



Minería de textos con redes sociales

Los mercados son conversaciones

Sangre y muerte cada año en la #FeriaDeCali matando toros y caballitos. Espectáculo grotesco, para psicoanalizar al que lo disfruta.

#FeriaDeCali dicen q en mi otra vida debí ser colombiana... una mexicana que baila como caleña Prueba superada!!!

10:01 - Jueves 20 de Octubre 2016

Candidatura de Claudia López muestra la fragmentación de campaña 2018. Quien vaya unido desde el inicio, gana la presidencia @CeDemocratico

5:40 - 20 dic. 2016

Perfectas.



Reviews & Ratings for "Breaking Bad" More at IMDbPro >

Write review

Filter: Best Hide Spoilers

Page 1 of 93: [1] [2] [3] [4] [5] [6] [7] [8] [9] [10] [11] ▶

Index 921 reviews in total

705 out of 653 people found the following review useful:

Stunning Comedy-Drama ★★★★★

Author: sgtdobkes from United States
13 March 2008

I probably haven't been hooked to a TV show like I am to Breaking Bad before. This beautiful piece of art is incredibly well written and furthermore the actors are doing a tremendous job! I've read a few remarks about the pace of the show, saying it is too slow. I completely disagree, there are so many aspects which get their place in the series, every single one in its own way and it would be terrible if it was a too fast paced show containing ridiculous cliffhangers like we see them on the major networks. Because this way you can entirely fall in with the show, the characters and every tiny detail of the story and the best part of it, it is unbelievably addictive and makes you stay with it.

mercado libre

Reputación

100% de sus compradores lo recomiendan | 1350 ventas concluidas | Excelente vendedor ¡Es uno de los mejores del sitio! | 1 AÑO vendiendo en Mercado Libre

Ver más datos de Nikon

Calificaciones

7 compradores lo recomendaron por este producto

+ "Muy cumplidos, el producto me llegó a tiempo y en perfecto estado" DIJUANES200... 03 de abril

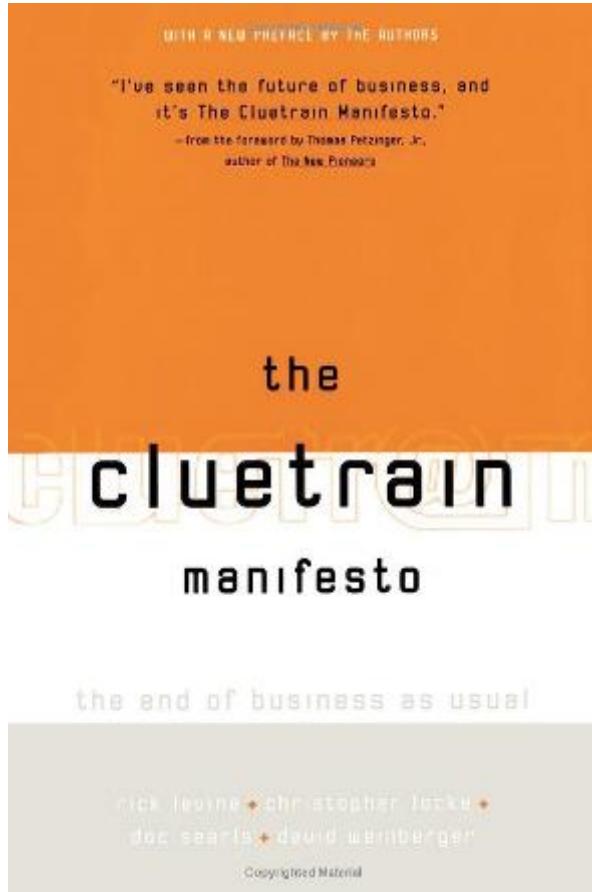
+ "fue amable y cumplido." BAMA75483... 11 de marzo

+ "Producto en excelente estado, entrega muy eficiente y rápida. Proceso de compra excelente." STARFLOWE... 11 de marzo

+ "Lo recomiendo 100%" SADI610... 24 de febrero

+ "El producto y la atención en perfecto estado, recomendable Bendiciones!" AL CANIN 19 de febrero

Ver más calificaciones



1 Los mercados son conversaciones.

2 Los mercados consisten de seres humanos, no de sectores demográficos.

3 Las conversaciones entre seres humanos suenan humanas. Se conducen en una voz humana.

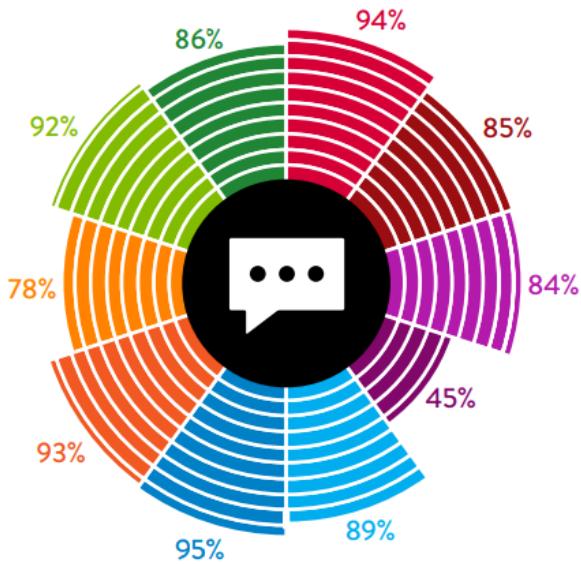
Se publica el Cluetrain Manifesto, un listado de 95 tesis que de forma visionaria empiezan a trazar el futuro del Internet que vendrá



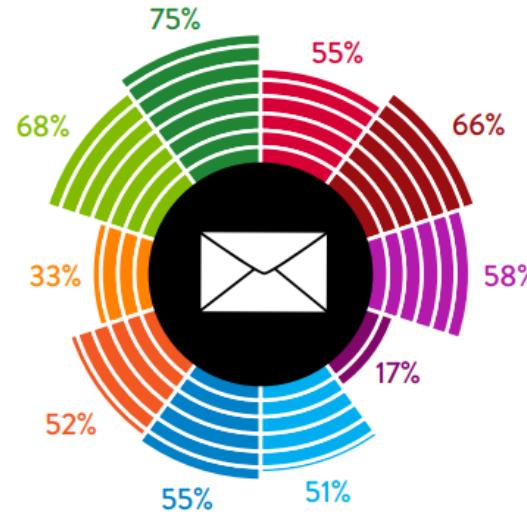
¿Qué hacemos con nuestros smartphones?

5

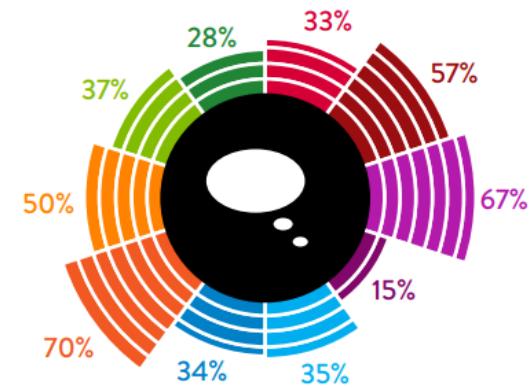
TEXT MESSAGING / SMS



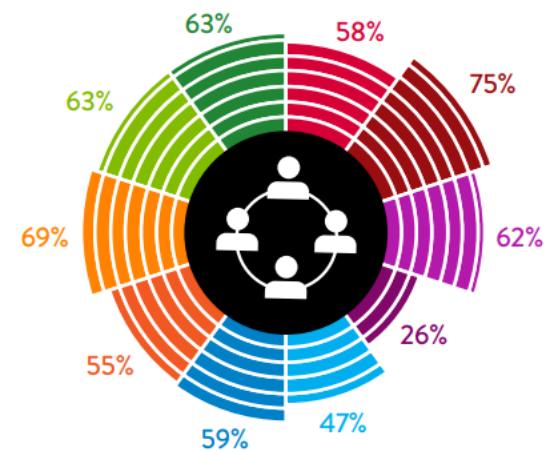
EMAIL



INSTANT MESSAGING



SOCIAL NETWORKING



AUSTRALIA ●
BRAZIL ●

CHINA ●
INDIA ●

ITALY ●
RUSSIA ●

SOUTH KOREA ●
TURKEY ●

UNITED KINGDOM ●
UNITED STATES ●

913.242.000

Número de textos que se envían
cada día en el mundo (sin incluir
los mensajes app-to-app)

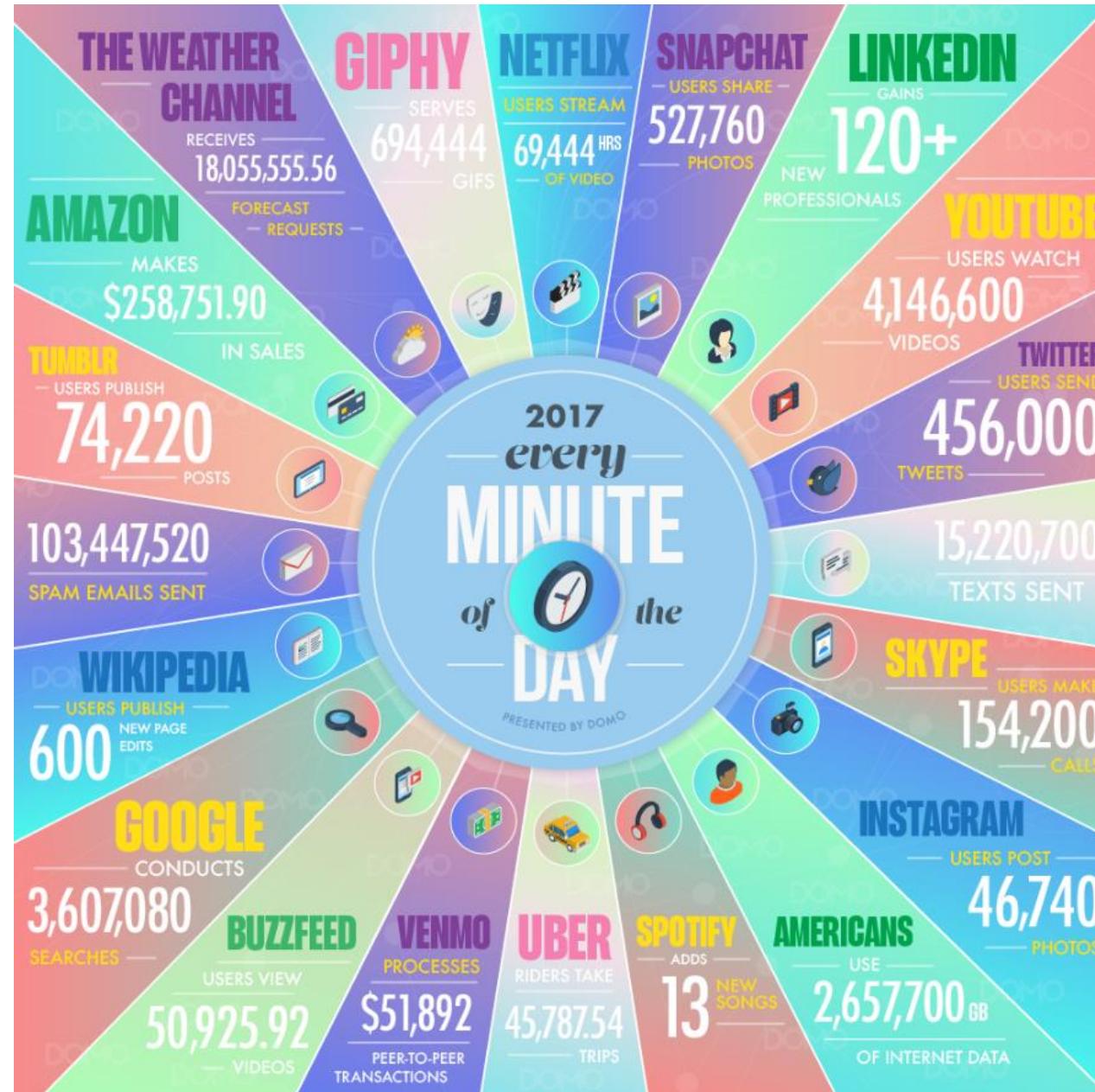
60.000.000.000

Número de textos enviados
cada día por Whatsapp y
Messenger

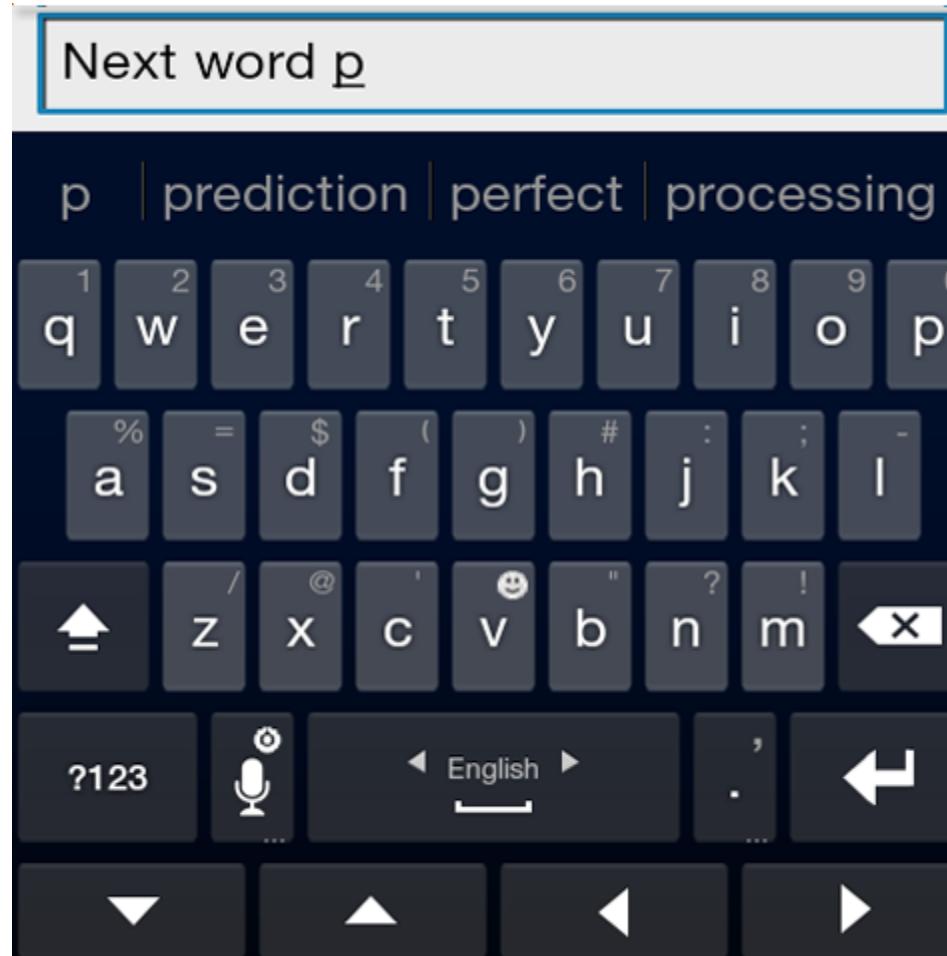
<https://www.domo.com/learn/data-never-sleeps-5>

<https://www.theverge.com/2016/4/12/11415198/facebook-messenger-whatsapp-number-messages-vs-sms-f8-2016>





Ejemplos de aplicaciones web que usan text mining





que

- que significa tbt
- que es la otan
- que son los derechos humanos
- que dios te bendiga
- que es la etica
- que es un ensayo
- que es el amor
- que es democracia
- que es emprendimiento
- que es el respeto

Buscar con Google Me siento con suerte

Denunciar predicciones ofensivas

Ejemplos de aplicaciones web que usan text mining

10

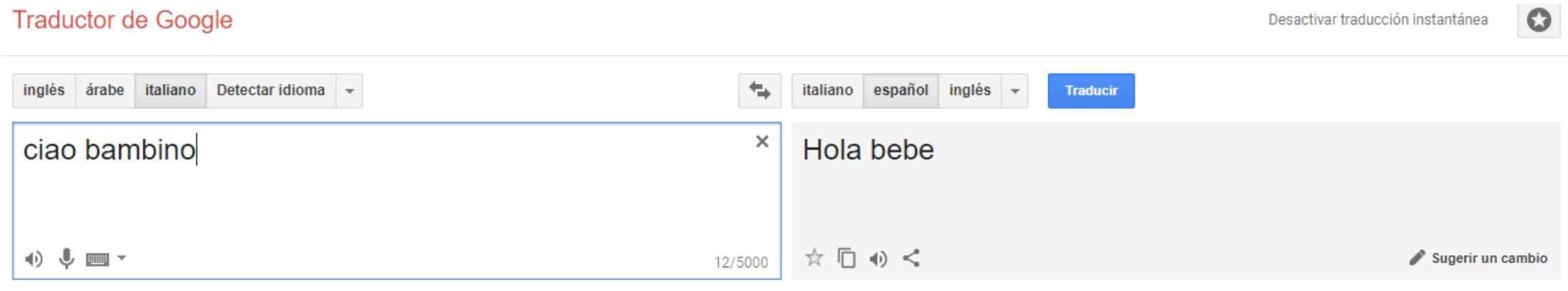
Traductor de Google

Desactivar traducción instantánea 

inglés árabe italiano Detectar idioma ▾  italiano español inglés ▾ 

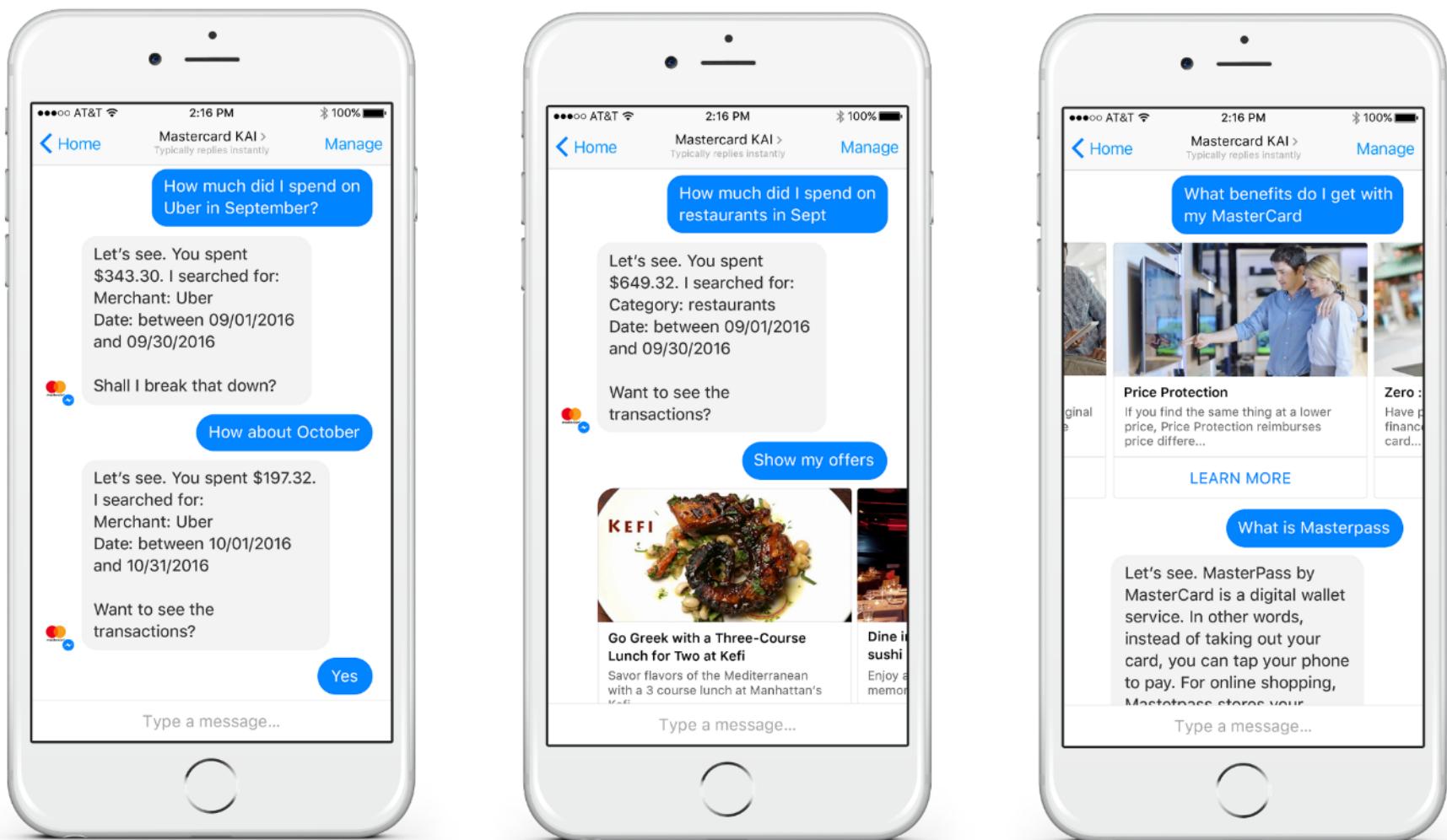
ciao bambino  12/5000         Sugerir un cambio

Hola bebe



Ver también
bambino

Ejemplos de aplicaciones web que usan text mining



Ejemplos de aplicaciones web que usan text mining

12

in:spam

□ C Más

Offerta automotive	Scopri la gamma Jeep: richiedi il tuo preventivo gratuito.
Anobii	I 20 libri più letti e tante statistiche su Anobii
Generali	Guida con stile e risparmia sulla polizza auto
Italiamac Forum	Aggiornamento privacy policy e termini di utilizzo di Italiamac
Adele Ninja Academy	Impara in 60 ore tutto ciò che devi sapere sul Digital Marketing!
Encuesta Bancolombia	Encuesta Bancolombia
Dyckman Green Chief	New calls
Adele Ninja Academy	Il prezzo del Master eCommerce sta per aumentare: corri ad iscriverti!
Investigación Industrial	Ofrezca sus comentarios y participe para ganar un premio en efectivo
Prestito Personale	Con il cambio rata Findomestic il Prestito e' piu' sereno
Adele Ninja Academy	Accedi al Social Media LIVE Program: candidature aperte fino a mezzanotte!
Magnum Photos	This week on Magnum Photos: 1968's Power, Protest and Politics, Susan Meiselas on Femal...
Mirko Pallera Ninja Ac.	Partecipa con noi alla prossima avventura dei Ninja
Rajeev Adrian<>	Re: Investment Capital Without Tears.....
Italiamac Forum	TrenDevice Days: 6 giorni di grandi sconti, mai visti prima.
Mike Baillie - Avaaz	Fermiamo le navi del massacro di animali
Nital Informa	Stupisci la tua mamma con la tecnologia by Nital
Lanieri via Offertaperte	Crea la tua camicia su misura online
Luca La Mesa Ninja Aca.	Riaprono le candidature al Social Media LIVE Program

Ejemplos de aplicaciones web que usan text mining





INFORMATION EXTRACTION

Son técnicas que permiten extraer información específica desde grandes cantidades de datos textuales

Named entity recognition



INFORMATION RETRIEVAL

Es la actividad de obtener información buscando dentro de colecciones de documentos digitales

Motores de búsqueda



NATURAL LANGUAGE PROCESSING

Se ocupa de las interacciones entre las computadoras y los lenguajes humanos

Chatbot
Text prediction



DATA MINING

Se centra en extraer patrones de datos de textos con la finalidad de describir o prever

Opinion mining
Document classification

1

DATA EXTRACTION



2

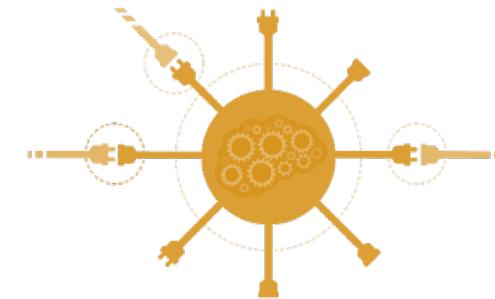
TEXT MINING





WEB SCRAPING

Es un método que nos permite utilizar el código fuente de páginas web para tomar datos y organizarlos en base de datos



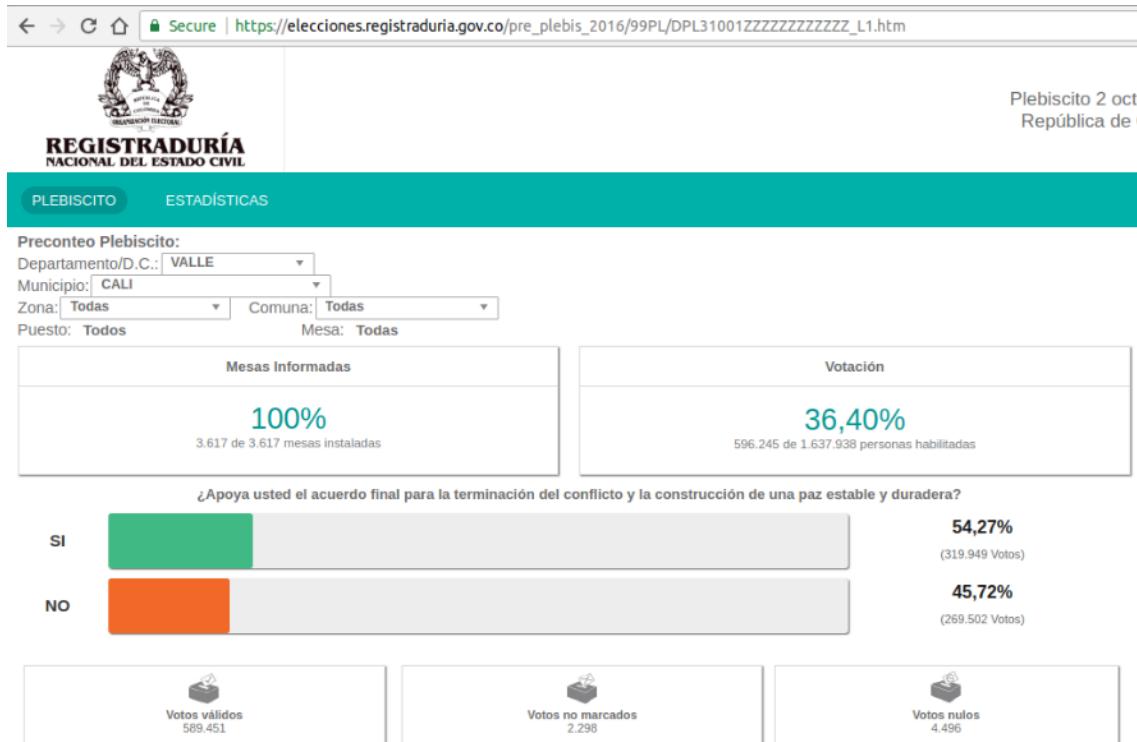
API

Application Programming Interface son un conjunto de reglas y especificaciones que las aplicaciones pueden seguir para comunicarse entre ellas

Web scraping



17



	A	B	C	D	E	F	G	H	I	J
1	cod_mpio	name_mpio	cod_dep	name_dep	vote.si	vote.no	vote.nm	vote.nl	vote.el	
2	01001	MEDELLIN	01	ANTIOQUIA	253.548	431.173	3.309	6.754	1.508.418	
3	01004	ABEJORRAL	01	ANTIOQUIA	1.296	2.105	194	52	15.274	
4	01007	ABRIAQUI	01	ANTIOQUIA	213	332	14	13	1.938	
5	01010	ALEJANDRIA	01	ANTIOQUIA	966	696	51	22	4.577	
6	01013	AMAGA	01	ANTIOQUIA	2.084	4.139	53	94	22.028	
7	01016	AMALFI	01	ANTIOQUIA	1.310	2.304	132	72	16.928	
8	01019	ANDES	01	ANTIOQUIA	3.193	5.729	78	93	34.842	
9	01022	ANGELOPOLIS	01	ANTIOQUIA	553	1.004	9	23	5.679	
10	01025	ANGOSTURA	01	ANTIOQUIA	705	1.089	57	84	9.154	
11	01028	ANORI	01	ANTIOQUIA	1.764	953	227	40	11.144	
12	01031	ANTIOQUIA	01	ANTIOQUIA	2.041	3.300	52	102	19.732	
13	01034	ANZA	01	ANTIOQUIA	363	740	31	16	5.715	
14	01035	APARTADO	01	ANTIOQUIA	14.195	12.936	352	478	89.196	
15	01037	ARBOLETES	01	ANTIOQUIA	3.691	1.937	181	57	21.448	
16	01039	ARGELIA	01	ANTIOQUIA	650	1.351	41	23	6.627	
17	01040	ARMENIA	01	ANTIOQUIA	301	676	23	26	4.844	
18	01043	BARBOSA	01	ANTIOQUIA	4.200	8.666	287	277	38.541	
19	01046	BELMIRA	01	ANTIOQUIA	687	897	16	17	4.807	
20	01049	BELLO	01	ANTIOQUIA	43.642	76.122	509	1.577	318.411	
21	01052	BETANIA	01	ANTIOQUIA	466	1.110	11	22	8.372	
22	01055	BETULIA	01	ANTIOQUIA	1.005	1.919	50	49	13.377	
23	01058	BOLIVAR	01	ANTIOQUIA	1.644	3.875	57	60	24.029	
24	01061	BURITICA	01	ANTIOQUIA	474	890	22	20	6.829	
25	01062	BRICEÑO	01	ANTIOQUIA	1.367	604	62	45	6.169	
26	01064	CACERES	01	ANTIOQUIA	2.506	1.358	57	83	19.788	
27	01067	CAICEDO	01	ANTIOQUIA	593	921	43	43	6.091	
28	01070	CALDAS	01	ANTIOQUIA	9.292	16.149	123	397	60.580	
29	01073	CAMPAMENTO	01	ANTIOQUIA	619	604	32	41	6.360	
30	01076	CAÑASGORDAS	01	ANTIOQUIA	1.911	2.739	113	121	13.402	
31	01078	CARACOLI	01	ANTIOQUIA	578	593	16	12	4.856	
32	01079	CARAMANTA	01	ANTIOQUIA	568	790	33	26	5.231	
33	01080	CAREPA	01	ANTIOQUIA	5.327	6.057	159	190	41.461	
34	01082	CARMEN DE VIBORAL	01	ANTIOQUIA	5.528	10.787	163	272	33.070	
35	01085	CAROLINA	01	ANTIOQUIA	464	641	22	23	4.081	
36	01088	CAUCASIA	01	ANTIOQUIA	8.801	11.495	247	368	66.133	
37	01091	CISNEROS	01	ANTIOQUIA	1.084	1.837	33	63	11.434	

¿Que es un API?

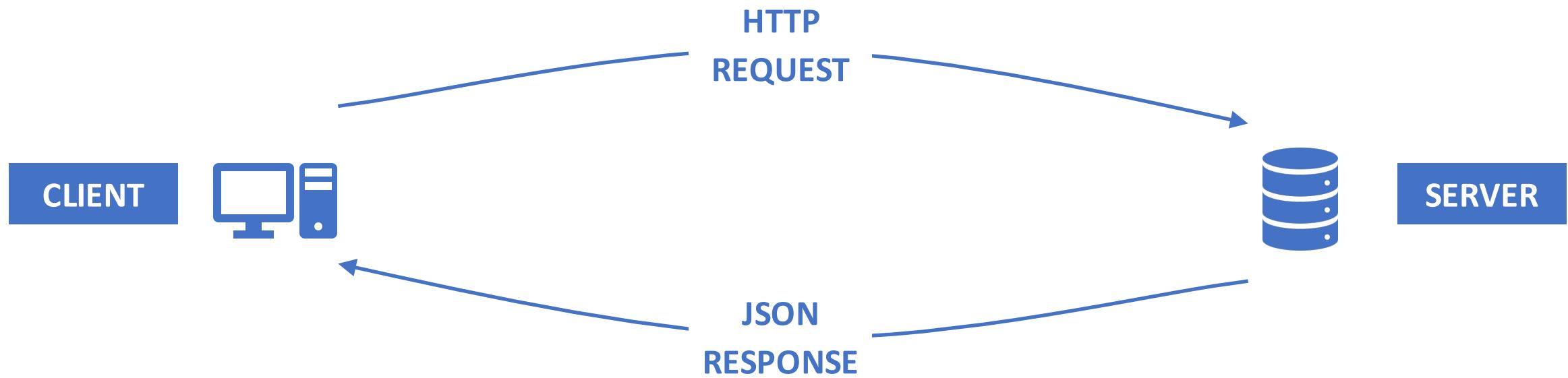
[Enlace](#)

¿Como funciona un API REST?

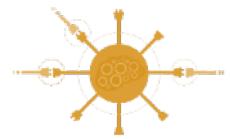


19

Un API es una posibilidad que nos ofrece una plataforma para poder dialogar con sus servidores y obtener servicios a cambio
La mayoría de plataformas de redes sociales proveen API de tipo REST
(REpresentational State Transfer)



Request -> response



20

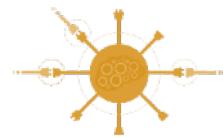
RESPONSE (JSON)

REQUEST (HTTP)

<http://maps.googleapis.com/maps/api/geocode/json?address=cali&sensor=false>

```
{
  "results" : [
    {
      "address_components" : [
        {
          "long_name" : "Cali",
          "short_name" : "Cali",
          "types" : [ "locality", "political" ]
        },
        {
          "long_name" : "Cali",
          "short_name" : "Cali",
          "types" : [ "administrative_area_level_2", "political" ]
        },
        {
          "long_name" : "Valle del Cauca",
          "short_name" : "Valle del Cauca",
          "types" : [ "administrative_area_level_1", "political" ]
        },
        {
          "long_name" : "Colombia",
          "short_name" : "CO",
          "types" : [ "country", "political" ]
        }
      ],
      "formatted_address" : "Cali, Valle del Cauca, Colombia",
      "geometry" : {
        "bounds" : {
          "northeast" : {
            "lat" : 3.505554,
            "lng" : -76.4608248999999
          },
          "southwest" : {
            "lat" : 3.2845748,
            "lng" : -76.590503
          }
        },
        "location" : {
          "lat" : 3.4516467,
          "lng" : -76.5319854
        },
        "location_type" : "APPROXIMATE",
        "viewport" : {
          "northeast" : {
            "lat" : 3.505554,
            "lng" : -76.4608248999999
          },
          "southwest" : {
            "lat" : 3.2845748,
            "lng" : -76.590503
          }
        }
      },
      "place_id" : "ChIJ8bNLzPCmMI4RaGGUum1Dx8",
      "types" : [ "locality", "political" ]
    }
  ],
  "status" : "OK"
}
```





<https://developer.twitter.com/en.html>

Sito desarrolladores de Twitter

<https://developer.twitter.com/en/docs>

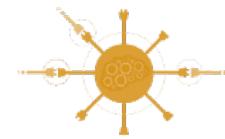
Documentación sobre APIs

<https://developer.twitter.com/en/docs/tweets/search/api-reference/get-search-tweets.html>

Documentación sobre Search API

<https://developer.twitter.com/en/docs/api-reference-index>

Documentación sobre la mayoría de APIs



1



AUTHENTICATION

Tenemos que tener unas credenciales de acceso.

Esa se crean teniendo una cuenta de Twitter y creando una app.

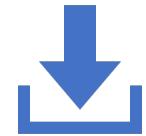
2



REQUEST

Tenemos que formular una **query** escogiendo el **endpoint** y los **parámetros** adecuados. Para poderla hacer de forma ágil vamos a utilizar un paquete de R apto para esto

3



RESPONSE

Hay que recibir el archivo en formato json y organizarlo en un formato que se pueda analizar fácilmente. La cantidad y tiempos de descarga depende del **rate limit**. Cada endpoint tiene uno propio. Siempre lo haremos mediante un paquete de R



<https://apps.twitter.com>

Create an application

1



AUTHENTICATION

Application Details

Name *

Your application name. This is used to attribute the source of a tweet and in user-facing authorization screens. 32 characters max.

Description *

Your application description, which will be shown in user-facing authorization screens. Between 10 and 200 characters max.

Website *

Your application's publicly accessible home page, where users can go to download, make use of, or find out more information about your application. This fully-qualified URL is used in the source attribution for tweets created by your application and will be shown in user-facing authorization screens.
(If you don't have a URL yet, just put a placeholder here but remember to change it later.)

Callback URLs

Where should we return after successfully authenticating? OAuth 1.0a applications must explicitly specify their oauth_callback URL(s) here, as well as include the one of the URLs below in the request token step. To restrict your application from using callbacks, leave this field blank.

Add a Callback URL

Developer Agreement

Yes, I have read and agree to the [Twitter Developer Agreement](#).

Create your Twitter application



1



AUTHENTICATION

TWOMR ←

Details Settings **Keys and Access Tokens** Permissions

Application Settings

Keep the "Consumer Secret" a secret. This key should never be human-readable in your application.

Consumer Key (API Key)	F5 [REDACTED] Yo
Consumer Secret (API Secret)	9wPRE [REDACTED] Kb5zmdx
Access Level	Read and write (modify app permissions)
Owner	cocometrico
Owner ID	84 [REDACTED] 52

← ← ←

Application Actions

Regenerate Consumer Key and Secret Change App Permissions

Your Access Token

You haven't authorized this application for your own account yet.

By creating your access token here, you will have everything you need to make API calls right away. The access token generated will be assigned your application's current permission level.

Token Actions

Create my access token

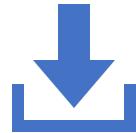


2

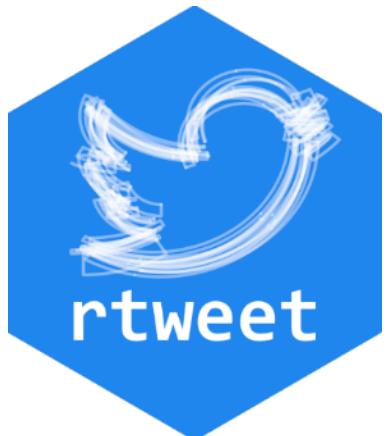


REQUEST

3



RESPONSE



Para hacer estas dos operaciones vamos a usar el paquete de R “rtweet”

<http://rtweet.info/>

<https://github.com/mkearney/rtweet>

1

DATA EXTRACTION



2

TEXT MINING





1

DB MANIPULATION



2

STRING
MANIPULATION



1

DOCUMENT TERM
MATRIX



2

DATA ANALYSIS





El paquete “data.table” fue escrito en el 2008 y propone una forma avanzada de data frame mas flexible, rápida y con menos impacto en cuanto a uso de la RAM

DT[i, j, by]

i <=> where

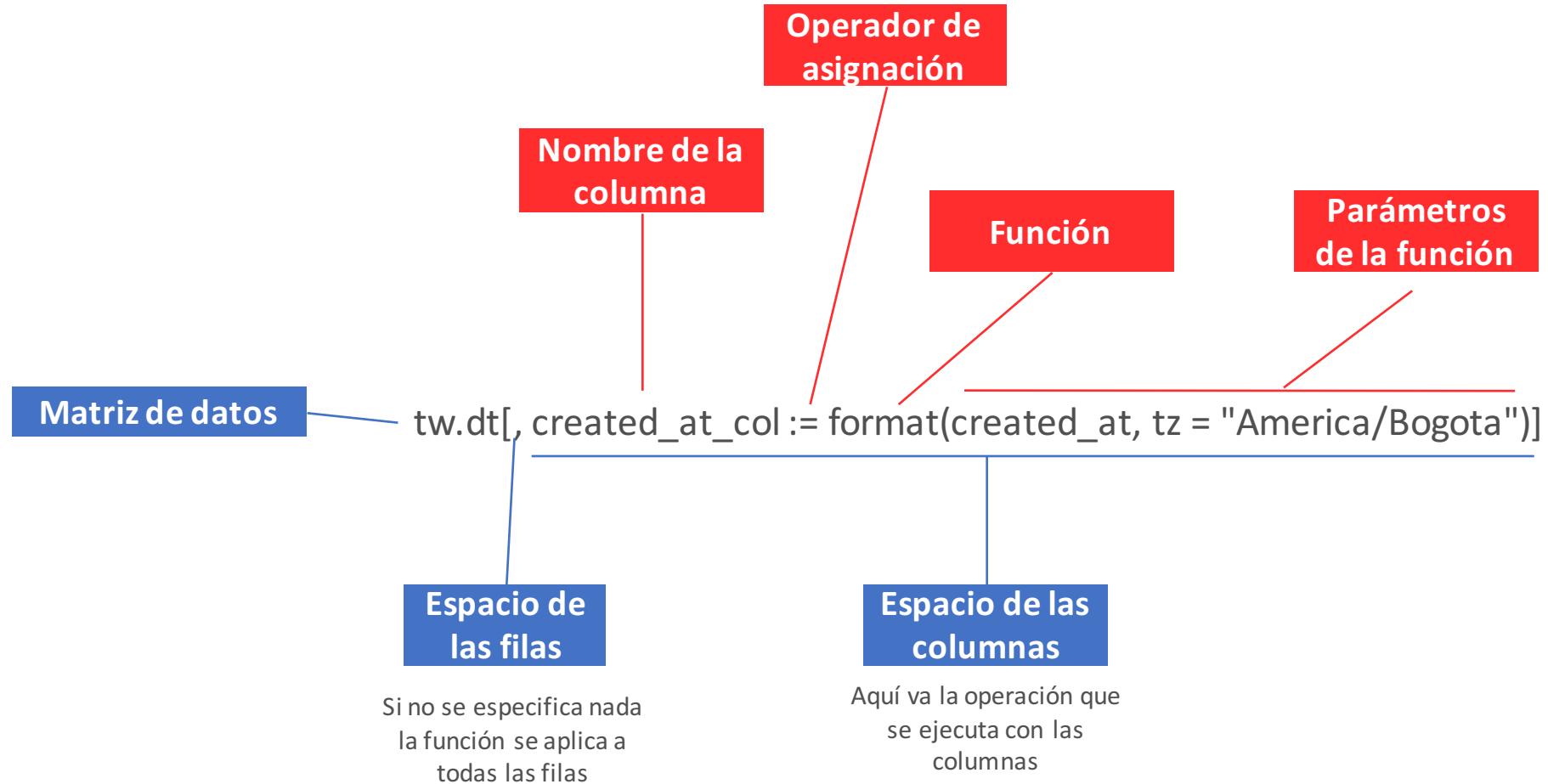
En este espacio va el indicador de la fila o la condición que hay que aplicar a las filas

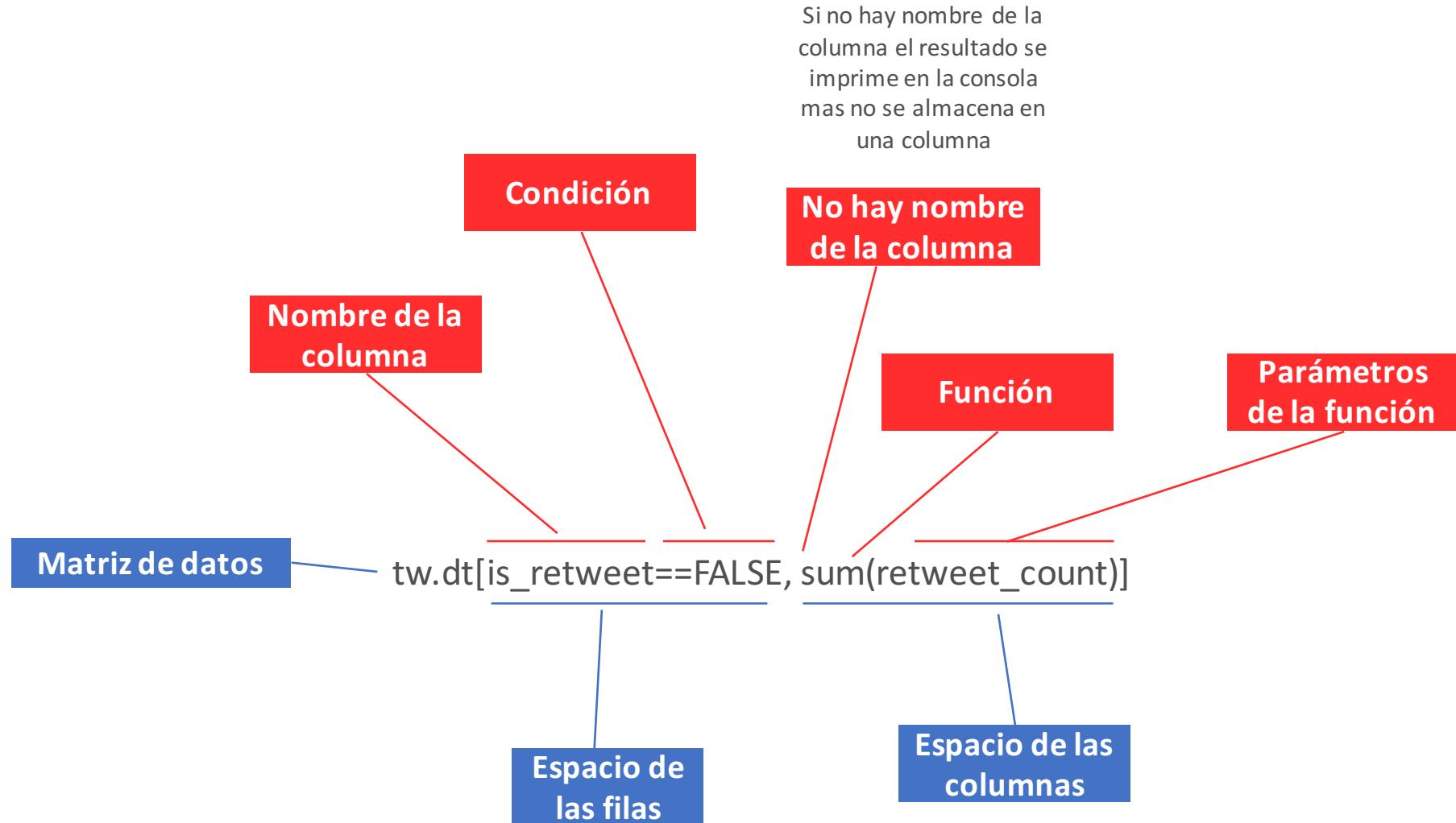
j <=> select

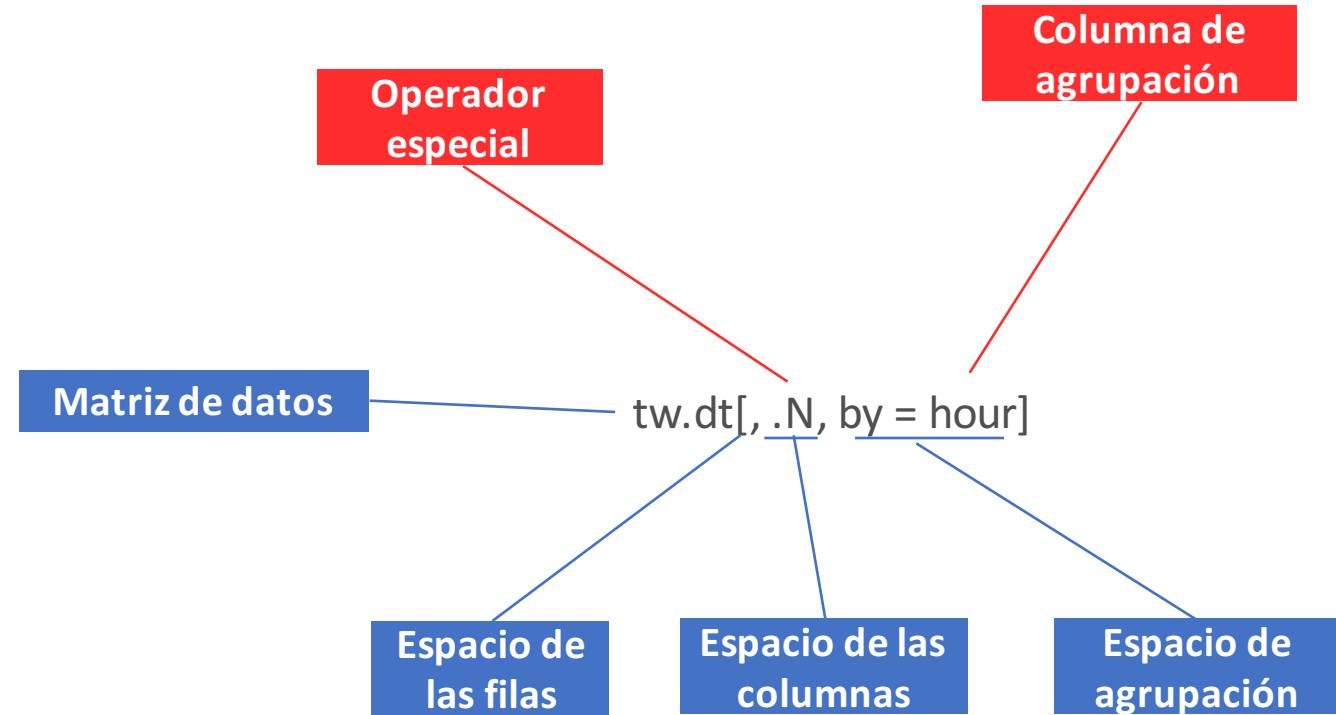
En este espacio va el indicador de la columna o la condición que hay que aplicar a la columnas

by <=> group_by

En este espacio va la eventual variable según la cual efectuar la agrupación









```
tw.dt[, .N, by = hour]
```

El resultado aparece en la consola

```
tw.dt[, .N, by = hour]
```

El resultado aparece en la consola

```
tweet_x_horas <- tw.dt[, .N, by = hour]
```

El resultado no aparece en la consola pero
queda asignado a otro objeto



- 1 Número de tweets
- 2 Número de usuarios
- 3 Número de trinos por día
- 4 Número de trinos por horas
- 5 Los tres trinos más virales (retweets + favoritos)
- 6 Los cinco usuarios más activos (los que más estuvieron trinando)



Las **EXPRESIONES REGULARES** son una serie de reglas y operadores que nos permiten volver extremamente flexible y eficiente las búsquedas y sustituciones de cadenas de texto



(a b)	'a' o 'b'
^	comienzo de la frase
\$	final de la frase
?	0 o 1 ocurrencias de la letra o conjunto de letras anteriores
+	1 o mas ocurrencias
[ab]	'a' o 'b' es otra forma de escribir (a b)
[a-z]	todas las letras minúsculas
[0-9]	todos los números
[a-z0-9]	todas las letras minúsculas y los números



1

En cuantos trinos aparece la palabra ...

hidroituango

presa

epm

rusia2018

colombia

alemania

registraduria

fraude

registrador

netflixlat

stranger things

sense8

avianca

retraso

vuelo

senalcolombia

ley 174

giroxseñal

BAG OF WORDS

	Palabra 1	Palabra 2	Palabra 3	Palabra 4	Palabra 5	Palabra 6
Documento 1	x	x		x		
Documento 2		x	x		x	
Documento 3		x				x
Documento 4	x		x	x		x
Documento 5	x	x	x	x	x	
Documento 6		x	x			x
Documento 7	x		x		x	



Es el proceso de romper una cadena de texto en tokens, es decir palabras, frases u otros elementos textuales dotados de sentido



Héctor Bega
@Guyogie

Seguir



#ReporteLluvia Primero se fue la luz y sobre las 12:16h llueve sobre el Sur fuerte con rayos en la Urb. Universal II Etapa.



Es el proceso de romper una cadena de texto en tokens, es decir palabras, frases u otros elementos textuales dotados de sentido

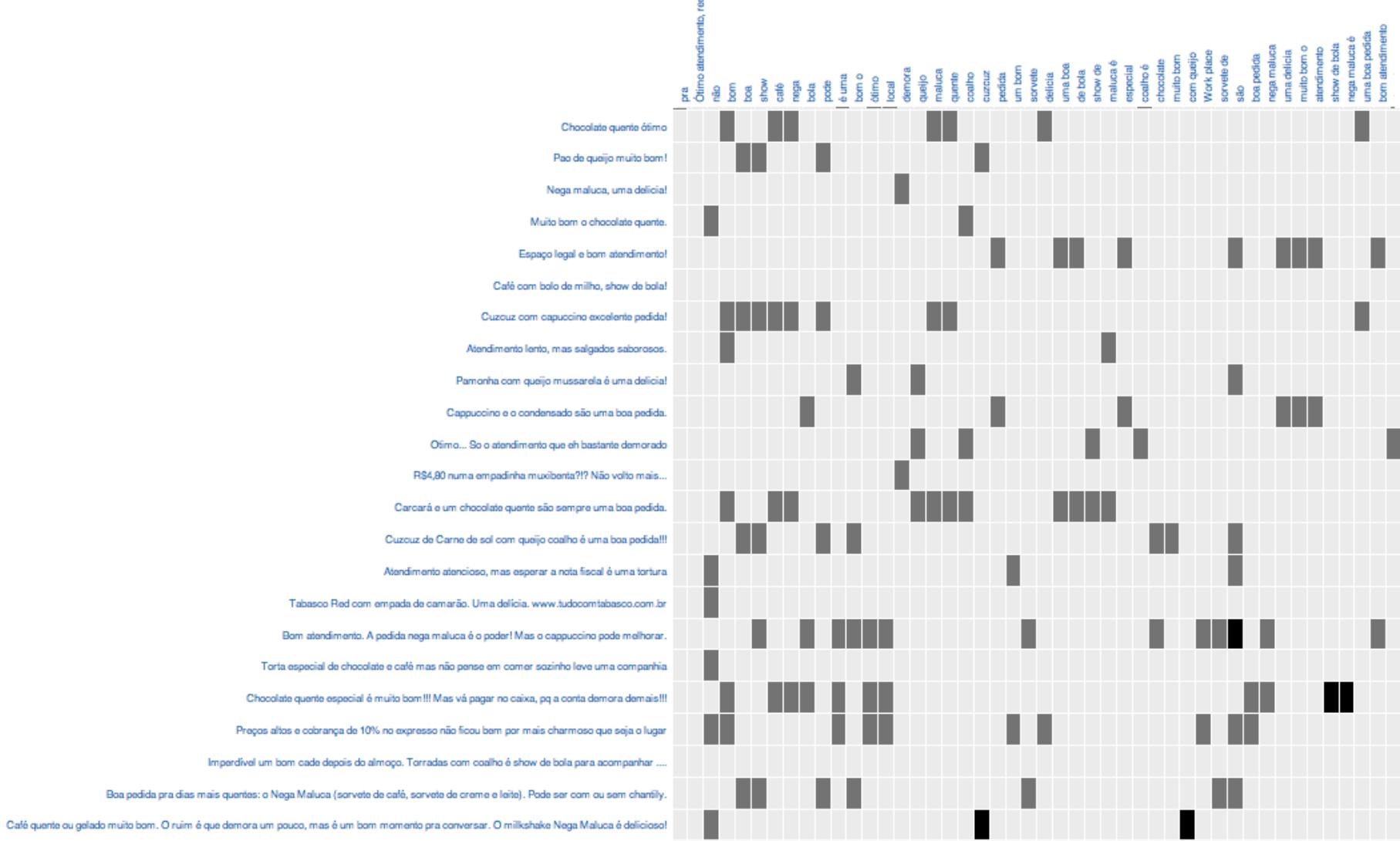


Héctor Bega
@Guyogie

Seguir

#ReporteLluvia|Primero|se|fue|la|luz|y|sobre|
las|12:16h|llueve|sobre|el|Sur|fuerte|con|
rayos|en|la|Urb.|Universal||Etapa.

Sparse matrix issues





Stop words o palabras vacías, son palabras que no añaden valor agregado a una frase. Normalmente son las palabras mas comunes dentro un determinado idioma.



Héctor Bega
@Guyogie

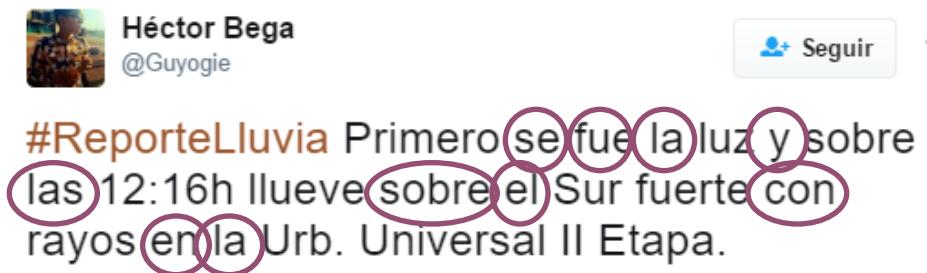
 Seguir



#ReporteLluvia Primero se fue la luz y sobre las 12:16h llueve sobre el Sur fuerte con rayos en la Urb. Universal II Etapa.



Stop words o palabras vacías, son palabras que no añaden valor agregado a una frase. Normalmente son las palabras mas comunes dentro un determinado idioma.



A screenshot of a Twitter post from user Héctor Bega (@Guyogie). The post contains the following text:
#ReporteLluvia Primero se fue la luz y sobre
las 12:16h llueve sobre el Sur fuerte con
rayos en la Urb. Universal II Etapa.

The word "se" is highlighted with a purple oval, and the words "la", "y", "sobre", "el", and "con" are also highlighted with purple ovals, demonstrating common stop words in Spanish.



Son una serie de operaciones por medio de las cuales normalizamos algunos grafemas o palabras para poderlos tratar mas fácilmente.



Abreviaciones y acrónimos

Números

Jergas y formas dialectales

White spaces

Errores ortográficos

Caracteres especiales

Puntuación

Grafemas especiales

Mayúsculas/minúsculas



Son una serie de operaciones por medio de las cuales normalizamos algunos grafemas o palabras para poderlos tratar mas fácilmente.



Héctor Bega
@Guyogie

Seguir ▾

#ReporteLluvia Primero se fue la luz y sobre las 12:16h llueve sobre el Sur fuerte con rayos en la Urb. Universal II Etapa.

Horas

Segunda

Urbanización



Es el proceso por medio del cual se infieren las raíces morfológicas de las palabras, generalmente reduciendo las flexiones verbales (conjugación) y nominales (genero, número)



Primero	→	Primer
Primera	→	Primer
Primeros	→	Primer
Primeras	→	Primer

1



Seleccionamos los tweets que queremos analizar

2



Creo el corpus

3



Tokenization + normalización de textos

4



Stopwords

a) Tokenization

Remover
puntuación,
caracteres
separadores,
números,
símbolos,

b)

5



Stemming

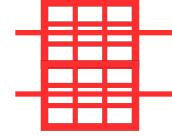
Encontrar las
raíces de las
palabras

6



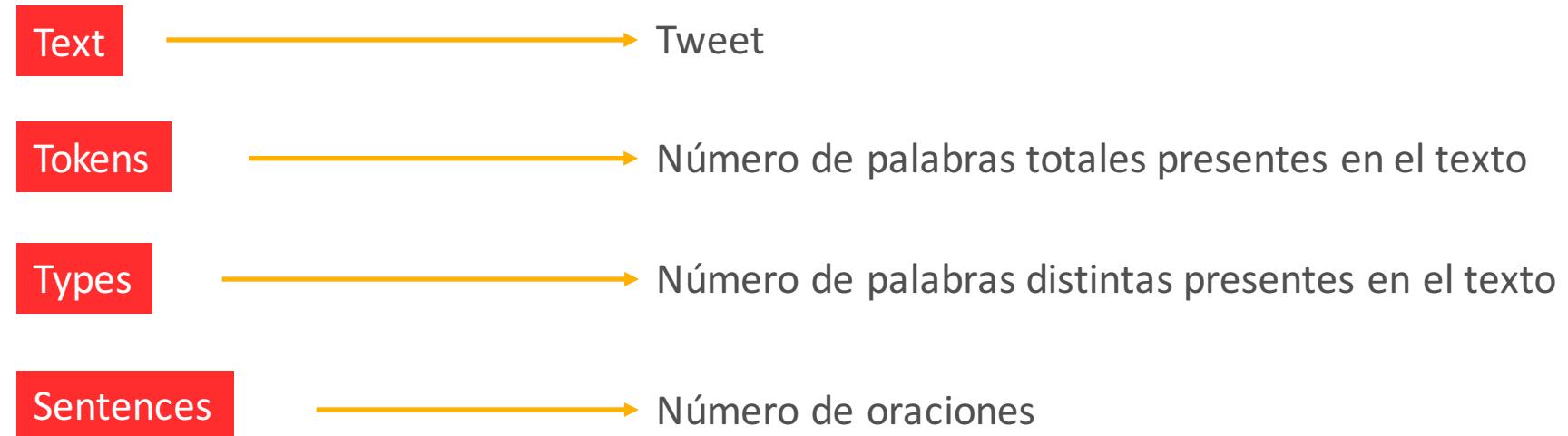
Eliminamos stems
cortos y nos
quedamos con los
stems mas
frecuentes

7



Eliminamos los
documentos
(tweets) que no
tienen relación
con ningún stem

head(corpus)





1

Crean la document term matrix con los pasos básicos (1,2,3,4)

2

Encuentren las 10 palabras mas frecuentes

3

Encuentren las correlaciones entre palabras frecuentes



1

Encuentren los 10 hashtags mas frecuentes

2

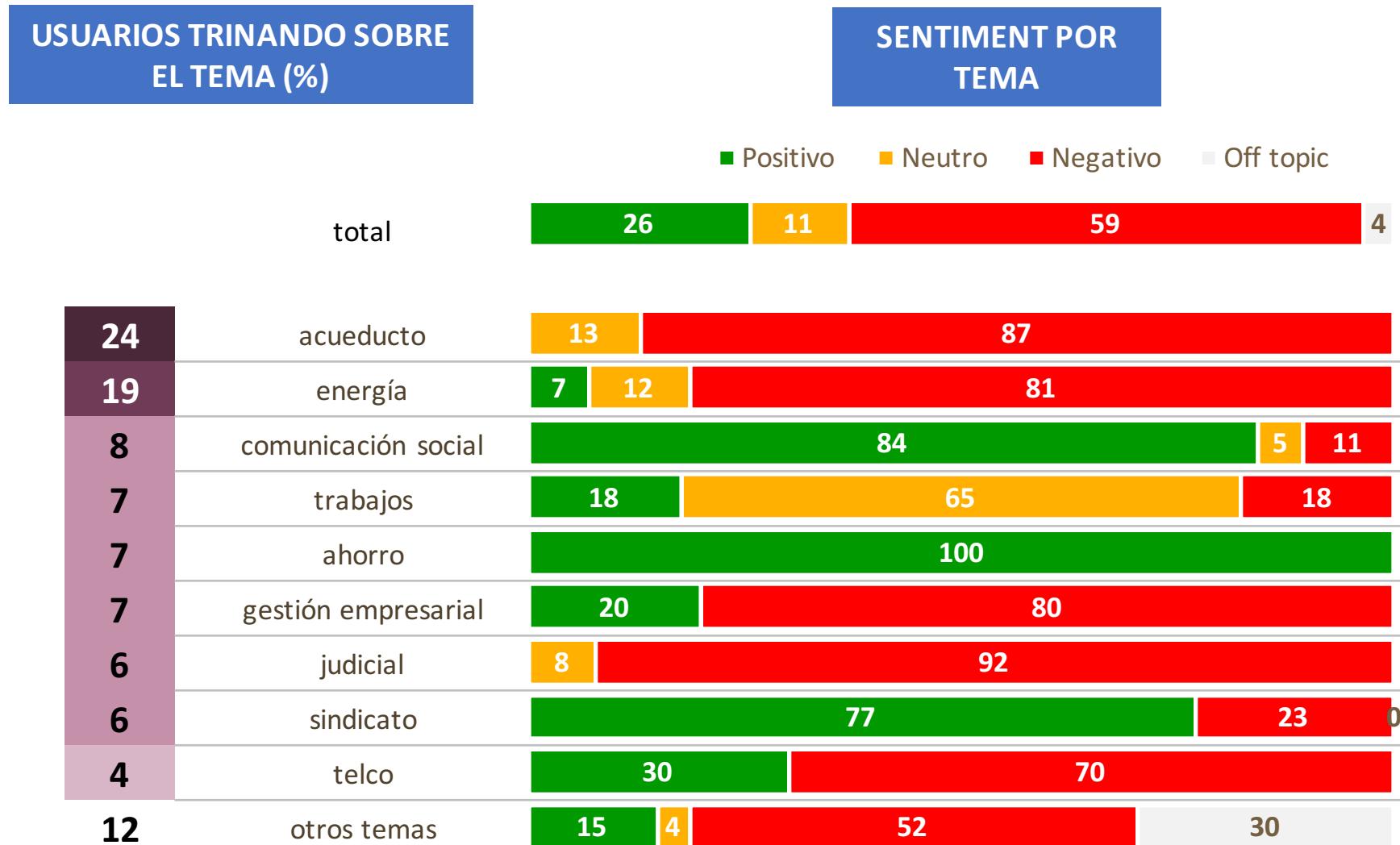
Encuentren las 10 cuentas mas mencionadas

3

Encuentren las 5 secuencias de dos palabras mas frecuentes

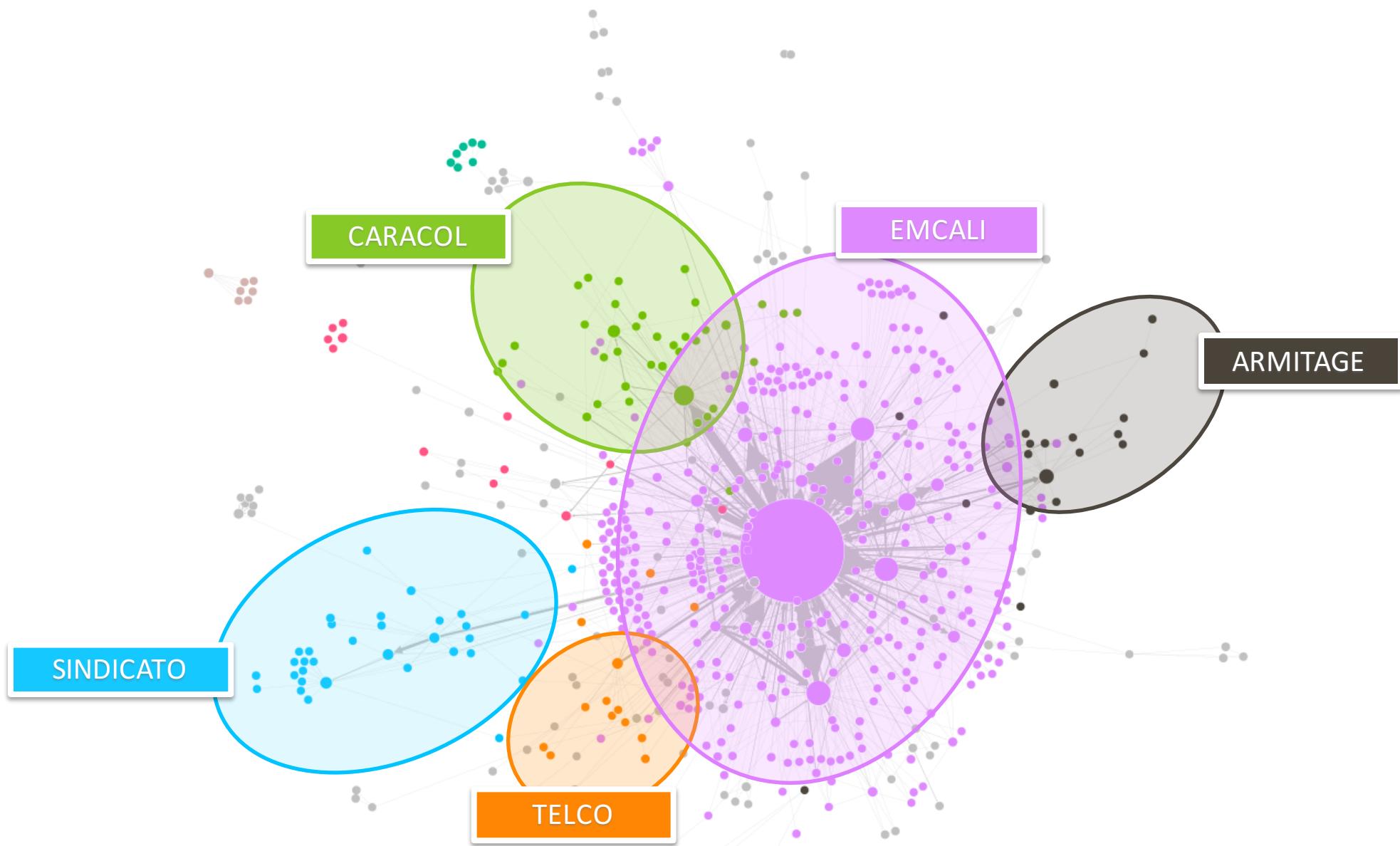
4

Wordcloud

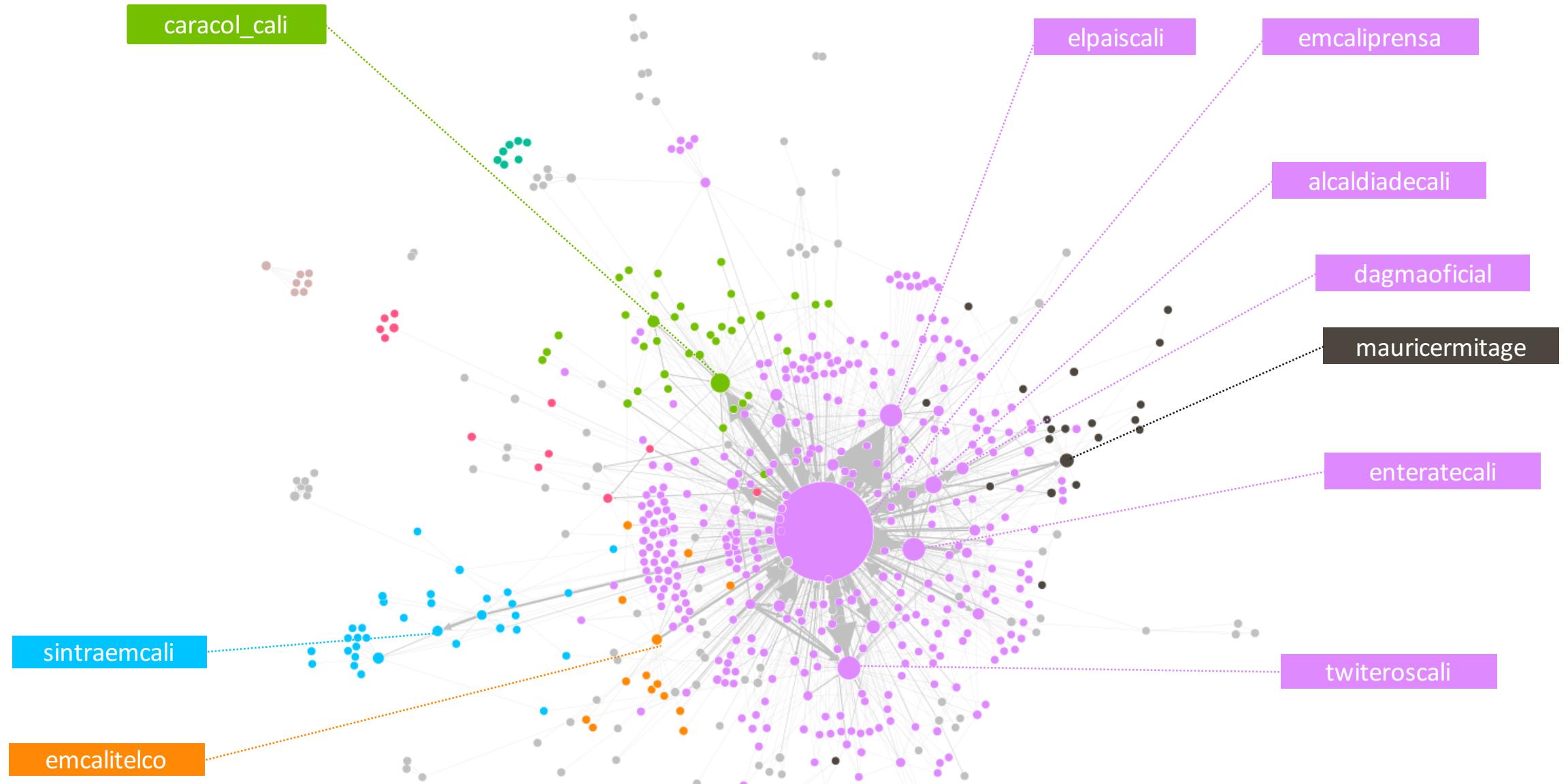


Ailment	Allergies	Depression	Aches/Pains	Cancer	Obesity	Flu	Dental
Prior Frequency	Allergies 6.4%	Anxiety 5.8%	Back Pain 10.8%	Breast Cancer 8.0%	Diabetes 2.3%	Flu 8.1%	Oral Health 4.6%
General Words	allergies stop eyes allergic	help dont body depression	body head need hurts	cancer pray mom shes	blood doctor high meds	flu “swine flu” “flu shot” dont	meds killers dentist teeth
Symptoms	sneezing cold coughing	pain anxiety stomach	pain aches stomach	pain sad “breast cancer”	pressure “high blood pressure”	fever cold “sore throat”	pain toothache sore
Treatments	medicine benadryl claritin	surgery treatment plastic	massage “hot bath” ibuprofen	surgery hospital “heart surgery”	hospital diet exercise	hospital vaccine medicine	braces surgery antibiotics

Word	#	Ent.	Most Common Ailments
Pain Relief Medication			
tylenol	1807	1.57	HA (39%), IN (30%), Cold (9%)
ibuprofen	1125	1.54	HA (37%), DN (21%), Aches (17%)
advil	1093	1.08	HA (61%), Cold (6%), DN (5%)
aspirin	885	1.04	HA (69%), IN (10%), Aches (10%)
vicodin	505	1.33	DN (61%), Injuries (11%), HA (10%)
codeine	406	1.94	Cold (25%), DN (19%), HA (17%)
morphine	206	1.17	DN (59%), Infection (22%), Aches (9%)
aleve	183	1.10	HA (62%), IN (15%), DN (14%)
Allergy Medication			
benadryl	871	1.24	Allergies (64%), Skin (13%), IN (12%)
claritin	417	0.54	Allergies (88%), HA (5%)
zyrtec	386	0.49	Allergies (90%)
sudafed	298	1.61	Allergies (39%), Cold (21%), HA (20%)

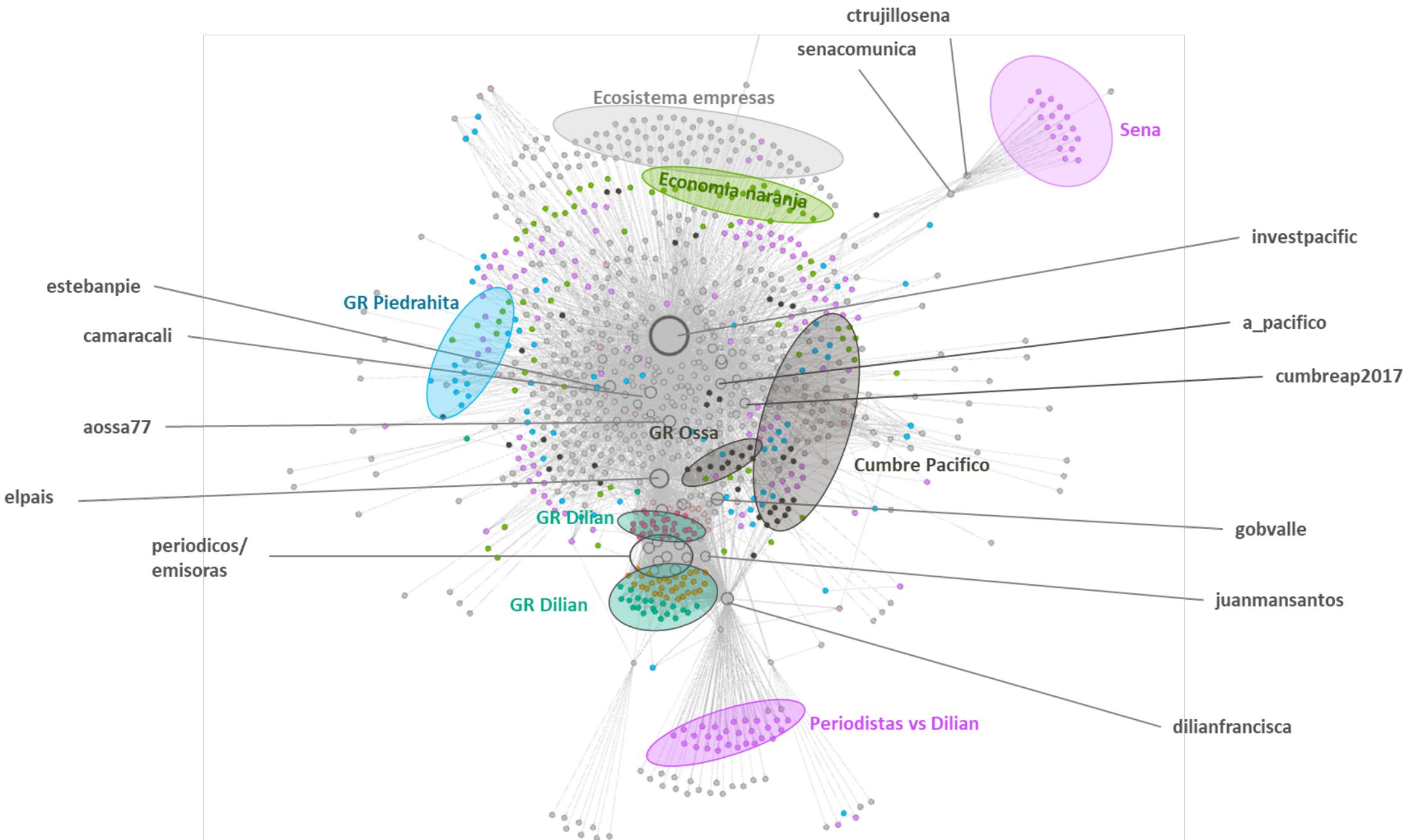


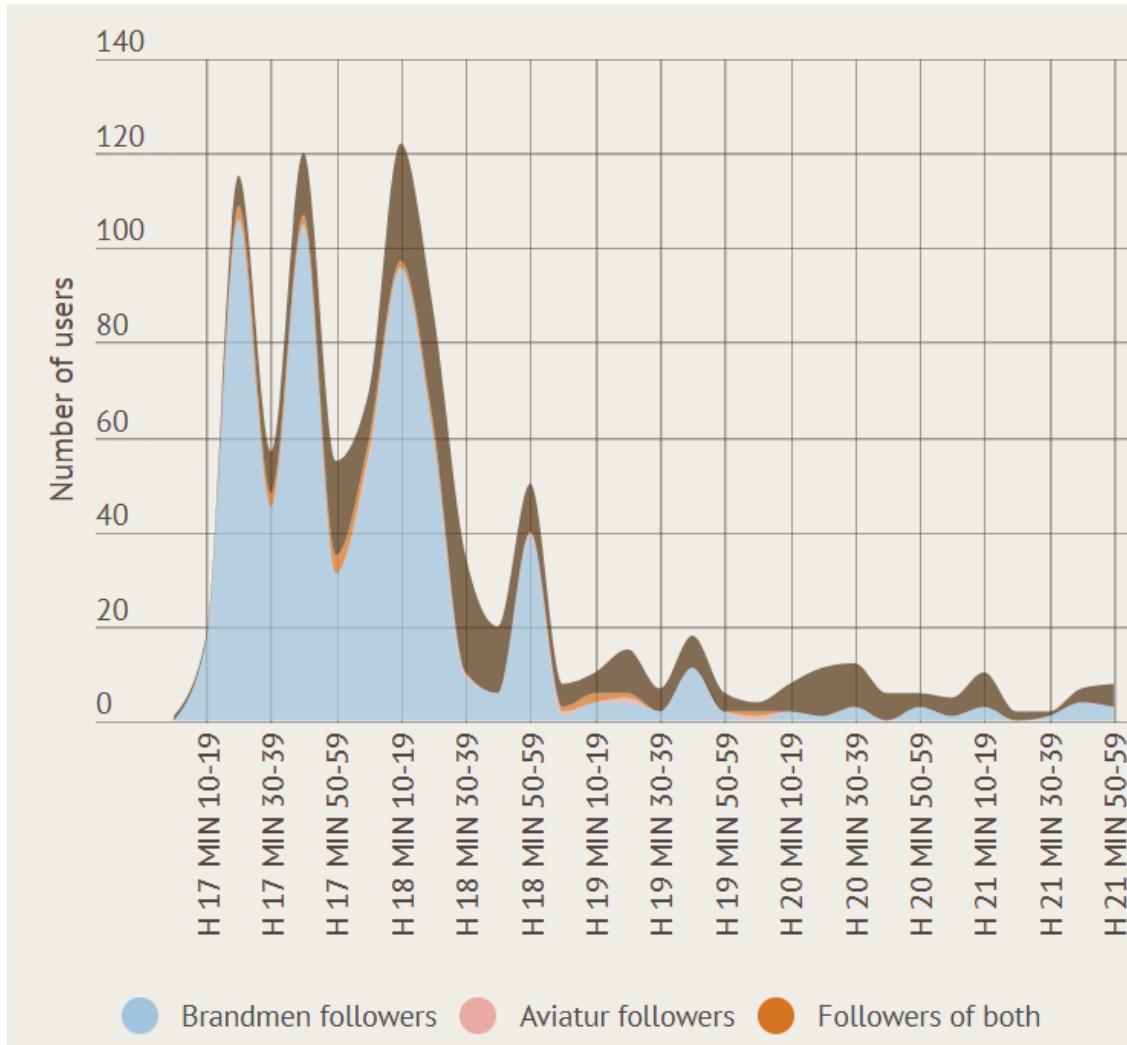
Ejemplos de uso de data mining y datos de redes sociales



Ejemplos de uso de data mining y datos de redes sociales

56

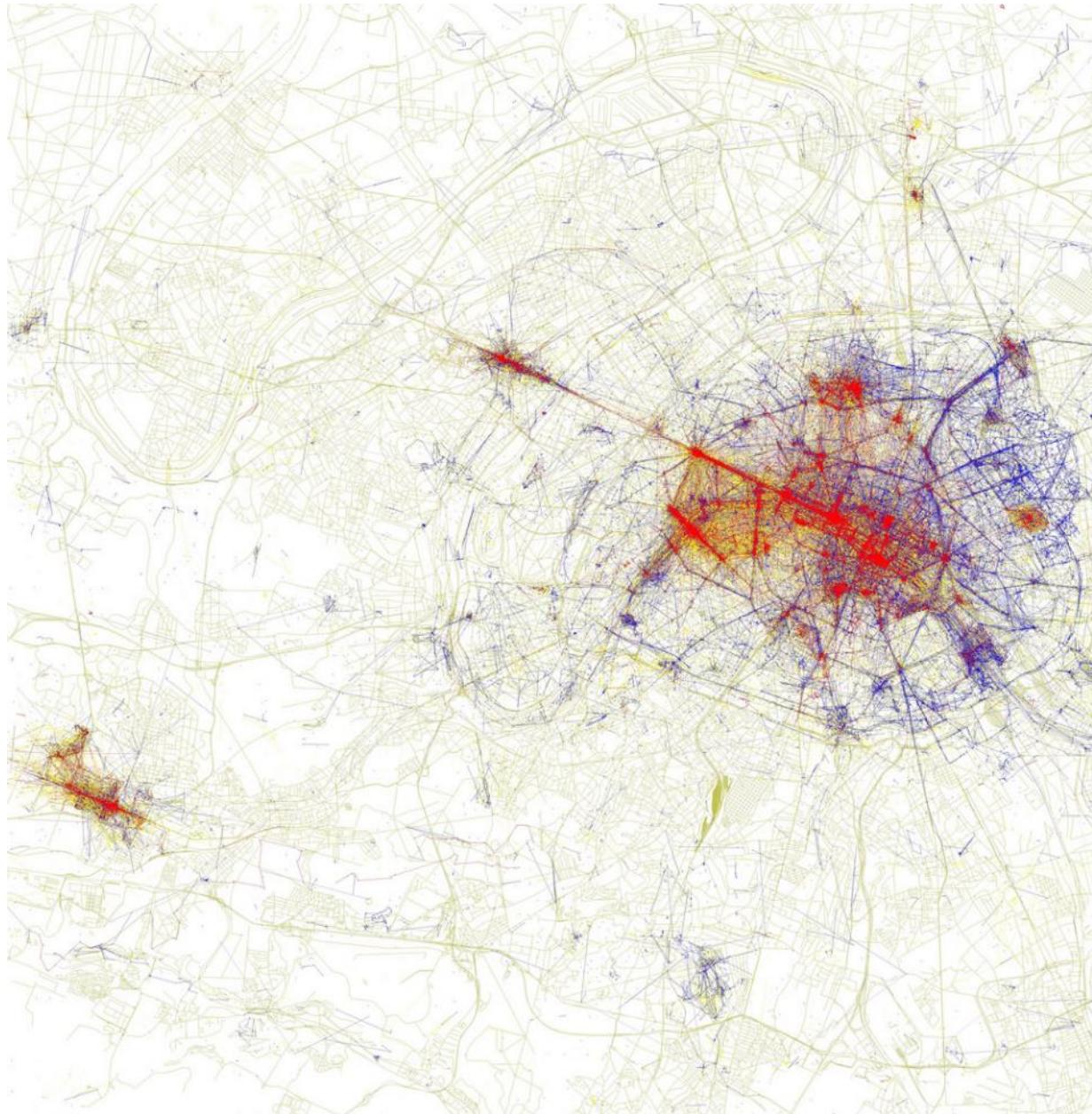




<https://www.youtube.com/watch?v=au1h5v6Xsbg>

Ejemplos de uso de data mining y datos de redes sociales

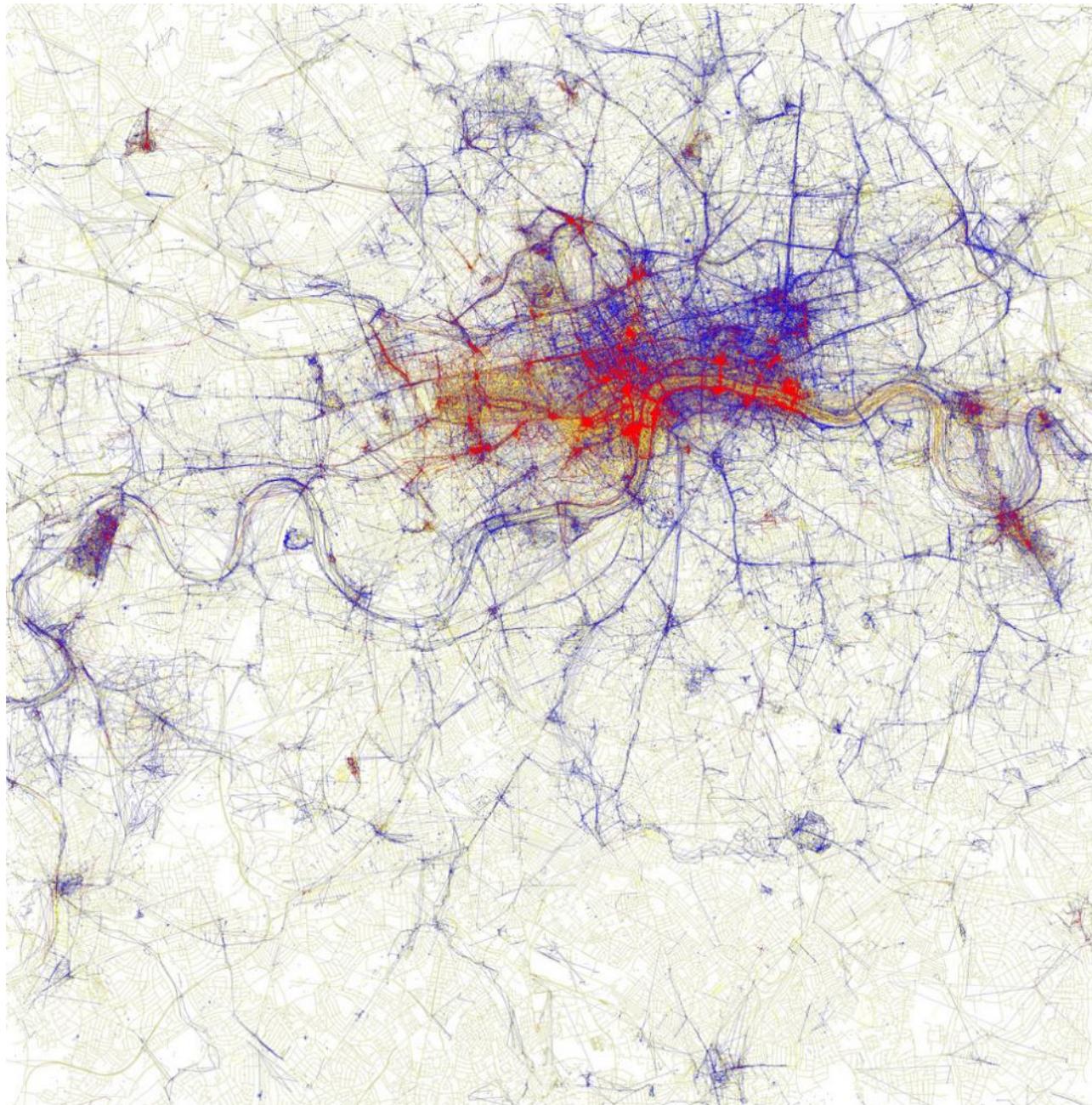
PARIS



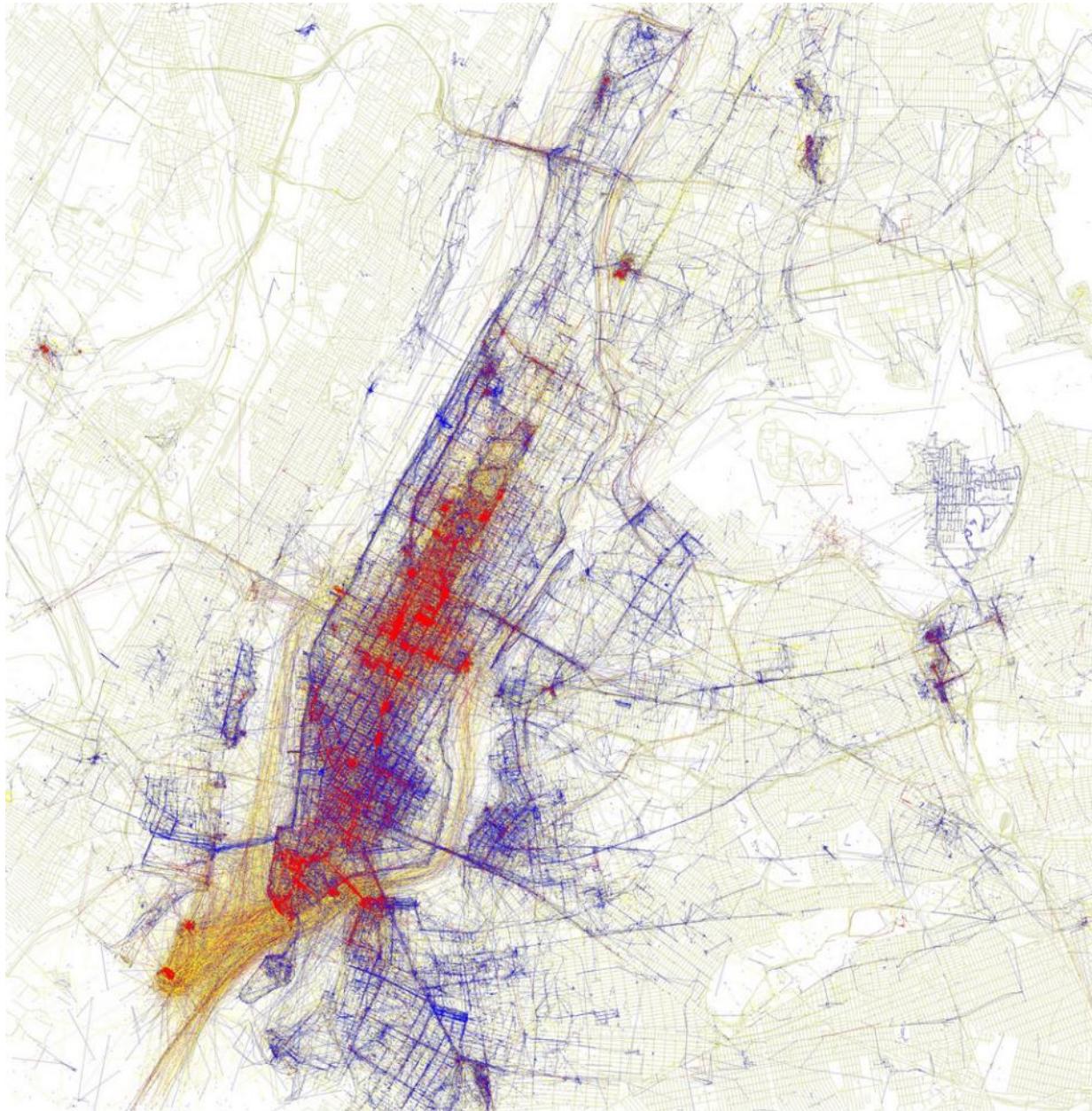
Ejemplos de uso de data mining y datos de redes sociales

59

LONDRES



NEW YORK



Hinchadas colombianas

[Enlace](#)

Final liga aguila Deportivo Cali - DIM

[Enlace](#)

Final UCL Real - Atlético Madrid

[Enlace](#)

Crisis Ucrania - Rusia

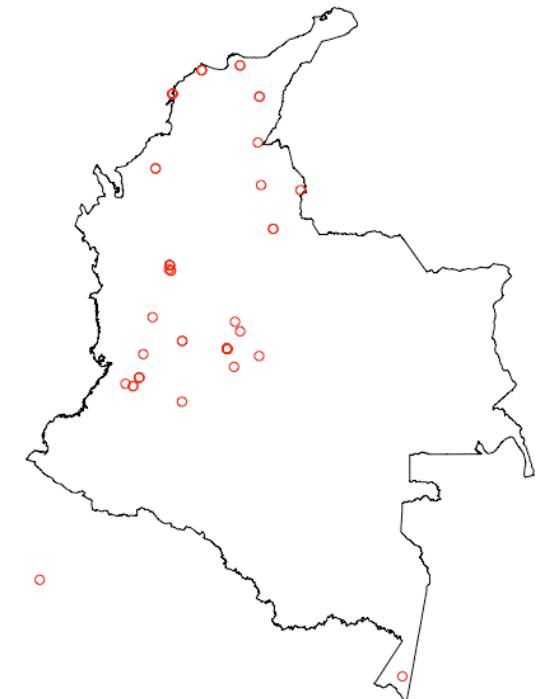
[Enlace](#)

Beyoncé álbum release

[Enlace](#)

familiar consumo pues mes menos media nacional alvarouribevel pago clase iba afip tener dice ricos solo santos ver iván aumento libro empresas gt así bajar
canasta impuesto igual presidente pagó gravar financiamiento lafm quieren cobro toda productos reforma decir favor billones
mascotas hoy igual corte tal presidente pagó gravar financiamiento lafm quieren cobro toda productos reforma decir favor billones
gobierno hoy igual corte tal presidente pagó gravar financiamiento lafm quieren cobro toda productos reforma decir favor billones
impuesto hoy igual corte tal presidente pagó gravar financiamiento lafm quieren cobro toda productos reforma decir favor billones
colombia hoy igual corte tal presidente pagó gravar financiamiento lafm quieren cobro toda productos reforma decir favor billones
duque hoy igual corte tal presidente pagó gravar financiamiento lafm quieren cobro toda productos reforma decir favor billones
ivan hoy igual corte tal presidente pagó gravar financiamiento lafm quieren cobro toda productos reforma decir favor billones

```
topfeatures(dtm_2gram, 10)
canasta familiar alimento mascotas      suprema corte ley financiamiento
        486           101           76           73
    clase media     comida mascotas reforma tributaria      corte avala
        62            52            50            45
cobro alimento alimentos mascotas
        42            37
```



Ejemplos de uso de data mining y datos de redes sociales: Consultas sobre @DilianFrancisca

63



> topfeatures(dtm_sw, 20)

diliani	francisca	mauricearmitage	cali	gobernadora	seguridad
1085		585	323	248	201
alcalde		ciudad	gobvalle	valle	alcaldia
185		161	154	113	decali
ivanduque		robertoortizu	calisosvos	usted	natiibedoya
91		75	69	68	67

