
'behind',
'the',
'tapestry',
```

'as',
'soon',
'as',
'the',
'company',
'left',
'the',
'table',
'her',
'intention',
'was',
'to',
'be',
'the',
'last',
'to',
'speak',
'and',
'so',
'to',
'have',
'the',
'power',
'of',
'counteracting',
'as',
'far',
'as',
'possible',
'any',
'evil',
'which',
'the',
'old',
'fairy',
'might',
'do',
'presently',
'the',
'fairies',
'began',
'to',
'bestow',
'their',
'gifts',
'upon',
'the',
'princess',

'the',
'youngest',
'ordained',
'that',
'she',
'should',
'be',
'the',
'most',
'beautiful',
'person',
'in',
'the',
'world',
'the',
'next',
'that',
'she',
'should',
'have',
'the',
'temper',
'of',
'an',
'angel',
'the',
'third',
'that',
'she',
'should',
'do',
'everything',
'with',
'wonderful',
'grace',
'the',
'fourth',
'that',
'she',
'should',
'dance',
'to',
'perfection',
'the',
'fifth',
'that',
'she',
'should',

'sing',
'like',
'a',
'nightingale',
'and',
'the',
'sixth',
'that',
'she',
'should',
'play',
'every',
'kind',
'of',
'music',
'with',
'the',
'utmost',
'skill',
'it',
'was',
'now',
'the',
'turn',
'of',
'the',
'aged',
'fairy',
'shaking',
'her',
'head',
'in',
'token',
'of',
'spite',
'rather',
'than',
'of',
'infirmity',
'she',
'declared',
'that',
'the',
'princess',
'should',
'prick',
'her',
'hand',

'with',
'a',
'spindle',
'and',
'die',
'of',
'it',
'a',
'shudder',
'ran',
'through',
'the',
'company',
'at',
'this',
'terrible',
'gift',
'all',
'eyes',
'were',
'filled',
'with',
'tears',
'but',
'at',
'this',
'moment',
'the',
'young',
'fairy',
'stepped',
'forth',
'from',
'behind',
'the',
'tapestry',
'take',
'comfort',
'your',
'majesties',
'she',
'cried',
'in',
'a',
'loud',
'voice',
'your',
'daughter',

'shall',
'not',
'die',
'my',
'power',
'it',
'is',
'true',
'is',
'not',
'enough',
'to',
'undo',
'all',
'that',
'my',
'aged',
'kinswoman',
'has',
'decreed',
'the',
'princess',
'will',
'indeed',
'prick',
'her',
'hand',
'with',
'a',
'spindle',
'but',
'instead',
'of',
'dying',
'she',
'shall',
'merely',
'fall',
'into',
'a',
'profound',
'slumber',
'that',
'will',
'last',
'a',
'hundred',
'years',

```
'at',
'the',
'end',
'of',
'that',
'time',
'a',
'kings',
'son',
'shall',
'come',
'to',
'awaken',
'her',
'the',
'king',
'in',
'an',
'attempt',
'to',
'avert',
'the',
'unhappy',
'doom',
'pronounced',
'by',
'the',
'old',
'fairy',
'at',
'once',
'published',
'an',
'edict',
'forbidding',
'all',
'persons',
'under',
'pain',
'of',
'death',
'to',
'use',
'a',
'spinning',
'wheel',
'or',
'keep',
```

'a',
'spindle',
'in',
'the',
'house',
'at',
'the',
'end',
'of',
'fifteen',
'or',
'sixteen',
'years',
'the',
'king',
'and',
'queen',
'happened',
'one',
'day',
'to',
'be',
'away',
'on',
'pleasure',
'bent',
'the',
'princess',
'was',
'running',
'about',
'the',
'castle',
'and',
'going',
'upstairs',
'from',
'room',
'to',
'room',
'she',
'came',
'at',
'length',
'to',
'a',
'garret',
'at',

```
'the',
'top',
'of',
'a',
'tower',
'where',
'an',
'old',
'serving',
'woman',
'sat',
'alone',
'with',
'her',
'distaff',
'spinning',
'this',
'good',
'woman',
'had',
'never',
'heard',
'speak',
'of',
'the',
'kings',
'proclamation',
'forbidding',
'the',
'use',
'of',
'spinning',
'wheels',
'what',
'are',
'you',
'doing',
'my',
'good',
'woman',
'asked',
'the',
'princess',
'i',
'am',
'spinning',
'my',
'pretty',
```

'child',
'replied',
'the',
'dame',
'not',
'knowing',
'who',
'she',
'was',
'oh',
'what',
'fun',
'rejoined',
'the',
'princess',
'how',
'do',
'you',
'do',
'it',
'let',
'me',
'try',
'and',
'see',
'if',
'i',
'can',
'do',
'it',
'equally',
'well',
'partly',
'because',
'she',
'was',
'too',
'hasty',
'partly',
'because',
'she',
'was',
'a',
'little',
'heedless',
'but',
'also',
'because',

'the',
'fairy',
'decree',
'had',
'ordained',
'it',
'no',
'sooner',
'had',
'she',
'seized',
'the',
'spindle',
'than',
'she',
'pricked',
'her',
'hand',
'and',
'fell',
'down',
'in',
'a',
'swoon',
'in',
'great',
'alarm',
'the',
'good',
'dame',
'cried',
'out',
'for',
'help',
'people',
'came',
'running',
'from',
'every',
'quarter',
'to',
'the',
'princess',
'they',
'threw',
'water',
'on',
'her',

'face',
'chafed',
'her',
'with',
'their',
'hands',
'and',
'rubbed',
'her',
'temples',
'with',
'the',
'royal',
'essence',
'of',
'hungary',
'but',
'nothing',
'would',
'restore',
'her',
'then',
'the',
'king',
'who',
'had',
'been',
'brought',
'upstairs',
'by',
'the',
'commotion',
'remembered',
'the',
'fairy',
'prophecy',
'feeling',
'certain',
'that',
'what',
'had',
'happened',
'was',
'inevitable',
'since',
'the',
'fairies',
'had',

```
'decreed',
'it',
'he',
'gave',
'orders',
'that',
'the',
'princess',
'should',
'be',
'placed',
'in',
'the',
'finest',
'apartment',
'in',
'the',
'palace',
'upon',
'a',
'bed',
'embroidered',
'in',
'gold',
'and',
'silver',
'you',
'would',
'have',
'thought',
'her',
'an',
'angel',
'so',
'fair',
'was',
'she',
'to',
'behold',
'the',
'trance',
'had',
'not',
'taken',
'away',
'the',
'lovely',
'color',
```

```
'of',
'her',
'complexion',
'her',
'cheeks',
'were',
'delicately',
'flushed',
'her',
'lips',
'like',
'coral',
'her',
'eyes',
'indeed',
'were',
'closed',
'but',
'her',
'gentle',
'breathing',
'could',
'be',
'heard',
'and',
'it',
'was',
'therefore',
'plain',
'that',
'she',
'was',
'not',
'dead',
'the',
'king',
'commanded',
'that',
'she',
'should',
'be',
'left',
'to',
'sleep',
'in',
'peace',
'until',
'the',
```

'hour',
'of',
'her',
'awakening',
'should',
'come',
'when',
'the',
'accident',
'happened',
'to',
'the',
'princess',
'the',
'good',
'fairy',
'who',
'had',
'saved',
'her',
'life',
'by',
'condemning',
'her',
'to',
'sleep',
'a',
'hundred',
'years',
'was',
'in',
'the',
'kingdom',
'of',
'mataquin',
'twelve',
'thousand',
'leagues',
'away',
'she',
'was',
'instantly',
'warned',
'of',
'it',
'however',
'by',
'a',

```
                    'little',
                    'dwarf',
                    'who',
                    ...]

In [33]: dict = Counter(tales_tokens)
         dict.most_common()

Out[33]: [('the', 1887),
          ('and', 1170),
          ('to', 864),
          ('a', 588),
          ('her', 581),
          ('she', 499),
          ('he', 491),
          ('of', 485),
          ('was', 445),
          ('that', 402),
          ('in', 383),
          ('had', 344),
          ('his', 318),
          ('it', 293),
          ('you', 272),
          ('with', 269),
          ('they', 255),
          ('for', 231),
          ('as', 219),
          ('but', 206),
          ('i', 205),
          ('him', 197),
          ('not', 192),
          ('all', 177),
          ('at', 175),
          ('said', 173),
          ('so', 167),
          ('on', 166),
          ('them', 151),
          ('little', 150),
          ('this', 139),
          ('into', 135),
          ('be', 133),
          ('is', 133),
          ('when', 131),
          ('were', 128),
          ('one', 128),
          ('have', 128),
          ('out', 127),
          ('who', 126),
```

```
('my', 118),
('then', 118),
('which', 115),
('from', 105),
('went', 102),
('up', 101),
('by', 100),
('will', 99),
('there', 96),
('very', 91),
('king', 89),
('no', 89),
('me', 89),
('what', 87),
('would', 84),
('their', 84),
('came', 83),
('could', 81),
('if', 79),
('time', 77),
('good', 74),
('more', 73),
('do', 71),
('are', 71),
('go', 68),
('see', 66),
('beautiful', 63),
('some', 62),
('again', 61),
('did', 60),
('two', 60),
('down', 57),
('where', 56),
('been', 56),
('well', 55),
('made', 54),
('great', 54),
('wife', 54),
('than', 53),
('your', 53),
('cried', 53),
('princess', 51),
('day', 51),
('took', 51),
('children', 50),
('an', 50),
('mother', 48),
('poor', 48),
```

```
('how', 47),
('after', 47),
('cinderella', 47),
('come', 46),
('way', 46),
('now', 45),
('bed', 44),
('put', 44),
('away', 43),
('home', 43),
('here', 42),
('once', 41),
('or', 41),
('father', 41),
('red', 41),
('kings', 40),
('off', 40),
('about', 39),
('saw', 39),
('however', 38),
('just', 38),
('house', 38),
('long', 38),
('must', 38),
('take', 37),
('other', 37),
('we', 37),
('queen', 36),
('because', 35),
('herself', 35),
('ogre', 35),
('gave', 34),
('old', 34),
('should', 34),
('ran', 34),
('such', 34),
('girl', 34),
('before', 33),
('gold', 33),
('young', 33),
('son', 33),
('fell', 33),
('grandmother', 33),
('wolf', 33),
('can', 32),
('only', 32),
('began', 32),
('shall', 32),
```

```
('cat', 32),
('upon', 31),
('every', 31),
('asked', 31),
('while', 31),
('himself', 31),
('last', 30),
('might', 30),
('never', 30),
('heard', 30),
('am', 30),
('much', 30),
('large', 30),
('told', 30),
('make', 30),
('door', 30),
('daughter', 29),
('over', 29),
('thought', 29),
('give', 29),
('like', 29),
('eyes', 29),
('castle', 28),
('nothing', 28),
('know', 28),
('cap', 28),
('woman', 27),
('brought', 27),
('prince', 27),
('blue', 27),
('hand', 26),
('these', 26),
('three', 26),
('us', 26),
('got', 26),
('dress', 26),
('done', 25),
('without', 25),
('found', 25),
('soon', 25),
('through', 25),
('having', 25),
('night', 25),
('even', 25),
('golden', 24),
('left', 24),
('silver', 24),
('immediately', 24),
```

```
('beard', 24),
('thumb', 24),
('everything', 23),
('seen', 23),
('most', 23),
('replied', 23),
('too', 23),
('until', 23),
('same', 23),
('back', 23),
('clothes', 23),
('skin', 23),
('years', 22),
('has', 22),
('let', 22),
('first', 22),
('called', 22),
('get', 22),
('godmother', 22),
('child', 21),
('big', 21),
('fairy', 20),
('being', 20),
('room', 20),
('oh', 20),
('find', 20),
('forest', 20),
('able', 20),
('morning', 20),
('sister', 20),
('our', 20),
('answered', 20),
('brothers', 20),
('sisters', 20),
('horse', 20),
('thumbling', 20),
('head', 19),
('happened', 19),
('man', 19),
('still', 19),
('look', 19),
('better', 19),
('eat', 19),
('opened', 19),
('maiden', 19),
('donkey', 19),
('marquis', 19),
('tree', 19),
```

```
('each', 18),
('fine', 18),
('any', 18),
('world', 18),
('moment', 18),
('hands', 18),
('always', 18),
('another', 18),
('cook', 18),
('looked', 18),
('master', 18),
('bring', 18),
('heart', 18),
('ring', 18),
('seven', 17),
('help', 17),
('thus', 17),
('small', 17),
('many', 17),
('wanted', 17),
('true', 16),
('work', 16),
('passed', 16),
('say', 16),
('whole', 16),
('open', 16),
('glass', 16),
('quite', 16),
('hair', 16),
('ball', 16),
('pick', 16),
('tell', 15),
('turn', 15),
('returned', 15),
('set', 15),
('far', 15),
('next', 15),
('under', 15),
('sat', 15),
('taken', 15),
('boots', 15),
('dear', 15),
('coming', 15),
('love', 15),
('rest', 15),
('eaten', 15),
('own', 15),
('blood', 15),
```

```
('ill', 15),
('run', 15),
('looking', 15),
('lord', 15),
('soup', 15),
('hans', 15),
('shoe', 15),
('beauty', 14),
('whom', 14),
('dance', 14),
('going', 14),
('sleep', 14),
('asleep', 14),
('ever', 14),
('ate', 14),
('ordered', 14),
('moral', 14),
('husband', 14),
('both', 14),
('foot', 14),
('anne', 14),
('ashes', 14),
('cut', 14),
('right', 14),
('people', 13),
('place', 13),
('voice', 13),
('indeed', 13),
('also', 13),
('jumped', 13),
('gone', 13),
('evening', 13),
('daughters', 13),
('sun', 13),
('its', 13),
('carabas', 13),
('along', 13),
('drink', 13),
('woods', 13),
('bride', 13),
('days', 12),
('others', 12),
('behind', 12),
('filled', 12),
('threw', 12),
('hour', 12),
('turned', 12),
('follow', 12),
```

```
('knew', 12),
('things', 12),
('everyone', 12),
('leave', 12),
('afraid', 12),
('spoke', 12),
('white', 12),
('often', 12),
('may', 12),
('cake', 12),
('pulled', 12),
('key', 12),
('frightened', 12),
('inside', 12),
('wood', 11),
('lived', 11),
('held', 11),
('fall', 11),
('upstairs', 11),
('face', 11),
('since', 11),
('wish', 11),
('pleased', 11),
('lost', 11),
('already', 11),
('court', 11),
('sent', 11),
('though', 11),
('become', 11),
('received', 11),
('hold', 11),
('begged', 11),
('together', 11),
('dont', 11),
('full', 11),
('cloak', 11),
('finally', 11),
('allkindsoffur', 11),
('frau', 11),
('country', 10),
('palace', 10),
('diamonds', 10),
('dead', 10),
('speak', 10),
('hundred', 10),
('horses', 10),
('whose', 10),
('story', 10),
```

```
('supper', 10),
('dressed', 10),
('bread', 10),
('ogres', 10),
('steward', 10),
('animal', 10),
('order', 10),
('riding', 10),
('mill', 10),
('taking', 10),
('ground', 10),
('coach', 10),
('yourself', 10),
('lay', 10),
('remained', 10),
('why', 10),
('became', 10),
('sit', 10),
('kitchen', 10),
('yes', 10),
('stepmother', 10),
('joy', 10),
('dirty', 10),
('tailor', 10),
('words', 9),
('those', 9),
('magnificent', 9),
('knife', 9),
('kind', 9),
('die', 9),
('enough', 9),
('end', 9),
('keep', 9),
('doing', 9),
('water', 9),
('certain', 9),
('lying', 9),
('seeing', 9),
('themselves', 9),
('continued', 9),
('sight', 9),
('almost', 9),
('less', 9),
('half', 9),
('married', 9),
('easily', 9),
('meat', 9),
('anything', 9),
```

```
('mind', 9),
('hood', 9),
('stay', 9),
('arrived', 9),
('saying', 9),
('among', 9),
('longer', 9),
('god', 9),
('gate', 9),
('marry', 9),
('meadow', 9),
('further', 9),
('finished', 9),
('mouth', 9),
('want', 9),
('fathers', 9),
('slipper', 9),
('cannot', 9),
('light', 9),
('walked', 9),
('beneath', 9),
('finger', 9),
('reel', 9),
('holle', 9),
('tried', 8),
('fairies', 8),
('invited', 8),
('table', 8),
('believed', 8),
('tears', 8),
('instead', 8),
('alone', 8),
('closed', 8),
('life', 8),
('fire', 8),
('wand', 8),
('boys', 8),
('ready', 8),
('neither', 8),
('fear', 8),
('carried', 8),
('towards', 8),
('path', 8),
('entered', 8),
('reached', 8),
('slept', 8),
('scarcely', 8),
('served', 8),
```

```
('later', 8),
('word', 8),
('care', 8),
('yet', 8),
('pot', 8),
('six', 8),
('money', 8),
('hall', 8),
('rich', 8),
('haste', 8),
('picked', 8),
('times', 8),
('happy', 8),
('forced', 8),
('feet', 8),
('alas', 8),
('thieves', 8),
('corner', 8),
('danced', 8),
('woodcutter', 8),
('chest', 8),
('fur', 8),
('lazy', 8),
('stag', 8),
('company', 7),
('placed', 7),
('hid', 7),
('third', 7),
('loud', 7),
('use', 7),
('spinning', 7),
('running', 7),
('top', 7),
('try', 7),
('except', 7),
('nor', 7),
('nobody', 7),
('believe', 7),
('surprised', 7),
('met', 7),
('side', 7),
('several', 7),
('ladies', 7),
('standing', 7),
('trembling', 7),
('drew', 7),
('awoke', 7),
('manner', 7),
```

```
('loved', 7),
('think', 7),
('although', 7),
('felt', 7),
('greatest', 7),
('few', 7),
('making', 7),
('led', 7),
('return', 7),
('brother', 7),
('dared', 7),
('early', 7),
('fast', 7),
('grandmothers', 7),
('best', 7),
('closet', 7),
('covered', 7),
('against', 7),
('anyone', 7),
('quickly', 7),
('part', 7),
('stone', 7),
('spring', 7),
('wedding', 7),
('delighted', 7),
('fit', 7),
('tired', 7),
('parents', 7),
('birds', 7),
('something', 7),
('promise', 7),
('heaven', 7),
('request', 7),
('moon', 7),
('strange', 7),
('cubbyhole', 7),
('stood', 7),
('wolfs', 7),
('stranger', 7),
('bird', 7),
('pigeons', 7),
('honor', 6),
('enter', 6),
('orders', 6),
('impossible', 6),
('creature', 6),
('teeth', 6),
('gift', 6),
```

```
('youngest', 6),
('declared', 6),
('stepped', 6),
('length', 6),
('royal', 6),
('color', 6),
('thousand', 6),
('touched', 6),
('distance', 6),
('seemed', 6),
('wine', 6),
('chamber', 6),
('near', 6),
('especially', 6),
('telling', 6),
('hunting', 6),
('marriage', 6),
('boy', 6),
('died', 6),
('afterwards', 6),
('madam', 6),
('neck', 6),
('yard', 6),
('carry', 6),
('certainly', 6),
('wolves', 6),
('crying', 6),
('show', 6),
('butter', 6),
('live', 6),
('flowers', 6),
('knocked', 6),
('girls', 6),
('lady', 6),
('short', 6),
('floor', 6),
('vain', 6),
('rock', 6),
('softly', 6),
('body', 6),
('fellow', 6),
('handsome', 6),
('bag', 6),
('fortune', 6),
('pass', 6),
('decided', 6),
('millers', 6),
('till', 6),
```

```
('pearls', 6),
('thing', 6),
('appeared', 6),
('drove', 6),
('new', 6),
('garden', 6),
('showed', 6),
('ah', 6),
('dropped', 6),
('climbed', 6),
('fresh', 6),
('wept', 6),
('cart', 6),
('wished', 6),
('shake', 6),
('andsaid', 6),
('peep', 6),
('banquet', 5),
('person', 5),
('play', 5),
('music', 5),
('forth', 5),
('death', 5),
('knowing', 5),
('sooner', 5),
('kingdom', 5),
('instantly', 5),
('within', 5),
('guards', 5),
('needed', 5),
('round', 5),
('trees', 5),
('during', 5),
('force', 5),
('peasant', 5),
('glasses', 5),
('fallen', 5),
('tender', 5),
('considering', 5),
('dawn', 5),
('second', 5),
('wealth', 5),
('difficulty', 5),
('husbands', 5),
('summoned', 5),
('tomorrow', 5),
('ogress', 5),
('wicked', 5),
```

```
('determined', 5),
('kill', 5),
('rode', 5),
('waited', 5),
('sure', 5),
('advice', 5),
('hear', 5),
('does', 5),
('tap', 5),
('shut', 5),
('greatly', 5),
('kinds', 5),
('women', 5),
('houses', 5),
('desired', 5),
('dancing', 5),
('gentleman', 5),
('journey', 5),
('friends', 5),
('promised', 5),
('embraced', 5),
('happiness', 5),
('thinking', 5),
('eldest', 5),
('myself', 5),
('grave', 5),
('air', 5),
('majesty', 5),
('thank', 5),
('cry', 5),
('noise', 5),
('belongs', 5),
('astonished', 5),
('form', 5),
('mouse', 5),
('followed', 5),
('elder', 5),
('bad', 5),
('else', 5),
('earth', 5),
('entirely', 5),
('matter', 5),
('alive', 5),
('hurried', 5),
('loudly', 5),
('crowns', 5),
('speaking', 5),
('piece', 5),
```

```
('heads', 5),
('hard', 5),
('sought', 5),
('shining', 5),
('farm', 5),
('servants', 5),
('servant', 5),
('washed', 5),
('fingers', 5),
('mothers', 5),
('dresses', 5),
('stars', 5),
('around', 5),
('parlor', 5),
('baked', 5),
('crept', 5),
('apples', 5),
('obeyed', 5),
('shoes', 5),
('hazeltree', 5),
('maid', 5),
('sleeping', 4),
('due', 4),
('means', 4),
('bestow', 4),
('sitting', 4),
('tower', 4),
('nearby', 4),
('evil', 4),
('presently', 4),
('spindle', 4),
('terrible', 4),
('slumber', 4),
('wheel', 4),
('quarter', 4),
('remembered', 4),
('finest', 4),
('apartment', 4),
('fair', 4),
('commanded', 4),
('twelve', 4),
('pair', 4),
('sevenleague', 4),
('step', 4),
('drawn', 4),
('awakened', 4),
('stables', 4),
('courtyard', 4),
```

```
('mistress', 4),
('whenever', 4),
('vast', 4),
('whatever', 4),
('family', 4),
('tale', 4),
('thither', 4),
('catch', 4),
('devoured', 4),
('hardly', 4),
('none', 4),
('appearance', 4),
('four', 4),
('straight', 4),
('truth', 4),
('courtiers', 4),
('fetch', 4),
('expected', 4),
('sauce', 4),
('answer', 4),
('hiding', 4),
('rage', 4),
('command', 4),
('hidden', 4),
('satisfied', 4),
('recognized', 4),
('instant', 4),
('sorry', 4),
('strong', 4),
('im', 4),
('wait', 4),
('matters', 4),
('village', 4),
('working', 4),
('bobbin', 4),
('latch', 4),
('stool', 4),
('sweet', 4),
('furniture', 4),
('ugly', 4),
('bear', 4),
('wives', 4),
('stayed', 4),
('send', 4),
('jewels', 4),
('walls', 4),
('wash', 4),
('clean', 4),
```

('fail', 4),
('pardon', 4),
('beg', 4),
('dust', 4),
('sheep', 4),
('prepared', 4),
('strike', 4),
('miller', 4),
('concerned', 4),
('imagine', 4),
('mice', 4),
('killed', 4),
('pity', 4),
('ahead', 4),
('ask', 4),
('change', 4),
('caused', 4),
('walking', 4),
('rat', 4),
('worth', 4),
('win', 4),
('younger', 4),
('bringing', 4),
('pitcher', 4),
('thats', 4),
('heavens', 4),
('escaped', 4),
('appear', 4),
('wear', 4),
('slender', 4),
('front', 4),
('started', 4),
('pumpkin', 4),
('coachman', 4),
('slippers', 4),
('midnight', 4),
('alighted', 4),
('unknown', 4),
('glad', 4),
('sound', 4),
('ten', 4),
('pebbles', 4),
('gathered', 4),
('slipped', 4),
('farther', 4),
('high', 4),
('candle', 4),
('hollow', 4),

```
('smell', 4),
('branch', 4),
('entire', 4),
('sky', 4),
('disappeared', 4),
('eye', 4),
('sick', 4),
('size', 4),
('deceased', 4),
('glistened', 4),
('huntsmen', 4),
('soot', 4),
('approached', 4),
('furry', 4),
('bowl', 4),
('trough', 4),
('flew', 4),
('stuck', 4),
('pitch', 4),
('tothe', 4),
('hut', 4),
('coat', 4),
('courage', 4),
('horns', 4),
('axe', 4),
('scythe', 4),
('withhim', 4),
('lentils', 4),
('hat', 4),
('amongst', 4),
('cow', 4),
('grand', 3),
('imaginable', 3),
('laid', 3),
('aged', 3),
('between', 3),
('seated', 3),
('bestowed', 3),
('tapestry', 3),
('dying', 3),
('awaken', 3),
('forbidding', 3),
('pain', 3),
('pleasure', 3),
('serving', 3),
('pretty', 3),
('feeling', 3),
('embroidered', 3),
```

```
('behold', 3),
('lovely', 3),
('lips', 3),
('gentle', 3),
('therefore', 3),
('plain', 3),
('leagues', 3),
('cover', 3),
('approaching', 3),
('gentlemen', 3),
('footmen', 3),
('beside', 3),
('partridges', 3),
('grew', 3),
('brambles', 3),
('magic', 3),
('throne', 3),
('pursue', 3),
('waits', 3),
('resolved', 3),
('direction', 3),
('notice', 3),
('brave', 3),
('gaze', 3),
('silence', 3),
('men', 3),
('animals', 3),
('perceived', 3),
('staircase', 3),
('shoulders', 3),
('waiting', 3),
('curtains', 3),
('seems', 3),
('hours', 3),
('succeeded', 3),
('hungry', 3),
('rise', 3),
('played', 3),
('need', 3),
('black', 3),
('nature', 3),
('secret', 3),
('bore', 3),
('whispered', 3),
('war', 3),
('appointed', 3),
('dinner', 3),
('raw', 3),
```

```
('serve', 3),
('plainly', 3),
('arms', 3),
('burst', 3),
('hide', 3),
('quarters', 3),
('search', 3),
('possibly', 3),
('save', 3),
('given', 3),
('leaving', 3),
('appetite', 3),
('hope', 3),
('huge', 3),
('vat', 3),
('cast', 3),
('free', 3),
('late', 3),
('comes', 3),
('delay', 3),
('dangerous', 3),
('talk', 3),
('whos', 3),
('somewhat', 3),
('ears', 3),
('charming', 3),
('quiet', 3),
('polite', 3),
('ones', 3),
('gilded', 3),
('neighbors', 3),
('backwards', 3),
('forwards', 3),
('concluded', 3),
('consequence', 3),
('rooms', 3),
('ceased', 3),
('nearly', 3),
('broke', 3),
('road', 3),
('advantage', 3),
('considered', 3),
('name', 3),
('today', 3),
('sign', 3),
('cloud', 3),
('green', 3),
('grass', 3),
```

```
         ...]

In [34]: dtm = cv.fit_transform(tales)
         dtm

Out[34]: <16x2978 sparse matrix of type '<class 'numpy.int64'>'
                 with 6385 stored elements in Compressed Sparse Row format>

In [35]: desparse = dtm.toarray()

         word_list = cv.get_feature_names()

         dtm_tb = Table(word_list).with_rows(desparse).drop(stop_words)
         dtm_tb

Out[35]: abandon | abandoned | able | absence | absolutely | abundance | accepted | accident |
         0       | 0         | 1    | 0       | 0          | 0          | 0         | 0        |

In [36]: toks_tab = Table()
         toks_tab.append_column(label="Word List", values=word_list)
         toks_tab.append_column(label="Frequency", values=sum(desparse))
         row_sums = np.sum(desparse, axis=1)
         normed = desparse/row_sums[:,None]
         dtm_tb = Table(word_list).with_rows(normed)
         dtm_tb

Out[36]: abandon | abandoned  | able       | absence     | absolutely  | abundance  | accepted
```

```
In [37]: cv = CountVectorizer(stop_words = stop_words)
         dtm = cv.fit_transform(tales)
         tt = TfidfTransformer(norm='l1',use_idf=False)
         dtm_tf = tt.fit_transform(dtm)

         word_list = cv.get_feature_names()
         dtm_array = dtm_tf.toarray()
         Table(word_list).with_rows(dtm_array)
```

Out[37]: abandon | abandoned | able | absence | absolutely | abundance | accepted

```
In [38]: dist_matrix = pairwise.pairwise_distances(dtm_tf, metric='euclidean')
         title_list = ['Story' +str(i+1) for i in range(len(tales))]
         Table(title_list).with_rows(dist_matrix)

         kmeans = KMeans(n_clusters=2)
         kmeans.fit_predict(dist_matrix)
```
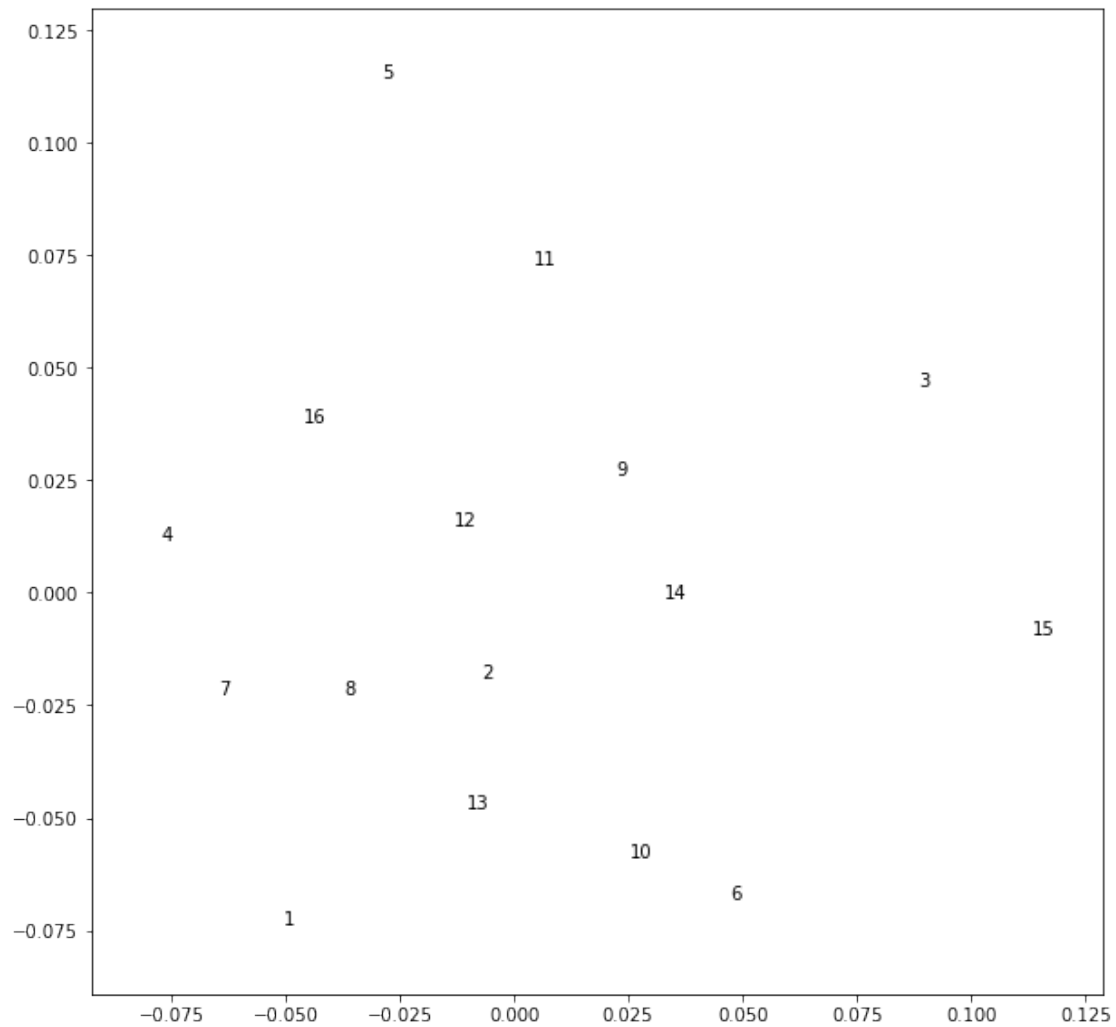
Out[38]: array([0, 0, 1, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0], dtype=int32)

In [39]: tales[1][:22], tales[9][:21], tales[15][:15]

Out[39]: ('The Sleeping Beauty in', '"The Master Cat; or, ', 'Blue Beard,"The')

```
In [40]: mds = MDS(n_components = 2, dissimilarity="precomputed")
         embeddings = mds.fit_transform(dist_matrix)

         _, ax = plt.subplots(figsize=(10,10))
         ax.scatter(embeddings[:,0], embeddings[:,1], alpha=0)
         for i in range(16):
             ax.annotate(i+1, ((embeddings[i,0], embeddings[i,1])))
```

```
In [41]: perrault_object = TextBlob(perrault2)
         [s for s in perrault_object.sentences]
         p_sentiment = 0
         for s in perrault_object.sentences:
             p_sentiment = p_sentiment+ s.sentiment.polarity

         p_sentiment / len(perrault_object.sentences)

Out[41]: 0.10361619597628512

In [42]: grimm_object = TextBlob(grimm2)
         [h for h in grimm_object.sentences]
         g_sentiment = 0
         for h in grimm_object.sentences:
             g_sentiment = g_sentiment+ h.sentiment.polarity
```

```
        g_sentiment / len(grimm_object.sentences)
```

Out[42]: 0.08507620694715379