# Monitoreo de Marca en Redes Sociales y Medios Digitales

Una alternativa usando R-Shiny







# Overview

- 1. Problema de Negocio: Monitor Marca Grupo Coomeva.
- 2. Análisis de Redes Sociales en R.
- 3. Data en Medios Digitales: Web Scraping vs Webhose API.
- 4. NLP & Modelación Predictiva de Polaridades.
- **5.** Shiny App: Solución Integradora.

# 1. Monitor Marca Grupo Coomeva:

# ¿Qué es Coomeva?





Grupo Empresarial, más de 13 empresas: Coomeva Eps, Bancoomeva, Coomeva Cooperativa.



250.000+ Asociados en la Cooperativa



4+ billones en Patrimonio.\*

# 1. Monitor Marca Grupo Coomeva: Problemas a Solucionar

Categoría	Preguntas
Redes Sociales & Medios Digitales	¿Qué se dice de la marca en redes sociales (Twitter,Facebook)? ¿Opiniones o Comunicados Positivos/Negativas/Neutros? ¿Qué se publica acerca de Coomeva en Medios Digitales (periódicos,blog,etc)?
Monitoreo Competencia	¿Qué campañas está promocionando el entorno competitivo del Grupo Coomeva? ¿Quién publica con mayor frecuencia?



- → 6 millones Usuarios en Colombia\*
- → Canal de interacción oficial de marcas como Xbox, Movistar, Avianca.
- → Sus usuarios se concentran en el rango de 24 a 34 años\*



- → 2da plataforma con mayores usuarios activos en el mundo\*\*
- → Principal repositorio de video de las grandes marcas.



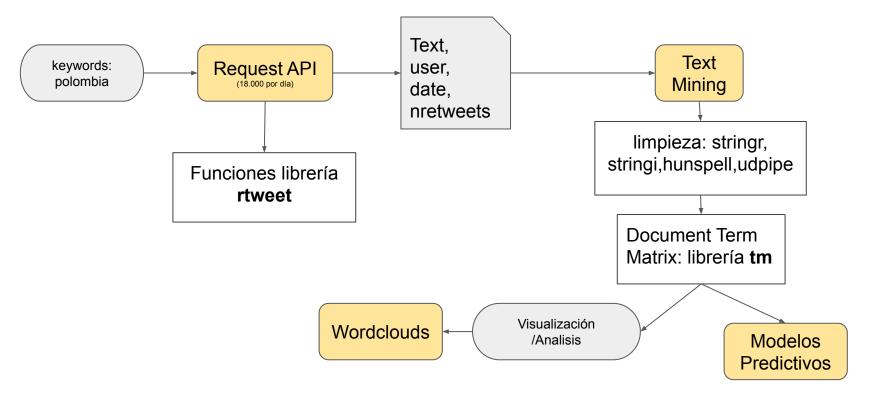
#### **Funcionamiento API**

```
library(rtweet)
token <- create_token(
   app="YOUR_APP_NAME",
   consumer_key = "YOUR_CONSUMER_KEY",
   consumer_secret = "YOUR_CONSUMER_SECRET",
   access_token <- "YOUR_ACCESS_TOKEN",
   access_secret <- "YOUR_ACCESS_SECRET"
)
tweets=search_tweets("polombia -filter:replies lang:es",n = 100,include_rts = FALSE)</pre>
```

user_id	status_id	created_at	screen_name	text
1,1755E+18	1,2259E+18	7/02/2020 21:57	momographic	No es Colombia es Polombia https://t.co/1gvyZOj5FO
144012645	1,22589E+18	7/02/2020 21:09	mechitamayorga	Y si Aída merlano es delincuente en Polombia, entonces que es para Polombia el corrupto terrorista asesino el señor julio Borges que se fue a ese país huyendo de la justicia venezolana https://t.co/5KG1JFju5b
453251447	1,22588E+18	7/02/2020 20:42	DeAvilaM2	#YoLeCreoAMerlano No se puede usar chaqueta roja por ya eres chavista, si señores esto es Polombia. https://t.co/zrtYHHdJeG
1399914055	1,22588E+18	7/02/2020 20:40	LuisEdu620829	Polombia, hijita. Polombia https://t.co/ucpnMJzO4x
195537073	1,22588E+18	7/02/2020 20:21	carolinaguerram	¡Gracias Polombia! Nos vamos de prohibiciones y no de soluciones <u+0001f695><u+0001f644> https://t.co/SQgaFYnFGI</u+0001f644></u+0001f695>



#### **Análisis de Tweets**





## **Posibles Análisis**

Nube de palabras usando la librería wordcloud2

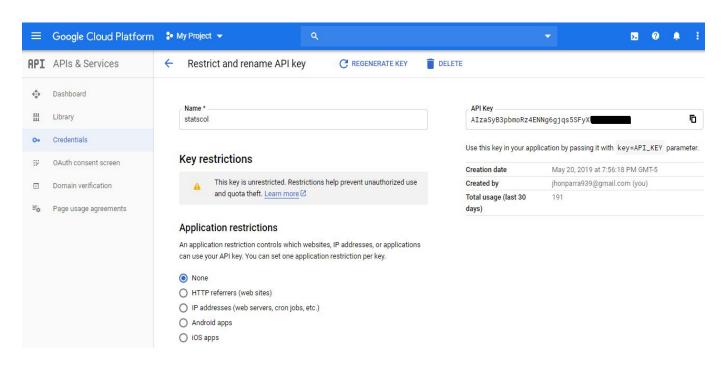




#### **Esenciales**

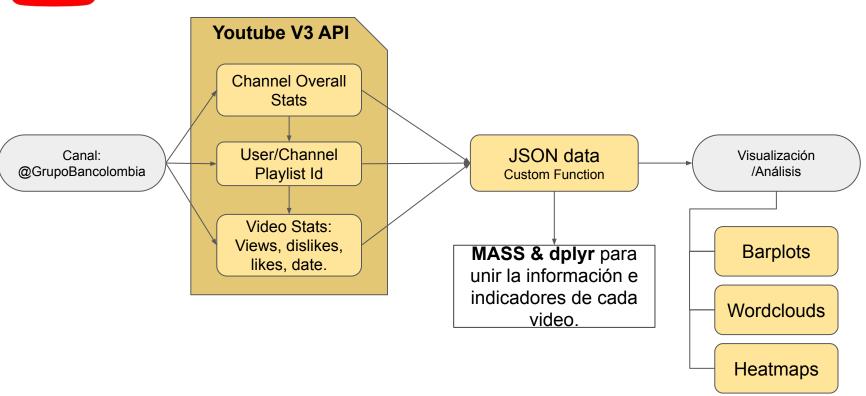
#### YOUTUBE V3 API

- → Registrarse en Google Cloud, en API's & Services buscar Youtube V3 API
- → Obtener una API Key





#### **Estructura API**





#### **Estadísticas de Un Canal**

→ Obtener el id del canal que deseas analizar.

Nombre Canal	ld Canal	API_key="PUT YOUR API KEY HERE"	
		getstats_canal<- <b>function</b> (id,API_key){	
Bancolombia	UCczkYQFOUOgsg958lfKMB_Q	url=paste0('https://www.googleapis.com/youtube/v3/channels?part=snippet'entDetails%2Cstatistics&id=',id,'&key=',API_key)	%2Ccont
Davivienda	UCoh2JDafVSw-0eBuZo4ecpQ	result <- from JSON(txt=url) return(data.frame(name=result\$items\$snippet\$title,result\$items\$statistics,result\$items\$contentDetails\$relatedPlaylists))	pl_list_id=
Banco de Bogota	UCHZS1Ghu4pU-rofAZUPoHFQ	}	
		<pre>getstats_canal(id = "UCczkYQFOUOgsg958lfKMB_Q",API_key = API_key)</pre>	)

name	viewCount	commen tCount	subscriberCount	hiddenSubscri berCount	videoCount	pl_list_id.uplo ads
Grupo Bancolombia	56802555	0	188000	FALSE	934	UUczkYQFOUOg sg958lfKMB_Q



# Obtener videos de un playlist

→ Después de tener el playlist\_id, se puede obtener el listado de videos y su id usando la función get\_playlist\_canal()

API key="PUT YOUR API KEY HERE"	videold	videoPublishedAt
get_playlist_canal<-function(id,API_key,topn=15){	6Vlk1bpeW-E	2020-01-30T20:56:45.000Z
url=paste0('https://www.googleapis.com/youtube/v3/playlistItems?part=contentDetails&play	9J0id95iEbl	2020-01-17T14:36:13.000Z
listId=',id,'&key=',API_key,'&maxResults=',topn)	_dD0yuBRbJc	2019-12-17T16:08:27.000Z
result=fromJSON(txt=url) return(data.frame(result\$items\$contentDetails)) }	oivrF9es4GE	2019-12-05T19:43:18.000Z
	BgAiYH7dG-c	2019-11-30T00:21:39.000Z
<pre>get_playlist_canal(id = "UUczkYQFOUOgsg958lfKMB_Q",API_key = API_key,topn = 5))</pre>		



## Estadísticas de los videos de un playlist

→ Al tener cada **videold** es posible obtener desde las estadísticas generales, hasta el título y descripción del mismo. Sin embargo la función getall\_channels() puede realizar todo el procedimiento para una lista de playlist id's.

#### Para cada video se obtiene:

- → Conteo de Views
- → Like Count
- → Dislike Count
- → Comments Count
- → Título
- → Fecha Publicación

```
API_key="PUT YOUR API KEY HERE"

getstats_video<-function(video_id,API_key){

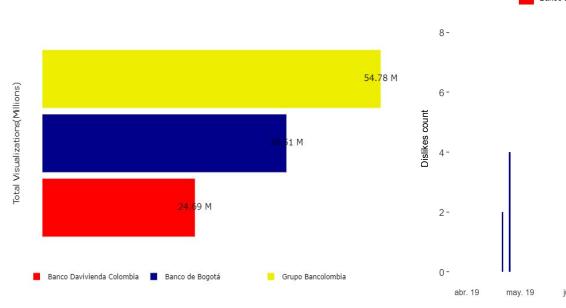
url=paste0("https://www.googleapis.com/youtube/v3/videos?part=snippet,statistics&id=",video_id,"&key=",API_key)
    result <- fromJSON(txt=url)
    salida=list()
    return(data.frame(name=result$items$snippet$channelTitle,
    result$items$statistics,title=result$items$snippet$title,date=result$items$snippet$publishedAt,descrip=result$items
$snippet$description))
}

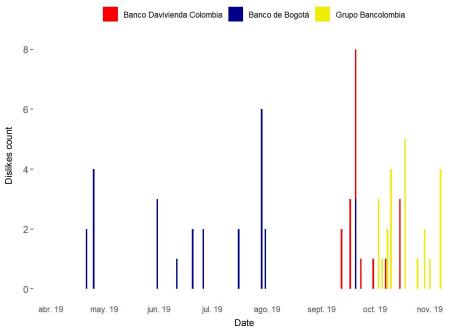
getall_channels<-function(ids,API_key,topn=5){
    videos=lapply(ids,FUN=get_playlist_canal,API_key=API_key,topn=topn) %>% bind_rows()
    stats=lapply(videos[,1],FUN=getstats_video,API_key=API_key)
    stats=bind_rows(stats)
    stats$vid_id=videos[,1]
    return(stats)
}
```



Código disponible en https://www.rpubs.com/statscol/youtube\_data\_in\_r

#### **Posibles Análisis**





# **Conceptos**

#### **Web Scraping**

Obtención de información de un sitio web a través de software. En R se puede realizar con xlm2-Rvest-RSelenium

- Para extraer información de medios digitales, se debería estructurar un código para cada página.
- La estructura del sitio web debe ser estática (puede solucionarse con Rselenium).
- Pueden presentarse bloqueos de IP al hacer muchos request a un sitio web.

# Webhose.io API

API para extraer noticias publicadas sobre un tema en periódicos, blogs, medios digitales en general.

- → Extrae el texto, fecha e interacciones en redes de cada artículo encontrado.
- → 1000 request en free tier:
- → No tiene librería nativa en R. Sin embargo existe una alternativa en github a la medida **github.com/hrbrmstr/webhose**
- → Devuelve diferentes campos: link, título, texto,fecha, Facebook & LinkedIn shares, imágenes externas utilizadas, entre otros.

## Webhose.io

Ir a la página web https://webhose.io/ y registrarse.

webhose.io • API	Dashboard	API Playground	Archived Data	Help Upgrade		juan pe
NEWS, BLOGS & DISCUSSIONS REVIEWS CYBER						
				Archived Dat	a	
Define the query						
stock market language:english +Add Filter				Q		
Format: USON ▼ Crawled Since: ② 3 Days Ago ▼ Sort By: ③ Crawl Date ▼		,	View Past Queries ▼	Example Queries	•	
Endpoint: http://webhose.io/filter/WebContent?token=9b880f4b-8884-4735-b8b2-26798075a85c8format=jpon6sort=crawled&q=stock%20market%20language%	3Aenglish			Q		
RUN	-					

## **Funcionamiento API**

```
library(webhose)
secret="YOUR_TOKEN_HERE"
res=filter_posts("(paro AND nacional AND colombia) language:spanish",token =
list(WEBHOSE TOKEN=secret))
```

	encabezado	fechapub 🏺	sitio	link
1	Petro pide a Comité del Paro adelantar manifestación nacional para el 24 de febrero	2020-02-04	elespectador.com	https://www.elespectador.com/noticias/politica/petro-pide-comite-del-paro-adelantar-manifestacion-nacional-para-el-24-de-febrero-articulo-902925
2	Centro Nacional de Memoria sale de la red internacional de Sitios de Consciencia	2020-02-03	elespectador.com	https://www.elespectador.com/colombia2020/justicia/verdad/centro- nacional-de-memoria-sale-de-la-red-internacional-de-sitios-de- consciencia-articulo-902862
3	Somalia declara la "emergencia nacional" por la plaga de langostas -	2020-02-02	elespectador.com	https://www.elespectador.com/noticias/el-mundo/somalia-declara- la-emergencia-nacional-por-la-plaga-de-langostas-articulo-902657
4	Colombia tiene una nueva universidad pública, Unitrópico deja de ser una institución privada - ELESPECTADOR.COM	2020-01-07	elespectador.com	https://www.elespectador.com/noticias/educacion/colombia-tiene- una-nueva-universidad-publica-unitropico-deja-de-ser-una- institucion-privada-articulo-898745
5	Conductor de taxi pide a mujeres que manejan Uber irse a "planchar o a cocinar"	2020-01-13	elespectador.com	https://www.elespectador.com/noticias/nacional/conductor-de-taxi- pide-mujeres-que-manejan-uber-irse-planchar-o-cocinar-articulo- 899576

## 4. NLP & Modelación Predictiva de Polaridad.

#### **Conceptos**

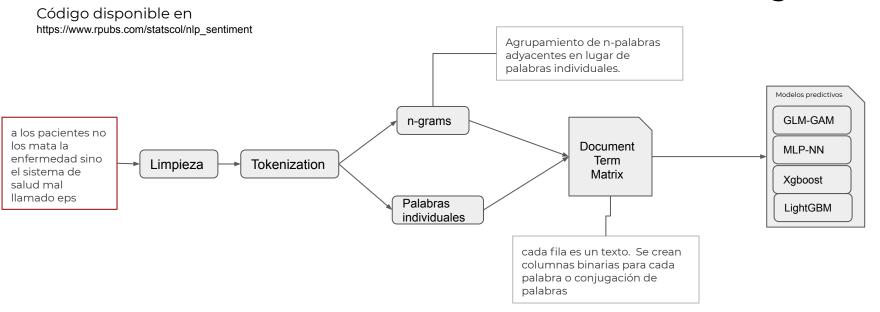
#### Procesamiento Lenguaje Natural (NLP):

Lograr que los cumputadores entiendan, interpreten y manipulen el lenguaje humano.

texto	fecha	opinión
bajo su responsabilidad pongo la vida y salud de mi gineth marin diabetica q esta en hospital de engativa y se encuentra a espera examen urgente colangioresonancia magnetica solicitada x hospital de engativa	4/06/2019 14:52	-1
edicionimpresa las eps no acaban con el viacrucis de los pacientes con cancer en el cesar	4/06/2019 13:50	-1
el mal servicio y la mala atencion en las eps es con la finalidad obligar a los cotizantes a adquirir los famosos planes complementarios por que la salud es un negocio rentable no les basta con lo que aporta uno obligatoriamente	4/06/2019 14:37	-1
gracias por escribirnos nos comunicaremos lo mas pronto posible saludos	4/06/2019 13:35	1
gracias por escribirnos nos encontramos revisando el caso saludos	30/05/2019 20:12	1
a los pacientes no los mata la enfermedad sino el sistema de salud mal llamado eps	4/06/2019 14:20	-1

### 4. NLP & Modelación Predictiva de Polaridad.

## Word n-grams

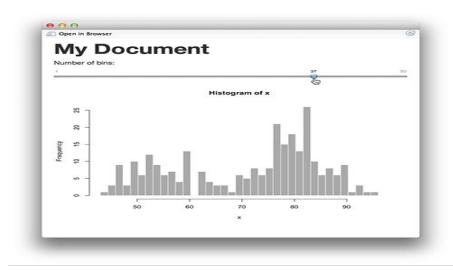


enfermedad	enfermedad_sino	eps	llamado	pacientes	 salud_mal	sino	sino_sistema	sistema	sistema_salud	Opinion
1	1	1	1	1	1	1	1	1	1	-1

# 5. Shiny App: Solución Integradora

#### **Esenciales**

Framework de R para realizar apps web interactivas.



- Compatible con css/html
- Integración de todo tipo de resultados de R de forma interactiva
- Sintaxis sencilla e intuitiva
- Gran comunidad en la web con respuesta a dudas de implementación
- Deployment en servidores de R sin costo adicional



## 5. Shiny App: Solución Integradora

## ¿Cómo Integrar las Soluciones?

Mediante el uso de una Shiny App y R-Markdown se puede dar solución a:

- → Combinar simultáneamente el análisis de redes sociales y de medios digitales, estimando la polaridad de las opiniones.
  - Xgboost de polaridad aplicado en tiempo real a las opiniones recolectadas por la API de Twitter
- → Encontrar campañas de la competencia en diferentes segmentos de negocio de Coomeva: solidario, financiero, cooperativo.
  - Mapear cuentas en redes sociales de cada competidor (si existe)
  - Encontrar palabras clave de detección de una campaña (keywords)
- → Generar un reporte compartible con la fuerza comercial/agencia publicidad
  - ♦ Integrar Shiny con R Markdown



negativo

negativo

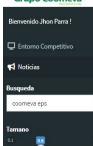
2020-

02-06

34 Link

23 Link

71 Link





₫ Descargar Reporte





		,	
encabezado		sitio	link
IPS Universitaria adeuda salarios a sus empleados	2020-02-08	telemedellin.tv	Link
Supersalud multó a Coomeva por negar procedimiento a un adulto mayor	2020-02-08	telemedellin.tv	Link
Supersalud verificó en Cauca condiciones de atención a usuarios traslados por Ef liquidadas	2020-02-07	periodicovirtual.com	Link
Sanción de la Superintendencia de Salud a Coomeva EPS – Salud	2020-02-06	noticierovenezuela.com	Link
Multa de más de \$870 millones a EPS Coomeva por no atender a un adulto mayor	2020-02-06	larepublica.co	Link
EPS Coomeva, multada con más de \$870 millones por no atender a un adulto ma	yor - 2020-02-06	elespectador.com	Link
Detrimento patrimonial sería de \$ 17.986 millones	2020-02-06	laopinion.com.co	Link
Deathletto parimonaristia de 0 11500 millones	2020 02 00	toopillioniconiiico	

Search:



mayor aduciendo problemas administrativos con los hospitales habilitados para

Información importante sobre Coomeva EPS, https://t.co/UrvLqCjnQz

mayor https://t.co/eBVzy9q3NT https://t.co/OSC0Jbq2Xq

EPS Coomeva, multada con más de \$870 millones por no atender a un adulto

la cirugía. https://t.co/QQdh7AVzYq







Search:



nombre	\$ viewCount 🔻	likeCount	\$	dislikeCount	\$ commentCount	† titulo	∳ fe	cha	<b>♦ link</b>	\$
Coasmedas Cooperativa de ahorro y crédito	295236	1	,	0	0	FERIA DE VEHÍCULO COASMEDAS 2019	201	9-09-13	Link	
Coasmedas Cooperativa de ahorro y crédito	255400	4		0	0	¡Feliz mes Mamá! - Coasmedas	201	9-05-15	Link	
Coasmedas Cooperativa de ahorro y crédito	209836	1		0	0	Feliz Mes Súper Papás	201	9-05-31	Link	
Coasmedas Cooperativa de ahorro y crédito	198476	3		2	0	Sorteo Sobre Ruedas con Coasmedas	201	9-05-17	Link	
Grupo Coomeva	133042	0		2	2	CDT Premio Bancoomeva	201	9-11-26	Link	
Showing 1 to 5 of 75 entries						Previous	1 2 3	4 5	15	Next

## **Further Work**

- Predicción mejorada de polaridad: BERT, XLNet.
- → Alertas de menciones a la marca por parte de personas con alto potencial de influenciación: score de influenciación.
- → Clasificador de categoría de campaña de los competidores: tarjeta de crédito, viajes, seguro de auto, etc.
- → Historial: Guardar en un repositorio el histórico de las opiniones/publicaciones encontradas en cada ejecución.
- → Monitoreo 24/7/365: alertas en tiempo real.
- → Añadir otras Redes Sociales: Facebook, Instagram.

Gracias por su atención.

¿Preguntas?





