

# Análisis de Redes Sociales Twitter

Parametría

19/06/2020

---

El análisis de redes sociales es una buena herramienta para conocer y entender la opinión pública usuaria de redes sociales. Cabe señalar que, dado el contexto socioeconómico en nuestro país y la brecha tecnológica característica de los países en desarrollo, este grupo recoge características específicas que merece un tratamiento diferenciado de los trabajos demoscópicos realizados en vivienda o por teléfono.

En México solo 70 por ciento de la población tiene acceso a internet (80 millones de mexicanos). Este segmento es el más urbano, más escolarizado y con mayores recursos económicos. Por ello, no es representativo de la opinión pública en su conjunto. Por su perfil es muy probable que este segmento represente solo a la clase media. Sin embargo estos segmentos pueden anticipar frecuentemente lo que será un tema de conversación social más general en el futuro.

Del total de usuarios de internet casi 90 % usa redes sociales; un poco más del 85 % usa WhatsApp, ligeramente por abajo está Facebook (84 %). Youtube lo utiliza sólo una tercera parte de los usuarios de Internet (31 %), Instagram ligeramente menos (27 %). Finalmente, Twitter sólo 10 % de los usuarios de internet.

Entre el público que accede a redes sociales la mayor parte de ellos dedican su tiempo a temas como deportes o espectáculos (alrededor de 85 %). Es solo una minoría los que emplean las redes sociales para debatir los temas de vida pública. Twitter es probablemente la red social con mayor frecuencia de opiniones en vida pública. Por ello, nuestro análisis se enfoca en esta red social.

El presente trabajo se realiza con interfaces basadas en la API de Twitter, donde se filtra- en tiempo real- los tweets de interés para crear una base de datos propia analizable. A partir de este análisis se identifican los topics, hashtags y utilizadores más importantes para el seguimiento. Este análisis se basa en un scrapping personal de tweets en México que hacen mención de COVID-19.

# Tiempo

---

Los datos sobre tweets hasta 2020-06-18 10:35:57 tienen 580187 tweets.

El gráfico de tiempo presenta el número de tweets que mencionan el #COVID19MX. Estos tweets se muestran según la hora de su publicación desde el 30 de marzo al 18 de junio de 2020.

Es importante remarcar que los tweets tienen ciclos normales de creación determinado por las horas de descanso, así como la escalada esperada en los fines de semana. Este ejercicio permite evaluar la relevancia de un tema particular para la opinión pública usuaria de redes sociales y, subsecuentemente, entender su relación con eventos observados en un periodo de tiempo.

En la actualización de este reporte se incluyeron los tuits a partir del día 22 de mayo de 2020. Durante este periodo, se observa que entre los días 03 y el 04 de junio y 11 al 12 de junio hay un incremento en el número de tweets acumulados —sobre todo en el último periodo mencionado.

El primer punto, correspondiente a los días 03 y 04 de junio coincide con el anuncio de 1,092 defunciones al conteo de muertes por COVID-19. Si bien, estas defunciones no se contabilizaron en un solo día, éste fue el mayor incremento registrado a lo largo de la pandemia. Además, durante este periodo se superaron los 100 mil casos confirmados.

El segundo periodo, entre el 11 y 12 de junio podría tener varias explicaciones e implicaciones. Estos fueron los días previos a que se realizara una reapertura parcial de ciertas actividades económicas. Respecto a la comunicación oficial, las contradicciones entre el vocero de la pandemia, el subsecretario López-Gatell, y el presidente López Obrador pudieron haber contribuido a la confusión de la población: mientras el primero insistía en que las medidas de sana distancia y el quédate en casa debían continuar a pesar de la reapertura, el segundo hacía un llamado a que no se tuviera miedo de salir a la calle. En el plano local, el 12 de junio fue el día en el que se anunció la reapertura escalonada de algunas actividades en la capital del país —a pesar de que esta entidad concentraba la mayor proporción de casos activos y se encontraba en semáforo rojo. Las principales medidas anunciadas fueron la finalización del programa Hoy No Circula extendido, la reapertura total de estaciones de Metro y Metrobús y la reincorporación de las industrias manufacturera y aquellas relacionadas con actividades de construcción, minería y producción automotriz.

Otro elemento que vale la pena remarcar es que la tendencia respecto a este tema previo al 10 de junio tendía a la baja. Sin embargo, a partir de este día la discusión se reactivó a partir de un estímulo ocasionado por varios comunicados de autoridades públicas.

[Link a esta gráfica](#)

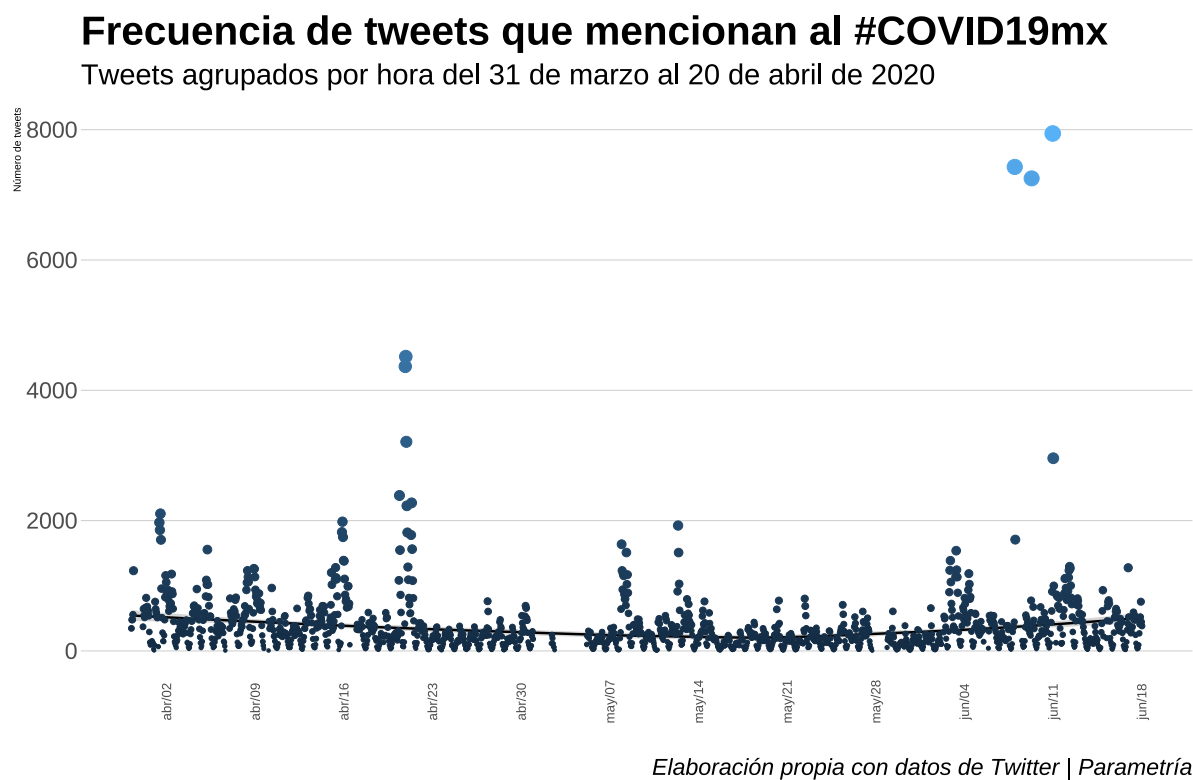


Figura 1: Línea de Tiempo de Tweets sobre COVID-19 en México - Twitter

## Los primeros días vs los últimos días

---

Más que una nueva normalidad, la discusión del COVID-19 en redes sociales es ya la normalidad. Si bien la frecuencia de tuits relacionados con el virus ha bajado considerablemente, como ya se presentó en la gráfica anterior, las palabras, menciones y temas relacionados con éste han tenido cierto dinamismo. Para ilustrar este ejercicio, se realizará una comparación entre los términos más mencionados en las primeras dos semanas de cuarentena y las últimas dos.

Los hashtags de las primeras dos semanas son encabezados por #QuédateEnCasa, término que no perdió el liderato en ningún momento del periodo de analizado, y #SusanaDistancia. Destaca que #Tlalnepantla se encuentra dentro de los primeros diez hashtags más frecuentes, debido a haber sido el lugar de uno de los primeros brotes hospitalarios. Además, términos como #CuandoTermineLaCuarentena retratan una especie de nostalgia de la abandonada normalidad.

En cambio, los hashtags del último mes se relacionan más con las actividades que ahora son cotidianas en el encierro: #PontePilaEnCasa, #YoConPontePila y #DeporteComunitario. Además, existe un boom en el reconocimiento del personal médico que se dedica a atender a pacientes COVID-19: #DíaDeLaEnfermería, #SalvaVidas y #GraciasPorCuidarnos.

La comparación de menciones a cuentas públicas entre ambos periodos de tiempo también arroja resultados interesantes. En las primeras dos semanas se observa que la Secretaría de Salud federal (SSalud) era la cuenta más mencionada; el subsecretario López-Gatell se hallaba en tercer lugar, después de la cuenta del presidente López Obrador. Además, varios funcionarios públicos federales —el canciller Ebrard, el director del IMSS Zoé Robledo y la secretaria de trabajo Luisa Alcalde— se encuentran en las primeras 20 posiciones debido a su rol central en la atención de distintas áreas de la epidemia. En materia local, sólo los gobernadores Enrique Alfaro (Jalisco) y Alfredo del Mazo (EdoMex) logran entrar dentro de las cuentas con mayor número de menciones.

En contraste, en las últimas semanas la cuenta más mencionada, como era de esperarse, es del vocero oficial del gobierno: Hugo López-Gatell; le sigue Ebrard y Obrador. En general, se tiende a mencionar menos las cuentas institucionales y se menciona más a las y los titulares de cada institución. También vale la pena destacar que varias cuentas relacionadas con el gobierno de la CDMX cobran relevancia; posiblemente debido a las intervenciones locales que se han hecho tanto en lo económico como en materia de salud pública en paralelo con lo establecido por el gobierno federal.

## Hashtags, Menciones y Temas

---

Las nubes de palabras de hashtags, menciones y temas presentan las palabras y alusiones con mayor número de menciones en los tweets publicados del 30 de marzo al 18 de junio, relacionados al #COVID19MX. El tamaño de cada palabra en la nube alude a la mayor o menor frecuencia en que fue mencionada en las publicaciones.

Este ejercicio permite conocer los verbatismos más asociados con el tema de análisis y, con ello, analizar los constructos sociales que se crean en redes sociales respecto al mismo. Asimismo, esta información permite identificar aquellas figuras públicas o autoridades, así como movimientos detrás de los hashtags, a los que más alude la población usuaria de redes sociales con el fin de comunicar su interés particular en un tema.

Hashtags	Tweets	Temas	Tweets	Arrobas	Tweets
#QuedateEnCasa	102913	COVID	433712	@hlgatell	122433
#SusanaDistancia	23824	SANA	158032	@lopezobrador_	73457
#Chimalhuacan	16292	DISTANCIA	152503	@ssalud_mx	50070
#CDMX	12885	HLGATELL	116689	@m_ebrard	16736
#Mexico	12642	MX	104426	@gobcdmx	16283
#MedicosEnPeligro	9096	COVIDMX	103356	@alfredodelmazo	13253
#EmergenciaSanitariaEdomex	7824	MÉXICO	82946	@edomex	11427
#SalvaVidas	7641	LOPEZOBRADOR	73911	@s	10906
#ExigimosInsumosDeCalidad	6994	CASOS	66229	@gobiernomx	8832
#DiaDeLaEnfermeria	6892	HOY	63143	@tu_imss	6944
#Fase3	6032	CONTAGIOS	62282	@susanadistancia	6661
#SanaDistancia	6021	NACIONAL	60651	@javier_alatorre	5178
#SNTEsalud	6011	QUÉDATEENCASA	57208	@zoerobledo	4966
#ULTIMAHORA	4953	COVIDMEXICO	55682	@enriquealfaror	4406
#Edomex	4594	JORNADA	52599	@metrocdmx	4181
#CuidemosTodosDeTodos	4469	MEDIDAS	50474	@mbarbosamx	3541
#EdomexEnSemaforoRojo	4302	SSALUD	50153	@ssaludcdmx	3399
#YoConPontePila	4063	SI	49318	@who	3174
#QuedateEnCasaYa	3988	GOBIERNO	47765	@felipecalderon	3043
#ProtegeteYProtegeALosDemas	3726	SALUD	42856	@imagentvmex	2789

La actualización de esta gráfica indica que las palabras relacionadas con las estrategias gubernamentales para prevenir la propagación del virus —tales como Sana Distancia— siguen vigentes al día de hoy. La palabra “fase” se incorpora al top 10, debido a la entrada a la fase 3 en el país.



6

## Hashtags

---

Definición: Es una etiqueta de metadatos precedida de un carácter especial (normalmente un numeral #) con el fin de ser identificada de forma rápida. Se usa en servicios web tales como Twitter, Telegram, FriendFeed, Facebook, Google+, Instagram, Weibo para señalar o hacer hincapié en un tema sobre el que gira cierta conversación.

Para este periodo de tiempo, la discusión pública en Twitter es dominada por aquellos hashtags relacionados con las recomendaciones gubernamentales para contener el contagio del virus (1. #QuedateEnCasa, 2. #SusanaDistancia). Sin embargo, el primer hashtag ya cuatruplica en número de menciones al segundo. La insistencia de que las personas no abandonen sus hogares, a pesar de la reapertura, para evitar la propagación del virus se ha convertido en la principal estrategia de comunicación política. En este tema destaca que Chimalhuacán se posicione como el tercer hashtag más mencionado, debido a un brote de COVID-19 en el Hospital Materno Infantil del municipio.



Figura 3: Menciones más comunes en Tweets sobre COVID-19 en México - Twitter

Definición: Es una etiqueta de metadatos precedida de un carácter especial (normalmente una arroba @) con el fin de relacionar a un usuario específico en una conversación y notificarlo de la misma. Se usa en servicios web tales como Twitter, Telegram, FriendFeed, Facebook, Google+, Instagram, Weibo para establecer relación entre una conversación y un usuario o para marcar la respuesta hacia un mensaje específico de un usuario. Resulta interesante observar que la discusión pública en redes se relaciona fuertemente con las autoridades correspondientes. Para el periodo de tiempo analizado, el subsecretario Hugo López Gatell se coloca en la primera posición con 72,376 menciones. Ni siquiera el presidente, quien se posiciona en segundo lugar con un total de 42,496, se le acerca en magnitud. Esta amplia socialización como el principal responsable durante la pandemia, podría jugarle para bien o para mal al subsecretario, dependiendo de los resultados que las y los ciudadanos perciban cuando la emergencia pase.



Link a esta gráfica



## Análisis de Sentimiento

El gráfico de Análisis de sentimiento presenta un promedio de la orientación positiva o negativa de los tweets publicados del 30 de marzo al 18 de junio, relacionados al #COVID19MX.

Esta gráfica podría clasificarse en tres etapas: la incertidumbre de los primeros días, el optimismo generado por la promesa de una nueva normalidad, y la desilusión porque esta nueva normalidad no cumplió con las expectativas esperadas. Al día de hoy, vivir en durante una pandemia global se ha normalizado, pero esto no implica que necesariamente lo asumamos con gusto. Nuestros hábitos han cambiado, como lo muestra nuestra última medición telefónica nacional, pero la nostalgia de la vida antes de pandemia es algo que, al menos en el sentimiento de las y los tweeteros, se puede observar.

El texto de cada tweet puede ser positivo una vez que menciona palabras con este carácter, tales como: bueno, recuperación, cura, entre otras.

Por el contrario, el texto de cada tweet puede ser negativo una vez que menciona palabras con este carácter, tales como: malo, muerte, tristeza, entre otros. Este sistema automatizado permite asignar un puntaje para cada palabra en los tweets de la base de datos, lo cual obtiene un puntaje general por tweet. Con estos puntajes se obtiene un promedio de puntuación de los tweets por hora, y subsecuentemente construir el “sentimiento promedio” del día.

El puntaje se posiciona en una escala de números reales  $(-\infty, +\infty)$ , donde los valores positivos significan un “sentimiento promedio positivo”, los números negativos significan un “sentimiento promedio negativo” y el número cero significa un “sentimiento promedio neutral”. Cada círculo en el gráfico representa un día y su tamaño representa el volumen de tweets encontrados.

### Análisis de sentimiento en tweets de COVID19

Cada círculo representa un día; el tamaño del círculo indica la cantidad de tweets encontrados por día. Una puntuación mayor a cero representa un sentimiento promedio positivo; una menor, un negativo.

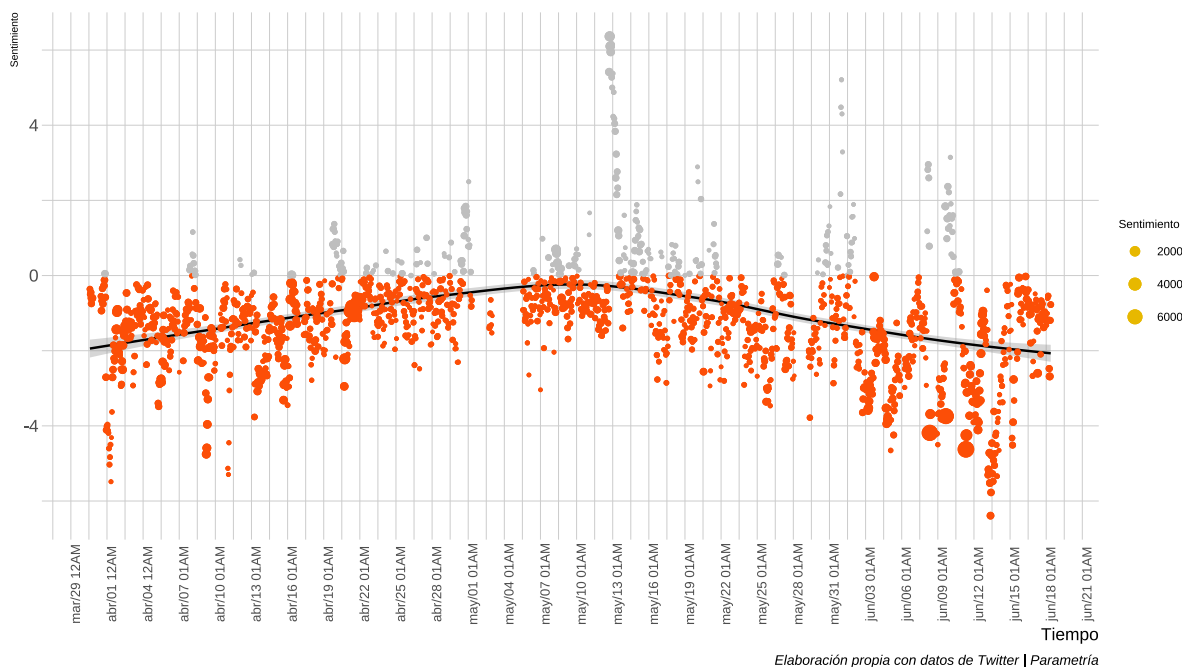


Figura 5: Análisis de Sentimiento en Tweets sobre COVID-19 en México - Twitter

[Link a esta gráfica](#)

## Diez palabras positivas y 10 negativas

Ejemplo de Diccionario para Sentimiento

Negativas	Puntuación	Positivas	Puntuación
bastardo	-5.0	excepcional	5
negro	-5.0	magnífico	5
perro	-5.0	asombroso	4
pinchazo	-5.0	celestial	4
tragón	-5.0	eufórico	4
puto	-4.7	exuberante	4
burro	-4.0	fabuloso	4
catastrófico	-4.0	fantástico	4
culo	-4.0	ganador	4
defraudador	-4.0	gracioso	4

## ¡Emojis!

Debido a la propia extensión de un tweet, una forma no estructurada de expresión reside en las figuras que pueden acompañar el texto. Esta información se muestra en el gráfico de Emojis. Este gráfico presenta el número de veces en que el emoji específico fue usado en los tweets publicados del 30 de marzo al 18 de junio, relacionados al #COVID19MX. Este ejercicio es un primer acercamiento a entender la expresión no estructurada como un termómetro de la opinión pública en redes sociales en un tema específico.

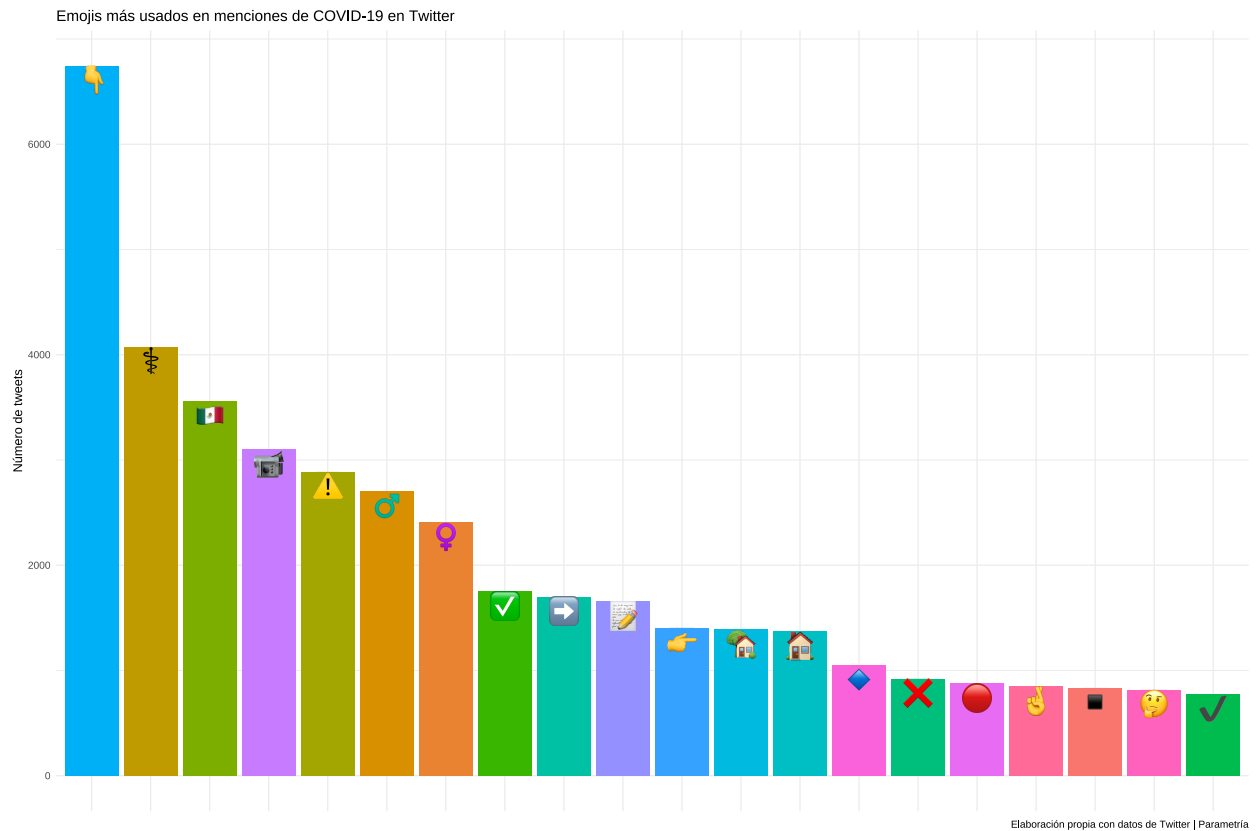


Figura 6: Emojis más comunes en Tweets sobre COVID-19 en México - Twitter

[Link a esta gráfica](#)

Descargue los archivo RMarkdown para crear este reporte aquí [Rmarkdown](#).

Consulta el artículo publicado en Milenio por nuestro director, Francisco Abundis [aquí](#)