

# **SOCIAL AND BEHAVIORAL NETWORKS**

## **Final Project**

Cecilia Martinez Oliva

## **Abstract**

*The purpose of this work is to analyze tweets supporting or not the 4th december italian referendum. In the first part the aim is to find groups of terms capturing the collective attention of the two different parties. In the second part, instead, the target is to search for the more influent users for each party, trying to identify them also thanks to the previously found politicians and the words they used. Finally it analyse the spread of influence of the two parties.*

## Contents

<b>Introduction</b>	<b>3</b>
<b>1 Temporal Analysis</b>	<b>3</b>
1.1 . . . . .	3
1.2 . . . . .	4
1.3 . . . . .	7
1.4 . . . . .	13
1.5 . . . . .	24
<b>2 Identify mentions of candidates or YES/NO supporter</b>	<b>24</b>
2.1 . . . . .	24
2.2 . . . . .	25
2.3 . . . . .	27
<b>3 Spread of Influence</b>	<b>32</b>
3.1 . . . . .	33
3.2 . . . . .	33

## Introduction

Starting from 18850000 tweets and 2287781 users in total collected from 26/11/2016 to 06/12/2016.

First of all I created an index for all the collected tweets.

- *date* is a LongField and contains the date in the timestamp format.
- *id* is also a LongField and contains the twitter id of the user.
- *screenname* is a StringField, user's Twitter screen name.
- *name* is a TextField, the user's name. Special characters are removed.
- *followers* is a LongField, follower's number of the user.
- *text* is a TextField, the text of the tweet. In this case there is a stronger analyzer and there is the possibility to use stemming or not.
- *hashtags* is a TextField, it's applied lower case and tokenization.
- *mentions* is a TextField, mentioned users in the tweet. Only tokenization.

I used the SimpleAnalyzer in the name field in order to remove special characters. In the text field we can choose to use the StopAnalyzer (with italian stopwords) or to use also the stemmer; in the latter case the analyzer is the ItalianAnalyzer. For hashtags and mentions is used the WhitespaceAnalyzer and in the first case is applied also the lower case.

However finally for this work I decided to use the ItalianAnalyzer for the tweet's text.

Then I also created another index, from the first one, for all the collected users.

- *id* is a LongField.
- *screenname* is a StringField.
- *name* is a TextField, again using the SimpleAnalyzer.
- *followers* is a LongField.

Since screen names are unique, starting from all the screen names saved in the previous index, I query each screen name and get only the first result to get the informations needed.

## 1 Temporal Analysis

### 1.1

I manually searched on the web for politicians and journalists. To obtain bigger lists of yes or no supporters, I automatically searched for their screen name and their support.

Using the users index, I query the name and I suppose that the real screen name is associated to the one with more followers.

I then searched for the more indicative hashtags as "bastaunsi", "iovotesi", "iodicono", "ioivotono", to classify the users. If they use more yes-hashtags are classified as yes-users, if they use more no-hashtags they are classified as no-users, otherwise they are not classified and are not considered in the lists.

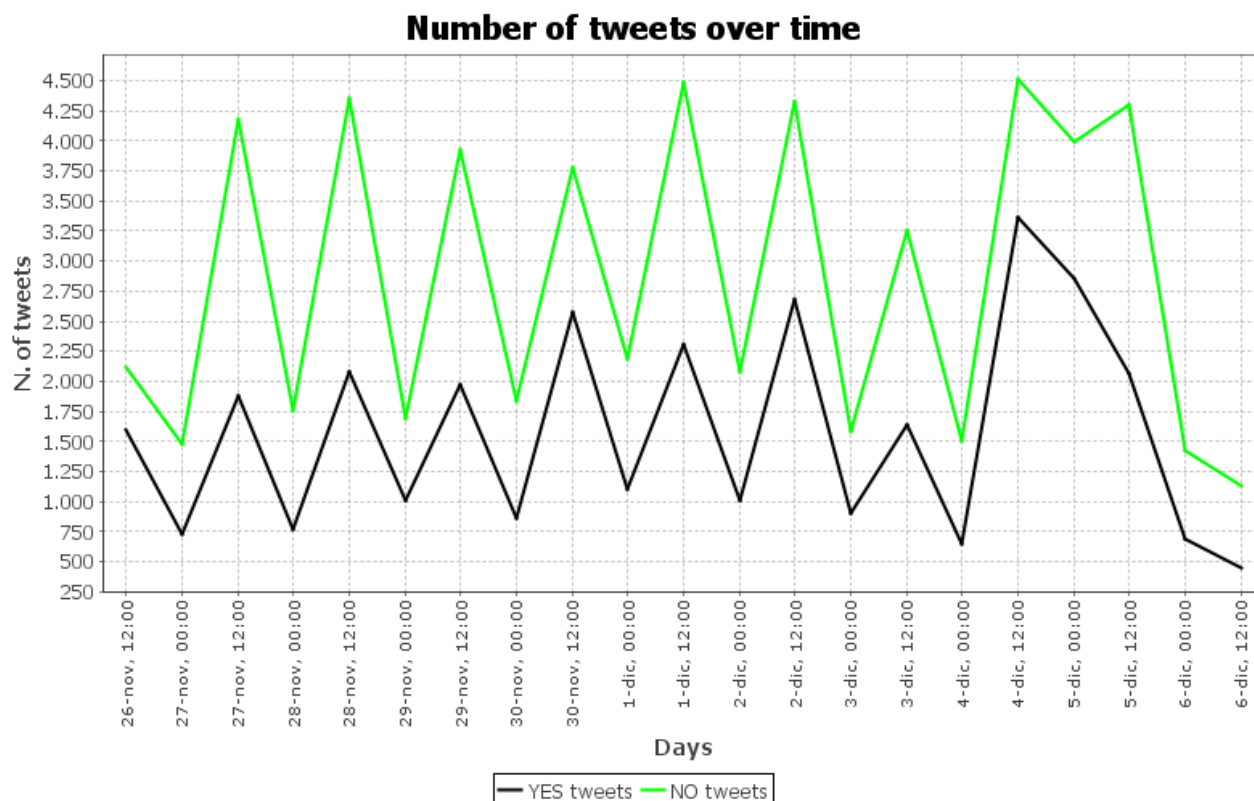
Those lists are saved in two files: yes-p.txt and no-p.txt. The easier way to have a fast access to the tweets we are interested to is to create two new indeces: TwitterIndexYES, TwitterIndexNO. In Table 1 it's shown the real number of people and tweets found for each index.

	YES	NO
n.users	497	605
n.tweets	33200	59887

**Table 1:** YES/NO supporters

The number of tweets over time is represented in Figure 1. As we can imagine, the number of tweets

decreases from 00:00:00 to 12:00:00.



**Figure 1:** Number of tweet over time

## 1.2

We now want to collect the top 1000 words by frequencies that expose the typical pattern that capture the collective attention representative for YES and NO subset of users.

What I am doing is to first collect all terms in the index (excluding terms not containing any alphabet character and shorter than 3 letters), ordered by frequency, and then, using a range filter on dates, I do a query for each term to build SAX. In this way we can save the first 1000 terms matching the pattern of collective attention (one or two picks).

In this case SAX was built with 24h grain and two alphabet characters. Once the term is chosen, I rebuild SAX with grain 12h and numbers instead of alphabet characters.

I then wrote a kmeans algorithm to cluster terms. In order to find terms where SAX behavior is very similar I chose  $k=20$  and 10000 randomizations (plots for the clusters are in the folder).

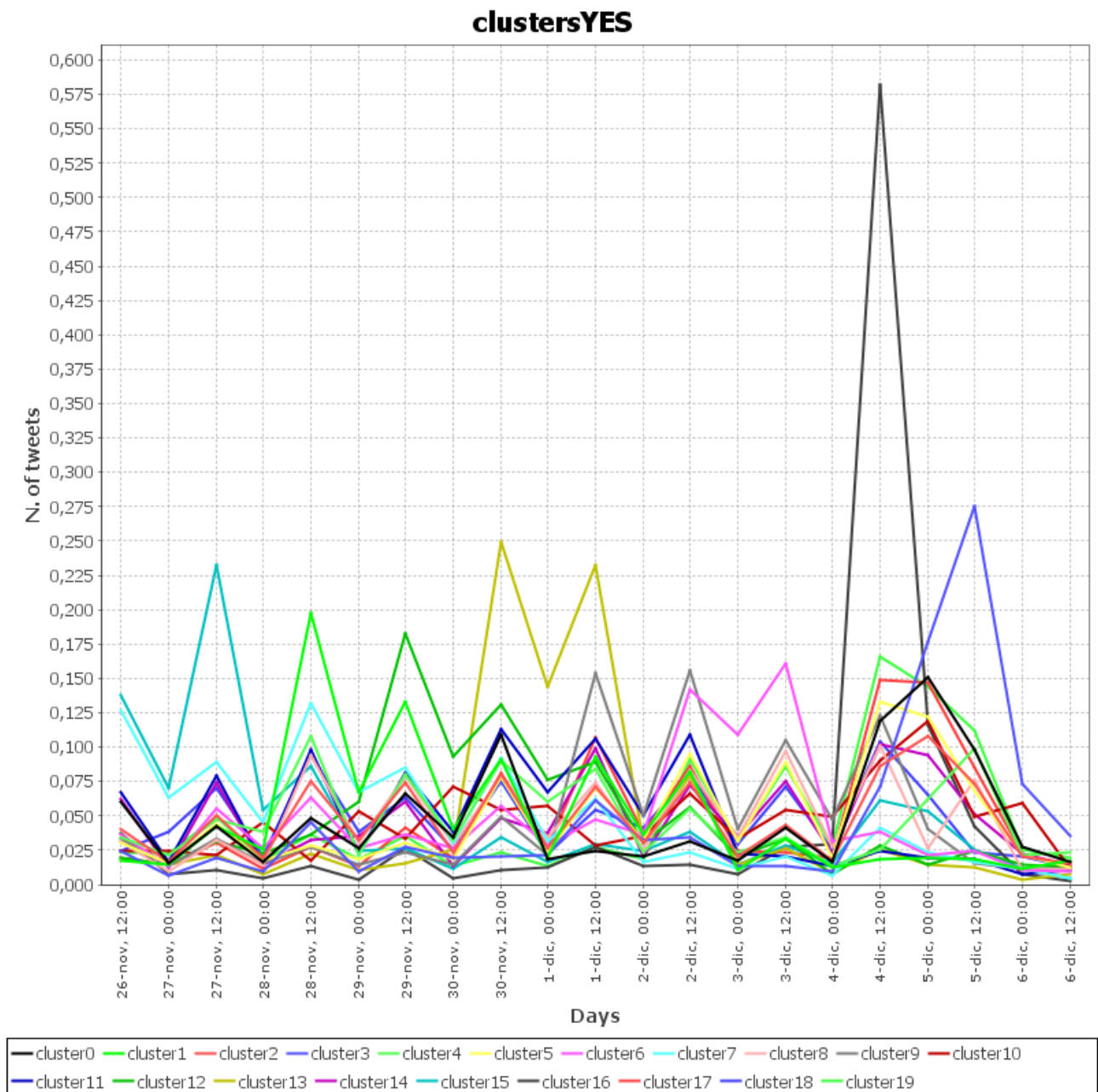
I also saved 100 tweets for each term to have a check of what was going on.

If we choose  $k$  too high, we can find more similar temporal behaviors for the terms but we will not be able to find many co-occurrences in the same cluster.

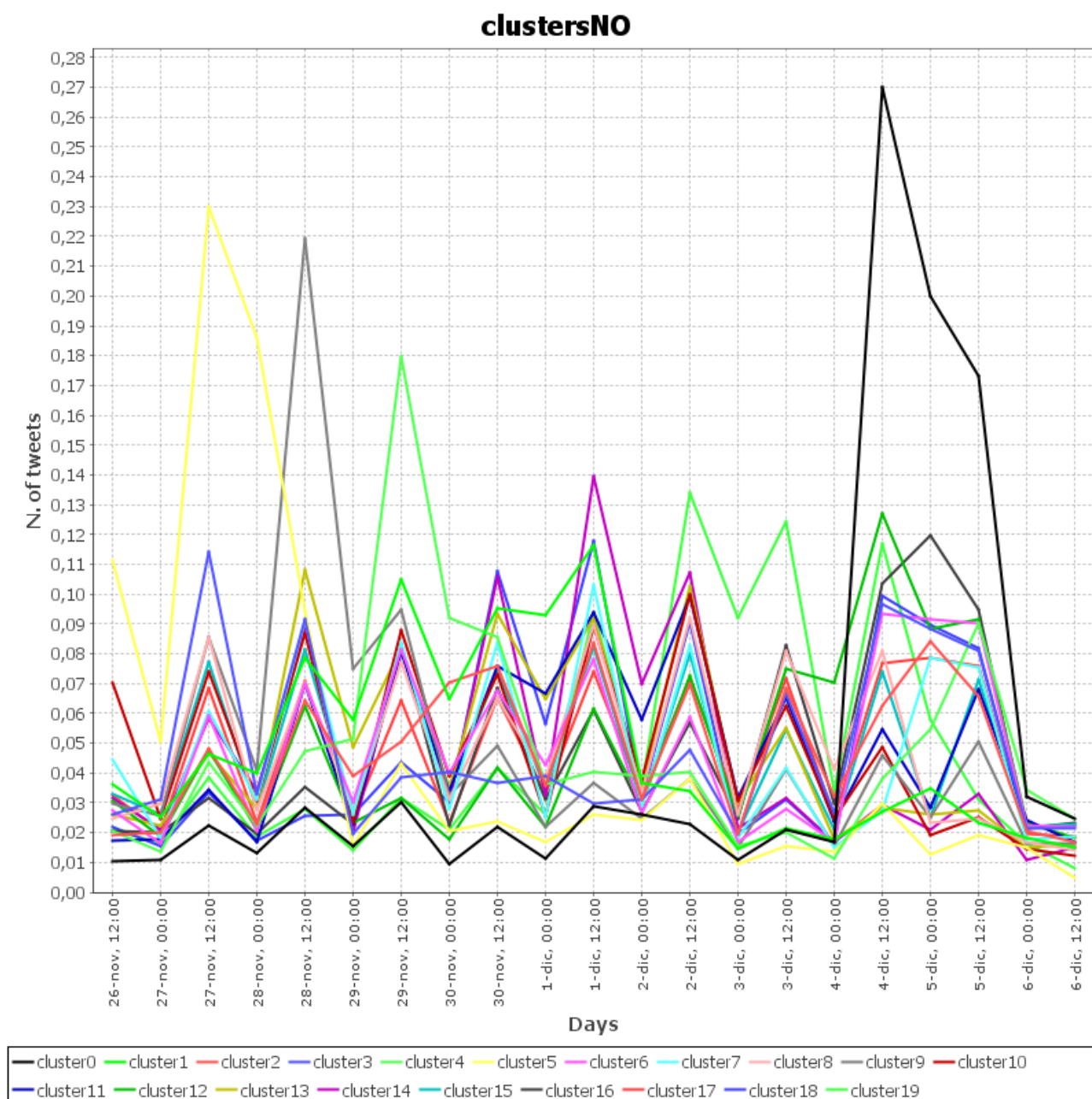
**Example:** TemporalAnalysis/resuts2/New/ClustersYES/cluster-0.txt

```
centroid:          a a a a a a b a b a a a a a a b b b a a
average euclidean distance: 1.5317794667299742
orgogl:           a a a a b a b a b a a a a a b a b b b a a
corrido:          a a a a a a a a b a a a a a b a b b b a a
instant:          b b b a b a a a b a a a a a b a b b b a a
uambrosol:        b a a a a a a a b a a a a a a a b b b a a
qualc:            a a a a a a a a b a a a a a a a b b b a a
saper:            a a b a a a b a b a a a a b b b b b a a
segretar:         a a a a a a b a b a b a a a a a b b b b a
putin:            a a a a a a b b b a a a a a a a b b b b a
posso:            a a a a a b b a b a a a b a a a b b a a a
profond:          b a b a a b b a b a a a a a a a a b b b a a
subit:            a a a a a a b a b a a a a a a a b b b b b
speranz:          a a b b a a a a b a b a a a a a b b b b a
sappiam:          a a a b a a a a b a a a a a a a b b b a a
perder:           b a b a a a a a a a a a a a a a a b b b a a
palle:            a a a a a a b a b a a a a a a a b b b a a
onest:            a a b a b b b a b a a a a a a a a b b a a
rivoluzione:      b a b a b b b a b a a a a a a a a b b b a a
class:            b a b a b a a a b a a a a a a a a b b b b a
pensav:           b a a a a a a a b a b a a b a a b b b b a
fortunatolin:     a a a a a a b a b a a a a a b a b b b a a
detto:            b a a a a a b a b a a a b a a a b b b a a
bosc:             b a a a a a b a b a a a a a b a b b b a a
giorn:            a a b a b a a b b a b a a a a a b b b b a a
elettor:          a a b a a a a a b a a a b a a a b b b a a
sbagl:            b a b a a b a a b a a b a a b a b b b b b
ottim:            b a a a a a b a b b a a a a a a a b b b a a
traged:           a a a a a a b b b a a a b a a a b b b a a
apre:             b a a a a a a a b b a a a a a a a a b b b a b
fredd:            a a a a b b b a b a a a a a a a b b b a a
vice:             a a b a a a b a b a a a a a a a b b a b a a
perso:            a a a a a a a a b a a a a a a a b b b a a
colvieux:         a a a a a a a a b a a a a a a a a b b b a a
andreamarcucc:    b a b a b a b b b a a a a a a a a a b b b a a
speriam:          a b a a a a b a b a a a a a b a b b b a a
elezion:          a a b a a a b a b a a a a a a b a b b b a a
andar:            a a b a b a a a b a a a a a a a a b b b b a a
attenzion:        b b a a a a a a b b a a a a a a a a b a b a a
antgentil:        b a a a a a a a b a a a a a b a a a a b a a a
obama:            a a b a a a b a a a a a a a a a a b b b b a
cadut:            b a a a a b a a b a b a a a a a a b b b b a
aria:             a a a a a a b a b a a b a b a a a b b b a a
primar:           a a b a b a a a b a a a a a a a a b b b a a
averl:            b a b a a a b b b a a a a a a a a a b b b b a
spiegat:          b a b a a b a b b a a b a a a a b a b a a
previst:          b a a a b a b a b a a b a a a a a b b b b a a
fedeargel:        a a a a a a b a b a a a a a a a a b b b b b a
padri:            b a a a a a a b b a a a a b b a b b b a a
fate:             a a a a a a b b b a a a b a a a b b b a a
solit:            a a a a a a b a b a a a b a a a b b b a a
ripartir:         a a a a a a b b b a a a a a a a a a b b b a
```

For each term of a cluster it has been computed the number of tweets containing that term over time, divided by the total frequency of that term (a normalized time series for the term). It is then computed the sum of those values, divided by the number of terms in the cluster (a normalized time series for the cluster) (see Figure 2 and 3).



**Figure 2:** YES clusters



**Figure 3:** NO clusters

### 1.3

For each cluster we want to find terms used in the same context. I built a graph where terms are the nodes. The number of co-occurrences of two terms cannot overcome the frequency of the less used term. So, as a threshold to decide whether to put an edge between two terms, I used the minimum frequency of the two terms times 0.5. This means that the less used term appears more than 50% of times near to the other term.

I then searched for all the connected components of the graph. In this way I am excluding terms which do not frequently appear in the same tweets with any other term of the same cluster. Now we want to find groups of terms homogeneously connected. If the connected component is composed by only two terms, we can be sure of it, otherwise a good way to ensure it can be to find the innermost core.

**Results:** [TemporalAnalysis/resuts3/New/YesWords.txt](#)



	cluster	connected components	innermost core
0	[orgogl, corrido, intant, uambrosol, qualc, saper, segretar, putin, posso, profond, subit, speranz, sappiam, perder, palle, onest, rivoluzion, class, pensav, fortunatolin, detto, bosc, giorn, elettor, sbagl, ottim, traged, apre, fredd, vice, perso, colvieux, andreamarcucc, speriam, elezion, andar, attenzion, antgentil, obama, cadut, aria, primar, averl, spiegat, previst, fedearangel, padri, fate, solit, ripartir]		
1	[teresabellanov, elegger, la_dant, vittim, simil, investment, raiun, evoluzion, mondial, risponder, student, progett, report, thanks, appell, tg1referendum, false, montecitor, ragionidels, info, aula, andreaorlandosp, pension, honest, gabanell, indagin, tema, principal, trasparenz, cambiat, vino, efficient, cambier, quintacolonn, discuter, leggedibilanc, andreaoliver, palerm, mantov, scuol, deputatipd, mannin, glgallett, paoloigna1, attravers, noist, campidogl, camer, approvat, novit, matteispond, competenz]	gabanell report	gabanell report
2	[rainews, dittatur, paese, maggior, certo, stabilit, spettacol, altriment, indietr, idea, arriv, mondo, mani, nuove, due, mai, aspettiam, vera, iniz, https, peggior, ben, parte, mercat, vita, contezero76, t.co, process, davant, interessant, rispost, lavorar, far, farlo, video, referendum, fatto, elett, infatt, prossim, diventat, diventar, tropp, rifiut, ecco, città, immagin, propost, buon, fassin, tutta, quind, direzion, milan, milion, avant, via]	rainews video maggior referen- dum ecco città https milan milion mercat via t.co interessant	via ecco referen- dum video inter- essant t.co https rainews
3	[tristezz, moltissim, dati, dimess, mps, singol, belli, guidocrosett, industr, ripartiam, iniziati, salvo, dunqu, minoranz, congress, pbersan, viva, strad, manovr, bersan, addio, padoan, italicum, analis, mante, napolitan, borsa, slide, messin, kcgfbwnnrh, lasc, prepar, prest, postreferendum, piazzapulitala7, palazzo_chig, scenar, mattarell, peraltr, eni, dimission, piazzapulit]	viva singol belli slide kcgfbwnnrh  ripartiam iniziati	kcgfbwnnrh slide viva belli singol  iniziati ripartiam
4	[posizion, continuar, luigibobb, alcun, divers, lariachetir, bilanc, fake, voterann, cines, dicembr, sostener, venerd, cgilnazional, dalem, serviz, veder, approvazion, domand, maio, deriv, misur, iron, partir, gesù, wiredital, tagl, ilfoglio.it, corso, mostr, tecnic, incontr, servon, prem, nextquotidian, riman, grupp, costruir, regional, soldii, controll, confront, union, chieder, sera, m5s]		
5	[frase, finir, person, trovat, ancor, fascist, dovrebb, spero, propr, vinca, populist, passion, quali, alfonsofuggett, tackledur, ripet, prend, libri, credo, banca, red_mask9, bene, soli, vatican.it, sole, figl, provat, sentit, cancellar, beh, spess, stell, ritorn, chius, verament, cambiam, fogl, governar, casa, esatt, voto, ital, pena, eugeniocard, graz, discussion, amo, benvenut, popol, premier, difes, cmq, bellissim, vincer, vuoi, voce, appunt, brexit, linea, elettorale, sotto, concord, sostenut, aspettar, the, vittoriozuccon, visto, fare]		
6	[livorn, censis, scriver, messa, intervist, silenz, sbonaccin, alliev, quotidian, pacat, rompe, maron, stazion, legend, votos, stefanoesposit, robertapinott, forchiell, final, dfhbk4cddk, claud, calabr, chiud, annunziat, emergenz, proseguir, firenz, affrontar, serviziocivil, piena, attacc, signor, sogno, entusiasms, procur, unità, indecis, decisiv, quint, federmeccanic, silenzioelettoral]	silenz silenzioelet- toral proseguir rompe	silenzioelettoral proseguir rompe silenz
3	[tristezz, moltissim, dati, dimess, mps, singol, belli, guidocrosett, industr, ripartiam, iniziati, salvo, dunqu, minoranz, congress, pbersan, viva, strad, manovr, bersan, addio, padoan, italicum, analis, mante, napolitan, borsa, slide, messin, kcgfbwnnrh, lasc, prepar, prest, postreferendum, piazzapulitala7, palazzo_chig, scenar, mattarell, peraltr, eni, dimission, piazzapulit]	viva singol belli slide kcgfbwnnrh  ripartiam iniziati	kcgfbwnnrh slide viva belli singol  iniziati ripartiam
7	[corruzion, morte, fidelcastr, volantinagg, potenz, oro, luigi, pura, raitr, nuti, sindac, metalmeccanic, bastaunsindac, nazional, contratt, devon, plcastag- nett, indagat, cuba, settor, profil, consult, bicameralism, incapac, iscritt, fotograf, danielecin, cagliar, qualit, riccard, firmefals, basilicat, violenz, raggi, frances, novembr, madia, virgin, german]	raggi virgin  metalmeccanic contratt	virgin raggi  contratt metalmec- canic
8	[fondamental, viagg, messi, tedesc, certa, innovazion, gratuit, prodott, cOnvey, sondagg, uso, scopert, mario, venut, salut, votiam, sandrogoz, angelinascanu, croce, scusa, preved, sembr, web, imprenditor, presentat, news, comunit, ricord, onlin, cultur, brava, bianc, giocator, ufficial, concret, ricevut, special, buono, stefan, massim, canzon, sale]		
9	[sprec, europe, accus, referendar, dover, terremotocentroital, jungle1970, possiam, ogni,	abuso siren	siren abuso

	societ, unitaonlin, livell, primo, sostien, dariofranc, viminal, abuso, siren, sai, cambiar, important, fiduc, futur, vota, holland, han, domenic, torin, campagn, muri, decider, radical, battist, segnal, insiem, forte, urna, vergogn, casapound, basso, scriviam, passo, clicc, chiusur, iovoton]	insiem cambiar forte important futur possiam ogni scriviam passo clicc primo sostien referendar chiusur campagn	passo insiem im- portant primo ogni possiam chiusur campagn referendar
10	[dentr, giornat, casin, intelligent, rassegnat, capir, reale, sentir, szampa56, vecc, tesor, acqua, stamattin, pompe, insopportabil, piu, buongiorno, radio3tweet, contribut, editorial, consequenz]		
11	[riform, balle, paesi, costi, facebook, fnicodem, spaz, abruzz, rispar, clementif, region, chiar, immunit, parlam, energ, meb, bloccat, 4dicembr, lavorator, terremot, soemiad, ugual, senatoripd, quesit, alessiamoran, legget, iovot, patto, pari, direttament, autonom, chiti, liceo, parlamen- tar, twantonio, bambin, sinistraxils, pizza, anno, volet, art, camb, sostegn, rapport, consiglier, giudic, sport, cittadin, accord, bastauns, darionardell, senator, magistrat, intes, blocc, annaascan, ragion, scopr, universit, buro- craz, giuliadivit, sistem]	4dicembr bastauns	bastauns 4dicembr
12	[perdon, giustiz, siti, tarant, chapecoens, lebufaledeln, dimarted, ilva, inchiest, internet, preferisc, dedicat, disinformazion, buzzfeed, ilmessengeroit, dialog, napol, riguard, albertonardell, nomfup, gioved, digital, sette, alepp, motivazion, siria, assemble, turism, bufal, rete, luca, prof, dipendent, propagand, pubblic, cosedilavor, lapo, mozion, comun, orfin, bufaledeln, amiant, buone, testo, troll, lapoelkann, ambient, micheleemilian, comunicazion, pratic]	nomfup cosedila- vor lebufaledeln bufaledeln bufal	cosedilavor nom- fup bufaledeln bufal lebufaledeln
13	[istat, ancon, suveret, statal, roman, ottobr, amica, produzion, siono, specialereferendum, ripart, reportersett, wpc16it, ivanscalfarott, reato, ricostruzion, bruxelles, presidenz, eliminar, session, liberi_s, jxgiulian, paspron, liberis, rrvellin, zanard, costituzionalist, tramont, _geniodelmal, yppi2011, crescit, joyotmzlau, sblocc, vinciam, strateg, contrattop, fondi, uilmnazional, mariannamad, prodi, sanitar, darioballin, lavoltabuon, notiz, credulit, mariaelenabozz1, votand, osservator, disoccupazione, sali, tribunal, matrix, ulivo, istat_it]	sali joyotmzlau disoccupazione ot- to contrattop mariannamad liberi_s liberis roman ulivo prodi session wpc16it jxgiulian yppi2011 mariaelenabozz1	sali joyotmzlau disoccupazione ot- to mariannamad con- trattop liberi_s liberis ulivo prodi roman session wpc16it mariaelenabozz1 ypipi2011 jxgiulian
14	[liber, stefanoceccant, seren, vicin, nord, vai, guida, devi, devo, amore, sapev, schif, flaviotositw, sicurament, stess, succed, passat, vive, oggi, bel, civil, direi, mater, annaleonardi1, general, carissim, opportunit, tipo, terremotat, marted, moment, modell, pieno, sogni, laura, esercit, entrat, sorris, cioè, giunt, mente, dolce, dovut, buonanott, crescer, poche]		
15	[olio, buca, quart, maurizio_lup, violent, bologn, rosberg, manifestazion, matteoricc, dittator, nunziachinnic, attiv, sfide, facolt, riesc, carabinieri, dirigent, temi, corte, donne, superament, fillon, pierbiagio1, nonunadimen, avvent, domenicaliv, pagina_99, inmezzor, paoloall, paritar, super, render, spieg, leragionidels, arma, fidel, argoment, assum, flop, castr, sezion, almen, preferit]	superament paritar arma bologn cara- binier fidel castr maurizio_lup paoloall	paritar superament carabinieri arma bologn castr fidel maurizio_lup
16	[nuvol, a_viscom, livespinoz, austriac, pro_torr, lazio, poll, reggioemil, antonella180262, van, manu57371839, op_supernov, tota_mif, matitecancellabil, segg, dukana2, cento, cancellabil, fbianco91, raffaelepizzat, chimic, sconfitt, copiativ, iohovotatos, libe12andr, saggezz, gomma, pfrascar, occhiocritico12, carta, glover406p, matit, codacons, maratonamentan, dubbi, palazz, alexa5313, der, laziorom, officiasrom, lottiluc, sorpres, labbufal, pelù, affluenz, piero, asocialesemp, pieropelu, hyperbros, emmoverimm, verde, proiezion, you_trend, seggi, exit, chigi, austr, minim, hofer, nazionalist, mariarosariac, aggiung, dimett, mezzanott, sardegn, exitpoll, vanderbellen, sopra, bellen]	you_trend affluenz maratonamentan exitpoll sorpres labbufal emmoverimm fbianco91 raf- faelepizzat pro_torr alexa5313 antonella180262 libe12andr oc- chiocritico12 mariarosariac asocialesemp	you_trend affluenz exitpoll maratona- mentan labbufal sorpres mariarosariac occhiocritico12 raffaelepizzat emmoverimm pro_torr asociale- semp alexa5313 libe12andr fbianco91 an- tonella180262

		cancellabil copiativ matit	copiativ matit can- cellabil
		saggezz segg	saggezz segg
		verde vanderbellen bellen der scon- fitt austr van hofer nazionalist	bellen austr der van
		pelù piero	piero pelù
		palazz chigi	chigi palazz
		poll exit	exit poll
		codacons tota_mif	codacons tota_mif
17	[mdpennalung, candidat, dato, cosa, scie, tanto, ora, allor, qui, salvin, puoi, discors, consigl, vanno, democraz, minut, fino, esce, ettore_rosat, vediam, ferma, dite, crede, francofontana43, contenut, president, dipend, pensat, divis, cambiament, ospedal, ieri, qualcun, claudiovelard, matteosalvinim, dicev, perde, front, alta, compliment, mancan, lavorat, leader, europ, disse, mezzo, battaglia, riuscit, risultat, la7, potrebb, claudioceras, ovviament, annunc, area, post, concors, skytg24, popolar, coerenza]		
18	[viene, serie, barbar, aziend, uomin, econom, verso, date, lotta, appuntament, treno, semplic, qualunqu, normal, rinnov, nuova, siceid, venez, scena, nemmen, sindacat, usa, ridurr, augur, decis, votar, primi, prova, altra, deriuregin, anti, cara, clima, tasse, mica, giul, calc, svilupp, informazion, stefaniapezzop, forma, tempo, francobassanin, alber, ridicol, vista, passa, bellezz, iovotos]		
19	[riott, error, ecc, brunett, salvat, internazional, serio, difficil, magari, vittor, populism, quirinal, trova, vincitor, opposizion, mentan, derby, vinto, punti, stamp, orii151, foto, valeriafedel, votat, sola, trionf, marco, condivid, offes, letta, comunque, pare, pensa, vandab, assur, andrà, guard, bgrisaf, sbagliat, credut, brogl, andiam, colpa, color, dura, sicur, huffpostital, marcomariani_lit, cnel, volev, pensand, certezz, mariocalabres, enric, orgoglios, pacerenat, potrà, ester, lenelepr, urne, esito, christianrocc, cpetruccioli, fedemell]	enric mentan	enric mentan
		pensand votat	pensand votat
		vandab orii151 bgrisaf pacerenat lenelepr	lenelepr pacerenat bgrisaf vandab orii151

## Results: TemporalAnalysis/results3/New/NoWords.txt

	cluster	connected compo- nents	innermost core
0	[renzian, poll, dimess, salvat, dimetters, van, portaaport, vittor, discors, quirinal, segg, ciao, cancellabil, sconfitt, subit, vinto, renxit, iohovotaton, zuccon, renziacas, matit, maratonamentan, cacciar, impres, der, laziorom, havintolademocraz, analis, giornalista, pelù, affluenz, esperienz, piero, fallit, perso, postreferendum, radioanch, verde, popol, risultat, seggi, ultimeno- tiz, definitiv, exit, mattarell, austr, brexit, festa, dimission, laurapuppat, piazzapulit, ermannokilgor]	matit cancellabil	matit cancellabil
		van der austr	austr der van
		poll exit	exit poll
		piero pelù	piero pelù
1	[terrorism, colleg, criminal, vicin, msi, tarant, gasparritour, ilva, semplic, firmat, gratis, divent, morir, mese, carmentpf, film, terremot, arrestat, banca, gioved, rtl1025, turism, agropol, bancaetrur, general, ventaglidiparol, aument, arrest, tutel, giornalism, nomin, schiav, lapo, rifiut, entrar, madre, pil, famigl, vincol, contrattosubit, meravigliosafrittur, agorara, forner, dollar, matrix, tribunale, novembr, german]	agropol merav- igliosafrittur	meravigliosafrittur agropol
2	[blog, propr, populist, ricordar, qualc, ilmanifest, chiar, chiuder, riesc, fratell, idiot, maria, partigian, condivid, vedo, restan, gr_grim, neanc, camb, pessim, seguir, finanz, atto, recuper, ovunque, ognun, cit, chiamat]		

3	[rainews, seren, nord, amici, londr, puoi, frankcapiton, annalisa3073, porta, marisaladestr, splendid, fiorella0003, ferma, angel, chiunque, mario, vento, arrogante, sole, primi, molti, patrigiancott, signific, lasciat, voglion, campan, officialasrom, mcontroguerr, llagamiteut, potev, barbaralu66, responsabilit, dica, direttore, restar, però, toscan, caro, cara, rubino7004, abbracc, set-timan, eliolannutt, gastefy, potrà, bellissim, forte, dolce, primar, cuore, gol]		
4	[vittim, chapecoens, dimarted, tedesc, cinem, intern, valang, terzo, ron-dolin, pbersan, richiest, sequestr, mercat, figur, spread, coffeekla7, scusa, italicum, borse, borsa, sbagli, croll, rialz, schaeubl, decider, schaubl, capital, gioca, regional, santor, osservator, emilian, lapoelkann]		
5	[morte, fidelcastr, triest, requisit, minacc, avvamaestr, o.d, prorog, cammin, temporal, buonadomenic, indag, chiediam, speranz, pugl, termin, congress, monti, romapescar, genoajuv, cuba, adrianacioc, corte, marialuisazocc1, donne, barbaresc, piano, trump, tasse, cristallizzazione, firenz, nonunadimen, vbenetell, crozzatweet, giulianavesp, mondodivers, arenagilett, lauratam-agn, opzion, restitutionday, simonebaldell, violenz, castr, esclus, giure99, madia, omnibusla7, odproroga2018]	opzion prorog mar- ialuisazocc1	opzion mari- aluisazocc1 prorog
		vbenetell o.d donne odpro- donne temporal esclus chiediam speranz cristal- izzazione termin odproroga2018 temporal o.d	donne odpro- roga2018 esclus vbenetell cristal- izzazione termin speranz chiediam
6	[dato, parla, dovreb, europe, vince, dentr, ora, attes, pdnetwork, democraz, male, prevision, jobsact, parlar, grillin, valy_s, meno, kidstu, partit, solo, venut, pensar, matteorenz, sentit, beh, detto, palazz, prova, fuoco, anzi, appartien, continu, prezz, renzi, parlamentar, adess, brutt, referendum, aspett, provinc, ascoltar, prest, pieno, stata, andat, riman, diciam, debit, davver, chigi, minim, lusipatriz, giù, micheleemilian, avant, volut]	chigi palazz	chigi palazz
7	[continuan, promess, cattiv, magari, region, mln, verit, finta, segretar, m5s_camer, silv, quotidian, banc, cura, princip, nulla, mesi, energ, lavor, dire, mov5stell, gruber, montecitor, firma, dallavostrapart, padoan, stipend, luca, barbararaval, chiud, dice, bugie, merit, parlament, sostegn, immens, immagin, propost, centrodestr, ragion, finisc, mandat, amministrazion, mi-lan, sera, passo, govern]		
8	[aziend, condann, link, arrivan, costi, sprec, sito, cremaschig, serio, emanue-laneri14, immunit, poliz, grappatommas, animal, tratt, emozion, entro, ogni, limit, crede, vale, streaming, quesit, votar, votan, htt, sola, primo, falsa, legger, ben, ennesim, bufal, rende, riformacostituzional, altra, da-vant, far, vota, volet, gran, canal, domenic, scuol, brava, bastauns, medic, spieg, indecis, gioco, cina, votand, contribut, padre, ambient, guardar, stor, bellezz, iovoton]		
9	[vincenz, ciroarmen, vinca, deputat, oro, casalettor, squadr, serviz, restituit, reportrai3, malat, impegn, libro, report, milenagabanell, sms, pezzo, cara-binier, prego, convint, aereo, manovr, bce, titol, iacopo_jacjon, gabanell, perde, profil, quintacolonn, maxdanton, rivcristian, tecnic, acqua, incapac, ferit, zampierichiar, fidel, bancar, mercoled, grotond, presentazion, milen, massim, stud, frances, guerr, difend]	grotond rivcristian	grotond rivcristian
		pezzo gabanell milen perde report milenagabanell	milen perde ga- banell pezzo mile- nagabanell report
10	[riform, scompars, simonetoscano86, spiegar, risparm, mani, investment, iniziativ, scriv, petizion, librer, 4dicembr, francesc, chilhavistorai3, posti, napol, robert, cerca, schiform, false, votiam, difficult, firme, san, donna, more, 5bc32772e3fb467, bicameralism, bambin, attacc, luigidima, truff, chies, territor, costituzional, youtub, cinqu, campion, iodicon, poke_dem, controll, guardat, fatev]	iodicon 5bc32772e3fb467	iodicon 5bc32772e3fb467
		false firme simonetoscano86 librer	firme false librer simone- toscano86

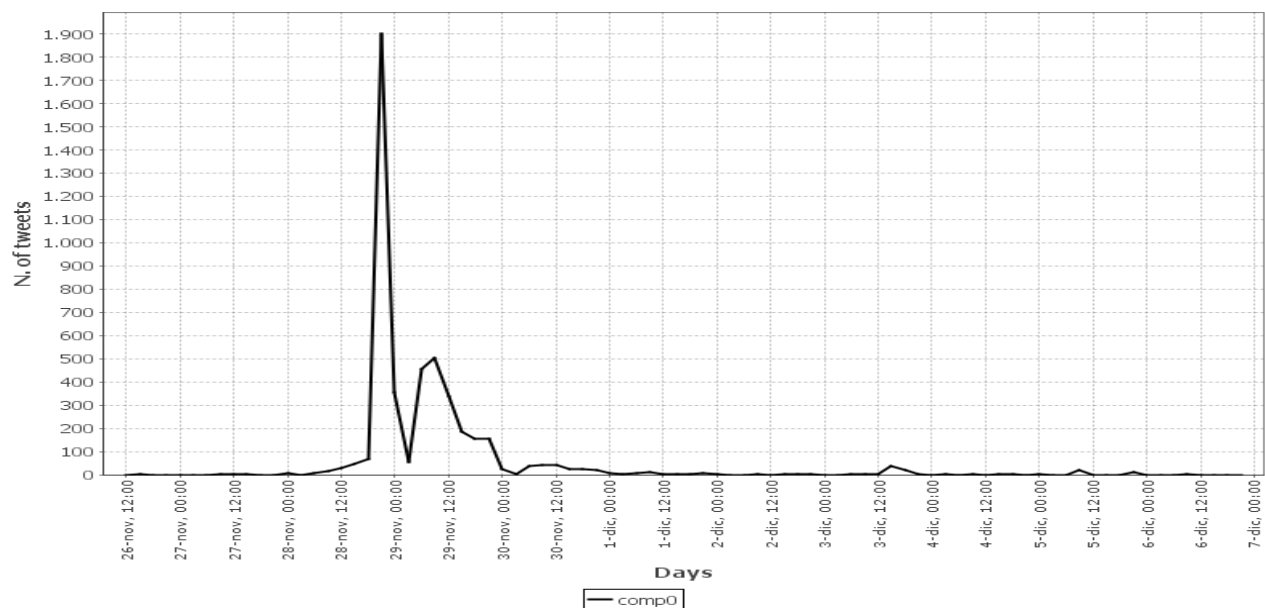
11	[scamb, boldrin, corruzion, statal, tav, data, event, referendar, natal, pdsin-istr, sanno, simpatic, etrur, ascolt, emanuelecariot, gliel, publicat, possa, danilotoninell, prend, schif, gruppoficamer, culo, incend, pensier, bruxelles, patrizia6119, santa, alepp, marco, memor, bugiard, ansa_econom, proiezion-idb, guard, dipendent, farlo, pur, perc, cani, dura, giro, imbrogli, metod, terremotat, apre, fredd, anno, presa, calo, lvoir, press, giudic, milanofinanz, regal, spot, tolto, vuol, mente, esist, soldi, averl, canzon, arrivat, poche]		
12	[error, finit, partecipazion, trastevererm, apert, vai, appen, leganordpadan, onda, dellorco85, albertobagna, derby, tocca, clabellocc, augur, votat, maestr, sentir, viva, diritt, arysedic, compliment, alto, andiam, voto, hash-tag, lastamp, record, speriam, sole24or, soprattutt, obama, giamma71, borghi_claud, rimast, urne, vincer, messagg, testa, editorial, padri, chiusur, popolar]		
13	[finir, costa, sanit, spesa, alessandr, ottobr, facebook, treno, inchiest, in-teress, dicembr, mld, quali, tappa, puntat, referendum4dicembr, reato, student, sindac, immigrat, vertic, miliard, final, voter, civil, tagl, interes-sant, donneinart, fogg, crescit, pensionat, tagliar, public, elett, palerm, letter, ascoltat, fondi, lagelon, arsenalekapp, motiv, t.c, quattr, inter, pagar, notiz, risors, sara, azion, passa, cronicar]		
14	[informat, tetto, voterann, appuntament, juncker, rinnov, gazebo, zero, veron, venerd, parliam, carlo, quellaschedanonesist, ale_mussolini., present, sindacat, perfett, comitaton, banda, natur, stefanoesposit, contratt, con-vincer, digital, process, torna, verament, mafia, comitatodeln, condannat, possibilit, fregar, sorpres, paura, prof, emergenz, destin, radiocusan, propa-gand, disse, incontr, consiglier, trenotour, seguit, migrant, cultur, accord, sicurez, senator, arturo_scott, unità, firmefals, ministr, approvat, eni, gen-itor, casta, confront, abolit, matteispond, ridicol, milion]	rinnov contratt	contratt rinnov
15	[viene, balle, giustiz, arrivar, morti, econom, infinit, amico, dare, tante, conferm, senat, mare, intant, aggiornament, ore, intervist, povert, amore, semplicement, vanno, serat, dibattit, sondagg, posto, usa, societ, leggi, liv-ell, berluscon, alecoscin, soli, raccont, info, siria, vita, condizion, confermat, comunist, capo, stare, anni, calc, piena, art, arte, francetomm, qualit, tre, ecco, ricevut, vergogn, chiedono, conosc, articol, vuoi, esemp, sotto, staser, dobbiam, commission, volte, sevenseasmarin, via]	aggiornament via	via aggiornament
16	[abbandonat, sandrobrizzola4, pront, qualunqu, onore, idee, ricordat, venet, doman, assolut, arroganz, pensav, succed, bel, dubbi, coglion, comunqu, uscit, francobechis, dico, bosc, napolitan, traged, affett, campagn, cena, ester, solidariet, dovrebbero, cadut, orfin, durant, marcotravagl, buone, gioia, lezion, chieder, tornat, ringraz]		
17	[miglior, ancor, cosa, cose, certo, soluzion, ciao, proposit, qui, ormai, metà, cazzar, dukana2, comun, nuova, mano, possiam, posso, social, vuole, veder, dite, perder, esser, molto, https, bene, stess, onest, oltre, fattoquotidian, moviment, antonio_bordin, terra, salut, qualcun, purtropp, altro, prima, t.co, addio, possibil, unica, dopp, campo, poter, pensi, futur, nonostant, ilmattofog, voti, color, fatto, comprar, tropp, causa, carloaltereg, pegg, piu, repubblic, cialtron, bravo, elettor, giust, scritt, avere, forza, volta, bisogn, ultim, fare]	color proposit t.co qui addio https piu comun nuova fat- toquotidian movi- ment	color addio movi- ment fattoquotid- ian nuova qui t.co https proposit
18	[emil, trovat, dati, brunett, spettacol, mztin, base, massimoleaderpd, comment, finalment, dignit, simil, alfan, pagin, decision, c_appendin, axd3izubcn, 24emil, senso, alema, patriziaramett, veri, cambiament, lu-cianobarracar, sportelloindecis, bersan, cespaz, assolt, armandocesar, alta, occupat, letta, pensa, elett, paga, cnel, regg, preso, document, nicolam-orra63, ariachetir, fondo, massa, attenzion, prodi, aria, esito, special, live, cgil, pericol, disoccupazione, sopra, tornar, disoccupat]	emil regg	regg emil
19	[dittatur, censis, divers, difendiam, siono, silenz, benign, girar, mentan, lunga, figli, dini, appell, carta, lucia, quorum, travagl, compleann, cos-tituent, ragazz, sched, visit, edicol, ferrarin, comitat, case, brogl, huffposti-tal, denunc, segret, scegl, raffiner, alber, voce, annaxnar, silenzioelettoral, mezzanott, iovotos]		

## 1.4

Previously we obtained groups of token, both for yes and for no users.

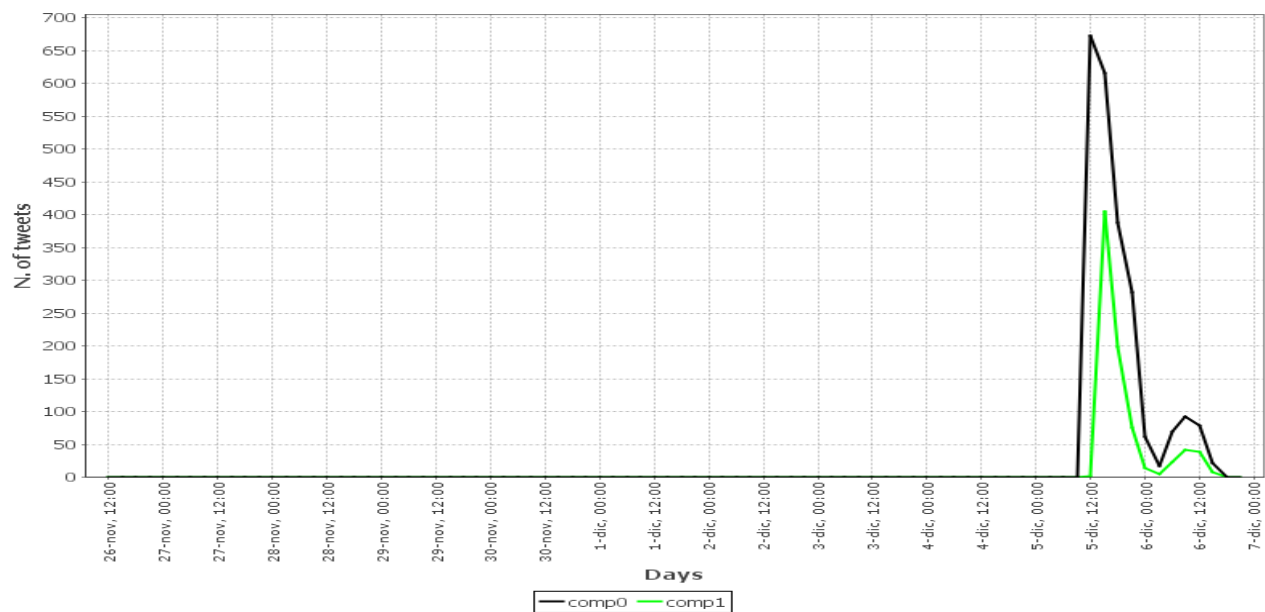
Now I set the grain to 3h and from the main index I checked up for each group of each cluster the number of tweets (containing terms of the considered group) distribution over time. I also saved text files for each component, for each interval (100 tweets for file).

```
cluster1
comp0: gabanell freq: 7952 report freq: 12746
tfreq: 4778
```



**Figure 4: YES cluster1**

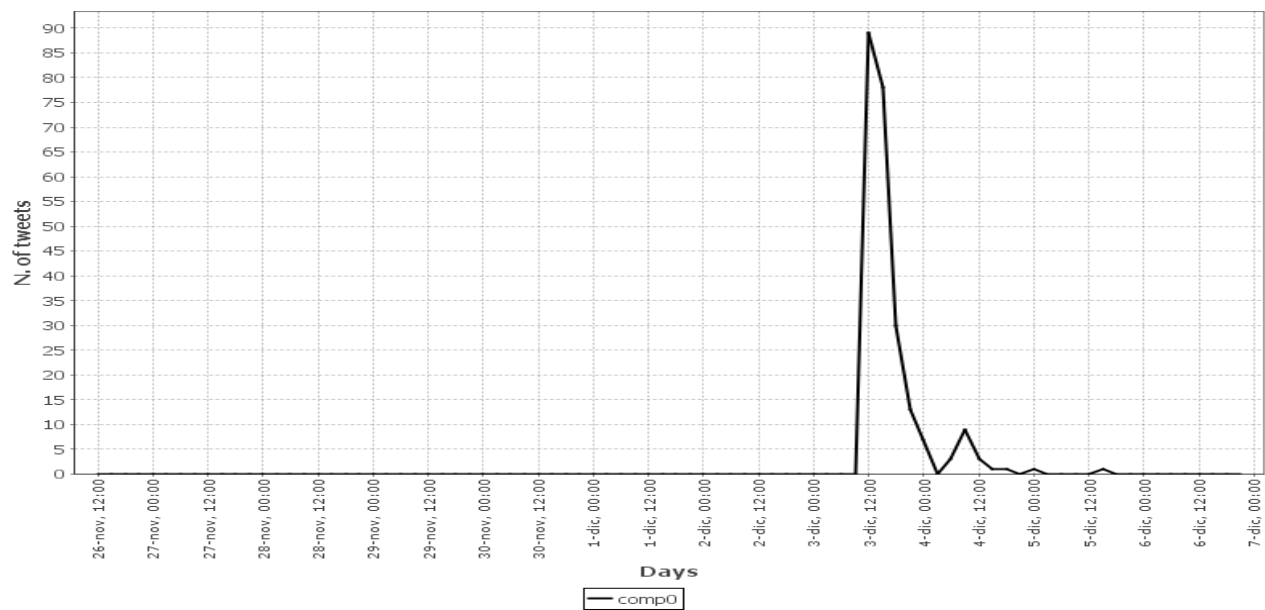
```
cluster3
comp0: kcgfbwnnrh freq: 2298 slide freq: 4055 viva freq: 28518 belli freq: 30527 singol
freq: 23949
tfreq: 2298
comp1: iniziati freq: 31282 ripartiam freq: 2129
tfreq: 812
```



**Figure 5: YES cluster3**

cluster6

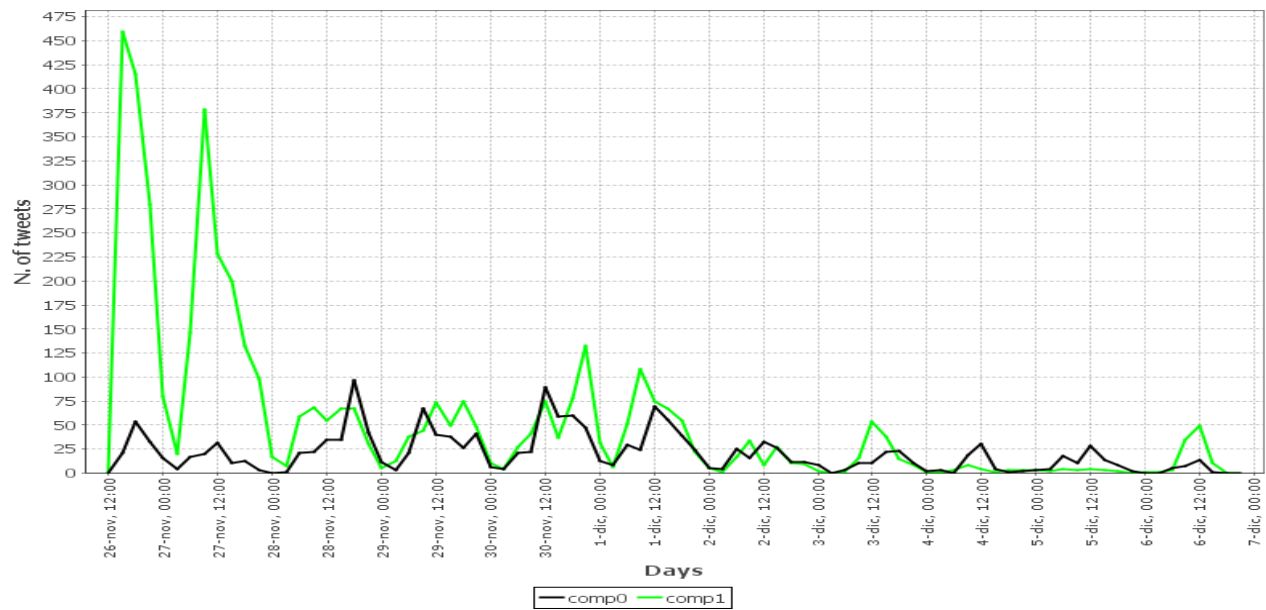
comp0: silenzioelettoral freq: 13012 proseguir freq: 1324 rompe freq: 4420 silenz freq: 43883  
tfreq: 236



**Figure 6: YES cluster6**

cluster7

comp0: virgin freq: 5019 raggi freq: 16710  
tfreq: 1691  
comp1: contratt freq: 24206 metalmeccanic freq: 7351  
tfreq: 4361



**Figure 7: YES cluster7**

cluster9

comp0: siren freq: 2244 abuso freq: 4382

tfreq: 540

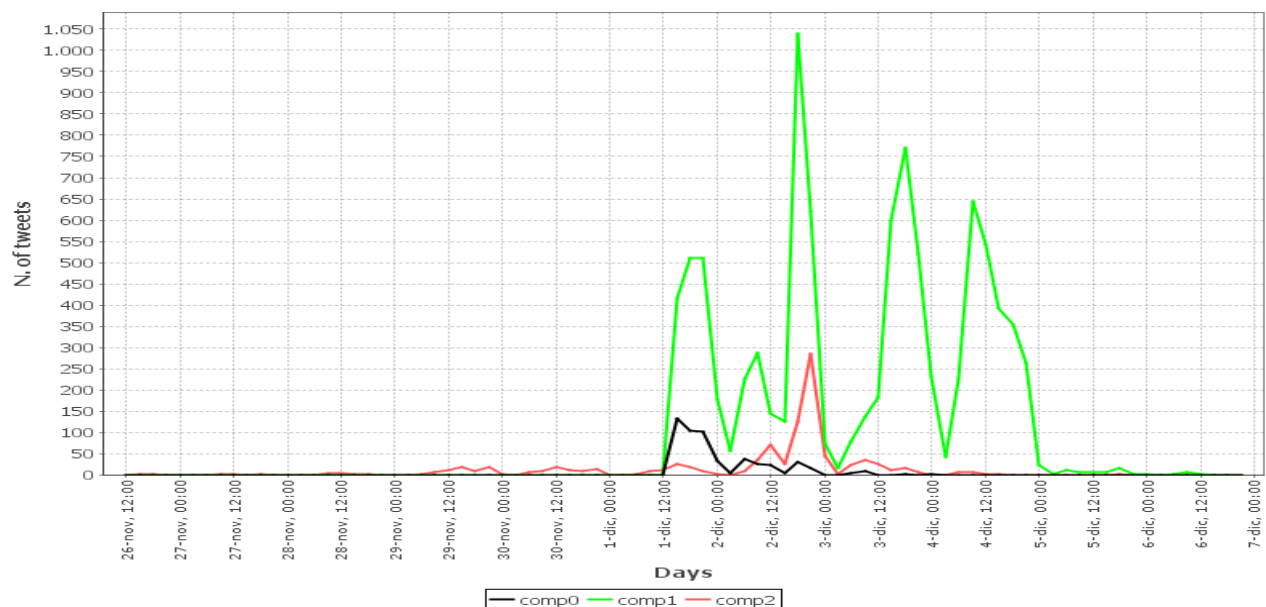
comp1: passo freq: 34545 insiem freq: 101951 important freq: 69525 primo freq: 107405

ogni freq: 181402 possiam freq: 42396

tfreq: 9284

comp2: chiusur freq: 9439 campagn freq: 40405 referendar freq: 12031

tfreq: 1001



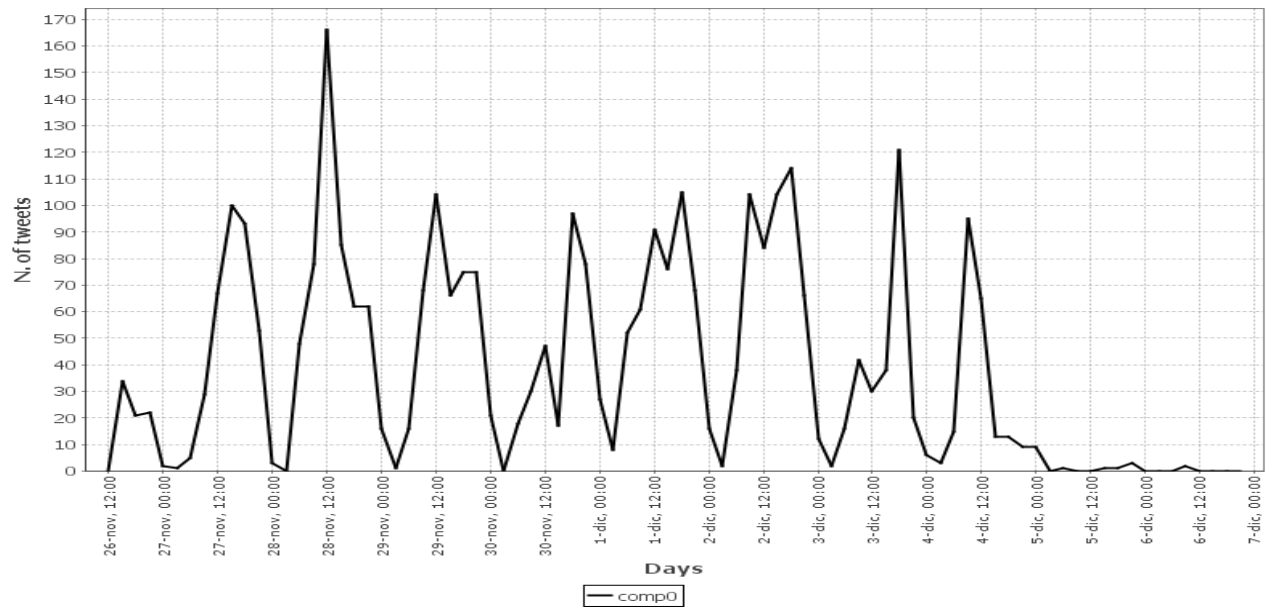
**Figure 8: YES cluster9**

cluster11

comp0: bastauns freq: 95151 4dicembr freq: 24769

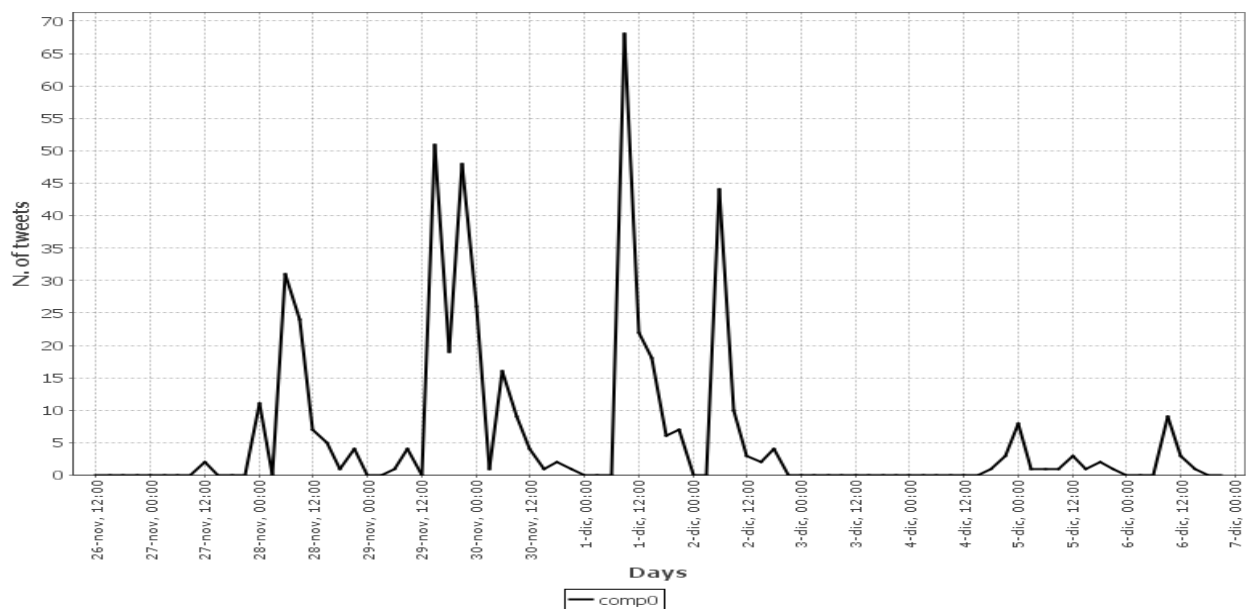


tfreq: 3193



**Figure 9:** YES cluster11

cluster12  
 comp0: cosedilavor freq: 711 nomfup freq: 2448  
 tfreq: 487



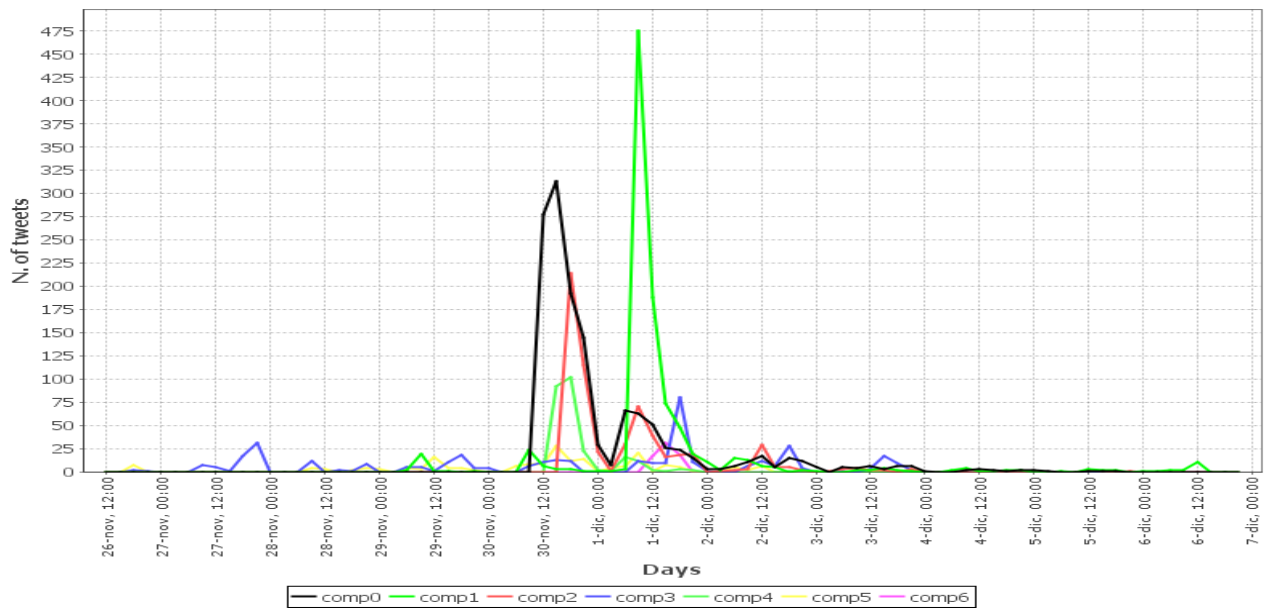
**Figure 10:** YES cluster12

cluster13  
 comp0: sali freq: 3230 joyotmzlau freq: 1336  
 tfreq: 1336  
 comp1: disoccupazione freq: 7743 ottobr freq: 7315  
 tfreq: 975

```

comp2: mariannamad freq: 3832 contrattop freq: 1350
tfreq: 614
comp3: liberis freq: 715 liberis freq: 909
tfreq: 398
comp4: ulivo freq: 1679 prodi freq: 30743 roman freq: 22519
tfreq: 262
comp5: session freq: 3719 wpc16it freq: 1392
tfreq: 201
comp6: mariaelenabozz1 freq: 219 yppi2011 freq: 364 jxgiulian freq: 596
tfreq: 68

```

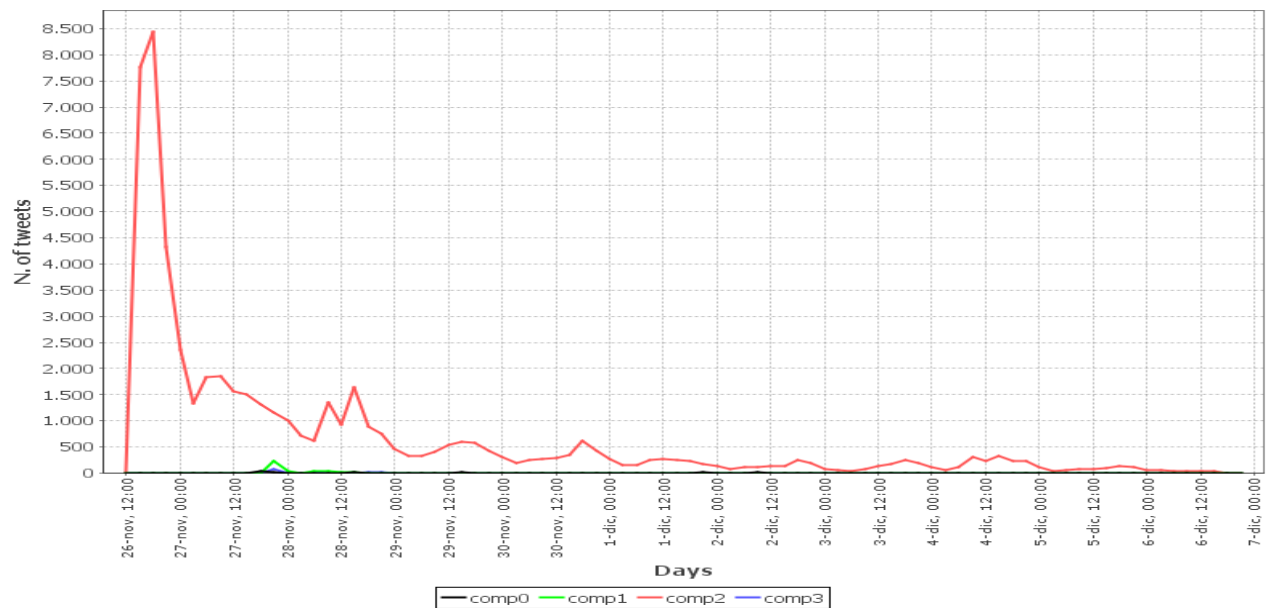


**Figure 11: YES cluster13**

```

cluster15
comp0: paritar freq: 1887 superament freq: 1019
tfreq: 184
comp1: carabinieri freq: 10462 arma freq: 5347 bologn freq: 39143
tfreq: 436
comp2: castr freq: 66881 fidel freq: 83777
tfreq: 54238
comp3: paoloall freq: 241 maurizio.lup freq: 931
tfreq: 138

```

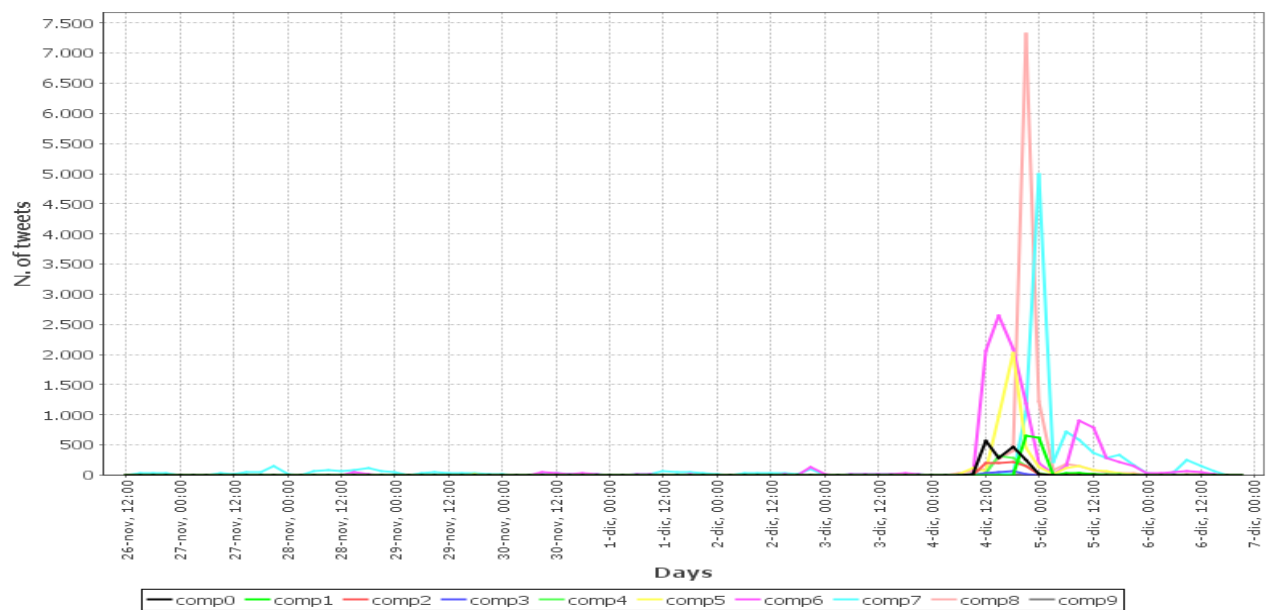


**Figure 12:** YES cluster15

```

cluster16
comp0: youtrend freq: 4658 affluenz freq: 26978
tfreq: 1611
comp1: exitpoll freq: 4546 maratonamentan freq: 78042
tfreq: 1388
comp2: labbufal freq: 3550 sorpres freq: 25351
tfreq: 840
comp4: copiativ freq: 4637 matit freq: 47066 cancellabil freq: 11153
tfreq: 196
comp5: saggezz freq: 5206 segg freq: 16358
tfreq: 962
comp6: bellen freq: 5339 austr freq: 19534 der freq: 8120 van freq: 10240
tfreq: 4386
comp7: piero freq: 20272 pelù freq: 15446
tfreq: 11604
comp8: chigi freq: 11511 palazz freq: 27937
tfreq: 11276
comp9: exit freq: 11922 poll freq: 10478
tfreq: 10182
comp10: codacons freq: 1644 totamif freq: 163
tfreq: 28

```



**Figure 13:** YES cluster16

cluster19

comp0: enric freq: 8540 mentan freq: 17196

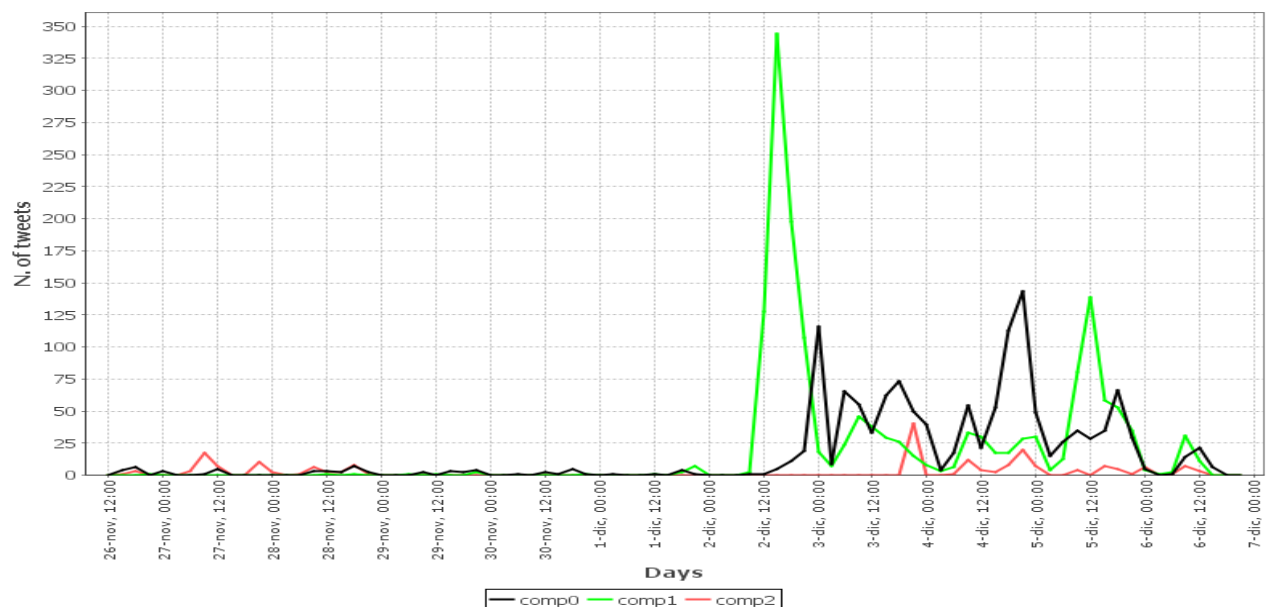
tfreq: 1338

comp1: pensand freq: 14416 votat freq: 97648

tfreq: 1603

comp2: lenelepr freq: 523 pacerenat freq: 324 bgrisaf freq: 728 vandab freq: 2115 orii151  
freq: 529

tfreq: 191

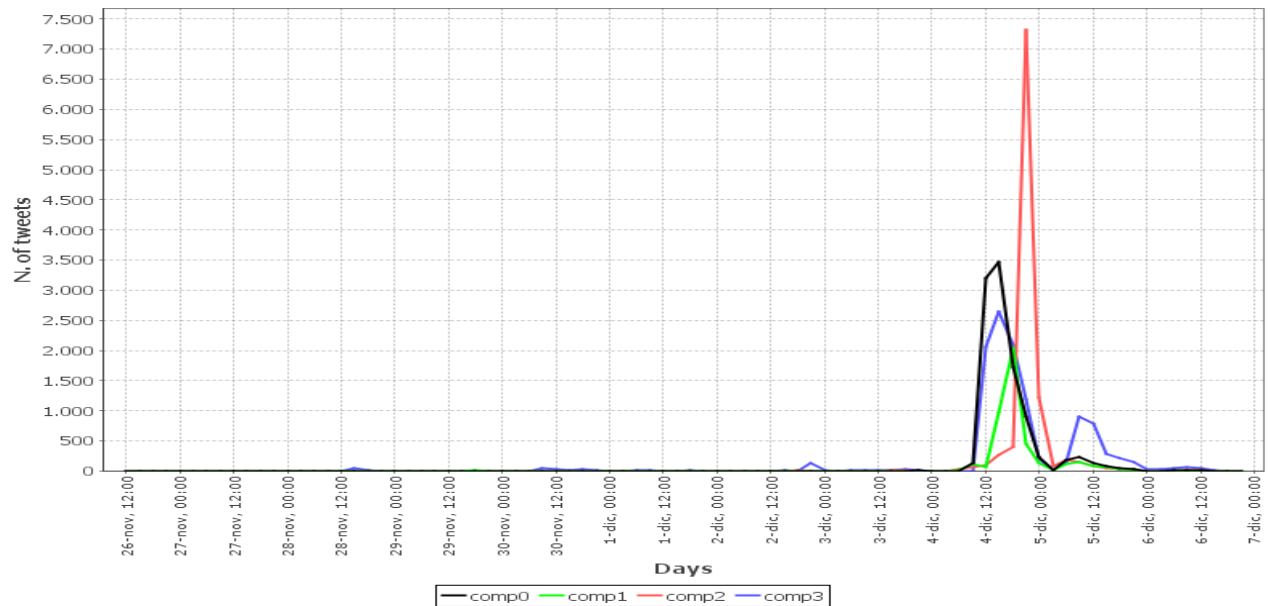


**Figure 14:** YES cluster19

```

cluster0
comp0:  matit freq:  47066 cancellabil freq:  11153
tfreq:  10494
comp1:  austr freq:  19534 der freq:   8120 van freq:  10240
tfreq:  4413
comp2:  exit freq:  11922 poll freq:  10478
tfreq:  10182
comp3:  piero freq:  20272 pelù freq:  15446
tfreq:  11604

```

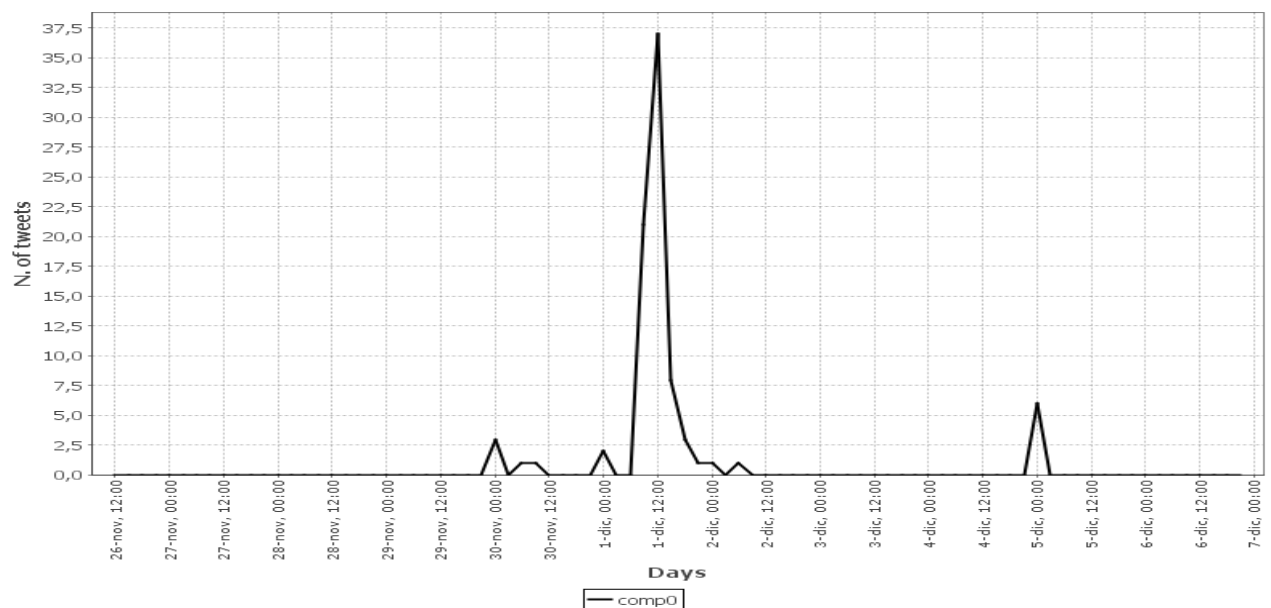


**Figure 15:** NO cluster0

```

cluster1
comp0:  meravigliosafrittur freq:  189 agropol freq:  2019
tfreq:  85

```



**Figure 16: NO cluster1**

cluster5

comp0: opzion freq: 7745 marialuisazoccl freq: 2157 prorog freq: 6159

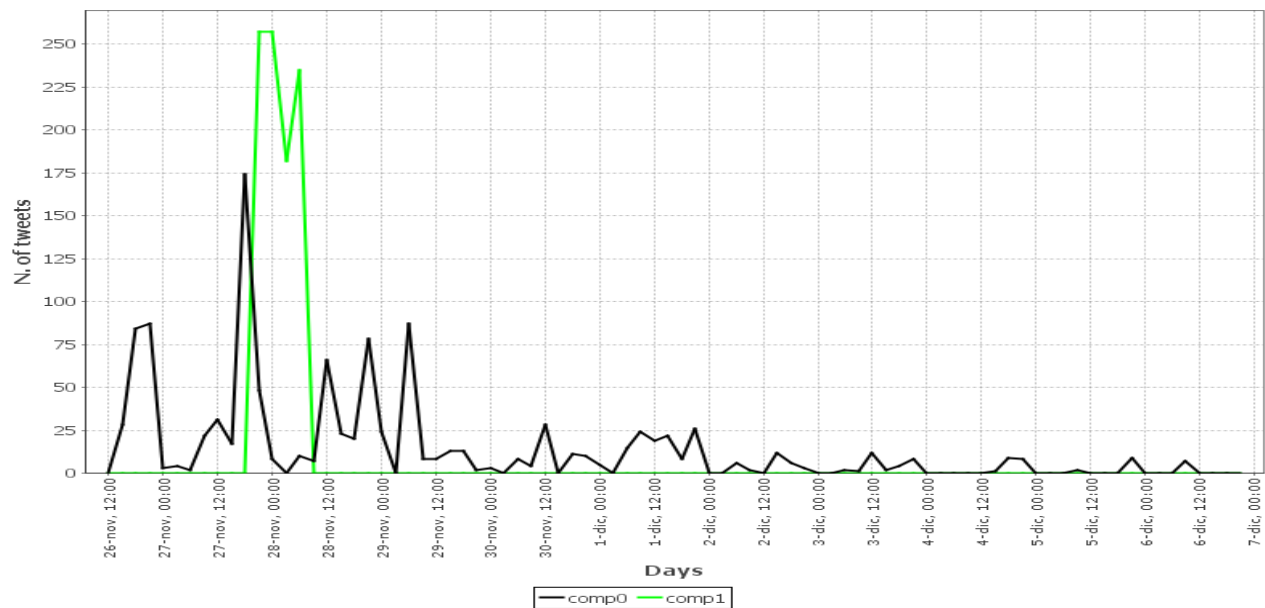
tfreq: 1144

comp1: donne freq: 65864 odproroga2018 freq: 4535 esclus freq: 6406 vbenetell freq: 2228

cristalizzazion freq: 1092 termin freq: 16627 speranz freq: 28526 chiediam freq: 6391 temporal

freq: 3105 o.d freq: 1098

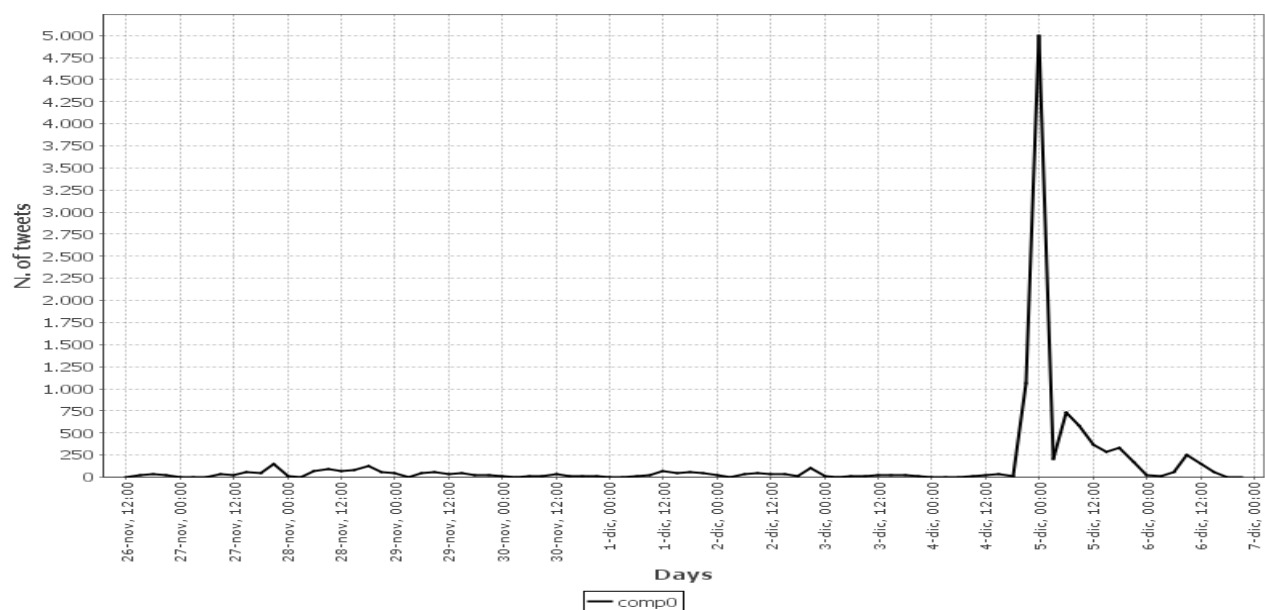
tfreq: 931

**Figure 17: NO cluster5**

cluster6

comp0: chigi freq: 11511 palazz freq: 27937

tfreq: 11276

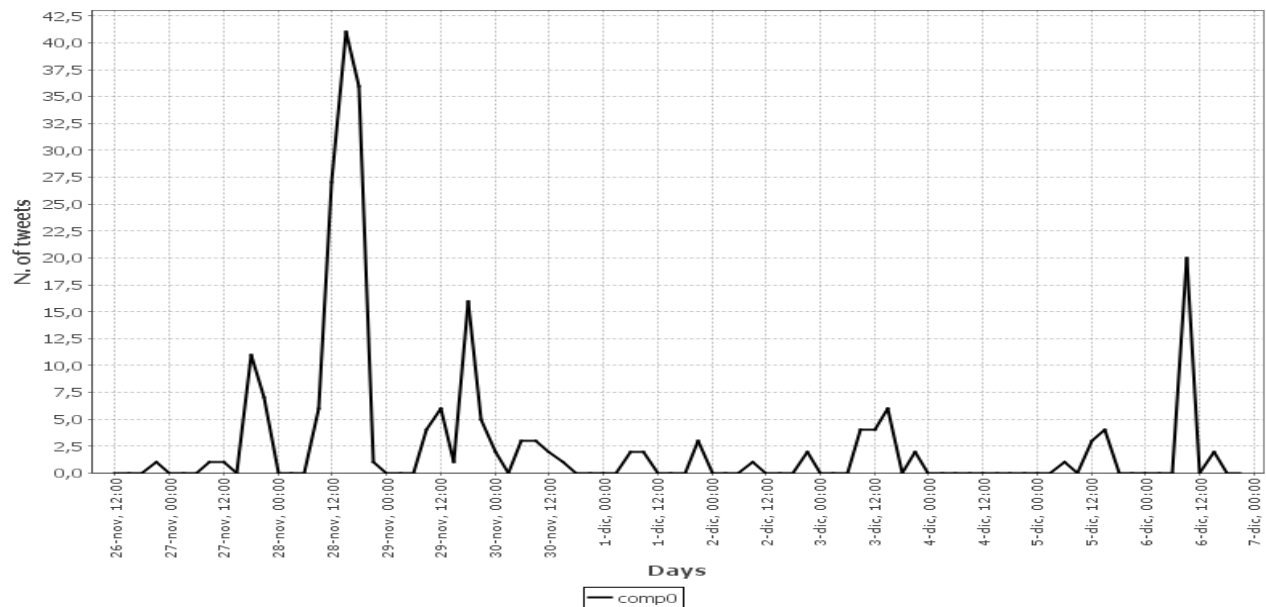


**Figure 18: NO cluster6**

```

cluster9
comp0: grotond freq: 1020 rivcristian freq: 563
tfreq: 231
comp1: milen freq: 6994 perde freq: 19790 gabanell freq: 7952 pezzo freq: 18367 milenagabanel
freq: 1989 report freq: 12746
tfreq: 0

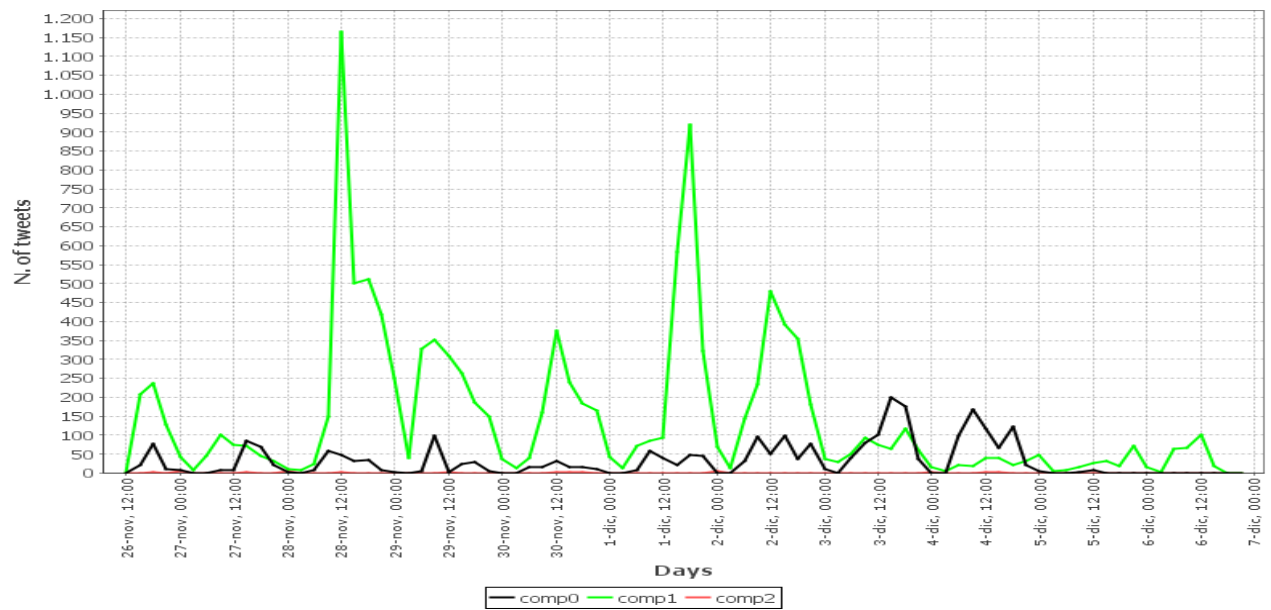
```

**Figure 19: NO cluster9**

```

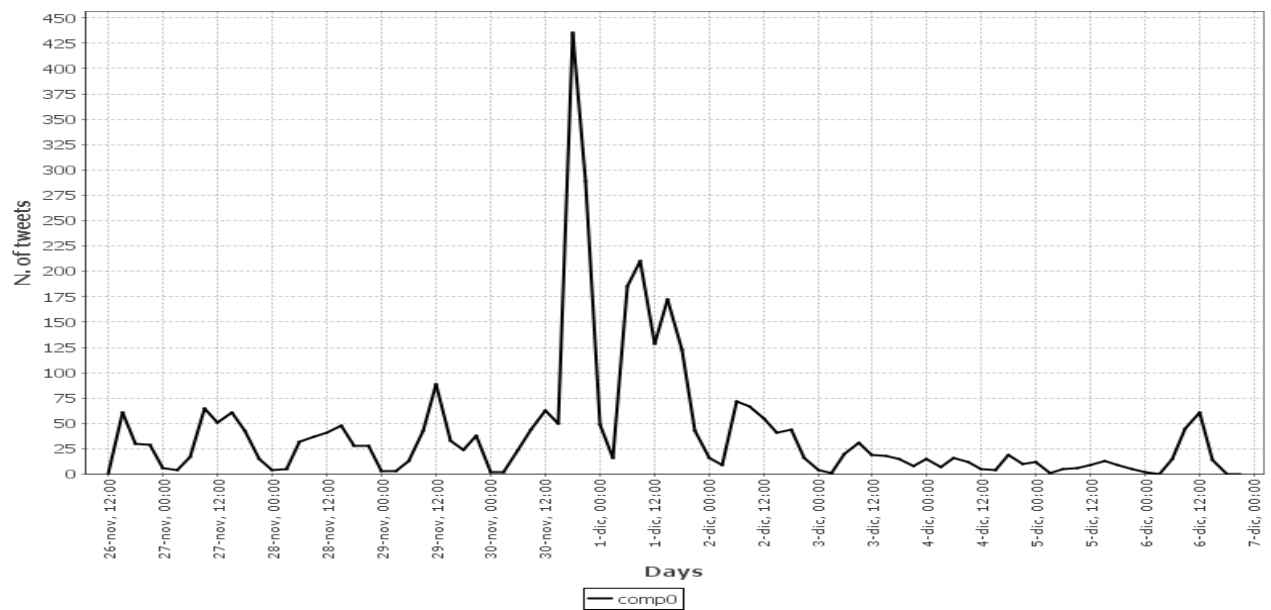
cluster10
comp0: iodicon freq: 60142 5bc32772e3fb467 freq: 4278
tfreq: 2646
comp1: firme freq: 19791 false freq: 20097
tfreq: 12086
comp2: librer freq: 7352 simonetoscano86 freq: 286
tfreq: 59

```



**Figure 20:** NO cluster10

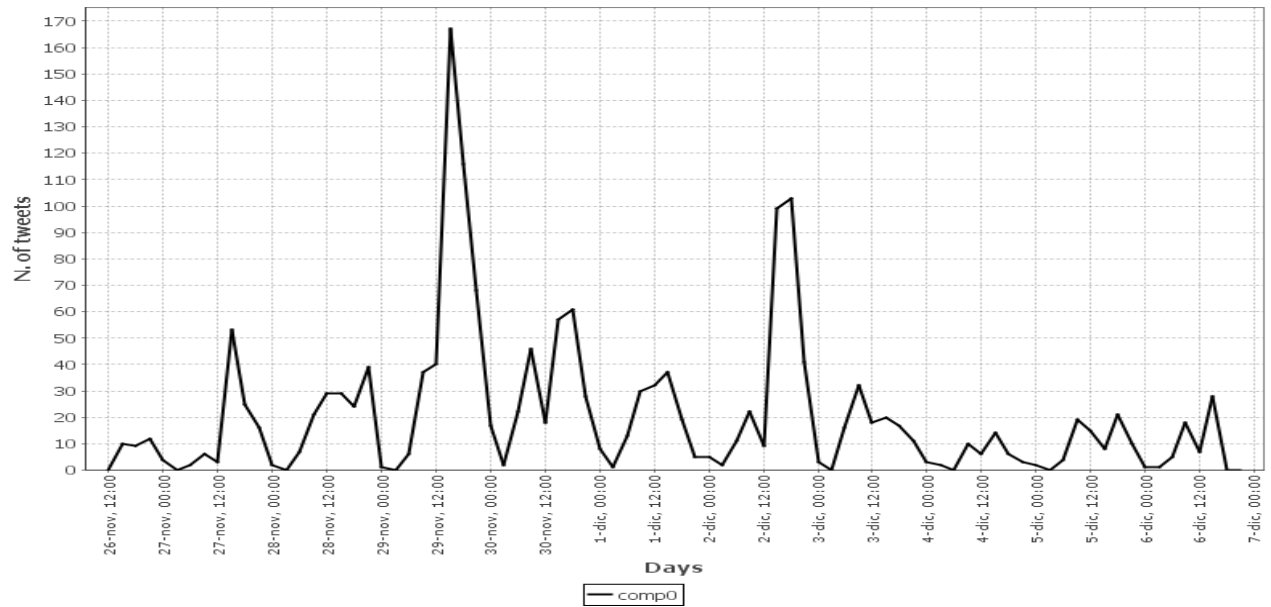
```
cluster14
comp0: contratt freq: 24206 rinnov freq: 12167
tfreq: 3408
```



**Figure 21:** NO cluster14

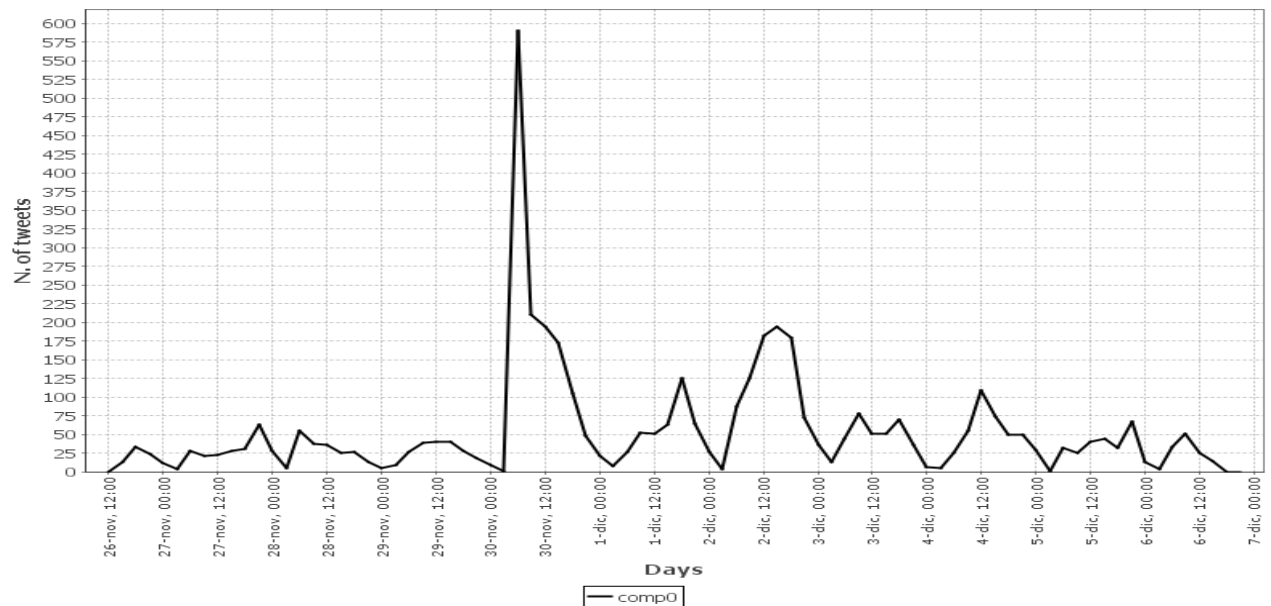
```
cluster15
comp0: via freq: 353733 aggiornament freq: 14979
tfreq: 1714
```





**Figure 22:** NO cluster15

```
cluster18
comp0: regg freq: 11170 emil freq: 12910
tfreq: 4526
```



**Figure 23:** NO cluster18

## 1.5

I finally searched for hashtags containing “yes” or “no” and then manually selected the relevant ones.

## 2 Identify mentions of candidates or YES/NO supporter

### 2.1

From a sampled network of 450193 twitter users we want to find the YES or NO supporters. For each user I counted the number of mentions of the previously found politicians/journalists for yes or

no. Then I made the same search with the number of hashtags, and again with the number of tweets containing all the terms of the previously found groups of tokens. If at least one of this counts overcomes a threshold(=5), I computed the score:

$$score = (nhashtags\_y * 2 + npoliticians\_y + ncomponents\_y) - (nhashtags\_n * 2 + npoliticians\_n + ncomponents\_n)$$

If the score is bigger than 0, I assigned the user to the yes supporters, if it is smaller I assigned him to the no supporters.

```
M yes:          5541
M no:           9455
N. of yes tweets: 1255221
N. of yes tweets: 2223702
```

No supporters are at least two times of yes supporters.

## 2.2

Given the subgraph induced by the previously found users (both yes or no supporters), I found the largest connected component and computed HITS algorithm.

```
lcc size:        14431
```

Thanks to the two lists created before, I could classify the hubs and authorities in yes or no supporters.

```
N yes authorities: 187
N no authorities:  813
N yes hubs:        253
N no hubs:         747
```

### authorities:

```
name:             pino habitat
screen name:       31239a839db94c9
followers:         1866
in-degree:         437
n mentions:        703
n hashtags-yes:    7
n hashtags-no:     83
n politicians-yes: 14
n politicians-no:  44
n components-yes:  6
n components-no:   11
```

```
name:             furio
screen name:       Giangiaagainst
followers:         5027
in-degree:         503
n mentions:        682
n hashtags-yes:    1
n hashtags-no:     5
n politicians-yes: 2
n politicians-no:  10
n components-yes:  1
n components-no:   2
```

```
name:             david
screen name:       Didi1648
followers:         1188
in-degree:         223
n mentions:        14
n hashtags-yes:    0
n hashtags-no:     8
n politicians-yes: 2
n politicians-no:  2
n components-yes:  1
```

n components-no: 3

name: leonardo  
screen name: leonardolao  
followers: 2193  
in-degree: 498  
n mentions: 478  
n hashtags=yes: 1  
n hashtags=no: 26  
n politicians=yes: 15  
n politicians=no: 14  
n components=yes: 0  
n components=no: 0

name: domenico formica  
screen name: domenicoformic3  
followers: 1746  
in-degree: 302  
n mentions: 26  
n hashtags=yes: 1  
n hashtags=no: 7  
n politicians=yes: 1  
n politicians=no: 0  
n components=yes: 0  
n components=no: 0

#### hubs:

name: il fatto quotidiano  
screen name: fattoquotidiano  
followers: 1654598  
in-degree: 54  
n mentions: 34749  
n hashtags=yes: 1  
n hashtags=no: 20  
n politicians=yes: 0  
n politicians=no: 25  
n components=yes: 22  
n components=no: 27

name: matteo renzi  
screen name: matteorenzi  
followers: 2754423  
in-degree: 28  
n mentions: 128706  
n hashtags=yes: 14  
n hashtags=no: 0  
n politicians=yes: 5  
n politicians=no: 0  
n components=yes: 8  
n components=no: 1

name: beppe grillo  
screen name: beppe.grillo  
followers: 2244415  
in-degree: 20  
n mentions: 34293  
n hashtags=yes: 3  
n hashtags=no: 159  
n politicians=yes: 4  
n politicians=no: 23  
n components=yes: 0  
n components=no: 0

name: la repubblica  
screen name: repubblicait

```
followers:      2506673
in-degree:      5
n mentions:     81986
n hashtags-yes: 10
n hashtags-no:  11
n politicians-yes: 19
n politicians-no: 8
n components-yes: 137
n components-no: 114
```

```
name:           sky tg24
screen name:     SkyTG24
followers:       2636921
in-degree:       12
n mentions:      19808
n hashtags-yes:  1
n hashtags-no:   25
n politicians-yes: 10
n politicians-no: 0
n components-yes: 24
n components-no: 13
```

## 2.3

I now created a root of 100 users supporting yes and a root of 100 users supporting no. The yes root is composed by the users with the higher yes score (computed as said before). Given the root, I also included in the graph users in the yes list connected to the root users. I finally computed HITS on this graph. I did the same for users supporting no.

**Results:** [YNsupporters/results3/New/yes-root.txt](#)

```
name:           midispiaceno
screen name:     MiDispiaceNO
followers:       66
n hashtags-yes:  2061
n hashtags-no:   56
n politicians-yes: 177
n politicians-no: 30
n components-yes: 6
n components-no:  3
SCORE:           4160
```

```
name:           angelinascanu votosi
screen name:     angelinascanu
followers:       1667
n hashtags-yes:  1267
n hashtags-no:   20
n politicians-yes: 749
n politicians-no: 55
n components-yes: 144
n components-no:  40
SCORE:           3292
```

```
name:           daniele dellavedova
screen name:     danieledvpd
followers:       1376
n hashtags-yes:  603
n hashtags-no:   14
n politicians-yes: 742
n politicians-no: 81
n components-yes: 129
n components-no:  46
SCORE:           1922
```

```

name:          senatori pd
screen name:   SenatoriPD
followers:     41645
n hashtags=yes: 281
n hashtags=no: 2
n politicians=yes: 727
n politicians=no: 6
n components=yes: 59
n components=no: 10
SCORE:        1328

```

```

name:          jenfa
screen name:   Jenfakoko
followers:     83
n hashtags=yes: 451
n hashtags=no: 1
n politicians=yes: 410
n politicians=no: 1
n components=yes: 7
n components=no: 0
SCORE:        1316

```

### Results: YNsupporters/results3/New/no-root.txt

```

name:          schiforma
screen name:   Schiforma
followers:     192
n hashtags=yes: 0
n hashtags=no: 4180
n politicians=yes: 31
n politicians=no: 426
n components=yes: 0
n components=no: 0
SCORE:        8755

```

```

name:          marino29b
screen name:   marino29b
followers:     1708
n hashtags=yes: 85
n hashtags=no: 1676
n politicians=yes: 179
n politicians=no: 697
n components=yes: 70
n components=no: 97
SCORE:        3727

```

```

name:          rino siconolfi
screen name:   rinosiconolfil
followers:     41167
n hashtags=yes: 0
n hashtags=no: 1690
n politicians=yes: 0
n politicians=no: 110
n components=yes: 2
n components=no: 0
SCORE:        3488

```

```

name:          demian yexil dimaio
screen name:   demian.yexil
followers:     590
n hashtags=yes: 69
n hashtags=no: 1241
n politicians=yes: 253

```

```
n politicians-no:      678
n components-yes:      65
n components-no:       70
SCORE:                 2774

name:                  alessandro niemiz
screen name:           niemiz1964
followers:             1529
n hashtags-yes:        16
n hashtags-no:         821
n politicians-yes:     59
n politicians-no:     1151
n components-yes:      17
n components-no:       20
SCORE:                 2705
```

### Results: YNsupporters/results3/New/yes.txt

```
starting size:         5541
|R|:                   100
|B|:                   2735
|S|=|B U R|:          2743
```

### Results: YNsupporters/results3/New/no.txt

```
starting size:         9455
|R|:                   100
|B|:                   5626
|S|=|B U R|:          5631
```

### yes authorities:

```
name:                  carla dossi
screen name:           DossiCarla
followers:             4229
n hashtags-yes:        79
n hashtags-no:         3
n politicians-yes:     73
n politicians-no:     18
n components-yes:      8
n components-no:       7

name:                  renzo seccia
screen name:           Renzoseccia
followers:             3446
n hashtags-yes:        47
n hashtags-no:         3
n politicians-yes:     52
n politicians-no:     22
n components-yes:      11
n components-no:       7

name:                  danielecina
screen name:           danielecina
followers:             10892
n hashtags-yes:        24
n hashtags-no:         3
n politicians-yes:     56
n politicians-no:     11
n components-yes:      18
n components-no:      48
```

```

name:                emiliano liberati
screen name:         ELiberati
followers:           2287
n hashtags=yes:      94
n hashtags=no:        0
n politicians=yes:    17
n politicians=no:     2
n components=yes:     5
n components=no:      3

```

```

name:                laura cesaretti
screen name:         lauracesaretti1
followers:           6572
n hashtags=yes:      15
n hashtags=no:        2
n politicians=yes:    43
n politicians=no:     9
n components=yes:     11
n components=no:      4

```

### yes hubs:

```

name:                matteo renzi
screen name:         matteorenzi
followers:           2754423
n hashtags=yes:      14
n hashtags=no:        0
n politicians=yes:    5
n politicians=no:     0
n components=yes:     8
n components=no:      1

```

```

name:                partito democratico
screen name:         pdnetwork
followers:           200416
n hashtags=yes:      203
n hashtags=no:        1
n politicians=yes:    267
n politicians=no:     4
n components=yes:     39
n components=no:      9

```

```

name:                la repubblica
screen name:         repubblicait
followers:           2506673
n hashtags=yes:      10
n hashtags=no:        11
n politicians=yes:    19
n politicians=no:     8
n components=yes:     137
n components=no:     114

```

```

name:                deputati pd
screen name:         Deputatipd
followers:           57302
n hashtags=yes:      248
n hashtags=no:        0
n politicians=yes:    339
n politicians=no:     6
n components=yes:     34
n components=no:     10

```

```

name:                graziano delrio
screen name:         graziano.delrio
followers:           193037
n hashtags=yes:      14

```

```
n hashtags-no:      0
n politicians-yes:  19
n politicians-no:   3
n components-yes:   1
n components-no:    0
```

#### no authorities:

```
name:               oinot49
screen name:        oinot49
followers:          1634
n hashtags-yes:     22
n hashtags-no:      153
n politicians-yes:   56
n politicians-no:    52
n components-yes:    6
n components-no:     5
```

```
name:               pino habitat
screen name:        31239a839db94c9
followers:          1866
n hashtags-yes:     7
n hashtags-no:      83
n politicians-yes:   14
n politicians-no:    44
n components-yes:    6
n components-no:    11
```

```
name:               furio
screen name:        Giangiaagainst
followers:          5027
n hashtags-yes:     1
n hashtags-no:      5
n politicians-yes:   2
n politicians-no:    10
n components-yes:    1
n components-no:     2
```

```
name:               leonardo
screen name:        leonardolao
followers:          2193
n hashtags-yes:     1
n hashtags-no:      26
n politicians-yes:   15
n politicians-no:    14
n components-yes:    0
n components-no:     0
```

```
name:               sandro&maria
screen name:        SandroPop1
followers:          1501
n hashtags-yes:     2
n hashtags-no:      17
n politicians-yes:   2
n politicians-no:    9
n components-yes:    1
n components-no:     0
```

#### no hubs:

```
name:               beppe grillo
screen name:        beppe_grillo
followers:          2244415
n hashtags-yes:     3
n hashtags-no:      159
n politicians-yes:   4
```



```
n politicians-no: 23
n components-yes: 0
n components-no: 0

name: il fatto quotidiano
screen name: fattoquotidiano
followers: 1654598
n hashtags-yes: 1
n hashtags-no: 20
n politicians-yes: 0
n politicians-no: 25
n components-yes: 22
n components-no: 27

name: lello esposito
screen name: LelloEsposito5
followers: 92540
n hashtags-yes: 21
n hashtags-no: 192
n politicians-yes: 125
n politicians-no: 103
n components-yes: 43
n components-no: 23

name: ?.friendsm5s.com
screen name: MPenikas
followers: 46121
n hashtags-yes: 31
n hashtags-no: 922
n politicians-yes: 104
n politicians-no: 692
n components-yes: 40
n components-no: 51

name: luigi di maio
screen name: luigidimaio
followers: 147034
n hashtags-yes: 0
n hashtags-no: 9
n politicians-yes: 0
n politicians-no: 1
n components-yes: 0
n components-no: 0
```

### 3 Spread of Influence

Finally we want to estimate the spread of influence of the two parties using the Label Propagation Algorithm. We have to first assign a label to each node of the graph (users):

0: if we do not know the party of the user.

1: if he is a yes supporter.

2: if he is a no supporter.

#### **initList:**

list is the list containing all the nodes which we want to change the label.

If the label is different from 0, we already know their label and we do not need to add the node to the list.

We do not need to add a node also if we already know that it has no other nodes that can influence it (indegree = 0).

#### **bestLabel:**

First of all we do not want 0 labels to propagate, so we have to remove them from the array of the neighbors.

If the length of the new array is 0, we can only return the 0 label.

If there is only one element left, it will return it.

Otherwise we can start counting the occurrences of the labels.

maxCount contains the label occurred more frequently.

If two labels have the same frequency, the algorithm will chose the new best randomly.

If  $(maxCount - counter) > (nneighborhood.length - i)$  we can be sure that the last best label is the label that we are searching for.

In the first case, M is divided in 5541 yes supporters and 9455 no supporters. And the results of the LPA are:

```
YES:          203090
NO:           219965
```

Obviously the results can change a little bit with the order of the list, which is random.

In the second case I chose the best 500 hubs for yes supporters and 500 for no supporters. And the results are:

```
YES:          241317
NO:           181392
```

### 3.1

I then tried to change the direction of the edges:

```
g.add(mapper.getId(edge[1]), mapper.getId(edge[0]), w);
```

And I used as M' the authorities instead of the hubs:

**Results:** SpreadOfInfluence/New/LPv1(authorities).txt

```
-----M-----
YES:          9225
NO:           378581
```

```
-----M'-----
YES:          2547
NO:           384289
```

But I also tried to use the hubs:

**Results:** SpreadOfInfluence/New/LPv1(hubs).txt

```
-----M'-----
YES:          3893
NO:           383006
```

Now the number of no-supporters is always bigger than the number of yes-supporters.

### 3.2

In LPv2 we can have 3 options: yes, no or not supporting. In this case anyone can change his idea thanks to his neighbors.

**Results:** SpreadOfInfluence/New/LPv2.txt

-----M-----

YES:	203090
NO:	219965
remaining:	576945

-----M'-----

YES:	241317
NO:	181392
remaining:	577291