

Análisis de Redes Sociales Twitter

Parametría

07/04/2020

Análisis de Sentimiento y Temas

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. Sin embargo, dada la continua expansión de las redes sociales y su más frecuente uso como medio de información hace imprescindible su mayor análisis. Sin duda, hoy en día, el tweet es la forma de expresión más rápida, en tiempo real y efectiva usada por un importante grupo de población.

Etapas de Análisis

Iniciando con una extracción directa de todos los tweets donde hay menciones o utilizan las palabras clave predeterminadas creamos un diagnóstico preliminar. Esta descarga se realiza con interfaces basadas en la API de Twitter, donde filtramos en tiempo real los tweets de interés para crear una base de datos propia analizable.

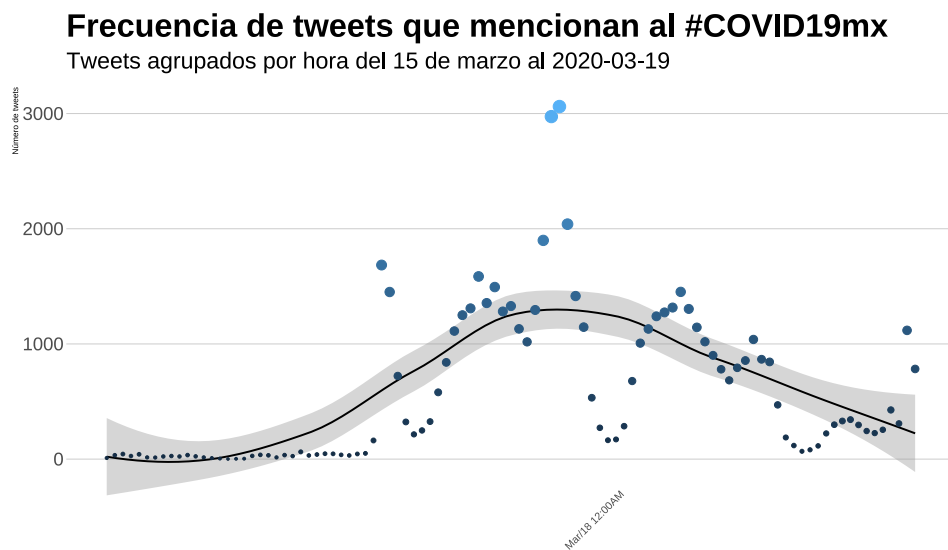
A partir de este análisis se identifican los topics, hashtags y utilizadores más importantes para el seguimiento. Para Parametría es poco relevante los seguimientos a tiempo real, dado que es imposible ex ante conocer los futuros temas o tendencias coyunturales. Este análisis se basa en un scrapping personal de tweets en México que hacen mención de COVID-19.

Los datos sobre tweets hasta 2020-03-19 16:09:59 tienen 59821 tweets y 23 variables.

Tiempo

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 15 al 19 de marzo de 2020. La imagen inferior que acompaña el gráfico permite elegir la temporalidad específica de su interés.

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.



Elaboración propia con datos de Twitter | Parametría

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

[Link a esta gráfica](#)

Los gráficos de Hashtags, menciones y temas presentan las palabras y alusiones con mayor número de menciones en los tweets publicados del 15 al 20 de marzo de 2020, 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.

Esta gráfica indica que la palabras relacionadas con las consecuencias del virus —tales como cuarentena, efecto y casos— son recurrentes en la discusión pública. También cabe destacar que “hoy” fue la palabra más mencionada —con un conteo total de 8,695—, lo anterior sugiere que las y los usuarios de Twitter están preocupados por el seguimiento diario de la evolución del tema.

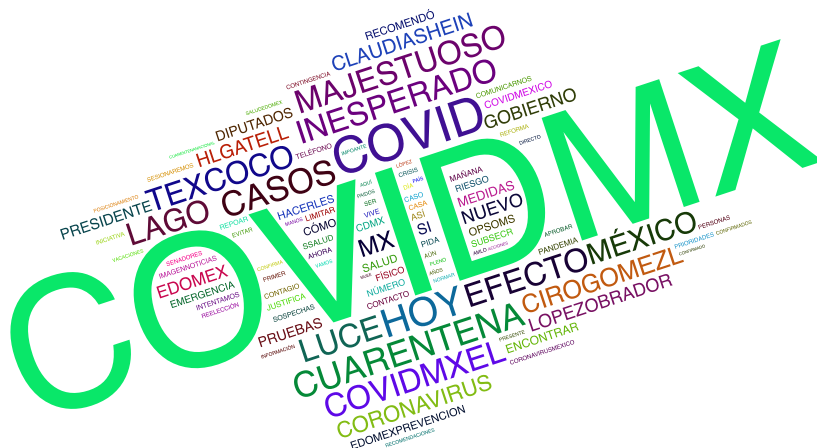
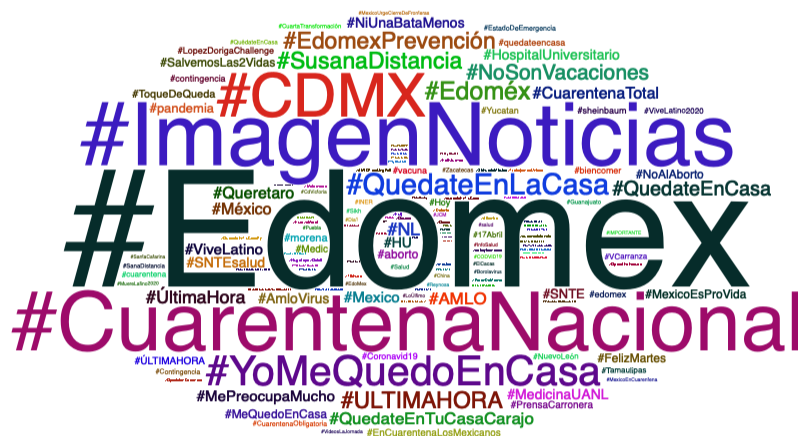


Figura 2: Menciones más comunes en Tweets sobre COVID-19 en México - Twitter

Los hashtags que se relacionaron con mayor frecuencia a los tweets hacen referencia al #COVID19mx pueden dividirse en tres grandes conglomerados. Primero, aquellos relacionados con algunas entidades federativas que ya han implementado estrategias específicas. Segundo, los hashtags relacionados con las recomendaciones para contener el contagio del virus (CuarentenaNacional, QuedateEnCasa). Finalmente, términos comúnmente usados por agencias de noticias para promover su contenido.



7

Menciones

Resulta interesante observar que la discusión pública en redes se relaciona fuertemente con las autoridades correspondientes. En particular, llama la atención que la jefa de gobierno Claudia Sheinbaum (4,342) y el subsecretario Hugo López-Gatell (4,292) superen la frecuencia de menciones del presidente López Obrador (3,923). Lo anterior pudiera ser indicativo de una atribución de responsabilidad sofisticada por parte de las y los usuarios de Twitter: no se responsabiliza o asocia en su totalidad el manejo de la crisis al jefe del ejecutivo, sino que se relaciona el manejo de la crisis con autoridades locales y/o especializadas en salud. Adicionalmente, este comportamiento es congruente con la misma estrategia de comunicación que el gobierno ha implementado en este tema. Por otro lado, el comunicador Ciro Gómez-Leyva fue puntero en la frecuencia de menciones con un total de 5,631.

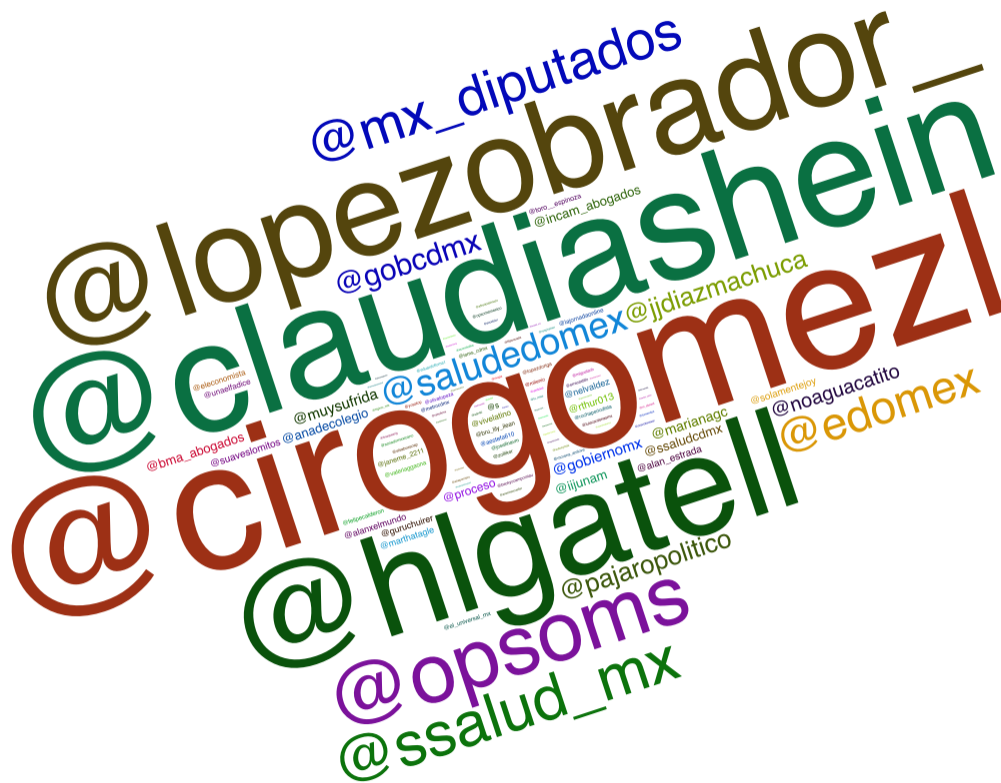


Figura 4: Menciones más comunes en Tweets sobre COVID-19 en México - Twitter

Link a esta gráfica

Análisis de Sentimiento

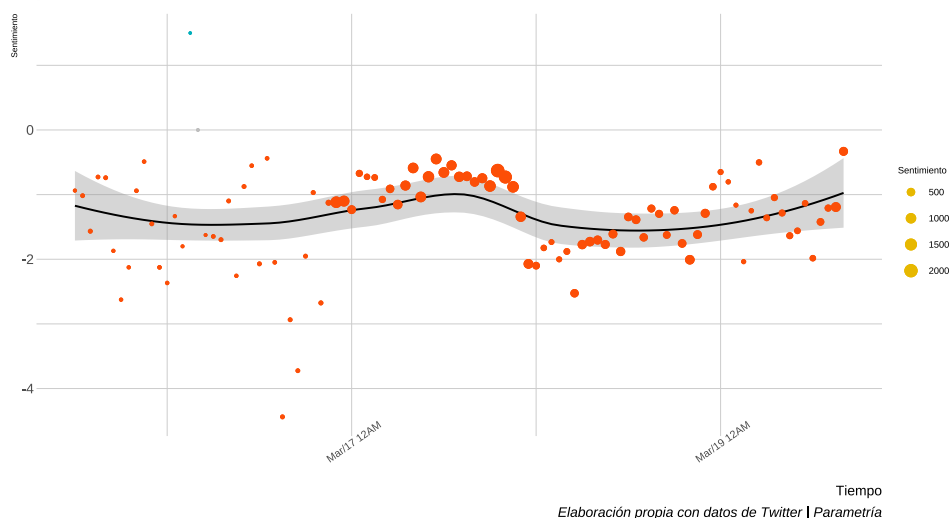
El gráfico de Análisis de sentimiento presenta un promedio de la orientación positiva o negativa de los tweets publicados del 15 al 20 de marzo de 2020, relacionados al #COVID19MX. El texto de cada tweet puede ser positivo una vez que menciona palabras con este carácter, tales como: bueno, recuperación, 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.

La imagen inferior que acompaña el gráfico permite elegir la temporalidad específica de su interés. Este tipo de ejercicio permite completar el primer análisis de lenguaje no estructurado con el fin entender los constructos y sentimientos que acompañan los textos de la opinión pública en redes sociales en un tema específico.

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.



Elaboración propia con datos de Twitter | Parametría

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

[Link a esta gráfica](#)

¡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 15 al 20 de marzo de 2020, 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.

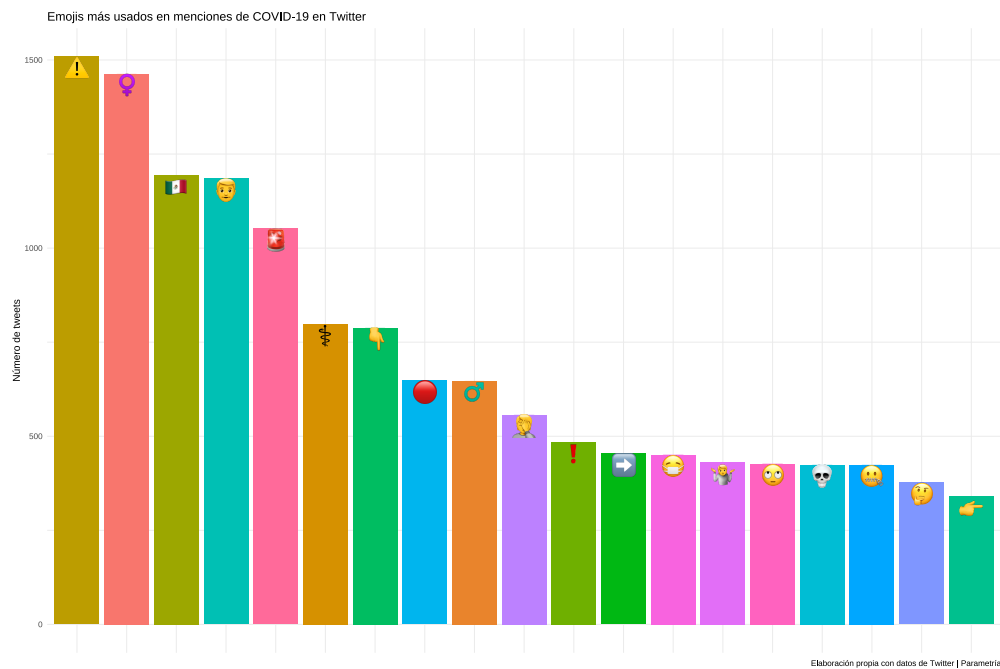


Figura 6: Emojis más comunes en Tweets sobre COVID-19 en México - Twitter

Link a esta gráfica

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