Corso di Web Intelligence - A.A. 2016-2017

RELAZIONE

SECONDA CONSEGNA WEB INTELLIGENCE

A cura di:

Benetello Francesco 845626

PREMESSA

I risultati degli algoritmi(FP-Growth e Random Sampling) non sono stati scritti interamente, per la mole dei risultati prodotti.

Gli script sono stati eseguiti sia su file contenenti 10M di Tweets sia su Sample.json entrambi su macchine diverse di cui Notebook personale, Server VPS, Server di dipartimento assegnato, PC di laboratorio e Amazon AWS.

I risultati (frequents itemset e tempi di esecuzione) presi in considerazione sono riguardanti al notebook personale.

PROCESSO PRODUTTIVO

i.   Parsing del file JSon da 30 GB (10M Tweets) su cluster utilizzando SQL Context;
ii. Semplificazione del file selezionando la colonna "Text", corrispondente al campo tweet: il nuovo file ha dimensione di circa 800MB;
iii. Ulteriore semplificazione del file attraverso la rimozione delle stopwords inglesi, e la pulizia di altri caratteri come per esempio i trattini, i punti sostituiti da uno spazio. Il file finale da analizzare ha dimensione attorno ai 520MB;
iv. Applicazione dell'algoritmo FPGrowth sul file di testo finale contenente i tweets;
v.   Applicazione dell'algoritmo "Random Sampling" sul file finale;

I seguenti risultati sono stati ottenuti impostando:
supporto = 0.1 per Random Sampling
supporto = 0.25 per Random Sampling
supporto = 0.01 per FPGrowth
supporto = 0.001 per FPGrowth

OSSERVAZIONE

Eseguendo gli script sulla stessa macchina, si hanno avuti tempi di esecuzione diversi per ogni supporto inserito e per ogni Algoritmo eseguito.

Gli output di FPGrowth sono molto simili all'output di Random Sampling, mentre Random Sampling analizza un campione preso casualmente e quindi non vengono presi in considerazioni tutti gli items, ma solo quelli all'interno del campione e FPGrowth prende i valori del FP-Tree al di sopra del supporto prestabilito, questo spiega la leggera diversità tra gli esiti di un algoritmo e l'altro. Mentre come tempi di esecuzione FPGrowth è più efficiente in esecuzione e per avere più risultati occorre impostare valori di supporto molto più bassi rispetto al Random Sampling.

OUTPUT DELL'ALGORITMO "RandomSampling.py" e "fpgrowth.py"

Random Sampling Algorithm, supporto 0.1

Tempo di esecuzione: circa 33 min
Tempo di esecuzione: 1993.18 sec

FREQUENT ITEMSETS =====> OCCORRENZE

💪 ========> 24

8 ========> 8

🙆 ========> 35

💛 ========> 7

😍 ========> 25

😩 ========> 13

💁 ========> 9

😔 ========> 9

😐 ========> 9

👌 ========> 13

😏 ========> 9

: ========> 24

👀 ========> 8

💜 ========> 36

/ ========> 23

👩 ========> 34

🌍 ========> 34

😂 ========> 47

🙏 ========> 45

😊 ========> 20

🖤 ========> 21

😭 ========> 60

😳 ========> 7

♥ ========> 15

😴 ========> 7

♡ ========> 11

👑 ========> 15

💙 ========> 9

2 =======> 44

😘 =======> 11

4 =======> 27

◀ =======> 12

▶ =======> 12

5 =======> 21

👍 =======> 7

è =======> 16

😞 =======> 11

✌ =======> 14

1 =======> 23

3 =======> 28

6 =======> 10

😉 =======> 8

('world', '🌎') ========> 35

('happy', 'yourself!') ========> 25

('today', 'wind') ========> 18

('life', 'queen') ========> 9

(':)', 'life') ========> 9

('song!', 'year,') ========> 25

:(========> 9

('2015', 'vote') ========> 28

('@madisonellebeer', 'follow') ========> 13

('http://t.co/1domayesip', 'shop') ========> 14

😂😂 ========> 27

('0.0', 'mph') ========> 7

('http://t.co/8tttpri3ag', '👉abraham') ========> 18

tv ========> 9

('hpa,', 'rain') ========> 9

('follow', '◀') ========> 13

ur ========> 22

('idol', 'killing') ========> 19

si ========> 10

('(@alejandradextre)', 'pls') ========> 8

('#jobs', '#nursing') ========> 7

('know,', 'miles') ========> 13

('#viña2015', 'mateo') ========> 18

('mateo', '👉abraham') ========> 18

('knowing?', 'me?') ========> 13

('hpa,', 'temperature') ========> 7

('follow', 'u.') ========> 13

('fantastically', 'well!') ========> 8

('idol', 'me♥') ========> 9
('barometer', 'hpa,') ========> 7
('in.', 'today') ========> 7
('#viña2015', '☞abraham') ========> 18
(@ ========> 17
('in.', 'mph') ========> 7
('harassed', 'relentlessly') ========> 8
('#retail', '#tweetmyjobs') ========> 8
('ignore', '👩🏽') ========> 35
('humbly', 'stealing') ========> 19
('jabong,12.5%on', 'myntra.') ========> 14
('babe', 'follow') ========> 8
('(la', 'estar') ========> 18
('#viña2015', 'girls') ========> 18
('good', 'night') ========> 9
('flipkart,12%on', 'upto15%on') ========> 14
('dream?😍', 'mad') ========> 8
('!!', 'feeling:') ========> 14
('#job', '#nursing') ========> 9
('america', 'latin') ========> 8
('@madisonellebeer', 'idol') ========> 9
('**pic', 'heavy**') ========> 12
('dreams,', 'yourself!') ========> 25
('10', 'words') ========> 14
('killing', 'notice') ========> 19
('#jobs', 'manager') ========> 7
('cashback', 'myntra.') ========> 14
('@feriafitur', 'vote') ========> 28
('photo', 'posted') ========> 7
('girls', '☞abraham') ========> 18
('away?', 'knowing?') ========> 13
('mph', 'temperature') ========> 11
('away?', 'me?') ========> 13
('love', 'queen') ========> 9
('@justinbieber', 'know,') ========> 13
('#job', 'manager') ========> 7
('retweet', 'travelblogger') ========> 28
('follow', 'year,') ========> 29
('#3', '#getgarytonoonooswedding') ========> 8
('know,', 'love') ========> 13
('(la', 'mateo') ========> 18
('batman,', 'saved') ========> 13
('birthday', 'happy') ========> 36
('dreams,', 'follow') ========> 25
('http://t.co/8tttpri3ag', 'mateo') ========> 18
('review', 'vendor') ========> 12
('la):', 'mateo') ========> 18
('idol', 'queen') ========> 9

(':)', 'queen') ========> 9
('me?', 'miles') ========> 13
('song!', 'yourself!') ========> 25
('support', 'year,') ========> 25
('away?', 'know,') ========> 13
😍😍 ========> 8
('flipkart,12%on', 'snapdeal.') ========> 14
('#getgarytonoonooswedding', 'days') ========> 8
('@real_liam_payne', 'stealing') ========> 19
('indian', 'vendor') ========> 12
('@justinbieber', 'knowing?') ========> 13
la ========> 40
('#viña2015', '(la') ========> 18
('love', 'words') ========> 14
('en', 'http://t.co/8tttpri3ag') ========> 18
('!!', 'hunter".') ========> 14
('dreams,', 'support') ========> 25
('girls', 'http://t.co/8tttpri3ag') ========> 18
('mm.', 'today') ========> 10
('10', 'letters') ========> 14
('know,', 'u.') ========> 13
('miles', '♥♥♥♥how') ========> 13
(":'(", 'follow') ========> 8
('"2', 'words') ========> 14
('love', 'me♥') ========> 9
rn ========> 16
('@feriafitur', 'nelsoncarvalheiro') ========> 28
('mm.', 'temperature') ========> 9
('en', 'la):') ========> 18
yo ========> 11
(',realize', 'dream?😍') ========> 8
(':)', 'follow') ========> 9
('away?', 'follow') ========> 13
('https://t.co/1bds2lfusd', 'support') ========> 25
('me♥', 'queen') ========> 9
('killing', 'stealing') ========> 19
('mm.', 'wind') ========> 10
('#jobs', 'alert:') ========> 7
('travelblogger', 'vote') ========> 28
('away?', 'miles') ========> 13
('happy', 'make') ========> 8
('love', 'people') ========> 10
('@madisonellebeer', 'queen') ========> 9
('la', 'merese') ========> 18
('@feriafitur', 'world') ========> 28
('@justinbieber', 'person') ========> 13
('follow', 'yourself!') ========> 25
('"2', 'love') ========> 14
('@justinbieber', 'miles') ========> 13

('rain', 'temperature') =======> 17
('**pic', 'indian') =======> 12
('nelsoncarvalheiro', 'travelblogger') =======> 28
('#viña2015', 'en') =======> 18
('song!', 'support') =======> 25
('#jobs', 'health') =======> 7
('flat', 'jabong,12.5%on') =======> 14
('@justinbieber', '♥♥♥♥how') =======> 13
('barometer', 'today') =======> 17
('nelsoncarvalheiro', 'retweet') =======> 28
('harassed', 'make') =======> 8
('10', 'love') =======> 15
('temperature', 'today') =======> 17
('#gaanaapp', '@gaana') =======> 11
('@real_liam_payne', 'humbly') =======> 19
('fantastically', 'hope') =======> 8
('good', 'morning') =======> 12
('follow', 'knowing?') =======> 13
('@justinbieber', '♡') =======> 8

('email', 'remyexporter@gmail.com') =======> 12
('2015', 'fitur') =======> 28
.. =======> 15
('https://t.co/1bds2lfusd', 'year,') =======> 25
('@feriafitur', 'fitur') =======> 28
('happy', 'hope') =======> 7
('@real_liam_payne', 'killing') =======> 19
('@matthunter123', 'feeling:') =======> 14
('love', 'smile') =======> 8
('falling', 'rain') =======> 7
('humbly', 'killing') =======> 19
('in.', 'rain') =======> 7
('2015', 'nelsoncarvalheiro') =======> 28
('relentlessly', 'tweeted') =======> 8
('love', '♡') =======> 8

('@matthunter123', 'letters') =======> 14
('know,', 'knowing?') =======> 13
('love', 'you.') =======> 8
🙏🙏 =======> 14
('#getgarytonoonooswedding', 'relentlessly') =======> 8
('know,', 'person') =======> 13
('feeling:', 'love') =======> 14
('notice', 'ta') =======> 19
('merese', '👉abraham') =======> 18
('#getgarytonoonooswedding', 'harassed') =======> 8
se =======> 10
('#job', 'alert:') =======> 10
"2 =======> 13
('@justinbieber', 'follow') =======> 28

('flat', 'myntra.') ========> 14

('know,', '♥♥♥♥how') ========> 13

('en', '👉abraham') ========> 18

('follow', 'song!') ========> 25

('10', 'hunter".') ========> 14

('cashback', 'jabong,12.5%on') ========> 14

('dreams,', 'https://t.co/1bds2lfusd') ========> 25

🌿💜 ========> 10

('follow', '❤') ========> 8

#3 ========> 7

('@justinbieber', 'me?') ========> 13

❤️ ========> 22

(':)', '@madisonellebeer') ========> 9

('cashback', 'flat') ========> 14

('humbly', 'notice') ========> 19

('otrat', 'screams') ========> 8

('follow', 'https://t.co/1bds2lfusd') ========> 25

('feeling:', 'hunter".') ========> 14

('superhuman', '✌') ========> 13

('@garybarlow', 'days') ========> 8

('@matthunter123', 'hunter".') ========> 14

('hpa,', 'today') ========> 7

:) ========> 46

(":'(", 'love') ========> 8

('#3', 'relentlessly') ========> 8

em ========> 10

('letters', 'matt') ========> 14

('follow', 'important') ========> 14

('brazil', '😊') ========> 8

('follow', 'support') ========> 25

('harassed', 'tweeted') ========> 8

('cashback', 'of10.5%on') ========> 14

('@garybarlow', 'harassed') ========> 8

('princess', '👑') ========> 15

Random Sampling Algorithm, supporto 0.25

Tempo di esecuzione: 231 sec
Tempo di esecuzione: circa 4 min

FREQUENT ITEMSETS =====> OCCORRENZE
1 ========> 23

😂 =======> 47

/ ========> 23
2 ========> 44

😊 ========> 20

: ========> 24

🙏 ========> 45

4 ========> 27
3 ========> 28

💜 ========> 36

5 ========> 21

💪 ========> 24

😭 ========> 60

😍 ========> 25

🙋 ========> 35

🖤 ========> 21

🌍 ========> 34

è ========> 16

🙍 ========> 34

('follow', '😭') ========> 35

('@real_liam_payne', 'killing') ========> 19
('follow', 'happy') ========> 26
('(la', 'la):') ========> 18
('support', 'yourself!') ========> 25
('nelsoncarvalheiro', 'world') ========> 28
('humidity', 'today') ========> 17
('idol', 'notice') ========> 19
('@real_liam_payne', 'notice') ========> 19
('#viña2015', 'girls') ========> 18
('fitur', 'world') ========> 28
('#job', '#jobs') ========> 30

😂😂 ========> 27

('dreams,', 'year,') ========> 25
('en', 'merese') ========> 18
('vote', 'world') ========> 28
('fitur', 'retweet') ========> 28
('@real_liam_payne', 'idol') ========> 19
('nelsoncarvalheiro', 'vote') ========> 28
('@feriafitur', 'retweet') ========> 28
('nelsoncarvalheiro', 'retweet') ========> 28
('dream', 'realize') ========> 24
('en', 'girls') ========> 18

w/ ========> 17
('travelblogger', 'vote') ========> 28
im ========> 49
('barometer', 'wind') ========> 17
('birthday', 'happy') ========> 36
en ========> 20
('dreams,', 'follow') ========> 25
('girls', 'la') ========> 18
('humidity', 'temperature') ========> 16
('(la', 'http://t.co/8tttpri3ag') ========> 18
('@justinbieber', 'love') ========> 30
('!!', 'love') ========> 16
('(la', 'girls') ========> 18
('happy', 'song!') ========> 25
('humbly', 'notice') ========> 19
('temperature', 'today') ========> 17
('killing', 'ta') ========> 19

('la):', '👉abraham') ========> 18
('estar', 'girls') ========> 18
('idol', 'stealing') ========> 19
('@feriafitur', 'nelsoncarvalheiro') ========> 28
('https://t.co/1bds2lfusd', 'yourself!') ========> 25
('@feriafitur', 'fitur') ========> 28
rt ========> 23
:) ========> 46
('http://t.co/8tttpri3ag', '👉abraham') ========> 18
('follow', 'song!') ========> 25
('(la', 'merese') ========> 18
('http://t.co/8tttpri3ag', 'merese') ========> 18

❤️ ========> 22
('estar', 'http://t.co/8tttpri3ag') ========> 18
('barometer', 'humidity') ========> 17
('happy', 'support') ========> 26
('@justinbieber', 'follow') ========> 28
('#viña2015', 'merese') ========> 18
('notice', 'stealing') ========> 19
('2015', 'world') ========> 28
('#viña2015', '(la') ========> 18

('en', '👉abraham') ========> 18
('fitur', 'nelsoncarvalheiro') ========> 28
('ignore', '👩🏽') ========> 35
('en', 'la') ========> 18
('en', 'mateo') ========> 18
('en', 'estar') ========> 18
('#viña2015', 'http://t.co/8tttpri3ag') ========> 18
('#viña2015', 'la') ========> 18
('travelblogger', 'world') ========> 28
('humidity', 'wind') ========> 18
('la):', 'merese') ========> 18
('killing', 'notice') ========> 19
('follow', 'yourself!') ========> 25
('support', 'year,') ========> 25

('happy', 'year,') ========> 25
('rain', 'wind') ========> 19
('barometer', 'today') ========> 17
('2015', 'vote') ========> 28
('http://t.co/8tttpri3ag', 'la):') ========> 18
('estar', 'mateo') ========> 18
('(la', 'mateo') ========> 18
('barometer', 'rain') ========> 16
di ========> 27
('@real_liam_payne', 'stealing') ========> 19
rn ========> 16
('follow', 'love') ========> 53
('world', '🌍') ========> 35
('#viña2015', 'mateo') ========> 18
('#viña2015', '👉abraham') ========> 18
('(la', 'la') ========> 18
('estar', '👉abraham') ========> 18
('#viña2015', 'en') ========> 18
('fitur', 'travelblogger') ========> 28
('dreams,', 'song!') ========> 25
('(la', '👉abraham') ========> 18
('estar', 'merese') ========> 18
('girls', '👉abraham') ========> 18
('rain', 'today') ========> 19
('http://t.co/8tttpri3ag', 'la') ========> 18
(@ ========> 17
('girls', 'merese') ========> 18
('happy', 'yourself!') ========> 25
('#viña2015', 'estar') ========> 18
('stealing', 'ta') ========> 19
ur ========> 22
('humidity', 'rain') ========> 17
('humbly', 'idol') ========> 19
ya ========> 16
('song!', 'support') ========> 25
('killing', 'stealing') ========> 19
('nelsoncarvalheiro', 'travelblogger') ========> 28
('retweet', 'world') ========> 28
('idol', 'ta') ========> 19
('follow', 'year,') ========> 29
('today', 'wind') ========> 18
('@madisonellebeer', 'quote') ========> 17
('ilysfm', '💜') ========> 35
('la):', 'mateo') ========> 18
('2015', '@feriafitur') ========> 28
ta ========> 21
('la', 'merese') ========> 18
('humbly', 'stealing') ========> 19
('@real_liam_payne', 'humbly') ========> 19
('2015', 'nelsoncarvalheiro') ========> 28
('@feriafitur', 'world') ========> 28
('dreams,', 'happy') ========> 25

ma ========> 16
('2015', 'fitur') ========> 28
('idol', 'killing') ========> 19
('https://t.co/1bds2lfusd', 'song!') ========> 25
('retweet', 'travelblogger') ========> 28
('happy', 'https://t.co/1bds2lfusd') ========> 25
('follow', 'https://t.co/1bds2lfusd') ========> 25
('barometer', 'temperature') ========> 16
('en', 'http://t.co/8tttpri3ag') ========> 18
('@feriafitur', 'vote') ========> 28
('2015', 'retweet') ========> 28
!! ========> 33
('fitur', 'vote') ========> 28
('https://t.co/1bds2lfusd', 'year,') ========> 25
('@feriafitur', 'travelblogger') ========> 28
('dreams,', 'https://t.co/1bds2lfusd') ========> 25
('humbly', 'killing') ========> 19
('@real_liam_payne', 'ta') ========> 19
('year,', 'yourself!') ========> 25
la ========> 40
('2015', 'travelblogger') ========> 28
('en', 'la):') ========> 18
('girls', 'http://t.co/8tttpri3ag') ========> 18
('girls', 'mateo') ========> 18
('follow', 'support') ========> 25
('dreams,', 'support') ========> 25
('estar', 'la):') ========> 18
('rain', 'temperature') ========> 17
('song!', 'year,') ========> 25
('notice', 'ta') ========> 19
('#jobs', '#tweetmyjobs') ========> 28
('follow', 'me?') ========> 59
('song!', 'yourself!') ========> 25
('girls', 'la):') ========> 18
('merese', '👉abraham') ========> 18
('http://t.co/8tttpri3ag', 'mateo') ========> 18
('https://t.co/1bds2lfusd', 'support') ========> 25
('(la', 'estar') ========> 18
('#viña2015', 'la):') ========> 18
('dreams,', 'yourself!') ========> 25
('la', '👉abraham') ========> 18
('#job', '#tweetmyjobs') ========> 21
('me?', '😭') ========> 35
('mateo', '👉abraham') ========> 18
('temperature', 'wind') ========> 17
('humbly', 'ta') ========> 19
('estar', 'la') ========> 18
('mateo', 'merese') ========> 18
('retweet', 'vote') ========> 28
('la', 'la):') ========> 18
10 ========> 32
('(la', 'en') ========> 18
('love', 'me?') ========> 18

('la', 'mateo') ========> 18
('@feriafitur', 'retweet', 'vote') ========> 28
('2015', 'retweet', 'travelblogger') ========> 28

('girls', 'merese', '☝abraham') ========> 18
('follow', 'https://t.co/1bds2lfusd', 'year,') ========> 25
('2015', 'retweet', 'world') ========> 28
('@feriafitur', 'fitur', 'retweet') ========> 28

('http://t.co/8tttpri3ag', 'la', '☝abraham') ========> 18
('humbly', 'idol', 'notice') ========> 19
('http://t.co/8tttpri3ag', 'la', 'mateo') ========> 18
('en', 'mateo', 'merese') ========> 18
('@real_liam_payne', 'humbly', 'stealing') ========> 19
('girls', 'la):', 'merese') ========> 18
('fitur', 'travelblogger', 'world') ========> 28
('@feriafitur', 'fitur', 'nelsoncarvalheiro') ========> 28
('temperature', 'today', 'wind') ========> 17
('dreams,', 'follow', 'support') ========> 25
('(la', 'mateo', 'merese') ========> 18
('(la', 'http://t.co/8tttpri3ag', 'la):') ========> 18
('retweet', 'travelblogger', 'world') ========> 28
('en', 'estar', 'merese') ========> 18
('https://t.co/1bds2lfusd', 'song!', 'yourself!') ========> 25
('@real_liam_payne', 'humbly', 'notice') ========> 19
('fitur', 'vote', 'world') ========> 28
('@feriafitur', 'nelsoncarvalheiro', 'travelblogger') ========> 28
('2015', 'fitur', 'travelblogger') ========> 28
('(la', 'http://t.co/8tttpri3ag', 'la') ========> 18
😂😂😂 ========> 16
('(la', 'en', 'http://t.co/8tttpri3ag') ========> 18
('2015', 'fitur', 'vote') ========> 28
('dreams,', 'https://t.co/1bds2lfusd', 'year,') ========> 25
('(la', 'en', 'la):') ========> 18
('follow', 'happy', 'year,') ========> 25
('la', 'la):', 'merese') ========> 18
('http://t.co/8tttpri3ag', 'la', 'la):') ========> 18
('#viña2015', 'estar', 'mateo') ========> 18
('#viña2015', '(la', 'en') ========> 18
('en', 'la):', 'merese') ========> 18
('(la', 'en', 'mateo') ========> 18
('#viña2015', 'la', 'mateo') ========> 18
('#viña2015', 'mateo', 'merese') ========> 18
('girls', 'http://t.co/8tttpri3ag', 'mateo') ========> 18
('(la', 'girls', 'la') ========> 18
('(la', 'girls', 'mateo') ========> 18
('killing', 'notice', 'ta') ========> 19
('en', 'estar', 'girls') ========> 18
('dreams,', 'happy', 'support') ========> 25

FPGrowth Algorithm, supporto 0.01

Tempo di esecuzione: 0:00:03.452330 sec

FreqItemset(items=[u'@madisonellebeer'], freq=190)
FreqItemset(items=[u'@madisonellebeer', u''], freq=190)
FreqItemset(items=[u'people'], freq=161)
FreqItemset(items=[u'people', u''], freq=161)
FreqItemset(items=[u'me?'], freq=124)
FreqItemset(items=[u'me?', u''], freq=124)
FreqItemset(items=[u''], freq=12184)
FreqItemset(items=[u'today'], freq=153)
FreqItemset(items=[u'today', u''], freq=153)
FreqItemset(items=[u'love'], freq=507)
FreqItemset(items=[u'love', u''], freq=507)
FreqItemset(items=[u'lol'], freq=176)
FreqItemset(items=[u'lol', u''], freq=176)
FreqItemset(items=[u'follow'], freq=465)
FreqItemset(items=[u'follow', u''], freq=465)
FreqItemset(items=[u'world'], freq=175)
FreqItemset(items=[u'world', u''], freq=175)
FreqItemset(items=[u'time'], freq=146)
FreqItemset(items=[u'time', u''], freq=146)
FreqItemset(items=[u'make'], freq=170)
FreqItemset(items=[u'make', u''], freq=170)
FreqItemset(items=[u'back'], freq=146)
FreqItemset(items=[u'back', u''], freq=146)
FreqItemset(items=[u'happy'], freq=217)
FreqItemset(items=[u'happy', u''], freq=217)
FreqItemset(items=[u'birthday'], freq=127)
FreqItemset(items=[u'birthday', u''], freq=127)
FreqItemset(items=[u'good'], freq=215)
FreqItemset(items=[u'good', u''], freq=215)
FreqItemset(items=[u'day'], freq=163)
FreqItemset(items=[u'day', u''], freq=163)

FPGrowth Algorithm, supporto 0.001

Tempo di esecuzione: 0:00:03.523217 sec

Duration: 0:00:03.523217
FreqItemset(items=[u'0.0'], freq=14)
FreqItemset(items=[u'0.0', u''], freq=14)
FreqItemset(items=[u'send'], freq=29)
FreqItemset(items=[u'send', u''], freq=29)
FreqItemset(items=[u'lot'], freq=34)
FreqItemset(items=[u'lot', u''], freq=34)
FreqItemset(items=[u'save'], freq=15)
FreqItemset(items=[u'save', u''], freq=15)
FreqItemset(items=[u'hey'], freq=25)
FreqItemset(items=[u'hey', u''], freq=25)
FreqItemset(items=[u'ready'], freq=43)
FreqItemset(items=[u'ready', u''], freq=43)
FreqItemset(items=[u'body'], freq=16)
FreqItemset(items=[u'body', u''], freq=16)
FreqItemset(items=[u'da'], freq=18)
FreqItemset(items=[u'da', u''], freq=18)
FreqItemset(items=[u'show'], freq=47)
FreqItemset(items=[u'show', u''], freq=47)
FreqItemset(items=[u'paris'], freq=17)
FreqItemset(items=[u'paris', u''], freq=17)
FreqItemset(items=[u'waking'], freq=15)
FreqItemset(items=[u'waking', u''], freq=15)
FreqItemset(items=[u'part'], freq=31)
FreqItemset(items=[u'part', u''], freq=31)
FreqItemset(items=[u'nice'], freq=67)
FreqItemset(items=[u'nice', u''], freq=67)
FreqItemset(items=[u'college'], freq=18)
FreqItemset(items=[u'college', u''], freq=18)
FreqItemset(items=[u'talk'], freq=37)
FreqItemset(items=[u'talk', u''], freq=37)
FreqItemset(items=[u'app'], freq=20)
FreqItemset(items=[u'app', u''], freq=20)
FreqItemset(items=[u'giving'], freq=13)
FreqItemset(items=[u'giving', u''], freq=13)
FreqItemset(items=[u'#jobs'], freq=83)
FreqItemset(items=[u'#jobs', u''], freq=83)
FreqItemset(items=[u'\U0001f60f'], freq=18)
FreqItemset(items=[u'\U0001f60f', u''], freq=18)
FreqItemset(items=[u'\u2764\ufe0f'], freq=40)
FreqItemset(items=[u'\u2764\ufe0f', u'love'], freq=21)
FreqItemset(items=[u'\u2764\ufe0f', u'love', u''], freq=21)
FreqItemset(items=[u'\u2764\ufe0f', u''], freq=40)
FreqItemset(items=[u'understand'], freq=21)
FreqItemset(items=[u'understand', u''], freq=21)
FreqItemset(items=[u'https://t.co/1bds2lfusd'], freq=52)
FreqItemset(items=[u'https://t.co/1bds2lfusd', u'year,'], freq=52)
FreqItemset(items=[u'https://t.co/1bds2lfusd', u'year,', u'support'], freq=52)
FreqItemset(items=[u'https://t.co/1bds2lfusd', u'year,', u'support', u'follow'], freq=52)
FreqItemset(items=[u'https://t.co/1bds2lfusd', u'year,', u'support', u'follow', u''], freq=52)
FreqItemset(items=[u'https://t.co/1bds2lfusd', u'year,', u'support', u'happy'], freq=52)
FreqItemset(items=[u'https://t.co/1bds2lfusd', u'year,', u'support', u'happy', u'follow'], freq=52)
FreqItemset(items=[u'https://t.co/1bds2lfusd', u'year,', u'support', u'happy', u'follow', u''], freq=52)
FreqItemset(items=[u'https://t.co/1bds2lfusd', u'year,', u'support', u'happy', u''], freq=52)
FreqItemset(items=[u'https://t.co/1bds2lfusd', u'year,', u'support', u''], freq=52)

FreqItemset(items=[u'https://t.co/1bds2lfusd', u'year,', u'follow'], freq=52)
FreqItemset(items=[u'https://t.co/1bds2lfusd', u'year,', u'follow', u''], freq=52)
FreqItemset(items=[u'https://t.co/1bds2lfusd', u'year,', u'happy'], freq=52)
FreqItemset(items=[u'https://t.co/1bds2lfusd', u'year,', u'happy', u'follow'], freq=52)
FreqItemset(items=[u'https://t.co/1bds2lfusd', u'year,', u'happy', u'follow', u''], freq=52)
FreqItemset(items=[u'https://t.co/1bds2lfusd', u'year,', u'happy', u''], freq=52)
FreqItemset(items=[u'https://t.co/1bds2lfusd', u'year,', u''], freq=52)
FreqItemset(items=[u'https://t.co/1bds2lfusd', u'support'], freq=52)
FreqItemset(items=[u'https://t.co/1bds2lfusd', u'support', u'follow'], freq=52)
FreqItemset(items=[u'https://t.co/1bds2lfusd', u'support', u'follow', u''], freq=52)
FreqItemset(items=[u'https://t.co/1bds2lfusd', u'support', u'happy'], freq=52)
FreqItemset(items=[u'https://t.co/1bds2lfusd', u'support', u'happy', u'follow'], freq=52)
FreqItemset(items=[u'https://t.co/1bds2lfusd', u'support', u'happy', u'follow', u''], freq=52)
FreqItemset(items=[u'https://t.co/1bds2lfusd', u'support', u'happy', u''], freq=52)
FreqItemset(items=[u'https://t.co/1bds2lfusd', u'support', u''], freq=52)
FreqItemset(items=[u'https://t.co/1bds2lfusd', u'follow'], freq=52)
FreqItemset(items=[u'https://t.co/1bds2lfusd', u'follow', u''], freq=52)
FreqItemset(items=[u'https://t.co/1bds2lfusd', u'song!'], freq=52)
FreqItemset(items=[u'https://t.co/1bds2lfusd', u'song!', u'year,'], freq=52)
FreqItemset(items=[u'https://t.co/1bds2lfusd', u'song!', u'year,', u'support'], freq=52)
FreqItemset(items=[u'https://t.co/1bds2lfusd', u'song!', u'year,', u'support', u'follow'], freq=52)
FreqItemset(items=[u'https://t.co/1bds2lfusd', u'song!', u'year,', u'support', u'follow', u''], freq=52)
FreqItemset(items=[u'https://t.co/1bds2lfusd', u'song!', u'year,', u'support', u'happy'], freq=52)
FreqItemset(items=[u'https://t.co/1bds2lfusd', u'song!', u'year,', u'support', u'happy', u'follow'], freq=52)
FreqItemset(items=[u'https://t.co/1bds2lfusd', u'song!', u'year,', u'support', u'happy', u'follow', u''], freq=52)
FreqItemset(items=[u'https://t.co/1bds2lfusd', u'song!', u'year,', u'support', u'happy', u''], freq=52)
FreqItemset(items=[u'https://t.co/1bds2lfusd', u'song!', u'year,', u'support', u''], freq=52)
FreqItemset(items=[u'https://t.co/1bds2lfusd', u'song!', u'year,', u'follow'], freq=52)
FreqItemset(items=[u'https://t.co/1bds2lfusd', u'song!', u'year,', u'follow', u''], freq=52)
FreqItemset(items=[u'https://t.co/1bds2lfusd', u'song!', u'year,', u'happy'], freq=52)
FreqItemset(items=[u'https://t.co/1bds2lfusd', u'song!', u'year,', u'happy', u'follow'], freq=52)
FreqItemset(items=[u'https://t.co/1bds2lfusd', u'song!', u'year,', u'happy', u'follow', u''], freq=52)
FreqItemset(items=[u'https://t.co/1bds2lfusd', u'song!', u'year,', u'happy', u''], freq=52)
FreqItemset(items=[u'https://t.co/1bds2lfusd', u'song!', u'year,', u''], freq=52)
FreqItemset(items=[u'https://t.co/1bds2lfusd', u'song!', u'support'], freq=52)
FreqItemset(items=[u'https://t.co/1bds2lfusd', u'song!', u'support', u'follow'], freq=52)
FreqItemset(items=[u'https://t.co/1bds2lfusd', u'song!', u'support', u'follow', u''], freq=52)
FreqItemset(items=[u'https://t.co/1bds2lfusd', u'song!', u'support', u'happy'], freq=52)
FreqItemset(items=[u'https://t.co/1bds2lfusd', u'song!', u'support', u'happy', u'follow'], freq=52)
FreqItemset(items=[u'https://t.co/1bds2lfusd', u'song!', u'support', u'happy', u'follow', u''], freq=52)
FreqItemset(items=[u'https://t.co/1bds2lfusd', u'song!', u'support', u'happy', u''], freq=52)
FreqItemset(items=[u'https://t.co/1bds2lfusd', u'song!', u'support', u''], freq=52)
FreqItemset(items=[u'https://t.co/1bds2lfusd', u'song!', u'follow'], freq=52)
FreqItemset(items=[u'https://t.co/1bds2lfusd', u'song!', u'follow', u''], freq=52)
FreqItemset(items=[u'https://t.co/1bds2lfusd', u'song!', u'happy'], freq=52)
FreqItemset(items=[u'https://t.co/1bds2lfusd', u'song!', u'happy', u'follow'], freq=52)
FreqItemset(items=[u'https://t.co/1bds2lfusd', u'song!', u'happy', u'follow', u''], freq=52)
FreqItemset(items=[u'https://t.co/1bds2lfusd', u'song!', u'happy', u''], freq=52)
FreqItemset(items=[u'https://t.co/1bds2lfusd', u'song!', u''], freq=52)
FreqItemset(items=[u'https://t.co/1bds2lfusd', u'happy'], freq=52)
FreqItemset(items=[u'https://t.co/1bds2lfusd', u'happy', u'follow'], freq=52)
FreqItemset(items=[u'https://t.co/1bds2lfusd', u'happy', u'follow', u''], freq=52)
FreqItemset(items=[u'https://t.co/1bds2lfusd', u'happy', u''], freq=52)
FreqItemset(items=[u'https://t.co/1bds2lfusd', u''], freq=52)
FreqItemset(items=[u'flipkart,12%on'], freq=22)
FreqItemset(items=[u'flipkart,12%on', u'upto15%on'], freq=22)
FreqItemset(items=[u'flipkart,12%on', u'upto15%on', u''], freq=22)
FreqItemset(items=[u'flipkart,12%on', u''], freq=22)
FreqItemset(items=[u'brother'], freq=19)
FreqItemset(items=[u'brother', u''], freq=19)
FreqItemset(items=[u'posted'], freq=20)

FreqItemset(items=[u'posted', u'photo'], freq=14)
FreqItemset(items=[u'posted', u'photo', u''], freq=14)
FreqItemset(items=[u'posted', u''], freq=20)
FreqItemset(items=[u'ride'], freq=13)
FreqItemset(items=[u'ride', u''], freq=13)
FreqItemset(items=[u'jabong,12.5%on'], freq=22)
FreqItemset(items=[u'jabong,12.5%on', u'flat'], freq=22)
FreqItemset(items=[u'jabong,12.5%on', u'flat', u''], freq=22)
FreqItemset(items=[u'jabong,12.5%on', u'myntra.'], freq=22)
FreqItemset(items=[u'jabong,12.5%on', u'myntra.', u'flat'], freq=22)
FreqItemset(items=[u'jabong,12.5%on', u'myntra.', u'flat', u''], freq=22)
FreqItemset(items=[u'jabong,12.5%on', u'myntra.', u''], freq=22)
FreqItemset(items=[u'jabong,12.5%on', u'cashback'], freq=22)
FreqItemset(items=[u'jabong,12.5%on', u'cashback', u'flat'], freq=22)
FreqItemset(items=[u'jabong,12.5%on', u'cashback', u'flat', u''], freq=22)
FreqItemset(items=[u'jabong,12.5%on', u'cashback', u'myntra.'], freq=22)
FreqItemset(items=[u'jabong,12.5%on', u'cashback', u'myntra.', u'flat'], freq=22)
FreqItemset(items=[u'jabong,12.5%on', u'cashback', u'myntra.', u'flat', u''], freq=22)
FreqItemset(items=[u'jabong,12.5%on', u'cashback', u'myntra.', u''], freq=22)
FreqItemset(items=[u'jabong,12.5%on', u'cashback', u''], freq=22)
FreqItemset(items=[u'jabong,12.5%on', u'of10.5%on'], freq=22)
FreqItemset(items=[u'jabong,12.5%on', u'of10.5%on', u'flat'], freq=22)
FreqItemset(items=[u'jabong,12.5%on', u'of10.5%on', u'flat', u''], freq=22)
FreqItemset(items=[u'jabong,12.5%on', u'of10.5%on', u'myntra.'], freq=22)
FreqItemset(items=[u'jabong,12.5%on', u'of10.5%on', u'myntra.', u'flat'], freq=22)
FreqItemset(items=[u'jabong,12.5%on', u'of10.5%on', u'myntra.', u'flat', u''], freq=22)
FreqItemset(items=[u'jabong,12.5%on', u'of10.5%on', u'myntra.', u''], freq=22)
FreqItemset(items=[u'jabong,12.5%on', u'of10.5%on', u'cashback'], freq=22)
FreqItemset(items=[u'jabong,12.5%on', u'of10.5%on', u'cashback', u'flat'], freq=22)
FreqItemset(items=[u'jabong,12.5%on', u'of10.5%on', u'cashback', u'flat', u''], freq=22)
FreqItemset(items=[u'jabong,12.5%on', u'of10.5%on', u'cashback', u'myntra.'], freq=22)
FreqItemset(items=[u'jabong,12.5%on', u'of10.5%on', u'cashback', u'myntra.', u'flat'], freq=22)
FreqItemset(items=[u'jabong,12.5%on', u'of10.5%on', u'cashback', u'myntra.', u'flat', u''], freq=22)
FreqItemset(items=[u'jabong,12.5%on', u'of10.5%on', u'cashback', u'myntra.', u''], freq=22)
FreqItemset(items=[u'jabong,12.5%on', u'of10.5%on', u'cashback', u''], freq=22)
FreqItemset(items=[u'jabong,12.5%on', u'of10.5%on', u''], freq=22)
FreqItemset(items=[u'jabong,12.5%on', u''], freq=22)
FreqItemset(items=[u'2015'], freq=94)
FreqItemset(items=[u'2015', u'world'], freq=59)
FreqItemset(items=[u'2015', u'world', u''], freq=59)
FreqItemset(items=[u'2015', u''], freq=94)
FreqItemset(items=[u'screams'], freq=14)
FreqItemset(items=[u'screams', u'latin'], freq=13)
FreqItemset(items=[u'screams', u'latin', u'america'], freq=13)
FreqItemset(items=[u'screams', u'latin', u'america', u''], freq=13)
FreqItemset(items=[u'screams', u'latin', u''], freq=13)
FreqItemset(items=[u'screams', u'america'], freq=13)
FreqItemset(items=[u'screams', u'america', u''], freq=13)
FreqItemset(items=[u'screams', u''], freq=14)
FreqItemset(items=[u'head'], freq=26)
FreqItemset(items=[u'head', u''], freq=26)
FreqItemset(items=[u'nelsoncarvalheiro'], freq=58)
FreqItemset(items=[u'nelsoncarvalheiro', u'retweet'], freq=58)
FreqItemset(items=[u'nelsoncarvalheiro', u'retweet', u'2015'], freq=58)
FreqItemset(items=[u'nelsoncarvalheiro', u'retweet', u'2015', u'world'], freq=58)
FreqItemset(items=[u'nelsoncarvalheiro', u'retweet', u'2015', u'world', u''], freq=58)
FreqItemset(items=[u'nelsoncarvalheiro', u'retweet', u'2015', u''], freq=58)
FreqItemset(items=[u'nelsoncarvalheiro', u'retweet', u'world'], freq=58)
FreqItemset(items=[u'nelsoncarvalheiro', u'retweet', u'world', u''], freq=58)
FreqItemset(items=[u'nelsoncarvalheiro', u'retweet', u'vote'], freq=58)
FreqItemset(items=[u'nelsoncarvalheiro', u'retweet', u'vote', u'2015'], freq=58)
FreqItemset(items=[u'nelsoncarvalheiro', u'retweet', u'vote', u'2015', u'world'], freq=58)
FreqItemset(items=[u'nelsoncarvalheiro', u'retweet', u'vote', u'2015', u'world', u''], freq=58)

FreqItemset(items=[u'nelsoncarvalheiro', u'retweet', u'vote', u'2015', u''], freq=58)
FreqItemset(items=[u'nelsoncarvalheiro', u'retweet', u'vote', u'world'], freq=58)
FreqItemset(items=[u'nelsoncarvalheiro', u'retweet', u'vote', u'world', u''], freq=58)
FreqItemset(items=[u'nelsoncarvalheiro', u'retweet', u'vote', u''], freq=58)
FreqItemset(items=[u'nelsoncarvalheiro', u'retweet', u''], freq=58)
FreqItemset(items=[u'nelsoncarvalheiro', u'2015'], freq=58)
FreqItemset(items=[u'nelsoncarvalheiro', u'2015', u'world'], freq=58)
FreqItemset(items=[u'nelsoncarvalheiro', u'2015', u'world', u''], freq=58)
FreqItemset(items=[u'nelsoncarvalheiro', u'2015', u''], freq=58)
FreqItemset(items=[u'nelsoncarvalheiro', u'@feriafitur'], freq=58)
FreqItemset(items=[u'nelsoncarvalheiro', u'@feriafitur', u'retweet'], freq=58)
FreqItemset(items=[u'nelsoncarvalheiro', u'@feriafitur', u'retweet', u'2015'], freq=58)
FreqItemset(items=[u'nelsoncarvalheiro', u'@feriafitur', u'retweet', u'2015', u'world'], freq=58)
FreqItemset(items=[u'nelsoncarvalheiro', u'@feriafitur', u'retweet', u'2015', u'world', u''], freq=58)
FreqItemset(items=[u'nelsoncarvalheiro', u'@feriafitur', u'retweet', u'2015', u''], freq=58)
FreqItemset(items=[u'nelsoncarvalheiro', u'@feriafitur', u'retweet', u'world'], freq=58)
FreqItemset(items=[u'nelsoncarvalheiro', u'@feriafitur', u'retweet', u'world', u''], freq=58)
FreqItemset(items=[u'nelsoncarvalheiro', u'@feriafitur', u'retweet', u'vote'], freq=58)
FreqItemset(items=[u'nelsoncarvalheiro', u'@feriafitur', u'retweet', u'vote', u'2015'], freq=58)
FreqItemset(items=[u'nelsoncarvalheiro', u'@feriafitur', u'retweet', u'vote', u'2015', u'world'],
freq=58)
FreqItemset(items=[u'nelsoncarvalheiro', u'@feriafitur', u'retweet', u'vote', u'2015', u'world', u''],
freq=58)
FreqItemset(items=[u'nelsoncarvalheiro', u'@feriafitur', u'retweet', u'vote', u'2015', u''], freq=58)
FreqItemset(items=[u'nelsoncarvalheiro', u'@feriafitur', u'retweet', u'vote', u'world'], freq=58)
FreqItemset(items=[u'nelsoncarvalheiro', u'@feriafitur', u'retweet', u'vote', u'world', u''], freq=58)
FreqItemset(items=[u'nelsoncarvalheiro', u'@feriafitur', u'retweet', u'vote', u''], freq=58)
FreqItemset(items=[u'nelsoncarvalheiro', u'@feriafitur', u'retweet', u''], freq=58)
FreqItemset(items=[u'nelsoncarvalheiro', u'@feriafitur', u'2015'], freq=58)
FreqItemset(items=[u'nelsoncarvalheiro', u'@feriafitur', u'2015', u'world'], freq=58)
FreqItemset(items=[u'nelsoncarvalheiro', u'@feriafitur', u'2015', u'world', u''], freq=58)
FreqItemset(items=[u'nelsoncarvalheiro', u'@feriafitur', u'2015', u''], freq=58)
FreqItemset(items=[u'nelsoncarvalheiro', u'@feriafitur', u'world'], freq=58)
FreqItemset(items=[u'nelsoncarvalheiro', u'@feriafitur', u'world', u''], freq=58)
FreqItemset(items=[u'nelsoncarvalheiro', u'@feriafitur', u'vote'], freq=58)
FreqItemset(items=[u'nelsoncarvalheiro', u'@feriafitur', u'vote', u'2015'], freq=58)
FreqItemset(items=[u'nelsoncarvalheiro', u'@feriafitur', u'vote', u'2015', u'world'], freq=58)
FreqItemset(items=[u'nelsoncarvalheiro', u'@feriafitur', u'vote', u'2015', u'world', u''], freq=58)
FreqItemset(items=[u'nelsoncarvalheiro', u'@feriafitur', u'vote', u'2015', u''], freq=58)
FreqItemset(items=[u'nelsoncarvalheiro', u'@feriafitur', u'vote', u'world'], freq=58)
FreqItemset(items=[u'nelsoncarvalheiro', u'@feriafitur', u'vote', u'world', u''], freq=58)
FreqItemset(items=[u'nelsoncarvalheiro', u'@feriafitur', u'vote', u''], freq=58)
FreqItemset(items=[u'nelsoncarvalheiro', u'@feriafitur', u''], freq=58)
FreqItemset(items=[u'nelsoncarvalheiro', u'world'], freq=58)
FreqItemset(items=[u'nelsoncarvalheiro', u'world', u''], freq=58)
FreqItemset(items=[u'nelsoncarvalheiro', u'vote'], freq=58)
FreqItemset(items=[u'nelsoncarvalheiro', u'vote', u'2015'], freq=58)
FreqItemset(items=[u'nelsoncarvalheiro', u'vote', u'2015', u'world'], freq=58)
FreqItemset(items=[u'nelsoncarvalheiro', u'vote', u'2015', u'world', u''], freq=58)
FreqItemset(items=[u'nelsoncarvalheiro', u'vote', u'2015', u''], freq=58)
FreqItemset(items=[u'nelsoncarvalheiro', u'vote', u'world'], freq=58)
FreqItemset(items=[u'nelsoncarvalheiro', u'vote', u'world', u''], freq=58)
FreqItemset(items=[u'nelsoncarvalheiro', u'vote', u''], freq=58)
FreqItemset(items=[u'nelsoncarvalheiro', u''], freq=58)
FreqItemset(items=[u'tho'], freq=23)
FreqItemset(items=[u'tho', u''], freq=23)
FreqItemset(items=[u'@madisonellebeer'], freq=190)
FreqItemset(items=[u'@madisonellebeer', u'follow'], freq=25)
FreqItemset(items=[u'@madisonellebeer', u'follow', u'love'], freq=17)
FreqItemset(items=[u'@madisonellebeer', u'follow', u'love', u''], freq=17)
FreqItemset(items=[u'@madisonellebeer', u'follow', u''], freq=25)
FreqItemset(items=[u'@madisonellebeer', u'love'], freq=23)
FreqItemset(items=[u'@madisonellebeer', u'love', u''], freq=23)

FreqItemset(items=[u'@madisonellebeer', u''], freq=190)
FreqItemset(items=[u'blue'], freq=15)
FreqItemset(items=[u'blue', u''], freq=15)
FreqItemset(items=[u'text'], freq=24)
FreqItemset(items=[u'text', u''], freq=24)
FreqItemset(items=[u'tickets'], freq=13)
FreqItemset(items=[u'tickets', u''], freq=13)
FreqItemset(items=[u'join'], freq=16)
FreqItemset(items=[u'join', u''], freq=16)
FreqItemset(items=[u'job'], freq=31)