

Exploración de indicadores que reflejen la incidencia de COVID-19 en Venezuela usando fuentes de datos no oficiales

Eliana Bonalde¹, Alfonso Navas²

¹Universidad Industrial de Santander

²Universidad Nacional de Colombia

20 de mayo de 2021

Índice

1. Introducción	2
2. Metodología	2
2.1. IHME	2
2.2. GoFundMe	3
2.3. Twitter API	8
2.4. Google Trends API	11
3. Resultados y Análisis	12
4. Conclusiones y Recomendaciones	16

Resumen

El objetivo principal de este trabajo es hacer una comparación mediante series de tiempo, correspondientes al periodo comprendido entre marzo de 2020 y mayo de 2021, entre datos numéricos oficiales del virus COVID-19 en Venezuela y datos no oficiales extraídos de las plataformas de GoFundMe, Twitter y Google Trends. Tomaremos como datos oficiales los estimados por el modelo del Instituto de Métricas y Evaluación de la Salud (IHME, por sus siglas en inglés), ya que según estimaciones de organizaciones nacionales e internacionales las cifras oficiales reportadas por el gobierno de Venezuela no corresponden a la realidad. Por medio del análisis de los datos extraídos, se muestran los mejores indicadores que reflejan la incidencia de Covid-19 en Venezuela. Python es el lenguaje de programación utilizado para extraer, almacenar, visualizar y analizar estos resultados, los cuales van a permitir tener una idea de la realidad que se está viviendo en Venezuela con respecto al virus.

1. Introducción

A un año del inicio de la pandemia de COVID-19, sigue avanzando por todo el mundo con millones de muertos e infectados. El IHME, centro de investigación de salud global independiente de la Universidad de Washington, señala que hay falta de información y anomalías de informes presentes en las estadísticas del virus que se informan a diario [1], por ejemplo, no todos los casos tienen acceso a pruebas confirmatorias, ya que se hacen pruebas especialmente a personas con síntomas específicos o cierta gravedad de la enfermedad. En Venezuela, la situación es mucho peor, la actual crisis que vive, las pocas pruebas confirmatorias que se han realizado por millón de habitantes y la posible manipulación de los datos han generado la incertidumbre en torno al número real de casos [2].

Actualmente, el impacto de Internet y las redes sociales nos facilita la interacción a nivel mundial, y debido a esto se puede usar minería de datos y valerse de esta valiosa herramienta para extraer información relevante para su posterior análisis, que nos permita obtener indicadores de la incidencia del virus en Venezuela.

Por un lado, la crisis en Venezuela en todos los sectores, especialmente en el de salud, ha llevado al venezolano a la necesidad de pedir ayuda económica externa para poder adquirir suministros y atención médica. El crowdfunding, aprovecha el poder de las redes sociales e Internet para proporcionar medios de recaudación de fondos. Una de las plataformas de preferencia es GoFundMe [3], se ha convertido en una forma de financiación alternativa para familias con gastos médicos que son incapaces de asumir por cuenta propia [20], esto se evidencia en las redes sociales donde se encuentran numerosas campañas de ayuda.

Por otro lado, la red social Twitter [4] es una fuente importante para el marketing, los negocios y la investigación, muchas personas expresan, por medio esta plataforma, estados, sentimientos y opiniones sobre cualquier tema, y parte de esta data se puede obtener por medio de las interfaces de programación de aplicaciones (API, por sus siglas en inglés) [5], las cuales facilitan la extracción de cualquier información que sea pública en tiempo real y a través de la historia en Twitter. Por último, diariamente hay miles de solicitudes a los buscadores web para encontrar información sobre cualquier tema, hoy en día Google Trend [6] permite el acceso a las tendencias de búsqueda en ese buscador desde el año 2004 hasta 36 horas antes de la búsqueda y, debido a que se gestionan millones de búsquedas diariamente, se puede hacer un análisis confiable.

El presente trabajo se divide en 4 secciones. La primera corresponde al análisis de los datos obtenidos del sitio web de GoFundMe. En la segunda y tercera sección se trabaja la obtención y análisis de las publicaciones de Twitter y las tendencias de búsqueda de Google Trend, respectivamente. Finalmente, se hace un análisis comparativo de los resultados obtenidos con los datos presentados por el IHME.

2. Metodología

2.1. IHME

A finales de enero de 2021, el IHME lanza un nuevo modelo para el pronóstico de COVID-19, éste toma en consideración tres métricas: infecciones, hospitalización de infecciones, y muertes de infecciones. A partir de ellas, genera una estimación sobre el número de infecciones, también toma en

consideración las estadísticas de seroprevalencia y utiliza métodos para predecir la tasa de detección de infectados (IDR), la tasa de infección-hospitalización (IHR) y la tasa de infección-letalidad (IFR) en entornos sin encuestas de seroprevalencia. Este modelo genera estimaciones para muchos países del mundo, incluyendo Venezuela. Mayores detalles del modelo de IHME se pueden encontrar en [8].

Los datos se decargaron de la página de Our World in Data el 11 de mayo de 2021 [9], quienes toman como referencia los datos del IHME. Otros modelos, el ICL (Imperial College London) y el LSHTM (The London School of Hygiene & Tropical Medicine), por ejemplo, asumen que el número de muertes confirmadas es igual al número real, éste no es el caso en Venezuela, como mencionamos anteriormente, y por lo tanto, subestiman los valores reales. El modelo YYG (Youyang Gu) fue creado y optimizado para EE.UU, para otros países las estimaciones del modelo pueden ser menos precisas. El modelo IHME se ajusta más a la realidad, adicionalmente, se sigue actualizando continuamente desde sus inicios hasta la actualidad.

La Figura 1 muestra gráficas de series de tiempo de los datos dados por Our Word in Data para Venezuela, quienes también proveen los datos oficiales del gobierno de Venezuela. Los casos confirmados de infectados representan entre el (18 %) al (24 %) de los casos estimados por IHME [10]. El comportamiento cualitativo de los picos de infecciosos concuerdan tanto en los casos reportados como estimados, sin embargo, estos últimos son mucho mayores, especialmente el segundo pico. también se observa que las franjas de incertidumbre (95 %) de las estimaciones son mucho más grandes en estos picos. Debido a esto, ratificamos la desconfianza en los datos oficiales.

2.2. GoFundMe

Los datos de las campañas de GoFundMe son producto de una colaboración con Harvey Maddocks y pueden encontrarse en [11]. Estos datos fueron colectados mediante técnicas de webscrapping. A continuación, se expondrán brevemente los pasos que se siguieron en este procedmiento:

1. Como es común en muchos sitios web, GoFundMe tiene un motor de búsqueda interno que permite al usuario consultar campañas de crowdfunding a partir de palabras clave. Para usar esta herramienta de la mejor manera posible, se propusieron las palabras clave que nos pudieran conducir al mayor número de campañas de interés. En nuestro caso, las palabras usadas fueron: “Venezuela”, “Covid”.
2. Considerando el formato de url del sitio web de GoFundMe, se crea la dirección correspondiente a la primera página de la búsqueda de las palabras clave escogidas en el punto anterior.
3. Posteriormente, usando el navegador virtual Selenium [22], compatible con el lenguaje de programación Python, fue posible acceder a la página web de GoFundMe usando la url creada en el punto anterior. Esta herramienta permite además extraer el archivo HTML correspondiente a las páginas web.
4. Luego, se utilizó la biblioteca Beautiful Soup [23] para extraer los datos de los archivos HTML de las páginas web. Por ejemplo, al usar esta función sobre las páginas del resultado de búsqueda, se pudo crear una lista con todas las direcciones url de los sitios de las campañas que aparecían como resultado.

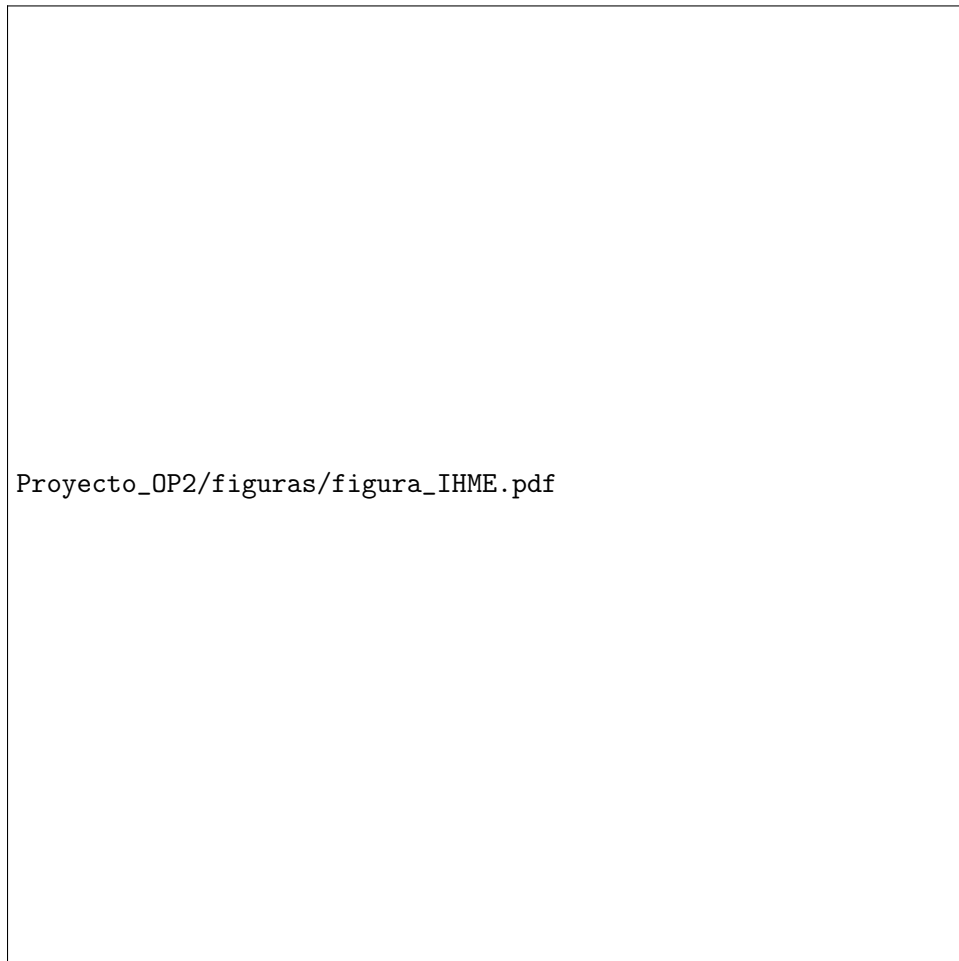


Figura 1: Estimaciones del número real de infecciones. Las líneas superior e inferior muestran los límites de un intervalo de incertidumbre del 95 %

5. Entonces, usando las bibliotecas Selenium, para acceder a las páginas web de las campañas, y BeautifulSoup, para recolectar datos; se recorrió la lista de direcciones url y se extrayeron los datos de interés para cada una de ellas en esta página. En este caso, la información que se recolectó de cada campaña fue:

- La dirección url de la página donde se encuentra alojada.
- El título.
- las etiquetas (tags), o la categoría a la que pertenece cada una.
- El monto actual recaudado.
- El monto total a recaudar o meta de recaudación.
- La descripción.
- La fecha de lanzamiento.
- El país de origen.
- El número de donaciones.
- Y una columna indicando si la campaña se trata de caridad o no.

Estos datos se iban añadiendo en una tabla de tal forma que cada fila era una campaña diferente.

6. Por último, se itera este procedimiento sobre todas las páginas del resultado de búsqueda de las palabras clave seleccionadas.

Como se menciona en la sección anterior, del proceso de webscrapping se obtuvo una tabla con 11 columnas y 2042 filas. Los nombres de las columnas, que corresponden a cada una a una de las variables mencionadas en el punto anterior, son respectivamente: 'url', 'title', 'tags', 'current amount', 'total_amount', 'description', 'created', 'launch_date', 'country', 'donation_count', 'is_charity'.

Ahora bien, antes de realizar el análisis, la tabla de datos crudos pasó por un proceso de limpieza. A continuación, se describirá brevemente las etapas de este procedimiento:

1. Se garantizó que todas las variables estuvieran en el formato adecuado. Por ejemplo, que la columna 'launch_date' (fecha de lanzamiento) estuviese en formato de fecha; o que la columna 'donation_count' (recuento de donaciones) tenga formato numérico.
2. Se seleccionaron datos cuya fecha sea posterior al 1 de enero del 2020.
3. Luego, por motivos éticos, se eliminaron de la tabla de datos las columnas 'url', 'title', 'description', ya que ponían en riesgo la privacidad de las personas involucradas en las campañas.
4. Por último, al considerar que en las columnas 'donation_count' y 'current_amount' el dinero se expresa en monedas diferentes, se implementaron rutinas con las que se pudo:
 - Colocar el valor numérico del dinero y el tipo de moneda usada en columnas diferentes.

- Se añadieron dos columnas 'current amount in dollars' y 'total amount in dollars', en donde se transformaron todos los montos monetarios a dólares. En este paso se usó la biblioteca `currency converter` [24].
- Aprovechando que las entradas en 'current amount in dollars' y 'total amount in dollars' entre todas las campañas son comparables, se agregó una columna: 'percentage of completion', la cual cuantifica cuanto ha avanzado la campaña en términos monetarios.

El notebook que presenta todo el código utilizado para la exploración y limpieza de los datos obtenidos lo encontramos en [12].

Para entender mejor los datos de GoFundMe, se hicieron gráficas en donde se puede observar la distribución de las variables. La descripción detallada de cada una de estas gráficas se encuentra en [13]. Aquí, discutiremos los resultado más relevantes. Por un lado, vemos que, como el covid es una enfermedad, no sorprende que la inmensa mayoría de campañas que concuerdan con las palabras de búsqueda "Venezuela", "Covid" estén en la categoría *Medical, Illness & Healing*. Sin embargo, debe considerarse que, al hacer una revisión manual se encontró que no todas las campañas que estaban en esta sección se hallaban directamente relacionadas a casos de covid-19. Adicionalmente, entre los datos había campañas que sí estaban relacionadas a casos de COVID-19 bajo otros tags, por ejemplo, *Accidents & Emergencies*. Lo anterior representa una limitación de los datos, la cual puede superarse mediante incorporación de técnicas de minería de textos, ya que a través de éstas se podría identificar las campañas asociadas directamente a casos de COVID-19, lo que a su vez permitiría la construcción de un indicador más robusto.

Por otra parte, para las distribuciones de las variables numéricas: 'monto actual', 'monto total', 'porcentaje de avance' y 'recuento de donaciones'; se encontró en todos los casos una asimetría positiva con una moda muy cercana a cero y la presencia de algunas campañas con valores relativamente altos para las variables que se distribuyen.

Por último, como se buscan indicadores que reflejen el estado de la pandemia, nos fijamos especialmente en el número de las campañas lanzadas diariamente. Para esta variable, se presentan las series de tiempo de 3 subconjuntos distintos de los datos. Entre los 4 subconjuntos que se presentan, se observa que los datos de la categoría *Medical, Illness & Healing* son los que guardarían una mejor correlación con el número de casos de la pandemia en Venezuela. Por tal motivo, escogemos solamente a este subconjunto como indicador.

- Todas las campañas.
- Las campañas con las etiquetas: *Medical, Illness & Healing*.
- Las campañas con las etiquetas: *Accidents & Emergencies*.
- La unión de los dos subconjuntos anteriores: *Medical, Illness & Healing + Accidents & Emergencies*

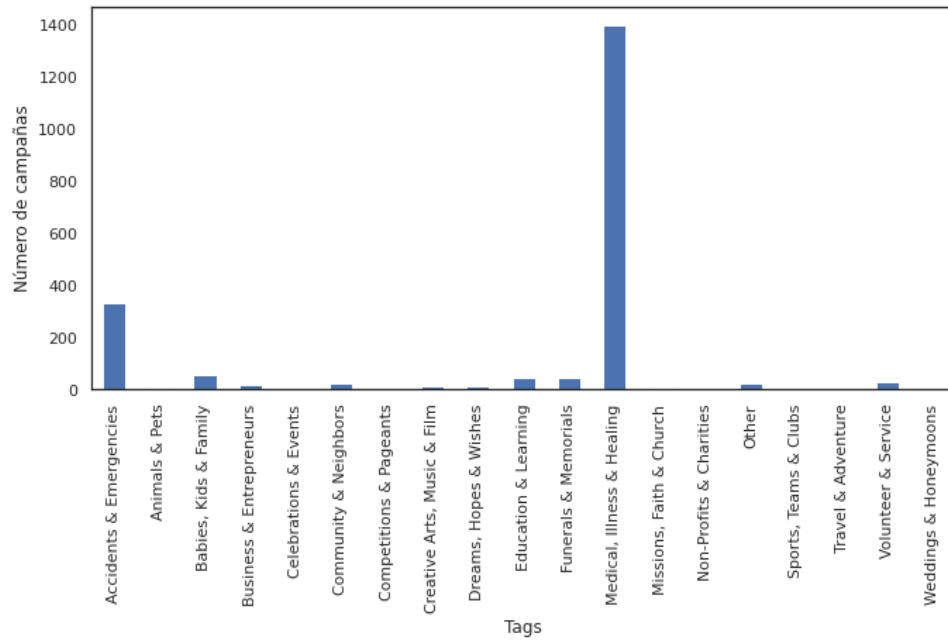


Figura 2: Número de campañas de GoFundMe por categorías

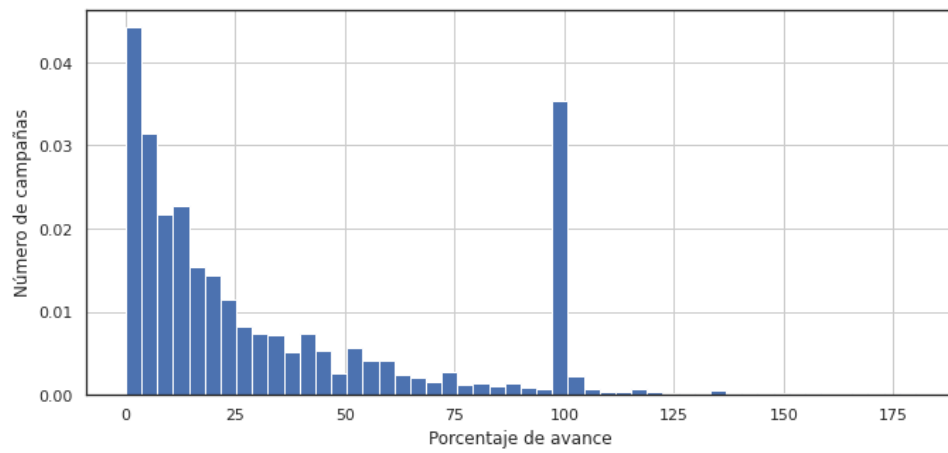


Figura 3: Distribución del porcentaje de avance (i.e monto actual sobre monto total) de las campañas de GoFundMe

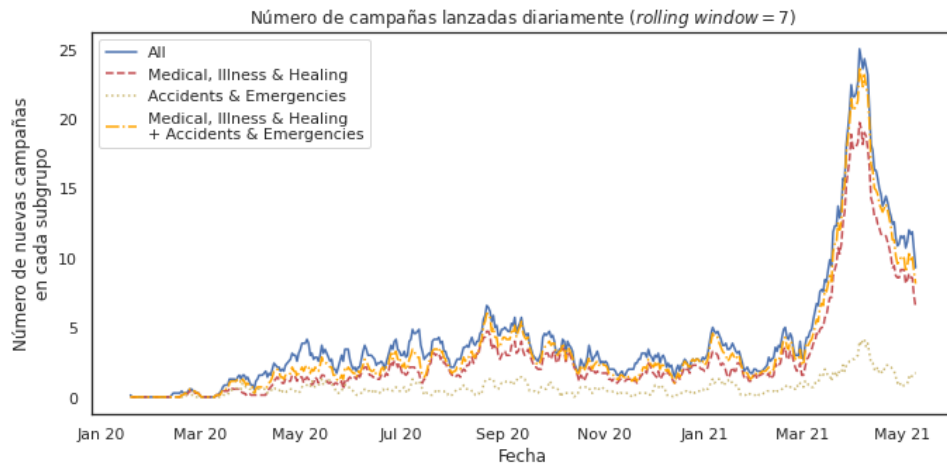


Figura 4: Número de campañas de GoFundMe lanzadas diariamente en los subgrupos seleccionados

2.3. Twitter API

Los datos de Twitter se obtienen con el uso de una API dedicada que provee los tweets públicos del archivo completo de Twitter (desde marzo 2006). La API permite hacer búsquedas relacionadas a palabras claves y operadores que hacen coincidir unos atributos al tweet.

La consulta está comprendida entre marzo de 2020 hasta el 10 de mayo de 2021, ésta coincide con con las palabras claves:

- covid-19/coronavirus
- infectado covid-19
- muerte covid-19
- (bombona/fluómetro/saturación) de oxígeno
- dexametasona
- GoFundMe venezuela covid-19
- recolectar dinero
- ayuda covid-19
- tratamiento covid-19
- UCI covid-19
- servicio público covid-19

Las consultas se especificaron con y sin hastahtag(`#`), ya que con este operador se realiza una coincidencia exacta; las palabras claves se especificaron con y sin tilde diacrítica, debido a que hay distinción entre las palabras que la posean y las que no, no obstante, no se hace distinción entre mayúsculas y minúsculas. Se filtraron todos aquellos tweets que eran retweets y, posteriormente, se eliminaron los repetidos. Adicionalmente, la búsqueda de todas las palabras claves, a excepción de las relacionadas con "GoFundMe", se hizo local, es decir, se restringió la búsqueda al país [18].

Los operadores en la búsqueda son:

- `-is:retweet`. Negación del operador `is:retweet`, usada para buscar solo tweets originales.
- `place_country:VE`. Usada para obtener tweets proveniente del país Venezuela.
- `has link`. Con link adjunto.

Las consultas con los operadores `'place_country :VE'` y `'-is:retweet'` son:

- "covid-19 OR coronavirus"
- "infectado covid-19 OR muerte covid-19"
- "oxígeno OR dexametasona"
- "bombona de oxígeno OR dexametasona"
- "flujómetro de oxígeno OR saturación de oxígeno"
- "flujometro de oxigeno OR saturacion de oxigeno"
- "#flujmetroOxigeno OR #saturacionOxigeno "
- "recolectar dinero covid-19 OR ayuda covid-19"
- "#RecolectarDinero covid-19 OR covid-19"
- "#ServicioPúblico covid-19 OR #ServicioPublico covid-19"
- "servicio público covid-19 OR servicio publico covid-19"
- "ServicioPublico covid-19 OR ServicioPúblico covid-19"
- "servicio publico OR servicio publico"
- "UCIS OR unidad de cuidados intensivos"
- "UCIS "
- "tratamiento covid-19"

Las consultas con el operador `'has link'` son:

- “GoFundMe venezuela covid-19 OR GoFundMe venezuela covid-19”
- “#GoFundMe venezuela covid-19 OR #GoFundMe venezuela covid-19”

El Notebook que contiene todo el código utilizado para la extracción de datos de Twitter se puede conseguir en [15].

Las políticas y acuerdos para desarrolladores de las API de Twitter [19], señalan que se permite hacer análisis de su contenido siempre y cuando no se almacenen y presenten datos personales, debido a esto, de toda la información que se recauda, se presenta solo la correspondiente al cuerpo del twitter “text” y la fecha de publicación “created_ad”.

La consulta de las palabras “covid-19” y “coronavirus”, es la que más tweets da por día, tan solo para el primer día del periodo arrojó 430 tweets, sin embargo, se descarta debido a que estos tweets están relacionado a cualquier noticia o comentarios sobre el COVID-19 desde el inicio de la pandemia, no necesariamente sobre infectados.

El resto de consultas se divide en 6 grupos, con un total de tweets recaudados de 3307. (ver Tabla 1).

Cuadro 1: Clasificación de las consultas hechas en Twitter con sus respectivos números de tweets

Grupo 1	Palabras claves 1	Tweets recaudados
1	infectado covid-19 muerte covid	365
2	(bombona/fluómetro/saturación) de oxígeno dexametasona	295
3	GoFundMe venezuela covid-19 recolectar dinero covid-19 ayuda covid-19	757
4	servicio público	1409
5	UCI unidad de cuidados intensivos	77
6	tratamiento covid-19	325

Inicialmente, se hizo la consulta sobre la palabra “oxígeno”, la cual no fue un buen indicador; ésta fue cambiada por “bombona/fluómetro/saturación de oxígeno”. “GoFundMe venezuela covid-19” con el operador ‘has link’ arroja en su mayoría todos los tweets que son de campañas de apoyo para enfermos de COVID-19. Esta consulta no se hizo en el país Venezuela debido a que es muy difícil abrir una campaña sin tener una cuenta en el exterior [20]. Las consultas relacionadas a “servicio público covid-19” fue la que más tweets arrojó, se esperaba que el número de publicaciones en la franja de abril/mayo de 2021, donde se encuentra el pico más alto de infecciones, según el modelo IHME, fuera más alto, ya que en Venezuela se suele usar la palabra clave “servicio público” para ayudas médicas en redes sociales.

El número de tweets extraídos reflejan posibles casos de infección, necesidad de insumos y ayuda económica para luchar contra el COVID-19. En las Figuras 5 y 6 se observa el comportamiento individual por grupo y general, respectivamente. Los picos guardan relación a los casos estimados,

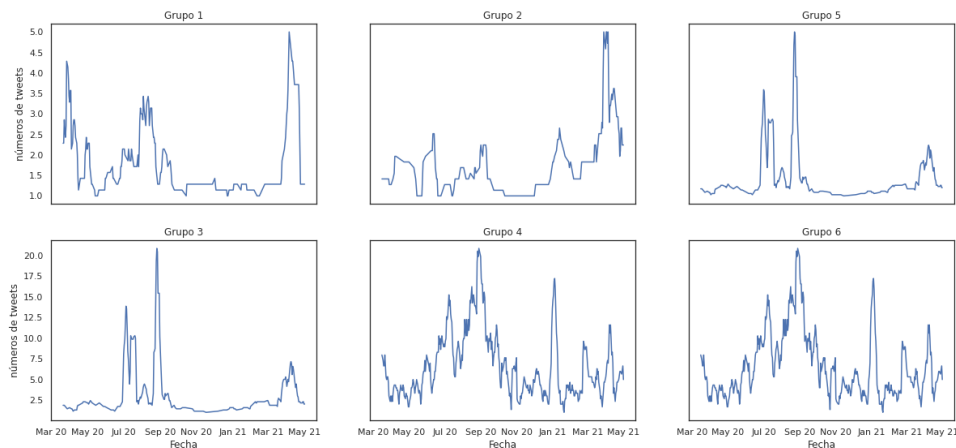


Figura 5: Comportamiento de los grupos de palabras claves en Twitter

sin embargo, éstos no se encuentran en la misma proporción. Hay que tomar en cuenta que las publicaciones de Twitter pueden ser un buen indicador del estado de la pandemia, pero hay que corregir las consultas y los tiempos de ejecución.

El Notebook que contiene todo el código utilizado para el análisis y la construcción de las series de tiempo de los datos extraídos de Twitter se puede conseguir en [14].

2.4. Google Trends API

Las consultas se hacen por medio de pytrend, una API no oficial de Google Trend que permite una interfaz simple para automatizar los informes de descargas [21]. Entre los métodos que provee se pueden encontrar: 'interés a lo largo de tiempo', 'interés histórico por horas', 'interés por región', 'tendencias de búsqueda', 'consultas relacionados', etc. Utiliza parámetros tales como: 'palabras claves', 'categorías', 'país', 'tiempo' y 'grupos' (imágenes, noticias, etc.).

Las palabras claves utilizadas son: “oxígeno”, “GoFundMe”, “ayuda económica”, “servicio público”, “tratamiento covid-19” y “dexametasona”. Éstas se especificaron con y sin tilde diacrítica, arrojando distintos resultados en ambos casos, sin embargo, no hay distinción entre mayúsculas y minúsculas. Todas las búsquedas se hicieron en el país Venezuela y en el intervalo de tiempo de 15 de mayo de 2020 hasta 15 de mayo de 2021.

Los tipos de solicitudes que se hacen son:

- 'related_queries', la cual muestra las búsquedas relacionadas a las palabras claves, donde se les asocian valores en una escala del 0 al 100. El número 100 le corresponde a la tendencia más alta y a medida que va disminuyendo el número, va disminuyendo la tendencia.
- 'interest_over_time', representada en una serie de tiempo que nos muestra el cambio de popularidad con el tiempo; los valores representan el interés de búsqueda relativo al punto más alto en el gráfico, siendo 100 el pico de popularidad.

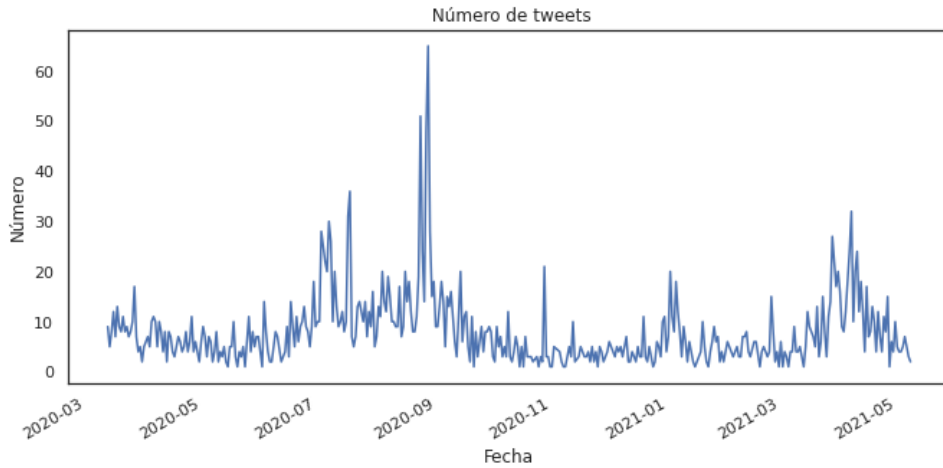


Figura 6: Número de total tweets recaudados

Las consultas más significativas fueron las palabras claves: “oxígeno” (sin tilde), “GoFundMe”, “tratamiento covid” y “dexametasona”, ya que sus series de tiempo tienen un comportamiento muy similar al de los datos aportados por el IHME. En la Tabla Figura 7 y Tabla 2, se muestran las búsquedas relacionadas de cada una y sus gráficas de series de tiempo, respectivamente.

Cuadro 2: Clasificación de las consultas hechas en Twitter con sus respectivos números de tweets

Palabra clave	Búsqueda relacionada	Valor
oxígeno	saturación de oxígeno	100
	concentrador de oxígeno	63
	bombona de oxígeno	48
GoFundMe	GoFundMe venezuela	100
tratamiento covid	sin búsquedas relacionadas	
dexametasona	sin búsquedas relacionadas	

Para más detalles, el Notebook que contiene todo el código utilizado para la extracción y presentación de los datos de Google Trends se puede consueguir en [16].

3. Resultados y Análisis

Las gráficas de serie de tiempo de los tres conjuntos de datos extraídos muestran similitud con el comportamiento del modelo IHME, por lo tanto, como se mencionó anteriormente, el análisis comparativo toma como referencia este modelo, sin tomar en cuenta los datos reportados por el gobierno de Venezuela.

En la figuras 8, 9 y 10 se evidencia lo siguiente:

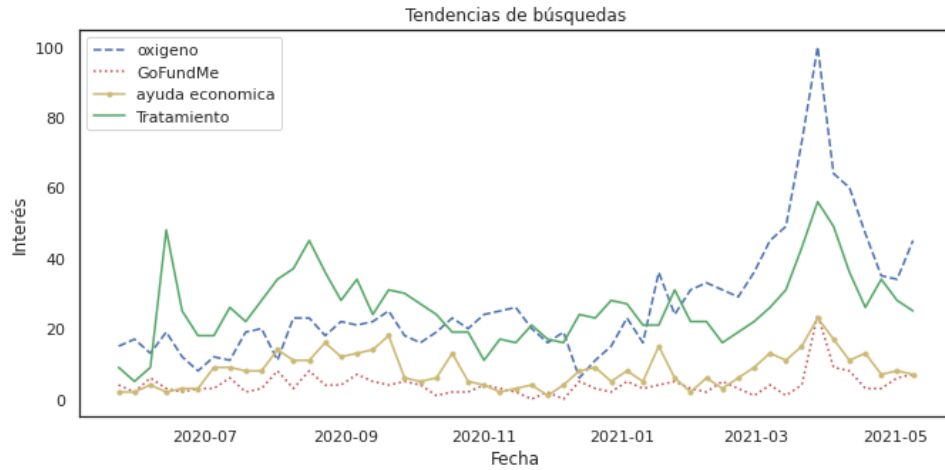


Figura 7: Recuento de las palabras claves en Google Trend

- Las campañas de GoFundMe aumentaron considerablemente en abril de 2021, coincidiendo con el pico más alto de casos reales estimados del modelo IHME. Esto evidencia que el número de campañas aumentan a medida que el número de infectados crece, lo que sugiere que éste podría ser un buen indicador para la incidencia del virus, y esto se confirma al calcular la correlación cruzada entre estas series de tiempo. De estos resultados, presentados en (9), se propone que existe un retraso de aproximadamente 5 días entre el aumento del número de casos y el incremento del número de campañas.
- Los datos extraídos de Twitter presentan un comportamiento irregular a lo largo de la pandemia, sin embargo, se puede notar que en septiembre de 2020 existe un número de publicaciones considerablemente más alto que en otras fechas. También tenemos un incremento notable en agosto de 2020 y abril de 2021. En comparación con la serie de tiempo base, los picos están en las mismas fechas, pero en distinta proporción. Al calcular la correlación cruzada entre estos datos y la estimación media de casos reales del IHME, se obtiene una correlación menor que la que se obtuvo con los datos de GoFundMe, lo que confirma que el ruido presente en éstos es mayor. Otra diferencia que arroja este análisis es que este indicador estaría anticipando el incremento de casos reales del COVID-19.
- En este caso, el análisis de la correlación cruzada permitió discernir que los mejores indicadores obtenidos a partir de Google Trend son las búsquedas de “tratamiento covid” y “dexametasona”. Adicionalmente, este análisis sugiere que un incremento en las búsquedas de “oxígeno” y “dexametasona” anticipan un aumento en el número de casos reales, mientras que las búsquedas de “tratamiento covid” y “GoFundMe” coinciden más. Cabe aclarar que la determinación de los retrasos entre las series de tiempo se ve bastante limitada en este caso porque se trabajó con unos datos de menor frecuencia (los datos de Google Trends son semanales y no diarios).

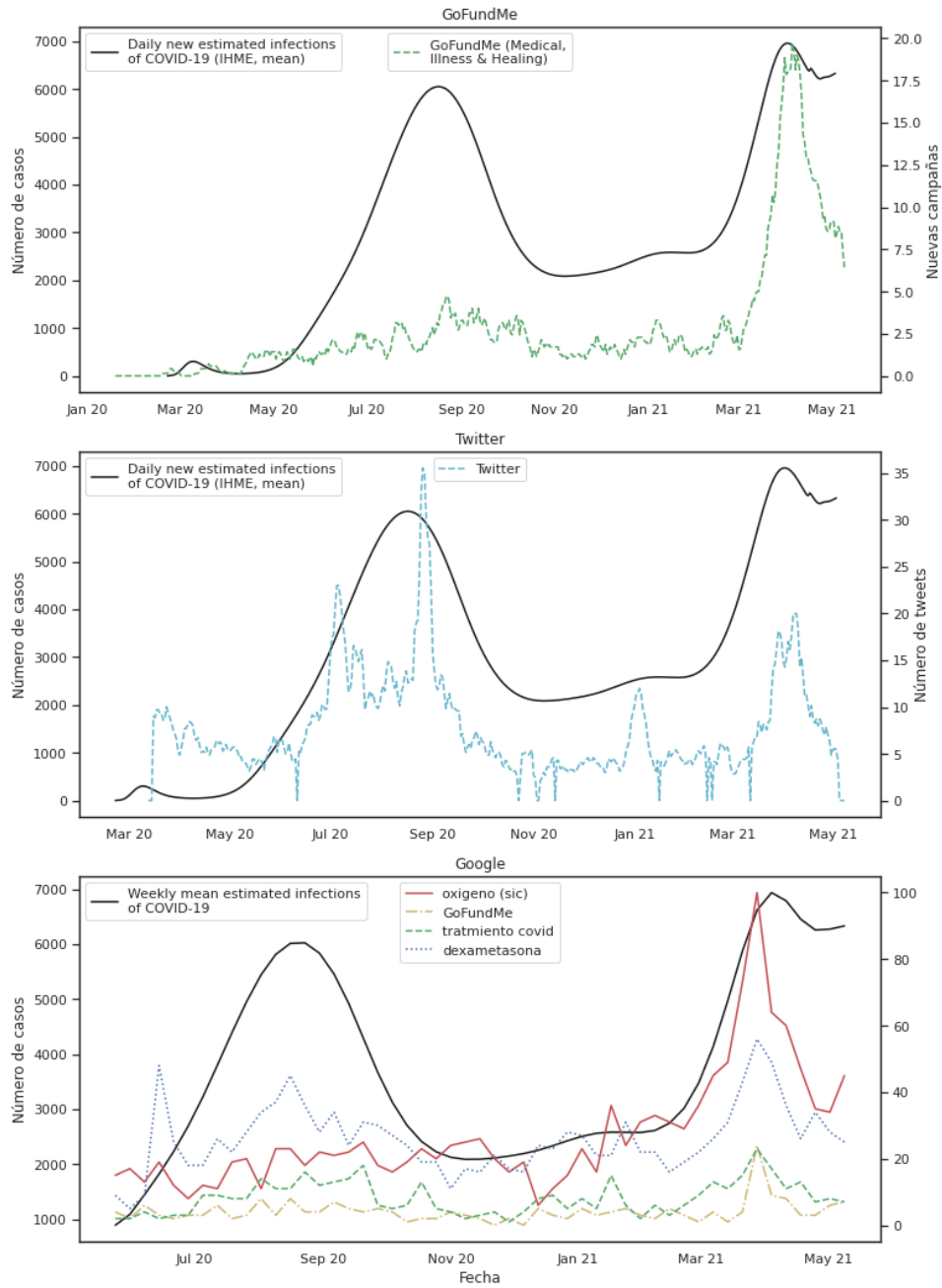


Figura 8: Comparación de los indicadores escogidos con las estimación media del IHME

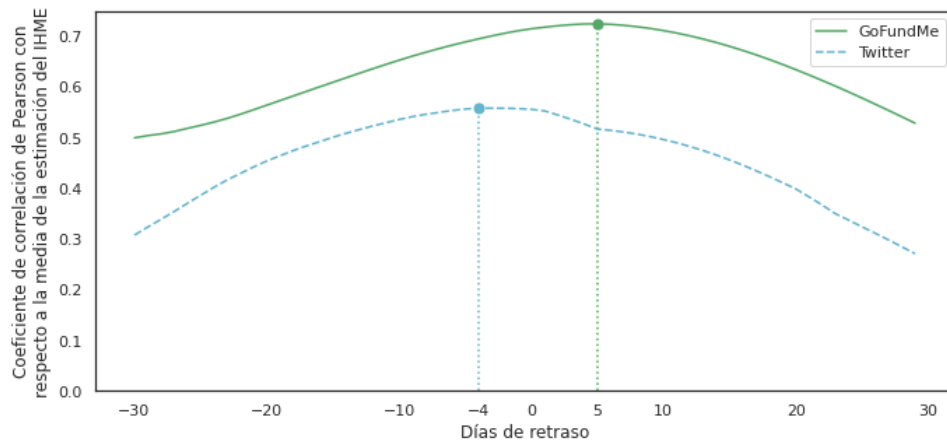


Figura 9: Correlación cruzada de entre la serie de tiempo de la estimación media de casos reales de covid-19 del IHME y los indicadores extraídos de GoFundMe y Twitter

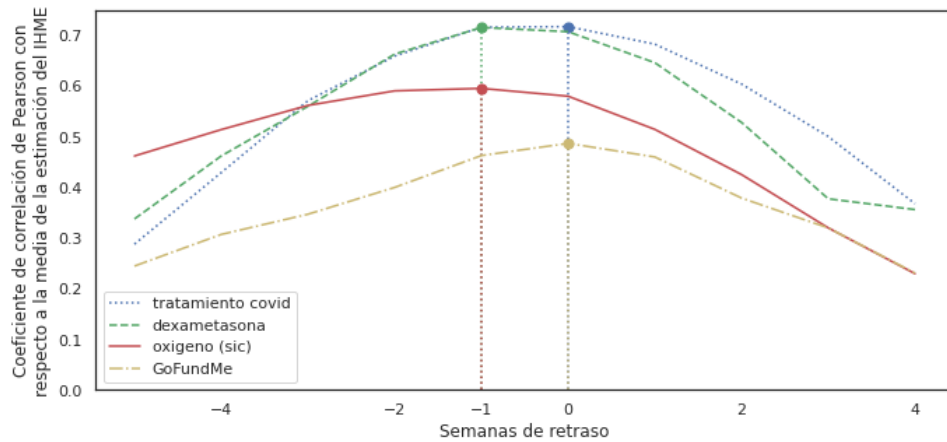


Figura 10: Correlación cruzada de entre la serie de tiempo de la estimación media de casos reales de covid-19 del IHME y los indicadores extraídos de Google Trends

4. Conclusiones y Recomendaciones

Our World in Data provee, entre otras cosas, los resultados de un modelo que estima el número real de infectados de COVID-19 en Venezuela (IHME), además de los datos del número de infectados reportados por el gobierno. Al analizar las gráficas de series de tiempo de estos datos y la de los datos extraídos de las distintas plataformas, se observa que hay cierta congruencia al encontrar dos picos, uno en septiembre de 2020 y otro abril de 2021, aproximadamente, siendo éste último el más pronunciado. Sin embargo, el número de infectados reportados por el gobierno de Venezuela no muestra el incremento abrupto en estos picos, reforzando la hipótesis de que las cifras dadas por el gobierno de Venezuela no son reales.

Estos indicadores parecen ser prometedores y con un estudio más exhaustivo, podrían ser una fuente de información clave para predecir o detectar incidencias de brotes en la población. Sin embargo, se recalca que los datos que se tienen hasta ahora son insuficientes para presentar interpretaciones concluyentes. Este tipo de ejercicio podría tener diversas aplicaciones en otros casos de fenómenos que tengan manifestaciones en redes sociales o Internet, que por alguna u otra razón no cuentan con datos oficiales confiables.

Referencias

- [1] The Institute for Health Metrics and Evaluation (IHME). COVID-19: Estimating the historical time series of infections.
- [2] Ospina-Valencia, J. Apr. 21, 2021. DW. Made for minds. La pandemia en Venezuela: manipulación de la información, cementerios llenos y aún sin plan de vacunación. <https://www.dw.com/es/la-pandemia-en-venezuela-manipulaci%C3%B3n-de-la-informaci%C3%B3n-cementerios-llenos-y-a%C3%BA-sin-plan-de-vacunaci%C3%B3n/a-57258234>
- [3] GoFundMe. <https://es.gofundme.com/>
- [4] Twitter. <https://twitter.com>
- [5] Wikipedia. https://es.wikipedia.org/wiki/Interfaz_de_programaci%C3%B3n_de_aplicaciones
- [6] Google Trends. <https://trends.google.es>
- [7] Soporte de Google Trends. <https://support.google.com/trends/answer/4365533?hl=es>
- [8] Friedman J, Liu P, Troeger CE, Carter A, Reiner RC, Barber RM, Collins J, Lim SS, Pigott DM, Vos T, Hay SI, Murray CJL, Gakidou E. Predictive performance of international COVID-19 mortality forecasting models. Nature Communications. 10 May. 2021. doi: 10.1038/s41467-021-22457-w.
- [9] Our World in Data. <https://ourworldindata.org/grapher/daily-new-estimated-covid-19-infections-ihme-model?country=~VEN>

- [10] Giattino G. Aug. 24, 2020. How epidemiological models of COVID-19 help us estimate the true number of infections. <https://ourworldindata.org/covid-models>
- [11] Harvey Maddocks. May. 18, 2021. gofundme_webscraper. https://github.com/HarveyMaddocks/gofundme_webscraper
- [12] Eliana Bonalde. May. 18, 2021. Proyecto_OP2. https://github.com/bonaldee/Proyecto_OP2/blob/main/codigo/GoFundMe_datos.ipynb
- [13] Eliana Bonalde. May 18, 2021. Proyecto_OP2. https://github.com/bonaldee/Proyecto_OP2/blob/main/codigo/GoFoundMe_analisis.ipynb
- [14] Eliana Bonalde. May. 18, 2021. Proyecto_OP2. https://github.com/bonaldee/Proyecto_OP2/blob/main/codigo/Twitter_analisis.ipynb
- [15] Eliana Bonalde. May. 18, 2021. Proyecto_OP2. https://github.com/bonaldee/Proyecto_OP2/blob/main/codigo/Twitter_datos.ipynb
- [16] Eliana Bonalde. May. 18, 2021. Proyecto_OP2. https://github.com/bonaldee/Proyecto_2/blob/main/codigo/GoogleTrends_datos.ipynb
- [17] Eliana Bonalde. May. 18, 2021. Proyecto_OP2. https://github.com/bonaldee/Proyecto_OP2/blob/main/codigo/analisis_general.ipynb
- [18] Search Tweets. Building queries for Search Tweets. <https://developer.twitter.com/en/docs/twitter-api/tweets/search/integrate/build-a-query#examples>
- [19] Developer terms. Developer Agreement and Policy. <https://developer.twitter.com/en/developer-terms>
- [20] Espinoza, A. Apr 15, 2021. Caracas Chronicles. <https://www.caracaschronicles.com/2021/04/15/gofundme-venezuelans-lifesaver-against-covid-19/>
- [21] Pytrend 4.7.3. <https://pypi.org/project/pytrends/>
- [22] Selenium with Python. <https://selenium-python.readthedocs.io/>
- [23] Beautiful Soup Documentation. <https://www.crummy.com/software/BeautifulSoup/bs4/doc/>
- [24] CurrencyConverter 0.16.1. <https://pypi.org/project/CurrencyConverter/>