

CS 751: Assignment 4

Bharath Kongara

Spring 2015

Contents

1	Question 1	2
	1.1 Solution	2
	1.2 Code Listing	2
	1.3 Examples	5
2	Question 2	7
	2.1 Solution	7
3	Question 3	11
	3.1 Solution	11
4	Question 4	24
	4.1 Solution	24

1 Question 1

- Using the pages from A3 that boilerpipe successfully processed, download those representations again & reprocess them with boilerpipe.:
- Document the time difference (e.g., $\text{Time}(A4) - \text{Time}(A3)$).
- Compute the Jaccard Distance x for each pair of pages (i.e., $P(A3)$ & $P(A4)$) for:
 - Unique terms (i.e., unigrams)
 - Bigrams
 - Trigrams
- See:http://en.wikipedia.org/wiki/Jaccard_index
- For each of the 3 cases (i.e., 1-, 2-, 3-grams) build a Cumulative Distribution Function that shows the % change on the x-axis & the % of the population on the x-axis
- See:http://en.wikipedia.org/wiki/Cumulative_distribution_function
- Give 3-4 examples illustrating the range of change that you have measured.

1.1 Solution

- I used the same script 'getText.py' from A3 which uses 'jusText'[2] to remove templates from HTML files of 10000 URIs.
- 'jusText'[2] is a tool for removing boilerplate content such as navigation links, headers and footers from HTML pages.
- The time difference between processing boilerpipe of A3 and A4 was 27 days.

1.2 Code Listing

Here is the python code that generate n-grams and computes Jaccard Distance.

```
1 from nltk.util import ngrams
2 import json
3 import sys
4 import os
5 import string
6
7 def jaccardDistance(list1, list2):
8     union = set(list1).union(list2)
9     intersect = set(list1).intersection(list2)
10    dist = (len(union) - len(intersect)) * 1.0 / len(union)
11    return dist
12
13 def wordsFromFile(filePath):
14     text = ''
15     print filePath
16     with open(filePath, 'r') as textFile:
17         for line in textFile:
18             text = text + extractWords(line).strip() + ' '
19     words = text.strip().split(' ')
20     if '' in words:
21         words.remove('')
22     return words
23
24 def calculateNgrams():
25     finalGrams = ngrams(sentence.split(), n)
26     for grams in finalGrams:
27         print grams
28
29 def extractWords(s):
30     valid_characters = string.ascii_letters
31     valid_following_character = "'-_"
32     letters = []
```

```

32     for letter in s:
33         if letter in valid_characters:
34             letters.append(letter)
35         elif len(letters) > 0 and \
36             letter in valid_following_character and \
37             letters[-1] in valid_characters:
38             letters.append(letter)
39         elif len(letters) > 0 and letters[-1] != ' ':
40             letters.append(' ')
41     return ''.join(letters)
42
43 mementosDir = 'Texts'
44 mementosDir2 = 'Texts2'
45 outputfile = open('DifferenceUrl', 'w')
46 outputfile1 = open('UnigramJackardDistance', 'w')
47 outputfile2 = open('BigramJackardDistance', 'w')
48 outputfile3 = open('TrigramJackardDistance', 'w')
49 filesList = os.listdir(mementosDir)
50 filesList2 = os.listdir(mementosDir2)
51 for item in filesList:
52     if item in filesList2:
53         path1 = mementosDir+'/'+item
54         path2 = mementosDir2+'/'+item
55         words1 = wordsFromFile(path1)
56         words2 = wordsFromFile(path2)
57         words1.sort()
58         words2.sort()
59         bigrams1 = []
60         bigrams2 = []
61         trigrams1 = []
62         trigrams2 = []
63         for i in range(0, len(words1)):
64             if i >= 1:
65                 bigrams1.append(words1[i-1] + ' ' + words1[i])
66             if i >= 2:
67                 trigrams1.append(words1[i-2] + ' ' + words1[i-1] + ' ' + words1[i])
68         for i in range(0, len(words2)):
69             if i >= 1:
70                 bigrams2.append(words2[i-1] + ' ' + words2[i])
71             if i >= 2:
72                 trigrams2.append(words2[i-2] + ' ' + words2[i-1] + ' ' + words2[i])
73         jacardDistanceUni = jaccardDistance(words1, words2)
74         jacardDistanceUni2 = jaccardDistance(bigrams1, bigrams2)
75         jacardDistanceUni3 = jaccardDistance(trigrams1, trigrams2)
76         outputfile1.write(str(jacardDistanceUni)+ '\n')
77         outputfile2.write(str(jacardDistanceUni2)+ '\n')
78         outputfile3.write(str(jacardDistanceUni3)+ '\n')
79     else:
80         outputfile.write(item+ '\n')

```

Listing 1: Python program to find unigrams, bigrams, trigrams

I referred to a sample of code that constructs CDF from r-bloggers[1] website.
Here is the R code which plots CDF for unigrams,bigrams,trigrams.

```

1 library(RColorBrewer)
2 data <- read.table("TrigramJaccardDistance.txt")
3 finalData = rep(data[,1])
4 title <- "CDF of Change in Unigrams"
5 xlab <- "Jaccard Distance"
6 ylab <- "Probability"
7 png("Trigramq1.png")
8 par(mar=c(4,4,2.5,2) + 0.1)
9 set.seed(1)
10 mp <- plot(finalData, 1:length(finalData), type="l", col="red", main=title, xlab=xlab, ylab=
11         ylab, yaxt="n")
12 axis(side=2, at=seq(1,2735,length.out=5), labels=seq(0.2,1,by=0.2))

```

Listing 2: R program to plot unigrams, bigrams, trigrams change

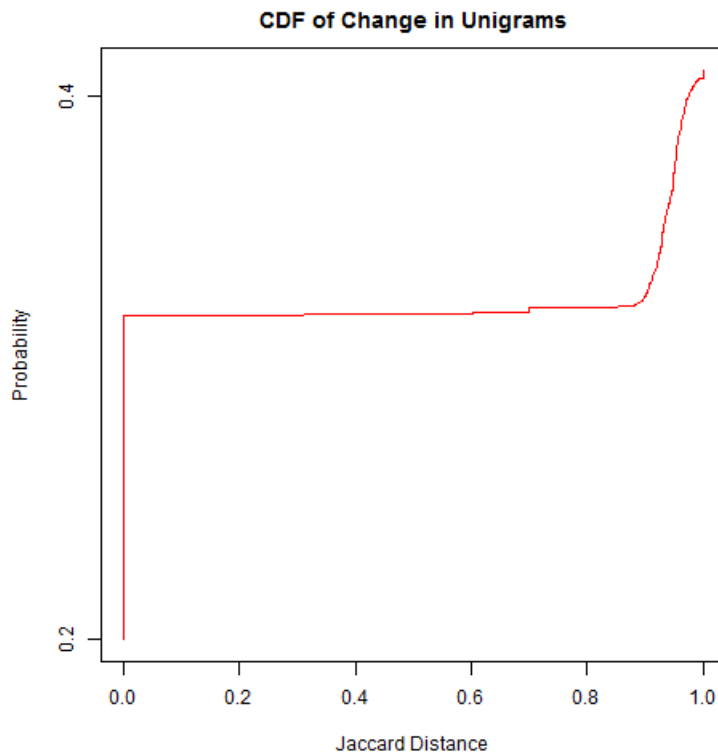


Figure 1: CDF of change in unigrams

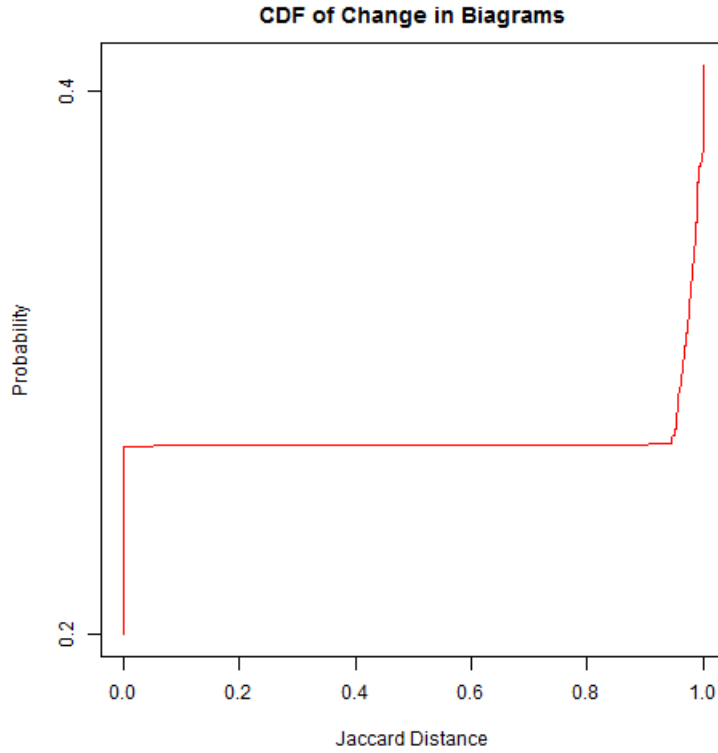


Figure 2: CDF of change in bigrams

1.3 Examples

Below are 3 examples with data that shows range of change.

- http://www.foxnews.com/us/2015/02/09/judge-rules-utah-woman-accused-leaving-baby-in-trash-is-incom?cmpid=cmt_twitter_fn

Unigrams	Bigrams	Trigrams
0.971	0.700	0.789

- http://www.ebay.com/itm/like/301515875453?item=301515875453&lgeo=1&utm_medium=twitter&utm_source=dlvr.it&vectorid=229466&rmvSB=true

Unigrams	Bigrams	Trigrams
0.471	0.600	0.689

- <http://du3a.org>

Unigrams	Bigrams	Trigrams
0.771	0.890	0.880

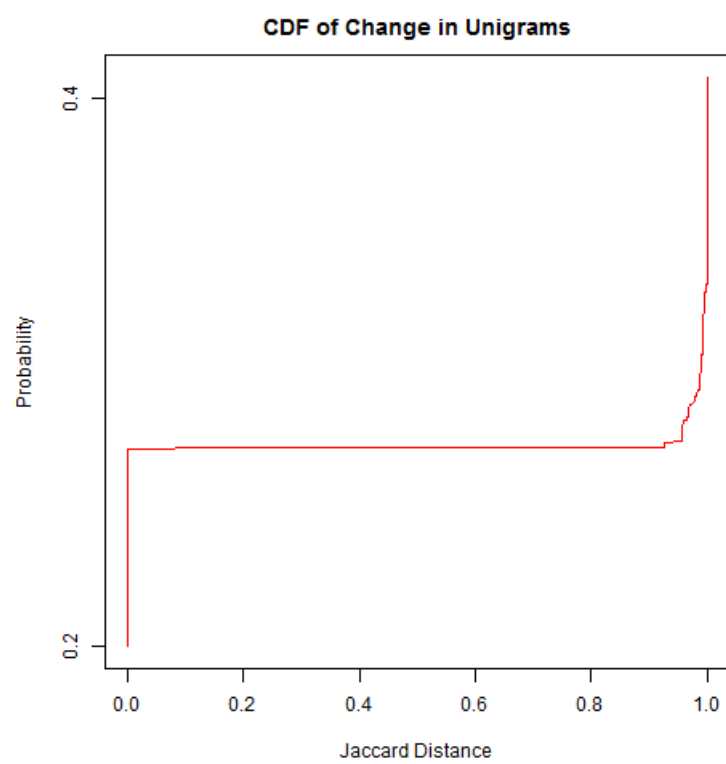


Figure 3: CDF of change in trigrams

2 Question 2

- Using the pages from Q1 (A4), download all TimeMaps (including TimeMaps with 404 responses, i.e. empty or null TimeMaps)
 - Upload all the TimeMaps to github
- Build a CDF for # of mementos for each original URI (i.e., x-axis = # of mementos, y-axis = % of links)
- See: <http://timetravel.mementoweb.org/guide/api/>

2.1 Solution

- I used 'timelapse'[3] library to retrieve TimeMaps of each of the URIs.
- 'timelapse'[3] library extracts memento created date along with memento.
- Out of 9,800 URIs, TimeMaps are retrieved for 8200 which includes empty files.
- I ran multiple instances to retrieve TimeMaps as aggregator was taking time.
- Few times retrieval stopped because of request time out error. I waited for a few minutes before running it again.
- I have uploaded TimeMaps of each of the URIs to Github.

```
1  # -*- coding: utf-8 -*-
2  #!/usr/bin/env python
3  from getConfig import getConfigParameters
4  import commands
5  import time
6  import datetime
7  import sys
8  import argparse, os
9  import subprocess
10 import hashlib
11 import tldextract
12 import urlparse
13 import glob
14 import json
15 import requests
16
17
18 globalMementoUrlDateTimeDelimiter = "+++"
19
20 def getMementosPages(url):
21
22     pages = []
23     url = url.strip()
24     if(len(url)>0):
25
26         firstChoiceAggregator = getConfigParameters('mementoAggregator')
27         timemapPrefix = firstChoiceAggregator + url
28         #timemapPrefix = 'http://mementoproxy.cs.odu.edu/aggr/timemap/link/1/' + url
29
30         '''
31
32         The CS memento aggregator payload format:
33             [memento, ..., memento, timemap1]; timemap1 points to next
34             page
35
36         The LANL memento aggregator payload format:
37             1. [timemap1, ..., timemapN]; timemapX points to mementos
38                list
39             2. [memento1, ..., mementoN]; for small payloads
40
41         For LANL Aggregator: The reason the link format is used after
42             retrieving the payload
43
44             with json format is due to
45             the fact that the
46             underlying code is
47             based
```


*on the link format
structure. json format
was not always the norm*

```

'''
#select an aggregator - start
aggregatorSelector = ''

co = 'curl --silent -I ' + timemapPrefix
head = commands.getoutput(co)

indexOfFirstNewLine = head.find('\n')
if( indexOfFirstNewLine > -1 ):

    if( head[:indexOfFirstNewLine].split(' ')[1] != '200' ):
        firstChoiceAggregator = getConfigParameters('
            latentMementoAggregator')
        timemapPrefix = firstChoiceAggregator + url

    if( firstChoiceAggregator.find('cs.odu.edu') > -1 ):
        aggregatorSelector = 'CS'
    else:
        aggregatorSelector = 'LANL'

print '... using aggregator:', aggregatorSelector
#select an aggregator - end

#CS aggregator
if( aggregatorSelector == 'CS' ):
    while( True ):
        #old: co = 'curl --silent ' + timemapPrefix
        #old: page = commands.getoutput(co)

        page = ''
        r = requests.get(timemapPrefix)
        print 'status code:', r.status_code
        if( r.status_code == 200 ):
            page = r.text

        pages.append(page)
        indexOfRelTimemapMarker = page.rfind('>;rel="timemap"')

        if( indexOfRelTimemapMarker == -1 ):
            break
        else:
            #retrieve next timemap for next page of mementos e.g
            #retrieve url from <http://mementoproxy.cs.odu.
            #edu/aggr/timemap/link/10001/http://www.cnn.com>;
            #rel="timemap"
            i = indexOfRelTimemapMarker -1
            timemapPrefix = ''
            while( i > -1 ):
                if( page[i] != '<' ):
                    timemapPrefix = page[i] +
                        timemapPrefix

                else:
                    break
                i = i - 1

else:
    #LANL Aggregator
    #old: co = 'curl --silent ' + timemapPrefix
    #old: page = commands.getoutput(co)

    page = ''
    r = requests.get(timemapPrefix)
    if( r.status_code == 200 ):
        page = r.text

    try:

```

```

103         payload = json.loads(page)
104
105         if 'timemap_index' in payload:
106
107             for timemap in payload['timemap_index']:
108
109                 timemapLink = timemap['uri'].replace('/timemap/json/', '/timemap/link/')
110                 #old: co = 'curl --silent ' + timemapLink
111                 #old: page = commands.getoutput(co)
112                 #old: pages.append(page)
113                 r = requests.get(timemapLink)
114                 if( r.status_code == 200 ):
115                     pages.append(r.text)
116
117             elif 'mementos' in payload:
118                 #untested block
119                 timemapLink = payload['timemap-uri']['json-format'].
120                     replace('/timemap/json/', '/timemap/link/')
121                 #old: co = 'curl --silent ' + timemapLink
122                 #old: page = commands.getoutput(co)
123                 #old: pages.append(page)
124
125                 print 'timemap:', timemapLink
126                 r = requests.get(timemapLink)
127                 if( r.status_code == 200 ):
128                     pages.append(r.text)
129
130
131         except:
132             exc_type, exc_obj, exc_tb = sys.exc_info()
133             fname = os.path.split(exc_tb.tb_frame.f_code.co_filename)[1]
134             print(fname, exc_tb.tb_lineno, sys.exc_info() )
135
136
137     return pages
138
139
140 def getItemGivenSignature(page):
141
142     listOfItems = []
143     if( len(page) > 0 ):
144         page = page.splitlines()
145         for line in page:
146             if(line.find('memento';') != -1):
147                 #uriRelDateTime: ['<http://www.webcitation.org/64ta04WpM>',
148                     ' rel="first memento"', ' datetime="Mon, 23 Jan 2012
149                     02:01:29 GMT"', ']
150                 uriRelDateTime = line.split(';')
151                 if( len(uriRelDateTime) > 2 ):
152                     if( uriRelDateTime[0].find('/://') != -1 ):
153                         if( uriRelDateTime[2].find('datetime="') !=
154                             -1 ):
155
156                             uri = ''
157                             uri = uriRelDateTime[0].split('<')
158                             #print uri
159                             if( len(uri) > 1 ):
160                                 uri = uri[1].replace('>', ''
161                                     )
162                                 uri = uri.strip()
163
164                             datetime = ''
165                             datetime = uriRelDateTime[2].split('
166                                 "')

```

```

167         if( len(uri) != 0 and len(datetime)
168             != 0 ):
169             #print uri, '---', datetime
170             listOfItems.append(uri +
171                                 globalMementoUrlDateTimeDelimiter
172                                 + datetime)
173
174     return listOfItems
175
176 fh = open('allurls.txt','r+')
177 saveFile1 = open("memontoscount.txt", 'w')
178 count = 0
179 l = 48
180 count1 = 1
181 for line in fh:
182     data = json.loads(line)
183     count = 0
184     url = data['finalurl']
185     pages = getMementosPages(url)
186     saveFile = open("memontos"+str(l)+".txt", 'w')
187     for i in range(0, len(pages)):
188         mementos = getItemGivenSignature(pages[i])
189         print mementos
190         count += len(mementos)
191         saveFile.write(str(mementos) + '\n')
192     print 'total count of mementos:', count
193     saveFile1 = open("count.txt", 'a')
194     saveFile1.write(str(count) + '\n')
195     saveFile1.close()
196     l=l+1
197     count1 = count1 + 1
198     count = 0

```

Listing 3: Python program to retrieve TimeMap of each of the URIs

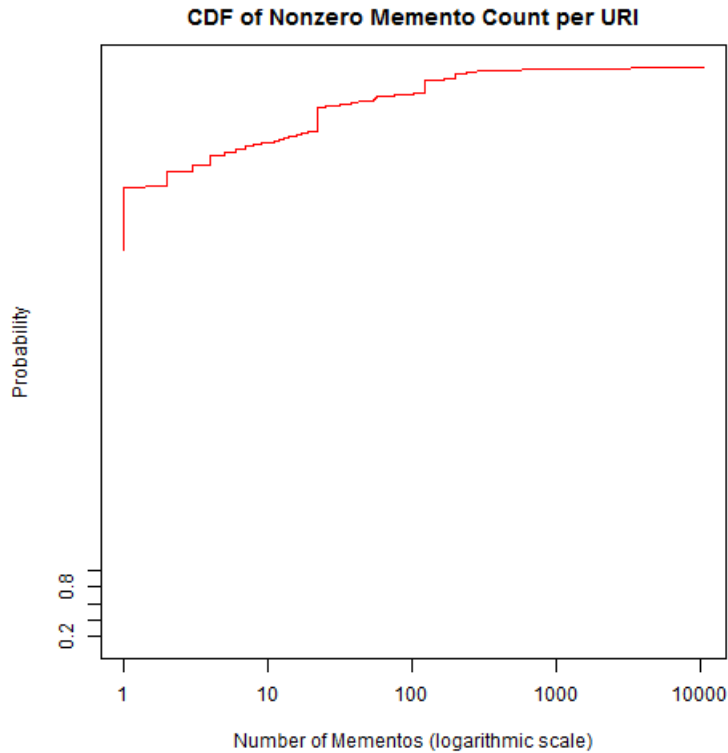


Figure 4: CDF of change in unigrams

3 Question 3

- Using 20 links that have TimeMaps
 - With ≥ 20 mementos
 - Have existed ≥ 2 years (i.e., Memento-Datetime of first memento is April XX, 2013 or older)
 - Note: select from Q1/Q2 links, else choose them by hand
- For each link, create a graph that shows Jaccard Distance, relative to the first memento, through time
 - x-axis: continuous time, y-axis: Jaccard Distance relative to the first memento

3.1 Solution

- I downloaded all mementos for each URI retrieved and processed with boilerpipe.
- With that data I have calculated Jaccard Distance relative to the first memento data.
- Few mementos are totally different which means Jaccard Distance is one and few of them have significant change through time.

Below are graphs that shows Jaccard Distance relative to the first memento through time.

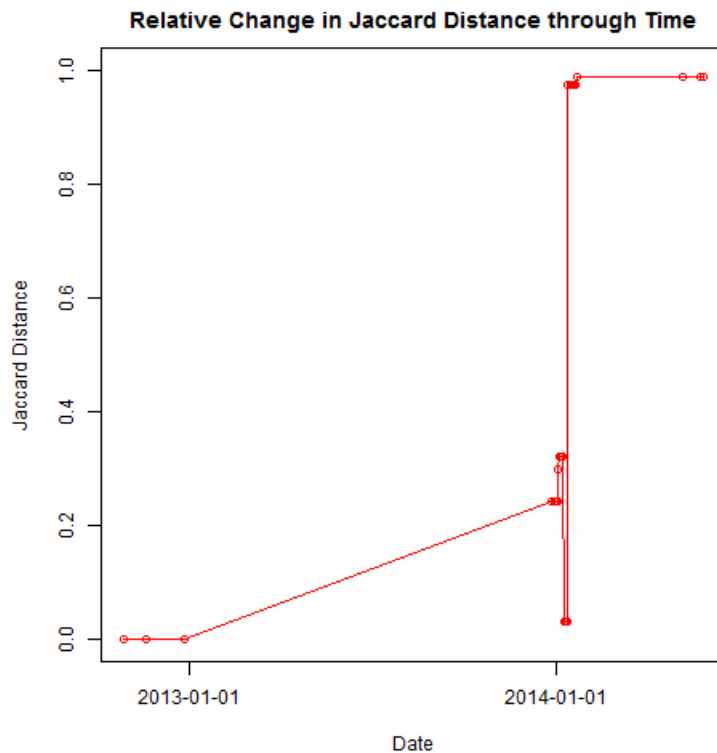


Figure 5: Jaccard Distance relative to the first memento URI# 1

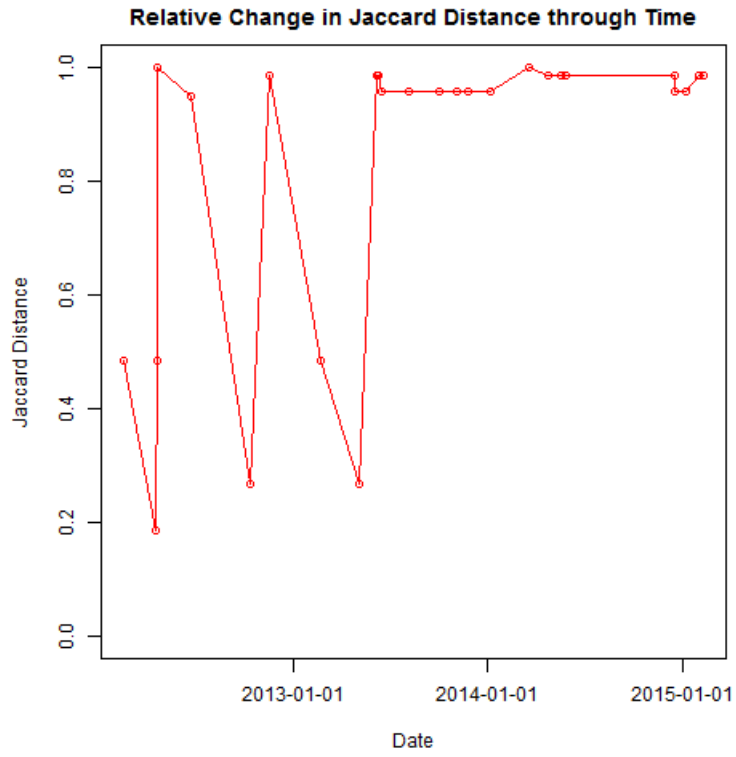


Figure 6: Jaccard Distance relative to the first memento URI# 2

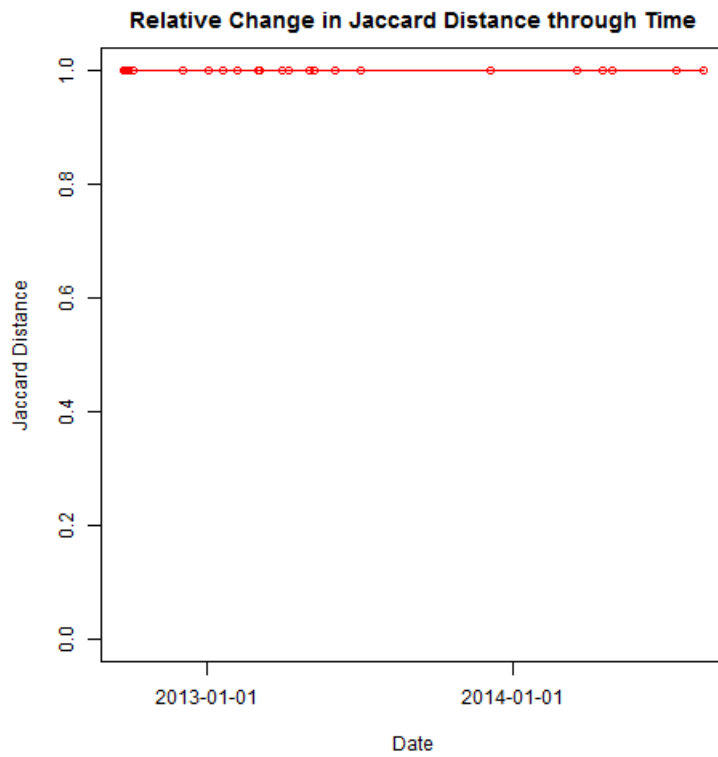


Figure 7: Jaccard Distance relative to the first memento URI# 3

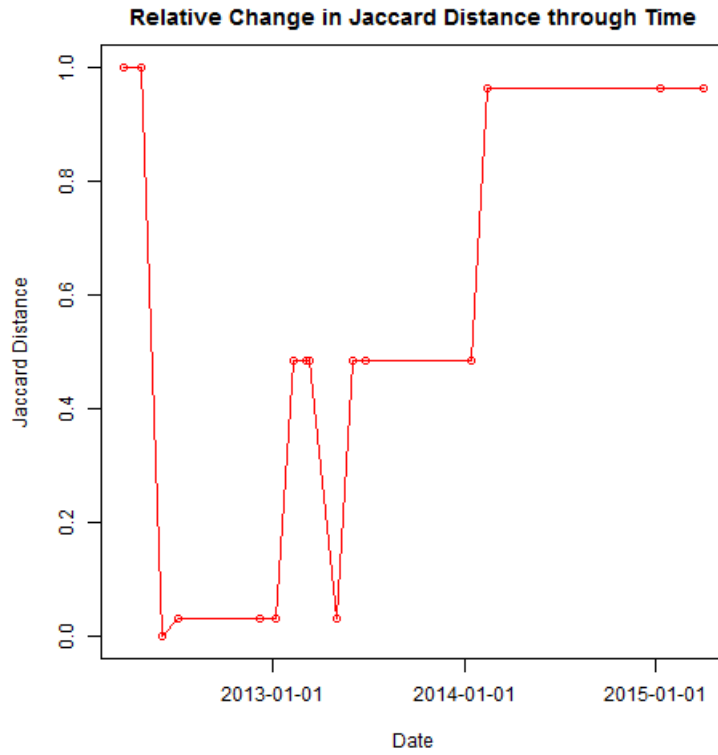


Figure 8: Jaccard Distance relative to the first memento URI# 4

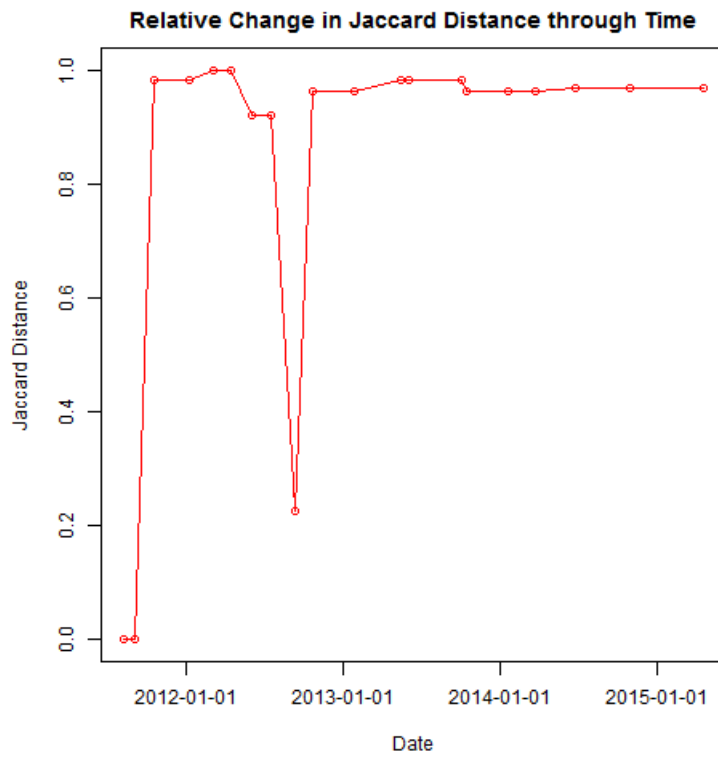


Figure 9: Jaccard Distance relative to the first memento URI# 5

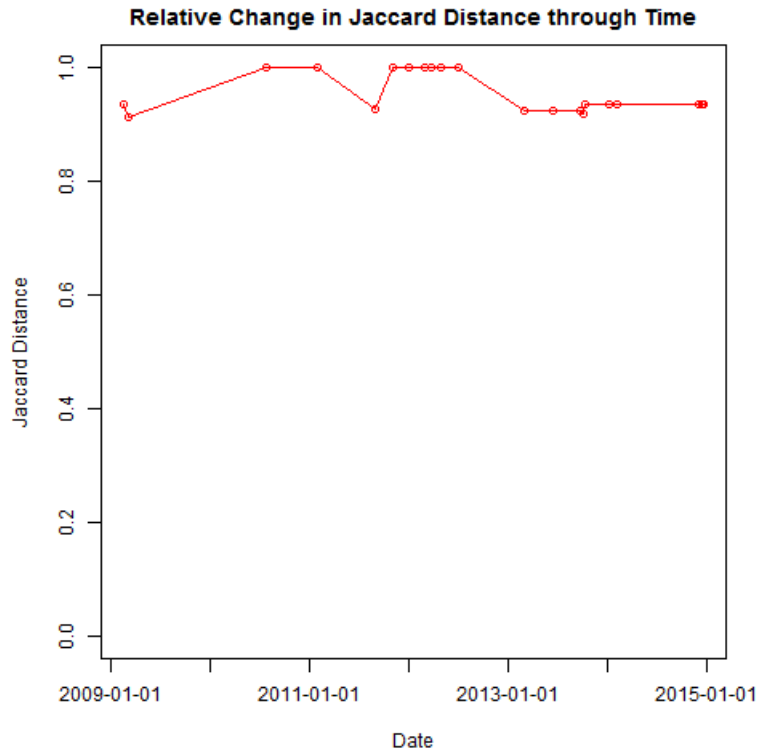


Figure 10: Jaccard Distance relative to the first memento URI# 5

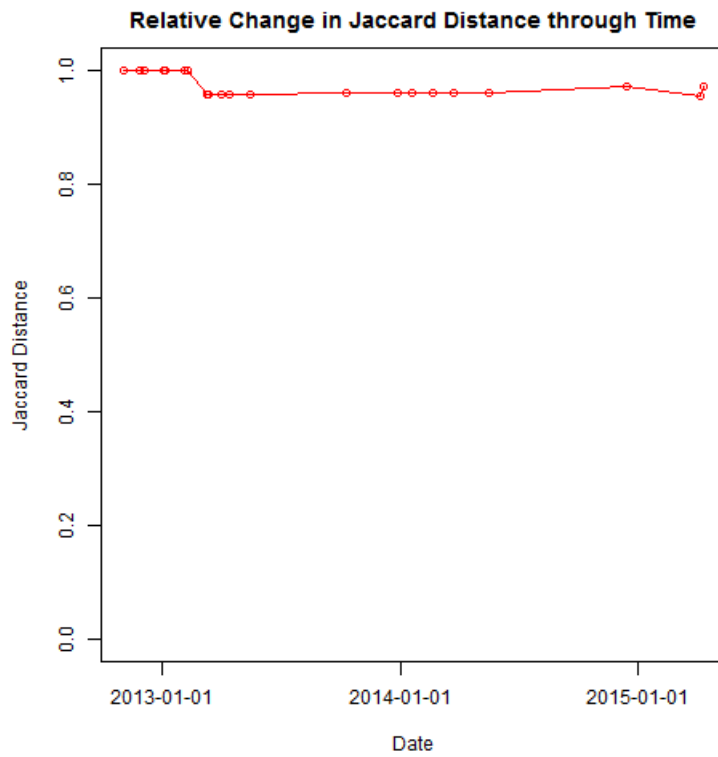


Figure 11: Jaccard Distance relative to the first memento URI# 7

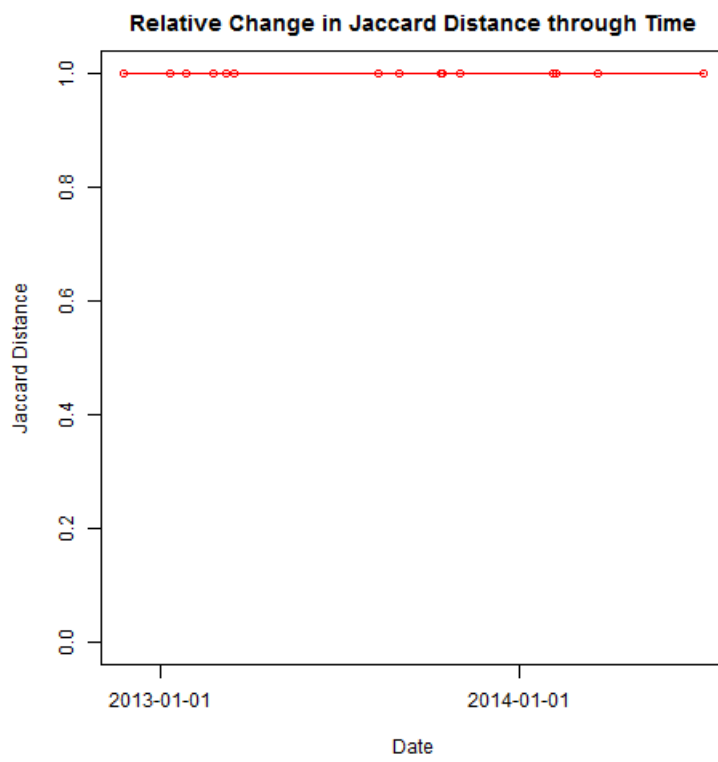


Figure 12: Jaccard Distance relative to the first memento URI# 8

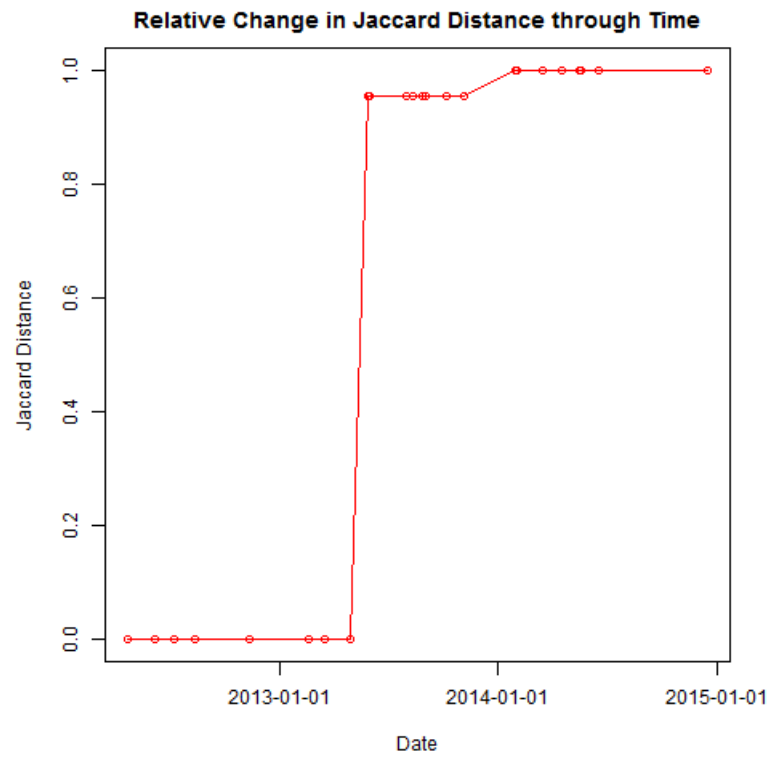


Figure 13: Jaccard Distance relative to the first memento URI# 9

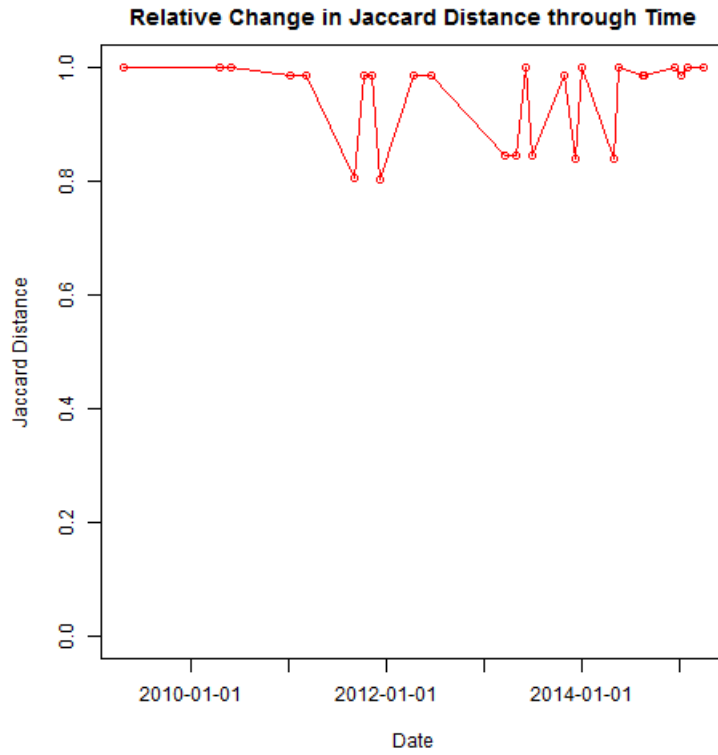


Figure 14: Jaccard Distance relative to the first memento URI# 10

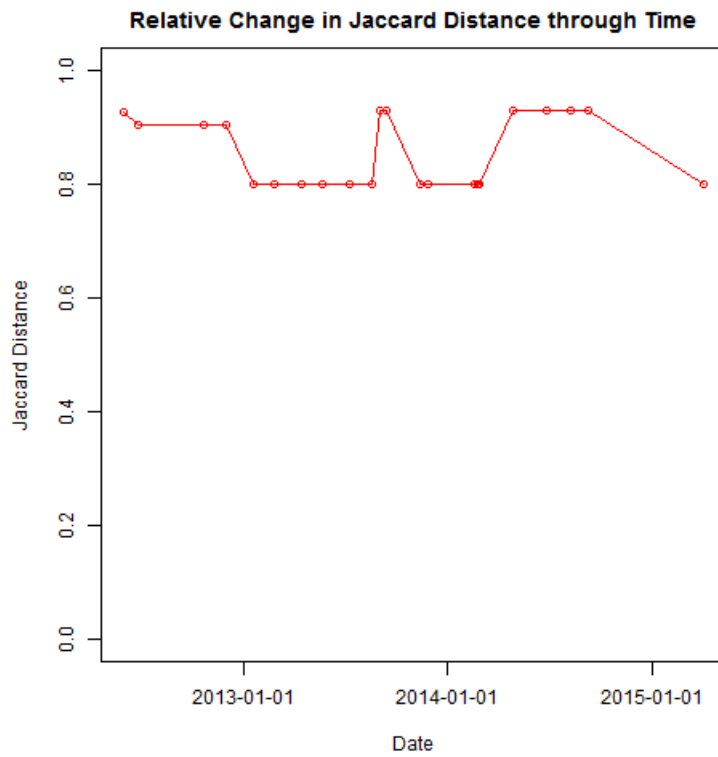


Figure 15: Jaccard Distance relative to the first memento URI# 11

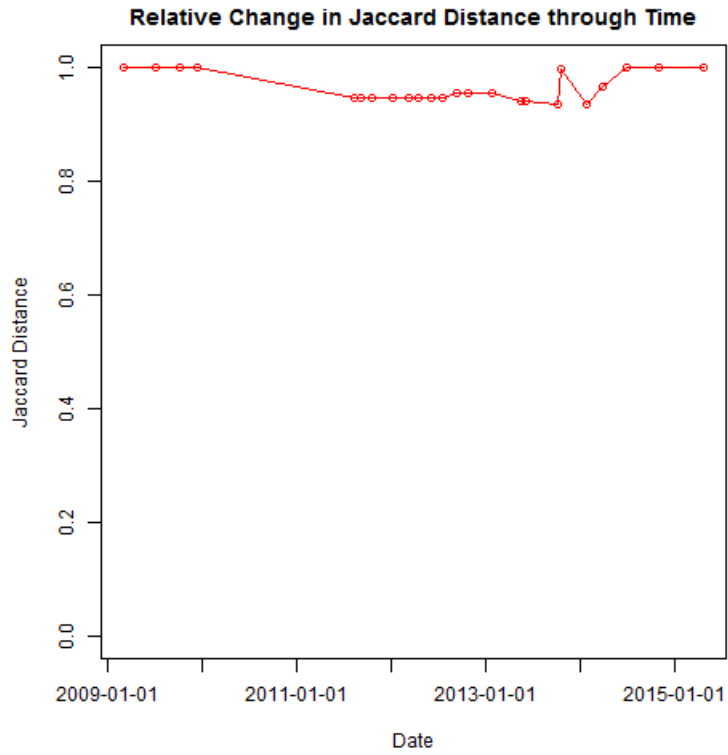


Figure 16: Jaccard Distance relative to the first memento URI# 12

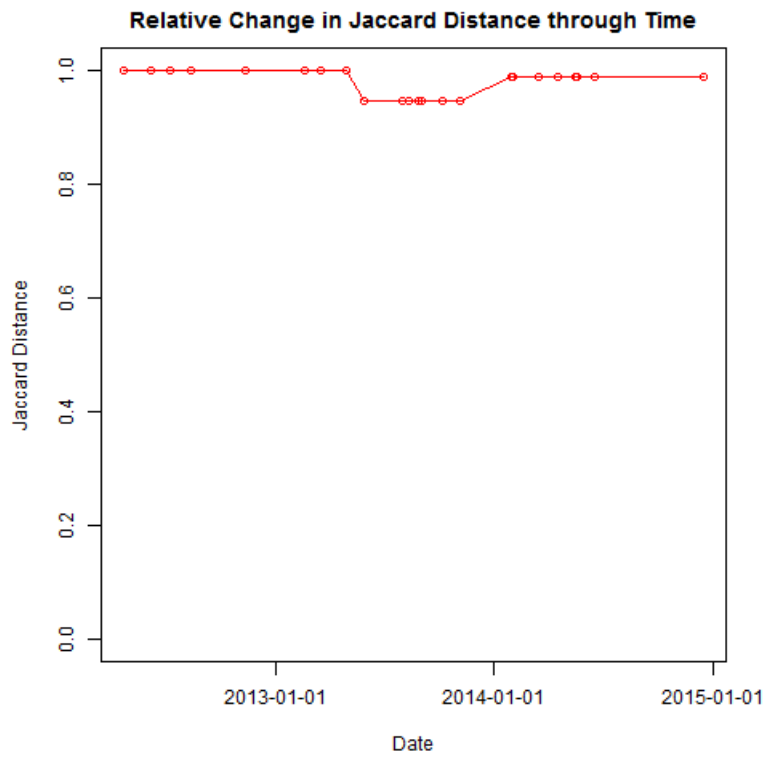


Figure 17: Jaccard Distance relative to the first memento URI# 13

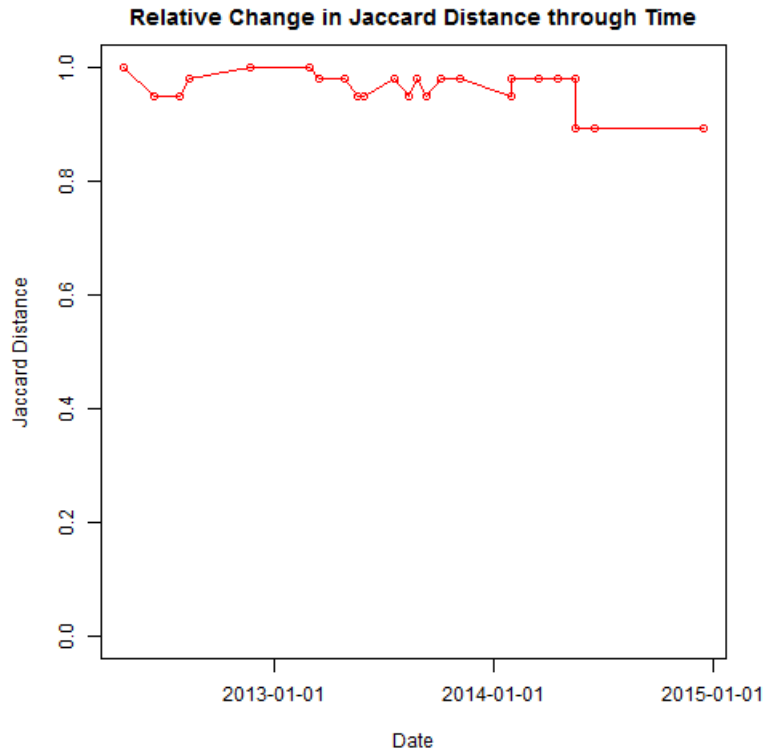


Figure 18: Jaccard Distance relative to the first memento URI# 14

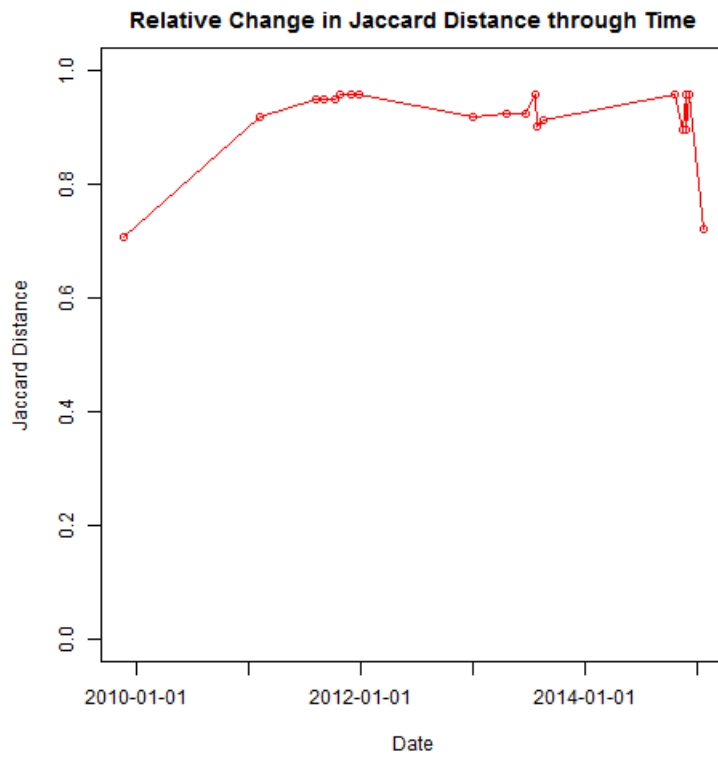


Figure 19: Jaccard Distance relative to the first memento URI# 15

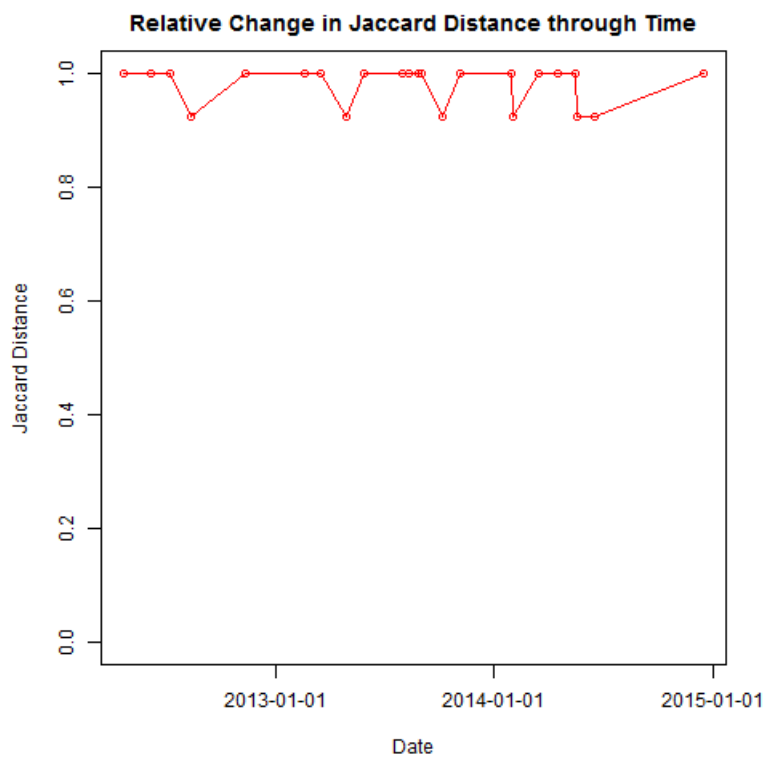


Figure 20: Jaccard Distance relative to the first memento URI# 16

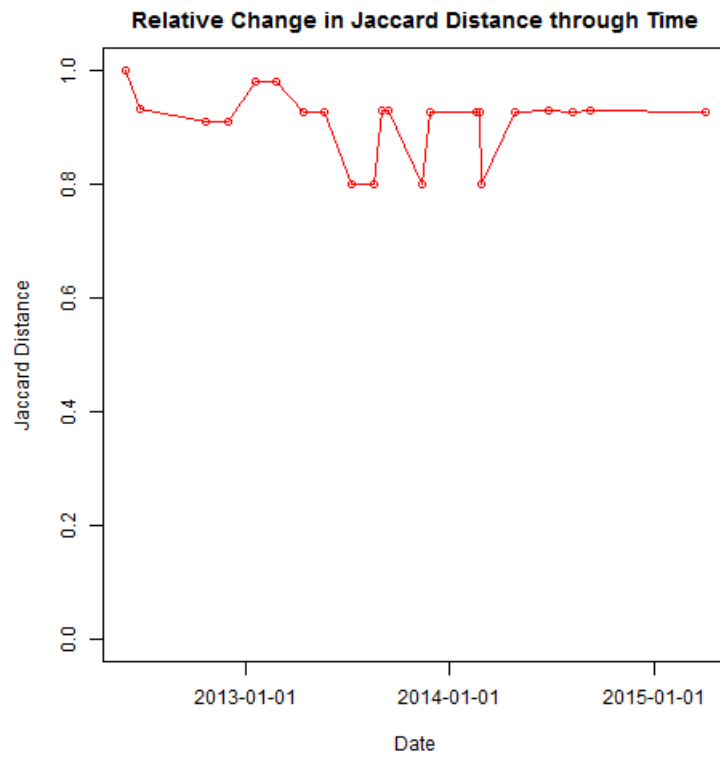


Figure 21: Jaccard Distance relative to the first memento URI# 17

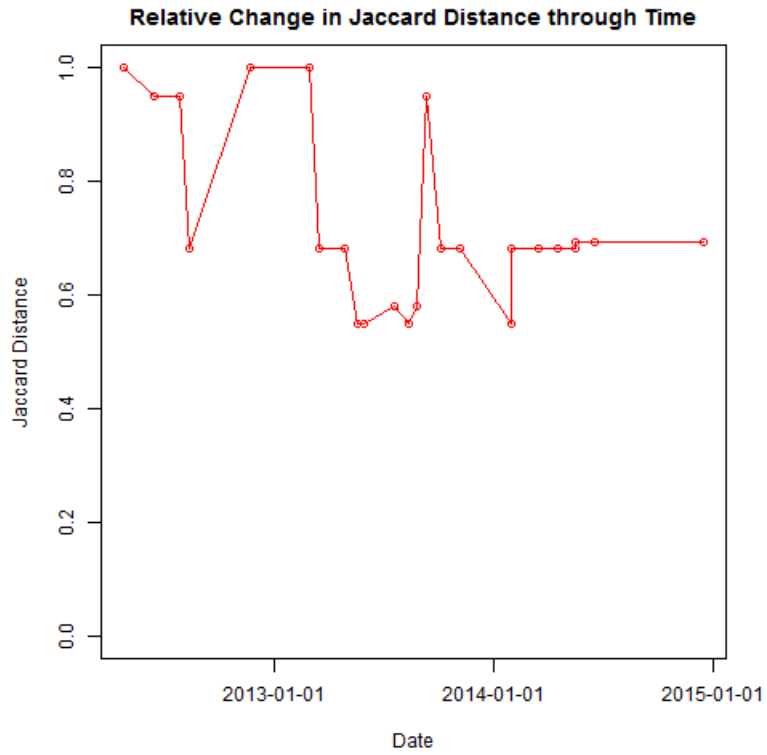


Figure 22: Jaccard Distance relative to the first memento URI# 18

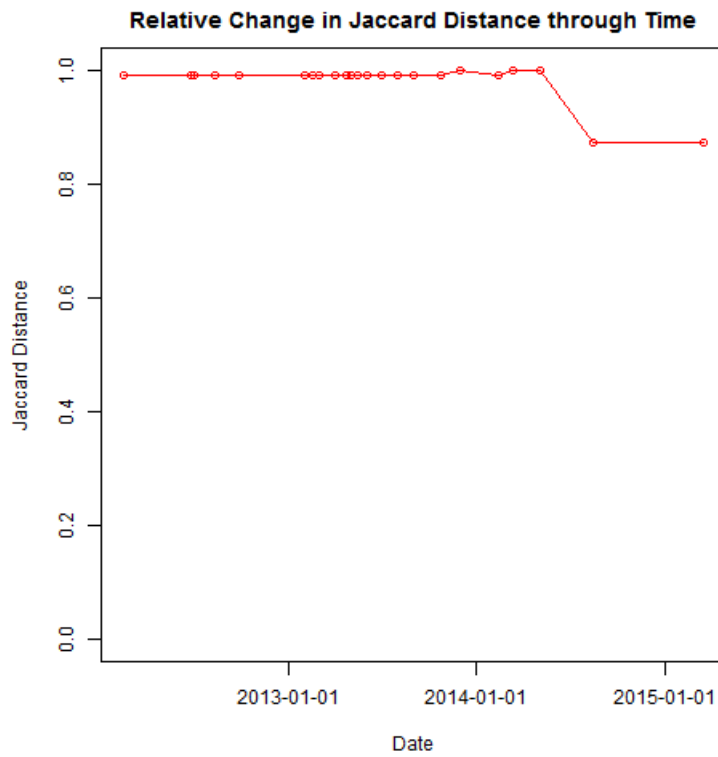


Figure 23: Jaccard Distance relative to the first memento URI# 19

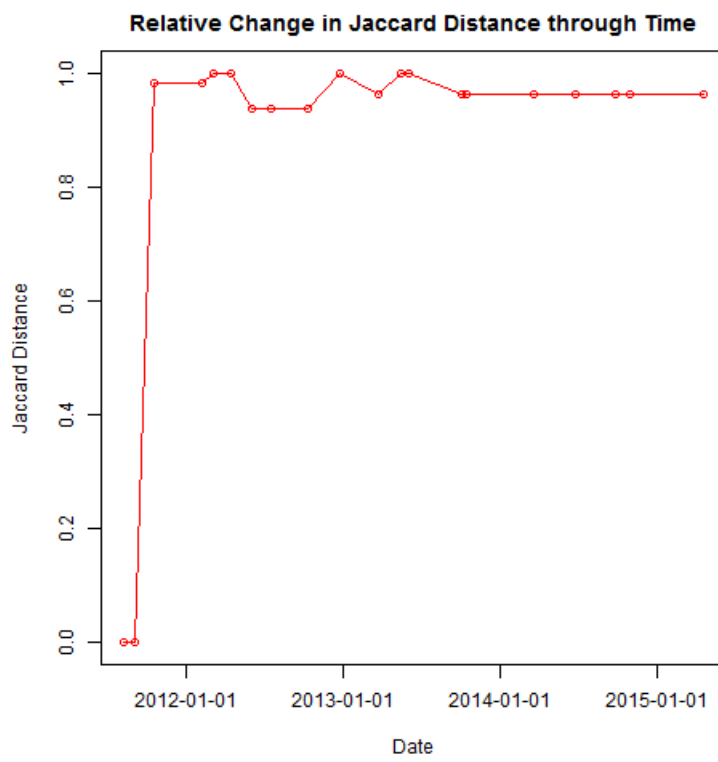


Figure 24: Jaccard Distance relative to the first memento URI# 20

4 Question 4

- Choose a news-related event
- Use twarc.py to collect 1000 tweets, every day for 5 different days
 - See: <https://github.com/edsu/twarc>
- For each day:
 - Create a wall
 - Build a tag/word cloud for each day
 - Create a map using GeoJSON & Github
 - * <https://help.github.com/articles/mapping-geojson-files-on-github/>
- Discuss in detail lessons learned, experiences, etc.

4.1 Solution

Below are screen shots of five day tweets wall and word cloud.

- Installed twarc[4] to retrieve 1000 tweets.
- I used twarc to build tweets wall, word cloud and to retrieve geojson.
- "World Earth day" was my news trending topic and I retrieved 1000 tweets daily.
- I found few interesting things with word cloud and GeoJSON.
- GeoJSON showed exact geographical locations where this topic is being discussed each day.
- Word cloud clubbed all words and displayed the words in different sizes based on the word frequency in the tweets. Each day earth was the word most frequently used.
- I have uploaded GeoJSON files for each of the day to Github. GeoJSON can be visualized at <https://github.com/bharath559/cs851-s15/tree/master/Assignment4/Q4/geojson>

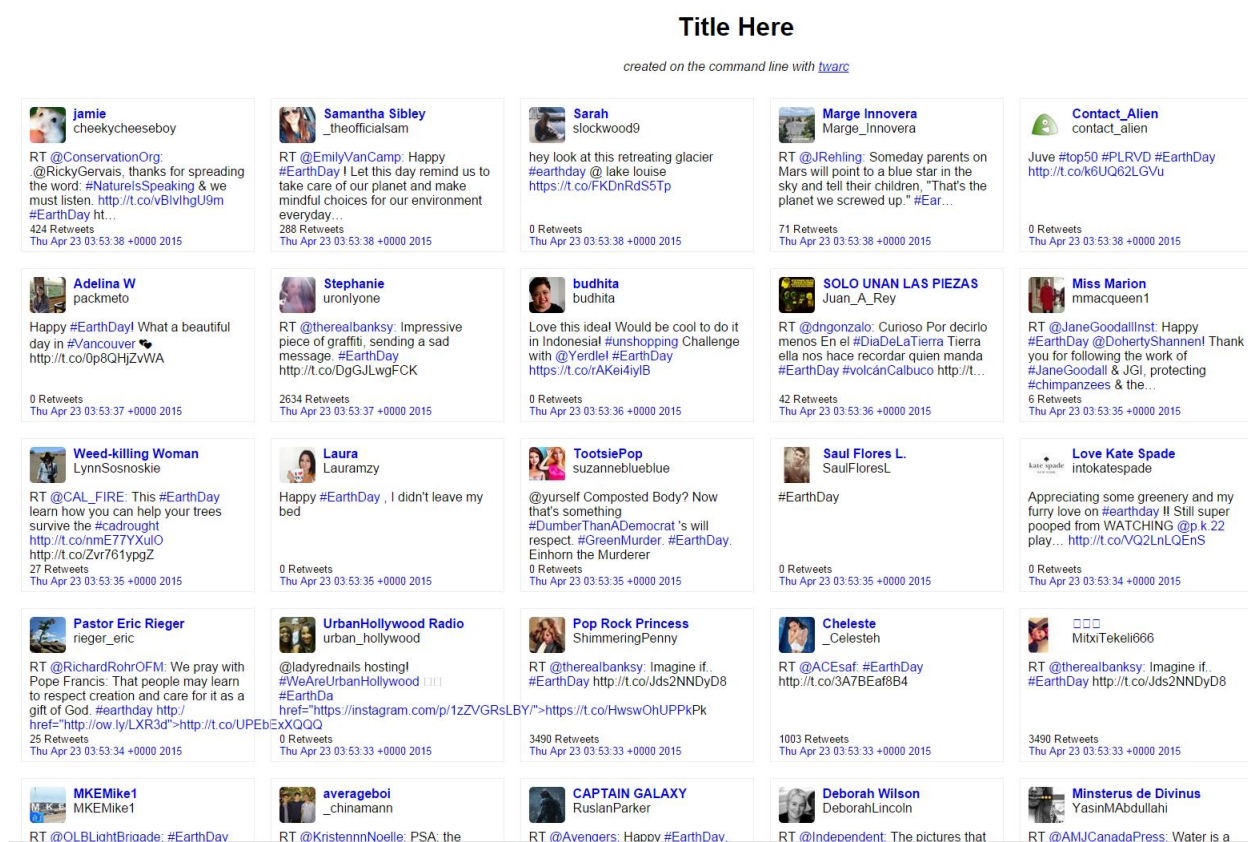


Figure 25: Tweets Wall Day1

Title Here

created on the command line with [tvarc](#)


























 Estelle cartidge Stelc81 RT @loveyogabears: Happy Earth Day to all Earthlings ♥ blessings from @loveyogabears @shaktistudios @bikramliverpool @Stelc81 #earthday ht... 2 Retweets Thu Apr 23 22:45:36 +0000 2015	 Sumicel USA sumicelusa #Samsung patents a lightweight #fitness tracking band (+photos) - http://t.co/g16thFDV0J #News #EarthDay 0 Retweets Thu Apr 23 22:45:34 +0000 2015	 Jessica Orwig orwigwrites RT @mashable: We know #EarthDay can be depressing. So here are 4 reasons to be optimistic about the planet http://t.co/7mMK1SqYEB http://t.co/... 65 Retweets Thu Apr 23 22:45:34 +0000 2015	 Bri BriThibodeaux RT @NigelBarker: It's #EarthDay & the perfect time for us to remember to look after & respect our beautiful planet #Roma #hugatree x http://t.co/... 61 Retweets Thu Apr 23 22:45:32 +0000 2015	 Consulado Ecu Queens CEcuQueens RT @ApoyaAlEcuador: Our voices are stronger together than apart. Help our brothers and sisters in the Amazons #WeDeserveJustice #EarthDay h... 12 Retweets Thu Apr 23 22:45:31 +0000 2015
 Irina Burkova IrinaBurkova RT @Jasmine_Soul: Happy #EarthDay 🌱🌱 http://t.co/OvaWf2QrmE 4 Retweets Thu Apr 23 22:45:26 +0000 2015	 Heather ourcircuitime Happy #EarthDay! We made a bird feeder during the craft segment on today's http://t.co/AvDOQj87BL video activity. http://t.co/opHxQZ38n 0 Retweets Thu Apr 23 22:45:26 +0000 2015	 Hello Enterprise HelloEnterpris RT @GoodsForGoodCo: BETTER LIFE - what-EVER! http://t.co/lcBhTLE2pQ #BetterLife #EcoFriendly #homecare #EarthDay #Online #shopping http://t.co/... 3 Retweets Thu Apr 23 22:45:23 +0000 2015	 bsbafflesbrains bsbafflesbrains RT @elliosch: @firedoglake nails it. Thankfuly @motherjones didn't realize their grave mistake earlier. http://t.co/TKQvW8yWNv http://t.co/... 3 Retweets Thu Apr 23 22:45:23 +0000 2015	 oops_im_amazing DestinyFugate #EarthDay 0 Retweets Thu Apr 23 22:45:21 +0000 2015
 Brian David Stockman Carlisleboy RT @WAGNERGIRLE: #APicturesWorthOneThousandWords Obama's empty #earthday speech 🌱 http://t.co/nlNk5oXZUsUs #tcot #pcra #ccot http://t.co/... 10 Retweets Thu Apr 23 22:45:21 +0000 2015	 Sean Nelson opinionatedwino #earthday http://t.co/stfprs3jyB 0 Retweets Thu Apr 23 22:45:20 +0000 2015	 r.e.e.n.a reenahuffman13 RT @JoeyGraceffa: HAPPY #EarthDay in the spirit I've been playing @TreeStorygame try it out its so cute! #HaveFunDoGood http://t.co/FjGdzY0j2y 373 Retweets Thu Apr 23 22:45:18 +0000 2015	 Lachlan Milt MiltPozjbk RT @4BillLewis: The most expensive #Starbucks drinks ever http://t.co/q8KUXVsp9S via @nypost #socialmedia #reteeet #coffee #StarbucksProble... 133 Retweets Thu Apr 23 22:45:18 +0000 2015	 DavidPMiranda™ DavidPMiranda Happy Earth Day! So excited I got up before Dawn. Fortunately Dawn sleeps until noon. #EarthDay 0 Retweets Thu Apr 23 22:45:17 +0000 2015
 Abbey Farkas AbFark RT @radicalbytes: This #EarthDay remember personal recycling & water conservation initiatives are designed to shift blame from corporations... 237 Retweets Thu Apr 23 22:45:15 +0000 2015	 sal levett RT @WAGNERGIRLE: #APicturesWorthOneThousandWords Obama's empty #earthday speech 🌱 http://t.co/nlNk5oXZUsUs #tcot #pcra #ccot http://t.co/... 10 Retweets Thu Apr 23 22:45:15 +0000 2015	 Jonathan Cohee joncohee RT @PaigeToccoTerra: Kudos to my daughter Emmalyn, and all who dedicate their careers to preserving wildlife habitats #EarthDay 2 Retweets Thu Apr 23 22:45:15 +0000 2015	 papadukeschaney KevinDChaneySr RT @mashable: We know #EarthDay can be depressing. So here are 4 reasons to be optimistic about the planet http://t.co/7mMK1SqYEB http://t.co/... 65 Retweets Thu Apr 23 22:45:15 +0000 2015	 Jaime Silva Jaime0 RT @FortuneMagazine: See Apple's glossy #EarthDay video http://t.co/1eXXRKjPcl http://t.co/aCmunNe6Yh 16 Retweets Thu Apr 23 22:45:14 +0000 2015
 Yordano YordanoNapa RT @NASA: Ah. #Earth. there's 	 SOLFARMERS SOLFARMERS Thursdav dinner idea 4 #EarthDay by 	 Abbey Farkas AbFark RT @radicalbytes: This #EarthDav 	 Keith Tharp KeithTharp Our Earth Dav haul at Hodason 	 Giovanna Quina Klapp gioquina #chile #volcanes esta es 1 fumarola

Figure 26: Tweets Wall Day2

Title Here

created on the command line with [twarc](#)


























 GreeneaseApp GreeneaseApp Our fabulous intern, @sevincsey, blending in some smoothies at GWU's #EarthDay http://t.co/0w17ULrReB 0 Retweets Fri Apr 24 13:40:21 +0000 2015	 Little Critterz LittleCritterz3 #Sloths are an integral part of tropical rain forest ecosystems #EarthDay #nature #environment #Earthdayeveryday 0 Retweets Fri Apr 24 13:40:21 +0000 2015	 Anjana AnjanaKandhan RT @Avengers: Happy #EarthDay, from Earth's Mightiest! http://t.co/IRBeo0kbT 1879 Retweets Fri Apr 24 13:40:21 +0000 2015	 CJ Ronneberg braintwat RT @deborahhallsted: As if on cue, snow strikes America as Pres. Obama talks global warming for #EarthDay [photos] http://t.co/bK4HYA4HL5 1 Retweets Fri Apr 24 13:40:16 +0000 2015	 Connected Universe ConnectdUnivrse Nature is a portal to deeper understanding ~ Malcom Carter, director of #ConnectedUniverse #EarthDay http://t.co/W5YNI7i9of 0 Retweets Fri Apr 24 13:40:15 +0000 2015
 OEM Off-Highway OEMOffHighway It was a busy social media week for us! Check out @intematparis coverage with #OEMatINTERMAT and #EarthDay coverage with #GreenOEM 0 Retweets Fri Apr 24 13:40:14 +0000 2015	 sevenbowie sevenbowie RT @ConservationOrg: @RickyGervais, thanks for spreading the word: #NatureIsSpeaking & we must listen. http://t.co/vB1vHgtU9m #EarthDay ht... 732 Retweets Fri Apr 24 13:40:12 +0000 2015	 Danny Ivatt Daniellvatt According to #earthday crap where I am, I need to eat less meat and be more vegan. Neither of which is conducive to being swole. #nothanks 0 Retweets Fri Apr 24 13:40:10 +0000 2015	 Ashley Ashleyash27 RT @Martha_x15: @LisaVanderpump >Important reminder on #EarthDay. <a href="http://t.co/CiObnwtkm2<...loll">http://t.co/CiObnwtkm2<...loll 10 Retweets Fri Apr 24 13:40:09 +0000 2015	 Chris cmdiehl29 RT @VictoriaJustice: How did you celebrate #EarthDay? Join me & @DoSomething by recycling clothes at @hmusal http://t.co/dFL8Ob0bAv 382 Retweets Fri Apr 24 13:40:07 +0000 2015
 gidgiegoomie gidgiegoomie RT @Martha_x15: @LisaVanderpump >Important reminder on #EarthDay. <a href="http://t.co/CiObnwtkm2<...loll">http://t.co/CiObnwtkm2<...loll 10 Retweets Fri Apr 24 13:40:04 +0000 2015	 Dolores Briseno mofodolo RT @Martha_x15: @LisaVanderpump >Important reminder on #EarthDay. <a href="http://t.co/CiObnwtkm2<...loll">http://t.co/CiObnwtkm2<...loll 10 Retweets Fri Apr 24 13:40:04 +0000 2015	 giveaway 100K coins coins_100k RT @SlashedBeauty: It's #EarthDay! Discover #ecofriendly #skincare brand, @derma_e & win their best products: http://t.co/tkQ1TyvA3J http://t.co/6rHY8MnjAA 19 Retweets Fri Apr 24 13:40:04 +0000 2015	 PinkGreen Creative WeArePinkGreen #FeatureFontFriday is here, with #earthday as the theme. @miradax goes green with Grafolia. #typography #design #... http://t.co/6rHY8MnjAA 0 Retweets Fri Apr 24 13:39:57 +0000 2015	 Agility CSR AgilityCSR Good #EarthDay memories indeed. https://t.co/YNFdUpE1y3 0 Retweets Fri Apr 24 13:39:55 +0000 2015
 Team Lupold Thoman LupoldThoman Green Team!!! Thank you for all your hard work! Wonderful video about recycling! @NH_Elementary #earthday http://t.co/3obgMhVwPY 0 Retweets Fri Apr 24 13:39:51 +0000 2015	 Ashwini Kaikini ashkaikini RT @Nriyandhi: The earth & I have a lot in common. We're made mostly of water & every year we both get hotter! ~@TheEllenShow @ashkai... 1 Retweets Fri Apr 24 13:39:48 +0000 2015	 Abbi Abi_Frann RT @Martha_x15: @LisaVanderpump >Important reminder on #EarthDay. <a href="http://t.co/CiObnwtkm2<...loll">http://t.co/CiObnwtkm2<...loll 10 Retweets Fri Apr 24 13:39:46 +0000 2015	 1victoria bailey gabjamom1 RT @Martha_x15: @LisaVanderpump >Important reminder on #EarthDay. <a href="http://t.co/CiObnwtkm2<...loll">http://t.co/CiObnwtkm2<...loll 10 Retweets Fri Apr 24 13:39:46 +0000 2015	 gareth lee-bell gingerkermi RT @LuxeTiffany: Happy #EarthDay https://t.co/Wmectr8IVb (Victoria Falls, Zambia) 3 Retweets Fri Apr 24 13:39:41 +0000 2015
 Mauri Solages MauriSolages RT @ColleenNGRID: Gearing up! 0 Retweets Fri Apr 24 13:39:41 +0000 2015	 Rita Rin2379 RT @findmeaifit: #Competition 0 Retweets Fri Apr 24 13:39:41 +0000 2015	 EZstorit.com ezstorit RT @PublicStorage: Inspired by 0 Retweets Fri Apr 24 13:39:41 +0000 2015	 Claire Basilaña ClaireBasilona Calling on Young People to Lead Us 0 Retweets Fri Apr 24 13:39:41 +0000 2015	 Burke County Georgia BurkeCountyGA Master Gardeners Plant Sale today at 0 Retweets Fri Apr 24 13:39:41 +0000 2015

Figure 27: Tweets Wall Day3

Title Here

created on the command line with [twarc](#)


























 aaron k. snidesnark RT @EricWolfson: Best. #ClimateChange. Tweet. EVER. (via @ScottWesterfeld) #ClimateAction #EarthDay #EarthDay2015 #EarthDayEveryDay http://... 1101 Retweets Sat Apr 25 15:14:48 +0000 2015	 Boulder Book Store boulderbooks RT @techtionary: Today #Boulder @alfalfasmarket B'Earthday #EarthDay Bash 11-4 https://t.co/yTlmVgHJm w/ @AsherBrewing @GrowingGardensB 2 Retweets Sat Apr 25 15:14:44 +0000 2015	 CHANDRA PRAKASH dv891798 RT @BrightHouseNow: April 25, Earthday Birthday, Central Florida Fairgrounds Join @1011wjrr & Bright House Networks! Lineup of bands: http://t.co/8gMoi5PpQ0 2 Retweets Sat Apr 25 15:14:44 +0000 2015	 Nita Pujiastuti nee_tai5 RT @tupperwareID: Sebelum mulai menanam #mangrove, jangan lupa berfoto bersama :) #EarthDay http://t.co/8HW81UZAHa 3 Retweets Sat Apr 25 15:14:39 +0000 2015	 CHANDRA PRAKASH dv891798 RT @BrightHouseNow: TODAY- Earthday Birthday, Central Florida Fairgrounds Join @1011wjrr & Bright House Networks! Lineup of bands: http://t.co/8gMoi5PpQ0 1 Retweets Sat Apr 25 15:14:30 +0000 2015
 पंडित जी Pandit ji dheerajrazdan @AgentSaffron #Nepal #NepalEarthquake God have mercy on the Nepalese population #IndiaWithNepal #EarthDay #earthquake 1 Retweets Sat Apr 25 15:14:29 +0000 2015	 Nita Pujiastuti nee_tai5 RT @tupperwareID: Ayo kita tanam 10.000 #mangrove sebagai bentuk cinta kita pada lingkungan bersama WWF & Tupperware! #EarthDay http://t.co/http://t.co/YwMD39TbwD 3 Retweets Sat Apr 25 15:14:29 +0000 2015	 haltonmastergardener HaltonMGs RT @henhere: All set up at Georgetown Market Place for #EarthDay Fair, till 4pm. @ImagineHH http://t.co/1DcdT4er05 1 Retweets Sat Apr 25 15:14:14 +0000 2015	 TheEcoVeganEvents TomasinaBarr RT @missrainbird: This #EarthDay, watch Cowspracy online for just \$1 and learn the truth about who's killing the planet! #documentary http://t.co/Dp28uvT96U 1 Retweets Sat Apr 25 15:14:08 +0000 2015	 Robert Rossachaj RRossachaj RT @the_sca: SCA CEO Jaime Matyas greets 400+ volunteers at today's #EarthDay service project w/ @AmericanExpress in #DC. http://t.co/0yhnkn... 2 Retweets Sat Apr 25 15:13:58 +0000 2015
 Oak Hills Sports OakHillsSports RT @kenmeibers: Giving back to the community. #earthday http://t.co/LpwwQ74Gyx 3 Retweets Sat Apr 25 15:13:57 +0000 2015	 Ahmed Numen AhmedNumen Sad to hear the terrible news of massive earthquake hit in Nepal. My prayers and thoughts are with the people of Nepal #EarthDay #Earthquake 0 Retweets Sat Apr 25 15:13:53 +0000 2015	 Primped and Improper primpdNimproper Pace Team Cleanup! #earthday #oakville #joshuacreek #realtor #realtorlife #realestate #century21... https://t.co/dg4De0l2Ox 0 Retweets Sat Apr 25 15:13:42 +0000 2015	 money honey g_money_honey29 RT @Martha_x15: @LisaVanderpump -Important reminder on #EarthDay. http://t.co/CiObnwtkm2<...lol 58 Retweets Sat Apr 25 15:13:40 +0000 2015	 Crystal Coast CrystalCoast_NC The #CrystalCoast #EarthDay Celebration is going on @ #FortMacon. Enjoy FREE kids activities, demos & more, 10am-2pm. http://t.co/J9vXiOnKsC 0 Retweets Sat Apr 25 15:13:38 +0000 2015
 Yenny T UNSWstudentleif RT @ThePatriot143: CLIMATE ACTIVISTS IN MELTDOWN AS 'SCEPTIC' GETS AUSTRALIAN GOVERNMENT FUNDING http://t.co/mv4xisaWuG #EarthDay 1246 Retweets Sat Apr 25 15:13:36 +0000 2015	 Estuk T_F_L RT @reason: News flash: #NorthKorea wins #EarthDay for record 45th time in a row! http://t.co/JALvCO3WT HT @Neoavataara http://t.co/MUynUG0... 1246 Retweets Sat Apr 25 15:13:24 +0000 2015	 Angelo Waglio_O RT @AlexineAudiard: #EarthDay http://t.co/gmoF0DireT 36 Retweets Sat Apr 25 15:13:24 +0000 2015	 Goran Kolasinac onikapova RT @MariaSharapova: Everyday is #EarthDay #HugATree http://t.co/SS2svT2AmA 424 Retweets Sat Apr 25 15:13:23 +0000 2015	 Ferguson-Florisant FergFlorSchools Great crowd @ our Little Creek Nature Area celebrating #EarthDay big thanks to @missioncontinue for building seating http://t.co/hPUtDsyAGJ 0 Retweets Sat Apr 25 15:13:20 +0000 2015
 Kim Rees kimba420 RT @WeeLittleHawk: @Stantec 	 li'l gal with aGlock liilglockgal Al Gore fact #EarthDay 	 mercedes.m LuisaM56_ RT @AlexineAudiard: #EarthDay 	 Warsame Abdi warsameeabdi RT @Stanford: Stanford's new 	 Barrett Henry Team bhenryteam Pace Team Cleanup! #oakville

Figure 28: Tweets Wall Day4

Title Here

created on the command line with [twarc](#)

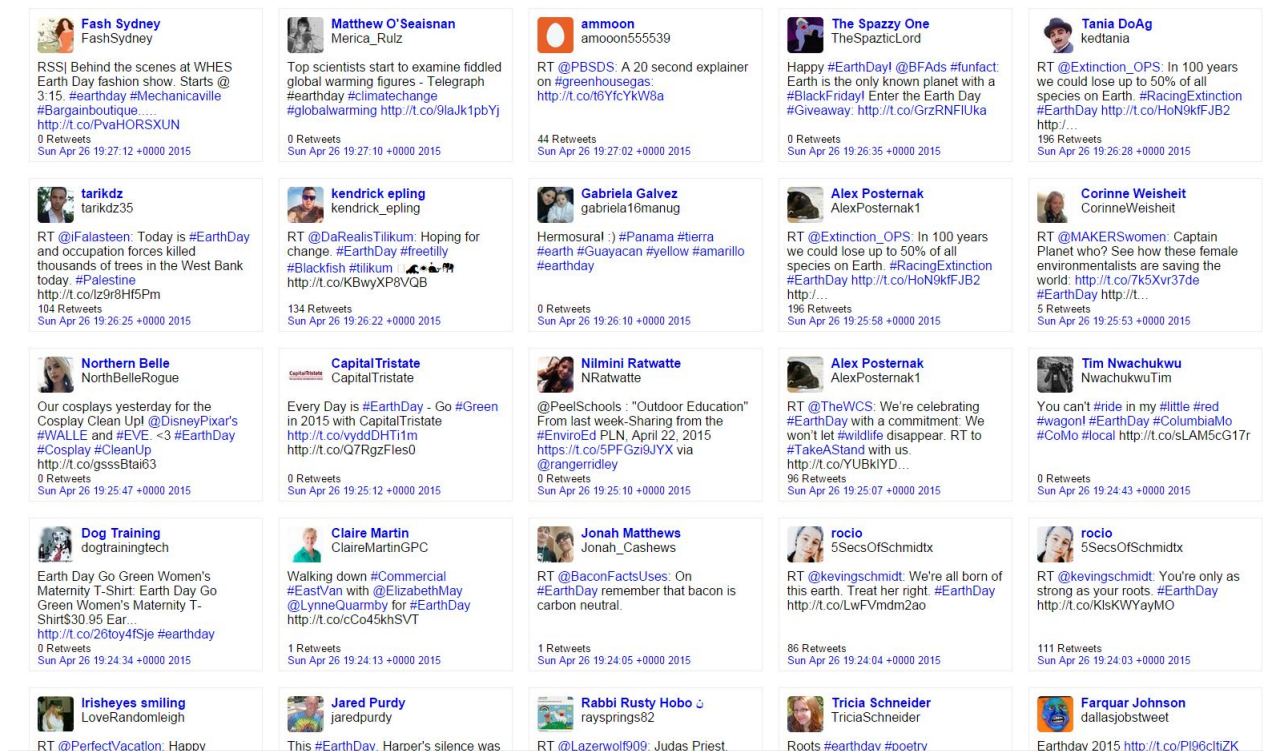


Figure 29: Tweets Wall Day5

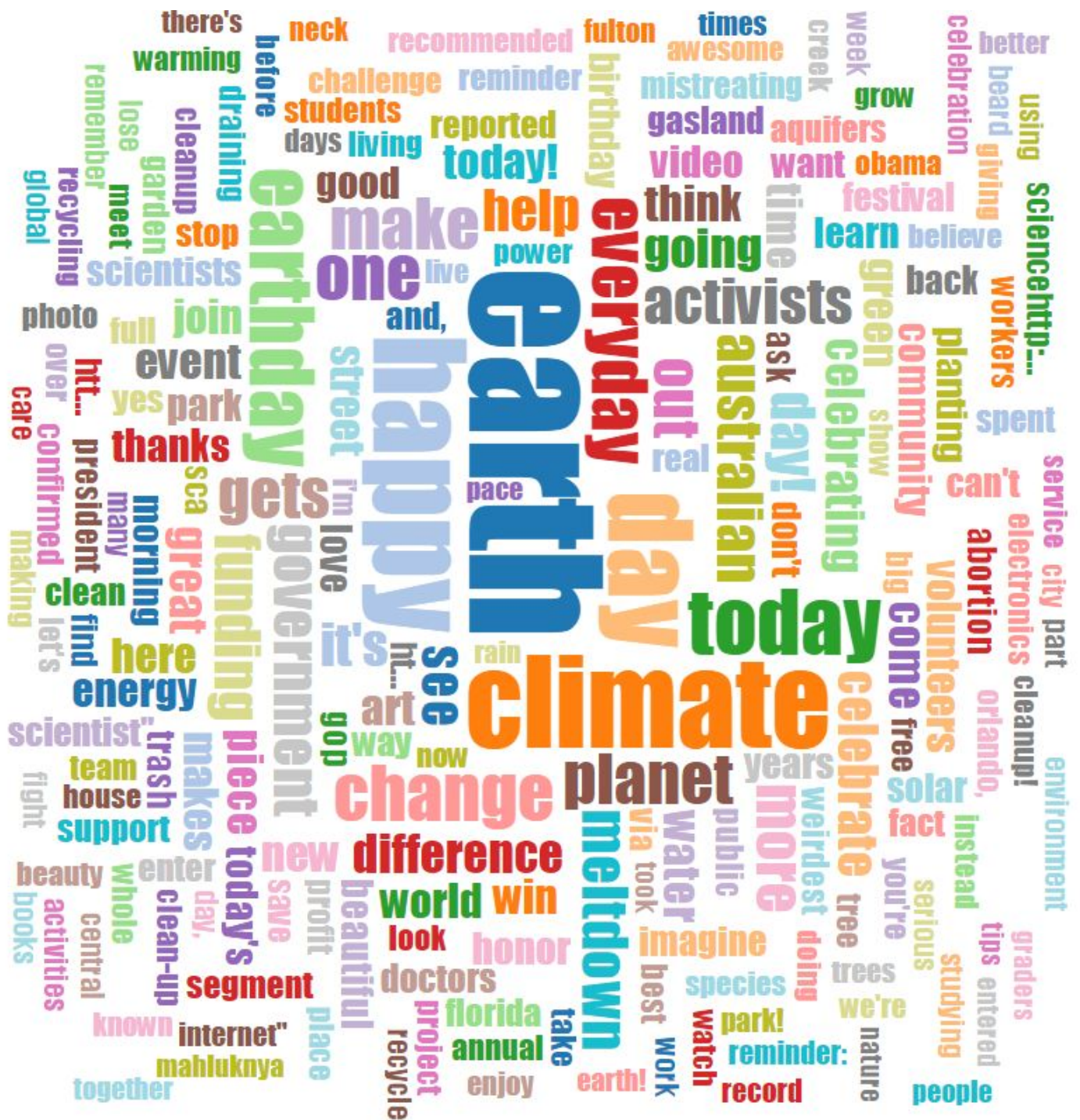


Figure 33: Tweets Word Cloud Day4

Bibliography

- [1] Cumulative. <http://www.r-bloggers.com>.
- [2] jusText. <https://pypi.python.org/pypi/jusText/2.1.0>.
- [3] timelapse. <https://github.com/burg/timelapse/>.
- [4] twarc. <https://github.com/edsu/twarc>.
- [5] Wordcount. <https://github.com/mlafeldt/google-python-class/blob/master/basic/solution/wordcount.py>.
Accessed: 2015-04-03.