

Estrategias de marketing y comunicación en la era Covid-19:

Uso de web scraping en RRSS para su análisis

Por:
María Nanton (mcnanton@gmail.com)
@bynans1

Adquisición de datos masivos

Herramientas y posibilidades

Adquisición de datos masivos para la investigación comercial o académica

¿Qué es el Web Scraping?



Adquisición de datos masivos para la investigación comercial o académica

¿Cómo se scrapea un sitio?



```
jupyter Twint_FINAL Last Checkpoint: 01/11/2020 (autosaved)
File Edit View Insert Cell Kernel Widgets Help Trusted Python 3

In [9]: import nest_asyncio
nest_asyncio.apply()

In [10]: for twitter_handle in twitter_list:
    c = twint.Config()
    # c.Lang = 'es'
    c.Search = 'from:' + twitter_handle
    c.Since = "2010-01-01"
    c.Store_csv = True
    c.Output = twitter_handle + ".csv"
    c.Count = True
    twint.run.Search(c)

1324741726163132419 2020-11-06 12:53:34 -0300 <CespedesLibros> @lucianacancer Ni que el correo trabajara para nosotrosxs!
132467836635848448 2020-11-06 08:41:48 -0300 <CespedesLibros> Miguel, quien se ha robado nuestro corazón ❤️
1322693947437785089 2020-10-31 21:16:26 -0300 <CespedesLibros> @megustaleerarg
1317576932976885768 2020-10-17 18:23:14 -0300 <CespedesLibros> @lannister2007 Hola, por favor escribimos por whatsapp para ve
n qué pudo haber pasado. Si la compra fue por Mercado libre con Mercado envíos, la logística es de la plataforma de venta y no
de la librería, no tenemos incidencia ni forma de modificarlo.
1316073099012513794 2020-10-13 14:47:32 -0300 <CespedesLibros> @chutass_ pero te vas a fundir ahorrando :)
1316070158507618112 2020-10-13 14:35:51 -0300 <CespedesLibros> Queridos clientes: este 14 y 15 de octubre con tarjetas de gba
ncoprovincia van a poder aprovechar 30% de reintegro y 3 cuotas sin interés en la compra de sus libritos para regalar y regal
arse. https://t.co/hnaCTL4b5
1315638569927684098 2020-10-12 10:00:53 -0300 <CespedesLibros> Hoy abrimos en nuestro horario habitual 10-18 horas. Les esper
amos!
1315057930855789312 2020-10-10 19:33:37 -0300 <CespedesLibros> @sebacampanario @mabiuso Abriremos salón directamente después
de la mudanza, hasta 3 personas y con todos los protocolos. Estamos trabajando en eso! 🍷
1315042824316952578 2020-10-10 18:33:36 -0300 <CespedesLibros> @sebacampanario @mabiuso Estamos abiertxs desde mediados de ab
ril. El salón no está habilitado para el ingreso, pero seguimos atendiendo desde la puerta 🍷
1314261026079158274 2020-10-08 14:47:01 -0300 <CespedesLibros> Pueden reservar sus ejemplares en nuestro whatsapp +5411300536
21
1314261024279846912 2020-10-08 14:47:00 -0300 <CespedesLibros> Vamos a tener en stock -Aranat -Averno -Iris Salvaje -Vita No

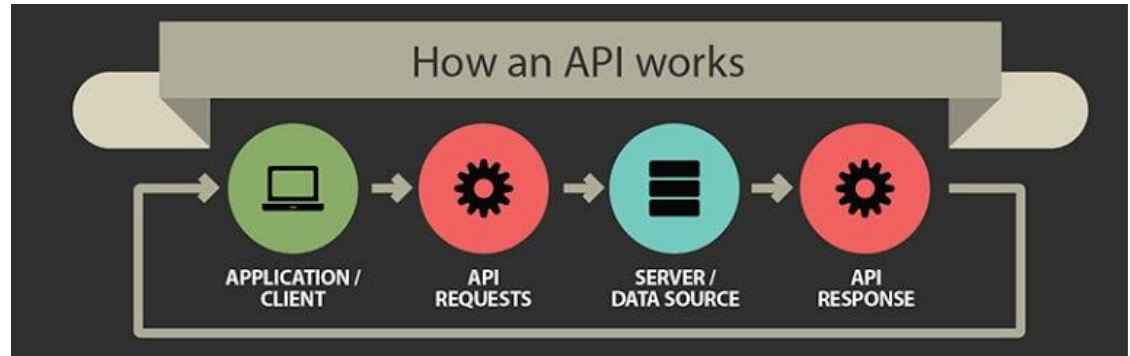
In [ ]: import twint
c = twint.Config()
c.Search = "#leydehumadalesya"
c.Since = "2020-01-01"
c.Until = "2020-02-28"
c.Show_hashtags = True
c.Followers = True
c.Custom["tweet"] = ["id", "conversation_id", "tweet", "username", "user_id", "date", "time", "likes_count", "replies_count", "re
c.Output = "nombrearchivo.csv"
c.Store_csv = True
twint.run.Search(c)
```

Adquisición de datos masivos para la investigación comercial o académica

¿Cuáles son las alternativas?



Instagram API



Adquisición de datos masivos para la investigación comercial o académica

¿Cuáles son las alternativas?



Brandwatch



socialbakers



sysomos

Now Meltwater Social

RRSS en contexto Covid-19

Análisis preliminar de estrategias de marketing
y comunicación de librerías y editoriales

Estrategias de marketing y comunicación en la era Covid-19

El corpus



25k de publicaciones de 26
usuarios de librerías y editoriales
que comercializan libros en
Argentina



41k de publicaciones de 40
usuarios de librerías y editoriales
que comercializan libros en
Argentina

Estrategias de marketing y comunicación en la era Covid-19

El corpus

```
c> Users > PC > Documents > Untitled Folder 2 > base_tw > eternacencia.csv
1 id,conversation_id,tweet,user_name,user_id,date,time,likes,count,retweets,count,hashtags,mentions,retweet,quote_url,user_rt_id,source,retweet_date
2 1322951803316921352,1322951803316921352,y_tsted https://t.co/BG5BEXthK,eternacencia,34765906,2020-11-01,14:18:10,0,30,4,5,[],[],False,...
3 132292488100959911,132292488100959911,"...Yo, han dicho que mis obras traducidas al español no han quedado muy bien, pero no sé, no hablo español"" https://t.co/QA98QWDPDI",eternacencia,34765906,2020-11-01,12:34:
4 13228804721149186849,13228804721149186849,"Mila Guerriero presenta su nuevo libro. La entrevista @AlaKaremina Fagnelli,eternacencia,34765906,2020-11-01,10:34:14,24,0,3,[],[],False,...
5 1322874781285392384,1322874781285392384,"Las dependencias, un perfil de la escritora Silvina Ocampo realizado en 1999 https://t.co/MBvYQCEPFA",eternacencia,34765906,2020-11-01,09:15:00,80,2,15,[],[],False,...
6 1322697362444185600,1322697362444185600,"...La cantidad de cuentos que logró publicar parece interminable y por ellos desfilan, literalmente, miles de personajes": tomado de Cuentos selectos (@dhasaArgentina), aq
7 1322664897956569088,1322664897956569088,"Mue admirada por Alexandina Storni y Rubén Darío https://t.co/XtJK0S8VWH",eternacencia,34765906,2020-10-31,19:21:00,88,
8 132264677845889312,132264677845889312,"...Yo soy una poeta que se mete, un poco intrusa, en otras áreas"" @MarlaneGroni https://t.co/h2dQIdoHn",eternacencia,34765906,2020-10-31,18:09:00,53,0,9,[],[],False,...
9 1322608725159976960,1322608725159976960,"...Nadie debería sentirse obligado a terminar un libro que no le gusta"" @FlaviaPetella https://t.co/AU1jnnshKf",eternacencia,34765906,2020-10-31,15:06:00,28,1,6,[],[],False,...
10 1322571029462437890,1322571029462437890,"...Poco a poco se iría transformando en una novela mía, y no sé si podría seguir tratándose de una traducción"" César Aira https://t.co/LXW0rKwU1",eternacencia,34765906,2020
11 1322560711365578561,1322560711365578561,"...Cuando yo era chico, una vez me regalaron una estación de bomberos de juguete, al que no me interesaba particularmente, pero encontré que, si usaba mi imaginación, yo era
12 1322518237785616904,1322518237785616904,"...Joyce Carol Oates: ""La soledad da origen a la literatura y otras veces puede ser un impedimento"" https://t.co/MXCEBfMy2",eternacencia,34765906,2020-10-31,09:39:00,22,0,3
13 1322379402766218966,1322379402766218966,"...El fin del concepto de publicación, tal como lo entendíamos hasta ahora", por @Gschweilens https://t.co/nMh3z3fTM",eternacencia,34765906,2020-10-31,00:26:54,13,0,6,[],
14 1322346298477932544,1322346298477932544,"...Mi mente es un suave espejo del mundo"". Dos poemas inéditos de Hécate Carliárescu en versiones de @AlaKaremina, https://t.co/F4bRIALF0",eternacencia,34765906,2020-10
15 1322314337994887169,1322314337994887169,"...ya he visto cómo se puede la carne se parece al color de algunas flores"" Tres poemas de Lila Biscia (@liluxb), tomados de La casa del tornado (@bajolaluna), https://t.c
16 1322299993462435842,1322299993462435842,"Virginia Woolf se negó a editar el Ulises: no quiso publicarla en el sello que mantenía con su marido, Hogarth Press. Las tiradas allí eran de 300 ejemplares. https://t.co/Ul
17 1322292460172739905,1322292460172739905,"Se viene la segunda! #leechild,eternacencia,34765906,2020-10-30,18:41:04,2,0,0,[],[],False,https://twitter.com/blattyrios/status/1322197737937448961,...
18 1322291880948846592,1322291880948846592,"La verdad increíble, de Tove Jansson, novedad de @AlaNavierallida recién llegada. #novedades https://t.co/rnsdWVR0CM",eternacencia,34765906,2020-10-30,18:38:46,15,0,0,[],[],False,...
19 1322277331710080129,1322277331710080129,"La tortá: ¿desde cuándo la muerte es bato en los libros para chicos? Sandra Comino deja una guía generosa de lectura alrededor de este tema. https://t.co/ngUKdHudd,eterna
20 132224085380733825,132224085380733825,"...La torta de naranja no tenía una receta legendaria, pero era rica y a los turistas les gustaba."" ""Torta de naranja"", un cuento de @talligolman, tomado de su libro Larga
21 132221665897132610,132221665897132610,"¿Qué novedad llegó de @las afueras? #novedades https://t.co/fotogty3dQg,eternacencia,34765906,2020-10-30,13:39:51,16,3,1,[],[],False,...
22 1322149954180965617,1322149954180965617,"...Camino @chiditadora, se sirve Oreo Manhattan? #novedades https://t.co/cy2wTud8",eternacencia,34765906,2020-10-30,13:33:15,32,2,2,[],[],False,...
23 132210322011357185,132210322011357185,"Esta semana, por suerte, salió el sol... #novedades https://t.co/ARUyPi5wF",eternacencia,34765906,2020-10-30,13:14:40,26,1,2,[],[],False,...
24 132208411782191104,132208411782191104,"...Nunca hice carrera de nada, ni siquiera de la escritura. Arranqué con nada en el mundo pero una cierta pasión, un deseo impulsor. No sé de dónde vino y no sé por qué"" https://t.co/R1jdKnDjR,eternacencia,34765906,2020-1
25 1322191019107763280,1322191019107763280,"...@orsenix En tal caso te dejo la sección de mi blog donde cargamos poesía cada miércoles. ¡Ojalá encuentres algo para vos! https://t.co/R1jdKnDjR,eternacencia,34765906,2020-1
26 1322185130555953152,1322185130555953152,"¿Qué libros son sus poemas favoritos vivos? ¿poetas favoritos? De cualquier parte del mundo, pregunto...".eternacencia,34765906,2020-10-30,11:34:34,41,43,4,[],[],False,...
27 1322178946184679429,1322178946184679429,"Novedades para escribir que #Poe dejó con pistas en sus obras. Un taller literario imprescindible en 7000 caracteres. https://t.co/TjS3dXW5X",eternacencia,34765906,2020
28 1322148746860371971,1322148746860371971,"Una conversación con @paulavazquez, escritora, abogada y también librera, que por estos días inaugura junto a su socio la segunda sucursal de @latapeinada, en Madrid. ""Sólo
29 1321983658962628416,1321983658962628416,"Lectura de marates: ""El principio de las cosas"", gran cuento gran de la autora inglesa Ali Smith, incluido en la antología La historia universal (@mordicaLibros). https://
30 1321951608592608263,1321951608592608263,"Un extracto de la novedad de @ampersand_ed: El libro expandido. Variaciones, materialidad y experimentos, de @maranorhorsk. https://t.co/HF3jDbImk1",eternacencia,347659
31 1321910678379257863,1321910678379257863,"Los consejos para escribir que #Poe dejó con pistas en sus obras. Un taller literario imprescindible en 7000 caracteres. https://t.co/TjS3dXW5X",eternacencia,34765906,2020
32 1321878214275737797,1321878214275737797,"...La lección mayor del infinito es dejar de ser a veces infinito..."" Tres poemas de Roberto Durrero, tomados de Poesía vertical (@visiorlibros). https://t.co/cnfYjC3K2A",ete
33 1321849021938316448,1321849021938316448,"Seguimos con fotos de nuestro paisaje favorito, las bibliotecas! 📖 Aquí la de Daniel Huergo, quien dice acerca de su orden: ""El plan original fue por género. Luego el azar y
34 1321815948617837824,1321815948617837824,"...Ya no sabemos nada sobre nuestro futuro personal, mucho menos sobre el futuro global. Lo que me parece seguro es que ya nada volverá a ninguna normalidad"". Lee la introducc
35 1321785108934254592,1321785108934254592,"Leé el arranque de El gueto interior, de Santiago Amigorena (@megustaleerarg) #nueva finalista de los tres grandes premios de las letras francesas en 2019 y traducida por @ma
36 1321621819438381696,1321621819438381696,"Nueva pregunta a Cristina García Morales. La escritora española, autora de Lectora fácil, responde nuestro cuestionario filio. https://t.co/CLR2blg2Qy",eternacencia,347659
37 1321591575500891392,1321591575500891392,"Nueve preguntas a @francovaccarini. El autor de más de ochenta títulos para un público infantil y juvenil responde nuestras preguntas de siempre. https://t.co/Fbf72boul1eter
38 1321558356994215936,1321558356994215936,"...El corazón de un granovaccarini. El autor de más de ochenta títulos para un público infantil y juvenil responde nuestras preguntas de siempre. https://t.co/mmqC03DFKv",eternacencia,34765906,2020-10-28,18:04:00,7,0,1,[],[],False,...
39 1321512939184865286,1321512939184865286,"...Los niños al nacer saben leer el mundo, no los libros; nos hacemos lectores viendo leer"" Mercedes Calvo sobre mediación lectora, poesía e infancias #U1 https://t.co/gi
40 1321516813084677511,1321516813084677511,"...Felices os, @margo glantz! 📖 https://t.co/8WmXh9ZD",eternacencia,34765906,2020-10-28,15:55:33,56,0,5,[],[],False,...
41 1321508475286985888,1321508475286985888,"...Un cambio de orgánico", un cuento de @ropilli https://t.co/cyMh38K0c",eternacencia,34765906,2020-10-28,14:14:00,20,0,3,[],[],False,...
42 1321493688348368993,1321493688348368993,"...@moralciraui ¡Nunca es tarde, Moral!,eternacencia,34765906,2020-10-28,13:47:00,4,1,0,[],[],False,...
43 1321493409123762177,1321493688348368993,"...@moralciraui ¡Nunca es tarde, Moral!,eternacencia,34765906,2020-10-28,13:47:00,4,1,0,[],[],False,...
44 1321486634009858848,1321486634009858848,"El observatorio de palabras de la @RAEinforma incluye el ""elle""; no determina la implementación de este término inclusivo, pero su incorporación abre el debate https://t.co
45 1321471781186785282,1321471781186785282,"...@santapaluca https://t.co/vYj1aX0y2,eternacencia,34765906,2020-10-28,12:19:50,6,0,0,[],[],False,...
46 132146952166442240,132146952166442240,"...¿Cuál es el primer libro que recuerdan haber leído?,eternacencia,34765906,2020-10-28,12:11:00,154,529,9,[],[],False,...
47 1321452660763947089,1321452660763947089,"...Si la no violencia parece una posición ""débil", deberíamos preguntarnos: ¿qué se considera fuerza?"". Lee un fragmento de La fuerza de la no violencia, de @JudithButler (@pl
48 1321421958295265280,1321421958295265280,"...[te acordás de aquella vez que paraste en la mitad de un poema para contarme de una ciudad sin sombras?]"... Tres poemas de Marilía García, tomados de Cámara lenta (@2ind
49 1321258380836485126,1321258380836485126,"Compartimos uno de los textos incluidos en el Dossier descargable gratuito realizado en el marco de las Jornadas A Propósito de Juana Bignozzi (Buenos Aires, 1937-2015), orga
```

Estrategias de marketing y comunicación en la era Covid-19

Frecuencia de publicación

Comparación de publicaciones mensuales Año/Año

usuario	2020-01	2020-02	2020-03	2020-04	2020-05	2020-06	2020-07	2020-08	2020-09	2020-10
sbsargentina	-0.52	-0.41	0.16	2.27	0.08	3.22	10.33	2	1.16	1
notanpua	-0.43	-0.58	-0.5	-0.45	0.12	0	-0.47	-0.36	-0.5	-0.6
librosvuelvoalsur	0.07	0.61	-0.46	0.1	0.13	0.81	0.52	0.94	0.28	0.16
mandragora.librosycultura	1.33	-0.27	-0.44	-0.74	-0.84	-0.58	-0.73	-0.68	-0.58	-0.48
librosdelpasaje	-0.22	-0.29	-0.5	0	-0.9	0.43	-0.42	0	-0.38	-0.5
libreriatraslospasos	-0.11	-0.17	-0.57	0.37	-0.04	-0.17	-0.37	-0.39	-0.47	-0.32
libreriahomosapiens	-0.59	-0.89	-0.87	0	0.29	-0.18	0.36	-0.38	0.28	0.11
libreriadelfondoycc	-0.75	-0.5	0.17	-0.77	0.83	1.4	2.33	2	0.6	1.17
libreriacorneja	0.27	-0.52	-0.28	0.1	0.5	0	0.2	-0.22	-0.38	-0.17
libreriademujeresargentina	-0.76	-0.65	-0.84	-0.12	-0.81	-0.43	-0.72	-0.69	0.03	-0.34
libreriahammurabi		0.33	-0.07	0.62	-0.15	-0.17	0.01	-0.24	-0.03	0
libreriadefamilias	0.5	0	0.78	0	0.55	0.74	-0.11	-0.17	-0.34	-0.42
yenny_elateneo	-0.39	0.46	0.1	0.03	-0.2	0.31	-0.02	-0.04	0.02	0.12
libreriakier	Inf	11	12	2	3.33	Inf	1.33	9	11	Inf
libreria_guardia	-0.58	-0.82	-0.74	-0.89	-0.86	-0.8	-0.92	-0.93	-0.94	-0.85
zivals.ok	-0.5	-0.56	-0.11	0.75	-0.57	-0.38	0.07	-0.33	0	-0.3
libreriarodriguezalmagro	0.33	-0.9	-0.04	0.09	2.08	0.89	0.33	0.05	0.65	0.09
libreriadelconti	1.5	-0.6	-1	0	-0.86	-0.65	-0.55	-0.18	2.33	Inf
libreria.norte	0.55	0.4	0.5	-0.6	-0.78	0.39	0.65	0.33	-0.03	-0.11
libreria.micasa	0	-0.5	-0.36	0	1	0.36	-0.6	0.5	0.67	-0.27
libreria.afortismos	-1	0	-0.71	-0.64	1.67	2.33	1.67	Inf	0	1.75
estacion_libro	1	4	8	Inf	1.5	4	Inf	-0.38	-0.11	-0.44
cuspidelibros	1.17	0.79	0.92	-0.23	-0.42	-0.22	-0.68	-0.67	-0.76	-0.76
cespedeslibros	Inf	-1	-0.33	3	Inf	0.5	Inf	0.33	-0.75	0
editorialsanpabloargentina	0.04	0.08	0	0.28	0.65	0.95	1.13	-0.18	-0.91	-1
labolaeditora		0	5	-1	1.67	1.67	2	-0.33	0	8
labestiaequilatara_editorial	0.5	-0.38	-0.42	-0.29	-0.19	0.16	0.28	0.36	0.41	0.14
chinaeditora	Inf	Inf	Inf	Inf	Inf	Inf	-0.56	-0.5	-0.83	-0.25
ediciones.amperсанд	-0.17	-0.27	0.19	0.06	-0.21	-0.07	0.67	0.17	0	-0.5
editorial_sigilo	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf
oceanos_argentina	-0.6	-0.17	0.19	-0.09	0	0	0	-0.03	0.07	0.29
quipulibros	0.75	0.69	0.73	-0.25	0.02	0.16	-0.26	0.72	-0.34	-0.38
cajanegraeditora	2.75	1.12	2.5	0.5	0.86	0.69	0.83	0.35	0.81	0.27
ed_entropia	-0.75	-1	-0.78	-1	-1	-1	7.5	6.6	13.5	5
smiliteraturajjuvenil	-0.29	-0.07	0.42	-0.15	0.21	1.78	11.5	3.5	1	Inf
sigloxxiarg	-0.43	-0.8	0.12	0.1	0.08	-0.22	0.55	0.25	0	1
seix_barral	0.08	0	-0.21	0	-0.23	0	-0.08	0.15	0.56	0
planetadelibrosar	-0.82	-0.84	-0.88	-0.87	-0.88	-0.71	-0.24	-0.27	-0.5	0.57
anagramaeditor	3.43	3.14	3.29	3.29	0.74	4.8	1.46	1.29	1.83	0.22
alfaguaraes	0.88	1.2	2	2.38	0.61	3.57	9	Inf	3.8	4.25

factor(valor)

negativo

nulo

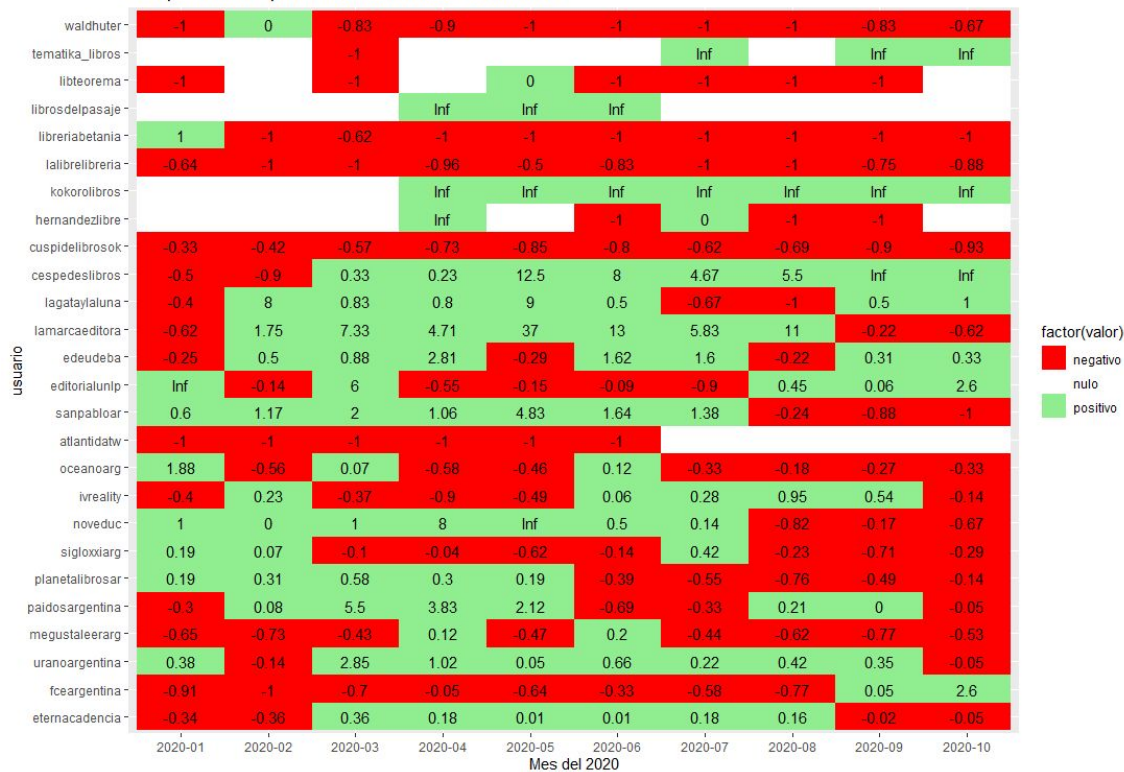
positivo



Estrategias de marketing y comunicación en la era Covid-19

Frecuencia de publicación

Comparación de publicaciones mensuales Año/Año

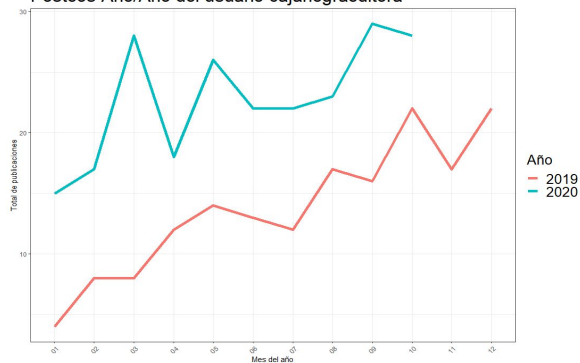


Estrategias de marketing y comunicación en la era Covid-19

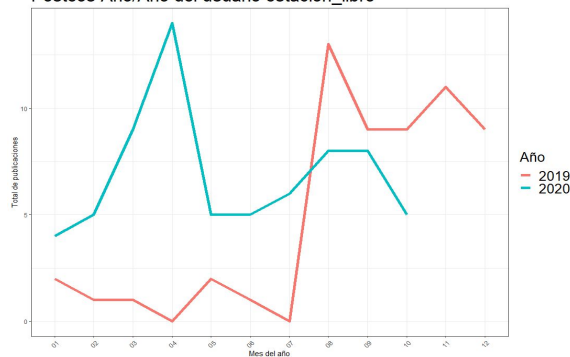
Frecuencia de publicación



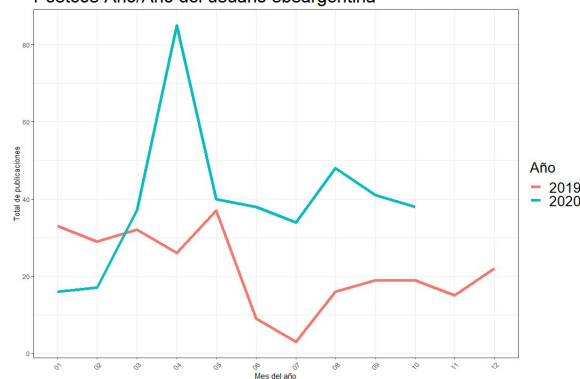
Posteos Año/Año del usuario cajanegraeditora



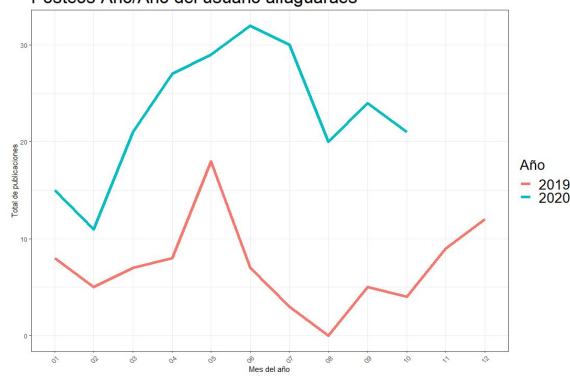
Posteos Año/Año del usuario estacion_libro



Posteos Año/Año del usuario sbsargentina



Posteos Año/Año del usuario alfaguaraes

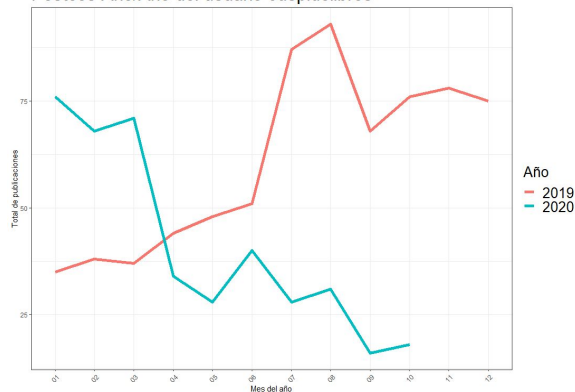


Estrategias de marketing y comunicación en la era Covid-19

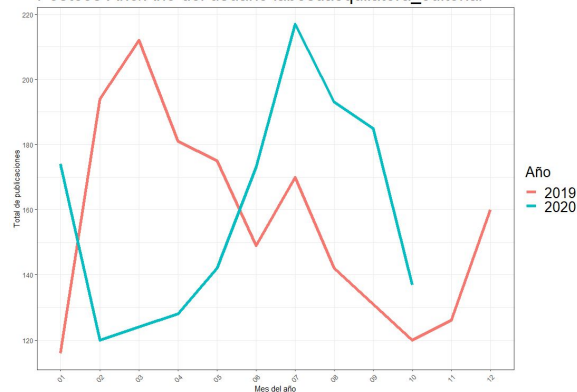
Frecuencia de publicación



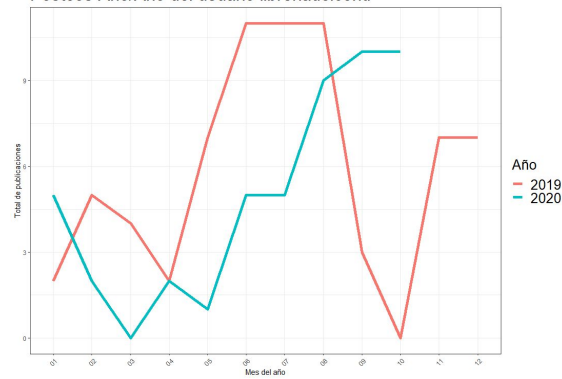
Posteos Año/Año del usuario cuspidelibros



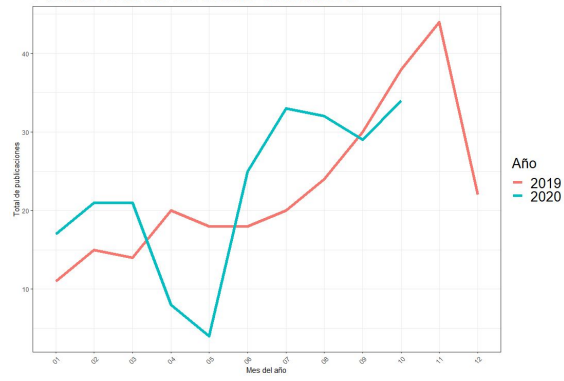
Posteos Año/Año del usuario labestiaequilatera_editorial



Posteos Año/Año del usuario libreriadelconti

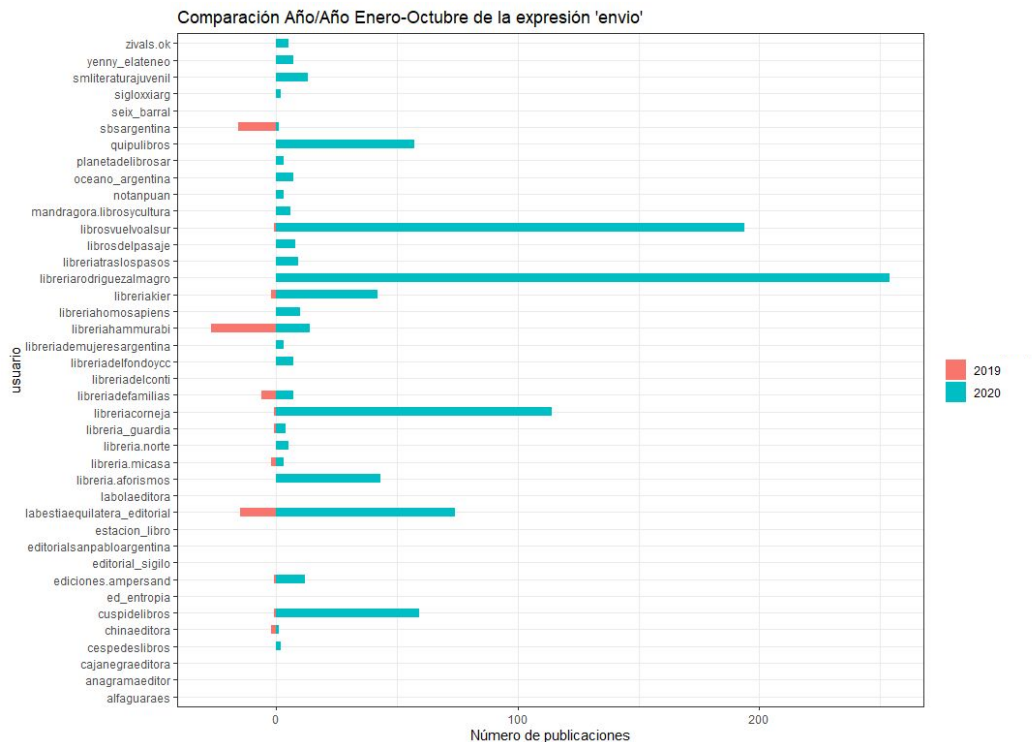


Posteos Año/Año del usuario libreria.norte



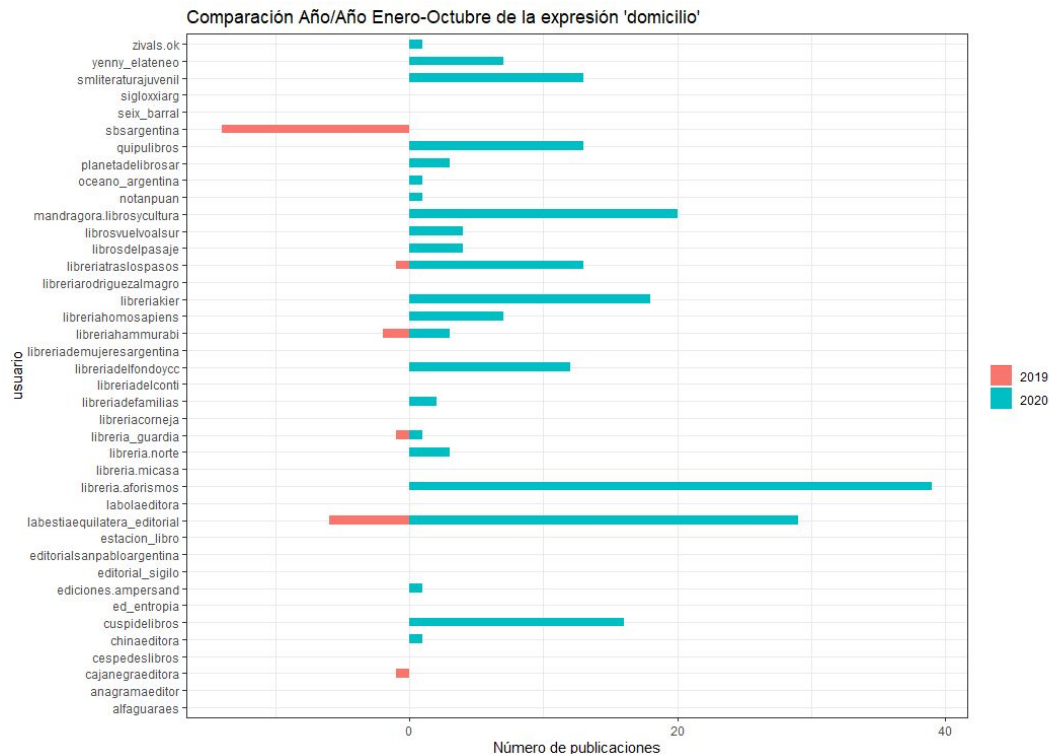
Estrategias de marketing y comunicación en la era Covid-19

Tácticas de marketing - uso de RRSS como proxy



Estrategias de marketing y comunicación en la era Covid-19

Tácticas de marketing

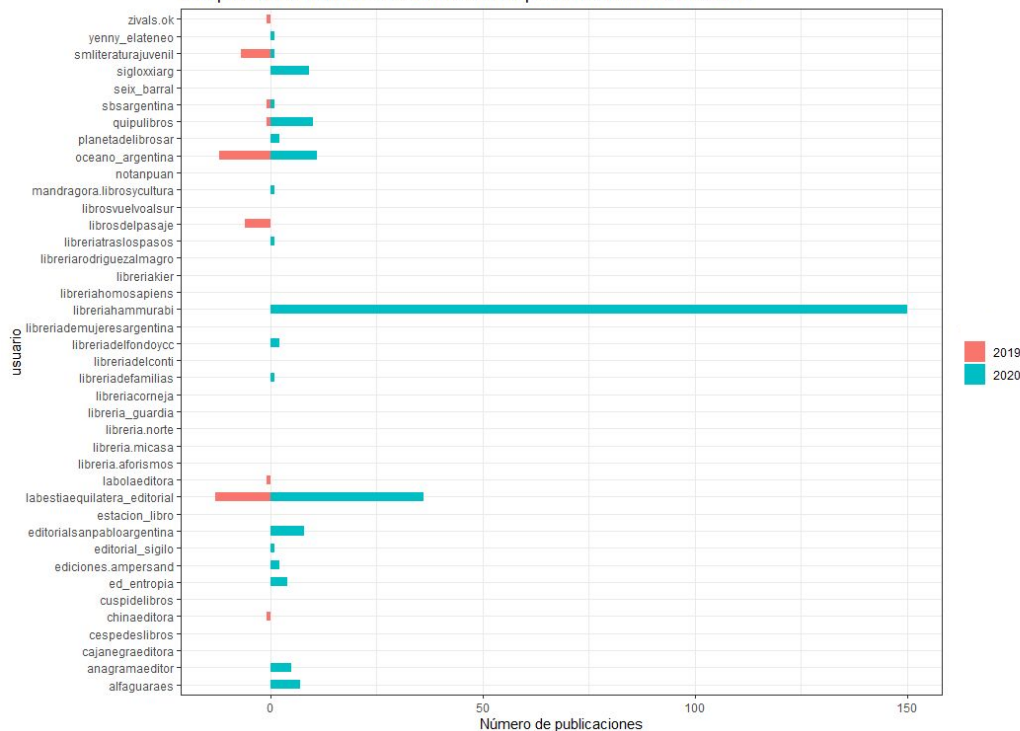


Estrategias de marketing y comunicación en la era Covid-19

Tácticas de marketing

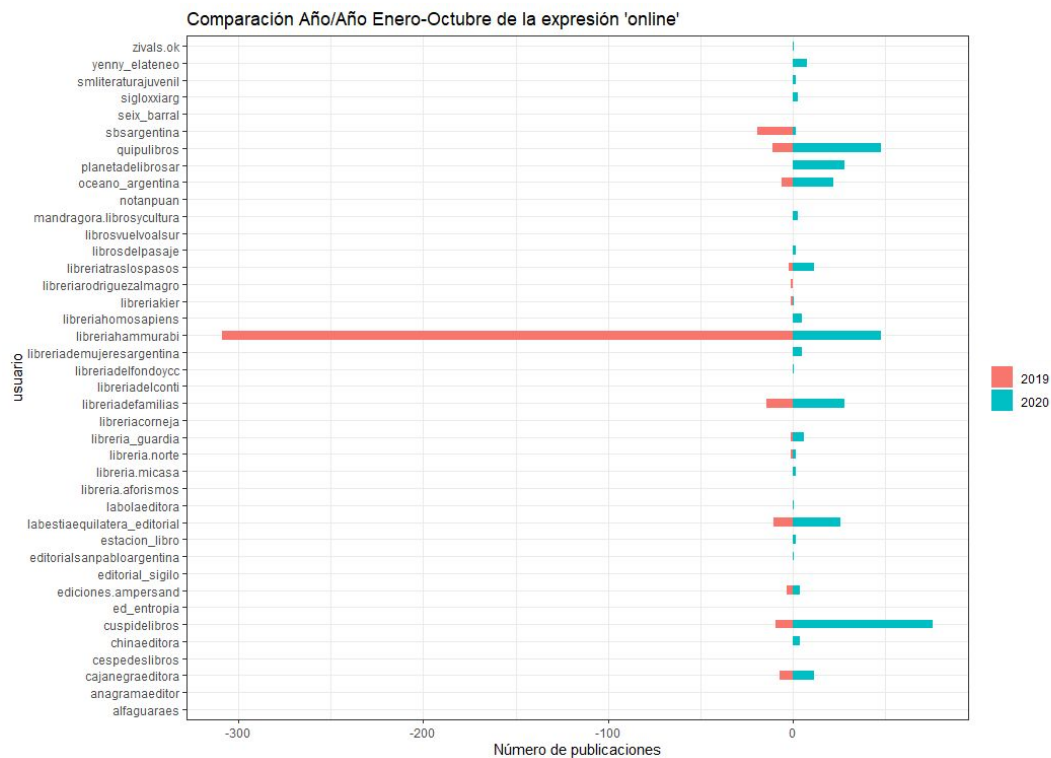


Comparación Año/Año Enero-Octubre de expresiones relativas a ebooks



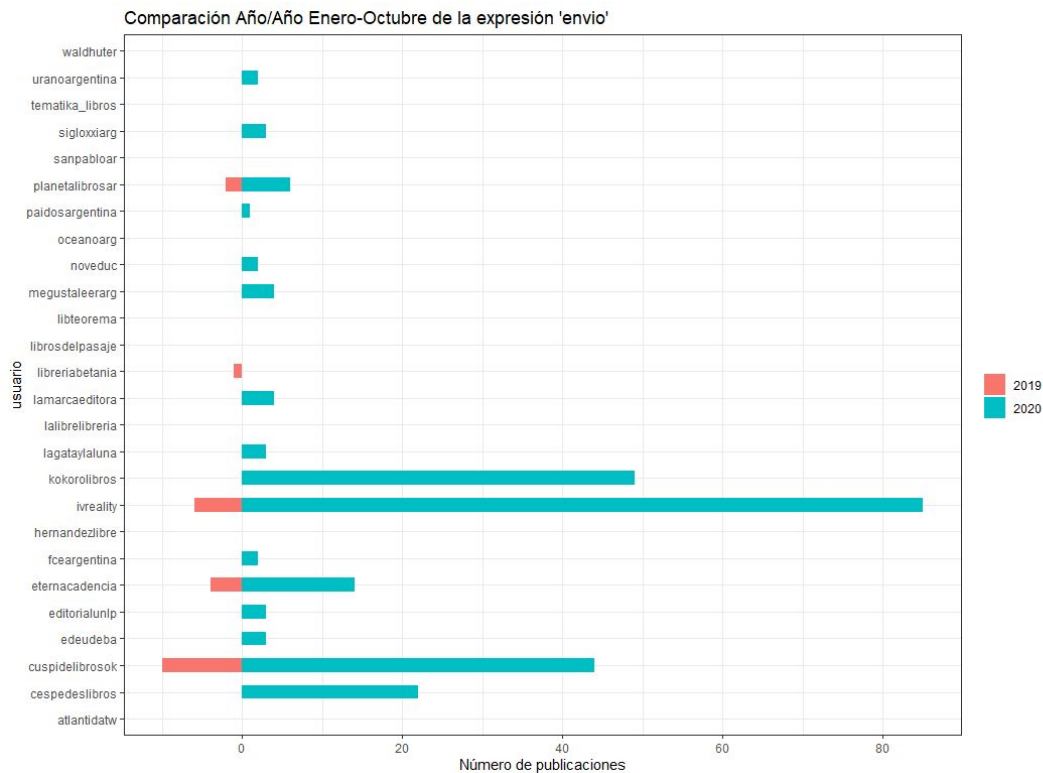
Estrategias de marketing y comunicación en la era Covid-19

Tácticas de marketing



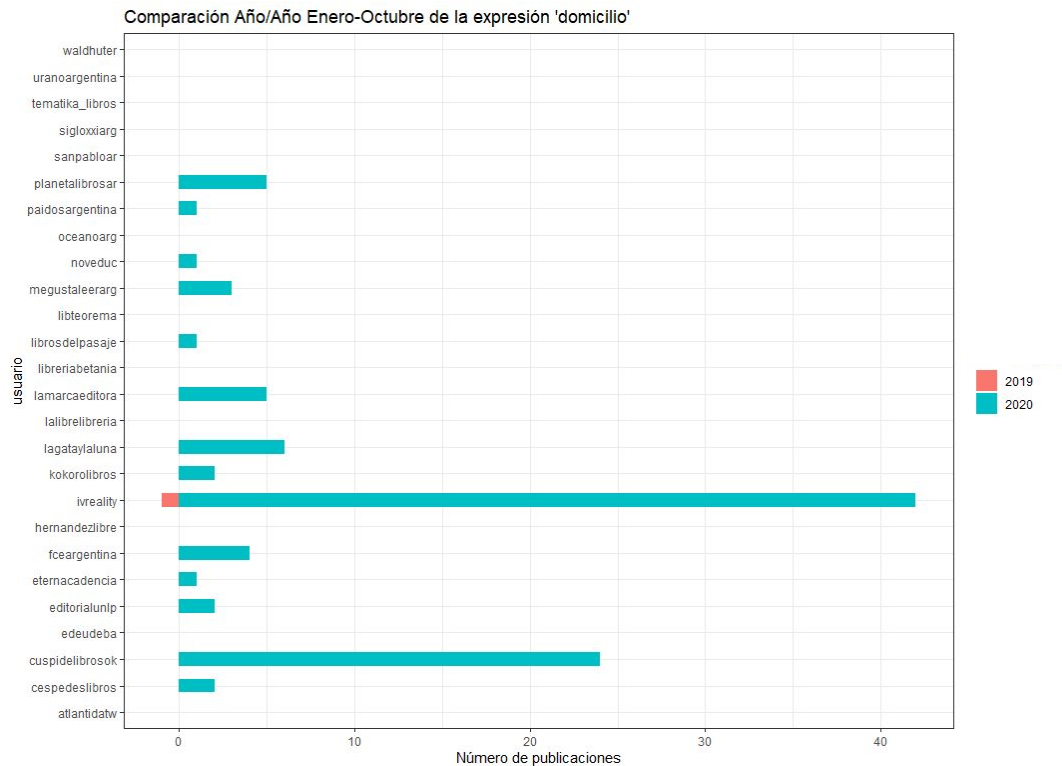
Estrategias de marketing y comunicación en la era Covid-19

Tácticas de marketing



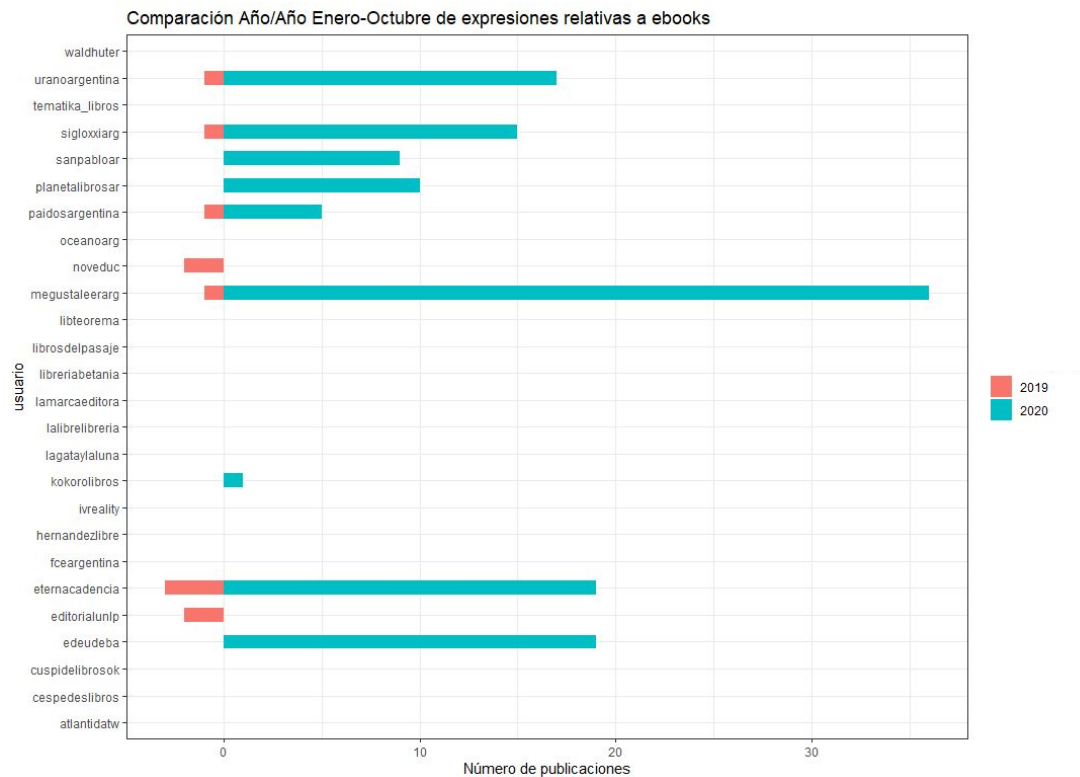
Estrategias de marketing y comunicación en la era Covid-19

Tácticas de marketing



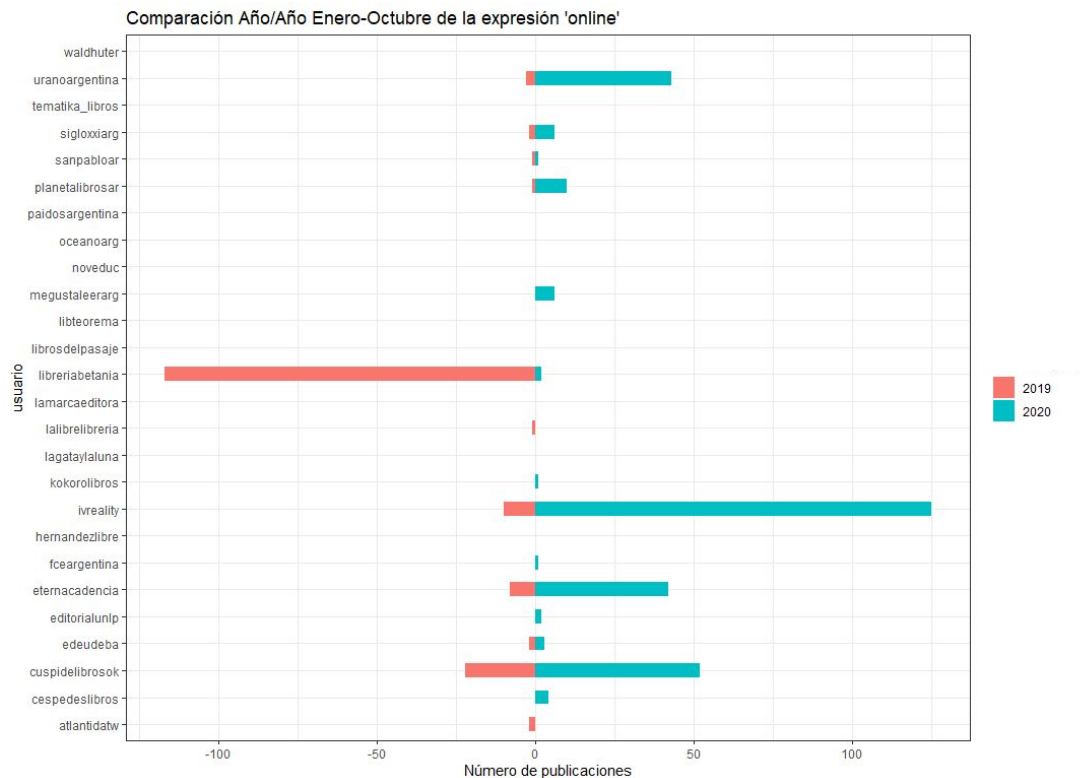
Estrategias de marketing y comunicación en la era Covid-19

Tácticas de marketing



Estrategias de marketing y comunicación en la era Covid-19

Tácticas de marketing





Estrategias de marketing y comunicación en la era Covid-19

Próximos pasos

- Construcción de segmentos de usuarios
- Análisis de la efectividad de las comunicaciones
- Incorporación de usuarios de Facebook al corpus analizado
- Incorporación de algoritmos de detección de texto en imágenes

¡Muchas gracias!



Código y datasets abiertos disponibles en Github

<https://github.com/bynans/congreso-tendencias-2020/>

María Nanton (mcnanton@gmail.com)
@bynans1